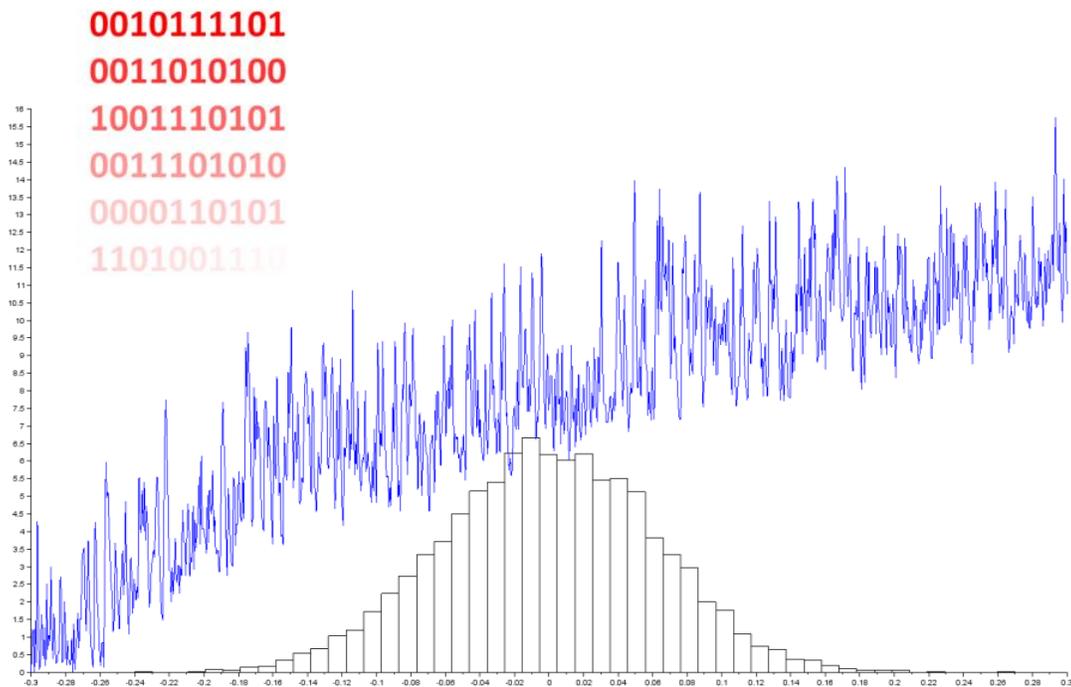


ΣΤΟΙΧΕΙΑ ΠΙΘΑΝΟΤΗΤΩΝ

ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΣΤΑΤΙΣΤΙΚΗ
ΚΑΙ ΤΗΝ ΠΛΗΡΟΦΟΡΙΚΗ

Γιάννης Κοντογιάννης
Σταύρος Τουμπής



Γιάννης Κοντογιάννης
Σταύρος Τουμπής

ΣΤΟΙΧΕΙΑ ΠΙΘΑΝΟΤΗΤΩΝ
ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΣΤΑΤΙΣΤΙΚΗ
ΚΑΙ ΤΗΝ ΠΛΗΡΟΦΟΡΙΚΗ

Αθήνα 2015

**ΣΤΟΙΧΕΙΑ ΠΙΘΑΝΟΤΗΤΩΝ
ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΣΤΑΤΙΣΤΙΚΗ
ΚΑΙ ΤΗΝ ΠΛΗΡΟΦΟΡΙΚΗ**

Συγγραφείς: Γιάννης Κοντογιάννης, Σταύρος Τουμπής
Γλωσσική επιμέλεια: Θεόφιλος Τραμπούλης
Τεχνική επιμέλεια: Σάββας Γαιτζένης
Κριτικός αναγνώστης: Πέτρος Δελλαπόρτας
Έκδοση: Νοέμβριος 2015

ISBN: 978-960-603-182-3

Copyright © Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών – ΣΕΑΒ, 2015



Το παρόν έργο αδειοδοτείται υπό τους όρους της άδειας Creative Commons Αναφορά
Δημιουργού – Μη Εμπορική Χρήση – Παρόμοια Διανομή 3.0

ΣΥΝΔΕΣΜΟΣ ΕΛΛΗΝΙΚΩΝ ΑΚΑΔΗΜΑΪΚΩΝ ΒΙΒΛΙΟΘΗΚΩΝ

Εθνικό Μετσόβιο Πολυτεχνείο
Ηρώων Πολυτεχνείου 9, 15780 Ζωγράφου
<http://www.kallipos.gr>

Πρόλογος

Η θεωρία πιθανοτήτων και οι εφαρμογές της αποτελούν, εδώ και σχεδόν έναν αιώνα, ακέραιο μέρος των μαθηματικών. Επιπλέον, τις τελευταίες δεκαετίες οι πιθανότητες έχουν αποκτήσει κεντρικό ρόλο σε μια πλειάδα σύγχρονων επιστημονικών και τεχνολογικών περιοχών, από τη γενετική και τα χρηματοοικονομικά μέχρι τη θεμελίωση της στατιστικής, και από τη μαθηματική περιγραφή της έννοιας της πληροφορίας ως τις ψηφιακές επικοινωνίες, την πληροφορική, ακόμη και τη μελέτη του ανθρώπινου εγκεφάλου μέσω της νευροεπιστήμης.

Βασικός σκοπός του βιβλίου αυτού είναι να προσφέρει μια εισαγωγή στα στοιχειώδη εργαλεία της κλασικής θεωρίας πιθανοτήτων, από μια σύγχρονη σκοπιά που να επισημαίνει και να συνδέει τη μαθηματική αυτή θεωρία με ενδιαφέρουσες εφαρμογές της σε διάφορα πεδία. Οι κεντρικές έννοιες των πιθανοτήτων και της τυχαιότητας αναπτύσσονται παράλληλα με τη μαθηματική θεωρία που τις περιβάλλει, με έμφαση σε τρεις βασικούς άξονες:

1. *Μαθηματική περιγραφή:* Έχουμε καταβάλει σημαντική προσπάθεια ώστε η θεωρία να παρουσιαστεί με μαθηματικά αυστηρό τρόπο, χωρίς όμως να είναι απαραίτητη η γνώση προχωρημένων και πιο εξειδικευμένων μαθηματικών εργαλείων και τεχνικών.
2. *Σύγχρονα παραδείγματα:* Μαζί με τη θεωρητική ανάπτυξη του αντικείμενου, εξίσου σημαντικός στόχος μας είναι η παρουσίαση, έστω και επιγραμματικά, των πρακτικών εφαρμογών της θεωρίας των πιθανοτήτων. Γι' αυτόν τον λόγο, έχουμε συμπεριλάβει ένα πλήθος παραδειγμάτων και ασκήσεων, με σκοπό την ανάδειξη της στενής σχέσης των πιθανοτήτων με σύγχρονες εφαρμογές τους.
3. *Πληροφορική και στατιστική:* Αν και έχουμε προσπαθήσει να αναδείξουμε με παραδείγματα τη σχέση των πιθανοτήτων με όλο το φάσμα των εφαρμογών – από κλασικά καθημερινά προβλήματα, όπως η ρίψη ενός νομίσματος ή μια εκλογική δημοσκόπηση, μέχρι την ανάλυση κάποιων εξειδικευμένων σύγχρονων αλγορίθμων στην επιστήμη υπολογιστών – έχει δοθεί μεγαλύτερη έμφαση στην πληροφορική και τη στατιστική. Ελπίζουμε αυτό το βιβλίο να μπορέσει να κεντρίσει το ενδιαφέρον κάποιων από τους αναγνώστες του προς αυτές τις κατευθύνσεις και να αποτελέσει έναν προθάλαμο που θα τους οδηγήσει στη βαθύτερη μελέτη των αντίστοιχων περιοχών.

Ειδικά για το σημείο (1) πρέπει να επισημάνουμε πως έχουμε συνειδητά αποφύγει να αναπτύξουμε εκτενώς τη *θεωρία μέτρου*, η οποία αποτελεί το φυσικό και θεμελιώδες μαθηματικό πλαίσιο της θεωρίας των πιθανοτήτων. Παρ' όλα αυτά, στο Κεφάλαιο 3 δίνουμε τον αυστηρό μαθηματικό ορισμό του μέτρου πιθανότητας, και στα Κεφάλαια 5 έως 9 παρουσιάζουμε τη θεωρία των διακριτών τυχαίων μεταβλητών και των ιδιοτήτων τους χωρίς την παραμικρή μαθηματική παράλειψη ή αναφορά στη βαθύτερη θεωρία μέτρου.

Αν και αυτό το βιβλίο προέκυψε από τη διδασκαλία δύο προπτυχιακών μαθημάτων στο τμήμα Πληροφορικής του Οικονομικού Πανεπιστημίου Αθηνών («Πιθανότητες» και «Εφαρμοσμένες πιθανότητες και προσομοίωση», κατά τη δεκαετία 2005-2015), ευελπιστούμε πως θα μπορέσει να χρησιμοποιηθεί στη διδασκαλία εισαγωγικών μαθημάτων πιθανοτήτων σε οποιοδήποτε τμήμα θετικών επιστημών ή πολυτεχνικών σχολών. Κατά συνέπεια, το κύριο ακροατήριο στο οποίο απευθυνόμαστε είναι προπτυχιακοί φοιτητές τέτοιων τμημάτων, και γι' αυτόν τον λόγο έχουμε φροντίσει οι μαθηματικές γνώσεις που απαιτούνται να μην ξεπερνούν κατά πολύ την ύλη μαθηματικών του Λυκείου (αν την ξεπερνούν).

Η έμφασή μας στη σχέση των πιθανοτήτων με την πληροφορική και τη στατιστική εξηγείται από τα προσωπικά επιστημονικά ενδιαφέροντα των συγγραφέων, από το γεγονός ότι τα μαθήματα που σε πρώτη φάση μάς οδήγησαν στη συγγραφή του βιβλίου διδάσκονταν σε τμήμα πληροφορικής, αλλά και από την εκτίμησή μας πως οι περιοχές της επιστήμης υπολογιστών και της στατιστικής προσφέρουν σημαντικότερο πεδίο επιστημονικής μελέτης για τα σύγχρονα εφαρμοσμένα μαθηματικά εν γένει.

Η ύλη η οποία περιέχεται σε αυτό το βιβλίο καλύπτει πλήρως το περιεχόμενο του εισαγωγικού μαθήματος «Πιθανότητες» και περίπου το ένα έκτο του πιο προχωρημένου μαθήματος «Εφαρμοσμένες πιθανότητες και προσομοίωση», τα οποία προαναφέραμε. Επιπλέον, κατά τη συγγραφή του βιβλίου ενσωματώσαμε και τα ακόλουθα στοιχεία:

1. Κάποια πιο προχωρημένα μέρη της ύλης του μαθήματος των «Πιθανοτήτων», τα οποία δεν παρουσιάζονταν στις διαλέξεις κάθε χρόνο ή διδάσκονταν προαιρετικά.
2. Αποδείξεις κάποιων θεωρημάτων που δεν παρουσιάζονταν στα μαθήματα και παρατίθενται εδώ για λόγους πληρότητας (σημειωμένες με *).
3. Κάποιες παραγράφους (επίσης σημειωμένες με *) που περιλαμβάνουν προχωρημένο υλικό. Όσοι αναγνώστες θέλουν να εμβαθύνουν στις αντίστοιχες περιοχές θα βρουν αυτές τις παραγράφους τις πιο ενδιαφέρουσες.
4. Πολλές ασκήσεις μαζί με τις αναλυτικές τους λύσεις στο τέλος του κειμένου, οι πιο δύσκολες εκ των οποίων είναι και αυτές σημειωμένες με *.

Υπογραμμίζουμε πως οι ασκήσεις και οι λύσεις τους αποτελούν σημαντικό και αμέριστο μέρος του βιβλίου. Στις ασκήσεις έχουμε συμπεριλάβει και κάποια σημαντικά αποτελέσματα, τα οποία άλλοτε δεν εντάσσονταν φυσικά στη ροή της ύλης του αντίστοιχου κεφαλαίου και άλλοτε κρίναμε πως αφορούσαν ερωτήματα τα οποία θα ήταν πιο χρήσιμο για τον αναγνώστη να έχει την ευκαιρία να τα εξετάσει μόνος του πριν του δοθούν οι απαντήσεις. Επιπλέον, η ίδια η φύση του αντικειμένου απαιτεί συστηματική εξοικείωση – με χαρτί και μολύβι – με τις έννοιες και τις τεχνικές που σταδιακά εισάγονται, και ίσως ο πιο αποτελεσματικός τρόπος να το επιτύχει κανείς αυτό είναι να αναπτύξει τις απαραίτητες δεξιότητες λύνοντας μια σειρά σχετικών ασκήσεων.

Είναι βέβαια μάλλον αυτονόητο, αλλά τονίζουμε πως δεν συνιστάται στους φοιτητές να χρησιμοποιήσουν το παρόν βιβλίο ως υποκατάστατο της παρακολούθησης των διαλέξεων.

Κλείνοντας, έχουμε τη χαρά να ευχαριστήσουμε τις δέκα περίπου «γενιές» φοιτητών που παρακολούθησαν τα μαθήματά μας για τη βοήθειά τους στη διαμόρφωση και την τροποποίηση της ύλης και για τις πολλές και χρήσιμες υποδείξεις τους.

Τέλος, ο πρώτος συγγραφέας αφιερώνει αυτό το βιβλίο στον ομορφότερο άνθρωπο που είχε την τύχη να γνωρίσει ποτέ στη ζωή του, τον γιο του Γιώργο. Ο δεύτερος συγγραφέας το αφιερώνει στην τριφυλλάρα ♣.

Γιάννης Κοντογιάννης, Σταύρος Τουμπής
Αθήνα, Σεπτέμβριος 2015

Περιεχόμενα

1	Εισαγωγή	1
1.1	Οι πιθανότητες ως μέρος των μαθηματικών	1
1.2	Ιστορική ανάπτυξη	2
1.3	Πιθανότητες και πληροφορική	5
2	Χώρος πιθανότητας και ενδεχόμενα	7
2.1	Προκαταρκτικά	7
2.2	Σύνολα	8
2.3	Χώρος πιθανότητας και ενδεχόμενα	12
2.4	Ασκήσεις	16
3	Μέτρο πιθανότητας	19
3.1	Ορισμός, παραδείγματα και ιδιότητες	19
3.2	Πέντε «κανόνες πιθανότητας»	23
3.3	* Γενικός ορισμός του μέτρου πιθανότητας	24
3.4	Ασκήσεις	26
4	Πιθανότητες και συνδυαστική	29
4.1	Διατάξεις, συνδυασμοί, επιλογές και πιθανότητες	29
4.2	Πέντε «κανόνες αρίθμησης»	37
4.3	Ασκήσεις	38
5	Ανεξαρτησία και δεσμευμένη πιθανότητα	41
5.1	Ανεξάρτητα ενδεχόμενα και δεσμευμένη πιθανότητα	41
5.2	Περαιτέρω ιδιότητες	48
5.3	Ο κανόνας του Bayes	55
5.4	Ακόμα πέντε «κανόνες πιθανότητας»	58
5.5	Ασκήσεις	59

6 Διακριτές τυχαίες μεταβλητές	63
6.1 Ορισμός και βασικές ιδιότητες	63
6.2 Μέση τιμή, διασπορά, ανεξαρτησία	70
6.3 * Μετρησιμότητα και άπειρες τιμές	77
6.3.1 Ορισμός τυχαίας μεταβλητής	77
6.3.2 Ορισμοί $E(X)$, $\text{Var}(X)$, και επιπλέον συνθήκες	78
6.4 Ασκήσεις	80
7 Διακριτές κατανομές	85
7.1 Κατανομές Bernoulli, διωνυμική και γεωμετρική	85
7.2 Υπεργεωμετρική και Poisson κατανομή	94
7.3 * Η γεωμετρική και συναφείς σειρές	101
7.4 * Η συνάρτηση e^x , δυναμοσειρές, τύπος του Stirling	102
7.5 Ασκήσεις	106
8 Παραδείγματα πιθανοκρατικής ανάλυσης αλγορίθμων	111
8.1 Επαλήθευση ισότητας πολυωνύμων	112
8.2 Εύρεση ελαχιστιαίου cut set σε γράφους	114
8.3 String matching	121
8.4 Γρήγορη ταξινόμηση δεδομένων	123
8.5 Ασκήσεις	129
9 Από κοινού κατανομή και ο N.M.A.	133
9.1 Ανισότητες Markov και Chebychev	133
9.2 Από κοινού κατανομή και συνδιακύμανση	135
9.3 Ο Νόμος των Μεγάλων Αριθμών	144
9.4 Ασκήσεις	150
10 Συνεχείς τυχαίες μεταβλητές	155
10.1 Συνεχείς T.M. και συνεχής πυκνότητα	156
10.2 Μέση τιμή και διασπορά	163
10.3 * Μετρησιμότητα και άπειρες τιμές	168
10.3.1 Ορισμός μιας συνεχούς T.M.	168
10.3.2 Ορισμοί $E(X)$, $\text{Var}(X)$, και επιπλέον συνθήκες	170
10.4 Ασκήσεις	171
11 Συνεχείς κατανομές και ο N.M.A.	175
11.1 Ομοιόμορφη και εκθετική κατανομή	175
11.2 Μετασχηματισμοί	181
11.3 Ανεξαρτησία, ανισότητες και ο N.M.A.	183
11.4 Ασκήσεις	187

12 Το Κεντρικό Οριακό Θεώρημα	191
12.1 Κανονική κατανομή	192
12.2 Το Κεντρικό Οριακό Θεώρημα	198
12.3 Παραδείγματα και απλές εφαρμογές	200
12.4 Πίνακες τιμών της τυπικής κανονικής κατανομής	210
12.5 Ασκήσεις	213
13 * Κ.Ο.Θ.: Λίγη θεωρία και αποδείξεις	217
13.1 Το γκαουσιανό ολοκλήρωμα	218
13.2 Κ.Ο.Θ. \Rightarrow Ν.Μ.Α.: Απόδειξη	219
13.3 Το θεώρημα de Moivre-Laplace	221
13.4 Το θεώρημα του Lindeberg	224
13.5 Ασκήσεις	232
14 Παραδείγματα εφαρμογών στη στατιστική	235
14.1 Διαστήματα εμπιστοσύνης	236
14.2 Έλεγχοι υποθέσεων	243
14.2.1 Έλεγχος παραμέτρου Bernoulli	245
14.2.2 Έλεγχος ανεξαρτησίας	248
14.3 Μείωση διασποράς	252
14.4 Ασκήσεις	261
15 Συνεχής από κοινού κατανομή	265
15.1 Από κοινού πυκνότητα	265
15.2 Μέση τιμή, διασπορά και συνδιακύμανση	274
15.3 Ανεξαρτησία	277
15.4 * Μετρησιμότητα και άπειρες τιμές	281
15.5 Ασκήσεις	283
A' Παράρτημα	287
A'.1 Διπλά ολοκληρώματα: Συνοπτική επισκόπηση	287
A'.2 Ασκήσεις Κεφαλαίου 2	293
A'.3 Ασκήσεις Κεφαλαίου 3	299
A'.4 Ασκήσεις Κεφαλαίου 4	305
A'.5 Ασκήσεις Κεφαλαίου 5	311
A'.6 Ασκήσεις Κεφαλαίου 6	323
A'.7 Ασκήσεις Κεφαλαίου 7	343
A'.8 Ασκήσεις Κεφαλαίου 8	357
A'.9 Ασκήσεις Κεφαλαίου 9	365
A'.10 Ασκήσεις Κεφαλαίου 10	379
A'.11 Ασκήσεις Κεφαλαίου 11	393

A'.12 Ασκήσεις Κεφαλαίου 12	403
A'.13 Ασκήσεις Κεφαλαίου 13	417
A'.14 Ασκήσεις Κεφαλαίου 14	423
A'.15 Ασκήσεις Κεφαλαίου 15	433
Βιβλιογραφία	451
Ευρετήριο	453
Απόδοση αγγλικών όρων	457

Αρκτικόλεξα

Κ.Ο.Θ.: Κεντρικό Οριακό Θεώρημα

Ν.Μ.Α.: Νόμος των Μεγάλων Αριθμών

Τ.Μ.: Τυχαία Μεταβλητή

Κεφάλαιο 1

Εισαγωγή

1.1 Οι πιθανότητες ως μέρος των μαθηματικών

Ιστορικά, έχουν υπάρξει δύο βασικές κινητήριες δυνάμεις για την ανάπτυξη νέων μαθηματικών: Η ανθρώπινη πνευματική περιέργεια και η ευρύτερη επιστημονική ή κοινωνική αναγκαιότητα της κάθε εποχής. Για παράδειγμα, οι πρακτικές ανάγκες της μέτρησης εδαφών και αποστάσεων στην αρχαιότητα αποτέλεσαν σημαντικό κίνητρο για την ανάπτυξη της επίπεδης (Ευκλείδειας) γεωμετρίας. Παρομοίως, η ανάγκη για την κατανόηση και την πρόβλεψη της κίνησης των στερεών σωμάτων – όπως, π.χ., των πλανητών ή των βλημάτων που χρησιμοποιούνταν σε πολεμικές μάχες – ήταν ένα από τα βασικότερα κίνητρα για την ανάπτυξη του διαφορικού λογισμού από τον Νεύτωνα και τον Leibniz.

Ένα πιο πρόσφατο, και ίσως πιο οικείο, παράδειγμα είναι η ανάπτυξη μιας νέας μαθηματικής θεωρίας για την περιγραφή και την ακριβή μέτρηση της «πληροφορίας». Στην εποχή μας, η έννοια της πληροφορίας βρίσκεται παντού – από τις πληροφορίες που μεταφέρονται ως δεδομένα μέσω του διαδικτύου και των κινητών τηλεφώνων, μέχρι τη μελέτη των πληροφοριών που είναι αποθηκευμένες στον ανθρώπινο εγκέφαλο και στο DNA. Πώς μετρείται και περιγράφεται η πληροφορία, ως φυσικό μέγεθος, στην καθεμία από τις πιο πάνω περιπτώσεις; Το επιστημονικό πεδίο της θεωρίας πληροφορίας δίνει κάποιες πρώτες απαντήσεις σε αυτά τα ερωτήματα.

Μάλλον το σημαντικότερο (και αρχαιότερο) κίνητρο για την ανάπτυξη των πιθανοτήτων – δηλαδή μιας μαθηματικά αυστηρής θεωρίας για την κατανόηση τυχαίων φαινομένων και, γενικότερα, καταστάσεων στις οποίες υπάρχει ένα σημαντικό μέρος *αβεβαιότητας* – ήταν το ανθρώπινο πάθος για τον τζόγο. Γύρω στα μέσα και προς τα τέλη του 19ου αιώνα, είχε ωριμάσει αρκετά η συστηματική μελέτη των σχετικά απλών φυσικών φαινομένων, όπως για παράδειγμα η μελέτη της κίνησης δύο απομονωμένων πλανητών κάτω από την επίδραση της αμοιβαίας βαρυτικής τους έλξης, και είχε ξεκινήσει να αναπτύσσεται έντονο επιστημονικό ενδιαφέρον για τη μελέτη «πολύπλοκων» συστημάτων.

Για παράδειγμα, ένα δωμάτιο περιέχει περίπου 10^{21} μόρια αέρα. Ακόμα κι αν γνωρίζουμε με ακρίβεια τους νόμους που διέπουν την κίνησή τους, είναι πρακτικά αδύνατο να λύσουμε ένα σύστημα 10^{21} διαφορικών εξισώσεων, ώστε να προβλέψουμε, π.χ., τη θερμοκρασία του αέρα στο δωμάτιο! Μια αποτελεσματικότερη προσέγγιση είναι να θεωρήσουμε τις θέσεις και τις ταχύτητες των μορίων τυχαίες και να αποπειραθούμε να κάνουμε μια στατιστική ανάλυση. Αυτή η προσέγγιση, η οποία αποτέλεσε την αφετηρία της σημαντικής νέας περιοχής της στατιστικής φυσικής, έδωσε την τελική ώθηση που απαιτούνταν ώστε οι πιθανότητες να αναπτυχθούν ως μια πλήρης μαθηματική θεωρία στο πρώτο μισό του 20ού αιώνα.

1.2 Ιστορική ανάπτυξη

Η αφετηρία της συστηματικής μελέτης των Πιθανοτήτων ως επιστημονικού πεδίου τοποθετείται στα μέσα του 17ου αιώνα, και συγκεκριμένα στην αλληλογραφία μεταξύ δύο σημαντικών μαθηματικών της εποχής, του Pascal και του Fermat, με αντικείμενο την κατανόηση ενός τυχερού παιχνιδιού.



Σχήμα 1.1: Ο Blaise Pascal (1623-1662) και ο Pierre de Fermat (1601-1665). [Οι εικόνες αποτελούν κοινό κτήμα και διέπονται από το καθεστώς υλικού που ανήκει στο public domain. Τα πρωτότυπα αρχεία βρίσκονται στις τοποθεσίες των συνδέσμων που δίνονται στο τέλος του κεφαλαίου.]

Μετά τη θεμελίωση των βασικών εννοιών από τους Pascal-Fermat, η σκυτάλη πέρασε σε έναν από τους σημαντικότερους μαθηματικούς όλων των εποχών, τον Gauss. Στα χέρια του Gauss, οι πιθανότητες έπαψαν να αποτελούν ένα συνονθύλευμα μεμονωμένων παραδειγμάτων και απλών τεχνικών. Ο Gauss διατύπωσε και απέδειξε μια σειρά από θεμελιώδη αποτελέσματα, τα οποία

αποτελούν τη βάση ολόκληρης της σύγχρονης θεωρίας πιθανοτήτων – αλλά και της στατιστικής – έως και σήμερα. Το σημαντικότερο από αυτά τα αποτελέσματα είναι το *Κεντρικό Οριακό Θεώρημα* (Κ.Ο.Θ.), του οποίου η μελέτη και η χρήση αποτελούν κεντρικούς στόχους αυτού του βιβλίου.



Σχήμα 1.2: Ο Johann Carl Frederich Gauss (1777-1866). [*Η εικόνα αποτελεί κοινό κτήμα και διέπεται από το καθεστώς υλικού που ανήκει στο public domain. Το πρωτότυπο αρχείο βρίσκεται στην τοποθεσία του συνδέσμου που δίνεται στο τέλος του κεφαλαίου.*]

Με απλά λόγια, το Κ.Ο.Θ. μας λέει δύο πράγματα: Πρώτον, πως μέσα από την πλήρη αταξία μερικές φορές γεννιέται τάξη. Για παράδειγμα, αν στρίψουμε ένα νόμισμα δυο-τρεις φορές, είναι απολύτως αδύνατο να προβλέψουμε τι θα συμβεί: αν, ας πούμε, θα φέρουμε πρώτα Γράμματα και μετά Κορώνα ή το αντίστροφο. Αλλά, αν στρίψουμε το νόμισμα χίλιες ή δέκα χιλιάδες φορές, τότε είναι σχεδόν βέβαιο ότι το ποσοστό των φορών που φέραμε Κορώνα θα είναι μεταξύ 49% και 51%. Επιπλέον, το Κ.Ο.Θ. μάς επιτρέπει να υπολογίσουμε, κατά προσέγγιση, πόσο μικρή είναι η πιθανότητα το ποσοστό από Κορώνες να *μην* είναι μεταξύ 49% και 51%.

Έτσι, από τις πολλές επαναλήψεις του τυχαίου και απρόβλεπτου, προκύπτει τάξη και προβλεψιμότητα. Όπως θα δούμε σε επόμενα κεφάλαια, η βασική ιδιότητα πάνω στην οποία στηρίζεται αυτή η συμπεριφορά, είναι η *ανεξαρτησία*, δηλαδή το γεγονός ότι τα αποτελέσματα των διαδοχικών ρίψεων του νομίσματος είναι ανεξάρτητα το ένα από το άλλο.

Ως την εποχή του Gauss και μέχρι μερικές δεκαετίες αργότερα, η μελέτη των πιθανοτήτων βασιζόταν σχεδόν εξολοκλήρου στην υπόθεση της ανεξαρτησίας. Για παράδειγμα, σε μια ιατρική μελέτη, είναι λογικό να υποθέσουμε ότι το πόσο αποτελεσματικά δρα ένα φάρμακο έχει διακυμάνσεις από ασθενή σε ασθενή, αλλά είναι εξίσου λογικό να υποθέσουμε ότι η αποτελεσματικότητα του φαρμάκου είναι ανεξάρτητη από τον εκάστοτε ασθενή. Παρομοίως, αν κάνουμε μια δημοσκόπηση διαλέγοντας τυχαία μέλη ενός πληθυσμού, είναι λογικό να υποθέσουμε πως το να επιλέξουμε

έναν συγκεκριμένο άνθρωπο για τη δημοσκόπηση δεν θα επηρεάσει τις πολιτικές προτιμήσεις κάποιου άλλου. Και στις δύο αυτές περιπτώσεις μπορούμε, λοιπόν, να υποθέσουμε πως διαδοχικά δείγματα – μετρήσεις της ανταπόκρισης των ασθενών σε ένα φάρμακο και προτιμήσεις ψηφοφόρων – είναι στατιστικά ανεξάρτητα.

Αλλά σε πιο πολύπλοκα φαινόμενα η υπόθεση της ανεξαρτησίας δεν είναι ρεαλιστική. Για παράδειγμα, ας πούμε πως έχουμε έναν αλγόριθμο επεξεργασίας κειμένου και θέλουμε να αναλύσουμε τη συνήθη συμπεριφορά του. Μια που δεν ξέρουμε εκ των προτέρων πάνω σε ποιο κείμενο θα εφαρμοστεί, λογικά θα καταφύγουμε στο να εξετάσουμε πώς συμπεριφέρεται σε κάποιο «τυχαίο» κείμενο. Αν όμως περιγράψουμε ένα τυχαίο κείμενο ως μια ακολουθία τυχαίων γραμμάτων, τότε σίγουρα δεν μπορούμε να θεωρήσουμε πως τα διαδοχικά γράμματα είναι ανεξάρτητα μεταξύ τους – εκτός κι αν είμαστε προετοιμασμένοι να δεχθούμε ως «κείμενο» μια ακολουθία γραμμάτων όπως η: ασλδ'Κ αΣλκδν οι θζαίις .θΔΙοαιξΜ-ΟΙα τρ88 Δλκσαξ λκλ,09ος ,σαμδ!



Σχήμα 1.3: Ο Andrei A. Markov (1856-1922) και ο Andrei N. Kolmogorov (1903-1987). [Οι εικόνες αποτελούν κοινό κτήμα και διέπονται από το καθεστώς υλικού που ανήκει στο public domain. Τα πρωτότυπα αρχεία βρίσκονται στις διαδικτυακές τοποθεσίες των συνδέσμων που δίνονται στο τέλος του κεφαλαίου.]

Ο πρώτος ερευνητής που μελέτησε συστηματικά τις τυχαίες ακολουθίες που αποτελούνται από όχι ανεξάρτητα αλλά *συσχετισμένα* μεταξύ τους δείγματα, ήταν ο Markov. Στην απλούστερη μορφή τους, οι ακολουθίες τέτοιων συσχετισμένων δειγμάτων ονομάζονται «αλυσίδες Markov», και η μελέτη τους αποτελεί κεντρικό μέρος πολλών ερευνητικών περιοχών της σύγχρονης επιστήμης και τεχνολογίας. Είναι αξιοσημείωτο πως ένα από τα βασικά κίνητρα του Markov ήταν η περιγραφή κειμένων φυσικής γλώσσας μέσω των πιθανοτήτων. Ακόμη και στη σημερινή εποχή του Internet, του Google και του YouTube, πολλοί από τους πιο δημοφιλείς αλγόριθμους που

χρησιμοποιούνται καθημερινά από εκατομμύρια ανθρώπους καθώς «σερφάρουν» στο διαδίκτυο, είναι βασισμένοι σε μοντέλα που περιγράφουν το περιεχόμενο των σελίδων του WWW μέσω των αλυσίδων Markov.

Ο πιο πρόσφατος μεγάλος σταθμός στην ιστορία των πιθανοτήτων είναι το 1933. Μέχρι τότε, παρά τη μεγάλη ώθηση που είχε πάρει η μελέτη τυχαίων φαινομένων στη φυσική και στο πρωτοεμφανιζόμενο τότε πεδίο της στατιστικής, οι πιθανότητες παρέμεναν μια μαθηματικά ακρόφημη επιστημονική περιοχή. Ο λόγος ήταν πως δεν είχαν ακόμα ενταχθεί, με την αυστηρή έννοια, στο κεντρικό κομμάτι των μαθηματικών. Δεν είχαν, δηλαδή, θεμελιωθεί αξιωματικά, όπως η γεωμετρία, η ανάλυση, η θεωρία συνόλων και όλες οι υπόλοιπες βασικές περιοχές των μαθηματικών. Αυτήν τη θεμελίωση κατάφερε το 1933 ο σπουδαίος Ρώσος μαθηματικός A.N. Kolmogorov, του οποίου η τεράστια επιρροή στην επιστημονική εξέλιξη του 20ού αιώνα είναι εξαιρετικά έντονα αισθητή ως τις μέρες μας.

1.3 Πιθανότητες και πληροφορική

Όπως αναφέρουν στην εισαγωγή του πρόσφατου βιβλίου τους *Probability and Computing* οι Mitzenmacher (Harvard) και Upfal (Brown):

Τις τελευταίες δύο δεκαετίες, η χρήση της θεωρίας πιθανοτήτων στην πληροφορική έχει ενταθεί σε πάρα πολύ μεγάλο βαθμό. Προχωρημένες και πολύπλοκες τεχνικές από τις πιθανότητες αναπτύσσονται και βρίσκουν εφαρμογή σε όλο και πιο ευρείες και δύσκολες περιοχές της επιστήμης υπολογιστών.

Συγκεκριμένα, τεχνικές και βασικές έννοιες των πιθανοτήτων παίζουν κεντρικό ρόλο, μεταξύ άλλων, στις εξής περιοχές:

- **Περιγραφή και προσομοίωση πολύπλοκων συστημάτων.** Π.χ., ένα μεγάλο δίκτυο που αποτελείται από πολλούς υπολογιστές (όπως το internet), ή ένα δίκτυο κινητής τηλεφωνίας, είναι αδύνατον να περιγραφεί με απόλυτη ακρίβεια. Νέοι υπολογιστές προστίθενται στο δίκτυο, κάποιοι αποσυνδέονται, ενώ και η συνδεσμολογία διαρκώς αλλάζει καθώς δημιουργούνται νέες συνδέσεις ή κάποιες υπάρχουσες παύουν να λειτουργούν. Επιπλέον, οι απαιτήσεις για τη μεταφορά δεδομένων αλλάζουν κάθε στιγμή με απρόβλεπτο τρόπο. Έτσι αναγκαστικά καταφεύγουμε σε μια πιθανοκρατική περιγραφή του δικτύου.
- **Πιθανοκρατική ανάλυση αλγορίθμων.** Συχνά παρατηρούμε ένας αλγόριθμος να έχει θεωρητικά απαγορευτικά μεγάλη πολυπλοκότητα, αλλά στην πράξη να είναι πολύ αποτελεσματικός. Αυτό συμβαίνει γιατί, ενώ η παραδοσιακή έννοια της πολυπλοκότητας βασίζεται στην ανάλυση της συμπεριφοράς του αλγορίθμου στη χειρότερη περίπτωση (σύμφωνα με τη λεγόμενη worst case analysis), στη μεγάλη πλειονότητα των περιπτώσεων μπορεί να είναι πολύ αποτελεσματικός. Η λεγόμενη πιθανοκρατική (ή average case) ανάλυση δίνει μια εξήγηση γι' αυτό το φαινόμενο: Αν θεωρήσουμε τα δεδομένα εισόδου τυχαία, τότε σε

πολλές περιπτώσεις μπορούμε να δείξουμε ότι, με πιθανότητα πολύ κοντά στο 100%, η πολυπλοκότητα του αλγορίθμου είναι πολύ σημαντικά μικρότερη από αυτήν της χειρότερης περίπτωσης.

- **Randomized αλγόριθμοι.** Υπάρχει μια κατηγορία αλγορίθμων, οι λεγόμενοι *randomized* ή *τυχαιοκρατικοί* αλγόριθμοι, οι οποίοι σε κάποια βήματα κατά την εκτέλεσή τους κάνουν «τυχαίες» επιλογές. Για παράδειγμα, το πρωτόκολλο επικοινωνίας του ethernet χρησιμοποιεί τυχαίους αριθμούς για να αποφασίσει πότε θα ξαναζητήσει πρόσβαση στο δίκτυο. Η χρήση της τυχειότητας – σε αυτόν και πολλούς άλλους σημαντικούς αλγορίθμους – όχι μόνο απλοποιεί τη δομή του αλγορίθμου, αλλά επιτυγχάνει σημαντικά καλύτερη συμπεριφορά του συστήματος. Το τίμημα που επιφέρει είναι πως πάντα υπάρχει μια μικρή πιθανότητα δυσλειτουργίας. Βάσει σωστού σχεδιασμού και προσεκτικής μαθηματικής ανάλυσης, αυτή η πιθανότητα μπορεί να καταστεί τόσο μικρή, ώστε το κέρδος από την άποψη της πολυπλοκότητας και της ευκολίας να είναι πολύ μεγαλύτερο.

Κλείνοντας, αναφέρουμε πως άλλες περιοχές της πληροφορικής στις οποίες χρησιμοποιούνται συστηματικά μέθοδοι των πιθανοτήτων περιλαμβάνουν, μεταξύ άλλων:

- το σχεδιασμό αλγορίθμων επεξεργασίας πολυμεσικών (multimedia) δεδομένων, π.χ., για τη συμπίεση ήχου, βίντεο και εικόνας,
- την ανάπτυξη μεθόδων μηχανικής μάθησης και ανάκτησης πληροφοριών,
- την κρυπτογραφία,
- τη θεωρητική θεμελίωση των βασικών εννοιών πολυπλοκότητας και υπολογισιμότητας (μηχανές Turing, NP-complete και NP-hard προβλήματα, κλπ.).

Σύνδεσμοι

Διαδικτυακές τοποθεσίες στις οποίες βρίσκονται οι εικόνες των Σχημάτων 1.1, 1.2 και 1.3.

Pascal:

commons.wikimedia.org/wiki/File:Blaise_pascal.jpg

Fermat:

commons.wikimedia.org/wiki/File:Pierre_de_Fermat.jpg

Gauss:

commons.wikimedia.org/wiki/File:Carl_Friedrich_Gauss.jpg

Markov:

commons.wikimedia.org/wiki/File:Andrei_Markov.jpg

Kolmogorov:

pronouncemath.blogspot.gr/2013/04/andrey-nikolaevich-kolmogorov.html

Κεφάλαιο 2

Χώρος πιθανότητας και ενδεχόμενα

2.1 Προκαταρκτικά

Έστω ότι κάποιος μας προτείνει να του δώσουμε δυόμισι ευρώ για να παίξουμε το εξής παιχνίδι: Θα στρίβουμε ένα νόμισμα μέχρι την πρώτη φορά που θα φέρουμε Κορώνα (Κ), κι όσο πιο αργά συμβεί αυτό, δηλαδή όσο πιο πολλές συνεχόμενες φορές φέρουμε Γράμματα (Γ) στην αρχή, τόσο πιο μεγάλο θα είναι το κέρδος μας: Αν το νόμισμα έρθει Κ στην πρώτη ρίψη, θα μας δώσει ένα ευρώ. Αν έρθει Γ και μετά Κ, θα μας δώσει δύο ευρώ. Γενικά, αν έρθει $(n - 1)$ φορές Γ και τη φορά n έρθει Κ, θα πάρουμε n ευρώ.

Αμέσως γεννιούνται μερικά προφανή ερωτήματα:

- Μας συμφέρει να παίξουμε;
- Πόσο πιθανό είναι να κερδίσουμε πιο πολλά χρήματα από όσα δώσαμε για να παίξουμε;
- Αν παίξουμε πολλές φορές, τελικά τι είναι πιο πιθανό, να βγούμε κερδισμένοι ή χαμένοι;
- Είναι «δίκαιη» η τιμή των 2.5 ευρώ;
- Τι θα πει ακριβώς «δίκαιη» τιμή;

Όλα αυτά τα ερωτήματα θα απαντηθούν με συστηματικό τρόπο στα επόμενα κεφάλαια. Προς το παρόν, αυτό που παρατηρούμε είναι η αναγκαιότητα να δώσουμε μια μαθηματική περιγραφή στο πιο πάνω παιχνίδι. Να ορίσουμε, πρώτα από όλα, τι θα πει «πιθανότητα» και να βρούμε τρόπους να υπολογίζουμε ποσοτικά και με ακρίβεια τις απαντήσεις σε ερωτήματα όπως τα πιο πάνω. Αυτό στα μαθηματικά είναι η διαδικασία κατά την οποία περιγράφουμε ένα πραγματικό φαινόμενο μέσω ενός μαθηματικού «μοντέλου». Σε κάποιες περιπτώσεις, αυτή η διαδικασία μάς είναι τόσο οικεία που ούτε καν της δίνουμε σημασία – για παράδειγμα, όταν βλέπουμε σε ένα χάρτη

μια ευθεία γραμμή να αναπαριστά ένα δρόμο, δεν σκεφτόμαστε «Α, βέβαια, εδώ επικαλούμαι την προσεγγιστική αναπαράσταση ενός μέρους της επιφάνειας του πλανήτη Γη μέσω του μοντέλου της επίπεδης γεωμετρίας»!

Η μοντελοποίηση φαινομένων που περιέχουν στοιχεία τυχαιότητας, και η εξοικείωση με αυτήν τη διαδικασία αποτελούν δύο από τους κεντρικούς μας στόχους.

Αν και δεν είναι ο μόνος, μάλλον ο πιο συνηθισμένος τρόπος για να προσεγγίσουμε κατ' αρχήν διαισθητικά την έννοια της πιθανότητας είναι μέσω της έννοιας της «συχνότητας». Π.χ., αν στρίψουμε ένα «δίκαιο» νόμισμα N φορές και φέρουμε k φορές Κορώνα (Κ), για μεγάλα N συχνά παρατηρούμε ότι,

$$\frac{k}{N} = \text{'ποσοστό από Κ'} \approx \frac{1}{2} \text{ ή } 50\%.$$

Και όσο μεγαλώνει το πλήθος N των ρίψεων, αντιστοίχως μεγαλώνει και το πλήθος k των φορών που φέραμε Κ, έτσι ώστε, μακροπρόθεσμα,

$$\frac{k}{N} \rightarrow \frac{1}{2}, \text{ καθώς το } N \rightarrow \infty.$$

Υπό αυτή την έννοια, λέμε ότι «η πιθανότητα το νόμισμα να έρθει Κ είναι $1/2$ ».

2.2 Σύνολα

Ένα μεγάλο μέρος του μαθηματικού λεξιλογίου που θα χρησιμοποιήσουμε βασίζεται στα βασικά στοιχεία της θεωρίας συνόλων. Ξεκινάμε υπενθυμίζοντας κάποιους γνωστούς ορισμούς:

- Ορισμός 2.1 (Πράξεις συνόλων)** 1. Ένα **σύνολο** είναι μια συλλογή στοιχείων. Για παράδειγμα, τα $A = \{-1, +1\}$, $B = \{3, 5, 9\}$, $\Gamma = \mathbb{Z} = \{\dots, -1, 0, 1, 2, \dots\}$ = οι ακέραιοι αριθμοί, $\Delta = \mathbb{R}$ = οι πραγματικοί αριθμοί, $E = \{A, B, 5, \{5\}, \mathbb{R}\}$ είναι όλα σύνολα.
2. Όταν κάποιο στοιχείο α **ανήκει** σε κάποιο σύνολο A , γράφουμε $\alpha \in A$. Αν το α **δεν ανήκει** στο A , γράφουμε $\alpha \notin A$. Π.χ., πιο πάνω έχουμε, $3 \in B$, αλλά, $0 \notin A$.
3. Το A είναι **υποσύνολο** του B αν κάθε στοιχείο του A ανήκει και στο B , οπότε γράφουμε $A \subset B$ ή $A \subseteq B$.
4. Το **κενό σύνολο** \emptyset ή $\{\}$ έχει την ιδιότητα ότι δεν περιέχει κανένα στοιχείο, δηλαδή $\alpha \notin \emptyset$ για οποιοδήποτε α .

Στις πιθανότητες, ανάλογα με το πρόβλημα που θα εξετάζουμε, όλα τα σύνολα που μας ενδιαφέρουν θα είναι υποσύνολα ενός βασικού συνόλου, το οποίο συνήθως συμβολίζεται ως Ω .

5. Η **ένωση** $A \cup B$ δύο συνόλων A, B είναι το σύνολο που αποτελείται από όλα τα στοιχεία που ανήκουν στο A ή στο B (ή και στα δύο). Γενικότερα, η ένωση ενός πεπερασμένου πλήθους συνόλων A_1, A_2, \dots, A_N συμβολίζεται ως,

$$A_1 \cup A_2 \cup \dots \cup A_N = \bigcup_{i=1}^N A_i,$$

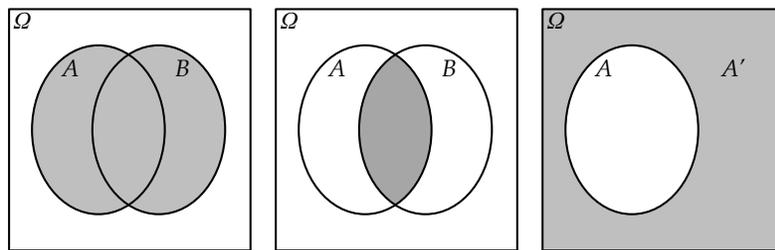
και περιέχει όλα τα στοιχεία του A_1 , τα στοιχεία του A_2 κλπ. Βλ. Σχήμα 2.1.

6. Η **τομή** $A \cap B$ δύο συνόλων A, B είναι το σύνολο που αποτελείται από όλα τα στοιχεία που ανήκουν και στο A και στο B . Γενικότερα, η τομή ενός πεπερασμένου πλήθους συνόλων A_1, A_2, \dots, A_N συμβολίζεται ως,

$$A_1 \cap A_2 \cap \dots \cap A_N = \bigcap_{i=1}^N A_i,$$

και αποτελείται από τα στοιχεία που περιέχονται σε όλα τα A_i . Βλ. Σχήμα 2.1.

7. Το **συμπλήρωμα** A' ενός συνόλου A που είναι υποσύνολο του βασικού συνόλου Ω , αποτελείται από όλα τα στοιχεία του Ω που δεν ανήκουν στο A . Βλ. Σχήμα 2.1.



Σχήμα 2.1: Γραφική αναπαράσταση της ένωσης, της τομής και του συμπληρώματος συνόλων.

Παράδειγμα 2.1 Έστω Ω το σύνολο όλων των δυνατών αποτελεσμάτων από τη ρίψη δύο νομισμάτων, δηλαδή,

$$\Omega = \{KK, ΚΓ, ΓΚ, ΓΓ\}.$$

Η περίπτωση του να φέρουμε Κ την πρώτη φορά μπορεί να περιγραφεί ως το σύνολο,

$$A = \{K \text{ την πρώτη φορά}\} = \{KK, ΚΓ\},$$

το οποίο είναι ένα υποσύνολο του Ω . Παρατηρούμε ότι το A μπορεί και να εκφραστεί ως,

$$A = \{KK\} \cup \{ΚΓ\} = \{ΓΓ, ΓΚ\}' = A \cap \Omega.$$

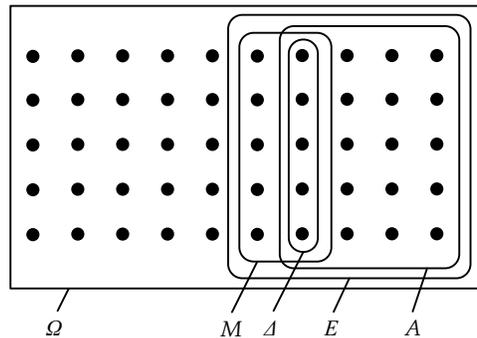
Παράδειγμα 2.2 Από 50 φοιτητές που βρίσκονται σε μια αίθουσα, οι 20 έχουν αυτοκίνητο, οι 10 έχουν μοτοσυκλέτα, και οι 25 δεν έχουν κανένα από τα δύο. Επιλέγουμε έναν φοιτητή στην τύχη.

Εδώ μπορούμε να ορίσουμε τα εξής σύνολα. Βλ. Σχήμα 2.2.

$$\begin{aligned}\Omega &= \text{Όλοι οι 50 φοιτητές} \\ A &= \text{Όσοι έχουν αυτοκίνητο} \\ M &= \text{Όσοι έχουν μοτοσυκλέτα} \\ E &= \text{Όσοι έχουν τουλάχιστον το ένα από τα δύο μέσα} \\ \Delta &= \text{Όσοι έχουν και τα δύο.}\end{aligned}$$

Θα απαντήσουμε στα εξής απλά ερωτήματα:

- (α') Πόσοι φοιτητές είναι στο E ;
- (β') Πόσοι φοιτητές είναι στο Δ ;
- (γ') Ποια είναι η πιθανότητα ο επιλεγμένος φοιτητής να έχει αυτοκίνητο;



Σχήμα 2.2: Γραφική αναπαράσταση των συνόλων στο Παράδειγμα 2.2.

Για το (α'), εφόσον είναι 50 συνολικά οι φοιτητές, δηλαδή το πλήθος των στοιχείων του Ω ισούται με 50, $\#\Omega = 50$, και αφού μας δίνεται ότι 25 φοιτητές δεν έχουν ούτε αυτοκίνητο ούτε μοτοσυκλέτα, εύκολα υπολογίζουμε ότι,

$$\begin{aligned}\#E &= \#\{\text{όσοι έχουν τουλάχιστον το ένα από τα δύο}\} \\ &= \#\{\{\text{όσοι δεν έχουν κανένα από τα δύο}\}'\} = 50 - 25 = 25,\end{aligned}$$

όπου πιο πάνω και σε ολόκληρο το βιβλίο, χρησιμοποιούμε τον συμβολισμό $\#A$ για το πλήθος των στοιχείων ενός οποιουδήποτε συνόλου A .

Για το (β'), από τη γραφική αναπαράσταση στο Σχήμα 2.2, παρατηρούμε πως,

$$\#(A \cup M) = \#A + \#M - \#\Delta,$$

όπου αφαιρούμε τα στοιχεία του συνόλου Δ για να μη μετρηθούν δύο φορές. Παρατηρούμε επίσης ότι $A \cup M = E$ και $\#E = 25$ από το (α'), ενώ μας δίνεται και ότι $\#A = 20$ και $\#M = 10$, άρα, $\#\Delta = 20 + 10 - 25 = 5$.

[*Παρένθεση.* Αν και δεν έχουμε ακόμα ορίσει την έννοια της πιθανότητας, μπορούμε να προσεγγίσουμε το ερώτημα (γ') διαισθητικά. Εφόσον η επιλογή του φοιτητή είναι τυχαία, η ζητούμενη πιθανότητα «ο επιλεγμένος φοιτητής να έχει αυτοκίνητο», δηλαδή η πιθανότητα να επιλέξουμε από όλο το Ω έναν φοιτητή που να είναι στο σύνολο A , λογικά μπορεί να υπολογιστεί ως η πιθανότητα του «ο επιλεγμένος φοιτητής να ανήκει στο A », δηλαδή, $\frac{\#A}{\#\Omega} = \frac{20}{50} = \frac{2}{5} = 0.4 = 40\%$.]

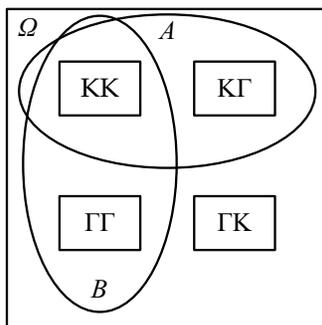
Συμβολισμός. Όπως θα δούμε στην επόμενη ενότητα, οποιαδήποτε ενδεχομένη έκβαση ενός τυχαίου πειράματος – είτε αυτή περιγράφεται περιφραστικά, π.χ., «ο επιλεγμένος φοιτητής να έχει αυτοκίνητο», είτε ως κάποιο υποσύνολο A του Ω όπως πιο πάνω – συχνά αναφέρεται απλά ως ένα *ενδεχόμενο*. Η πιθανότητα οποιουδήποτε ενδεχομένου μάς ενδιαφέρει σε κάποιο πρόβλημα συμβολίζεται ως Pr , από το αγγλικό «probability» που σημαίνει πιθανότητα.

Παράδειγμα 2.3 Ρίχνουμε ένα «δίκαιο» νόμισμα 2 φορές. Εδώ το σύνολο όλων των δυνατών αποτελεσμάτων είναι το,

$$\Omega = \{\text{ΚΚ}, \text{ΚΓ}, \text{ΓΚ}, \text{ΓΓ}\}.$$

Έστω A το ενδεχόμενο του να έρθει Κ την πρώτη φορά, και B το ενδεχόμενο να έρθει το ίδιο αποτέλεσμα δύο φορές, δηλαδή,

$$\begin{aligned} A &= \{\text{ΚΚ}, \text{ΚΓ}\} \\ B &= \{\text{ΚΚ}, \text{ΓΓ}\}. \end{aligned}$$



Σχήμα 2.3: Σχηματική αναπαράσταση των συνόλων στο Παράδειγμα 2.3.

[Παρένθεση. Και πάλι, αν και δεν έχουμε ακόμα ορίσει την έννοια της πιθανότητας, διαισθητικά μπορούμε να κάνουμε κάποιους απλούς υπολογισμούς. Παρατηρούμε ότι, εφόσον το νόμισμα είναι δίκαιο, είναι λογικό να υποθέσουμε ότι καθένα από τα τέσσερα δυνατά αποτελέσματα (που αντιστοιχούν στα 4 στοιχεία του Ω) έχουν την ίδια πιθανότητα, δηλαδή $1/4$. Άρα υπολογίζουμε εύκολα τις πιθανότητες, για παράδειγμα, των εξής ενδεχομένων:

$$\Pr(\{K \text{ την } 1\text{η φορά}\}) = \Pr(A) = \Pr(\{KK, K\Gamma\}) = \frac{\#A}{\#\Omega} = 2/4 = 1/2,$$

$$\Pr(\{\text{δύο φορές το ίδιο}\}) = \Pr(B) = \Pr(\{KK, \Gamma\Gamma\}) = \frac{\#B}{\#\Omega} = 2/4 = 1/2,$$

$$\Pr(\{\text{δύο φορές } K\}) = \Pr(\{KK\}) = \frac{\#\{KK\}}{\#\Omega} = 1/4,$$

και παρομοίως, η πιθανότητα να φέρουμε δύο φορές Γ είναι κι αυτή $1/4$.]

2.3 Χώρος πιθανότητας και ενδεχόμενα

Ορισμός 2.2 (Χώρος πιθανότητας και ενδεχόμενα)

1. Ο **χώρος πιθανότητας** ή **δειγματικός χώρος** Ω είναι το σύνολο όλων των δυνατών αποτελεσμάτων ενός τυχαίου πειράματος.
2. Οποιοδήποτε υποσύνολο $A \subset \Omega$ του χώρου πιθανότητας Ω ονομάζεται **ενδεχόμενο**.
3. Τα ενδεχόμενα που αποτελούνται από ένα μόνο στοιχείο, δηλαδή τα υποσύνολα $A \subset \Omega$ της μορφής $A = \{\omega\}$ για κάποιο $\omega \in \Omega$, λέγονται **στοιχειώδη ενδεχόμενα**.
4. Δύο ενδεχόμενα A, B είναι **ξένα** όταν δεν έχουν κανένα κοινό στοιχείο, δηλαδή αν και μόνο αν, $A \cap B = \emptyset$. Διαισθητικά, τα A, B είναι ξένα αν είναι αδύνατον να συμβούν συγχρόνως.

Παρατηρήσεις

1. Ο χώρος πιθανότητας μπορεί πάντα να εκφραστεί ως η ένωση τόσων στοιχειωδών ενδεχομένων όσα τα στοιχεία που περιέχει. Π.χ., αν $\Omega = \{\omega_1, \omega_2, \dots, \omega_N\}$, τότε,

$$\Omega = \{\omega_1\} \cup \{\omega_2\} \cup \dots \cup \{\omega_N\}.$$

Και γενικότερα, κάθε ενδεχόμενο μπορεί να εκφραστεί ως ένωση τόσων στοιχειωδών ενδεχομένων όσα τα στοιχεία που περιέχει. Επίσης σημειώνουμε πως δύο οποιαδήποτε στοιχειώδη ενδεχόμενα $\{\omega_1\}$ και $\{\omega_2\}$ είναι ξένα μεταξύ τους – αρκεί, βεβαίως, να μην είναι τα ίδια, δηλαδή το στοιχείο ω_1 να είναι διαφορετικό απ' το ω_2 .

2. Στο πιο πάνω παράδειγμα της ρίψης δύο δίκαιων νομισμάτων, ο χώρος πιθανότητας ήταν $\Omega = \{KK, K\Gamma, \Gamma K, \Gamma\Gamma\}$ και εξετάσαμε τα ενδεχόμενα $A = \{KK, K\Gamma\}$ και $B = \{KK, \Gamma\Gamma\}$, τα οποία μπορούν να εκφραστούν ως ενώσεις στοιχειωδών ενδεχομένων:

$$A = \{KK\} \cup \{K\Gamma\}, \quad B = \{KK\} \cup \{\Gamma\Gamma\}.$$

Παρατηρούμε ότι έχουμε τις πιθανότητες (όπως υπολογίστηκαν πιο πάνω), $\Pr(A) = 1/2$, $\Pr(\{KK\}) = 1/4$, και $(\Pr\{K\Gamma\}) = 1/4$. Άρα έχουμε τις «παράλληλες» σχέσεις:

$$A = \{KK\} \cup \{K\Gamma\} \quad \text{και} \quad \Pr(A) = \Pr(\{KK\}) + \Pr(\{K\Gamma\}).$$

Αργότερα θα δούμε πως, όταν κάποιο ενδεχόμενο A μπορεί να εκφραστεί ως ένωση δύο άλλων ενδεχομένων $A = B \cup \Gamma$, η μόνη περίπτωση κατά την οποία μπορούμε να είμαστε βέβαιοι ότι θα ισχύει και η αντίστοιχη σχέση για τις πιθανότητες, $\Pr(A) = \Pr(B) + \Pr(\Gamma)$ είναι όταν τα B, Γ είναι ξένα.

3. Όταν ένα ενδεχόμενο A περιγράφει την περίπτωση να συμβεί κάποιο γεγονός που μας ενδιαφέρει (π.χ. αν το A είναι το ενδεχόμενο του να φέρουμε την πρώτη φορά K ρίχνοντας ένα νόμισμα), τότε το συμπλήρωμά του A' περιγράφει το αντίθετο γεγονός, δηλαδή την περίπτωση να *μη* συμβεί το A (π.χ. πιο πάνω το A' αντιστοιχεί στο να φέρουμε την πρώτη φορά Γ).

Παρομοίως, η ένωση $A \cup B$ δύο ενδεχομένων A, B είναι το ενδεχόμενο του να συμβεί το A ή το B , και η τομή τους $A \cap B$ περιγράφει το ενδεχόμενο του να συμβούν και τα δύο.

Τέλος, παραθέτουμε κάποιες βασικές σχέσεις που ικανοποιούν οι πράξεις της ένωσης, της τομής και του συμπληρώματος συνόλων. Για οποιαδήποτε υποσύνολα A, B, Γ του Ω , έχουμε:

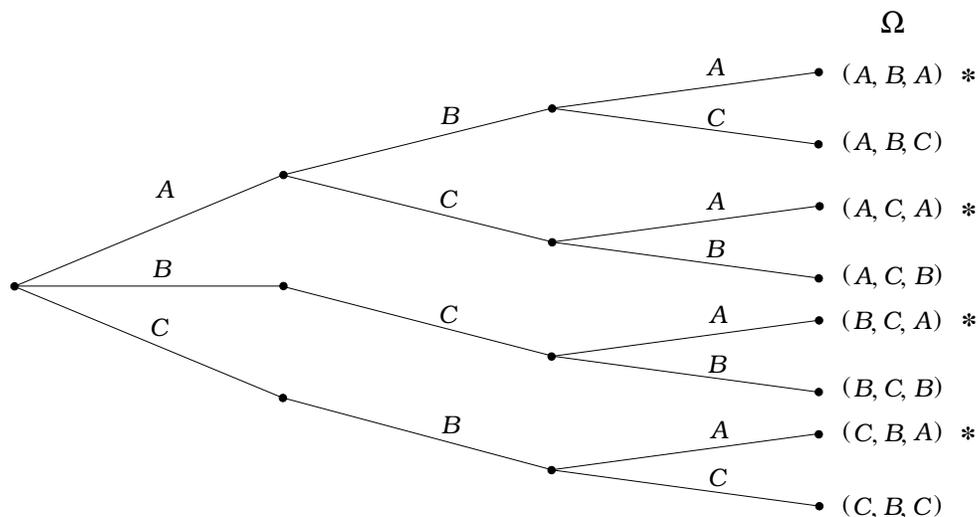
1. $\emptyset \subset A$
2. $A \cup A' = \Omega$
3. $A \cap A' = \emptyset$
4. $A \cap (B \cup \Gamma) = (A \cap B) \cup (A \cap \Gamma)$
5. $A \cup (B \cap \Gamma) = (A \cup B) \cap (A \cup \Gamma)$
6. $(A \cup B)' = A' \cap B'$
7. $(A \cap B)' = A' \cup B'$.

Κλείνουμε αυτό το κεφάλαιο με ένα ενδιαφέρον παράδειγμα το οποίο, αν και απλό, αν δεν το έχετε ξαναδεί, ίσως σας κινήσει ιδιαίτερα το ενδιαφέρον.

Παράδειγμα 2.4 (Παιχνίδι Monty Hall) Σε ένα τηλεπαιχνίδι ο διαγωνιζόμενος επιλέγει μία από τρεις κουρτίνες, αφού του πουν πως μία από αυτές κρύβει ένα δώρο και οι άλλες δύο δεν κρύβουν τίποτα (χωρίς, φυσικά, να του πουν πού είναι το δώρο). Αφού διαλέξει, ο παρουσιαστής τού ανοίγει μία από τις άλλες δύο κουρτίνες, του δείχνει ότι εκεί δεν υπάρχει τίποτα, και δίνει στον διαγωνιζόμενο τη δυνατότητα να κρατήσει την αρχική του κουρτίνα ή να διαλέξει την άλλη κουρτίνα της οποίας το περιεχόμενο παραμένει κρυφό. Ο διαγωνιζόμενος επιλέγει, και το παιχνίδι τελειώνει, είτε με νίκη του διαγωνιζόμενου (αν το δώρο βρίσκεται πίσω από την κουρτίνα της τελικής του επιλογής), είτε με ήττα του διαγωνιζόμενου (αν το δώρο δεν βρίσκεται πίσω από την κουρτίνα που επέλεξε).

Πώς μπορούμε να περιγράψουμε το χώρο πιθανότητας; Υπάρχουν διάφοροι τρόποι να περιγραφούν όλες οι δυνατές εκβάσεις του παιχνιδιού. Μια επιλογή είναι η ακόλουθη. Έστω πως ονομάζουμε κουρτίνα A την κουρτίνα όπου βρίσκεται το δώρο, και κουρτίνες B, C τις άλλες δύο. Μπορούμε να περιγράψουμε τα αποτελέσματα ως τριάδες της μορφής (X, X, X) , όπου τα X παίρνουν τιμές A, B ή C , και το πρώτο στοιχείο δείχνει την επιλογή του διαγωνιζόμενου, το δεύτερο την κουρτίνα που αποκαλύφθηκε, και το τρίτο την κουρτίνα που επέλεξε τελικά ο διαγωνιζόμενος.

Προφανώς υπάρχουν 3 επιλογές για το πρώτο στοιχείο. Αλλά για το δεύτερο στοιχείο υπάρχουν 2 επιλογές αν ο διαγωνιζόμενος έχει αρχικά επιλέξει την κουρτίνα με το δώρο, ενώ υπάρχει μόνο μία αν ο διαγωνιζόμενος έχει επιλέξει κενή κουρτίνα. Για το τρίτο στοιχείο, υπάρχουν πάντα δύο επιλογές. Ο αντίστοιχος χώρος πιθανότητας Ω περιέχει τις 8 δυνατές τριάδες (X, X, X) και έχει σχεδιαστεί στο Σχήμα 2.4.



Σχήμα 2.4: Ο χώρος πιθανότητας του Παραδείγματος 2.4. Στο πρώτο βήμα ο διαγωνιζόμενος επιλέγει μία κουρτίνα, στο δεύτερο ο παρουσιαστής ανοίγει μία από τις άλλες δύο, και στο τρίτο ο διαγωνιζόμενος αλλάζει αν θέλει την επιλογή του. Τα τέσσερα αποτελέσματα που αντιστοιχούν σε «νίκη» είναι σημειωμένα με *.

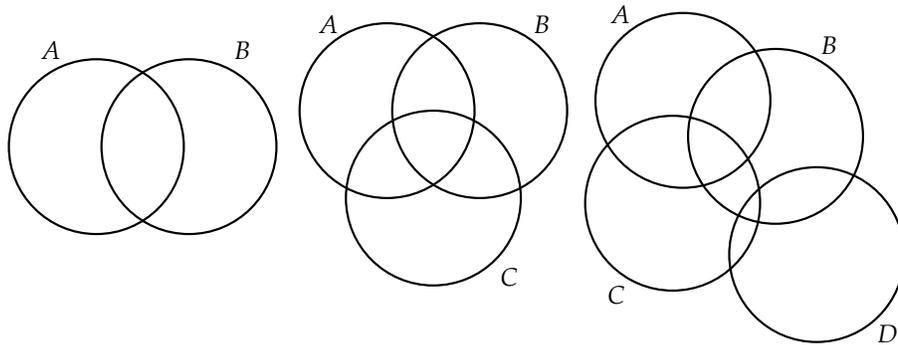
Αν θέλουμε τώρα να ορίσουμε, π.χ., το ενδεχόμενο $N = \text{«ο παίκτης κέρδισε το δώρο»}$, παρατηρούμε πως τα αποτελέσματα που καταλήγουν σε νίκη για τον διαγωνιζόμενο είναι εκείνα που έχουν τελευταίο στοιχείο το A , δηλαδή, $N = \{ABA, ACA, BCA, CBA\}$.

Σημείωση. Αυτό το παιχνίδι ήταν επί χρόνια τηλεπαιχνίδι στην Αμερική, γνωστό με το όνομα «Monty Hall». Τα βασικό ερώτημα, το οποίο θα εξετάσουμε αργότερα, είναι, «Ποια είναι η πιο συμφέρουσα στρατηγική για τον παίκτη – να κρατήσει την αρχική του κουρτίνα ή να αλλάξει;»

2.4 Ασκήσεις

1. *Τυχαία παιδιά.* Έστω πως εκτελείται το ακόλουθο πείραμα: Ένα ζευγάρι κάνει n παιδιά, καθένα εκ των οποίων μπορεί να είναι αγόρι ή κορίτσι. Περιγράψτε το χώρο πιθανότητας αυτού του πειράματος.
2. *Κι άλλα τυχαία παιδιά.* Έστω πως εκτελείται το ακόλουθο πείραμα: Ένα ζευγάρι κάνει παιδιά επ' άπειρο, μέχρι να κάνει το πρώτο κορίτσι, και μετά σταματάει. Περιγράψτε το χώρο πιθανότητας αυτού του πειράματος.
3. *Δύο διαδοχικές ζαριές.* Ρίχνουμε ένα ζάρι 2 φορές και καταγράφουμε τα δύο αποτελέσματα με τη σειρά που ήρθαν.
 - (α') Ποιος είναι ο χώρος πιθανότητας Ω ;
 - (β') Περιγράψτε τα ακόλουθα ενδεχόμενα ως υποσύνολα του Ω :
 - i. $A = \langle \text{Ζάρι 1} = \text{Ζάρι 2} \rangle$ (δηλαδή διπλές)
 - ii. $B = \langle \text{Άθροισμα 4} \rangle$
 - iii. $C = \langle \text{Πρώτο ζάρι 4} \rangle$
 - iv. $D = \langle \text{Άθροισμα 7} \rangle$
 - v. $E = \langle \text{Δεύτερο ζάρι 5} \rangle$
4. *Υπάρχουν και περίεργοι χώροι πιθανότητας.* Έστω πως ρίχνουμε ένα βελάκι σε ένα στόχο με σχήμα κύκλου, και ακτίνα 20 cm. Αν πετύχουμε το στόχο το βελάκι μένει καρφωμένο, και αν αστοχήσουμε το βελάκι πέφτει στο πάτωμα και το κλέβει ο σκύλος μας. Ορίστε το χώρο πιθανότητας Ω ώστε να περιγράφει όλα τα δυνατά αποτελέσματα, δηλαδή όλες τις θέσεις στις οποίες μπορεί να καταλήξει το βελάκι μας, συμπεριλαμβανομένου του στόματος του σκύλου!
5. *Δύο ταυτόχρονες ζαριές.* Λύστε την Άσκηση 3, υποθέτοντας πως τα ζάρια ρίχνονται ταυτόχρονα, και δεν είμαστε σε θέση να τα ξεχωρίζουμε μεταξύ τους.
6. *Τρία νομίσματα.* Ρίχνουμε τρία νομίσματα. Περιγράψτε το χώρο πιθανότητας Ω του πειράματος και τα ενδεχόμενα $A = \langle \text{τρεις φορές το ίδιο αποτέλεσμα} \rangle$, $B = \langle \text{τις πρώτες δύο φορές Γράμματα} \rangle$, $C = \langle \text{περισσότερες Κορώνες από Γράμματα} \rangle$, ως υποσύνολα του Ω .
7. *Άσπρες και μαύρες μπάλες.* Ένα κουτί περιέχει μία άσπρη μπάλα και 3 πανομοιότυπες μαύρες μπάλες.
 - (α') Επιλέγουμε μια μπάλα στην τύχη και χωρίς να την ξαναβάλουμε στο κουτί επιλέγουμε άλλη μία (δηλαδή έχουμε επιλογή χωρίς επανατοποθέτηση). Περιγράψτε το χώρο πιθανότητας Ω_1 αυτού του πειράματος.

- (β') Αν η επιλογή άσπρης μπάλας μάς δίνει κέρδος 10 ευρώ και η επιλογή μαύρης μπάλας μάς δίνει κέρδος 5 ευρώ, περιγράψτε το ενδεχόμενο συνολικά στις δύο επιλογές να κερδίσουμε 10 ευρώ.
- (γ') Αν, αφού επιλέξουμε την πρώτη μπάλα, την ξαναβάλουμε στο κουτί πριν επιλέξουμε τη δεύτερη, έχουμε επιλογή με επανατοποθέτηση, και προκύπτει ένα διαφορετικό πείραμα. Περιγράψτε το χώρο πιθανότητας Ω_2 αυτού του πειράματος, και το ενδεχόμενο συνολικά στις δύο επιλογές να κερδίσουμε 15 ευρώ.
8. Λειτουργία δικτύου. Έστω τα ενδεχόμενα $A = \text{«Σήμερα θα πέσει το δίκτυο»}$, $B = \text{«Σήμερα είναι εργάσιμη μέρα»}$, $C = \text{«Σήμερα ο τεχνικός είναι στο εργαστήριο»}$. Να εκφραστούν τα πιο κάτω ενδεχόμενα ως σύνολα, συναρτήσει των συνόλων A, B, C :
- (α') $D = \text{«Σήμερα θα πέσει το δίκτυο και είναι εργάσιμη μέρα»}$
- (β') $E = \text{«Σήμερα είναι αργία και θα πέσει το δίκτυο»}$
- (γ') $F = \text{«Σήμερα θα πέσει το δίκτυο, είναι εργάσιμη, και ο τεχνικός δεν είναι στο εργαστήριο»}$
- (δ') $G = \text{«Σήμερα ή θα πέσει το δίκτυο και είναι αργία, ή θα πέσει το δίκτυο και ο τεχνικός είναι στο εργαστήριο, ή δεν θα πέσει το δίκτυο»}$
9. Απλά διαγράμματα ενδεχομένων. Στα τρία διαγράμματα του Σχήματος 2.5, να σκιαστούν (αντιστοίχως) τα τρία ενδεχόμενα $B \cap A'$, $(A \cup B') \cap C$, $(A' \cup B' \cup C') \cap D$.



Σχήμα 2.5: Άσκηση 9.

10. Τρεις ζαριές. Ρίχνουμε ένα ζάρι 3 φορές. Περιγράψτε το χώρο πιθανότητας Ω και τα ενδεχόμενα: $A = \text{«Την 1^η και 3^η φορά ήρθε 6»}$, $B = \text{«την 1^η φορά ήρθε 1 και τη 2^η και 3^η φορά ήρθε το ίδιο αποτέλεσμα»}$ και $C = \text{«τρεις φορές ήρθε το ίδιο ζυγό αποτέλεσμα»}$.

11. *Σταθερά και κινητά τηλέφωνα.* Ένα δίκτυο τηλεφωνίας αποτελείται από 400 σταθερά τηλέφωνα και 50 κινητά. Επιλέγουμε δύο τηλέφωνα στην τύχη, όπου στην 2^η επιλογή δεν επιτρέπουμε να επιλεγεί το ίδιο τηλέφωνο με την 1^η:
- Περιγράψτε το χώρο πιθανότητας. Επιπλέον, αν η επιλογή σταθερού τηλεφώνου έχει κόστος 1 ευρώ και η επιλογή κινητού 5 ευρώ, περιγράψτε τα ενδεχόμενα $A = \text{«συνολικά οι 2 επιλογές κόστισαν 6 ευρώ»}$ και $B = \text{«συνολικά οι 2 επιλογές κόστισαν 11 ευρώ»}$.
12. *Το πρόβλημα των τριών φυλακισμένων.* Σε μια φυλακή, ο διευθυντής αποφασίζει να απονείμει χάρη σε έναν από τους τρεις φυλακισμένους (η φυλακή είναι μικρή!) και να εκτελέσει τους άλλους δύο. Ένας από τους τρεις φυλακισμένους ζητά από τον δεσμοφύλακα να του αποκαλύψει ποιος από τους άλλους δύο κρατούμενους θα εκτελεστεί, με τη λογική ότι υπάρχει πάντοτε κάποιος τέτοιος. Ο δεσμοφύλακας το κάνει, και κατόπιν του παρέχει τη δυνατότητα να αλλάξει θέση με αυτόν του οποίου την τύχη δεν αποκάλυψε. Ο φυλακισμένος έχει την επιλογή να δεχθεί ή να αρνηθεί. Να περιγράψετε το χώρο πιθανότητας αυτού του τυχαίου πειράματος.
13. *Monty Hall 2.* Επαναλάβετε το Παράδειγμα 2.4 με την ακόλουθη τροποποίηση: Οι κουρτίνες έχουν πάρει το όνομά τους πριν τοποθετηθεί το δώρο, και έτσι τα αποτελέσματα είναι τετράδες, αντί για τριάδες.
14. *Monty Hall 3.* Επαναλάβετε το Παράδειγμα 2.4 με την ακόλουθη τροποποίηση: Ο διαγωνιζόμενος δεν αλλάζει ποτέ κουρτίνα.
15. *Monty Hall 4.* Επαναλάβετε το Παράδειγμα 2.4 με την ακόλουθη τροποποίηση: Ο διαγωνιζόμενος αλλάζει πάντα κουρτίνα.

Κεφάλαιο 3

Μέτρο πιθανότητας

3.1 Ορισμός, παραδείγματα και ιδιότητες

Σ' αυτό το σύντομο κεφάλαιο θα δώσουμε, για πρώτη φορά, έναν αυστηρά μαθηματικό ορισμό της έννοιας της πιθανότητας. Αν και, εκ πρώτης όψεως, ο ορισμός φαίνεται δυσνόητος και πολύ απομακρυσμένος από αυτό που διαισθητικά ονομάζουμε «πιθανότητα», όπως θα δούμε στα παραδείγματα που ακολουθούν, στην πράξη είναι πολύ απλός και εύχρηστος.

Ορισμός 3.1 (Μέτρο πιθανότητας) Έστω ένας χώρος πιθανότητας Ω και έστω \mathcal{F} το δυναμοσύνολο του Ω , δηλαδή το σύνολο που έχει ως στοιχεία όλα τα ενδεχόμενα $A \subset \Omega$ (συμπεριλαμβανομένου και του κενού συνόλου \emptyset). Ένα **μέτρο πιθανότητας** είναι μια συνάρτηση $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ η οποία ικανοποιεί τις παρακάτω ιδιότητες:

1. Για οποιοδήποτε ενδεχόμενο $A \in \mathcal{F}$ έχουμε: $\mathbb{P}(A) \geq 0$.
2. Πάντοτε έχουμε: $\mathbb{P}(\Omega) = 1$.
3. Αν δύο ενδεχόμενα $A, B \in \mathcal{F}$ είναι ξένα (δηλαδή $A \cap B = \emptyset$), τότε,

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B).$$

Και γενικότερα, αν A_1, A_2, \dots είναι μια οποιαδήποτε (πεπερασμένη ή όχι) ακολουθία ξένων ενδεχομένων (δηλαδή $A_i \cap A_j = \emptyset$ για κάθε $i \neq j$), τότε,

$$\mathbb{P}(A_1 \cup A_2 \cup \dots) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \dots .$$

Πριν εξετάσουμε τις συνέπειες του ορισμού, ας δούμε πώς ορίζεται το μέτρο πιθανότητας σε ένα πολύ απλό παράδειγμα.

Παράδειγμα 3.1 Ρίχνουμε ένα δίκαιο ζάρι. Όλα τα δυνατά αποτελέσματα περιγράφονται από τα στοιχεία του αντίστοιχου χώρου πιθανότητας $\Omega = \{1, 2, 3, 4, 5, 6\}$. Ορίζουμε τα ενδεχόμενα,

$$\begin{aligned} A &= \text{«ζυγό αποτέλεσμα»} = \{2, 4, 6\}, \\ B &= \text{«μονό αποτέλεσμα»} = \{1, 3, 5\}, \end{aligned}$$

και τα έξι στοιχειώδη ενδεχόμενα $E_i = \{i\}$, για κάθε $i = 1, 2, 3, 4, 5, 6$. Εφόσον το ζάρι είναι δίκαιο, απαιτούμε το αντίστοιχο μέτρο πιθανότητας \mathbb{P} που περιγράφει αυτό το πείραμα να δίνει την ίδια πιθανότητα, δηλαδή $1/6$, σε κάθε δυνατό αποτέλεσμα, δηλαδή να ικανοποιεί $\mathbb{P}(E_i) = 1/6$ για κάθε $i = 1, 2, 3, 4, 5, 6$.

Για να υπολογίσουμε την πιθανότητα του ενδεχομένου A (η οποία διαισθητικά είναι προφανώς ίση με $1/2$), παρατηρούμε όπως νωρίτερα πως το A μπορεί να εκφραστεί ως ένωση στοιχειωδών ενδεχομένων, δηλαδή,

$$A = \{2\} \cup \{4\} \cup \{6\} = E_2 \cup E_4 \cup E_6,$$

και πως όλα τα στοιχειώδη ενδεχόμενα είναι ξένα μεταξύ τους. Άρα από την τρίτη ιδιότητα του ορισμού ενός μέτρου πιθανότητας έχουμε,

$$\begin{aligned} \Pr(A) &= \mathbb{P}(A) = \mathbb{P}(E_2 \cup E_4 \cup E_6) = \mathbb{P}(E_2) + \mathbb{P}(E_4) + \mathbb{P}(E_6) \\ &= 1/6 + 1/6 + 1/6 = 1/2. \end{aligned}$$

Παρομοίως υπολογίζουμε την πιθανότητα του μονού αποτελέσματος,

$$\begin{aligned} \Pr(B) &= \mathbb{P}(B) = \mathbb{P}(E_1 \cup E_3 \cup E_5) = \mathbb{P}(E_1) + \mathbb{P}(E_3) + \mathbb{P}(E_5) \\ &= 1/6 + 1/6 + 1/6 = 1/2. \end{aligned}$$

Παρατηρήσεις

1. Το πιο πάνω παράδειγμα ανήκει σε μια ευρεία κατηγορία προβλημάτων όπου έχουμε *ισοπίθανα στοιχειώδη ενδεχόμενα*, και η οποία θα εξετασθεί πιο αναλυτικά στο επόμενο κεφάλαιο.
2. Τα στοιχειώδη ενδεχόμενα είναι *πάντα* ξένα μεταξύ τους. Κατά συνέπεια, αν το Ω είναι πεπερασμένο, για να οριστεί το μέτρο πιθανότητας για όλα τα ενδεχόμενα αρκεί να οριστεί για τα στοιχειώδη ενδεχόμενα. Ο λόγος είναι απλός. Έστω ένα οποιοδήποτε ενδεχόμενο A το οποίο αποτελείται από τα στοιχεία $\omega_1, \omega_2, \dots, \omega_k$. Εκφράζοντας το A ως την ένωση των *ξένων* ενδεχομένων,

$$A = \{\omega_1\} \cup \{\omega_2\} \cup \dots \cup \{\omega_k\},$$

και χρησιμοποιώντας την τρίτη ιδιότητα του ορισμού του μέτρου πιθανότητας, μπορούμε να υπολογίσουμε την πιθανότητα του A από τις πιθανότητες των στοιχειωδών ενδεχομένων $\{\omega_i\}$ ως,

$$\mathbb{P}(A) = \mathbb{P}(\{\omega_1\}) + \mathbb{P}(\{\omega_2\}) + \dots + \mathbb{P}(\{\omega_k\}).$$

3. Όταν δύο ενδεχόμενα A, B δεν είναι ξένα, τότε η πιθανότητα της ένωσής τους γενικά δεν ισούται με το άθροισμα των επιμέρους πιθανοτήτων. Π.χ., στο πιο πάνω παράδειγμα έχουμε ότι το $A = \{2, 4, 6\} = \{2, 4, 6\} \cup \{2\} = A \cup E_2$, αλλά φυσικά $\mathbb{P}(A) \neq \mathbb{P}(A) + \mathbb{P}(E_2)$ αφού $1/2 \neq 1/2 + 1/6!$

Παράδειγμα 3.2 Σε ένα εργαστήριο πληροφορικής λειτουργούν τρία δίκτυα, από τα οποία, κατά τις τελευταίες 200 μέρες:

- 30% των ημερών, τουλάχιστον ένα δίκτυο δεν λειτουργεί,
- 10% των ημερών, ακριβώς δύο δεν λειτουργούν,
- 5% των ημερών, δεν λειτουργεί κανένα δίκτυο.

Εξετάζουμε το τι συμβαίνει μια «τυχαία» μέρα. Έστω Ω το σύνολο όλων των διακοσίων ημερών, και έστω B_i το ενδεχόμενο του να λειτουργούν ακριβώς i από τα τρία δίκτυα, για $i = 0, 1, 2, 3$. Για παράδειγμα, το B_2 είναι το σύνολο των ημερών εκείνων κατά τις οποίες δύο δίκτυα λειτουργούν κι ένα όχι.

Από τις υποθέσεις μας έχουμε ότι $\Pr(B_0) = 5\% = 0.05$ και $\Pr(B_1) = 10\% = 0.1$. Επιπλέον, η πρώτη υπόθεση μας λέει ότι $\Pr(B'_3) = 30\% = 0.3$ (γιατί;). Αλλά ποιες είναι οι πιθανότητες των B_2 και B_3 ;

Για το B_3 παρατηρούμε ότι, εφόσον το «λειτουργούν και τα τρία δίκτυα» είναι το αντίθετο του «τουλάχιστον ένα δεν λειτουργεί», διαισθητικά περιμένουμε να ισχύει ότι,

$$\Pr(\{\text{λειτουργούν και τα τρία δίκτυα}\}) = 1 - \Pr(\{\text{τουλάχιστον ένα δεν λειτουργεί}\}),$$

δηλαδή ότι,

$$\Pr(B_3) = 1 - \Pr(B'_3) = 1 - 30\% = 0.7.$$

Πράγματι, αυτή η διαισθητική υπόθεση είναι σωστή, όπως θα δούμε αμέσως μετά το παράδειγμα.

Για το B_2 τώρα, παρατηρούμε ότι το ενδεχόμενο B'_3 του να μην λειτουργεί τουλάχιστον ένα δίκτυο μπορεί να εκφραστεί ως ένωση,

$$B'_3 = B_0 \cup B_1 \cup B_2,$$

όπου τα B_0, B_1 και B_2 είναι εξ ορισμού ξένα. Άρα,

$$0.3 = \Pr(B'_3) = \Pr(B_0) + \Pr(B_1) + \Pr(B_2) = 0.05 + 0.1 + \Pr(B_2),$$

οπότε βρίσκουμε πως $\Pr(B_2) = 0.15$ ή 15%.

Λήμμα 3.1 Για οποιοδήποτε ενδεχόμενο A και οποιοδήποτε μέτρο πιθανότητας \mathbb{P} , έχουμε:

$$\mathbb{P}(A') = 1 - \mathbb{P}(A). \quad (3.1)$$

Απόδειξη: Έστω ότι το A είναι υποσύνολο του χώρου πιθανότητας Ω , όπου, από τη δεύτερη ιδιότητα του ορισμού ενός μέτρου πιθανότητας, έχουμε $\mathbb{P}(\Omega) = 1$.

Προφανώς μπορούμε να εκφράσουμε το Ω ως την ένωση $\Omega = A \cup A'$, όπου εξ ορισμού τα A και A' είναι ξένα. Άρα, από την τρίτη ιδιότητα του ορισμού του μέτρου πιθανότητας, $1 = \mathbb{P}(\Omega) = \mathbb{P}(A) + \mathbb{P}(A')$, που αποδεικνύει τη ζητούμενη σχέση (3.1). ■

Παράδειγμα 3.3 Ρίχνουμε ένα δίκαιο ζάρι 2 φορές, οπότε ο χώρος πιθανότητας Ω αποτελείται από τα 36 δυνατά αποτελέσματα:

$$\begin{aligned} \Omega = \{ & 11, 12, 13, 14, 15, 16 \\ & 21, 22, \dots, 26 \\ & \vdots \quad \quad \quad \vdots \\ & 61, 62, \dots, 66\}. \end{aligned}$$

Όπως και στο Παράδειγμα 3.1, εφόσον το ζάρι είναι δίκαιο, λογικά υποθέτουμε ότι το καθένα από τα 36 στοιχειώδη ενδεχόμενα έχει την ίδια πιθανότητα, δηλαδή $1/36$. Θα υπολογίσουμε την πιθανότητα των εξής ενδεχομένων:

$$\begin{aligned} A &= \text{«ασόδυο»} = \{12, 21\}, \\ B &= \text{«εξάρες»} = \{66\}, \\ \Gamma &= \text{«6 την πρώτη φορά»} = \{61, 62, 63, 64, 65, 66\}, \\ \Delta &= \text{«σύνολο 6»} = \{15, 24, 33, 42, 51\}. \end{aligned}$$

Για το A έχουμε, από τις πιο πάνω υποθέσεις,

$$\begin{aligned} \Pr(A) &= \Pr(\{12, 21\}) = \Pr(\{12\} \cup \{21\}) = \Pr(\{12\}) + \Pr(\{21\}) \\ &= \frac{1}{36} + \frac{1}{36} = \frac{1}{18} \approx 0.0555, \end{aligned}$$

όπου και πάλι χρησιμοποίησαμε το γεγονός ότι τα στοιχειώδη ενδεχόμενα είναι πάντοτε ξένα μεταξύ τους. Για το B απλώς έχουμε, $\Pr(B) = \Pr\{66\} = 1/36$, άρα υπάρχει διπλάσια πιθανότητα να φέρουμε ασόδυο από το να φέρουμε εξάρες.

Με την ίδια λογική, για το Γ έχουμε,

$$\begin{aligned} \Pr(\Gamma) &= \Pr(\{61\} \cup \{62\} \cup \{63\} \cup \{64\} \cup \{65\} \cup \{66\}) \\ &= \Pr(\{61\}) + \Pr(\{62\}) + \Pr(\{63\}) + \Pr(\{64\}) + \Pr(\{65\}) + \Pr(\{66\}) \\ &= 6/36 = 1/6, \end{aligned}$$

δηλαδή μόλις αποδείξαμε το διαισθητικά προφανές – ότι η πιθανότητα του να φέρουμε 6 την πρώτη φορά είναι $1/6$. Και ακολουθώντας πάλι την ίδια λογική, εύκολα υπολογίζουμε ότι, εφόσον το Δ αποτελείται από 5 στοιχεία και όλα τα στοιχειώδη ενδεχόμενα είναι ισοπίθανα, $\Pr(\Delta) = 5/36$. Αυτή η παρατήρηση ισχύει πιο γενικά, όπως διατυπώνεται στο πιο κάτω λήμμα.

Λήμμα 3.2 Αν όλα τα στοιχειώδη ενδεχόμενα ενός (πεπερασμένου) χώρου πιθανότητας Ω είναι ισοπίθανα, τότε η πιθανότητα ενός οποιουδήποτε ενδεχομένου $A \subset \Omega$ είναι ίση με:

$$\Pr(A) = \frac{\#A}{\#\Omega} = \frac{\text{πλήθος στοιχείων του } A}{\text{πλήθος στοιχείων του } \Omega}.$$

Απόδειξη: Έστω ότι ο χώρος πιθανότητας $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ αποτελείται από τα $\#\Omega = n$ στοιχεία ω_i , για $i = 1, 2, \dots, n$, έστω \mathbb{P} το μέτρο πιθανότητας, και έστω p η πιθανότητα ενός οποιουδήποτε στοιχειώδους ενδεχομένου, δηλαδή $p = \Pr\{\omega_i\}$ για κάθε i . Από τις ιδιότητες του ορισμού του μέτρου πιθανότητας έχουμε,

$$1 = \mathbb{P}(\Omega) = \mathbb{P}\left(\bigcup_{i=1}^n \{\omega_i\}\right) = \sum_{i=1}^n \mathbb{P}(\{\omega_i\}) = np,$$

άρα έχουμε $p = 1/n$.

Έστω τώρα ένα οποιουδήποτε ενδεχόμενο $A = \{\alpha_1, \alpha_2, \dots, \alpha_k\}$ που αποτελείται από $\#A = k$ στοιχεία. Τότε έχουμε,

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcup_{i=1}^k \{\alpha_i\}\right) = \sum_{i=1}^k \mathbb{P}(\{\alpha_i\}) = kp = k \frac{1}{n} = \frac{\#A}{\#\Omega}.$$

■

3.2 Πέντε «κανόνες πιθανότητας»

Από τον ορισμό του μέτρου πιθανότητας και τα δύο λήμματα που αποδείξαμε πιο πάνω, προκύπτουν κάποιες βασικές ιδιότητες τις οποίες θα χρησιμοποιούμε συχνά. Για να αναφερόμαστε σε αυτές πιο εύκολα, τις παραθέτουμε περιληπτικά στην επόμενη σελίδα.

Απόδειξη: Έστω δύο ενδεχόμενα $A \subset B$. Το B μπορεί να εκφραστεί ως $B = A \cup B_1$ όπου το B_1 αποτελείται από τα στοιχεία του B που δεν περιέχονται στο A . Άρα τα A και B_1 είναι ξένα, και συνεπώς, από την τρίτη ιδιότητα του ορισμού του μέτρου πιθανότητας, $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B_1)$, αλλά αφού όλες οι πιθανότητες είναι εξ ορισμού μεγαλύτερες ή ίσες του μηδενός, αυτό συνεπάγεται ότι $\mathbb{P}(B) \geq \mathbb{P}(A)$, και μας δίνει την Ιδιότητα 2.

Εφόσον εξ ορισμού τα \emptyset και Ω είναι ξένα, από τον ορισμό του μέτρου πιθανότητας έχουμε,

$$1 = \mathbb{P}(\Omega) = \mathbb{P}(\Omega \cup \emptyset) = \mathbb{P}(\Omega) + \mathbb{P}(\emptyset) = 1 + \mathbb{P}(\emptyset),$$

άρα $\mathbb{P}(\emptyset) = 0$. Και εφόσον $\emptyset \subset A$ για κάθε A , από την Ιδιότητα 2, $\mathbb{P}(A) \geq \mathbb{P}(\emptyset) = 0$. Παρομοίως, οποιοδήποτε A είναι υποσύνολο του Ω , άρα, $\mathbb{P}(A) \leq \mathbb{P}(\Omega) = 1$, και έχουμε αποδείξει την Ιδιότητα 1.

Η Ιδιότητα 3 είναι μέρος του ορισμού, και οι Ιδιότητες 4 και 5 προκύπτουν από το Λήμμα 3.1 και το Λήμμα 3.2 αντίστοιχα. ■

Κανόνες πιθανότητας (1–5)

Για οποιοδήποτε μέτρο πιθανότητας \mathbb{P} :

1. $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$, και $0 \leq \mathbb{P}(A) \leq 1$, για κάθε ενδεχόμενο A .
2. Αν $A \subset B$, τότε $\mathbb{P}(A) \leq \mathbb{P}(B)$.
3. Αν τα ενδεχόμενα A_1, A_2, \dots είναι ξένα (δηλαδή $A_i \cap A_j = \emptyset$ για κάθε $i \neq j$), τότε:

$$\mathbb{P}(A_1 \cup A_2 \cup \dots) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \dots .$$

4. $\mathbb{P}(A') = 1 - \mathbb{P}(A)$, για κάθε ενδεχόμενο A .
5. Αν όλα τα στοιχειώδη ενδεχόμενα είναι ισοπίθανα, τότε για κάθε ενδεχόμενο A :

$$\Pr(A) = \frac{\#A}{\#\Omega} = \frac{\text{πλήθος στοιχείων του } A}{\text{πλήθος στοιχείων του } \Omega}.$$

3.3 * Γενικός ορισμός του μέτρου πιθανότητας

Αν ο χώρος πιθανότητας είναι διακριτός (δηλαδή είτε πεπερασμένος είτε άπειρος αλλά αριθμήσιμος), τότε ο ορισμός που δώσαμε για το μέτρο πιθανότητας σε αυτό το κεφάλαιο είναι μαθηματικά πλήρης και απολύτως επαρκής για όλες τις αντίστοιχες εφαρμογές. Αλλά για την περίπτωση άπειρων και μη αριθμήσιμων χώρων πιθανότητας, ο ορισμός αυτός είναι απαραίτητο να τροποποιηθεί. Ο λόγος μπορεί να εξηγηθεί από το εξής παράδειγμα.

Έστω ότι θέλουμε να ορίσουμε την έννοια ενός «τυχαίου πραγματικού αριθμού» στο διάστημα $[0, 1]$. Σε αυτή την περίπτωση κάθε ενδεχόμενο A είναι ένα υποσύνολο του διαστήματος $[0, 1]$, και το A περιγράφει το ενδεχόμενο ο τυχαίος αυτός αριθμός να ανήκει στο A . Για παράδειγμα, αν το $A = [0, 1/2]$, λογικά θα θέλαμε να ορίσουμε ένα μέτρο πιθανότητας \mathbb{P} που να μας λέει ότι η πιθανότητα ένας τυχαίος πραγματικός αριθμός μεταξύ 0 και 1 να είναι μικρότερος ή ίσος με $1/2$, ισούται με $1/2$. Με άλλα λόγια, να έχει $\mathbb{P}(A) = 1/2$. Γενικά, θα θέλαμε για κάθε υποδιάστημα B του $[0, 1]$, η τιμή του μέτρου πιθανότητας $\mathbb{P}(B)$ να ισούται με το μήκος αυτού του διαστήματος.

Εδώ συμβαίνει κάτι πραγματικά αξιοσημείωτο και μάλλον απροσδόκητο. Μπορεί να αποδειχθεί πως είναι ΑΔΥΝΑΤΟΝ να οριστεί ένα μέτρο πιθανότητας στο $\Omega = [0, 1]$ το οποίο να ικανοποιεί τις τρεις συνθήκες του ορισμού μας και επίσης να δίνει, για κάθε υποδιάστημα $C \subset [0, 1]$,

$$\mathbb{P}(B) = \text{μήκος του } B.$$

Η βαθύτερη αιτία της δυσκολίας είναι η ύπαρξη κάποιων πολύ πολύπλοκων, κατά κάποιον τρόπο παθολογικών, υποσυνόλων B του $[0, 1]$.

Η λύση σε αυτό το πρόβλημα είναι να περιορίσουμε τα υποσύνολα του χώρου πιθανότητας Ω στα οποία απαιτούμε να ορίζεται το μέτρο πιθανότητας. Αυτή η παρατήρηση αποτελεί την αφετηρία μιας μεγάλης υποπεριοχής της μαθηματικής ανάλυσης, η οποία ονομάζεται *θεωρία μέτρου*, αλλά με την οποία δεν θα ασχοληθούμε περαιτέρω επί του παρόντος και γενικά σε αυτό το βιβλίο. Αρκεί να θυμάστε πως, όταν ο χώρος πιθανότητας δεν είναι αριθμήσιμος, υπάρχουν κάποια σπανιότατα παθολογικά ενδεχόμενα για τα οποία δεν μπορούμε να ορίσουμε την πιθανότητά τους.

Περισσότερες (πολύ περισσότερες) πληροφορίες γι' αυτό το ζήτημα και για τη θεωρία μέτρου εν γένει μπορείτε να βρείτε σε πιο προχωρημένα βιβλία μαθηματικής ανάλυσης ή πιθανοτήτων.

3.4 Ασκήσεις

1. *Η πιθανότητα της διαφοράς.* Να δείξετε, χρησιμοποιώντας τους κανόνες πιθανότητας αυτού του κεφαλαίου, ότι για οποιαδήποτε δύο ενδεχόμενα E, F :

$$\Pr(E \cap F') = \Pr(E) - \Pr(E \cap F).$$

Σημείωση. Το σύνολο $E \cap F'$ περιλαμβάνει όλα τα στοιχεία του E αν αφαιρέσουμε τα στοιχεία του F , γι' αυτό συχνά καλείται η *διαφορά του E από το F* , και συμβολίζεται $E - F$.

2. *Περίεργα ζάρια.* Έστω ένα (όχι απαραίτητα δίκαιο) ζάρι για το οποίο γνωρίζουμε ότι η πιθανότητα να έρθει 1 ή 2 είναι $1/3$, ενώ η πιθανότητα να έρθει 2 ή 3 είναι επίσης $1/3$. Ποια είναι η μέγιστη δυνατή και η ελάχιστη δυνατή τιμή για την πιθανότητα να έρθει 2; Δώστε από ένα συγκεκριμένο παράδειγμα ενός μέτρου πιθανότητας για την καθεμία από τις ακραίες δυνατές τιμές αυτής της πιθανότητας.
3. *Τυχαία συνάντηση.* Ο Σταύρος και ο Γιάννης έχουν ορίσει να συναντηθούν σε ένα μπαρ. Ο Σταύρος έχει έρθει στην ώρα του, αλλά για τον Γιάννη γνωρίζουμε ότι μπορεί να εμφανιστεί οποιαδήποτε στιγμή μέσα στις επόμενες δύο ώρες, χωρίς προτίμηση σε κάποια στιγμή ή διάστημα. Ορίζουμε τα ακόλουθα ενδεχόμενα: $A = \text{«Ο Γιάννης έρχεται εντός της πρώτης ώρας»}$, $B = \text{«Ο Γιάννης έρχεται εντός του τελευταίου μισάωρου»}$, $C = \text{«Ο Γιάννης έρχεται με τουλάχιστον μισή ώρα καθυστέρηση»}$.

(α') Ορίστε έναν κατάλληλο χώρο πιθανότητας και ένα μέτρο πιθανότητας, και υπολογίστε τις πιθανότητες των ενδεχόμενων $A, B, C, A \cap B, A \cap C, B \cap C$.

(β') Υπολογίστε τις πιθανότητες των ενδεχόμενων $A \cup B, A \cup C, B \cup C$.

4. *Ένωση τριών ενδεχόμενων.* Να δείξετε ότι, για οποιαδήποτε τρία ενδεχόμενα A, B και C ,

$$\Pr(A \cup B \cup C) = \Pr(A) + \Pr(A' \cap B) + \Pr(A' \cap B' \cap C). \quad (3.2)$$

5. *Ένας χώρος πιθανότητας με 3 στοιχεία.* Ένα τυχαίο πείραμα έχει χώρο πιθανότητας το σύνολο $\Omega = \{a, b, c\}$. Έστω πως κάποιο μέτρο πιθανότητας \mathbb{P} ικανοποιεί τις σχέσεις, $\mathbb{P}(\{a, c\}) = 9/16$ και $\mathbb{P}(\{a, b\}) = 3/4$. Χρησιμοποιήστε τους κανόνες πιθανότητας αυτού του κεφαλαίου για να υπολογίσετε τις πιθανότητες όλων των στοιχειωδών ενδεχομένων.
6. *Τι λένε οι πιθανοθεωρίστες στα παιδιά τους.* Στην Αττική, το 95% των εγκλημάτων συμβαίνει τη νύχτα, και το 54% συμβαίνει μέσα στην Αθήνα. Αν μόνο 2% των εγκλημάτων συμβαίνουν μέρα στην Αθήνα, τι ποσοστό συμβαίνει νύχτα στην Αθήνα; Τι ποσοστό συμβαίνει νύχτα έξω από την Αθήνα;
7. *Διαιρέτες.* Επιλέγουμε έναν τυχαίο αριθμό από το σύνολο $\Omega = \{1, 2, 3, \dots, 600\}$. Ποια η πιθανότητα να:

- (α') διαιρείται με το 2;
- (β') διαιρείται με το 3;
- (γ') διαιρείται και με το 2 και με το 3;
- (δ') διαιρείται με τουλάχιστον ένα από τα δύο;
- (ε') διαιρείται με το 2 αλλά όχι με το 3;

8. * Άλλη μια τυχαία συνάντηση. Ο Σταύρος και ο Γιάννης έχουν δώσει ραντεβού σε ένα μπαρ, και έχουν συμφωνήσει να συναντηθούν εντός μίας συγκεκριμένης ώρας. Καθένας όμως μπορεί να έρθει οποιαδήποτε χρονική στιγμή μέσα σε αυτή την ώρα, χωρίς να δείχνει κάποια προτίμηση σε κάποια στιγμή ή εύρος στιγμών, και χωρίς να επηρεάζεται από το τι θα κάνει ο άλλος. Μοντελοποιήστε το χώρο πιθανότητας αυτού του πειράματος, ορίστε κάποιο μέτρο πιθανότητας που να συμφωνεί με το πραγματικό πρόβλημα, και ακολούθως χρησιμοποιήστε αυτό το μέτρο για να υπολογίσετε ποια είναι η πιθανότητα να μην περιμένει ο πρώτος που θα έρθει τον δεύτερο για περισσότερο από ένα τέταρτο της ώρας. [Υπόδειξη. Μελετήστε το τετράγωνο $[0, 1] \times [0, 1] = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq 1\} \subset \mathbb{R}^2$.]

9. * Το μέτρο πιθανότητας είναι συνεχής συνάρτηση.

- (α') Έστω μια οποιαδήποτε ακολουθία ενδεχόμενων A_1, A_2, \dots σε κάποιο χώρο πιθανότητας Ω . Έστω, επιπλέον, ότι τα A_i «μεγαλώνουν», δηλαδή $A_1 \subset A_2 \subset A_3 \subset \dots$. Ορίζουμε ως A την (άπειρη) ένωσή τους, δηλαδή το σύνολο που αποτελείται από τα στοιχεία που ανήκουν σε τουλάχιστον ένα από τα A_i . Το A συμβολίζεται ως:

$$A = \bigcup_{i=1}^{\infty} A_i = \lim_{i \rightarrow \infty} A_i.$$

Να δείξετε ότι, για οποιοδήποτε μέτρο πιθανότητας \mathbb{P} ,

$$\lim_{i \rightarrow \infty} \mathbb{P}(A_i) = \mathbb{P}(\lim_{i \rightarrow \infty} A_i).$$

- (β') Έστω μια ακολουθία ενδεχόμενων B_1, B_2, \dots που μικραίνουν, δηλαδή $B_1 \supset B_2 \supset B_3 \supset \dots$, και έστω B η άπειρη τομή τους, δηλαδή το σύνολο B αποτελείται από όλα τα στοιχεία που ανήκουν σε όλα τα B_i . Το B συμβολίζεται ως:

$$B = \bigcap_{i=1}^{\infty} B_i = \lim_{i \rightarrow \infty} B_i.$$

Να δείξετε ότι, για οποιοδήποτε μέτρο πιθανότητας \mathbb{P} ,

$$\lim_{i \rightarrow \infty} \mathbb{P}(B_i) = \mathbb{P}(\lim_{i \rightarrow \infty} B_i).$$

Παρατηρήστε ότι οι πιο πάνω ιδιότητες δείχνουν ότι το όριο στο αριστερό σκέλος υπάρχει, και ότι μπορούμε να αλλάξουμε τη σειρά του ορίου και της πιθανότητας, εφόσον βέβαια ορίζεται το όριο των συνόλων. Αυτή είναι μια ιδιότητα που χαρακτηρίζει τις συνεχείς συναρτήσεις $f : \mathbb{R} \rightarrow \mathbb{R}$ και τα συνήθη όρια, αλλά η έννοια της συνέχειας μπορεί να γενικευτεί στο μέτρο πιθανότητας και τα πιο πάνω αποτελέσματα δείχνουν ότι η συνάρτηση \mathbb{P} είναι πράγματι «συνεχής».

Υπόδειξη. Για το πρώτο σκέλος, παρατηρήστε πως,

$$A = A_1 \cup (A_2 \cap A'_1) \cup (A_3 \cap A'_2) \cup \dots,$$

δηλαδή το σύνολο A μπορεί να γραφεί σαν την ένωση ξένων μεταξύ τους ενδεχόμενων. Για το δεύτερο σκέλος, χρησιμοποιήστε το πρώτο.

Κεφάλαιο 4

Πιθανότητες και συνδυαστική

Όπως είδαμε σε κάποια παραδείγματα των προηγούμενων κεφαλαίων, συχνά συναντάμε καταστάσεις όπου όλες οι δυνατές εκφάνσεις ενός τυχαίου πειράματος έχουν την ίδια πιθανότητα. Αυτά αποτελούν μια επιμέρους αλλά σημαντική κατηγορία προβλημάτων, και σε αυτό το κεφάλαιο θα δούμε πώς μπορούν να επιλυθούν εύκολα με τη χρήση κάποιων απλών αποτελεσμάτων της συνδυαστικής. Η αρετηρία μας είναι ο κανόνας πιθανότητας #5 τον οποίο είδαμε στο προηγούμενο κεφάλαιο:

Κανόνας πιθανότητας #5

Αν όλα τα στοιχειώδη ενδεχόμενα είναι ισοπίθανα, τότε, για κάθε ενδεχόμενο A :

$$\Pr(A) = \frac{\#A}{\#\Omega} = \frac{\text{πλήθος στοιχείων του } A}{\text{πλήθος στοιχείων του } \Omega}$$

Για να εφαρμοστεί αυτός ο κανόνας, προφανώς πρέπει να είμαστε σε θέση να υπολογίσουμε το πλήθος των στοιχείων που περιέχονται σε διάφορα σύνολα – συγκεκριμένα στο χώρο πιθανότητας Ω και στο ενδεχόμενο A το οποίο μας ενδιαφέρει σε κάθε περίπτωση. Η συνδυαστική είναι ο μαθηματικός τομέας που μας προσφέρει ακριβώς τα εργαλεία που χρειαζόμαστε για αυτούς τους υπολογισμούς. Πιο κάτω θα δούμε μια σειρά από σχετικά απλά αποτελέσματα της συνδυαστικής, και μέσα από παραδείγματα θα δείξουμε με ποιους τρόπους αυτά τα αποτελέσματα χρησιμοποιούνται για την απάντηση ερωτημάτων σε προβλήματα πιθανοτήτων.

4.1 Διατάξεις, συνδυασμοί, επιλογές και πιθανότητες

Ξεκινάμε υπενθυμίζοντας μια πολύ απλή ιδιότητα:

Ιδιότητα 4.1 Όταν συνδυάζονται δύο πειράματα, εκ των οποίων το πρώτο έχει N δυνατά αποτελέσματα και το δεύτερο έχει M δυνατά αποτελέσματα, τότε το νέο πείραμα έχει $M \times N$ δυνατά αποτελέσματα. Πιο αυστηρά μαθηματικά μιλώντας, αν το σύνολο A περιγράφει όλα τα δυνατά αποτελέσματα του πρώτου πειράματος και αντίστοιχα το B τα αποτελέσματα του δεύτερου, τότε όλα τα δυνατά αποτελέσματα του συνδυασμού των δύο πειραμάτων περιγράφεται από το καρτεσιανό τους γινόμενο,

$$A \times B = \{(a, b) : a \in A, b \in B\},$$

και το πλήθος των στοιχείων του προφανώς ικανοποιεί:

$$\#(A \times B) = (\#A)(\#B).$$

Παράδειγμα 4.1 (α') Εφόσον η ρίψη ενός ζαριού έχει 6 δυνατά αποτελέσματα, οι δύο διαδοχικές ρίψεις έχουν $6 \times 6 = 36$ δυνατά αποτελέσματα, οι τρεις διαδοχικές ρίψεις έχουν $6 \times 6 \times 6 = 216$ δυνατά αποτελέσματα, και γενικά οι k διαδοχικές ρίψεις έχουν 6^k δυνατά αποτελέσματα.

(β') Επιλέγουμε έναν από τους 5 υπολογιστές ενός εργαστηρίου (5 δυνατά αποτελέσματα) και αποφασίζουμε να του εγκαταστήσουμε λειτουργικό σύστημα windows ή linux (2 δυνατά αποτελέσματα). Συνολικά υπάρχουν $5 \times 2 = 10$ δυνατά αποτελέσματα.

Παράδειγμα 4.2 Τρία άτομα, ας τους πούμε A , B και Γ , τρέχουν σε έναν αγώνα 100 μέτρων. Υποθέτουμε ότι η τελική κατάταξη είναι εντελώς τυχαία. Ποια είναι η πιθανότητα να κερδίσει ο B ;

Κατ' αρχάς παρατηρούμε ότι ο χώρος πιθανότητας είναι το σύνολο όλων των δυνατών διατάξεων των A , B και Γ :

$$\Omega = \{123, 132, 213, 231, 312, 321\},$$

όπου, για παράδειγμα, το 132 μας λέει πως ο A βγήκε πρώτος, ο B τρίτος και ο Γ δεύτερος. Το ενδεχόμενο που μας ενδιαφέρει, δηλαδή το να κερδίσει ο B , αντιστοιχεί στο σύνολο $\{213, 312\}$. Εφόσον «η τελική κατάταξη είναι εντελώς τυχαία», υποθέτουμε ότι όλα τα στοιχειώδη ενδεχόμενα είναι ισοπίθανα, και από τον κανόνα πιθανότητας $\#5$ έχουμε,

$$\Pr(\text{«κέρδισε ο } B\text{»}) = \Pr(\{213, 312\}) = \frac{\#\{213, 312\}}{\#\Omega} = \frac{2}{6} = 1/3,$$

όπως είναι και διαισθητικά προφανές.

Εδώ απλώς απαριθμήσαμε όλες τις δυνατές κατατάξεις για τα 3 άτομα. Αλλά αν, αντί για τρεις, συμμετείχαν στον αγώνα 100 άνθρωποι, πόσες δυνατές κατατάξεις θα υπήρχαν; Μπορούμε να σκεφτούμε την τελική κατάταξη ως το αποτέλεσμα του συνδυασμού 100 επιμέρους «πειραμάτων»: Για την πρώτη θέση έχουμε 100 επιλογές. Έχοντας αποφασίσει ποιος είναι πρώτος, για τη δεύτερη

θέση έχουμε 99 επιλογές, κ.ο.κ. Έτσι, εφαρμόζοντας διαδοχικά την Ιδιότητα 4.1, για την τελική κατάταξη έχουμε,

$$100 \times 99 \times 98 \times \cdots \times 3 \times 2 \times 1 = 100!,$$

δυνατές κατατάξεις, δηλαδή 100! δυνατούς τρόπους που μπορούν να διαταχθούν 100 άτομα. Με το ίδιο ακριβώς σκεπτικό καταλήγουμε στο εξής συμπέρασμα:

Ιδιότητα 4.2 Υπάρχουν $N!$ δυνατές διατάξεις N αντικειμένων.

Παρατήρηση: Θυμίζουμε πως για κάθε ακέραιο αριθμό $N \geq 1$ το « N παραγοντικό» συμβολίζεται ως $N!$ και ορίζεται ως το γινόμενο $N! = N(N-1) \cdots 2 \cdot 1$. Επίσης, για λόγους ευκολίας ορίζουμε συμβατικά το $0! = 1$.

Παράδειγμα 4.3 Έστω ότι στις βουλευτικές εκλογές συμμετέχουν 42 κόμματα. Άρα υπάρχουν $42! \approx 10^{51}$ δυνατές κατατάξεις, αλλά πόσες δυνατές κατατάξεις έχουμε για τα 3 πρώτα κόμματα; Μπορούμε να σκεφτούμε το τελικό αποτέλεσμα ως το συνδυασμό τριών επιμέρους «πειραμάτων»: Για την πρώτη θέση έχουμε 42 επιλογές. Έχοντας αποφασίσει ποιο κόμμα είναι πρώτο, για τη δεύτερη θέση έχουμε 41, και παρομοίως για την τρίτη 40 επιλογές. Εφαρμόζοντας διαδοχικά την Ιδιότητα 4.1, το πλήθος των τελικών κατατάξεων για τα τρία πρώτα κόμματα είναι,

$$42 \times 41 \times 40 = 68880.$$

Γενικά, μπορούμε να ρωτήσουμε πόσες διαφορετικές διατάξεις μπορούμε να πετύχουμε, επιλέγοντας k από N αντικείμενα. Με το ίδιο σκεπτικό, έχουμε N επιλογές για το πρώτο, $(N-1)$ για το δεύτερο, κ.ο.κ., μέχρι το αντικείμενο k , για το οποίο έχουμε $(N-k+1)$ επιλογές. Άρα, από την Ιδιότητα 4.1, βρίσκουμε πως το πλήθος των τελικών διατάξεων είναι:

$$\begin{aligned} N(N-1) \cdots (N-k+2)(N-k+1) &= \frac{N(N-1) \cdots (N-k+1) \cdot (N-k)!}{(N-k)!} \\ &= \frac{N!}{(N-k)!}. \end{aligned}$$

Έχουμε έτσι αποδείξει το εξής:

Ιδιότητα 4.3 Το πλήθος όλων των δυνατών διατάξεων k αντικειμένων που επιλέγονται από N αντικείμενα ισούται με:

$$\frac{N!}{(N-k)!}.$$

Παράδειγμα 4.4 Έστω ότι έχουμε μια συνηθισμένη τράπουλα 52 φύλλων. Από την Ιδιότητα 4.2 υπάρχουν $52! \approx 8 \times 10^{67}$ δυνατές διατάξεις για τα φύλλα της τράπουλας!

Αν επιλέξουμε 3 φύλλα στην τύχη, πόσες δυνατές (διατεταγμένες) τριάδες υπάρχουν; Από την Ιδιότητα 4.3, το πλήθος τους είναι,

$$\frac{52!}{(52-3)!} = \frac{52 \times 51 \times 50 \times 49!}{49!} = 52 \times 51 \times 50 = 132600.$$

Αν υποθέσουμε τώρα ότι η επιλογή των τριών φύλλων είναι εντελώς τυχαία, ποια είναι η πιθανότητα του ενδεχομένου A να επιλέξουμε τρεις άσους; Εδώ ο χώρος πιθανότητας Ω αποτελείται από όλες τις δυνατές τριάδες φύλλων, που, όπως υπολογίσαμε, είναι $\#\Omega = 132600$, και εφόσον η επιλογή είναι εντελώς τυχαία υποθέτουμε ότι όλα τα στοιχειώδη ενδεχόμενα είναι ισοπίθανα. Ως υποσύνολο του Ω , το ενδεχόμενο A αποτελείται από όλες τις δυνατές τριάδες άσων. Η Ιδιότητα 4.3 λοιπόν μας λέει ότι, εφόσον εξετάζουμε τις διατάξεις $k = 3$ φύλλων που μπορούν να επιλεγθούν από $N = 4$ (δηλαδή από τους τέσσερις άσους), έχουμε $\#A = \frac{4!}{(4-3)!} = 4!/1! = 24$. Άρα, από τον κανόνα πιθανότητας $\#5$ έχουμε,

$$\Pr(\text{«επιλέξαμε 3 άσους»}) = \Pr(A) = \frac{\#A}{\#\Omega} = \frac{24}{132600} = \frac{1}{5525} \approx 0.02\%.$$

Παράδειγμα 4.5 Αν επιλέξουμε 5 άτομα από μια ομάδα 100 ατόμων, η Ιδιότητα 4.3 μας λέει πως υπάρχουν,

$$\frac{100!}{(100-5)!} \text{ διατάξεις δυνατών 5άδων.}$$

Αλλά αν δεν μας ενδιαφέρει η διάταξη, δηλαδή η σειρά με την οποία επιλέγουμε τα 5 άτομα, πόσες διαφορετικές πεντάδες υπάρχουν; Γενικά, πόσες ομάδες k αντικειμένων μπορούν να προκύψουν, όταν αυτά επιλέγονται από N αντικείμενα; Έστω ότι το πλήθος τους είναι x , δηλαδή υπάρχουν x μη διατεταγμένες ομάδες k αντικειμένων. Από την Ιδιότητα 4.2, κάθε τέτοια ομάδα μπορεί να διαταχθεί με $k!$ τρόπους. Άρα, το συνολικό πλήθος των διατεταγμένων ομάδων είναι $x \cdot k!$. Αλλά, από την Ιδιότητα 4.3, αυτό ισούται με $N!/(N-k)!$. Άρα έχουμε,

$$x \cdot k! = \frac{N!}{(N-k)!}, \quad \text{δηλαδή} \quad x = \frac{N!}{k!(N-k)!}.$$

Έχουμε λοιπόν αποδείξει:

Ιδιότητα 4.4 Το πλήθος όλων των δυνατών συνδυασμών (ή μη διατεταγμένων επιλογών) k αντικειμένων που επιλέγονται από N αντικείμενα ισούται με,

$$\binom{N}{k} = \frac{N!}{k!(N-k)!},$$

όπου το $\binom{N}{k}$ είναι ο συνήθης διωνυμικός συντελεστής.

Παράδειγμα 4.6 Όπως στο Παράδειγμα 4.5, επιλέγουμε 5 άτομα από 100. Αν δεν μας ενδιαφέρει η σειρά επιλογής, υπάρχουν,

$$\frac{100!}{5!(100-5)!} = \frac{100 \times 99 \times 98 \times 96 \times 96 \times 95!}{5!95!} = 75287520,$$

δυνατές πεντάδες που μπορούμε να επιλέξουμε.

Έστω τώρα ότι τα 100 άτομα αποτελούνται από 40 άνδρες και 60 γυναίκες, και ότι η επιλογή μας είναι εντελώς τυχαία. Θα εξετάσουμε τα εξής ερωτήματα:

- Πόσες πεντάδες μπορούν να σχηματιστούν με 2 άνδρες και 3 γυναίκες;
- Ποια η πιθανότητα να επιλέξουμε μόνο μία γυναίκα;
- Ποια η πιθανότητα να μην επιλέξουμε καμία γυναίκα;

Για το πρώτο ερώτημα παρατηρούμε ότι το πείραμα μπορεί να χωριστεί σε δύο μέρη. Βάσει της Ιδιότητας 4.4 μπορούμε να επιλέξουμε 2 άνδρες από τους 40 με $\binom{40}{2}$ τρόπους, και 3 γυναίκες από τις 60 με $\binom{60}{3}$ τρόπους. Άρα, ο συνδυασμός αυτών των δύο επιλογών, βάσει της Ιδιότητας 4.1, έχει,

$$\binom{40}{2} \binom{60}{3} = \frac{40!60!}{2!38!3!57!} = \frac{40 \times 39 \times 60 \times 59 \times 58}{2 \times 3 \times 2} = 26691600,$$

δυνατά αποτελέσματα.

Για τα άλλα δύο ερωτήματα, ορίζουμε το χώρο πιθανότητας Ω ως το σύνολο όλων των δυνατών (μη διατεταγμένων) επιλογών 5 ατόμων από 100 (εφόσον σε αυτό το πρόβλημα δεν μας απασχολεί η σειρά με την οποία επιλέγονται), οπότε βάσει της Ιδιότητας 4.4 βρίσκουμε όπως παραπάνω ότι $\#\Omega = \binom{100}{5} = 75287520$. Ορίζουμε επίσης και τα δύο ενδεχόμενα,

$$\begin{aligned} A &= \{\text{όλες οι πεντάδες που αποτελούνται από 4 άνδρες και μία γυναίκα}\}, \\ B &= \{\text{όλες οι πεντάδες που αποτελούνται μόνο από άντρες}\}. \end{aligned}$$

Με την ίδια συλλογιστική που χρησιμοποιήσαμε για το πρώτο ερώτημα έχουμε, από τις Ιδιότητες 4.4 και 4.1, ότι,

$$\begin{aligned} \#A &= \binom{40}{4} \binom{60}{1} = \dots = 5483400, \\ \#B &= \binom{40}{5} = \dots = 658008. \end{aligned}$$

[Παρατηρήστε πως στον υπολογισμό του $\#B$ δεν συμπεριλάβαμε την επιλογή της «καμίας γυναίκας από τις 60», αλλά αυτό δεν επηρεάζει το αποτέλεσμα διότι το πλήθος των τρόπων με τους οποίους μπορεί να γίνει αυτή η επιλογή ισούται με $\binom{60}{0} = \frac{60!}{0!60!} = 1$.]

Αφού θεωρούμε ότι η επιλογή γίνεται «εντελώς τυχαία», μπορούμε να εφαρμόσουμε τον πέμπτο κανόνα πιθανότητας, έτσι ώστε,

$$\begin{aligned}\Pr(\text{«επιλέξαμε μόνο μία γυναίκα»}) &= \Pr(A) = \frac{\#A}{\#\Omega} = \frac{5483400}{75287520} \approx 0.0728, \\ \Pr(\text{«δεν επιλέξαμε καμία γυναίκα»}) &= \Pr(B) = \frac{\#B}{\#\Omega} = \frac{658008}{75287520} \approx 0.0087.\end{aligned}$$

Παράδειγμα 4.7 Επιλέγουμε τυχαία 3 βιβλία από 10, που αποτελούνται από 5 συγγράμματα μαθημάτων και 5 εγχειρίδια (manual) υπολογιστών. Ποια η πιθανότητα να είναι όλα εγχειρίδια; Να είναι δύο εγχειρίδια κι ένα σύγγραμμα;

Ακριβώς όπως και στο προηγούμενο παράδειγμα, εφόσον δεν μας ενδιαφέρει η σειρά με την οποία επιλέγονται τα βιβλία, ορίζουμε,

$$\begin{aligned}\Omega &= \{\text{«όλες οι μη διατεταγμένες τριάδες βιβλίων»}\}, \\ A &= \{\text{«οι μη διατεταγμένες τριάδες εγχειριδίων»}\}, \\ B &= \{\text{«οι μη διατεταγμένες τριάδες με 2 εγχειρίδια και ένα σύγγραμμα»}\},\end{aligned}$$

και υπολογίζουμε,

$$\begin{aligned}\#\Omega &= \binom{10}{3} = 120, \\ \#A &= \binom{5}{3} = 10, \\ \#B &= \binom{5}{2} \binom{5}{1} = 50,\end{aligned}$$

οπότε έχουμε τις πιθανότητες,

$$\begin{aligned}\Pr(\text{«επιλέξαμε 3 manual»}) &= \Pr(A) = \frac{\#A}{\#\Omega} = \frac{10}{120} = \frac{1}{12}, \\ \Pr(\text{«επιλέξαμε 2 manual και ένα σύγγραμμα»}) &= \Pr(B) = \frac{\#B}{\#\Omega} = \frac{50}{120} = \frac{5}{12}.\end{aligned}$$

Παράδειγμα 4.8 Σε κάποιες εκλογές είναι υποψήφιοι 3 φοιτητές και 7 καθηγητές. Εκλέγονται τυχαία τρεις και ζητάμε την πιθανότητα να εκλεγούν τουλάχιστον ένας φοιτητής και τουλάχιστον ένας καθηγητής. Πάλι με την ίδια συλλογιστική όπως στα δύο παραπάνω παραδείγματα, εφόσον δεν μας ενδιαφέρει η διάταξη των τριών ατόμων, υπολογίζουμε τη ζητούμενη πιθανότητα ως,

$$\begin{aligned}&\Pr(\text{«εκλέγονται } \geq 1 \Phi \text{ και } \geq 1 K\text{»}) \\ &= \Pr(\{\text{«εκλέγονται 1 } \Phi \text{ και 2 } K\text{»}\} \cup \{\text{«εκλέγονται 2 } \Phi \text{ και 1 } K\text{»}\}) \\ &= \Pr(\text{«εκλέγονται 1 } \Phi \text{ και 2 } K\text{»}) + \Pr(\text{«εκλέγονται 2 } \Phi \text{ και 1 } K\text{»}),\end{aligned}$$

επειδή τα δύο ενδεχόμενα στη δεύτερη γραμμή παραπάνω είναι ξένα. Άρα, τελικά έχουμε,

$$\Pr(\text{«εκλέγονται } \geq 1 \Phi \text{ και } \geq 1 K\text{»}) = \frac{\binom{3}{1}\binom{7}{2}}{\binom{10}{3}} + \frac{\binom{3}{2}\binom{7}{1}}{\binom{10}{3}} = \frac{21}{40} + \frac{7}{40} = 0.7.$$

Ας υποθέσουμε τώρα πως, ανάλογα με τη σειρά εκλογής, αυτοί που εκλέγονται παίρνουν διαφορετικούς ρόλους σε μια επιτροπή – ο πρώτος γίνεται πρόεδρος, ο δεύτερος γραμματέας και ο τρίτος ταμίας. Ποια είναι η πιθανότητα να εκλεγεί φοιτητής πρόεδρος, και καθηγητές γραμματέας και ταμίας; Εφόσον εδώ μας απασχολεί και η σειρά με την οποία επιλέγονται τα «αντικείμενα» (δηλαδή τα μέλη της επιτροπής), υπολογίζουμε αυτή την πιθανότητα βάσει του χώρου πιθανότητας Ω , ο οποίος περιέχει όλες τις διατεταγμένες τριάδες, δηλαδή περιέχει $10!/(10-3)! = 720$ στοιχεία. Άρα, η πιθανότητα του ενδεχομένου που μας ενδιαφέρει, βάσει της Ιδιότητας 4.3 ισούται με:

$$\frac{\frac{3!}{(3-1)!} \cdot \frac{7!}{(7-2)!}}{720} = \frac{7}{40} \approx 0.175.$$

Παράδειγμα 4.9 Σε 3 επεξεργαστές πρέπει να καταμεληθούν 12 διεργασίες, δίνοντας 4 διεργασίες στον κάθε επεξεργαστή. Με πόσους τρόπους μπορεί να γίνει αυτός ο καταμερισμός;

Για να απαντήσουμε, χωρίζουμε το πρόβλημα σε τρία μέρη. Αρχικά επιλέγουμε 4 διεργασίες από τις 12 για τον πρώτο επεξεργαστή, πράγμα που (βάσει της Ιδιότητας 4.4) μπορεί να γίνει με $\binom{12}{4}$ τρόπους. Κατόπιν, επιλέγουμε 4 διεργασίες από τις υπόλοιπες 8 για τον δεύτερο επεξεργαστή, πράγμα που μπορεί να γίνει με $\binom{8}{4}$ τρόπους. Και τέλος οι 4 διεργασίες που απομένουν πηγαίνουν στον τρίτο επεξεργαστή. Χρησιμοποιώντας την Ιδιότητα 4.1, συνολικά αυτός ο καταμερισμός μπορεί να γίνει με:

$$\binom{12}{4} \binom{8}{4} = \frac{12!}{4!8!} \frac{8!}{4!4!} = \frac{12!}{4!4!4!} \text{ τρόπους.}$$

Στη γενική του μορφή, ακριβώς ο ίδιος συλλογισμός μάς δίνει:

Ιδιότητα 4.5 Για να μοιραστούν N αντικείμενα σε M ομάδες, όπου η πρώτη αποτελείται από k_1 αντικείμενα, η δεύτερη από k_2 αντικείμενα κ.ο.κ. ως την ομάδα M η οποία αποτελείται από k_M αντικείμενα, υπάρχουν,

$$\binom{N}{k_1 k_2 \dots k_M} = \frac{N!}{k_1! k_2! \dots k_M!} \text{ δυνατοί συνδυασμοί,}$$

όπου το $\binom{N}{k_1 k_2 \dots k_M}$ είναι ο πολυωνυμικός συντελεστής. [Δεδομένου φυσικά ότι το άθροισμα $k_1 + k_2 + \dots + k_M = N$].

Παρατήρηση: Αν έχουμε μόνο $M = 2$ ομάδες και N αντικείμενα, τότε για $k_1 = k$ αναγκαστικά θα έχουμε $k_2 = N - k$ και η Ιδιότητα 4.5 λέει πως υπάρχουν $\frac{N!}{k!(N-k)!}$ τρόποι να μοιράσουμε N αντικείμενα σε δύο ομάδες των k και $(N - k)$, αντίστοιχα. Αυτό είναι ταυτόσημο με το περιεχόμενο της Ιδιότητας 4.4, άρα η Ιδιότητα 4.5 αποτελεί γενίκευση της 4.4.

Παράδειγμα 4.10 Έχουμε 20 υπολογιστές, που αποτελούνται από 10 PC και 10 Apple, και τους μοιράζουμε τυχαία σε τρία clusters, που αποτελούνται από 10, 5 και 5 υπολογιστές αντίστοιχα. Ποιες είναι οι πιθανότητες των ενδεχομένων A και B πιο κάτω;

$$A = \{\text{«όλα τα PC στο ίδιο cluster»}\},$$

$$B = \{\text{«4 PC στο πρώτο cluster, 3 PC στο δεύτερο και 3 PC στο τρίτο»}\}.$$

Ο χώρος πιθανότητας Ω , που περιγράφει αυτό το πείραμα, αποτελείται από όλους τους δυνατούς τρόπους με τους οποίους 20 αντικείμενα μπορούν να χωριστούν σε τρεις ομάδες των 10, 5 και 5 αντικειμένων. Άρα, από την Ιδιότητα 4.5, έχουμε,

$$\#\Omega = \binom{20}{10 \ 5 \ 5} = \frac{20!}{10! \ 5! \ 5!}.$$

Επιπλέον, οι υπολογιστές κατατάσσονται σε clusters τυχαία, οπότε υποθέτουμε ότι όλα τα στοιχειώδη ενδεχόμενα είναι ισοπίθανα και θα χρησιμοποιήσουμε και πάλι τον πέμπτο κανόνα πιθανότητας για να υπολογίσουμε την πιθανότητα του A και του B .

Για το A παρατηρούμε πως «όλα τα PC στο ίδιο cluster» είναι ακριβώς ισοδύναμο με το «όλα τα PC στο πρώτο cluster». Άρα, το πλήθος των στοιχείων του A ισούται με το πλήθος των τρόπων που μπορούν τα 10 Apple να μοιραστούν σε δύο clusters με 5 το καθένα, δηλαδή $\binom{10}{5}$. Συνεπώς,

$$\Pr(A) = \Pr(\text{«όλα τα PC στο ίδιο cluster»}) = \frac{\binom{10}{5}}{\binom{20}{10 \ 5 \ 5}} = \dots \approx 10^{-6}.$$

Τέλος, για το B , έχουμε όλους τους δυνατούς τρόπους με τους οποίους τα 10 PC μπορούν να μοιραστούν σε 3 clusters με αναλογία 4-3-3, σε συνδυασμό με όλους τους τρόπους με τους οποίους τα 10 Apple μπορούν να μοιραστούν σε 3 clusters με αναλογία 6-2-2. Άρα:

$$\begin{aligned} \Pr(B) &= \Pr(\text{«4 PC στο πρώτο cluster, 3 PC στο δεύτερο και 3 PC στο τρίτο»}) \\ &= \frac{\#B}{\#\Omega} = \frac{\binom{10}{4 \ 3 \ 3} \binom{10}{6 \ 2 \ 2}}{\binom{20}{10 \ 5 \ 5}} \\ &= \frac{10! \ 10! \ 10! \ 5! \ 5!}{20! \ 4! \ 3! \ 3! \ 6! \ 2! \ 2!} = \dots \approx 0.195. \end{aligned}$$

4.2 Πέντε «κανόνες αρίθμησης»

Στο κεφάλαιο αυτό ως τώρα διατυπώσαμε κάποιες βασικές ιδιότητες της συνδυαστικής τις οποίες θα χρησιμοποιούμε συχνά. Για να αναφερόμαστε σε αυτές πιο εύκολα, τις παραθέτουμε περιληπτικά πιο κάτω.

Κανόνες αρίθμησης

1. Αν ένα πείραμα έχει N δυνατά αποτελέσματα και ένα άλλο M δυνατά αποτελέσματα, τότε ο συνδυασμός τους έχει $M \times N$ δυνατά αποτελέσματα.

2. Υπάρχουν $N!$ δυνατές διατάξεις N αντικειμένων.

3. Υπάρχουν

$$\frac{N!}{(N-k)!}$$

δυνατές διατάξεις k αντικειμένων που επιλέγονται από N .

4. Υπάρχουν

$$\binom{N}{k} = \frac{N!}{k!(N-k)!}$$

δυνατοί συνδυασμοί (ή μη διατεταγμένες επιλογές) k αντικειμένων που επιλέγονται από N αντικείμενα.

5. Υπάρχουν

$$\binom{N}{k_1 k_2 \cdots k_M} = \frac{N!}{k_1! k_2! \cdots k_M!}$$

δυνατοί συνδυασμοί βάσει των οποίων μπορούν να μοιραστούν N αντικείμενα σε M ομάδες, όπου η πρώτη αποτελείται από k_1 αντικείμενα, η δεύτερη από k_2 αντικείμενα κ.ο.κ. ως την ομάδα M η οποία αποτελείται από k_M αντικείμενα [για $k_1 + k_2 + \cdots + k_M = N$].

4.3 Ασκήσεις

1. *Μέτρημα.* Πόσες δυνατές εκδοχές υπάρχουν στο καθένα από τα παρακάτω πειράματα;
 - (α') Επιλέγουμε με τη σειρά 3 από 12 αντικείμενα, χωρίς επανατοποθέτηση.
 - (β') Στρίβουμε ένα νόμισμα 6 φορές.
 - (γ') Επιλέγουμε 20 από 100 άτομα για μια δημοσκόπηση, χωρίς επανατοποθέτηση.
 - (δ') Ρίχνουμε ένα ζάρι 7 φορές.
 - (ε') Βάζουμε 13 ανθρώπους να κάτσουν σε μια σειρά.
 - (ς') Μοιράζουμε με τη σειρά 8 φύλλα από μια συνηθισμένη τράπουλα 52 φύλλων.
 - (ζ') Ρίχνουμε ένα νόμισμα 4 φορές και ένα ζάρι 2 φορές.

2. * *Επιλογές με επανατοποθέτηση.* Ο κανόνας αρίθμησης #4 λέει πως, αν επιλέξουμε k αντικείμενα από n χωρίς επανατοποθέτηση, υπάρχουν $\binom{n}{k}$ ομάδες k αντικειμένων που μπορούμε να επιλέξουμε. Τι θα γινόταν αν επιλέγαμε k αντικείμενα ανάμεσα σε n , επιτρέποντας την επανατοποθέτηση (και χωρίς να έχει σημασία η σειρά επιλογής); Μερικά παραδείγματα:
 - (α') Για να φτιάξουμε μια πίτσα, επιλέγουμε 4 υλικά από 7 διαθέσιμα, χωρίς όμως να μας νοιάζει ποιο θα μπει πρώτο, και επιτρέπεται να βάλουμε πολλές δόσεις από κάτι, μπορούμε, για παράδειγμα, να βάλουμε διπλό τυρί.
 - (β') Πρέπει να αγοράσουμε 15 αρκουδάκια επιλέγοντας από 5 διαφορετικά είδη, και μπορούμε να αγοράσουμε πολλές φορές το ίδιο είδος.

3. *1-2-X.* Η ομάδα μας παίζει στην έδρα της 9 διαδοχικά παιχνίδια με αντίπαλες ομάδες. Με πόσους τρόπους μπορούμε να έχουμε 4 νίκες (δηλαδή «1»), 3 ισοπαλίες (δηλαδή «X»), και 2 ήττες (δηλαδή «2»);

4. *Τράπουλα.* Μοιράζουμε στην τύχη 10 φύλλα από μια συνηθισμένη τράπουλα 52 φύλλων. Ποια η πιθανότητα να περιέχει η μοιρασιά:
 - (α') κανέναν άσο;
 - (β') το πολύ τρεις άσους;

5. *Poker.* Ένας παίκτης του πόκερ παίρνει 5 φύλλα από μια κανονική τράπουλα 52 φύλλων. Ποια είναι η πιθανότητα να έχει:
 - (α') καρτέ (δηλαδή 4 ίδια φύλλα, για παράδειγμα 4 άσους ή 4 ντάμες);
 - (β') φουλ (δηλαδή ένα ζευγάρι και μία τριάδα, για παράδειγμα 3 άσους και 2 ρηγάδες);
 - (γ') χρώμα (δηλαδή όλα κούπες ή όλα σπαθιά ή όλα μπαστούνια ή όλα καρό);

6. * Ξενοδοχείο Ακρόπολις. Έξι φίλοι συμφωνούν να συναντηθούν στο ξενοδοχείο Ακρόπολις των Αθηνών. Συμβαίνει όμως να υπάρχουν 4 ξενοδοχεία με το ίδιο όνομα. Κάθε ένας από τους 6 φίλους διαλέγει στην τύχη να πάει σε ένα από αυτά.
- (α') Ποιος είναι εδώ ο χώρος πιθανότητας; Πόσα στοιχεία περιλαμβάνει;
 - (β') Ποια είναι η πιθανότητα να συναντηθούν ανά ζεύγη (εννοείται σε τρία διαφορετικά ξενοδοχεία);
 - (γ') Ποια είναι η πιθανότητα να βρεθούν δύο μόνοι τους και άλλοι τέσσερις σε δύο ζεύγη;
7. *Superleague*. Η Superleague έχει 16 ομάδες και όλες πρέπει να παίξουν με όλες, ακριβώς δύο φορές, μία φορά σε κάθε έδρα. Πόσοι αγώνες πρέπει να γίνουν συνολικά;
8. *Λόττο*. Για να κερδίσουμε το Λόττο πρέπει να προβλέψουμε 6 αριθμούς ανάμεσα στους 1, 2, 3, ..., 49, χωρίς διάταξη και χωρίς επανατοποθέτηση. Ποια είναι η πιθανότητα να κερδίσουμε παίζοντας μόνο μία στήλη;
9. *Εύκολες και δύσκολες ασκήσεις*. Σε ένα μάθημα οι φοιτητές χωρίζονται σε τρεις ομάδες, και στην κάθε ομάδα δίνονται 6 ασκήσεις οι οποίες επιλέγονται τυχαία, και χωρίς επανατοποθέτηση, από ένα σύνολο 18 ασκήσεων. Αν, από τις 18 ασκήσεις, οι 3 είναι εύκολες και οι 15 δύσκολες, ποια είναι η πιθανότητα και οι τρεις ομάδες να έχουν από μία εύκολη άσκηση;
10. *Μέτρηση αποτελεσμάτων*.
- (α') Πόσοι διαφορετικοί αναγραμματισμοί μπορούν να γίνουν με τα γράμματα της λέξης ΚΥΠΡΟΣ; Αντίστοιχα, πόσοι για τη λέξη ΣΤΑΥΡΟΣ; Για τη λέξη ΣΙΣΙΝΙ; (Οι αναγραμματισμοί δεν χρειάζεται να υπάρχουν στο λεξικό!)
 - (β') Μια ομάδα χορού περιλαμβάνει 12 γυναίκες και 7 άντρες. Με πόσους διαφορετικούς τρόπους μπορούμε να σχηματίσουμε 3 ζευγάρια, καθένα αποτελούμενο από μία γυναίκα και έναν άντρα;
 - (γ') Με την παραγγελία μιας πίτσας μπορούμε να επιλέξουμε 6 από 20 υλικά. Πόσοι συνδυασμοί υπάρχουν αν τα υλικά μπορούν να επαναλαμβάνονται; Πόσοι συνδυασμοί υπάρχουν αν δεν επαναλαμβάνονται; [Και στις δυο περιπτώσεις, δεν έχει σημασία η σειρά επιλογής των υλικών.]

Κεφάλαιο 5

Ανεξαρτησία και δεσμευμένη πιθανότητα

Ας πούμε πως ένας μετεωρολόγος μάς πληροφορεί ότι, με βάση τα ιστορικά στατιστικά στοιχεία του καιρού στην Αθήνα, βρέχει μία στις 9 μέρες. Αν για κάποιο λόγο μάς ενδιαφέρει τι καιρό κάνει τις Κυριακές (γιατί, π.χ., τις Κυριακές κάνουμε πικ-νικ στην Πάρνηθα), λογικά θα υποθέσουμε ότι μία στις 9 Κυριακές βρέχει. Αυτός ο συλλογισμός ισχύει γιατί έχουμε, έμμεσα, υποθέσει πως το αν βρέχει ή όχι σήμερα είναι *ανεξάρτητο* από το ποια μέρα της εβδομάδας είναι.

Αντίθετα, αν μας ενδιαφέρει ακριβώς τι καιρό κάνει τις μέρες που έχει συννεφιά, θα ήταν λάθος να υποθέσουμε ότι μία στις 9 συννεφιασμένες μέρες βρέχει – το ποσοστό θα είναι προφανώς μεγαλύτερο, γιατί η συννεφιά *δεν είναι ανεξάρτητη* από τη βροχή.

Το κεντρικό αντικείμενο αυτού του κεφαλαίου είναι η μαθηματική περιγραφή του πότε δύο ενδεχόμενα είναι *ανεξάρτητα*, και η μελέτη του πώς επηρεάζονται οι πιθανότητες δύο ενδεχομένων από το εάν αυτά είναι ή δεν είναι ανεξάρτητα. Παράλληλα, θα διατυπώσουμε και κάποιους ακόμα κανόνες πιθανότητας, οι οποίοι συμπληρώνουν εκείνους που είδαμε στο Κεφάλαιο 3.

5.1 Ανεξάρτητα ενδεχόμενα και δεσμευμένη πιθανότητα

Παράδειγμα 5.1 Ρίχνουμε ένα δίκαιο ζάρι 2 φορές. Όπως στο Παράδειγμα 3.3, ο χώρος πιθανότητας αποτελείται από τα 36 δυνατά αποτελέσματα των δύο ρίψεων και όλα τα στοιχειώδη ενδεχόμενα είναι ισοπίθανα. Ορίζουμε όπως πριν τα ενδεχόμενα,

$$\begin{aligned} B &= \text{«εξάρες»}, \\ \Gamma &= \text{«6 την πρώτη φορά»}, \end{aligned}$$

και επιπλέον το ενδεχόμενο $E = \text{«6 τη δεύτερη φορά»}$, όπου παρατηρούμε ότι το B μπορεί να εκφραστεί ως $B = \Gamma \cap E$.

Για τις πιθανότητες αυτών των ενδεχομένων, χρησιμοποιώντας τον πέμπτο κανόνα πιθανότητας, εύκολα υπολογίζουμε:

$$\Pr(B) = \frac{1}{36} \quad \text{και} \quad \Pr(\Gamma) = \Pr(E) = \frac{1}{6}.$$

Από αυτό το αποτέλεσμα προκύπτει η σχέση,

$$\frac{1}{36} = \Pr(B) = \Pr(\Gamma \cap E) = \Pr(\Gamma) \Pr(E) = \frac{1}{6} \cdot \frac{1}{6},$$

δηλαδή $\Pr(\Gamma \cap E) = \Pr(\Gamma) \Pr(E)$. Αυτή είναι μια πολύ σημαντική παρατήρηση, αλλά εξίσου σημαντικό είναι να παρατηρήσουμε πως αυτή η σχέση δεν ισχύει για οποιαδήποτε δύο ενδεχόμενα. Για παράδειγμα,

$$\frac{1}{36} = \Pr(B) = \Pr(B \cap \Gamma) \neq \Pr(B) \Pr(\Gamma) = \frac{1}{6} \cdot \frac{1}{36}.$$

Ορισμός 5.1 (Ανεξάρτητα ενδεχόμενα) Δύο ενδεχόμενα A και B είναι **στατιστικά ανεξάρτητα**, ή απλά, **ανεξάρτητα**, αν και μόνο αν:

$$\Pr(A \cap B) = \Pr(A) \Pr(B).$$

Ο κάθε κανόνας πιθανότητας και ο κάθε κανόνας αρίθμησης που έχουμε δει ως τώρα αντιστοιχούν σε κάποιο συγκεκριμένο μαθηματικό αποτέλεσμα (και το ίδιο ισχύει για τους κανόνες πιθανότητας #7–10, που θα δούμε στη συνέχεια). Αντίθετα, ο κανόνας πιθανότητας #6, που θα διατυπώσουμε τώρα, μας παρέχει έναν τρόπο για να «μοντελοποιούμε» κάποια προβλήματα, δηλαδή μας δίνει μια γέφυρα μεταξύ της διαισθητικής περιγραφής ενός παραδείγματος και της αντίστοιχης αυστηρά μαθηματικής διατύπωσής του.

Κανόνας Πιθανότητας #6. Όταν δύο ενδεχόμενα A, B είναι «λογικά ανεξάρτητα», δηλαδή το εάν συμβεί το ένα δεν επηρεάζει το εάν συμβεί το άλλο, τότε θεωρούμε πως είναι και (στατιστικά) ανεξάρτητα, δηλαδή πως $\Pr(A \cap B) = \Pr(A) \Pr(B)$.

Παράδειγμα 5.2 Ένα δίκτυο κινητής τηλεφωνίας εξυπηρετεί 30 συνδρομητές στο κέντρο μιας πόλης και 20 συνδρομητές στην περιφέρεια. Από τους 30 του κέντρου, οι 3 έχουν συσκευές Nokia και οι 27 Ericsson, ενώ στην περιφέρεια υπάρχουν 15 Nokia και 5 Ericsson. Επιλέγουμε τυχαία έναν συνδρομητή από την κάθε περιοχή, και εξετάζουμε τα ενδεχόμενα:

K = «το τηλ. από το κέντρο είναι Nokia»,

Π = «το τηλ. από την περιφέρεια είναι Nokia».

Εφόσον τα δύο ενδεχόμενα αντιστοιχούν σε δύο διαφορετικές επιλογές, είναι φυσικό να υποθέσουμε πως είναι λογικά ανεξάρτητα. Σε αυτό το απλό παράδειγμα, η ανεξαρτησία τους μπορεί εύκολα και να επιβεβαιωθεί μαθηματικά, χωρίς να χρειαζόμαστε τον κανόνα πιθανότητας #6.

Μια που οι επιλογές είναι τυχαίες, τα στοιχειώδη ενδεχόμενα που αντιστοιχούν σε αυτά τα δύο «πειράματα» είναι ισοπίθανα, οπότε, από τον κανόνα πιθανότητας #5, έχουμε,

$$\Pr(K) = \frac{3}{30} = \frac{1}{10}, \quad \Pr(\Pi) = \frac{15}{20} = \frac{3}{4}.$$

Αν όμως εξετάσουμε τις δύο επιλογές ως ένα πείραμα, τότε υπάρχουν συνολικά $30 \times 20 = 600$ δυνατές επιλογές, ενώ το ενδεχόμενο $K \cap \Pi$, δηλαδή να επιλέξουμε συσκευή Nokia και τις δύο φορές, περιέχει $3 \times 15 = 45$ στοιχεία. Άρα,

$$\Pr(K \cap \Pi) = \frac{45}{600} = \frac{3}{40} = \frac{1}{10} \cdot \frac{3}{4} = \Pr(K) \Pr(\Pi),$$

και συνεπώς τα K, Π είναι πράγματι ανεξάρτητα.

Παράδειγμα 5.3 Έστω ότι στρίβουμε 4 φορές ένα νόμισμα, το οποίο έρχεται Κορώνα με πιθανότητα p , για κάποιο (γνωστό σε μας) $p \in (0, 1)$. Εδώ ο χώρος πιθανότητας Ω αποτελείται από τα $2 \times 2 \times 2 \times 2 = 16$ δυνατά αποτελέσματα (π.χ., ΚΓΚΚ ή ΓΓΓΚ κ.ο.κ.) των τεσσάρων ρίψεων. Αλλά, αν το νόμισμα δεν είναι δίκαιο (δηλαδή αν το $p \neq 1/2$) τα στοιχειώδη ενδεχόμενα δεν είναι ισοπίθανα, οπότε δεν μπορούμε να χρησιμοποιήσουμε τον κανόνα πιθανότητας #5 για τους υπολογισμούς πιθανοτήτων.

Από την άλλη, λογικά μπορούμε να θεωρήσουμε ότι οι διαδοχικές ρίψεις είναι ανεξάρτητες μεταξύ τους, άρα μπορούμε να υπολογίσουμε την πιθανότητα ενός στοιχειώδους ενδεχομένου, όπως π.χ. του $\{ΚΓΚΚ\}$, εφαρμόζοντας τον κανόνα πιθανότητας #6, ως,

$$\begin{aligned} & \Pr(\{K \text{ την πρώτη φορά}\} \cap \{\Gamma \text{ τη } 2\eta\} \cap \{K \text{ την } 3\eta\} \cap \{K \text{ την } 4\eta\}) \\ &= \Pr(\{K \text{ την πρώτη φορά}\}) \cdot \Pr(\{\Gamma \text{ τη } 2\eta\}) \cdot \Pr(\{K \text{ την } 3\eta\}) \cdot \Pr(\{K \text{ την } 4\eta\}) \\ &= p \cdot (1 - p) \cdot p \cdot p = (1 - p)p^3, \end{aligned}$$

όπου θεωρούμε ότι τα τέσσερα παραπάνω ενδεχόμενα είναι (λογικά, και συνεπώς στατιστικά) ανεξάρτητα.

Υπενθυμίζουμε την Παρατήρηση 2 που ακολουθούσε το Παράδειγμα 3.1 στην αρχή του Κεφαλαίου 3, βάσει της οποίας (σε έναν διακριτό χώρο πιθανότητας), για να υπολογίσουμε την πιθανότητα οποιουδήποτε ενδεχομένου της μορφής $A = \{\omega_1, \omega_2, \dots, \omega_k\}$, αρκεί να γνωρίζουμε την πιθανότητα των στοιχειωδών ενδεχομένων $\{\omega_i\}$, οπότε:

$$\Pr(A) = \Pr(\{\omega_1\}) + \Pr(\{\omega_2\}) + \dots + \Pr(\{\omega_k\}).$$

Συνδυάζοντας αυτή την παρατήρηση με το γεγονός ότι, στο παράδειγμά μας, η ανεξαρτησία μάς επιτρέπει να υπολογίσουμε την πιθανότητα για κάθε στοιχειώδες ενδεχόμενο, καταλήγουμε στο ότι μπορούμε να υπολογίσουμε την πιθανότητα οποιουδήποτε ενδεχομένου.

Για παράδειγμα, αν $A = \text{«τέσσερις φορές το ίδιο αποτέλεσμα»}$, τότε έχουμε:

$$\begin{aligned} \Pr(A) &= \Pr(\{KKKK, ΓΓΓΓ\}) \\ &= \Pr(\{KKKK\} \cup \{ΓΓΓΓ\}) \\ &= \Pr(\{KKKK\}) + \Pr(\{ΓΓΓΓ\}) \\ &= p^4 + (1-p)^4. \end{aligned}$$

Με τον ίδιο τρόπο θα προσεγγίσουμε και το επόμενο παράδειγμα.

Παράδειγμα 5.4 Έστω ότι υπάρχουν τρεις συνδέσεις σε ένα δίκτυο, και η καθεμία ενεργοποιείται, ανεξάρτητα από τις άλλες δύο, με πιθανότητα $1/4$. Εξετάζουμε ποιες συνδέσεις είναι ενεργές μία δεδομένη στιγμή, οπότε ο χώρος πιθανότητας Ω αποτελείται από τα 8 στοιχεία που περιγράφουν τις αντίστοιχες 8 δυνατές καταστάσεις. Για παράδειγμα, το στοιχείο EEA του Ω περιγράφει την κατάσταση όπου οι δύο πρώτες συνδέσεις είναι ενεργές και η τρίτη ανενεργή.

Ποια είναι η πιθανότητα του ενδεχομένου $B = \text{«ακριβώς μία σύνδεση είναι ενεργή»}$; Εφόσον τα στοιχειώδη ενδεχόμενα εδώ δεν είναι ισοπίθανα, θα ακολουθήσουμε την ίδια λογική όπως και στο προηγούμενο παράδειγμα:

$$\begin{aligned} \Pr(B) &= \Pr(\{EAA, AEA, AAE\}) = \Pr(\{EAA\} \cup \{AEA\} \cup \{AAE\}) \\ &= \Pr(\{EAA\}) + \Pr(\{AEA\}) + \Pr(\{AAE\}). \end{aligned}$$

Και δεδομένου ότι η κάθε σύνδεση είναι ενεργή (με πιθαν. $1/4$) ή ανενεργή (με πιθαν. $3/4$) ανεξάρτητα από τις άλλες, έχουμε,

$$\Pr(B) = \frac{1}{4} \cdot \frac{3}{4} \cdot \frac{3}{4} + \frac{3}{4} \cdot \frac{1}{4} \cdot \frac{3}{4} + \frac{3}{4} \cdot \frac{3}{4} \cdot \frac{1}{4} = \frac{27}{64}.$$

Παρατήρηση: Όταν δύο ενδεχόμενα A, B είναι ανεξάρτητα, η ανεξαρτησία τους μας επιτρέπει να εκφράσουμε την πιθανότητα της τομής τους $\Pr(A \cap B)$ ως το γινόμενο των επιμέρους πιθανοτήτων τους, $\Pr(A) \cdot \Pr(B)$. Στη γενική περίπτωση όπου τα A, B μπορεί να μην είναι ανεξάρτητα, η αντίστοιχη ιδιότητα εκφράζεται με τη χρήση της δεσμευμένης πιθανότητας:

Ορισμός 5.2 (Δεσμευμένη πιθανότητα)

Για οποιαδήποτε δύο ενδεχόμενα A, B έχουμε,

$$\Pr(A \cap B) = \Pr(A) \Pr(B|A),$$

όπου η δεσμευμένη πιθανότητα του B δεδομένου του A ορίζεται αντιστοίχως ως

$$\Pr(B|A) = \frac{\Pr(A \cap B)}{\Pr(A)},$$

όποτε το ενδεχόμενο A έχει μη μηδενική πιθανότητα.

Παρατηρήσεις

1. Αφού προφανώς $\Pr(A \cap B) = \Pr(B \cap A)$ για κάθε A, B , από τον ορισμό προκύπτει πως πάντοτε έχουμε:

$$\Pr(A \cap B) = \Pr(A) \Pr(B|A) = \Pr(B) \Pr(A|B). \quad (5.1)$$

Όπως θα δούμε σε αρκετά από τα πιο κάτω παραδείγματα, η σχέση (5.1) θα μας φανεί πολύ συχνά χρήσιμη. Γι' αυτόν το λόγο καταγράφεται στην περίληψη στο τέλος του κεφαλαίου ως «κανόνας πιθανότητας #7».

2. Τα ενδεχόμενα A, B είναι εξ ορισμού ανεξάρτητα αν και μόνο αν $\Pr(A \cap B) = \Pr(A) \Pr(B)$. Αλλά από τη σχέση (5.1) βλέπουμε πως αυτό ισχύει αν και μόνο αν $\Pr(A|B) = \Pr(A)$, ή, ισοδύναμα, αν και μόνο αν $\Pr(B|A) = \Pr(B)$.
3. Γενικά $\Pr(A|B) \neq \Pr(B|A)$, όπως, λόγου χάρη, στο αμέσως επόμενο παράδειγμα.

Παράδειγμα 5.5 Μια εταιρία πληροφορικής απασχολεί 40 Έλληνες και 30 αλλοδαπούς εργαζόμενους, εκ των οποίων κάποιοι είναι τεχνικοί και κάποιοι προγραμματιστές:

	Έλληνες	αλλοδαποί
τεχνικοί	22	25
προγραμματιστές	18	5

Επιλέγουμε τυχαία έναν από τους 70 εργαζομένους, και εξετάζουμε τα ενδεχόμενα:

$$\begin{aligned} T &= \text{«επελέγη τεχνικός»}, \\ E &= \text{«επελέγη Έλληνας»}. \end{aligned}$$

Εφόσον η επιλογή είναι τυχαία, από τον κανόνα πιθανότητας #5 έχουμε,

$$\Pr(T) = \frac{22 + 25}{70} = \frac{47}{70} \quad \text{και} \quad \Pr(E) = \frac{40}{70} = \frac{4}{7}.$$

Παρομοίως, η πιθανότητα να επιλέξουμε έναν Έλληνα τεχνικό είναι,

$$\Pr(E \cap T) = \frac{22}{70},$$

το οποίο προφανώς δεν ισούται με $(47/70) \times (4/7)$, άρα τα ενδεχόμενα E και T δεν είναι ανεξάρτητα.

Έστω τώρα πως γνωρίζουμε ότι το άτομο που επελέγη είναι τεχνικός. Ποια η πιθανότητα να είναι Έλληνας; Από τον ορισμό της δεσμευμένης πιθανότητας, αυτό ισούται με:

$$\Pr(E|T) = \frac{\Pr(E \cap T)}{\Pr(T)} = \frac{22/70}{47/70} = \frac{22}{47}.$$

Σε αυτό το παράδειγμα μπορούμε να ελέγξουμε αν το αποτέλεσμα συμφωνεί με τη διαίσθησή μας για το τι θα πει «πιθανότητα ενός ενδεχομένου δεδομένου ενός άλλου» ως εξής: Εφόσον γνωρίζουμε πως το επιλεγμένο άτομο είναι τεχνικός, διαισθητικά το πείραμά μας είναι ισοδύναμο με μια τυχαία επιλογή μεταξύ των $22 + 25 = 47$ τεχνικών. Και εφόσον από αυτούς τους 47 οι 22 είναι Έλληνες, λογικά θα περιμέναμε η πιθανότητα του να επιλέξουμε έναν Έλληνα να ισούται με $22/47$, το οποίο πράγματι επιβεβαιώνεται από τον προηγούμενό μας υπολογισμό, ο οποίος έγινε βάσει του ορισμού.

Τέλος, μπορούμε να εξετάσουμε την «αντίστροφη» περίπτωση: Ποια είναι η πιθανότητα να επιλέξαμε τεχνικό δεδομένου ότι επιλέξαμε κάποιον Έλληνα; Όπως και πριν, από τον ορισμό της δεσμευμένης πιθανότητας βρίσκουμε,

$$\Pr(T|E) = \frac{\Pr(E \cap T)}{\Pr(E)} = \frac{22/70}{40/70} = \frac{22}{40} = \frac{11}{20}.$$

Στο επόμενο λήμμα καταγράφουμε μια απλή ιδιότητα της δεσμευμένης πιθανότητας, η οποία είναι ανάλογη του κανόνα πιθανότητας #4 για τις απλές πιθανότητες που είδαμε στο Κεφάλαιο 3.

Λήμμα 5.1 Για οποιαδήποτε δύο ενδεχόμενα A, B , έχουμε:

$$\Pr(A'|B) = 1 - \Pr(A|B).$$

Απόδειξη: Ξεκινώντας από τον ορισμό της δεσμευμένης πιθανότητας, βρίσκουμε,

$$\begin{aligned} \Pr(A'|B) &= \frac{\Pr(A' \cap B)}{\Pr(B)} \\ &= \frac{\Pr(A' \cap B) + \Pr(A \cap B) - \Pr(A \cap B)}{\Pr(B)} \\ &= \frac{\Pr((A' \cap B) \cup (A \cap B)) - \Pr(A \cap B)}{\Pr(B)} \\ &= \frac{\Pr(B) - \Pr(A \cap B)}{\Pr(B)} \\ &= 1 - \frac{\Pr(A \cap B)}{\Pr(B)} \\ &= 1 - \Pr(A|B), \end{aligned}$$

όπου στο πρώτο βήμα εφαρμόσαμε τον Ορισμό 5.2, στο τρίτο βήμα χρησιμοποιήσαμε τον κανόνα πιθανότητας #3, στο τελευταίο βήμα εφαρμόσαμε τον Ορισμό 5.2, και στο τέταρτο βήμα χρησιμοποιήσαμε την προφανή ιδιότητα ότι το ενδεχόμενο B μπορεί να εκφραστεί ως «τα στοιχεία του B που ανήκουν στο A μαζί με τα στοιχεία του B που δεν ανήκουν στο A », δηλαδή, $B = (A \cap B) \cup (A' \cap B)$, και το γεγονός ότι τα ενδεχόμενα $(A \cap B)$ και $(A' \cap B)$ είναι εξ ορισμού ξένα. ■

Παράδειγμα 5.6¹ Έστω πως, σε έναν πληθυσμό 10 εκατομμυρίων ανθρώπων, 20,000 άτομα είναι φορείς του ιού HIV. Επιλέγεται ένα άτομο τυχαία και του γίνεται μια εξέταση για HIV, η οποία έχει ποσοστό σφάλματος 5%, δηλαδή,

$$\begin{aligned} \Pr(\text{αποτέλεσμα εξέτασης θετικό} \mid \text{όχι φορέας του HIV}) &= 0.05, \\ \text{και } \Pr(\text{αποτέλεσμα εξέτασης αρνητικό} \mid \text{φορέας του HIV}) &= 0.05. \end{aligned}$$

Η σημαντική ερώτηση εδώ για τον εξεταζόμενο είναι: Αν το αποτέλεσμα της εξέτασης είναι θετικό (δηλαδή υποστηρίζει πως ο εξεταζόμενος είναι φορέας του ιού), ποια είναι η πιθανότητα ο εξεταζόμενος να είναι πράγματι φορέας;

Ίσως φαίνεται εκ πρώτης όψεως «προφανές» πως η απάντηση είναι 95%, αλλά, όπως θα δούμε, η σωστή απάντηση είναι πολύ διαφορετική.

Για να προσεγγίσουμε το πρόβλημα συστηματικά, ορίζουμε τα ενδεχόμενα,

$$\begin{aligned} \Theta &= \text{«αποτέλεσμα εξέτασης θετικό»}, \\ \text{και } H &= \text{«ο εξεταζόμενος είναι φορέας του HIV»}, \end{aligned}$$

και καταγράφουμε τα δεδομένα του προβλήματος:

$$\Pr(H) = \frac{20000}{10000000} = \frac{2}{1000} = 0.002, \quad \Pr(\Theta|H') = \Pr(\Theta'|H) = 0.05. \quad (5.2)$$

Η ζητούμενη πιθανότητα είναι η $\Pr(H|\Theta)$.

Ξεκινάμε υπολογίζοντας την πιθανότητα του ενδεχομένου Θ , δηλαδή του να βγει θετικό το αποτέλεσμα της εξέτασης ενός τυχαία επιλεγμένου ατόμου. Εκφράζοντας το Θ ως την ένωση δύο ξένων ενδεχομένων, $\Theta = (\Theta \cap H) \cup (\Theta \cap H')$, έχουμε,

$$\Pr(\Theta) = \Pr(\Theta \cap H) + \Pr(\Theta \cap H'),$$

και χρησιμοποιώντας τη σχέση (5.1), που προέκυψε από τον ορισμό της δεσμευμένης πιθανότητας,

$$\begin{aligned} \Pr(\Theta) &= \Pr(\Theta|H) \Pr(H) + \Pr(\Theta|H') \Pr(H') \\ &= (1 - \Pr(\Theta'|H)) \Pr(H) + \Pr(\Theta|H')(1 - \Pr(H)), \end{aligned}$$

όπου στο δεύτερο βήμα χρησιμοποιήσαμε το Λήμμα 5.1 και τον κανόνα πιθανότητας #4. Αντικαθιστώντας τις τιμές των πιθανοτήτων (5.2),

$$\Pr(\Theta) = (1 - 0.05) 0.002 + 0.05 (1 - 0.002) \approx 0.0518. \quad (5.3)$$

¹Στην πρώτη του ανάγνωση, αυτό το παράδειγμα ίσως φαίνεται δυσκολότερο και πιο πολύπλοκο από τα προηγούμενα. Όπως θα δούμε πιο κάτω σε αυτό το κεφάλαιο, αποτελεί ειδική περίπτωση της χρήσης ενός πολύ σημαντικού αποτελέσματος που θα διατυπώσουμε λεπτομερώς στον κανόνα πιθανότητας #10, τον λεγόμενο «κανόνα του Bayes».

Τώρα είμαστε σε θέση να απαντήσουμε το βασικό μας ερώτημα: Δεδομένου ότι η εξέταση βγήκε θετική, ποια η πιθανότητα να είναι φορέας του HIV ο εξεταζόμενος; Από τον ορισμό της δεσμευμένης πιθανότητας και τη σχέση (5.1), με τον ίδιο συλλογισμό όπως προηγουμένως βρίσκουμε:

$$\Pr(H|\Theta) = \frac{\Pr(\Theta \cap H)}{\Pr(\Theta)} = \frac{\Pr(\Theta|H) \Pr(H)}{\Pr(\Theta)} = \frac{[1 - \Pr(\Theta'|H)] \Pr(H)}{\Pr(\Theta)}.$$

Τέλος, αντικαθιστώντας τις τιμές των πιθανοτήτων από τα δεδομένα (5.2) και τον υπολογισμό (5.3), έχουμε,

$$\Pr(H|\Theta) = \frac{(1 - 0.05) 0.002}{0.0518} \approx 0.0367 = 3.67\%.$$

Βλέπουμε, λοιπόν, πως η πιθανότητα ο εξεταζόμενος να είναι φορέας του HIV δεδομένου πως το αποτέλεσμα της εξέτασης είναι θετικό, είναι *σημαντικά μικρότερη* από την πρώτη μας διασθητική απάντηση που ήταν βασισμένη στο σκεπτικό ότι για το 95% των περιπτώσεων η εξέταση δίνει το σωστό αποτέλεσμα! Η διαφορά αυτή προκύπτει από το ότι αρχικά δεν λάβαμε υπόψη μας πως ο εξεταζόμενος επιλέχθηκε *τυχαία* από έναν πληθυσμό στον οποίο πολύ σπάνια συναντάμε έναν φορέα του HIV. Η αρχική πιθανότητα (2 στους χίλιους) να είναι φορέας φυσικά αυξάνεται (στο 3.76%) λόγω του θετικού αποτελέσματος της εξέτασης, αλλά δεν φτάνει ως το 95% όπως αρχικά φανταζόμασταν.

5.2 Περαιτέρω ιδιότητες

Παράδειγμα 5.7 Σε ένα εργαστήριο υπάρχουν 150 PC, εκ των οποίων τα 120 είναι συνδεδεμένα στο internet, τα 45 είναι συνδεδεμένα με ένα δίκτυο εκτυπωτών, και 30 είναι και στα δύο δίκτυα. Επιλέγουμε ένα PC στην τύχη:

- (i) Αν διαπιστώσουμε ότι είναι συνδεδεμένο με το δίκτυο εκτυπωτών, ποια η πιθανότητα να είναι και στο internet;
- (ii) Ποια η πιθανότητα να είναι συνδεδεμένο σε τουλάχιστον ένα από τα δύο δίκτυα;

Εδώ ο χώρος πιθανότητας Ω αποτελείται από τα 150 PC, και τα ενδεχόμενα που μας ενδιαφέρουν για το πρώτο ερώτημα είναι τα I =«το επιλεγμένο PC είναι στο internet» και E =«το επιλεγμένο PC είναι στο δίκτυο εκτυπωτών». Εφόσον η επιλογή είναι τυχαία, εφαρμόζοντας τον κανόνα πιθανότητας #5 βρίσκουμε τις πιθανότητες $\Pr(I) = 120/150 = 4/5$, $\Pr(E) = 45/150 = 3/10$ και $\Pr(I \cap E) = 30/150 = 1/5$.

Για το πρώτο ερώτημα, από τον ορισμό της δεσμευμένης πιθανότητας έχουμε:

$$\Pr(I|E) = \frac{\Pr(I \cap E)}{\Pr(E)} = \frac{1/5}{3/10} = \frac{2}{3}.$$

[Σε αυτό το απλό παράδειγμα, να ελέγξουμε την απάντηση και διασθητικά: Εφόσον η επιλογή μας γίνεται «δεδομένου ότι το PC έχει εκτυπωτή», είναι σαν να ζητάμε την πιθανότητα να επιλέξουμε,

από τα 45 PC που έχουν εκτυπωτή, ένα από εκείνα τα 30 που είναι και στο internet, συνεπώς η ζητούμενη πιθανότητα θα ισούται με $30/45 = 2/3$.]

Για το ερώτημα (ii), έχουμε,

$$\Pr(I \cup E) = \frac{\#(I \cup E)}{\#\Omega}.$$

Για να υπολογίσουμε το πλήθος των στοιχείων του συνόλου $I \cup E$, παρατηρούμε πως ισχύει ότι,

$$\#I + \#E = \#(I \cup E) + \#(I \cap E),$$

δηλαδή προσθέτοντας τα στοιχεία του I και του E έχουμε τα στοιχεία της ένωσής τους, αλλά τα στοιχεία που ανήκουν και στα δύο σύνολα (δηλαδή στην τομή τους) συμπεριλαμβάνονται δύο φορές στο $\#I + \#E$, γι' αυτόν το λόγο τα προσθέτουμε άλλη μία φορά στο δεξί μέρος της πιο πάνω σχέσης. Άρα έχουμε,

$$\begin{aligned} \Pr(I \cup E) &= \frac{\#(I \cup E)}{\#\Omega} \\ &= \frac{\#I + \#E - \#(I \cap E)}{\#\Omega} \\ &= \frac{\#I}{\#\Omega} + \frac{\#E}{\#\Omega} - \frac{\#(I \cap E)}{\#\Omega} \\ &= \Pr(I) + \Pr(E) - \Pr(I \cap E) \\ &= \frac{120}{150} + \frac{45}{150} - \frac{30}{150} = \frac{2}{3}. \end{aligned}$$

Παρατηρούμε πως ο πιο πάνω υπολογισμός αποδεικνύει και τη σχέση,

$$\Pr(I \cup E) = \Pr(I) + \Pr(E) - \Pr(I \cap E),$$

η οποία ισχύει και γενικά:

Κανόνας Πιθανότητας #8. Για οποιαδήποτε ενδεχόμενα A, B :

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B). \quad (5.4)$$

Απόδειξη: Αρκεί απλά να εφαρμόσουμε δύο φορές τον κανόνα πιθανότητας #3: Η ένωση $A \cup B$ μπορεί να εκφραστεί ως η ένωση των δύο ξένων ενδεχομένων $A \cup (A' \cap B)$, οπότε,

$$\Pr(A \cup B) + \Pr(A \cap B) = \Pr(A \cup (A' \cap B)) + \Pr(A \cap B) = \Pr(A) + \Pr(A' \cap B) + \Pr(A \cap B),$$

και αφού εξ ορισμού τα $(A' \cap B)$ και $(A \cap B)$ είναι ξένα,

$$\Pr(A \cup B) + \Pr(A \cap B) = \Pr(A) + \Pr((A' \cap B) \cup (A \cap B)) = \Pr(A) + \Pr(B),$$

που μας δίνει την (5.4). ■

Παρατήρηση: Αν τα ενδεχόμενα A, B είναι ξένα, τότε $A \cap B = \emptyset$, και αφού εξ ορισμού $\Pr(\emptyset) = 0$, η σχέση (5.4) γίνεται, $\Pr(A \cup B) = \Pr(A) + \Pr(B)$. Άρα μπορούμε να θεωρήσουμε τον κανόνα πιθανότητας #8 ως μια γενίκευση του κανόνα πιθανότητας #3, τον οποίο είδαμε στο Κεφάλαιο 3.

Παράδειγμα 5.8 Από ένα στοκ 200 routers, μεταξύ των οποίων 3 είναι ελαττωματικοί, επιλέγουμε τυχαία 20 για ένα δίκτυο. Ποια είναι η πιθανότητα να επιλέξουμε ακριβώς έναν ελαττωματικό;

Εστω A το ενδεχόμενο που περιγράφει την πιο πάνω περίπτωση. Το A είναι υποσύνολο του χώρου πιθανότητας Ω , ο οποίος αποτελείται από όλες τις (μη διατεταγμένες) εικοσάδες routers τις οποίες μπορούμε να διαλέξουμε από τους 200. Από τον κανόνα αρίθμησης #4 έχουμε $\#\Omega = \binom{200}{20}$, και συνδυάζοντας τους κανόνες αρίθμησης #1 και #4 βρίσκουμε $\#A = \binom{3}{1} \binom{197}{19}$. Όπως στο Κεφάλαιο 4, χρησιμοποιώντας τον κανόνα πιθανότητας #5 και αντικαθιστώντας τις πιο πάνω τιμές, έχουμε,

$$\Pr(A) = \frac{\#A}{\#\Omega} = \frac{\binom{3}{1} \binom{197}{19}}{\binom{200}{20}} = \dots \approx 0.245.$$

Τώρα ας υποθέσουμε ότι παρατηρούμε το δίκτυο να λειτουργεί προβληματικά (άρα έχουμε επιλέξει τουλάχιστον έναν ελαττωματικό router). Ποια είναι η πιθανότητα να έχουμε επιλέξει ακριβώς έναν ελαττωματικό router; Αν ορίσουμε το ενδεχόμενο,

$$B = \text{«επιλέξαμε τουλάχιστον έναν ελαττωματικό router»},$$

τότε η ζητούμενη πιθανότητα είναι,

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A)}{\Pr(B)},$$

αφού το $A \cap B = A$ (γιατί;). Η $\Pr(A)$ υπολογίστηκε πιο πάνω, και παρομοίως έχουμε,

$$\begin{aligned} \Pr(B) &= 1 - \Pr(B') \\ &= 1 - \Pr(\text{«δεν επιλέξαμε κανέναν ελαττωματικό router»}) \\ &= 1 - \frac{\#B'}{\#\Omega} \\ &= 1 - \frac{\binom{197}{20}}{\binom{200}{20}} \\ &\approx 0.2722, \end{aligned}$$

συνεπώς,

$$\Pr(A|B) = \frac{\Pr(A)}{\Pr(B)} \approx \frac{0.245}{0.2722} \approx 0.9 = 90\%.$$

Σημείωση. Ο κανόνας πιθανότητας #3 του Κεφαλαίου 3 μάς επιτρέπει να υπολογίσουμε την πιθανότητα $\Pr(A \cup B)$ για δύο ξένα ενδεχόμενα A, B , και ο κανόνας πιθανότητας #8, που είδαμε πιο πάνω, μάς δίνει ένα γενικότερο αποτέλεσμα που ισχύει και όταν τα A, B δεν είναι ξένα. Το επόμενο μας αποτέλεσμα είναι μια σχετική ανισότητα, η οποία είναι πολύ χρήσιμη και επίσης ισχύει γενικά. Για παραδείγματα εφαρμογών του δείτε την Ενότητα 8.3 του Κεφαλαίου 8 και τις αντίστοιχες εκεί ασκήσεις.

Λήμμα 5.2 (Φράγμα ένωσης) Για οποιαδήποτε ενδεχόμενα A_1, A_2, \dots, A_N (είτε είναι ξένα, είτε όχι) ισχύει η ανισότητα:

$$\begin{aligned} \Pr(A_1 \cup A_2 \cup \dots \cup A_N) &= \Pr\left(\bigcup_{i=1}^N A_i\right) \\ &\leq \sum_{i=1}^N \Pr(A_i) = \Pr(A_1) + \Pr(A_2) + \dots + \Pr(A_N). \end{aligned}$$

Απόδειξη: Παρατηρούμε ότι η περίπτωση $N = 2$ προκύπτει άμεσα από τον κανόνα πιθανότητας #8: Εφόσον όλες οι πιθανότητες είναι μεγαλύτερες ή ίσες του μηδενός,

$$\Pr(A_1 \cup A_2) = \Pr(A_1) + \Pr(A_2) - \Pr(A_1 \cap A_2) \leq \Pr(A_1) + \Pr(A_2).$$

Για την περίπτωση $N = 3$, χρησιμοποιούμε την περίπτωση $N = 2$ δύο φορές:

$$\begin{aligned} \Pr(A_1 \cup A_2 \cup A_3) &= \Pr(A_1 \cup (A_2 \cup A_3)) \\ &\leq \Pr(A_1) + \Pr(A_2 \cup A_3) \\ &\leq \Pr(A_1) + \Pr(A_2) + \Pr(A_3). \end{aligned}$$

Η γενική περίπτωση αποδεικνύεται με τον ίδιο τρόπο, επαγωγικά. ■

Παράδειγμα 5.9 Έστω ότι έχουμε δύο νομίσματα: ένα δίκαιο με $\Pr(K) = 1/2$, και ένα που έχει πιθανότητα $2/3$ να έρθει Κορώνα. Επιλέγουμε ένα από τα δύο στην τύχη και το στρίβουμε δύο φορές:

1. Ποια η πιθανότητα να φέρουμε Γ την πρώτη φορά;
2. Ποια η πιθανότητα να φέρουμε $\Gamma\Gamma$;
3. Δεδομένου ότι φέραμε $\Gamma\Gamma$, ποια η πιθανότητα να έχουμε επιλέξει το δίκαιο νόμισμα;

Για το πρώτο ερώτημα, παρατηρούμε ότι δεν μπορεί να υπολογιστεί άμεσα η ζητούμενη πιθανότητα αν δεν γνωρίζουμε ποιο νόμισμα έχει επιλεγεί. Ορίζουμε λοιπόν τα ενδεχόμενα:

$$\begin{aligned} A_1 &= \text{«φέραμε } \Gamma \text{ την πρώτη φορά}, \\ A_2 &= \text{«φέραμε } \Gamma \text{ τη δεύτερη φορά}, \\ B &= \text{«φέραμε 2 φορές } \Gamma \text{»} = A_1 \cap A_2, \\ \Delta &= \text{«επιλέξαμε το δίκαιο νόμισμα»}. \end{aligned}$$

Από το πρόβλημα μας δίνεται ότι $\Pr(A_1|\Delta) = \Pr(A_2|\Delta) = \frac{1}{2}$ και $\Pr(A_1|\Delta') = \Pr(A_2|\Delta') = \frac{1}{3}$. Για να υπολογίσουμε τη ζητούμενη πιθανότητα $\Pr(A_1)$, εκφράζουμε το A_1 ως την ένωση δύο ξένων ενδεχομένων, $A_1 = (A_1 \cap \Delta) \cup (A_1 \cap \Delta')$, οπότε,

$$\begin{aligned} \Pr(A_1) &= \Pr(A_1 \cap \Delta) + \Pr(A_1 \cap \Delta') && [\text{καν. πιθ. \#3}] \\ &= \Pr(A_1|\Delta) \Pr(\Delta) + \Pr(A_1|\Delta') \Pr(\Delta') && [\text{σχέση (5.1)}] \\ &= \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{2} = \frac{5}{12}. \end{aligned} \tag{5.5}$$

Με τον ίδιο ακριβώς τρόπο βρίσκουμε πως και $\Pr(A_2) = 5/12$, και, για το δεύτερο ερώτημα, παρομοίως έχουμε,

$$\begin{aligned} \Pr(B) &= \Pr(B \cap \Delta) + \Pr(B \cap \Delta') \\ &= \Pr(B|\Delta) \Pr(\Delta) + \Pr(B|\Delta') \Pr(\Delta') \\ &= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{2} \\ &= \frac{13}{72}. \end{aligned}$$

[Παρενθετικά αναφέρουμε την εξής ενδιαφέρουσα παρατήρηση: Αν γνωρίζουμε την πιθανότητα $p = \Pr(K)$ με την οποία ένα νόμισμα έρχεται Κορώνα, τότε τα αποτελέσματα διαδοχικών ρίψεων είναι ανεξάρτητα. Αλλά αν δεν γνωρίζουμε την p , η ανεξαρτησία δεν ισχύει! Π.χ., με το τυχαία επιλεγμένο νόμισμα αυτού του παραδείγματος, η πιθανότητα να φέρουμε Γράμματα είναι $5/12$ αλλά η πιθανότητα να φέρουμε δύο φορές Γράμματα είναι $13/72$ το οποίο δεν ισούται με $(5/12)^2$. Ο λόγος είναι πως, όταν δεν γνωρίζουμε τα στατιστικά χαρακτηριστικά του νομίσματος, κάθε ρίψη μάς δίνει κάποια πληροφορία για το νόμισμα, η οποία επηρεάζει την πρόβλεψή μας για τα αποτελέσματα των μελλοντικών ρίψεων.]

Πριν εξετάσουμε το τρίτο ερώτημα, σημειώνουμε πως, για οποιαδήποτε δύο ενδεχόμενα A, B , εκφράζοντας το A ως $(A \cap B) \cup (A \cap B')$ όπως στην (5.5), και εφαρμόζοντας τον κανόνα πιθανότητας #3 σε συνδυασμό με τον ορισμό της δεσμευμένης πιθανότητας, όπως πιο πάνω, βρίσκουμε:

Κανόνας πιθανότητας #9: «Κανόνας συνολικής πιθανότητας».

Για οποιαδήποτε ενδεχόμενα A, B :

$$\Pr(A) = \Pr(A|B) \Pr(B) + \Pr(A|B') \Pr(B').$$

Γενικότερα, αν τα ενδεχόμενα B_1, B_2, \dots, B_N είναι ξένα (δηλ., $B_i \cap B_j = \emptyset$ για κάθε $i \neq j$) και καλύπτουν όλο το Ω (δηλ., $B_1 \cup B_2 \cup \dots \cup B_N = \Omega$), τότε:

$$\Pr(A) = \Pr(A|B_1) \Pr(B_1) + \Pr(A|B_2) \Pr(B_2) + \dots + \Pr(A|B_N) \Pr(B_N).$$

[Άσκηση. Αποδείξτε την τελευταία παραπάνω γενική περίπτωση.]

Παράδειγμα 5.9, συνέχεια Το τρίτο ερώτημα του παραδείγματος ρωτά, δεδομένου ότι φέραμε ΓΓ, ποια είναι η πιθανότητα $\Pr(\Delta|B)$ να έχουμε επιλέξει το δίκαιο νόμισμα; Χρησιμοποιώντας τον ορισμό της δεσμευμένης πιθανότητας δύο φορές, βλέπουμε πως η ζητούμενη πιθανότητα $\Pr(\Delta|B)$ μπορεί να εκφραστεί ως προς τη γνωστή πιθανότητα $\Pr(B|\Delta) = (1/2)(1/2) = 1/4$:

$$\Pr(\Delta|B) = \frac{\Pr(B \cap \Delta)}{\Pr(B)} = \frac{\Pr(B|\Delta) \Pr(\Delta)}{\Pr(B)} = \frac{(1/4)(1/2)}{13/72} = \frac{8}{13} \approx 0.615. \quad (5.6)$$

Αρχικά η πιθανότητα να επιλέξουμε το δίκαιο νόμισμα είναι $1/2$. Αλλά, δεδομένου ότι το νόμισμα που επιλέξαμε έφερε 2 φορές Γράμματα, είναι λιγότερο πιθανό να επιλέξαμε το νόμισμα που 2 στις 3 φορές φέρνει Κορώνα και πιθανότερο να επιλέξαμε το δίκαιο. Ο υπολογισμός (5.6) μάς λέει ακριβώς πόσο πιο πιθανό είναι να έχουμε επιλέξει το δίκαιο νόμισμα.

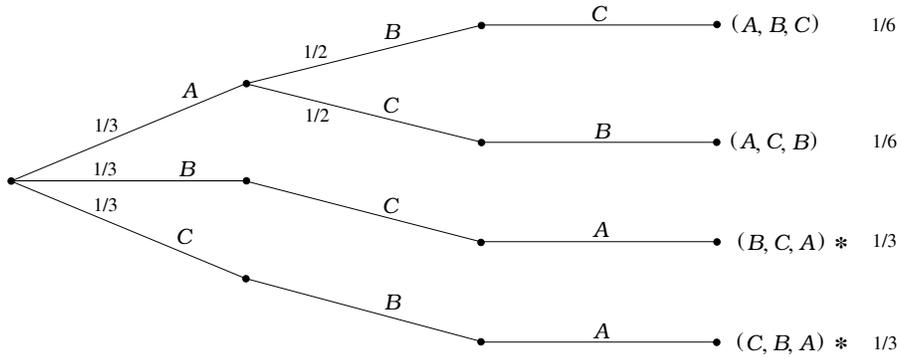
Η γενίκευση αυτού του παραδείγματος είναι το αντικείμενο της επόμενης ενότητας, αλλά, πριν το εξετάσουμε αναλυτικά, επανερχόμαστε στο Παράδειγμα 2.4 του τηλεπαιχνιδιού Monty Hall από το Κεφάλαιο 2.

Παράδειγμα 5.10 (Παιχνίδι Monty Hall, συνέχεια) Θυμηθείτε το πρόβλημα του τηλεπαιχνιδιού Monty Hall, που είδαμε στο Παράδειγμα 2.4 του Κεφαλαίου 2. Εδώ θα εξετάσουμε το βασικό ερώτημα που προκύπτει: Τι είναι πιο συμφέρον για τον παίκτη, να αλλάξει κουρτίνα ή να διατηρήσει την αρχική του επιλογή; Με άλλα λόγια, ποια από τις δύο στρατηγικές τού δίνει μεγαλύτερη πιθανότητα να κερδίσει το δώρο;

Αν και, διαισθητικά, θα φανταζόμασταν ίσως πως δεν υπάρχει διαφορά ανάμεσα στις δύο περιπτώσεις, η απάντηση ενδεχομένως να σας εκπλήξει.

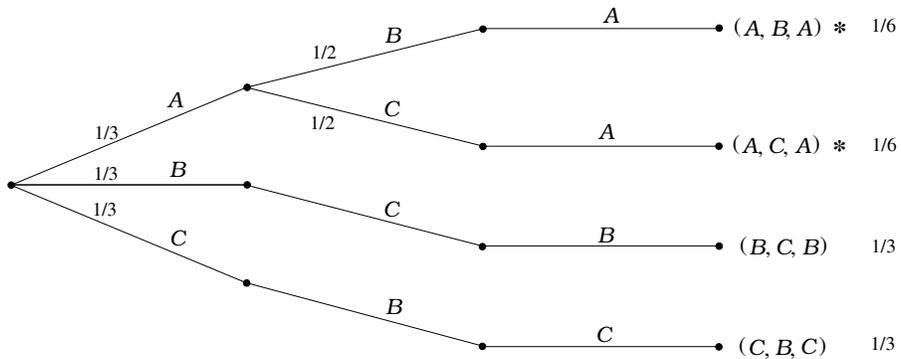
Όπως στο Παράδειγμα 2.4, θεωρούμε πως υπάρχουν τρεις κουρτίνες, οι A, B και C , και πως τα αποτελέσματα του παιχνιδιού (δηλαδή τα στοιχεία του χώρου πιθανότητας Ω) είναι της μορφής (X, X, X) , όπου το πρώτο X συμβολίζει την κουρτίνα που επιλέγει αρχικά ο παίκτης, το δεύτερο την κουρτίνα που αποκαλύπτεται, και το τρίτο την κουρτίνα που επιλέγει ο παίκτης στο τέλος. Υποθέτουμε ότι το δώρο είναι πίσω από την κουρτίνα A , και επίσης πως η αρχική επιλογή του παίκτη είναι τυχαία, δηλαδή πως τα ενδεχόμενα «ο παίκτης πρώτα επιλέγει την A », «ο παίκτης πρώτα επιλέγει την B » και «ο παίκτης πρώτα επιλέγει την C » έχουν πιθανότητα $1/3$ το καθένα.

Στρατηγική 1. Έστω πως ο παίκτης αλλάζει πάντα κουρτίνα. Τότε θα κερδίσει αν και μόνο αν αρχικά επιλέξει την κουρτίνα B ή C . Επειδή τα αντίστοιχα ενδεχόμενα είναι ξένα, η πιθανότητα να κερδίσει είναι $1/3 + 1/3 = 2/3$. Βλ. Σχήμα 5.1.



Σχήμα 5.1: Η πρώτη στρατηγική του Παραδείγματος 5.10. Στο πρώτο βήμα ο παίκτης επιλέγει μία κουρτίνα με πιθανότητα $1/3$. Στο δεύτερο βήμα, αν έχει επιλέξει την κουρτίνα A με το δώρο, ο παρουσιαστής ανοίγει μία από τις δύο άλλες τυχαία (με πιθ. $1/2$), ενώ αν έχει επιλέξει την B ή την C ο παρουσιαστής ανοίγει την άλλη κουρτίνα που δεν περιέχει το δώρο. Στο τρίτο βήμα ο παίκτης πάντα αλλάζει την αρχική επιλογή του. Δεξιά φαίνονται οι τέσσερις δυνατές καταλήξεις του παιχνιδιού μαζί με τις αντίστοιχες πιθανότητές τους. Αυτές στις οποίες ο παίκτης κερδίζει είναι σημειωμένες με *. Άρα, η πιθανότητα να κερδίσει είναι $1/3 + 1/3 = 2/3$.

Στρατηγική 2. Ο παίκτης δεν αλλάζει ποτέ κουρτίνα. Τότε κερδίζει αν και μόνο αν επιλέξει αρχικά την κουρτίνα A , δηλαδή με πιθανότητα μόνο $1/3$! Βλ. Σχήμα 5.2.



Σχήμα 5.2: Η δεύτερη στρατηγική του Παραδείγματος 5.10. Τα δύο πρώτα βήματα είναι τα ίδια με πριν, ενώ στο τρίτο βήμα ο παίκτης διατηρεί την αρχική επιλογή του. Δεξιά πάλι φαίνονται οι τέσσερις δυνατές καταλήξεις και οι πιθανότητές τους. Αυτές στις οποίες ο παίκτης κερδίζει είναι σημειωμένες με *, οπότε εδώ η πιθανότητα να κερδίσει είναι μόνο $1/6 + 1/6 = 1/3$!

Συνεπώς, συμφέρει τον παίκτη να αλλάζει πάντα κουρτίνα, ανεξαρτήτως του τι θα κάνει ο τηλεπαρουσιαστής!

5.3 Ο κανόνας του Bayes

Σε πολλά επιστημονικά αλλά και καθημερινά προβλήματα, συχνά προκύπτει το εξής ερώτημα: Αν γνωρίζουμε την τιμή μιας δεσμευμένης πιθανότητας $\Pr(A|B)$, πώς μπορούμε να υπολογίσουμε την πιθανότητα $\Pr(B|A)$, όπου οι ρόλοι των ενδεχομένων A και B έχουν αντιστραφεί; Φερειπείν, στο Παράδειγμα 5.9 πιο πάνω, ήταν προφανές ότι η πιθανότητα να φέρουμε δύο φορές «Γ», δεδομένου ότι έχουμε ένα δίκαιο νόμισμα, είναι 25%, αλλά το ζητούμενο ήταν να εκτιμήσουμε πόσο πιθανό ήταν να είχαμε επιλέξει το δίκαιο νόμισμα δεδομένου ότι φέραμε δύο φορές «Γ».

Ο κανόνας του Bayes αποτελεί το βασικό εργαλείο για την επίλυση τέτοιου είδους προβλημάτων. Μια από τις σημαντικότερες και πιο συνηθισμένες τέτοιες περιπτώσεις αφορούν, όπως στο Παράδειγμα 5.6, προβλήματα που σχετίζονται με τεστ διαγνωστικών εξετάσεων. Εκεί, κατά κανόνα γνωρίζουμε τη στατιστική συμπεριφορά του τεστ κάτω από ελεγχόμενες συνθήκες, π.χ., ξέρουμε την πιθανότητα ένα τεστ εγκυμοσύνης να έχει θετικό αποτέλεσμα δεδομένου ότι η εξεταζόμενη είναι έγκυος. Αλλά για εκείνη που κάνει το τεστ το σημαντικό ερώτημα είναι «δεδομένου ότι το τεστ είναι θετικό, πόσο πιθανό είναι να είμαι έγκυος»; Παρομοίως, σε δικαστικές διαμάχες, συχνά παρουσιάζονται στοιχεία της εξής μορφής: Αν ο κατηγορούμενος ήταν αθώος, η πιθανότητα να βρισκόταν το DNA του στο σημείο του εγκλήματος είναι μικρή. Στην πραγματικότητα όμως το δικαστήριο το ενδιαφέρει η «αντίστροφη» πιθανότητα: Δεδομένου ότι η αστυνομία εντόπισε το DNA του εκεί, ποια η πιθανότητα να είναι αθώος;

Χρησιμοποιώντας τον ορισμό της δεσμευμένης πιθανότητας, όπως στη σχέση (5.6), και εφαρμόζοντας κατόπιν τον κανόνα συνολικής πιθανότητας, προκύπτει το εξής γενικό αποτέλεσμα:

Κανόνας πιθανότητας #10: «Κανόνας του Bayes». Για οποιαδήποτε ενδεχόμενα A, B :

$$\Pr(B|A) = \frac{\Pr(A|B) \Pr(B)}{\Pr(A)} = \frac{\Pr(A|B) \Pr(B)}{\Pr(A|B) \Pr(B) + \Pr(A|B') \Pr(B')}.$$

Γενικότερα, αν τα ενδεχόμενα B_1, B_2, \dots, B_N είναι ξένα (δηλ., $B_i \cap B_j = \emptyset$ για κάθε $i \neq j$) και καλύπτουν όλο το Ω (δηλ., $B_1 \cup B_2 \cup \dots \cup B_N = \Omega$), τότε:

$$\begin{aligned} \Pr(B_1|A) &= \frac{\Pr(A|B_1) \Pr(B_1)}{\Pr(A)} \\ &= \frac{\Pr(A|B_1) \Pr(B_1)}{\Pr(A|B_1) \Pr(B_1) + \Pr(A|B_2) \Pr(B_2) + \dots + \Pr(A|B_N) \Pr(B_N)}. \end{aligned}$$

Παρενθετικά αναφέρουμε ότι ο κανόνας του Bayes αποτελεί την αφετηρία μιας πολύ σημαντικής επιστημονικής περιοχής, της *στατιστικής κατά Bayes*, η οποία είναι εξαιρετικά ενεργή ερευνητικά, και η οποία τα τελευταία 20 περίπου χρόνια έχει βρει εφαρμογές σε κεντρικά θέματα σχεδόν ολόκληρου του φάσματος της επιστήμης και της τεχνολογίας.

Παράδειγμα 5.11 Σε κάποιον πληθυσμό, 2% των ανθρώπων πάσχει από μια ασθένεια. Ένας τυχαία επιλεγμένος άνθρωπος κάνει ένα διαγνωστικό τεστ γι' αυτή την ασθένεια, όπου οι πιθανότητες σφάλματος του τεστ είναι:

$$\begin{aligned}\Pr(\text{αρνητικό τεστ} \mid \text{ασθενής}) &= 0.5\%, \\ \Pr(\text{θετικό τεστ} \mid \text{όχι ασθενής}) &= 5\%.\end{aligned}$$

Δεδομένου ότι το τεστ είναι αρνητικό, ποια η πιθανότητα παρ' όλα αυτά να έχει την ασθένεια; Ορίζουμε τα ενδεχόμενα,

$$\begin{aligned}\Theta &= \text{«θετικό τεστ»}, \\ A &= \text{«ασθενής»},\end{aligned}$$

οπότε τα δεδομένα του προβλήματος είναι,

$$\Pr(A) = 0.02, \quad \Pr(\Theta' \mid A) = 0.005, \quad \Pr(\Theta \mid A') = 0.05,$$

και η ζητούμενη πιθανότητα είναι η $\Pr(A \mid \Theta')$. Από τον κανόνα του Bayes έχουμε,

$$\begin{aligned}\Pr(A \mid \Theta') &= \frac{\Pr(\Theta' \mid A) \Pr(A)}{\Pr(\Theta' \mid A) \Pr(A) + \Pr(\Theta' \mid A') \Pr(A')} \\ &= \frac{\Pr(\Theta' \mid A) \Pr(A)}{\Pr(\Theta' \mid A) \Pr(A) + [1 - \Pr(\Theta \mid A')][1 - \Pr(A)]},\end{aligned}$$

όπου στο δεύτερο βήμα χρησιμοποιήσαμε τον κανόνα πιθανότητας #4 και το Λήμμα 5.1. Αντικαθιστώντας τώρα τις τιμές των πιο πάνω πιθανοτήτων από τα δεδομένα μας, εύκολα υπολογίζουμε,

$$\Pr(A \mid \Theta') = \frac{0.005 \times 0.02}{0.005 \times 0.02 + 0.95 \times 0.98} \approx 10^{-4}.$$

Παράδειγμα 5.12 Ανακοινώνεται από την εταιρία Dell ότι, από τα PC ενός συγκεκριμένου μοντέλου που κυκλοφορούν στην αγορά:

- το 25% είναι ελαττωματικά,
- το 25% έχουν σημαντικό ρίσκο να παρουσιάσουν πρόβλημα,
- το 50% δεν έχουν κανένα πρόβλημα από τα δύο.

Διανέμεται από την εταιρία ένα πρόγραμμα διάγνωσης των προβλημάτων, το οποίο είναι 99% ακριβές, δηλαδή,

$$\Pr(\Theta|(E \cup R)') = \Pr(\Theta'|E) = \Pr(\Theta'|R) = 1\%,$$

όπου τα ενδεχόμενα Θ, E, R , ορίζονται ως,

$$\Theta = \text{«θετικό αποτέλεσμα διαγνωστικού τεστ»},$$

$$E = \text{«ελαττωματικό PC»},$$

$$R = \text{«PC με ρίσκο προβλήματος»}.$$

Κάνουμε στο PC μας αυτό το διαγνωστικό τεστ και βγαίνει αρνητικό. Ποια η πιθανότητα ο υπολογιστής μας πράγματι να μην έχει κανένα από τα δύο προβλήματα;

Εδώ ο κανόνας του Bayes μάς δίνει,

$$\begin{aligned} \Pr((E \cup R)'|\Theta') \\ = \frac{\Pr(\Theta'|(E \cup R)') \Pr((E \cup R)')}{\Pr(\Theta'|(E \cup R)') \Pr((E \cup R)') + \Pr(\Theta'|E) \Pr(E) + \Pr(\Theta'|R) \Pr(R)}, \end{aligned}$$

όπου παρατηρούμε ότι τα $E, R, (E \cup R)'$ είναι ξένα και η ένωσή τους ισούται με ολόκληρο το χώρο πιθανότητας Ω . Αντικαθιστώντας, όπως και στο προηγούμενο παράδειγμα, τις τιμές από τα δεδομένα του προβλήματος υπολογίζουμε,

$$\Pr((E \cup R)'|\Theta') = \frac{0.99 \times 0.5}{0.99 \times 0.5 + 0.01 \times 0.25 + 0.01 \times 0.25} = 0.99.$$

Άρα, σε κάποιες (πολύ ειδικές) περιπτώσεις, μπορεί να έχουμε και $\Pr(A|B) = \Pr(B|A)$, όπως εδώ έχουμε $\Pr((E \cup R)'|\Theta') = \Pr(\Theta'|(E \cup R)').$

Παράδειγμα 5.13 Έστω ότι έχουμε τρία ζάρια, όπου το πρώτο είναι δίκαιο, το δεύτερο δεν έρχεται ποτέ 6, και το τρίτο έχει $\Pr(\text{«6»}) = 2/3$. Επιλέγουμε ένα στην τύχη, το ρίχνουμε δύο φορές και φέρνουμε 66. Ποια η πιθανότητα να επιλέξαμε το δίκαιο ζάρι;

Ορίζουμε τα ενδεχόμενα Δ, A_2, A_3 ως «επιλέξαμε το δίκαιο ζάρι», «επιλέξαμε το δεύτερο ζάρι» και «επιλέξαμε το τρίτο ζάρι», αντίστοιχα. Επιπλέον ορίζουμε το ενδεχόμενο $E = \text{«φέραμε 66»}$. Βάσει του κανόνα του Bayes,

$$\begin{aligned} \Pr(\Delta|E) &= \frac{\Pr(E|\Delta) \Pr(\Delta)}{\Pr(E|\Delta) \Pr(\Delta) + \Pr(E|A_2) \Pr(A_2) + \Pr(E|A_3) \Pr(A_3)} \\ &= \frac{(1/6)(1/6)(1/3)}{(1/6)(1/6)(1/3) + 0 \times (1/3) + (2/3)(2/3)(1/3)} \\ &= \frac{1}{16}, \end{aligned}$$

πράγμα το οποίο συμφωνεί με τη διαίσθησή μας: Εφόσον φέραμε 2 φορές «6», είναι πολύ πιθανότερο να επιλέξαμε ένα ζάρι το οποίο έχει πιθανότητα $2/3$ να φέρει «6», παρά ένα δίκαιο ζάρι.

5.4 Ακόμα πέντε «κανόνες πιθανότητας»

Σ' αυτό το κεφάλαιο διατυπώσαμε ακόμα 5 κανόνες πιθανότητας, οι οποίοι συμπληρώνουν εκείνους του Κεφαλαίου 3. Για ευκολότερη αναφορά και χρήση, τους παραθέτουμε περιληπτικά πιο κάτω.

Κανόνες πιθανότητας (6–10)

Για οποιαδήποτε δύο ενδεχόμενα A, B :

6. Εάν τα ενδεχόμενα A, B είναι «λογικά ανεξάρτητα», τότε θεωρούμε πως είναι και (στατιστικά) ανεξάρτητα, δηλαδή:

$$\Pr(A \cap B) = \Pr(A) \Pr(B).$$

7. $\Pr(A \cap B) = \Pr(A) \Pr(B|A) = \Pr(B) \Pr(A|B)$, όπου η δεσμευμένη πιθανότητα ορίζεται ως,

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}.$$

8. $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$.

9. **Κανόνας συνολικής πιθανότητας:**

$$\Pr(A) = \Pr(A|B) \Pr(B) + \Pr(A|B') \Pr(B').$$

10. **Κανόνας του Bayes:**

$$\Pr(B|A) = \frac{\Pr(A|B) \Pr(B)}{\Pr(A)} = \frac{\Pr(A|B) \Pr(B)}{\Pr(A|B) \Pr(B) + \Pr(A|B') \Pr(B')}.$$

5.5 Ασκήσεις

1. *Ένωση τριών ενδεχομένων.* Να δείξετε ότι, για οποιαδήποτε τρία ενδεχόμενα A, B και C , η πιθανότητα της ένωσης $\Pr(A \cup B \cup C)$ ισούται με:

$$\Pr(A) + \Pr(B) + \Pr(C) - \Pr(A \cap B) - \Pr(A \cap C) - \Pr(B \cap C) + \Pr(A \cap B \cap C).$$

2. *Πιθανότητα ακριβώς ενός ενδεχομένου.* Δείξτε ότι η πιθανότητα να πραγματοποιηθεί ακριβώς ένα από τα ενδεχόμενα A, B είναι ίση με:

$$\Pr(A) + \Pr(B) - 2\Pr(A \cap B).$$

3. *Τράπουλα.* Μοιράζουμε στην τύχη 10 φύλλα από μια συνηθισμένη τράπουλα 52 φύλλων. Ποια η πιθανότητα η μοιρασιά να περιέχει τουλάχιστον έναν άσο και τουλάχιστον μία φιγούρα (δηλαδή βαλέ, ντάμα ή ρήγα);

4. *Ιδιότητες δεσμευμένης πιθανότητας.* Έστω τρία οποιαδήποτε ενδεχόμενα A, B και C .

(α') Να δείξετε ότι $\Pr(A \cap B \cap C) = \Pr(A) \Pr(B|A) \Pr(C|A \cap B)$.

(β') Να βρεθεί η δεσμευμένη πιθανότητα $\Pr(A|B)$ αν $A \cap B = \emptyset$, αν $A \subset B$, και αν $B \subset A$.

(γ') Να δείξετε ότι $\Pr(A|B) > \Pr(A)$ αν και μόνο αν $\Pr(B|A) > \Pr(B)$.

Βεβαιωθείτε ότι κατανοείτε διαισθητικά γιατί ισχύουν τα παραπάνω.

5. *Ρίψεις ζαριού.* Ρίχνουμε ένα δίκαιο ζάρι δύο φορές. Υποθέτουμε ότι οι ρίψεις είναι ανεξάρτητες μεταξύ τους. Έστω τα ακόλουθα ενδεχόμενα:

(α') $A =$ «το άθροισμα είναι ζυγό»,

(β') $B =$ «το άθροισμα είναι μονό»,

(γ') $C =$ «το άθροισμα είναι ίσο με 7»,

(δ') $D =$ «φέραμε δύο φορές το ίδιο»,

(ε') $E =$ «η πρώτη ζαριά ήρθε 6»,

(ς') $F =$ «το πρώτο ζάρι είναι 1»,

(ζ') $G =$ «το άθροισμα είναι ίσο με 6».

Είναι ή όχι τα πιο κάτω ζεύγη ενδεχομένων ανεξάρτητα;

(α') A, B , (β') B, C , (γ') D, E , (δ') C, F , (ε') G, F .

6. *Δύο δοχεία.* Έστω δύο διαφορετικά τμήματα, εκ των οποίων το πρώτο αποτελείται από m_1 φοιτητές και w_1 φοιτήτριες, ενώ το δεύτερο αποτελείται από m_2 φοιτητές και w_2 φοιτήτριες. Επιλέγουμε ένα τμήμα στην τύχη, χωρίς κάποια προτίμηση, και εκ των υστέρων επιλέγουμε τυχαία ένα άτομο από αυτό το τμήμα. Έστω F το ενδεχόμενο να επιλέξουμε το πρώτο τμήμα, και έστω M το ενδεχόμενο να επιλέξουμε φοιτητή.

Ποια είναι η σχέση που πρέπει να ικανοποιούν τα m_1 , m_2 , w_1 , w_2 ώστε να είναι ανεξάρτητα τα ενδεχόμενα F και M ;

7. *Poker με δεσμευμένες πιθανότητες.* Όπως στην Άσκηση 5 του Κεφαλαίου 4, ένας παίκτης του πόκερ παίρνει 5 φύλλα από μια κανονική τράπουλα 52 φύλλων. Χρησιμοποιώντας δεσμευμένες πιθανότητες και όχι τους κανόνες αρίθμησης του Κεφαλαίου 4, υπολογίστε τις πιθανότητες να έχει:

- (α) καρέ (δηλαδή 4 ίδια φύλλα, για παράδειγμα 4 άσους ή 4 ντάμες);
 (β) φουλ (δηλαδή ένα ζευγάρι και μία τριάδα, για παράδειγμα 3 άσους και 2 ρηγάδες);
 (γ) χρώμα (δηλαδή όλα κούπες ή όλα σπαθιά ή όλα μπαστούνια ή όλα καρό);

8. *Τεστ πολλαπλών απαντήσεων.* Ένας μαθητής απαντά μια ερώτηση πολλαπλών απαντήσεων, με N δυνατές απαντήσεις. Αν ο μαθητής δεν ξέρει την απάντηση, επιλέγει μία από τις δυνατές απαντήσεις στην τύχη, χωρίς κάποια προτίμηση. Αν η πιθανότητα να ξέρει την απάντηση είναι p , ποια είναι η δεσμευμένη πιθανότητα να ήξερε την απάντηση δεδομένου ότι απάντησε σωστά;

9. *Δεσμευμένο μέτρο πιθανότητας.* Έστω ενδεχόμενο B σε κάποιο χώρο πιθανότητας Ω , με $\Pr(B) > 0$. Να δείξετε ότι η συνάρτηση $\mathbb{P}(A)$, η οποία ορίζεται ως $\mathbb{P}(A) = \Pr(A|B)$ για κάθε ενδεχόμενο A , είναι ένα μέτρο πιθανότητας, δηλαδή ικανοποιεί τις συνθήκες:

- (α) $\Pr(A|B) \geq 0$ για κάθε ενδεχόμενο A .
 (β) $\Pr(\Omega|B) = 1$.
 (γ) Αν τα A_i , $i = 1, 2, \dots$ είναι ζένα ενδεχόμενα, τότε:

$$\Pr\left(\bigcup_{i=1}^{\infty} A_i \mid B\right) = \sum_{i=1}^{\infty} \Pr(A_i|B).$$

10. *Συνέπειες ανεξαρτησίας.* Να δείξετε ότι, αν τα A και B είναι ανεξάρτητα ενδεχόμενα, τότε τα ζεύγη A και B' , A' και B , και A' και B' είναι επίσης ανεξάρτητα. Μπορείτε να ερμηνεύσετε διαισθητικά το αποτέλεσμα; *Υπόδειξη.* Όπως στον κανόνα συνολικής πιθανότητας, το ενδεχόμενο A μπορεί να εκφραστεί ως ένωση δύο ζένων ενδεχομένων, $A = (A \cap B) \cup (A \cap B')$.

11. *Ανεξαρτησία και δέσμευση.* Έστω τρία οποιαδήποτε ενδεχόμενα A, B και C .
- (α') Να δείξετε ότι $\Pr(A \cup B \cup C) = 1 - \Pr(A'|B' \cap C') \Pr(B'|C') \Pr(C')$.
- (β') Να δείξετε ότι, αν το ενδεχόμενο A είναι ανεξάρτητο από τον εαυτό του, τότε αναγκαστικά θα έχουμε είτε $\Pr(A) = 0$ ή $\Pr(A) = 1$.
- (γ') Να δείξετε ότι αν $\Pr(A) = 0$ ή $\Pr(A) = 1$, τότε το A είναι ανεξάρτητο του B για οποιοδήποτε ενδεχόμενο B .
12. *Κι άλλη τράπουλα.* Επιλέγουμε 5 χαρτιά (ανεξαρτήτως της σειράς) στην τύχη από μια συνηθισμένη τράπουλα.
- (α') Ποια η πιθανότητα να επιλέξαμε 4 άσους και μία φιγούρα;
- (β') Ποια η πιθανότητα, δεδομένου ότι επιλέξαμε ακριβώς μία φιγούρα, τα άλλα 4 φύλλα να είναι άσοι;
- (γ') Ποια η πιθανότητα να επιλέξαμε 4 φιγούρες και έναν άσο;
- (δ') Είναι ή όχι τα ενδεχόμενα στο (α') και στο (γ') ανεξάρτητα; Αποδείξτε την απάντησή σας.
13. *Κανόνας δεσμευμένης συνολικής πιθανότητας.* Να δείξετε πως, για τρία οποιαδήποτε ενδεχόμενα E, F και G , έχουμε:
- $$\Pr(E|F) = \Pr(E|G \cap F) \Pr(G|F) + \Pr(E|G' \cap F) \Pr(G'|F).$$
14. *Μεσογειακή αναιμία.* Σε κάποιο πληθυσμό, το 8% των ατόμων έχει το στίγμα της μεσογειακής αναιμίας. Έστω ότι ένα τυχαία επιλεγμένο άτομο κάνει μια εξέταση για να διαπιστώσει αν έχει το στίγμα ή όχι. Η εξέταση δεν είναι απόλυτα ακριβής, και η πιθανότητα το αποτέλεσμα να βγει θετικό, ενώ δεν υπάρχει στίγμα, είναι 10%. Επιπλέον, η πιθανότητα να βγει το αποτέλεσμα αρνητικό, ενώ υπάρχει στίγμα, είναι 1%.
- (α') Ποια η πιθανότητα να έχει στίγμα κάποιος που κάνει την εξέταση και προκύπτει θετικό αποτέλεσμα;
- (β') Ποια είναι η αντίστοιχη πιθανότητα για κάποιον που κάνει την εξέταση δύο ανεξάρτητες φορές, και προκύπτει θετικό αποτέλεσμα την πρώτη φορά και αρνητικό αποτέλεσμα τη δεύτερη;
15. *Ανεπιθύμητες εγκυμοσύνες.* Από το σύνολο των γυναικών που κάνουν ένα τεστ εγκυμοσύνης, μόνο το 12% είναι έγκυες. Έστω ότι το τεστ έχει τις εξής πιθανότητες σφάλματος: $\Pr(\text{θετικό} | \text{όχι έγκυος}) = 1\%$ και $\Pr(\text{αρνητικό} | \text{έγκυος}) = 3\%$.
- (α') Ποια η πιθανότητα να είναι έγκυος μια γυναίκα η οποία κάνει το τεστ και βγαίνει θετικό;

(β') Ποια είναι η αντίστοιχη πιθανότητα για μια γυναίκα που κάνει το τεστ 2 ανεξάρτητες φορές και βγει την πρώτη θετικό και τη δεύτερη αρνητικό;

16. *Απάτη.* Έστω ότι υποπτευόμαστε πως κάποιος χρησιμοποιεί ένα κάλπικο ζάρι, το οποίο δεν φέρνει ποτέ «6» και μάλιστα στην πλευρά του «6» έχει σημειωμένο το «1». Αλλά δεν είμαστε σίγουροι, οπότε θεωρούμε ότι υπάρχουν ίσες πιθανότητες να χρησιμοποιεί είτε ένα τέτοιο κάλπικο ζάρι είτε ένα δίκαιο. Όταν τον κατηγορούμε ότι σε έξι συνεχόμενες ζαριές έφερε τα αποτελέσματα 1,5,2,1,1,2, τα οποία δεν περιέχουν ούτε ένα «6», εκείνος διατείνεται ότι «Σιγά! Αυτό μια χαρά μπορεί να συμβεί και με το δίκαιο ζάρι!». Εμείς όμως τον αποστομώνουμε υπολογίζοντας πως η πιθανότητα να είναι δίκαιο το ζάρι δεδομένων αυτών των αποτελεσμάτων είναι μόλις μία στις εννιά! Αποδείξτε ότι έχουμε δίκιο.

Κεφάλαιο 6

Διακριτές τυχαίες μεταβλητές

Σε σύνθετα προβλήματα των πιθανοτήτων, όπως π.χ. σε προβλήματα ανάλυσης πολύπλοκων δικτύων ή στη στατιστική ανάλυση μεγάλων δεδομένων, η λεπτομερής, στοιχείο-προς-στοιχείο περιγραφή του πλήρους χώρου πιθανότητας Ω είναι εξαιρετικά χρονοβόρα, και συχνά είναι και περιττή. Όπως θα δούμε σε αυτό το κεφάλαιο, η έννοια της *τυχαίας μεταβλητής* μάς επιτρέπει να παρακάμψουμε την πιο πάνω διαδικασία.¹

Παράδειγμα 6.1 Ρίχνουμε ένα δίκαιο νόμισμα N φορές, και ορίζουμε τις εξής τυχαίες ποσότητες: Για κάθε $i = 1, 2, \dots, N$, έστω:

$$X_i = \begin{cases} 1, & \text{αν φέραμε Κ τη φορά } i, \\ 0, & \text{αν φέραμε Γ τη φορά } i. \end{cases}$$

Τα X_i είναι *τυχαίες μεταβλητές*, και το κάθε X_i περιγράφει το αποτέλεσμα της ρίψης i .

Παρομοίως, αν ορίσουμε την ποσότητα Y ως το πλήθος από Κορώνες στις N ρίψεις, τότε το Y είναι μια *άλλη τυχαία μεταβλητή* η οποία μπορεί και να εκφραστεί ως,

$$Y = X_1 + X_2 + \dots + X_N.$$

6.1 Ορισμός και βασικές ιδιότητες

Ορισμός 6.1 (Τυχαία μεταβλητή)

1. (Διαισθητικά) *Μια τυχαία μεταβλητή (Τ.Μ.) είναι μια οποιαδήποτε ποσότητα που εξαρτάται από την έκβαση ενός τυχαίου πειράματος.*

¹Ορισμένα από τα αποτελέσματα που θα δούμε απαιτούν κάποιες επιπλέον τεχνικές υποθέσεις, οι οποίες όμως δεν επηρεάζουν την ουσία τους. Περαιτέρω λεπτομέρειες δίνονται στην Ενότητα [6.3](#).

2. (Μαθηματικά) Μια τυχαία μεταβλητή (Τ.Μ.) X είναι μια οποιαδήποτε συνάρτηση $X : \Omega \rightarrow \mathbb{R}$ για κάποιο χώρο πιθανότητας Ω .

Παράδειγμα 6.2 Στο Παράδειγμα 6.1, ο χώρος πιθανότητας Ω αποτελείται από όλες τις 2^N δυνατές N -άδες αποτελεσμάτων για τις N ρίψεις. Π.χ., το στοιχείο $KK \cdots K \in \Omega$ (N φορές) αντιστοιχεί στο να φέρουμε N φορές Κορώνα.

Αν θέλαμε να ορίσουμε τις τυχαίες μεταβλητές X_i απολύτως σχολαστικά, θα λέγαμε ότι, για κάθε $\omega \in \Omega$,

$$X_i(\omega) = \begin{cases} 1, & \text{αν στη θέση } i \text{ του } \omega \text{ έχουμε } K, \\ 0, & \text{αν στη θέση } i \text{ του } \omega \text{ έχουμε } \Gamma. \end{cases}$$

Με την ίδια λογική, η Τ.Μ. Y που περιγράφει το πλήθος από Κορώνες στις N ρίψεις θα μπορούσε να οριστεί σαν συνάρτηση στο Ω ως $Y(\omega) = \langle \text{πλήθος από } K \text{ στο } \omega \rangle$ ή, εναλλακτικά, όπως πριν,

$$Y(\omega) = X_1(\omega) + X_2(\omega) + \cdots + X_N(\omega).$$

Ορισμός 6.2 1. Το σύνολο όλων των δυνατών (πραγματικών) τιμών που μπορεί να πάρει μια τυχαία μεταβλητή X είναι το σύνολο τιμών της X και συμβολίζεται S_X ή απλά S .

2. Μια τυχαία μεταβλητή X είναι διακριτή αν το σύνολο τιμών της είναι είτε πεπερασμένο είτε άπειρο αλλά αριθμήσιμο.

3. Η πυκνότητα (ή διακριτή συνάρτηση πυκνότητας πιθανότητας) μιας διακριτής Τ.Μ. X με σύνολο τιμών S_X είναι η συνάρτηση $P : S_X \rightarrow [0, 1]$, που ορίζεται ως:

$$P(x) = \Pr(X = x), \quad \text{για κάθε } x \in S.$$

Παρατήρηση: Από αυστηρά μαθηματική σκοπιά, μια τυχαία μεταβλητή X που ορίζεται ως συνάρτηση σε κάποιο χώρο πιθανότητας Ω , έχει σύνολο τιμών το πεδίο τιμών της συνάρτησης $X : \Omega \rightarrow \mathbb{R}$, δηλαδή το σύνολο $S_X = \{x \in \mathbb{R} : f(\omega) = x \text{ για κάποιο } \omega \in \Omega\}$. Τότε, δεδομένου κάποιου μέτρου πιθανότητας \mathbb{P} στο Ω , η πυκνότητα της X ορίζεται ως,

$$P(x) = \mathbb{P}(\{\omega \in \Omega : X(\omega) = x\}), \quad \text{για κάθε } x \in S.$$

Επίσης σημειώνουμε ότι, σε μεγάλο μέρος της βιβλιογραφίας, η συνάρτηση $P(x)$ την οποία εδώ ονομάζουμε πυκνότητα μιας διακριτής Τ.Μ., συχνά λέγεται μάζα της X .

Παράδειγμα 6.3 Στο Παράδειγμα 6.1 όλες οι Τ.Μ. X_i έχουν σύνολο τιμών $S_{X_i} = \{0, 1\}$ και όλες έχουν την ίδια πυκνότητα P που είναι: $P(1) = 1/2$ και $P(0) = 1/2$. Η Τ.Μ. Y έχει σύνολο τιμών το $S_Y = \{0, 1, 2, \dots, N\}$, και αποτελεί ειδική περίπτωση μιας σημαντικής ομάδας τυχαίων

μεταβλητών – των τυχαίων μεταβλητών με «διωνυμική κατανομή» – τις οποίες θα εξετάσουμε αναλυτικά στο επόμενο κεφάλαιο.

Η πυκνότητα της Y μπορεί να υπολογιστεί σχετικά εύκολα μέσω των κανόνων αρίθμησης και πιθανότητας των προηγούμενων κεφαλαίων. Για παράδειγμα (θεωρώντας πως τα ενδεχόμενα που αφορούν τα αποτελέσματα των διαδοχικών ρίψεων είναι ανεξάρτητα) έχουμε,

$$\begin{aligned} P(0) &= \Pr(Y = 0) \\ &= \Pr(\text{«φέραμε } N \text{ φορές } \Gamma\text{»}) \\ &= \Pr(X_1 = 0 \text{ και } X_2 = 0 \text{ και } \cdots X_N = 0) \\ &= \Pr(X_1 = 0) \Pr(X_2 = 0) \cdots \Pr(X_N = 0) \\ &= \frac{1}{2} \cdot \frac{1}{2} \cdots \frac{1}{2} = 2^{-N}. \end{aligned}$$

Στο Κεφάλαιο 7 θα δούμε αναλυτικά τον λεπτομερή υπολογισμό των υπόλοιπων τιμών της πυκνότητας της Τ.Μ. Y .

Ιδιότητες. Η πυκνότητα P μιας διακριτής Τ.Μ. X με σύνολο τιμών S_X έχει τις εξής ιδιότητες:

1. $0 \leq P(x) \leq 1$, για κάθε $x \in S_X$.
2. $\sum_{x \in S_X} P(x) = 1$.
3. $\Pr(X \in T) = \sum_{x \in T} P(x)$, για οποιοδήποτε $T \subset S_X$.

Απόδειξη: Εφόσον κάθε $P(x)$ είναι μια πιθανότητα, η πρώτη ιδιότητα είναι προφανής.

Η τρίτη ιδιότητα είναι συνέπεια του κανόνα πιθανότητας #3 του Κεφαλαίου 3. Εφόσον τα ενδεχόμενα $\{X = x\}$ για διαφορετικά $x \in S_X$ είναι ξένα μεταξύ τους, έχουμε:

$$\Pr(X \in T) = \Pr\left(\bigcup_{x \in T} \{X = x\}\right) = \sum_{x \in T} \Pr(\{X = x\}) = \sum_{x \in T} P(x).$$

Τέλος η δεύτερη ιδιότητα είναι ειδική περίπτωση της τρίτης: Αν θέσουμε $T = S_X$, η τρίτη ιδιότητα μας δίνει,

$$\sum_{x \in S_X} P(x) = \Pr(X \in S_X),$$

το οποίο ισούται με 1, από τον ορισμό του συνόλου τιμών. ■

Ορισμός 6.3 Η συνάρτηση κατανομής μιας Τ.Μ. X είναι η συνάρτηση $F : \mathbb{R} \rightarrow [0, 1]$ που ορίζεται ως,

$$F(x) = \Pr(X \leq x),$$

για κάθε $x \in \mathbb{R}$.

Παρατηρήσεις

1. Από τον ορισμό της συνάρτησης κατανομής και την τρίτη ιδιότητα της πυκνότητας πιο πάνω, βλέπουμε πως η συνάρτηση κατανομής $F(x)$ μιας διακριτής Τ.Μ. μπορεί εύκολα να υπολογιστεί από την πυκνότητά της μέσω της σχέσης,

$$F(x) = \Pr(X \leq x) = \sum_{y \in \mathcal{S}: y \leq x} P(x). \quad (6.1)$$

2. Παρατηρούμε ότι, αν $x' > x$, τότε $\{X \leq x\} \subset \{X \leq x'\}$, και από τον κανόνα πιθανότητας #2,

$$F(x) = \Pr(X \leq x) \leq \Pr(X \leq x') = F(x'),$$

άρα η F είναι αύξουσα συνάρτηση.

Επίσης, αφού προφανώς $\Pr(X < \infty) = 1$, έχουμε $\lim_{x \rightarrow +\infty} F(x) = 1$. Παρομοίως, αφού προφανώς $\Pr(X \leq -\infty) = 0$, έχουμε $\lim_{x \rightarrow -\infty} F(x) = 0$. Για την αυστηρή τεκμηρίωση αυτών των δύο ορίων δείτε την Άσκηση 2 στο τέλος του Κεφαλαίου.

Παράδειγμα 6.4 Στρίβουμε ένα νόμισμα με $\Pr(K) = p$ για κάποιο $p \in (0, 1)$, σε συνεχόμενες, ανεξάρτητες ρίψεις. Όπως και στα προηγούμενα παραδείγματα, για κάθε $i = 1, 2, \dots, N$ ορίζουμε τις Τ.Μ.,

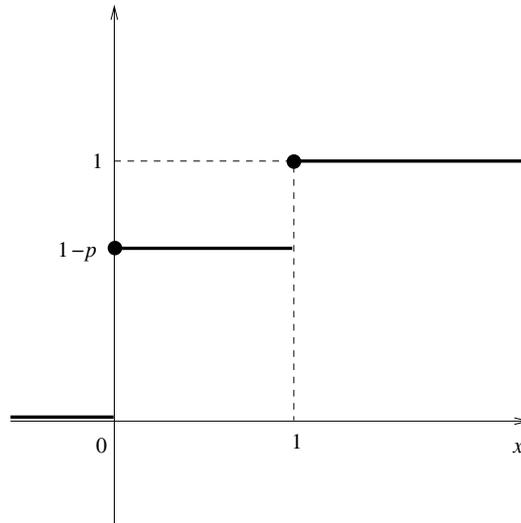
$$X_i = \begin{cases} 1, & \text{αν φέραμε Κ τη φορά } i, \\ 0, & \text{αν φέραμε Γ τη φορά } i, \end{cases}$$

όπου κάθε X_i περιγράφει το αποτέλεσμα της ρίψης i . Όπως πριν, όλες οι X_i έχουν το ίδιο σύνολο τιμών $S_{X_i} = S_X = \{0, 1\}$ και την ίδια πυκνότητα P που ικανοποιεί, $P(1) = p$ και $P(0) = 1 - p$. Στο Σχήμα 6.1 βλέπουμε τη γραφική αναπαράσταση της συνάρτησης κατανομής $F(x)$ των X_i . Εφόσον η X_i παίρνει μόνο τις τιμές 0 και 1, για $x < 0$ η πιθανότητα $F(x) = \Pr(X \leq x)$ ισούται με μηδέν, και για $x \geq 1$ έχουμε $F(x) = \Pr(X \leq x) = 1$. Τέλος αν το x ανήκει στο διάστημα $[0, 1)$, τότε,

$$F(x) = \Pr(X \leq x) = \Pr(X = 0) = P(0) = 1 - p.$$

Ορίζουμε επίσης την Τ.Μ. $Z = \langle \text{πρώτη φορά που θα έρθει Κ} \rangle$, δηλαδή, έχουμε $Z = m$ αν η πρώτη φορά που φέραμε Κ είναι στη ρίψη m :

$$Z = \begin{cases} 1, & \text{αν } X_1 = 1, \\ 2, & \text{αν } X_1 = 0, X_2 = 1, \\ 3, & \text{αν } X_1 = X_2 = 0, X_3 = 1, \\ \vdots & \vdots \\ m, & \text{αν } X_1 = X_2 = \dots = X_{m-1} = 0, X_m = 1. \end{cases}$$



Σχήμα 6.1: Γραφική αναπαράσταση της συνάρτησης κατανομής $F(x)$ των X_i .

Η Z έχει σύνολο τιμών όλους τους φυσικούς αριθμούς, $S_Z = \mathbb{N} = \{1, 2, \dots\}$ και η πυκνότητά της, P_Z , είναι,

$$\begin{aligned} P_Z(1) &= \Pr(Z = 1) = \Pr(X_1 = 1) = p, \\ P_Z(2) &= \Pr(Z = 2) = \Pr(X_1 = 0 \text{ και } X_2 = 1) = (1-p)p, \end{aligned}$$

και γενικά, για κάθε $m \geq 1$,

$$P_Z(m) = \Pr(Z = m) = \Pr(X_1 = \dots = X_{m-1} = 0 \text{ και } X_m = 1) = (1-p)^{m-1}p, \quad (6.2)$$

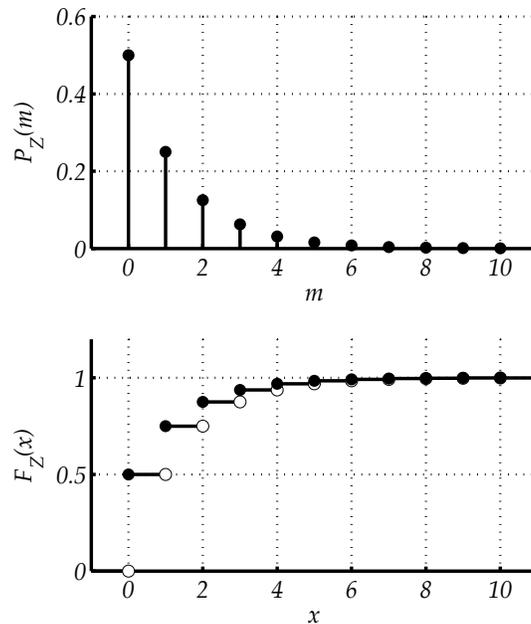
όπου φυσικά έχουμε χρησιμοποιήσει το γεγονός ότι τα αποτελέσματα των διαδοχικών ρίψεων είναι ανεξάρτητα. Η συνάρτηση κατανομής F_Z της Z μπορεί εύκολα να υπολογιστεί όπως πιο πάνω για τις X_i . Αφού η Z παίρνει μόνο τις τιμές $1, 2, 3, \dots$, για $x < 1$ έχουμε $F_Z(x) = \Pr(X \leq x) = 0$, και για οποιοδήποτε $x \geq 1$,

$$F_Z(x) = \Pr(X \leq x) = \Pr(X \leq \lfloor x \rfloor),$$

όπου το $\lfloor x \rfloor$ συμβολίζει το ακέραιο μέρος ενός πραγματικού αριθμού x . Συνεπώς, η $F_Z(x)$ ισούται με $1 - \Pr(X > \lfloor x \rfloor)$, και παρατηρούμε ότι το να έχουμε $X > k$ για κάποιο k είναι ισοδύναμο με τον να φέραμε k φορές Γ στις πρώτες k ρίψεις. Άρα, τελικά έχουμε,

$$F_Z(x) = 1 - \Pr(\lfloor x \rfloor \text{ φορές } \Gamma) = 1 - (1-p)^{\lfloor x \rfloor}, \quad \text{για κάθε } x \geq 1.$$

Οι γραφικές αναπαραστάσεις της πυκνότητας $P_Z(m)$ και της συνάρτησης κατανομής $F_Z(x)$ δίνονται στο Σχήμα 6.2.



Σχήμα 6.2: Γραφική αναπαράσταση της πυκνότητας και της συνάρτησης κατανομής της Τ.Μ. Z .

Παρατήρηση: Στο πιο πάνω παράδειγμα, τι θα συνέβαινε αν φέρναμε Γ επ' άπειρο; Αν, δηλαδή, όλα τα X_i ήταν ίσα με μηδέν, τότε θα είχαμε $Z = \infty$; Εκ πρώτης όψευς φαίνεται να έχουμε παραλείψει αυτή την εκδοχή, αλλά στην πραγματικότητα το ενδεχόμενο $\{Z = \infty\}$ έχει πιθανότητα μηδέν, όπως εύκολα αποδεικνύεται. Για παράδειγμα, από τον κανόνα πιθανότητας #2 βλέπουμε πως,

$$\Pr(Z = \infty) \leq \Pr(Z > m) = 1 - F_Z(m) = (1 - p)^m,$$

για οποιοδήποτε $m \geq 1$, και το δεξί μέρος πιο πάνω τείνει στο μηδέν καθώς $m \rightarrow \infty$. Άρα, αναγκαστικά έχουμε $\Pr(Z = \infty) = 0$ και κατά συνέπεια μπορούμε ασφαλώς να παραλείψουμε αυτό το ενδεχόμενο από το σύνολο τιμών της Z .

Παράδειγμα 6.5 Συνεχίζοντας το Παράδειγμα 6.4, έστω τώρα πως το νόμισμα είναι δίκαιο (δηλαδή $p = 1/2$), και πως κάποιος μας προτείνει να παίξουμε το εξής παιχνίδι: Θα του δώσουμε 2.50 ευρώ και θα στρίψουμε το νόμισμα μέχρι την πρώτη φορά που θα φέρουμε Κ. Αν φέρουμε Κ με την πρώτη, θα μας δώσει ένα ευρώ. Αν φέρουμε Γ και μετά Κ, θα μας δώσει δύο ευρώ. Και, γενικά, αν φέρουμε $(m - 1)$ φορές Γ και για πρώτη φορά φέρουμε Κ τη φορά m , θα μας δώσει m ευρώ (βλ. Ενότητα 2.1).

Είναι ή όχι δίκαιη η αμοιβή των δυόμισι ευρώ; Ποια θα ήταν μια δίκαιη αμοιβή; Πώς θα μπορούσαμε να την υπολογίσουμε;

Πριν απαντήσουμε, παρατηρούμε πως, γνωρίζοντας την πυκνότητα P_Z της Z από το προηγούμενο παράδειγμα, μπορούμε εύκολα να υπολογίσουμε κάποιες απλές πιθανότητες. Για παράδειγμα, η πιθανότητα να πάρουμε σ' αυτό το παιχνίδι μεταξύ 2 και 5 ευρώ είναι,

$$\begin{aligned} \Pr(2 \leq X \leq 5) &= \Pr(\{X = 2\} \cup \{X = 3\} \cup \{X = 4\} \cup \{X = 5\}) \\ &= \Pr(X = 2) + \Pr(X = 3) + \Pr(X = 4) + \Pr(X = 5) \\ &= P_Z(2) + P_Z(3) + P_Z(4) + P_Z(5) \\ &= \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^4 + \left(\frac{1}{2}\right)^5 \\ &= \frac{15}{32}. \end{aligned}$$

Επιστρέφοντας στο θέμα της δίκαιης αμοιβής, μια λογική σκέψη είναι η ακόλουθη. Ας φανταστούμε ότι παίζουμε το ίδιο παιχνίδι πολλές φορές συνεχόμενα. Αφού το αποτέλεσμα εξαρτάται από τις (τυχαίες) ρίψεις του νομίσματος, κάθε φορά θα παίρνουμε ένα διαφορετικό ποσό. Μακροπρόθεσμα, μια «δίκαιη» αμοιβή θα μπορούσε να θεωρηθεί ο μέσος όρος του ποσού που παίρνουμε ανά παιχνίδι.

Μεταφράζοντας την πιο πάνω σκέψη με λίγο μεγαλύτερη μαθηματική ακρίβεια, κατ' αρχάς παρατηρούμε πως το ποσό που θα πάρουμε σ' αυτό το τυχερό παιχνίδι ισούται με την Γ.Μ. Z του προηγούμενου παραδείγματος. Εκεί υπολογίσαμε την πυκνότητά της Z στη σχέση (6.2), οπότε θέτοντας $p = 1/2$, έχουμε,

$$P_Z(m) = \Pr(Z = m) = \Pr(\text{«παίρνουμε } m \text{ ευρώ»}) = \left(\frac{1}{2}\right)^m, \quad \text{για κάθε } m = 1, 2, \dots$$

Διαισθητικά, τη φράση «με πιθανότητα $1/2$ θα πάρουμε 1 ευρώ» μπορούμε να την ερμηνεύσουμε ως «μακροπρόθεσμα, τις μισές φορές θα πάρουμε 1 ευρώ» παρομοίως, για κάθε $m \geq 1$ περιμένουμε πως, μακροπρόθεσμα, το ποσοστό των φορών που θα πάρουμε m ευρώ θα είναι $P(m)$. Άρα, το κατά-μέσο-όρο-μακροπρόθεσμο ποσό που θα πάρουμε θα ισούται με,

$$1 \cdot P_Z(1) + 2 \cdot P_Z(2) + \dots + m \cdot P_Z(m) + \dots = \sum_{m=1}^{\infty} m P_Z(m) = \sum_{m=1}^{\infty} m \frac{1}{2^m}. \quad (6.3)$$

Αφού το άθροισμα της πιο πάνω σειράς είναι ίσο με 2, συμπεραίνουμε πως τα 2 ευρώ θα ήταν μια δίκαιη αμοιβή.²

Ο συλλογισμός που μας οδήγησε στον υπολογισμό (6.3) μας οδηγεί και στον πιο κάτω γενικό ορισμό.

²Για τον λεπτομερή υπολογισμό των σειρών $\sum_{m=1}^{\infty} m 2^{-m}$, $\sum_{m=1}^{\infty} m^2 2^{-m}$, και για άλλα σχετικά αποτελέσματα, βλ. την Ενότητα 7.3 του επόμενου κεφαλαίου.

6.2 Μέση τιμή, διασπορά, ανεξαρτησία

Ορισμός 6.4 Η μέση τιμή (ή αναμενόμενη τιμή, ή προσδοκώμενη τιμή) μιας διακριτής Τ.Μ. X με σύνολο τιμών το S και πυκνότητα $P(x)$, ορίζεται ως:

$$\mu = E(X) = \sum_{x \in S} xP(x). \quad (6.4)$$

Γενικότερα, για οποιαδήποτε συνάρτηση $f : S \rightarrow \mathbb{R}$, η μέση τιμή της νέας Τ.Μ. $f(X)$ ορίζεται ως:

$$E[f(X)] = \sum_{x \in S} f(x)P(x). \quad (6.5)$$

[Στην Άσκηση 1 στο τέλος του κεφαλαίου θα δείξετε πως ο δεύτερος ορισμός είναι στην πραγματικότητα συνέπεια του πρώτου.]

Παράδειγμα 6.6 Όπως υπολογίσαμε στο Παράδειγμα 6.5, η μέση τιμή της Τ.Μ. Z είναι,

$$E(Z) = \sum_{m=1}^{\infty} mP_Z(m) = \sum_{m=1}^{\infty} m2^{-m} = 2.$$

Αν, τώρα, το ποσό που παίρναμε στο τέλος του παιχνιδιού ήταν Z^2 και όχι Z , τότε θα μπορούσαμε να υπολογίσουμε τη δίκαιη αμοιβή για το νέο παιχνίδι από τη μέση τιμή του Z^2 , ως,³

$$E(Z^2) = \sum_{m=1}^{\infty} m^2P_Z(m) = \sum_{m=1}^{\infty} m^22^{-m} = 6.$$

Ορισμός 6.5 Η διασπορά μιας διακριτής Τ.Μ. X με μέση τιμή μ είναι:

$$\sigma^2 = \text{Var}(X) = E[(X - \mu)^2]. \quad (6.6)$$

Η τυπική απόκλιση της X είναι: $\sigma = \sqrt{\text{Var}(X)}$.

Παρατηρήσεις

1. Διαισθητικά, η μέση τιμή μ μιας Τ.Μ. X μας λέει ότι οι τιμές της X «τείνουν να κυμαίνονται» γύρω από την τιμή μ . Παρομοίως, η διασπορά σ^2 της X δίνει μια ποσοτική ένδειξη του πόσο μεγάλες τείνουν να είναι αυτές οι διακυμάνσεις. Συγκεκριμένα, από τον ορισμό βλέπουμε ότι η διασπορά είναι η «μέση τετραγωνική απόκλιση» της X από το μ , δηλαδή η μέση τιμή του τετραγώνου της απόστασης της τυχαίας τιμής X από τη μέση τιμή της.

2. Συνδυάζοντας τον ορισμό της διασποράς (6.6) με τον γενικό ορισμό της μέσης τιμής (6.5), προκύπτει ότι, για κάθε διακριτή Τ.Μ. X με σύνολο τιμών S και πυκνότητα $P(x)$, η διασπορά ισούται με:

$$\sigma^2 = \text{Var}(X) = \sum_{x \in S} (x - \mu)^2 P(x). \quad (6.7)$$

3. Από την έκφραση (6.7) αμέσως προκύπτει πως η διασπορά πάντοτε είναι $\text{Var}(X) \geq 0$. Επιπλέον, η μόνη περίπτωση η διασπορά να ισούται με μηδέν είναι αν όλοι οι όροι του άθροίσματος, $(x - \mu)^2 P(x)$, είναι μηδενικοί, πράγμα που μπορεί να συμβεί μόνο αν η Τ.Μ. X ισούται με τη σταθερά μ με πιθανότητα ένα! Στην Άσκηση 3 στο τέλος του κεφαλαίου θα δούμε μια ενδιαφέρουσα ιδιότητα τέτοιου είδους τετριμμένων «ντετερμινιστικών» Τ.Μ.
4. Ξεκινώντας από τη σχέση (6.7), έχουμε ότι η διασπορά $\text{Var}(X)$ μπορεί να εκφραστεί ως,

$$\begin{aligned} \sum_{x \in S} (x - \mu)^2 P(x) &= \sum_{x \in S} (x^2 - 2x\mu + \mu^2) P(x) \\ &= \sum_{x \in S} x^2 P(x) - 2\mu \sum_{x \in S} x P(x) + \mu^2 \sum_{x \in S} P(x). \end{aligned}$$

Το πρώτο άθροισμα εδώ ισούται με $E(X^2)$ από την (6.5), το δεύτερο άθροισμα ισούται με $E(X) = \mu$ εξ ορισμού, και το τρίτο άθροισμα ισούται με 1 από την αντίστοιχη ιδιότητα της πυκνότητας. Συνεπώς, $\text{Var}(X) = E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2$, και άρα έχουμε αποδείξει πως η διασπορά μιας οποιασδήποτε διακριτής Τ.Μ. X με μέση τιμή μ μπορεί εναλλακτικά να εκφραστεί:

$$\sigma^2 = \text{Var}(X) = E(X^2) - \mu^2. \quad (6.8)$$

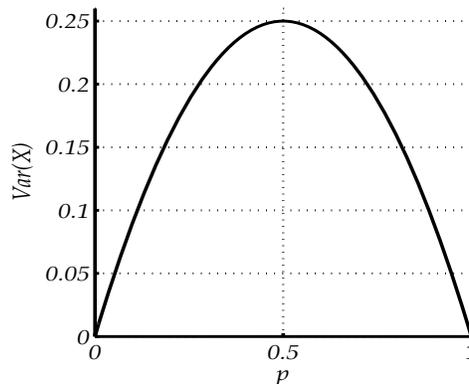
Παράδειγμα 6.7 Έστω μια Τ.Μ. X με την ίδια πυκνότητα όπως οι Τ.Μ. X_i του Παραδείγματος 6.4, δηλαδή $X = 1$ ή 0 , με πιθανότητα p ή $(1 - p)$, αντίστοιχα. Η X έχει σύνολο τιμών το $S = \{0, 1\}$ και πυκνότητα $P(1) = p$, $P(0) = 1 - p$, για κάποιο δεδομένο $p \in (0, 1)$.

Η μέση τιμή και η διασπορά της X εύκολα υπολογίζονται ως,

$$\begin{aligned} \mu &= E(X) = 0 \cdot P(0) + 1 \cdot P(1) = p, \\ \sigma^2 &= \text{Var}(X) = E(X^2) - \mu^2 = 0^2 \cdot P(0) + 1^2 \cdot P(1) - p^2 \\ &= p - p^2 = p(1 - p), \end{aligned}$$

όπου για τη διασπορά χρησιμοποιήσαμε την έκφραση (6.8).

Παρατηρούμε ότι τα αποτελέσματα αυτά συμφωνούν με τη διαίσθησή μας πως, όσο μεγαλώνει η πιθανότητα p με την οποία έχουμε $X = 1$, τόσο πιο πολύ τείνει η Τ.Μ. να παίρνει τη μεγαλύτερη από τις δύο τιμές της, και αντιστοίχως μεγαλώνει και η μέση τιμή της $E(X) = p$. Παρομοίως, στο Σχήμα 6.3 όπου δίνεται η γραφική αναπαράσταση της διασποράς της X ως συνάρτηση της παραμέτρου p , βλέπουμε πως η διασπορά είναι μεγαλύτερη όσο πιο κοντά είναι το p στο $1/2$, άρα έχουμε μεγάλη διασπορά όταν οι τιμές της Τ.Μ. είναι πιο «τυχαίες» ή λιγότερο «προβλέψιμες».



Σχήμα 6.3: Γραφική αναπαράσταση της διασποράς μιας δυαδικής Τ.Μ. X (με τιμές $X = 1$ με πιθαν. p και $X = 0$ με πιθαν. $(1 - p)$) ως συνάρτηση της παραμέτρου p .

Παράδειγμα 6.8 Για την Τ.Μ. Z στο Παράδειγμα 6.5 βρήκαμε πως η μέση τιμή της είναι $\mu = E(Z) = 2$, και στο Παράδειγμα 6.6 υπολογίσαμε πως $E(Z^2) = 6$. Συνεπώς, από τη σχέση (6.8), η διασπορά της Z ισούται με,

$$\sigma^2 = \text{Var}(Z) = E(Z^2) - [E(Z)]^2 = 6 - 2^2 = 2.$$

Παράδειγμα 6.9 Σε ένα τυχερό παιχνίδι, το κέρδος μας, έστω X , ισούται με 1 ευρώ ή με -3 ευρώ, με πιθανότητες $2/3$ και $1/3$ αντίστοιχα. Άρα η μέση τιμή του κέρδους μας είναι,

$$\mu = E(X) = 1 \cdot \frac{2}{3} + (-3) \cdot \frac{1}{3} = -\frac{1}{3},$$

και συμπεραίνουμε πως, αν παίξουμε αυτό το παιχνίδι πολλές φορές, μακροπρόθεσμα θα βγούμε χαμένοι. Επιπλέον, από τον τύπο (6.8), το X έχει διασπορά,

$$\text{Var}(X) = E(X^2) - \mu^2 = 1^2 \cdot \frac{2}{3} + (-3)^2 \cdot \frac{1}{3} - \left(-\frac{1}{3}\right)^2 = \frac{32}{9} \approx 3.5556,$$

και τυπική απόκλιση $\sigma = \sqrt{32/9} \approx 1.886$.

Στο Θεώρημα 6.1 θα εξετάσουμε τη μέση τιμή και τη διασπορά του αθροίσματος δύο τυχαίων μεταβλητών. Για τη διατύπωση των σχετικών ιδιοτήτων θα χρειαστούμε την εξής γενίκευση της έννοιας της ανεξαρτησίας:

Ορισμός 6.6 (Ανεξάρτητες Τ.Μ.) 1. Δύο διακριτές Τ.Μ. X και Y με σύνολα τιμών S_X, S_Y αντίστοιχα είναι **ανεξάρτητες** αν, για κάθε $x \in S_X, y \in S_Y$ τα ενδεχόμενα $\{X = x\}$ και $\{Y = y\}$ είναι ανεξάρτητα. Ισοδύναμα, οι X, Y είναι ανεξάρτητες αν και μόνο αν,

$$\Pr(X = x, Y = y) = \Pr(X = x) \Pr(Y = y),$$

για κάθε $x \in S_X, y \in S_Y$, όπου, χάριν συντομίας, συμβολίζουμε ως $\Pr(X = x, Y = y)$ την πιθανότητα του ενδεχομένου $\{X = x \text{ και } Y = y\}$.

2. Οι διακριτές Τ.Μ. X_1, X_2, \dots, X_N είναι **ανεξάρτητες** αν και μόνο αν,

$$\Pr(X_1 = x_1, X_2 = x_2, \dots, X_N = x_N) = \Pr(X_1 = x_1) \Pr(X_2 = x_2) \cdots \Pr(X_N = x_N),$$

για κάθε N -άδα τιμών $x_1 \in S_{X_1}, x_2 \in S_{X_2}, \dots, x_N \in S_{X_N}$, όπου S_{X_i} είναι το σύνολο τιμών της κάθε X_i .

3. Οι διακριτές Τ.Μ. X_i σε μια άπειρη ακολουθία X_1, X_2, \dots είναι **ανεξάρτητες**, αν οι Τ.Μ. X_1, X_2, \dots, X_N είναι ανεξάρτητες για κάθε $N \geq 1$.

Οι παρακάτω ιδιότητες αποτελούν κάποια από τα βασικότερα εργαλεία των πιθανοτήτων.

Θεώρημα 6.1 (Ιδιότητες μέσης τιμής και διασποράς)

Για οποιεσδήποτε διακριτές Τ.Μ. X, Y και σταθερές $a, b \in \mathbb{R}$:

$$1. E(aX + bY) = aE(X) + bE(Y).$$

$$2. \text{Var}(aX + b) = a^2 \text{Var}(X).$$

Αν οι Τ.Μ. X, Y είναι ανεξάρτητες, τότε:

$$3. E(XY) = E(X) E(Y).$$

$$4. \text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y).$$

$$5. \Pr(X \in T \text{ και } Y \in T') = \Pr(X \in T) \Pr(Y \in T'), \text{ για οποιαδήποτε υποσύνολα } T \subset S_X, T' \subset S_Y.$$

Παρατήρηση: Η πρώτη ιδιότητα του θεωρήματος εύκολα επεκτείνεται και για οποιοδήποτε (πεπερασμένο) πλήθος N τυχαίων μεταβλητών. Για παράδειγμα, για $N = 3$, χρησιμοποιώντας την Ιδιότητα 1 δύο φορές βρίσκουμε,

$$E(X_1 + X_2 + X_3) = E[(X_1 + X_2) + X_3] = E(X_1 + X_2) + E(X_3) = E(X_1) + E(X_2) + E(X_3).$$

Παρομοίως, έχουμε τη γενική περίπτωση,

$$E\left(\sum_{i=1}^N X_i\right) = \sum_{i=1}^n E(X_i).$$

[Στο ίδιο πλαίσιο, μια χρήσιμη άσκηση είναι η εξής: Διατυπώστε και αποδείξτε την προφανή γενίκευση της Ιδιότητας 4 για N τυχαίες μεταβλητές.]

Πριν δώσουμε την απόδειξη, θα δούμε ένα παράδειγμα της χρήσης κάποιων από αυτές τις ιδιότητες και ένα χρήσιμο τεχνικό αποτέλεσμα.

Παράδειγμα 6.10 Παίζουμε 5 φορές «Κορώνα-Γράμματα» με ένα δίκαιο νόμισμα, όπου κερδίζουμε 1 ευρώ κάθε φορά που έρχεται Κ και χάνουμε 1 ευρώ όποτε έρχεται Γ. Έστω οι Τ.Μ. X_1, X_2, X_3, X_4, X_5 όπου κάθε $X_i = \pm 1$ με πιθανότητα $1/2$. Εξετάζουμε 2 παραλλαγές του παιχνιδιού:

A. Παίζουμε 5 φορές κανονικά το παιχνίδι, με συνολικό κέρδος $Y = X_1 + X_2 + X_3 + X_4 + X_5$.

B. Παίζουμε μόνο μία φορά, αλλά κερδίζοντας ή χάνοντας 5 ευρώ αντί για ένα, οπότε έχουμε συνολικό κέρδος $Z = 5X_1$.

Από τον ορισμό των X_i έχουμε, για κάθε i ,

$$\begin{aligned} E(X_i) &= 1 \cdot (1/2) + (-1) \cdot (1/2) = 0, \\ \text{Var}(X_i) &= E(X_i^2) - 0^2 = 1^2 \cdot (1/2) + (-1)^2 \cdot (1/2) = 1. \end{aligned}$$

Χρησιμοποιώντας την πρώτη ιδιότητα του θεωρήματος, μπορούμε να υπολογίσουμε το «προσδοκώμενο» κέρδος,

$$\begin{aligned} E(Y) &= E(X_1 + X_2 + X_3 + X_4 + X_5) \\ &= E(X_1) + E(X_2) + E(X_3) + E(X_4) + E(X_5) \\ &= 0, \\ E(Z) &= E(5X_1) = 5E(X_1) \\ &= 0, \end{aligned}$$

άρα και στις δύο περιπτώσεις το μέσο κέρδος είναι μηδενικό.

Για τις διασπορές των δύο παιχνιδιών, από τις Ιδιότητες 2 και 4 βρίσκουμε, αντίστοιχα,

$$\begin{aligned} \text{Var}(Y) &= \text{Var}(X_1 + X_2 + X_3 + X_4 + X_5) \\ &= \text{Var}(X_1) + \text{Var}(X_2) + \text{Var}(X_3) + \text{Var}(X_4) + \text{Var}(X_5) \\ &= 5, \end{aligned}$$

και $\text{Var}(Z) = \text{Var}(5X_1) = 5^2 \text{Var}(X_1) = 25$. Παρατηρούμε λοιπόν πως, αν και τα δύο παιχνίδια έχουν την ίδια αναμενόμενη απόδοση, το δεύτερο έχει πολύ μεγαλύτερες διακυμάνσεις, συνεπώς και πολύ μεγαλύτερο ρίσκο.

Λήμμα 6.1 Για οποιεσδήποτε δύο διακριτές Τ.Μ. X, Y και για κάθε $x \in S_X$ έχουμε,

$$\Pr(X = x) = \sum_{y \in S_Y} \Pr(X = x, Y = y),$$

όπου, πάλι χάρην συντομίας, συμβολίζουμε την πιθανότητα του ενδεχομένου $\{X = x \text{ και } Y = y\}$ ως $\Pr(X = x, Y = y)$.

Απόδειξη του Λήμματος 6.1: Η απόδειξη είναι παρόμοια με εκείνη του κανόνα συνολικής πιθανότητας στο Κεφάλαιο 5. Έστω ένα δεδομένο στοιχείο $x \in S_X$. Για κάθε $y \in S_Y$ ορίζουμε το ενδεχόμενο,

$$E_y = \{X = x \text{ και } Y = y\},$$

και παρατηρούμε πως,

$$\{X = x\} = \bigcup_{y \in S_Y} \{X = x, Y = y\} = \bigcup_{y \in S_Y} E_y.$$

Εφόσον τα E_y είναι όλα ξένα μεταξύ τους, το αποτέλεσμα προκύπτει από τον κανόνα πιθανότητας #3. ■

Απόδειξη του Θεωρήματος 6.1: 1. Για να υπολογίσουμε τη μέση τιμή του $aX + bY$ παρατηρούμε ότι η νέα διακριτή Τ.Μ. $Z = aX + bY$ έχει σύνολο τιμών,

$$S_Z = \{ax + by : \text{για } x \in S_X, y \in S_Y\},$$

και παίρνει κάθε τιμή στο S_Z με αντίστοιχη πιθανότητα, $\Pr(X = x, Y = y)$. Άρα,

$$\begin{aligned} E(aX + bY) &= \sum_{x \in S_X, y \in S_Y} (ax + by) \Pr(X = x, Y = y) \\ &= a \sum_{x \in S_X, y \in S_Y} x \Pr(X = x, Y = y) \\ &\quad + b \sum_{x \in S_X, y \in S_Y} y \Pr(X = x, Y = y) \\ &= a \sum_{x \in S_X} \left[x \sum_{y \in S_Y} \Pr(X = x, Y = y) \right] \\ &\quad + b \sum_{y \in S_Y} \left[y \sum_{x \in S_X} \Pr(X = x, Y = y) \right] \\ &\stackrel{(a)}{=} a \sum_{x \in S_X} x \Pr(X = x) + b \sum_{y \in S_Y} y \Pr(Y = y) \\ &= aE(X) + bE(Y), \end{aligned}$$

όπου στο βήμα (a) χρησιμοποιήσαμε το Λήμμα 6.1.

2. Εφαρμόζοντας τη σχέση (6.8) στην Τ.Μ. $(aX + b)$ έχουμε,

$$\begin{aligned}
 \text{Var}(aX + b) &= E[(aX + b)^2] - (E[aX + b])^2 \\
 &\stackrel{(b)}{=} E[a^2X^2 + 2abX + b^2] - [aE(X) + b]^2 \\
 &\stackrel{(c)}{=} a^2E(X^2) + 2abE(X) + b^2 - a^2(E(X))^2 - 2abE(X) - b^2 \\
 &= a^2(E(X^2) - (E(X))^2) \\
 &\stackrel{(d)}{=} a^2\text{Var}(X),
 \end{aligned}$$

όπου στο βήμα (b) εφαρμόσαμε την Ιδιότητα 1 στον δεύτερο όρο του αθροίσματος, στο βήμα (c) εφαρμόσαμε την Ιδιότητα 1 στον πρώτο όρο, και στο βήμα (d) χρησιμοποιήσαμε πάλι τη σχέση (6.8).

Έστω τώρα ότι οι X, Y είναι ανεξάρτητες.

3. Με το ίδιο σκεπτικό όπως στην απόδειξη της πρώτης ιδιότητας, υπολογίζουμε τη μέση τιμή της Τ.Μ. XY ως,

$$\begin{aligned}
 E(XY) &= \sum_{x \in S_X, y \in S_Y} xy \Pr(X = x, Y = y) \\
 &\stackrel{(e)}{=} \sum_{x \in S_X, y \in S_Y} xy \Pr(X = x) \Pr(Y = y) \\
 &= \left[\sum_{x \in S_X} x \Pr(X = x) \right] \left[\sum_{y \in S_Y} y \Pr(Y = y) \right] \\
 &= E(X)E(Y),
 \end{aligned}$$

όπου στο βήμα (e) χρησιμοποιήσαμε τον ορισμό της ανεξαρτησίας δύο τυχαίων μεταβλητών.

4. Εφαρμόζοντας τη σχέση (6.8) στην Τ.Μ. $(X + Y)$ έχουμε,

$$\text{Var}(X + Y) = E[(X + Y)^2] - (E[X + Y])^2,$$

και από την Ιδιότητα 1,

$$\begin{aligned}
 \text{Var}(X + Y) &= E[X^2 + 2XY + Y^2] - (E(X) + E(Y))^2 \\
 &= E(X^2) + 2E(XY) + E(Y^2) - (E(X))^2 - 2E(X)E(Y) - (E(Y))^2 \\
 &\stackrel{(f)}{=} E(X^2) - (E(X))^2 + E(Y^2) - (E(Y))^2 \\
 &\stackrel{(g)}{=} \text{Var}(X) + \text{Var}(Y),
 \end{aligned}$$

όπου χρησιμοποιήσαμε στο βήμα (f) την Ιδιότητα 3 και στο βήμα (g) τη σχέση (6.8).

5. Ο υπολογισμός για την απόδειξη της τελευταίας ιδιότητας είναι παρόμοιος με εκείνον στην Ιδιότητα 3. Παρατηρώντας ότι τα ενδεχόμενα $\{X = x, Y = y\}$, που αντιστοιχούν σε διαφορετικά ζεύγη τιμών (x, y) , είναι ξένα και χρησιμοποιώντας τον ορισμό της ανεξαρτησίας,

$$\begin{aligned} \Pr(X \in T, Y \in T') &= \Pr\left(\bigcup_{x \in T, y \in T'} \{X = x, Y = y\}\right) \\ &= \sum_{x \in T, y \in T'} \Pr(X = x, Y = y) \\ &= \sum_{x \in T, y \in T'} \Pr(X = x) \Pr(Y = y) \\ &= \left[\sum_{x \in T} \Pr(X = x) \right] \left[\sum_{y \in T'} \Pr(Y = y) \right] \\ &= \Pr(X \in T) \Pr(Y \in T'), \end{aligned}$$

ολοκληρώνουμε την απόδειξη. ■

6.3 * Μετρησιμότητα και άπειρες τιμές

Κλείνουμε αυτό το κεφάλαιο με ορισμένες επιπλέον λεπτομέρειες, από τη σκοπιά της αυστηρής μαθηματικής προσέγγισης, για κάποιους από του ορισμούς και τα αποτελέσματα που είδαμε.

6.3.1 Ορισμός τυχαίας μεταβλητής

Στην Ενότητα 3.3 του τρίτου κεφαλαίου σημειώσαμε πως, σε κάποιες περιπτώσεις που ο χώρος πιθανότητας δεν είναι πεπερασμένος ή αριθμήσιμος, το μέτρο πιθανότητας δεν μπορεί να οριστεί για όλα τα υποσύνολα του Ω . Ένα αντίστοιχο φαινόμενο εμφανίζεται και στις τυχαίες μεταβλητές. Σε περιπτώσεις που ο χώρος πιθανότητας είναι μη αριθμήσιμος, μπορεί να υπάρχουν συναρτήσεις $X : \Omega \rightarrow \mathbb{R}$ για τις οποίες κάποιες πιθανότητες της μορφής $\Pr(X \leq a)$ να μην μπορούν να ορισθούν. Αυτό συμβαίνει διότι κάποια από τα αντίστοιχα ενδεχόμενα,

$$\{X \leq a\} = \{\omega \in \Omega \text{ τέτοια ώστε } X(\omega) \leq a\} \subset \Omega,$$

μπορεί να είναι παθολογικά (με την έννοια που περιγράψαμε στην Ενότητα 3.3), ή μη μετρήσιμα όπως λέγονται στα μαθηματικά. Τέτοιου είδους παθολογικές συναρτήσεις ονομάζονται «μη μετρήσιμες» και δεν τις θεωρούμε τυχαίες μεταβλητές.

Η γενική λύση σε αυτό το πρόβλημα είναι να περιορίσουμε τις συναρτήσεις $X : \Omega \rightarrow \mathbb{R}$ τις οποίες χρησιμοποιούμε ως τυχαίες μεταβλητές. Ο ακριβής προσδιορισμός αυτών των «καλών» συναρτήσεων αποτελεί μέρος της περιοχής της μαθηματικής ανάλυσης που ονομάζεται *θεωρία*

μέτρου, με την οποία δεν θα ασχοληθούμε περαιτέρω. Περισσότερες πληροφορίες για αυτά τα ζητήματα μπορείτε να βρείτε σε πιο προχωρημένα βιβλία μαθηματικής ανάλυσης ή πιθανοτήτων.

Ο τρόπος με τον οποίο εμείς θα αποφυγούμε τέτοιου είδους παθολογίες είναι ορίζοντας τις Τ.Μ. που θέλουμε να χρησιμοποιήσουμε μέσω της πυκνότητάς τους. Έστω, για παράδειγμα, πως θέλουμε να χρησιμοποιήσουμε μια Τ.Μ. X η οποία να παίρνει τις τιμές,

$$X = \begin{cases} 1, & \text{με πιθαν. } 1/2, \\ 2, & \text{με πιθαν. } 1/4, \\ 3, & \text{με πιθαν. } 1/8, \\ 4, & \text{με πιθαν. } 1/8, \end{cases}$$

δηλαδή να έχει σύνολο τιμών $S = \{1, 2, 3, 4\}$ και πυκνότητα $P(x)$ με $P(1) = 1/2$, $P(2) = 1/4$, $P(3) = P(4) = 1/8$. Πώς μπορεί να οριστεί μια τέτοια Τ.Μ. X ως συνάρτηση $X : \Omega \rightarrow \mathbb{R}$ σε κάποιο χώρο πιθανότητας Ω ; Απλά θέτοντας $\Omega = S$, ορίζοντας $X(\omega) = \omega$ για κάθε $\omega = 1, 2, 3, 4$, και ορίζοντας ένα μέτρο πιθανότητας που για τα στοιχειώδη ενδεχόμενα θα έχει,

$$\begin{aligned} \mathbb{P}(\{1\}) &= P(1) = 1/2, & \mathbb{P}(\{2\}) &= P(2) = 1/4, \\ \mathbb{P}(\{3\}) &= P(3) = 1/8, & \mathbb{P}(\{4\}) &= P(4) = 1/8. \end{aligned}$$

Τότε, για καθεμία από τις τιμές $k = 1, 2, 3, 4$ της X ,

$$\Pr(X = k) = \Pr(\{\omega \text{ τέτοια ώστε } X(\omega) = k\}) = \mathbb{P}(\{k\}) = P(k),$$

όπως και ήταν το ζητούμενο.

Γενικά, μια οποιαδήποτε διακριτή Τ.Μ. με δεδομένη πυκνότητα P σε κάποιο σύνολο τιμών S μπορεί να οριστεί ως συνάρτηση σε κάποιον διακριτό χώρο πιθανότητας Ω με τον ίδιο τρόπο. Θέτουμε $\Omega = S$, ορίζουμε το μέτρο πιθανότητας \mathbb{P} για όλα τα στοιχειώδη ενδεχόμενα ως,

$$\mathbb{P}(\{\omega\}) = P(\omega), \quad \text{για κάθε } \omega \in S = \Omega,$$

και τέλος θέτουμε $X(\omega) = \omega$. Έτσι, η X πράγματι παίρνει την κάθε τιμή $x \in S$ με πιθανότητα $P(x)$:

$$\Pr(X = x) = \Pr(\{\omega \text{ τέτοια ώστε } X(\omega) = x\}) = \mathbb{P}(\{x\}) = P(x).$$

Μια αντίστοιχη κατασκευή θα μας επιτρέψει να ορίσουμε και όλες τις συνεχείς Τ.Μ. που θα χρειαστούμε αργότερα, βλ. Ενότητα 10.3 του Κεφαλαίου 10.

6.3.2 Ορισμοί $E(X)$, $\text{Var}(X)$, και επιπλέον συνθήκες

Έστω μια διακριτή Τ.Μ. X με πυκνότητα P σε κάποιο σύνολο τιμών S . Αν το S δεν είναι πεπερασμένο (όπως π.χ. για την Τ.Μ. Z του Παραδείγματος 6.5), τότε η μέση τιμή $E(X)$ της X δίνεται από ένα άπειρο άθροισμα, το οποίο όμως, όπως γνωρίζουμε από τα βασικά αποτελέσματα

του διαφορικού λογισμού, μπορεί να παίρνει τις τιμές $\pm\infty$, ή και να μην ορίζεται καν. (Για ένα παράδειγμα μιας διακριτής Τ.Μ. με άπειρη μέση τιμή δείτε την Άσκηση 4 στο τέλος του κεφαλαίου.) Για να αποφύγουμε τεχνικά ζητήματα τα οποία ξεφεύγουν κατά πολύ από τα ζητούμενα του παρόντος βιβλίου, υιοθετούμε τις πιο κάτω συμβατικές υποθέσεις, οι οποίες θα παραμείνουν εν ισχύ σε όλα τα υπόλοιπα κεφάλαια.

Συμβάσεις

- Πάντοτε, όταν λέμε πως «η διακριτή Τ.Μ. X έχει μέση τιμή $\mu = E(X)$ », εμμέσως υποθέτουμε ότι το άθροισμα της σειράς που δίνει την $E(X)$ ορίζεται, και ότι η τιμή του είναι πεπερασμένη.
- Πάντοτε, όταν λέμε ότι «η διακριτή Τ.Μ. X έχει διασπορά $\sigma^2 = \text{Var}(X)$ », εμμέσως υποθέτουμε ότι η μέση τιμή $\mu = E(X)$ ορίζεται και είναι πεπερασμένη, και πως το άθροισμα της σειράς που δίνει την $\text{Var}(X)$ ορίζεται και δίνει πεπερασμένο αποτέλεσμα.
- Όποτε διατυπώνεται μια ιδιότητα για τη μέση τιμή (ή τη διασπορά), εμμέσως υποθέτουμε πως η αντίστοιχη μέση τιμή (αντίστοιχα, διασπορά) ορίζεται και είναι πεπερασμένη.

Για παράδειγμα, η εναλλακτική έκφραση για τη διασπορά που δώσαμε στη σχέση (6.8) ισχύει εφόσον η μέση τιμή μ και η διασπορά σ^2 ορίζονται και είναι πεπερασμένες. Παρομοίως, η πρώτη ιδιότητα του Θεωρήματος 6.1 ισχύει εφόσον οι μέσες τιμές $E(X)$ και $E(Y)$ ορίζονται και είναι πεπερασμένες. Αντίστοιχες υποθέσεις καλύπτουν και τις Ιδιότητες 2, 3 και 4.

Τέλος, σημειώνουμε πως η θεωρία μέτρου επιτρέπει τον ορισμό της μέσης τιμής και της διασποράς κάτω από πιο γενικές συνθήκες, και παρέχει μαθηματικά εργαλεία τα οποία μας επιτρέπουν να διατυπώσουμε γενικότερες μορφές των αντίστοιχων ιδιοτήτων. Αλλά αυτές οι γενικεύσεις αφορούν λεπτά μαθηματικά ζητήματα και επιπλέον δεν μας είναι απαραίτητες. Κατά συνέπεια, δεν θα επεκταθούμε προς αυτή την κατεύθυνση.

6.4 Ασκήσεις

1. ★ *Μέση τιμή συναρτήσεων T.M.* Στον Ορισμό 6.4 ορίσαμε τη μέση τιμή της $f(X)$ μέσω του τύπου (6.5). Αλλά η $Y = f(X)$ είναι και η ίδια μια τυχαία μεταβλητή, με τη δική της πυκνότητα, έστω $Q(y)$, και το δικό της σύνολο τιμών, έστω S' . Δείξτε πως, αν εφαρμόσουμε τον τύπο του ορισμού (6.4) για τη μέση τιμή της Y , το αποτέλεσμα ισούται με εκείνο του γενικότερου τύπου (6.5) για τη μέση τιμή της $f(X)$.
2. ★ *Η συνάρτηση κατανομής στο $\pm\infty$.* Χρησιμοποιώντας τα αποτελέσματα της Άσκησης 9 του Κεφαλαίου 3, αποδείξτε πως, για οποιαδήποτε διακριτή T.M. X με συνάρτηση κατανομής $F(x)$, πάντοτε έχουμε:

$$\lim_{x \rightarrow \infty} F(x) = 1, \quad \text{και} \quad \lim_{x \rightarrow -\infty} F(x) = 0.$$

Παρατηρήστε ότι ακριβώς η ίδια απόδειξη ισχύει για οποιαδήποτε T.M., όχι απαραίτητα διακριτή.

3. *Ντετερμινιστικές T.M.* Στην τρίτη παρατήρηση της Ενότητας 6.2 είδαμε πως μια T.M. έχει διασπορά $\text{Var}(X) = 0$, αν και μόνο αν παίρνει μία και μόνο μία τιμή, με πιθανότητα 1. Τέτοιες τετριμμένες περιπτώσεις T.M. ονομάζονται *ντετερμινιστικές*. Αποδείξτε πως μια ντετερμινιστική T.M. X είναι ανεξάρτητη από οποιαδήποτε άλλη διακριτή T.M. Y .
4. *Άπειρη μέση τιμή.* Έστω μια T.M. X με σύνολο τιμών το $S = \mathbb{N} = \{1, 2, \dots\}$ και πυκνότητα,

$$P(k) = \frac{C}{k^2}, \quad \text{για κάθε } k \geq 1.$$

(α') Υπολογίστε την τιμή της σταθεράς C . *Υπόδειξη.* $\sum_{k=1}^{\infty} (1/k^2) = \pi^2/6$.

(β') Αποδείξτε πως η μέση τιμή $E(X) = +\infty$.

5. ★ *Ανεξαρτησία και μέση τιμή.* Δώστε ένα παράδειγμα δύο διακριτών T.M. X, Y για τις οποίες ισχύει ότι,

$$E(XY) = E(X)E(Y),$$

αλλά δεν είναι ανεξάρτητες. [Σημείωση. Αυτή η άσκηση ίσως σας φανεί σημαντικά ευκολότερη αφού διαβάσετε το Κεφάλαιο 9.]

6. *Ανεξαρτησία και διασπορά.* Δώστε ένα παράδειγμα δύο διακριτών T.M. X, Y με διασπορές που να ικανοποιούν $\text{Var}(X) = \text{Var}(Y) > 0$, αλλά τέτοιες ώστε να έχουμε $\text{Var}(X+Y) = 0$.
7. *Δύο ζάρια.* Ρίχνουμε διαδοχικά δύο δίκαια ζάρια, και, όπως συνήθως, περιγράφουμε το αποτέλεσμα ως ένα από τα 36 στοιχεία του χώρου πιθανότητας:

$$\Omega = \{11, 12, \dots, 16, 21, \dots, 26, \dots, 61, \dots, 66\}.$$

Έστω X_1 και X_2 τα αποτελέσματα των δύο ρίψεων, όπου υποθέτουμε ότι οι ρίψεις είναι ανεξάρτητες μεταξύ τους. Επιπλέον, ορίζουμε τις τυχαίες μεταβλητές $Y = \max\{X_1, X_2\}$ και $Z = \min\{X_1, X_2\}$. Για καθεμία από τις Τ.Μ. X_1, X_2, Y, Z :

- (α') Βρείτε το σύνολο τιμών.
- (β') Περιγράψτε, για καθεμία από τις τιμές που μπορεί να πάρει η Τ.Μ., το αντίστοιχο ενδεχόμενο (π.χ. το $\{X_1 = 4\}$) ως υποσύνολο του χώρου πιθανότητας.
- (γ') Υπολογίστε την πυκνότητα και τη συνάρτηση κατανομής της.
- (δ') Υπολογίστε τη μέση τιμή και τη διασπορά της.

8. ★ *Πού να βρω γυναίκα να σου μοιάζει.* Σε κάποιο πάρτυ πηγαίνουν N ανδρόγυνα και χωρίζονται σε δύο διαφορετικά δωμάτια άντρες-γυναίκες. Φεύγοντας, οι γυναίκες πια έχουν μεθύσει, και η καθεμία παίρνει στην τύχη έναν από τους άντρες και πηγαίνει σπίτι της. Έστω X το πλήθος των σωστών ζευγαριών που φεύγουν από το πάρτυ. Να βρεθούν η μέση τιμή και η διασπορά της X . [Υπόδειξη. Ίσως σας φανεί βολικό να χρησιμοποιήσετε τις βοηθητικές τυχαίες μεταβλητές $X_i, i = 1, 2, \dots, N$, όπου $X_i = 1$ αν η γυναίκα i πάρει τον δικό της άντρα, και $X_i = 0$ στην αντίθετη περίπτωση.]

9. *Συνέλιξη.* Έστω δύο ανεξάρτητες διακριτές Τ.Μ. X, Y , που παίρνουν ακέραιες τιμές και έχουν πυκνότητες $P_X(x), P_Y(y)$, αντίστοιχα. Ναδειχθεί ότι η νέα Τ.Μ. $X + Y$ έχει πυκνότητα:

$$P_{X+Y}(m) = \sum_{k=-\infty}^{\infty} P_X(k)P_Y(m-k), \quad m \in \mathbb{Z}.$$

[Παρατήρηση. Η πιο πάνω έκφραση είναι γνωστή ως η συνέλιξη των δύο πυκνοτήτων $P_X(x)$ και $P_Y(y)$, και εμφανίζεται συχνά στα μαθηματικά. Στο Κεφάλαιο 15 θα δούμε και τη συνεχή εκδοχή της.]

10. *Άλλα δύο ζάρια.* Ρίχνουμε διαδοχικά δύο δίκαια ζάρια. Έστω X_1 και X_2 τα αποτελέσματα των δύο ρίψεων, οι οποίες υποθέτουμε πως είναι ανεξάρτητες. Έστω επίσης οι τυχαίες μεταβλητές $V = X_1 + X_2$ και $W = X_1 - X_2$.

Για τις V, W :

- (α') Βρείτε το σύνολο τιμών.
- (β') Περιγράψτε, για καθεμία από τις τιμές που μπορεί να πάρει η Τ.Μ., το αντίστοιχο ενδεχόμενο (π.χ. το $\{V = 2\}$) ως υποσύνολο του χώρου πιθανότητας.
- (γ') Υπολογίστε την πυκνότητα και τη συνάρτηση κατανομής της.
- (δ') Υπολογίστε τη μέση τιμή και τη διασπορά της, χωρίς να χρησιμοποιήσετε το Θεώρημα 6.1.

11. *Δέκα μπάλες.* Έστω πως από 10 μπάλες, οι οποίες είναι αριθμημένες από το 1 έως το 10, επιλέγουμε τυχαία 3 χωρίς επανατοποθέτηση. Έστω Y ο μέγιστος αριθμός από τις τρεις μπάλες που επιλέξαμε, και Z ο ελάχιστος. Να προσδιορίσετε την πυκνότητα και τη συνάρτηση κατανομής των Y και Z .
12. *Περιορισμοί στις παραμέτρους.* Μια Γ.Μ. X παίρνει τις τιμές $-2, -1, 0, 1,$ και 2 . Η πυκνότητά της $P(x)$ είναι:

$$P(-2) = P(2) = \frac{1}{10}, \quad P(-1) = P(1) = a, \quad P(0) = b.$$

- (α') Ποιες είναι οι επιτρεπτές τιμές του a και του b ; Ποια σχέση πρέπει να ικανοποιείται μεταξύ των a και b ;
- (β') Ποιο ζευγάρι τιμών (a, b) μεγιστοποιεί τη μέση τιμή της X ;
- (γ') Υπολογίστε τη διασπορά της X αν $a = 0$.
13. * *Επιζώντα ζευγάρια.* Έστω μια πόλη με n ζεύγη ανδρών-γυναικών (δηλαδή με συνολικό πληθυσμό $2n$ ατόμων.) Αν υποθέσουμε ότι m τυχαία άτομα πεθαίνουν, να υπολογιστεί η μέση τιμή του πλήθους των ζευγαριών που παραμένουν. [*Σημείωση.* Το πρόβλημα αυτό διατυπώθηκε για πρώτη φορά από τον Daniel Bernoulli, το 1768. *Υπόδειξη.* Όπως και στην Άσκηση 8, ίσως σας φανεί βολικό να χρησιμοποιήσετε κάποιες βοηθητικές τυχαίες μεταβλητές.]
14. *Τρία ζάρια.* Υπολογίστε την πυκνότητα και τη συνάρτηση κατανομής του αθροίσματος των αποτελεσμάτων τριών (ανεξάρτητων) δίκαιων ζαριών.
15. *Παιχνίδι.* Σε ένα παιχνίδι, ο παίκτης έχει τυχαίο κέρδος $X = -5, 0, 2,$ ή 10 ευρώ, με αντίστοιχες πιθανότητες $\frac{1}{6}, \frac{1}{2}, \frac{1}{6}, \frac{1}{6}$.

- (α') Υπολογίστε τη μέση τιμή και τη διασπορά του κέρδους του σε μία παρτίδα.
- (β') Υπολογίστε τη μέση τιμή και τη διασπορά του συνολικού του κέρδους σε τρεις ανεξάρτητες παρτίδες.

16. *Η μέθοδος της δεύτερης ροπής.* Έστω μια Γ.Μ. X με σύνολο τιμών το $S = \{0, 1, 2, \dots\}$, με μέση τιμή $\mu = E(X)$ και διασπορά $\sigma^2 = \text{Var}(X)$. Δείξτε πως, αν η τυπική απόκλιση σ της X είναι σημαντικά μικρότερη από τη μέση τιμή μ , τότε η πιθανότητα το X να ισούται με μηδέν είναι μικρή.

Συγκεκριμένα, αποδείξτε πως,

$$\Pr(X = 0) \leq \left(\frac{\sigma}{\mu}\right)^2.$$

17. *Εναλλακτική έκφραση της $E(X)$.* Έστω μια διακριτή τυχαία μεταβλητή X με σύνολο τιμών το $S = \{0, 1, 2, \dots\}$ και (πεπερασμένη) μέση τιμή $E(X)$. Αποδείξτε πως η $E(X)$ μπορεί εναλλακτικά να εκφραστεί ως,

$$E(X) = \sum_{k=1}^{\infty} \Pr(X \geq k).$$

18. * *Αποστάσεις μεταξύ πυκνοτήτων.* Έστω δύο πυκνότητες $P(x)$ και $Q(x)$ στο ίδιο πεπερασμένο σύνολο S . Η ℓ_1 -απόσταση μεταξύ της P και της Q ορίζεται ως,

$$d_{\ell_1}(P, Q) = \sum_{x \in S} |P(x) - Q(x)|,$$

και η χ^2 -απόσταση της Q από την P ορίζεται ως,

$$d_{\chi^2}(P, Q) = \sum_{x \in S} \frac{(P(x) - Q(x))^2}{Q(x)}.$$

- (α') Παρατηρήστε ότι και για τις δύο αυτές αποστάσεις πάντοτε έχουμε $d_{\ell_1}(P, Q) \geq 0$ και $d_{\chi^2}(P, Q) \geq 0$. Επιπλέον, παρατηρήστε ότι $d_{\ell_1}(P, Q) = 0$ αν και μόνο αν $P \equiv Q$, δηλαδή αν $P(x) = Q(x)$, για κάθε $x \in S$, και ότι το ίδιο ισχύει και για την $d_{\chi^2}(P, Q)$.
- (β') Έστω μια Τ.Μ. X με πυκνότητα Q στο S , και έστω η συνάρτηση $f(x) = \frac{|P(x) - Q(x)|}{Q(x)}$. Εξετάζοντας τη διασπορά $\text{Var}(f(X))$ της Τ.Μ. $f(X)$, αποδείξτε ότι η χ^2 -απόσταση είναι πιο «ισχυρή» από την ℓ_1 -απόσταση, υπό την έννοια ότι, για οποιοδήποτε δύο πυκνότητες P και Q , έχουμε,

$$d_{\ell_1}(P, Q) \leq \sqrt{d_{\chi^2}(P, Q)}.$$

Κεφάλαιο 7

Διακριτές κατανομές

Στο προηγούμενο κεφάλαιο είδαμε πως η έννοια της *τυχαίας μεταβλητής* (T.M.), δηλαδή μιας τυχαίας ποσότητας X που προσδιορίζεται από το σύνολο τιμών της S και την πυκνότητά της P , μας επιτρέπει να περιγράψουμε προβλήματα που μας ενδιαφέρουν με πιο σύντομο και σαφή τρόπο, χωρίς να αναγκάζομαστε κάθε φορά να ορίζουμε σχολαστικά τον πλήρη χώρο πιθανότητας Ω και το αντίστοιχο μέτρο πιθανότητας \mathbb{P} . Το επόμενο βήμα σε αυτήν τη διαδικασία της πιο αφαιρετικής περιγραφής, είναι η καταγραφή κάποιων σημαντικών τύπων τυχαίων μεταβλητών, ιδιαίτερα χρήσιμων στην πράξη, οι οποίοι εμφανίζονται συχνά σε βασικά προβλήματα των πιθανοτήτων.

Σ' αυτό το κεφάλαιο θα εντοπίσουμε πέντε τέτοιες κατηγορίες τυχαίων μεταβλητών και θα δείξουμε με παραδείγματα κάποιες από τις συνήθεις περιπτώσεις όπου χρησιμοποιούνται. Επιπλέον θα αποδείξουμε ορισμένες ιδιότητές τους (π.χ., θα υπολογίσουμε τη μέση τιμή και τη διασπορά τους), έτσι ώστε να μπορούμε να τις χρησιμοποιούμε χωρίς να απαιτείται η επανάληψη των ίδιων υπολογισμών σε κάθε επιμέρους πρόβλημα.

7.1 Κατανομές Bernoulli, διωνυμική και γεωμετρική

Η πιο απλή τυχαία μεταβλητή είναι εκείνη που παίρνει μόνο δύο τιμές, οι οποίες, στην απλούστερη περίπτωση είναι το «0» και το «1»:

Ορισμός 7.1 *Μια διακριτή T.M. X λέμε πως έχει κατανομή Bernoulli με παράμετρο p , για κάποιο $p \in (0, 1)$, αν έχει σύνολο τιμών το $S = \{0, 1\}$ και πυκνότητα P με τιμές $P(1) = p$ και $P(0) = 1 - p$. Για συντομία, αυτό συμβολίζεται: $X \sim \text{Bern}(p)$.*

Παρατηρήσεις

1. Ο πιο πάνω ορισμός απλά λέει πως μια T.M. X έχει κατανομή Bernoulli όταν παίρνει μόνο τις τιμές 0 και 1, και η παράμετρος p της κατανομής είναι η πιθανότητα το X να ισούται με 1, δηλαδή $p = \text{Pr}(X = 1) = P(1)$.

2. Οι πιο συχνές χρήσεις τυχαίων μεταβλητών με κατανομή Bernoulli είναι είτε για την περιγραφή δυαδικών δεδομένων (bits), είτε ως «δείκτες» που προσδιορίζουν αν κάποιο σημαντικό γεγονός έχει συμβεί ή όχι. Για παράδειγμα, μια Τ.Μ. X μπορεί να παίρνει την τιμή 1 αν κάποιο σύστημα παρουσιάζει σφάλμα, αν ο αλγόριθμός μας εκτελέστηκε κανονικά, αν μια κλήση σε ένα δίκτυο τηλεφωνίας ολοκληρώθηκε, αν κάποιος ασθενής θεραπεύτηκε, αν μια άλλη Τ.Μ. Y πάρει τιμή $Y \geq y$ κ.ο.κ., και στην αντίθετη περίπτωση να παίρνει την τιμή 0. Σε αρκετά παραδείγματα του προηγούμενου κεφαλαίου, π.χ, είδαμε τέτοιες Τ.Μ. να περιγράφουν τα αποτελέσματα διαδοχικών ρίψεων ενός νομίσματος. Στις Ασκήσεις 2 και 3 στο τέλος του κεφαλαίου θα δούμε κάποιες απλές ιδιότητες για τέτοιες δείκτριες, όπως συχνά ονομάζονται, τυχαίες μεταβλητές.
3. Όπως δείξαμε στο Παράδειγμα 6.7 του προηγούμενου κεφαλαίου, μια Τ.Μ. $X \sim \text{Bern}(p)$ έχει μέση τιμή $\mu = E(X) = p$ και διασπορά $\sigma^2 = \text{Var}(X) = p(1 - p)$. Η συνάρτηση κατανομής $F(x)$ της X υπολογίστηκε στο Παράδειγμα 6.4,

$$F(x) = \begin{cases} 0, & \text{αν } x < 0, \\ 1 - p, & \text{αν } 0 \leq x < 1, \\ 1, & \text{αν } x \geq 1. \end{cases}$$

και η γραφική της αναπαράσταση δίνεται στο Σχήμα 6.1.

4. Αν και ο αυστηρός ορισμός που δώσαμε πιο πάνω περιορίζει τις δυνατές τιμές της παραμέτρου p στο διάστημα $(0, 1)$, θα μπορούσαμε να επιτρέψουμε και τις ακραίες τιμές $p = 0$ και $p = 1$. Σε αυτές τις τετριμμένες περιπτώσεις, προφανώς η X είναι απλά μια σταθερά: $X = 0$ με πιθανότητα 1 όταν $p = 0$, και $X = 1$ με πιθανότητα 1 όταν $p = 1$.

Παράδειγμα 7.1 Έστω πως στρίβουμε N φορές ένα νόμισμα με $\text{Pr}(\text{Κορώνα}) = p$, για κάποιο $p \in (0, 1)$. Όπως στο Παράδειγμα 6.4, μπορούμε να περιγράψουμε τα αποτελέσματα των διαδοχικών ρίψεων ορίζοντας N ανεξάρτητες Τ.Μ. X_1, X_2, \dots, X_N , όπου η κάθε $X_i \sim \text{Bern}(p)$ και $X_i = 1$ αν φέραμε Κ τη φορά i , ενώ $X_i = 0$ αν φέραμε Γ.

Όπως στο Παράδειγμα 6.3, παρατηρούμε πως το πλήθος των φορών που φέραμε Κορώνα στις N ρίψεις, έστω Y , μπορεί να εκφραστεί ως:

$$Y = X_1 + X_2 + \dots + X_N = \sum_{i=1}^N X_i.$$

Η Τ.Μ. Y έχει σύνολο τιμών $S_Y = \{0, 1, \dots, N\}$. Για να υπολογίσουμε την πυκνότητά της, εξετάζουμε το χώρο πιθανότητας Ω του πειράματος, ο οποίος αποτελείται από στοιχεία της μορφής,

$$\omega = \text{ΓΚΚ} \dots \text{Κ},$$

που περιγράφουν τα αποτελέσματα των N ρίψεων. Π.χ., το πιο πάνω ω περιγράφει το αποτέλεσμα όπου φέραμε Γ την πρώτη φορά και K τις επόμενες $N - 1$ φορές.

Εφόσον το νόμισμα δεν είναι απαραίτητα δίκαιο (δηλαδή δεν μας δίνεται ότι $p = 1/2$), δεν μπορούμε να θεωρήσουμε ότι όλα τα στοιχειώδη ενδεχόμενα είναι ισοπίθανα. Για τον υπολογισμό των τιμών της πυκνότητας $P(k) = \Pr(Y = k)$ της Y θα χρησιμοποιήσουμε το γεγονός ότι οι διαδοχικές ρίψεις (δηλαδή οι Τ.Μ. X_i) είναι ανεξάρτητες. Όπως έχουμε ήδη παρατηρήσει, για να ορίσουμε ένα μέτρο πιθανότητας \mathbb{P} σε έναν πεπερασμένο χώρο πιθανότητας Ω , αρκεί να ορίσουμε το $\mathbb{P}(\{\omega\})$ για κάθε στοιχειώδες ενδεχόμενο $\{\omega\}$. Για παράδειγμα, για το στοιχειώδες ενδεχόμενο που αντιστοιχεί στο $\omega = \Gamma K K \cdots K$ όπως πιο πάνω, έχουμε,

$$\begin{aligned} \mathbb{P}(\{\omega\}) &= \Pr(\{\Gamma K K \cdots K\}) \\ &= \Pr(X_1 = 0, X_2 = 1, X_3 = 1, \dots, X_N = 1) \\ &= \Pr(X_1 = 0) \Pr(X_2 = 1) \Pr(X_3 = 1) \cdots \Pr(X_N = 1) \\ &= p(1-p)(1-p) \cdots (1-p) \\ &= p(1-p)^{N-1}, \end{aligned}$$

όπου στο τρίτο βήμα χρησιμοποιήσαμε την ανεξαρτησία των X_i . Παρομοίως, έστω A_k το υποσύνολο του Ω που αποτελείται από όλα τα $\omega \in \Omega$ που περιέχουν k φορές K και $N - k$ φορές Γ . Τότε έχουμε, με το ίδιο σκεπτικό όπως πριν,

$$\mathbb{P}(\{\omega\}) = p^k (1-p)^{N-k}, \quad \text{για κάθε } \omega \in A_k,$$

και αυτό ισχύει για κάθε δυνατή τιμή του $k = 0, 1, \dots, N$.

Έχοντας τώρα ορίσει το μέτρο πιθανότητας για όλα τα στοιχειώδη ενδεχόμενα, υπολογίζουμε την πυκνότητα της Y ως εξής. Για $k = 0, 1, \dots, N$,

$$\begin{aligned} P(k) &= \Pr(Y = k) \\ &= \Pr(\{\text{όλα τα } \omega \text{ με } k \text{ φορές } K \text{ και } (N - k) \text{ φορές } \Gamma\}) \\ &= \Pr(A_k) \\ &= \Pr\left(\bigcup_{\omega \in A_k} \{\omega\}\right), \end{aligned}$$

και αφού τα στοιχειώδη ενδεχόμενα είναι πάντα ξένα μεταξύ τους, και $\mathbb{P}(\{\omega\}) = p^k (1-p)^{N-k}$ για κάθε $\omega \in A_k$,

$$P(k) = \sum_{\omega \in A_k} \Pr(\{\omega\}) = (\#A_k) \times (p^k (1-p)^{N-k}).$$

Τέλος, παρατηρούμε ότι, από τον κανόνα αρίθμησης #4, υπάρχουν $\binom{N}{k}$ τρόποι να διατάξουμε k Κορώνες και $N - k$ Γράμματα, άρα $\#A_k = \binom{N}{k}$ και συνεπώς:

$$P(k) = \Pr(Y = k) = \binom{N}{k} p^k (1-p)^{N-k}, \quad \text{για κάθε } k \in S_Y = \{0, 1, \dots, N\}.$$

Ορισμός 7.2 Μια διακριτή Τ.Μ. Y λέμε πως έχει διωνυμική κατανομή με παραμέτρους N και p , για κάποια $N \geq 1$ και $p \in (0, 1)$, αν έχει σύνολο τιμών το $S = \{0, 1, \dots, N\}$ και πυκνότητα:

$$P(k) = \Pr(Y = k) = \binom{N}{k} p^k (1-p)^{N-k}, \quad \text{για κάθε } k \in S = \{0, 1, \dots, N\}.$$

Για συντομία, αυτό συμβολίζεται: $Y \sim \Delta\omega\upsilon\upsilon(N, p)$.

Παρατηρήσεις

1. Μια Τ.Μ. Y με διωνυμική κατανομή με παραμέτρους N και p , περιγράφει το πλήθος των «επιτυχιών» σε N διαδοχικά, ανεξάρτητα, όμοια πειράματα, όπου σε κάθε πείραμα η πιθανότητα επιτυχίας ισούται με p . Εδώ, βέβαια, «επιτυχία» θεωρείται το γεγονός που μας ενδιαφέρει στο εκάστοτε πρόβλημα – π.χ. το να φέρουμε K με κάποιο νόμισμα, την εμφάνιση σφάλματος σε κάποιο σύστημα, την ορθή εκτέλεση ενός αλγόριθμου κλπ.
2. Όπως είδαμε στο Παράδειγμα 7.1, μια Τ.Μ. $Y \sim \Delta\omega\upsilon\upsilon(N, p)$ μπορεί πάντα να εκφραστεί ως το άθροισμα $Y = X_1 + X_2 + \dots + X_N$, N ανεξάρτητων Τ.Μ. $X_i \sim \text{Bern}(p)$. Συνεπώς, χρησιμοποιώντας την Ιδιότητα 1 του Θεωρήματος 6.1, το Y έχει μέση τιμή,

$$\mu = E(Y) = E\left(\sum_{i=1}^N X_i\right) = \sum_{i=1}^N E(X_i),$$

και, αφού κάθε Bernoulli Τ.Μ. X_i με παράμετρο p έχει μέση τιμή $E(X_i) = p$, έχουμε,

$$\mu = E(Y) = Np.$$

Παρομοίως, εφόσον τα X_i είναι ανεξάρτητα μπορούμε να εφαρμόσουμε την Ιδιότητα 4 του Θεωρήματος 6.1, σε συνδυασμό με το γεγονός ότι κάθε X_i έχει $\text{Var}(X_i) = p(1-p)$, για να υπολογίσουμε τη διασπορά του Y ως,

$$\sigma^2 = \text{Var}(Y) = \text{Var}\left(\sum_{i=1}^N X_i\right) = \sum_{i=1}^N \text{Var}(X_i) = Np(1-p).$$

3. Όταν το $N = 1$, προφανώς η $Y = X_1$ είναι απλά μια Τ.Μ. με κατανομή $\text{Bern}(p)$. Και όπως στην περίπτωση της κατανομής Bernoulli, αν το p πάρει μία από τις δύο ακραίες τιμές $p = 0$ ή 1 , τότε με πιθανότητα 1 έχουμε $Y = 0$ ή $Y = N$, αντίστοιχα.

Μία από τις πιο συνηθισμένες περιπτώσεις χρήσης της διωνυμικής κατανομής είναι σε προβλήματα «επιλογών με επαναποθέτηση» όπως το ακόλουθο.

Παράδειγμα 7.2 Ένα mailbox περιέχει M μηνύματα email, εκ των οποίων τα k έχουν κάποιο ιό. Από τα M μηνύματα επιλέγουμε $M/3$ με επανατοποθέτηση. Έστω X το (τυχαίο) πλήθος των μηνυμάτων από αυτά που επιλέξαμε τα οποία έχουν τον ιό. Εδώ η X εκφράζει το πλήθος των «επιτυχιών» (όπου επιτυχία εδώ θεωρούμε την επιλογή ενός μηνύματος με τον ιό) ανάμεσα σε $N = M/3$ διαδοχικά πειράματα, τα οποία είναι εξ ορισμού ανεξάρτητα, και το καθένα έχει πιθανότητα επιτυχίας ίση με $p = k/M$. Άρα η $X \sim \text{Διων}(\frac{M}{3}, \frac{k}{M})$. Μπορούμε λοιπόν να κάνουμε μερικούς απλούς υπολογισμούς.

Ποια είναι η πιθανότητα να κολλήσουμε τον ιό από τα επιλεγμένα μηνύματα; Είναι η πιθανότητα να έχουμε επιλέξει τουλάχιστον ένα «μολυσμένο» μήνυμα, δηλαδή να έχουμε $X \geq 1$,

$$\begin{aligned} \Pr(X \geq 1) &= 1 - \Pr(X = 0) = 1 - P(0) = 1 - \binom{M/3}{0} \left(\frac{k}{M}\right)^0 \left(1 - \frac{k}{M}\right)^N \\ &= 1 - \left(1 - \frac{k}{M}\right)^{(M/3)}, \end{aligned}$$

όπου χρησιμοποιήσαμε τον τύπο της πυκνότητας μιας διωνυμικής Τ.Μ. Παρομοίως, εφαρμόζοντας τον αντίστοιχο τύπο για τη μέση τιμή έχουμε,

$$\text{«μέσο πλήθος μηνυμάτων με ιό σε αυτά που επιλέξαμε»} = E(X) = \frac{M}{3} \cdot \frac{k}{M} = \frac{k}{3},$$

το οποίο παρατηρούμε πως δεν εξαρτάται από το συνολικό πλήθος M των μηνυμάτων. Τέλος, αν έχουμε $k = 4$ μολυσμένα μηνύματα ανάμεσα σε $M = 120$, τότε, με $N = 120/3 = 40$ και $p = 4/120 = 1/30$, η πιθανότητα να έχουμε επιλέξει το πολύ δύο που να έχουν τον ιό είναι,

$$\begin{aligned} \Pr(X \leq 2) &= P(0) + P(1) + P(2) \\ &= \binom{40}{0} p^0 (1-p)^{40} + \binom{40}{1} p^1 (1-p)^{39} + \binom{40}{2} p^2 (1-p)^{38} \approx 0.852. \end{aligned}$$

Παράδειγμα 7.3 Όπως στο Παράδειγμα 6.4, στρίβουμε ένα νόμισμα με $\Pr(\text{Κορώνα}) = p$ για κάποιο $p \in (0, 1)$, σε συνεχόμενες, ανεξάρτητες ρίψεις, και για κάθε $i = 1, 2, \dots$, ορίζουμε τις ανεξάρτητες Τ.Μ.,

$$X_i = \begin{cases} 1, & \text{αν φέραμε Κ τη φορά } i, \\ 0, & \text{αν φέραμε Γ τη φορά } i, \end{cases}$$

όπου η κάθε $X_i \sim \text{Bern}(p)$. Έστω Z η πρώτη χρονική στιγμή που φέρνουμε Κ, δηλαδή το μικρότερο $i \geq 1$ τέτοιο ώστε $X_i = 1$. Στο Παράδειγμα 6.4 είδαμε πως η Z έχει σύνολο τιμών το $S_Z = \mathbb{N} = \{1, 2, \dots\}$ και πυκνότητα $P(m) = (1-p)^{m-1}p$, για κάθε ακέραιο $m \geq 1$.

Ορισμός 7.3 Μια διακριτή Τ.Μ. Z λέμε πως έχει γεωμετρική κατανομή με παράμετρο p , για κάποιο $p \in (0, 1)$, αν έχει σύνολο τιμών το $S = \mathbb{N} = \{1, 2, \dots\}$ και πυκνότητα:

$$P(k) = \Pr(Z = k) = (1-p)^{k-1}p, \quad \text{για κάθε } k \in S = \{1, 2, \dots\}.$$

Για συντομία, αυτό συμβολίζεται: $Z \sim \text{Γεωμ}(p)$.

Παρατήρηση: Μια Τ.Μ. Z με γεωμετρική κατανομή με παράμετρο p , περιγράφει τη χρονική στιγμή της πρώτης επιτυχίας σε μια σειρά από διαδοχικά, ανεξάρτητα, όμοια πειράματα, όπου σε κάθε πείραμα η πιθανότητα επιτυχίας ισούται με p . Αν τα αποτελέσματα των πειραμάτων περιγράφονται από μια σειρά ανεξάρτητων τυχαίων μεταβλητών $X_i \sim \text{Bern}(p)$ (όπου $X_i = 1$ αν έχουμε επιτυχία τη φορά i), τότε η Z μπορεί να εκφραστεί ως:

$$Z = \min\{i \geq 1 : X_i = 1\}.$$

Παράδειγμα 7.4 Κάθε πακέτο δεδομένων που περνάει από έναν συγκεκριμένο κόμβο κάποιου δικτύου, «χάνεται» με πιθανότητα $p = 0.01\%$, ανεξάρτητα από την τύχη των υπόλοιπων πακέτων. Έστω X το πρώτο πακέτο που χάνεται (ή, πιο σχολαστικά, ο αύξων αριθμός του πρώτου πακέτου που χάνεται), έτσι ώστε $X \sim \text{Γεωμ}(p)$. Για να διευκολυνθούμε στον παρακάτω υπολογισμό, ορίζουμε τις ανεξάρτητες Τ.Μ. W_1, W_2, \dots , όπου κάθε $W_i \sim \text{Bern}(p)$ περιγράφει αν το πακέτο i χάθηκε ($W_i = 1$) ή όχι ($W_i = 0$).

Ποια είναι η πιθανότητα το πρώτο πακέτο που χάθηκε να είναι μετά τα 10,000 πρώτα; Χρησιμοποιώντας την ανεξαρτησία των W_i :

$$\begin{aligned} \Pr(X > 10000) &= \Pr(W_i = 0 \text{ για κάθε } i = 1, 2, \dots, 10000) \\ &= \Pr(W_1 = 0) \Pr(W_2 = 0) \cdots \Pr(W_{10000} = 0) \\ &= (1 - p)^{10000} = (0.9999)^{10000} \approx 0.3679. \end{aligned}$$

Θεώρημα 7.1 (Ιδιότητες της γεωμετρικής κατανομής)

Έστω $X \sim \text{Γεωμ}(p)$. Η X έχει τις εξής ιδιότητες:

1. **Ουρά:** Για κάθε ακέραιο $m \geq 1$:

$$\Pr(X > m) = (1 - p)^m.$$

2. **Συνάρτηση κατανομής:** $F(x) = 0$ για $x < 1$ και:

$$F(x) = 1 - (1 - p)^{\lfloor x \rfloor}, \quad \text{για } x \geq 1.$$

3. **Ιδιότητα έλλειψης μνήμης:** Για κάθε ζευγάρι ακεραίων $m, n \geq 1$:

$$\Pr(X \geq m + n \mid X > n) = \Pr(X \geq m).$$

Δηλαδή, η πιθανότητα $\Pr(X \geq m + n \mid X > n)$ είναι ανεξάρτητη του n , και ίση με εκείνη που αντιστοιχεί στο $n = 0$, δηλαδή, $\Pr(X \geq m)$.

4. **Μέση τιμή:** $E(X) = \frac{1}{p}$.

5. **Διασπορά:** $\text{Var}(X) = \frac{1-p}{p^2}$.

Απόδειξη: Όπως στο Παράδειγμα 7.4, ορίζουμε τις ανεξάρτητες Τ.Μ. $W_i \sim \text{Bern}(p)$ και θέτουμε το X ίσο με το μικρότερο $i \geq 1$ τέτοιο ώστε $W_i = 1$.

Για την πρώτη ιδιότητα, παρατηρούμε πως το X είναι μεγαλύτερο από m αν και μόνο αν τα πρώτα m από τα W_i είναι όλα ίσα με 0, οπότε,

$$\begin{aligned}\Pr(X > m) &= \Pr(W_1 = 0, W_2 = 0, \dots, W_m = 0) \\ &= \Pr(W_1 = 0) \Pr(W_2 = 0) \cdots \Pr(W_m = 0) = (1 - p)^m.\end{aligned}$$

Η συνάρτηση κατανομής, ακριβώς όπως στην Ιδιότητα 2, έχει ήδη υπολογιστεί στο Παράδειγμα 6.4.

Για την Ιδιότητα 3, χρησιμοποιούμε τον ορισμό της δεσμευμένης πιθανότητας σε συνδυασμό με την Ιδιότητα 1:

$$\begin{aligned}\Pr(X \geq m + n | X > n) &= \frac{\Pr(X \geq m + n \text{ και } X > n)}{\Pr(X > n)} \\ &= \frac{\Pr(X \geq m + n)}{\Pr(X > n)} \\ &= \frac{\Pr(X > m + n - 1)}{\Pr(X > n)} \\ &= \frac{(1 - p)^{m+n-1}}{(1 - p)^n} \\ &= (1 - p)^{m-1} \\ &= \Pr(X > m - 1) \\ &= \Pr(X \geq m).\end{aligned}$$

Για τον υπολογισμό της μέσης τιμής θα χρησιμοποιήσουμε τον τύπο για την έκφραση μιας σειράς που σχετίζεται με μια άπειρη γεωμετρική πρόοδο. Λεπτομέρειες δίνονται στην Ενότητα 7.3. Συγκεκριμένα, εφαρμόζοντας τον τύπο (7.8) με $x = 1 - p$, βρίσκουμε:

$$\begin{aligned}E(X) &= \sum_{k=1}^{\infty} kP(k) = \sum_{k=0}^{\infty} k(1 - p)^{k-1}p = \frac{p}{1 - p} \sum_{k=0}^{\infty} k(1 - p)^k \\ &= \frac{p}{1 - p} \cdot \frac{1 - p}{[1 - (1 - p)]^2} = \frac{1}{p}.\end{aligned}$$

Παρομοίως, εφαρμόζοντας τον τύπο (7.9) με $x = 1 - p$, βρίσκουμε,

$$\begin{aligned}E(X^2) &= \sum_{k=1}^{\infty} k^2 P(k) = \sum_{k=0}^{\infty} k^2 (1 - p)^{k-1} p = \frac{p}{1 - p} \sum_{k=0}^{\infty} k^2 (1 - p)^k \\ &= \frac{p}{1 - p} \cdot \frac{(1 - p)(1 - p + 1)}{[1 - (1 - p)]^3} = \frac{2 - p}{p^2},\end{aligned}$$

και, τέλος, υπολογίζουμε,

$$\text{Var}(X) = E(X^2) - [E(X)]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2},$$

η οποία μας δίνει τη διασπορά της X , ολοκληρώνοντας την απόδειξη. ■

Παράδειγμα 7.5 Ένας φάκελος αρχείων σε κάποιο PC περιέχει 100 αρχεία, εκ των οποίων τα 70 είναι αρχεία pdf και τα 30 είναι mp3. Επιλέγουμε 3 στην τύχη με επανατοποθέτηση. Παρατηρούμε πως το πλήθος X των αρχείων pdf που επιλέξαμε έχει κατανομή $X \sim \text{Διων}(3, 70/100)$. Συνεπώς, η πιθανότητα να επιλέξαμε τουλάχιστον 2 αρχεία pdf είναι,

$$\begin{aligned} \Pr(X \geq 2) &= 1 - \Pr(X < 2) = 1 - P(0) - P(1) \\ &= 1 - \binom{3}{0} (0.7)^0 (0.3)^3 - \binom{3}{1} (0.7)^1 (0.3)^2 \\ &\approx 0.784. \end{aligned}$$

Παρομοίως, το μέσο πλήθος αρχείων pdf που επιλέξαμε είναι,

$$E(X) = 3 \times \frac{70}{100} = 2.1.$$

Τέλος, αν συνεχίζαμε να επιλέγουμε αρχεία (πάλι με επανατοποθέτηση) μέχρι την πρώτη φορά που θα είχαμε ένα pdf, τότε το συνολικό πλήθος, έστω Z , των επιλεγμένων αρχείων θα είχε κατανομή $Z \sim \text{Γεωμ}(70/100)$, και βάσει της παραπάνω ιδιότητας έλλειψης μνήμης θα μπορούσαμε να υπολογίσουμε,

$$\begin{aligned} \Pr(\text{θα επιλέξουμε τουλάχιστον 2 ακόμα} \mid \text{επιλέξαμε ήδη 3}) &= \Pr(Z \geq 3 + 2 \mid Z > 3) \\ &= \Pr(Z \geq 2) \\ &= \Pr(Z > 1) \\ &= \left(1 - \frac{70}{100}\right)^1 \\ &= 0.3, \end{aligned}$$

όπου χρησιμοποιήσαμε και την πρώτη ιδιότητα του Θεωρήματος 7.1.

Κλείνουμε αυτή την ενότητα με ένα διάσημο παράδειγμα.

Παράδειγμα 7.6 (Το πρόβλημα γενεθλίων) Ας υποθέσουμε πως βρίσκομαι σε ένα δωμάτιο μαζί με άλλα 29 άτομα. Ποια είναι η πιθανότητα κάποιος άλλος να έχει γενέθλια την ίδια μέρα με εμένα;

Εστω $n = 29$ Τ.Μ. X_i , όπου $X_i = 1$ αν το άτομο i έχει τα ίδια γενέθλια με εμένα, αλλιώς $X_i = 0$, για κάθε $i = 1, 2, \dots, 29$ (θεωρώντας ότι εγώ είμαι ο τριακοστός). Υποθέτουμε πως τα X_i είναι ανεξάρτητες Τ.Μ., προφανώς με κατανομή Bernoulli, και με παράμετρο,

$$p = \Pr(\text{ο άνθρωπος } i \text{ έχει τα ίδια γενέθλια με εμένα}) = \frac{1}{365}.$$

Άρα,

$$\begin{aligned} & \Pr(\text{υπάρχει κάποιος με τα ίδια γενέθλια}) \\ &= 1 - \Pr(\text{κανείς με τα ίδια γενέθλια}) \\ &= 1 - \Pr(X_1 = 0, X_2 = 0, \dots, X_n = 0) \\ &= 1 - \Pr(X_1 = 0) \Pr(X_2 = 0) \cdots \Pr(X_n = 0) \\ &= 1 - (1 - p)^n \\ &= 1 - \left(\frac{364}{365}\right)^{29} \\ &\approx 0.076 = 7.6\%. \end{aligned}$$

Ας εξετάσουμε τώρα μια απλή παραλλαγή αυτού του προβλήματος ρωτώντας ποια είναι η πιθανότητα, ανάμεσα στα 30 άτομα, να υπάρχουν τουλάχιστον δύο που θα έχουν γενέθλια την ίδια μέρα; Αν υποθέσουμε και πάλι πως τα γενέθλια των διαφορετικών ατόμων είναι ανεξάρτητα μεταξύ τους και πως η κάθε ημερομηνία έχει την ίδια πιθανότητα, δηλαδή $1/365$, βρίσκουμε πως,

$$\begin{aligned} & \Pr(\text{υπάρχουν τουλάχιστον 2 με ίδια γενέθλια}) \\ &= 1 - \Pr(\text{όλοι μεταξύ τους έχουν διαφορετικά γενέθλια}) \\ &= 1 - \Pr\left(\left\{\begin{array}{l} \text{ο 2ος διαφορετικά από τον 1ο} \\ \cap \{\text{ο 3ος διαφορετικά από τους δύο πρώτους}\} \\ \cap \cdots \cap \{\text{ο 30ος διαφορετικά από τους 29 πρώτους}\} \end{array}\right\}\right) \\ &= 1 - \frac{364}{365} \cdot \frac{363}{365} \cdots \frac{336}{365} \approx 0.7 = 70\%. \end{aligned}$$

Δηλαδή, με μόλις 30 άτομα σε ένα δωμάτιο, το πιθανότερο είναι να υπάρχουν τουλάχιστον δύο με τα ίδια γενέθλια! (Δοκιμάστε το την επόμενη φορά που θα βρεθείτε με μια αρκετά μεγάλη παρέα, είναι μια καλή ευκαιρία να κερδίσετε ένα γενναίο στοίχημα.)

Παρατήρηση: Το πρώτο ερώτημα του πιο πάνω παραδείγματος θα μπορούσαμε να το σκεφτούμε σαν ένα πρόβλημα επιλογής με επανατοποθέτηση: Βάζουμε τις 365 μέρες του χρόνου σε μια σακούλα, επιλέγουμε μία από τις 365 για τα γενέθλια του πρώτου ατόμου, μετά επιλέγουμε μία ημερομηνία για τον δεύτερο πάλι από όλες τις 365 κ.ο.κ.

Αντίθετα, ο υπολογισμός βάσει του οποίου απαντήθηκε το δεύτερο ερώτημα θυμίζει επιλογή χωρίς επανατοποθέτηση: Επιλέγουμε μία από τις 365 ημέρες του χρόνου για τα γενέθλια του πρώτου ατόμου, κατόπιν επιλέγουμε μία από τις υπόλοιπες 364 για τον δεύτερο, μετά μία από τις 363 που απομένουν για τον τρίτο, κλπ.

Όπως αναφέραμε πιο πάνω, στα προβλήματα επιλογής με επανατοποθέτηση συχνά μας είναι χρήσιμη η διωνυμική κατανομή. Στην επόμενη ενότητα θα δούμε την αντίστοιχη κατανομή – την υπεργεωμετρική – η οποία προκύπτει σε προβλήματα επιλογής χωρίς επανατοποθέτηση.

7.2 Υπεργεωμετρική και Poisson κατανομή

Παράδειγμα 7.7 Ξεκινάμε εξετάζοντας μια γενική μορφή του προβλήματος των τυχαίων επιλογών χωρίς επανατοποθέτηση. Έστω πως,

$$\text{από } N \text{ αντικείμενα όπου } \begin{cases} \text{τα } k \text{ είναι τύπου I,} \\ \text{τα } N - k \text{ είναι τύπου II,} \end{cases}$$

επιλέγουμε τυχαία n , χωρίς επανατοποθέτηση. [Υποθέτουμε ότι $N \geq k \geq n$.] Έστω η Τ.Μ.,

$$Y = \text{«πλήθος αντικειμένων τύπου I ανάμεσα σε αυτά που επιλέξαμε»,}$$

η οποία περιγράφει την ποσότητα που μας ενδιαφέρει εδώ. Η Y έχει, εξ ορισμού, σύνολο τιμών το $S = \{0, 1, \dots, n\}$, και η πυκνότητά της είναι εύκολο να υπολογιστεί:

Για οποιοδήποτε $m = 0, 1, \dots, n$, η $P(m) = \Pr(Y = m)$ είναι η πιθανότητα του ενδεχομένου «επιλέξαμε m αντικείμενα τύπου I και $(n - m)$ τύπου II». Από τους κανόνες αρίθμησης του Κεφαλαίου 4 σε συνδυασμό με τον κανόνα πιθανότητας #5 έχουμε,

$$P(m) = \frac{\binom{k}{m} \binom{N-k}{n-m}}{\binom{N}{n}}.$$

Ορισμός 7.4 Μια διακριτή Τ.Μ. Y λέμε πως έχει υπεργεωμετρική κατανομή με παραμέτρους N, k και n , για ακραίους $n \leq k \leq N$, αν έχει σύνολο τιμών το $S = \{0, 1, \dots, n\}$ και πυκνότητα:

$$P(m) = \Pr(Y = m) = \frac{\binom{k}{m} \binom{N-k}{n-m}}{\binom{N}{n}}, \quad \text{για κάθε } m \in S = \{0, 1, 2, \dots, n\}.$$

Για συντομία, αυτό συμβολίζεται: $Y \sim \text{Υπερ}(N, k, n)$.

Παράδειγμα 7.8 Έστω πως επιλέγουμε τυχαία τρία φύλλα από μια συνηθισμένη τράπουλα. Ποια είναι η πιθανότητα να επιλέξαμε ακριβώς μία φιγούρα;

Αν ορίσουμε την Τ.Μ. X ως το πλήθος από φιγούρες που επιλέξαμε, τότε (αφού η τράπουλα περιέχει 12 φιγούρες) η X έχει κατανομή Υπερ(52, 12, 3) και η ζητούμενη πιθανότητα ισούται με,

$$\Pr(X = 1) = P(1) = \frac{\binom{12}{1} \binom{52-12}{3-1}}{\binom{52}{3}} \approx 0.4235.$$

Παρατήρηση: Αν στο Παράδειγμα 7.7 επιλέγαμε τυχαία n αντικείμενα με επανατοποθέτηση, τότε η Y θα είχε κατανομή Διων($n, k/N$) αλλά, εφόσον εδώ δεν έχουμε επανατοποθέτηση, η Y έχει κατανομή Υπερ(N, k, n).

Ακολουθώντας κάποια από τα ίδια βήματα της μεθοδολογίας με την οποία εξετάσαμε ορισμένες από τις ιδιότητες της διωνυμικής κατανομής, μπορούμε να ορίσουμε κι εδώ n τυχαίες μεταβλητές Bernoulli, οι οποίες να εκφράζουν τα αποτελέσματα των διαδοχικών επιλογών: Έστω, για κάθε $i = 1, 2, \dots, n$:

$$X_i = \begin{cases} 1 & \text{αν στην επιλογή } i \text{ έχουμε αντικ. τύπου I,} \\ 0 & \text{αν στην επιλογή } i \text{ έχουμε αντικ. τύπου II.} \end{cases}$$

Τότε, όπως και στην περίπτωση της διωνυμικής κατανομής, η Y μπορεί να εκφραστεί,

$$Y = \sum_{i=1}^n X_i.$$

Κατ' αρχάς παρατηρούμε πως, σε αντίθεση με την περίπτωση της διωνυμικής κατανομής, οι Τ.Μ. X_i δεν είναι ανεξάρτητες. Για παράδειγμα, εύκολα υπολογίζουμε πως,

$$\begin{aligned} \Pr(X_2 = 1 | X_1 = 0) &= \frac{k}{N-1}, \\ \text{ενώ } \Pr(X_2 = 1 | X_1 = 1) &= \frac{k-1}{N-1}. \end{aligned} \quad (7.1)$$

Από την άλλη μεριά, λόγω της συμμετρίας του προβλήματος, όλες οι Bernoulli Τ.Μ. X_i έχουν την ίδια παράμετρο. Ένας τρόπος για να πεισθούμε γι' αυτό είναι να σκεφτούμε πως, αντί να επιλέξουμε n αντικείμενα χωρίς επανατοποθέτηση το ένα μετά το άλλο, ισοδύναμα πραγματοποιούμε το εξής πείραμα. Διατάσσουμε με τυχαίο τρόπο τα N αντικείμενα και μετά επιλέγουμε τα n πρώτα. Προφανώς, η πιθανότητα σε καθεμία από τις πρώτες n θέσεις να έχουμε αντικείμενο τύπου I είναι k/N , άρα, η παράμετρος p_i της κάθε X_i είναι ίδια:

$$p_i = \Pr(X_i = 1) = \frac{k}{N}, \quad \text{για κάθε } i = 1, 2, \dots, n.$$

Εφαρμόζοντας τώρα την πρώτη ιδιότητα του Θεωρήματος 6.1, βρίσκουμε την εξής έκφραση για τη μέση τιμή μιας Τ.Μ. $Y \sim \text{Υπερ}(N, k, n)$:

$$\mu = E(Y) = E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i) = \frac{nk}{N}. \quad (7.2)$$

Ένας παρόμοιος αλλά λίγο πιο πολύπλοκος υπολογισμός μάς επιτρέπει να υπολογίσουμε και τη διασπορά της Y :

$$\text{Var}(Y) = \frac{nk(N-k)(N-n)}{N^2(N-1)}. \quad (7.3)$$

Λεπτομέρειες για την απόδειξη της (7.3) δίνονται στην Άσκηση 18 στο τέλος του κεφαλαίου.

Παράδειγμα 7.9 Συνεχίζοντας το Παράδειγμα 7.8, μπορούμε να υπολογίσουμε το μέσο πλήθος από φυγούρες μεταξύ των 3 φύλλων που επιλέξαμε από τον τύπο (7.2) ως, $E(Y) = 3 \cdot \frac{12}{52} = \frac{9}{13}$.

Παράδειγμα 7.10 Έστω πως, από τα 10 εκατομμύρια άτομα ενός πληθυσμού, οι 100 χιλιάδες έχουν στο σπίτι τους σύνδεση internet μέσω γραμμής ADSL. Από αυτό τον πληθυσμό επιλέγουμε τυχαία, για μια έρευνα αγοράς, 150 άτομα, χωρίς επανατοποθέτηση. Τότε, το πλήθος, έστω X , των ανθρώπων με σύνδεση ADSL ανάμεσα στους επιλεγμένους έχει κατανομή $X \sim \text{Υπερ}(10^7, 10^5, 150)$.

Αν, αντίθετα, η επιλογή των ατόμων γίνει με επανατοποθέτηση, τότε το πλήθος Y των ανθρώπων με ADSL που θα επιλέξουμε θα έχει κατανομή $Y \sim \text{Διων}(150, 0.01)$, αφού η πιθανότητα να επιλεγεί κάποιος με ADSL είναι $p = 10^5/10^7 = 0.01$. Παρότι η πυκνότητα της Y μας είναι γνωστή, ο υπολογισμός των τιμών της πυκνότητας αυτής είναι αριθμητικά δύσκολος διότι απαιτεί τη χρήση τιμών της μορφής $n!$ για μεγάλα n . Π.χ.,

$$\Pr(Y = 10) = \binom{150}{10} p^{10} (1-p)^{140} = \frac{150!}{10! 140!},$$

όπου, για παράδειγμα, το $140!$ είναι μια ποσότητα της τάξης του 10^{241} .

Προκειμένου να αποφύγουμε τέτοιου είδους υπολογισμούς με αστρονομικά μεγέθη (οι οποίοι συχνά οδηγούν σε μεγάλα αριθμητικά σφάλματα στην πράξη), θα χρησιμοποιήσουμε την πιο κάτω προσέγγιση για το $n!$ όταν το n παίρνει μεγάλες τιμές. Η απόδειξή της δίνεται στην Ενότητα 7.4.

Λήμμα 7.1 (Τύπος του Stirling) Υπάρχει κάποια σταθερά $C > 0$ τέτοια ώστε, για μεγάλες τιμές του n , η τιμή του $n!$ μπορεί να προσεγγιστεί ως,

$$n! \approx C n^n \sqrt{n} e^{-n}.$$

Πιο συγκεκριμένα, έχουμε,

$$\frac{n!}{C n^n \sqrt{n} e^{-n}} \rightarrow 1, \quad \text{καθώς το } n \rightarrow \infty.$$

Σημείωση. Αν και δεν θα μας χρειαστεί, η τιμή της σταθεράς C στο Λήμμα 7.1 μπορεί να υπολογιστεί ακριβώς με κάποιο επιπλέον κόπο, και είναι $C = \sqrt{2\pi}$.

Παράδειγμα 7.11 Έστω πως σε ένα χρονικό διάστημα N δευτερολέπτων, όπου θεωρούμε πως το N είναι «μεγάλο», συμβαίνει ένα τυχαίο πλήθος Y γεγονότων, και κατά μέσο όρο συμβαίνουν $E(Y) = \lambda$ τέτοια γεγονότα. Πιο συγκεκριμένα, θεωρούμε πως για κάθε δευτερόλεπτο είτε συμβαίνει ένα γεγονός είτε όχι, και αυτά είναι ανεξάρτητα μεταξύ τους. Ορίζουμε λοιπόν N ανεξάρτητες Τ.Μ. $X_i \sim \text{Bern}(p)$, όπου η κάθε $X_i = 1$ αν έχουμε γεγονός το δευτερόλεπτο i , αλλιώς $X_i = 0$. Τότε, το συνολικό πλήθος Y ισούται με το άθροισμα των X_i , και συνεπώς $X \sim \text{Διων}(N, p)$.

Εφόσον υποθέτουμε πως $E(Y) = \lambda$ αλλά μια διωνυμική Τ.Μ. έχει $E(Y) = Np$, θα πρέπει να θέσουμε την πιθανότητα $p = \lambda/N$. Όπως παρατηρήσαμε στο τελευταίο παράδειγμα, η διωνυμική κατανομή είναι δύσχερηστη για μεγάλα N . Το ακόλουθο θεώρημα μας λέει πως, σε αυτή την περίπτωση, μπορεί να προσεγγιστεί αποτελεσματικά από μια άλλη, πολύ πιο εύχερηστη, κατανομή.

Θεώρημα 7.2 (Ορισμός κατανομής Poisson) Έστω $\lambda > 0$ μια δεδομένη σταθερά, και έστω για κάθε N μια Τ.Μ. Y_N με κατανομή Διων($N, \lambda/N$) και πυκνότητα $P_N(k) = \text{Pr}(Y_N = k)$, για $k = 0, 1, \dots, N$.

Καθώς το $N \rightarrow \infty$ οι τιμές της πυκνότητας $P_N(k)$ συγκλίνουν στις αντίστοιχες τιμές της πυκνότητας $P(k)$ μιας Τ.Μ. Z με κατανομή Poisson με παράμετρο λ , δηλαδή με σύνολο τιμών το $S = \{0, 1, 2, \dots\}$ και πυκνότητα,

$$P(k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad \text{για } k \in S = \{0, 1, 2, \dots\}.$$

Με άλλα λόγια, για κάθε $k = 0, 1, 2, \dots$, έχουμε,

$$P_N(k) = \binom{N}{k} \left(\frac{\lambda}{N}\right)^k \left(1 - \frac{\lambda}{N}\right)^{N-k} \rightarrow P(k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad (7.4)$$

καθώς το $N \rightarrow \infty$.

Σημείωση. Για συντομία, το ότι μια Τ.Μ. Z έχει κατανομή Poisson με παράμετρο λ συμβολίζεται ως εξής: $Z \sim \text{Poisson}(\lambda)$. Η κατανομή Poisson έχει δύο συνηθισμένες χρήσεις. Η μία είναι για να περιγράψει το πλήθος των γεγονότων που συμβαίνουν σε μια δεδομένη χρονική περίοδο, όταν αυτά συμβαίνουν με τον «πιο τυχαίο δυνατό τρόπο» και το μόνο που είναι γνωστό είναι το μέσο πλήθος τους. Η δεύτερη είναι προσέγγιση της διωνυμικής κατανομής. Οι ακριβείς συνθήκες κάτω από τις οποίες η κατανομή Poisson μάς δίνει πράγματι μια καλή προσέγγιση για τη διωνυμική δίνονται επιγραμματικά στο πιο κάτω πόρισμα, το οποίο είναι άμεση απόρροια του Θεωρήματος 7.2.

Πόρισμα 7.1 (Poisson προσέγγιση στη διωνυμική) Έστω πως μια τυχαία μεταβλητή Y έχει κατανομή Διων(N, p), με παραμέτρους N, p που να ικανοποιούν τις εξής συνθήκες:

- το N είναι αρκετά «μεγάλο», δηλαδή $N \geq 100$,
- το p είναι αρκετά «μικρό», δηλαδή $p < 1/25$,
- το γινόμενο Np είναι της τάξεως του 1.

Τότε η κατανομή της Y μπορεί να προσεγγιστεί από την κατανομή μιας $T.M.$ $Z \sim \text{Poisson}(\lambda)$ με $\lambda = Np$ υπό την έννοια ότι:

$$\Pr(Y = k) \approx \Pr(Z = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad \text{για κάθε } k = 0, 1, 2, \dots$$

Για την απόδειξη του Θεωρήματος 7.2 θα χρησιμοποιήσουμε τις πιο κάτω εκφράσεις για την εκθετική συνάρτηση $f(x) = e^x$. Η πρώτη μπορεί να ληφθεί ως ο ορισμός της e^x και τη θεωρούμε γνωστή. Λεπτομέρειες για την απόδειξη της τρίτης έκφρασης, η οποία είναι μια ισχυρότερη μορφή της δεύτερης, δίνονται στην Άσκηση 19 στο τέλος του κεφαλαίου.

Λήμμα 7.2 (Εκθετική συνάρτηση) Για κάθε πραγματική τιμή x , η εκθετική συνάρτηση e^x μπορεί να εκφραστεί ως:

1. Η σειρά:

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

2. Το όριο:

$$e^x = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n.$$

Επιπλέον, για μια οποιαδήποτε ακολουθία πραγματικών αριθμών $\{x_n\}$ που τείνουν στο $x \in \mathbb{R}$ καθώς το $n \rightarrow \infty$, έχουμε:

$$e^x = \lim_{n \rightarrow \infty} \left(1 + \frac{x_n}{n}\right)^n. \quad (7.5)$$

* **Απόδειξη του Θεωρήματος 7.2:** Έστω $\lambda > 0$ και $k \geq 0$ δεδομένα. Η ιδέα της απόδειξης είναι απλή: Θα εφαρμόσουμε τον τύπο του Stirling στην πυκνότητα $P_N(k)$ της διωνυμικής κατανομής. Θυμίζουμε ότι, όπως στο Λήμμα 7.1, η ακριβής διατύπωση του τύπου του Stirling μάς λέει πως υπάρχει μια σταθερά C και μια ακολουθία πραγματικών αριθμών $\{c_n\}$ που τείνει στο ένα, τέτοια ώστε, $n! = c_n C n^n \sqrt{n} e^{-n}$, για κάθε n . Αντικαθιστώντας αυτή την έκφραση στα

παραγοντικά $N!$ και $(N - k)!$ που εμφανίζονται στην $P_N(k)$,

$$\begin{aligned} P_N(k) &= \binom{N}{k} \left(\frac{\lambda}{N}\right)^k \left(1 - \frac{\lambda}{N}\right)^{N-k} \\ &= \frac{N!}{(N-k)!k!} \left(\frac{\lambda}{N}\right)^k \left(\frac{N-\lambda}{N}\right)^{N-k} \\ &= \frac{c_N C N^N \sqrt{N} e^{-N}}{c_{N-k} C (N-k)^{N-k} \sqrt{N-k} e^{-(N-k)}} \frac{1}{k!} \left(\frac{\lambda}{N}\right)^k \left(\frac{N-\lambda}{N}\right)^N \left(\frac{N}{N-\lambda}\right)^k. \end{aligned} \quad (7.6)$$

Θέτοντας $d_N = c_N/c_{N-k}$, παρατηρώντας ότι $d_N \rightarrow 1$ καθώς $N \rightarrow \infty$, και απλοποιώντας, βρίσκουμε,

$$\begin{aligned} P_N(k) &= \frac{1}{k!} e^{-k} \left(\frac{N}{N-k}\right)^N \sqrt{\frac{N}{N-k}} (N-k)^k \left(\frac{\lambda}{N}\right)^k \left(\frac{N-\lambda}{N}\right)^N \left(\frac{N}{N-\lambda}\right)^k d_N \\ &= \frac{1}{k!} e^{-k} \left(\frac{N-\lambda}{N-k}\right)^N \sqrt{\frac{1}{1-k/N}} \left(\frac{\lambda(N-k)}{N-\lambda}\right)^k d_N \\ &= \frac{1}{k!} e^{-k} \left[1 - \frac{1}{N} \left(\frac{\lambda-k}{1-k/N}\right)\right]^N \sqrt{\frac{1}{1-k/N}} \left[\lambda \cdot \frac{1-k/N}{1-\lambda/N}\right]^k d_N. \end{aligned}$$

Το τελευταίο παραπάνω γινόμενο αποτελείται από έξι όρους. Καθώς το $N \rightarrow \infty$, οι δύο πρώτοι είναι σταθεροί, και ο τέταρτος και ο έκτος τείνουν στο 1. Παρομοίως, ο πέμπτος όρος τείνει στο λ^k . Για τον βασικότερο όρο, τον τρίτο, χρησιμοποιώντας την τρίτη έκφραση (7.5) από το Λήμμα 7.2, με $x_N = \frac{\lambda-k}{1-k/N}$, και παρατηρώντας πως τα x_N τείνουν στο $x = \lambda - k$ καθώς το $N \rightarrow \infty$, έχουμε,

$$\left[1 - \frac{1}{N} \left(\frac{\lambda-k}{1-k/N}\right)\right]^N \rightarrow e^{k-\lambda}.$$

Τέλος, συνδυάζοντας όλα τα παραπάνω, καθώς το $N \rightarrow \infty$,

$$P_N(k) \rightarrow \frac{1}{k!} e^{-k} e^{k-\lambda} \lambda^k = e^{-\lambda} \frac{\lambda^k}{k!},$$

και έχουμε αποδείξει τη ζητούμενη σχέση (7.4). ■

Παρατήρηση: Το μόνο σημείο στην παραπάνω απόδειξη όπου εφαρμόστηκε ο τύπος του Stirling ήταν για να δεχθεί, στην έκφραση (7.6), ότι ο όρος $N!/((N-k)!N^k)$ τείνει στο 1 καθώς το $N \rightarrow \infty$. Αλλά αυτό μπορεί εύκολα να αποδειχθεί και απευθείας, παρατηρώντας ότι,

$$\frac{N!}{(N-k)!N^k} = \frac{N(N-1)\cdots(N-k+1)}{N^k},$$

το οποίο είναι ο λόγος δύο πολυωνύμων (ως προς N) βαθμού k , με πρώτο όρο N^k και στα δύο, και συνεπώς τείνει στο 1 καθώς το $N \rightarrow \infty$.

Παράδειγμα 7.12 Επανερχόμαστε στο Παράδειγμα 7.10, όπου το ζητούμενο ήταν να υπολογίσουμε την πιθανότητα $\Pr(Y = 10)$ για μια Τ.Μ. $Y \sim \text{Διων}(150, 0.01)$. Εδώ το γινόμενο $150 \times 0.01 = 1.5$, και προφανώς ικανοποιούνται και οι τρεις συνθήκες του Πορίσματος 7.1. Συνεπώς μπορούμε να προσεγγίσουμε την κατανομή της Y μέσω της κατανομής Poisson(λ) με $\lambda = 1.5$, έτσι ώστε:

$$\Pr(Y = 10) \approx e^{-1.5} \frac{(1.5)^{10}}{10!} \approx 3.5 \times 10^{-6}.$$

Στο επόμενο μας αποτέλεσμα υπολογίζονται η μέση τιμή και η διασπορά μιας Τ.Μ. με κατανομή Poisson.

Θεώρημα 7.3 Αν $Z \sim \text{Poisson}(\lambda)$, τότε:

$$E(Z) = \lambda \quad \text{και} \quad \text{Var}(Z) = \lambda.$$

Απόδειξη: Για τη μέση τιμή έχουμε,

$$\mu = E(Z) = \sum_{k=0}^{\infty} k P(k) = \sum_{k=0}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!}.$$

Το άθροισμα της πιο πάνω σειράς υπολογίζεται στην επόμενη ενότητα, και δίνεται από τον τύπο (7.10). Εφαρμόζοντάς τον, με $x = \lambda$, βρίσκουμε,

$$\mu = E(Z) = e^{-\lambda} \lambda e^{\lambda} = \lambda.$$

Παρομοίως υπολογίζεται και η μέση τιμή της Z^2 ως,

$$E(Z^2) = \sum_{k=0}^{\infty} k^2 e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!},$$

και, εφαρμόζοντας τον αντίστοιχο τύπο (7.11) με $x = \lambda$, προκύπτει πως,

$$E(Z^2) = e^{-\lambda} \lambda(1 + \lambda) e^{\lambda} = \lambda(1 + \lambda).$$

Συνεπώς, η διασπορά της Z είναι ίση με,

$$\sigma^2 = \text{Var}(Z) = E(Z^2) - [E(Z)]^2 = \lambda(1 + \lambda) - \lambda^2 = \lambda,$$

ολοκληρώνοντας την απόδειξη. ■

Κλείνουμε αυτή την ενότητα με ένα ακόμα παράδειγμα.

Παράδειγμα 7.13 Έστω πως σε καθένα από 1000 αντικείμενα δίνουμε τυχαία έναν αριθμό από το 1 ως το 1000 (επιτρέποντας σε δύο αντικείμενα να έχουν τον ίδιο αριθμό), και μας ενδιαφέρει το (τυχαίο) πλήθος X των αντικειμένων που έχουν τον αριθμό 1. Τότε, το X έχει κατανομή Διων(N, p) με παραμέτρους $N = 1000$ και $p = 1/1000 = 0.001$. Εφόσον το γινόμενο $N \cdot p = 1$, βάσει του Πορίσματος 7.1, η κατανομή του X μπορεί να προσεγγιστεί από την Poisson(λ), με $\lambda = N \cdot p = 1$.

Συνεπώς, μπορούμε, π.χ., να υπολογίσουμε (ή, πιο σωστά, να προσεγγίσουμε) την πιθανότητα να έχουν 2 ή 3 αντικείμενα τον αριθμό 1 ως,

$$\Pr(X = 2 \text{ ή } X = 3) = \Pr(X = 2) + \Pr(X = 3) \approx e^{-1} \frac{1^2}{2!} + e^{-1} \frac{1^3}{3!} = \frac{2}{3} e^{-1} \approx 0.245.$$

7.3 * Η γεωμετρική και συναφείς σειρές

Εδώ συγκεντρώνουμε μερικά απλά αποτελέσματα που συχνά φαίνονται χρήσιμα σε προβλήματα που σχετίζονται με Τ.Μ. με γεωμετρική κατανομή.

Ξεκινάμε από την απλή παρατήρηση ότι, για κάθε $x \in \mathbb{R}$ και κάθε $n \geq 1$,

$$(1 + x + x^2 + \dots + x^n)(1 - x) = 1 - x^{n+1}.$$

[Αν αυτή η σχέση δεν σας είναι προφανής, αποδείξτε την με επαγωγή.] Υποθέτοντας ότι $x \neq 1$ και διαιρώντας και τα δύο μέρη με το $(1 - x)$, έχουμε τον γνωστό τύπο για το άθροισμα των όρων μιας γεωμετρικής προόδου:

$$\sum_{k=0}^n x^k = \frac{1 - x^{n+1}}{1 - x}, \quad \text{για κάθε } x \neq 1.$$

Αν περιοριστούμε τώρα σε τιμές του x με $|x| < 1$, τότε προφανώς το $x^{n+1} \rightarrow 0$ καθώς το $n \rightarrow \infty$, άρα, περνώντας στο όριο, προκύπτει ο τύπος του αθροίσματος μιας γεωμετρικής σειράς,

$$g(x) = \sum_{k=0}^{\infty} x^k = \frac{1}{1 - x}, \quad \text{για κάθε } |x| < 1. \quad (7.7)$$

Από τον ορισμό της συνάρτησης $g(x)$ στον τύπο (7.7) έχουμε δύο διαφορετικές εκφράσεις για την $g(x)$, μία ως σειρά και μία ως το $1/(1 - x)$. Μπορούμε επομένως να υπολογίσουμε την παράγωγο $g'(x)$ με δύο τρόπους: Αφού η παράγωγος ενός αθροίσματος ισούται με το άθροισμα των παραγώγων,

$$g'(x) = \frac{d}{dx} \left(\sum_{k=0}^{\infty} x^k \right) = \sum_{k=0}^{\infty} \frac{d}{dx} (x^k) = \sum_{k=0}^{\infty} kx^{k-1},$$

αλλά και,

$$g'(x) = \frac{d}{dx} \left(\frac{1}{1 - x} \right) = \frac{1}{(1 - x)^2}.$$

Εξισώνοντας τις δύο παραπάνω εκφράσεις και πολλαπλασιάζοντας και τα δύο μέρη με το x , προκύπτει το αποτέλεσμα:

$$h(x) = \sum_{k=0}^{\infty} kx^k = \frac{x}{(1-x)^2}, \quad \text{για κάθε } |x| < 1. \quad (7.8)$$

Επαναλαμβάνοντας την ίδια διαδικασία, μπορούμε να υπολογίσουμε την παράγωγο της νέας συνάρτησης $h(x)$ με δύο τρόπους ως,

$$\begin{aligned} h'(x) &= \frac{d}{dx} \left(\sum_{k=0}^{\infty} kx^k \right) = \sum_{k=0}^{\infty} \frac{d}{dx} (kx^k) = \sum_{k=0}^{\infty} k^2 x^{k-1} \\ \text{ή } h'(x) &= \frac{d}{dx} \left(\frac{x}{(1-x)^2} \right) = \frac{(1-x)^2 + 2x(1-x)}{(1-x)^4} = \frac{1+x}{(1-x)^3}. \end{aligned}$$

Εξισώνοντας όπως πριν τις δύο εκφράσεις για την παράγωγο και πολλαπλασιάζοντας και τα δύο μέρη με το x , προκύπτει:

$$\sum_{k=0}^{\infty} k^2 x^k = \frac{x(1+x)}{(1-x)^3}, \quad \text{για κάθε } |x| < 1. \quad (7.9)$$

7.4 * Η συνάρτηση e^x , δυναμοσειρές, τύπος του Stirling

Σε αυτή την ενότητα θα υπολογίσουμε τις τιμές δύο σειρών που σχετίζονται με την κατανομή Poisson, και θα αποδείξουμε τον τύπο του Stirling όπως περιγράφεται στο Λήμμα 7.1.

Ξεκινάμε θυμίζοντας, όπως στο Λήμμα 7.2, ότι, για κάθε πραγματικό αριθμό x , η εκθετική συνάρτηση μπορεί να εκφραστεί μέσω του αναπτύγματος του Taylor ως μια δυναμοσειρά:

$$f(x) = e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}, \quad \text{για κάθε } x \in \mathbb{R}.$$

Εφόσον για την εκθετική συνάρτηση ισχύει πως $f'(x) = f(x)$, παίρνοντας την παράγωγο ως προς x και στα δύο μέρη της πιο πάνω σχέσης έχουμε,

$$e^x = \sum_{k=0}^{\infty} k \frac{x^{k-1}}{k!} = \frac{1}{x} \sum_{k=0}^{\infty} k \frac{x^k}{k!},$$

όπου χρησιμοποιήσαμε τη γνωστή ιδιότητα πως η παράγωγος ενός αθροίσματος ισούται με το άθροισμα των παραγώγων. Άρα έχουμε αποδείξει πως, για κάθε $x \in \mathbb{R}$,

$$xe^x = \sum_{k=0}^{\infty} k \frac{x^k}{k!}. \quad (7.10)$$

Επαναλαμβάνοντας την ίδια διαδικασία ακόμη μία φορά, παίρνοντας παραγώγους έχουμε,

$$\frac{d}{dx}(xe^x) = (x+1)e^x = \frac{d}{dx} \left(\sum_{k=0}^{\infty} k \frac{x^k}{k!} \right) = \sum_{k=0}^{\infty} k \frac{d}{dx} \left(\frac{x^k}{k!} \right),$$

και συνεπώς,

$$(x+1)e^x = \sum_{k=0}^{\infty} k^2 \frac{x^{k-1}}{k!} = \frac{1}{x} \sum_{k=0}^{\infty} k^2 \frac{x^k}{k!},$$

ή, απλοποιώντας,

$$x(x+1)e^x = \sum_{k=0}^{\infty} k^2 \frac{x^k}{k!}. \quad (7.11)$$

* **Απόδειξη του Λήμματος 7.1:** Κατ' αρχάς παρατηρούμε πως, παίρνοντας λογαρίθμους, αρκεί να αποδείξουμε πως, για κάποια σταθερά $C > 0$,

$$\log n! - \left(n + \frac{1}{2} \right) \log n + n \rightarrow \log C, \quad \text{καθώς το } n \rightarrow \infty, \quad (7.12)$$

όπου με \log συμβολίζουμε τον φυσικό λογάριθμο με βάση το e (ο οποίος μερικές φορές συμβολίζεται και ως \ln).

Για κάθε $t > 0$ ορίζουμε το ολοκλήρωμα,

$$I(t) = \int_0^t \log t \, dt = t \log t - t,$$

και επίσης (για $k \geq 1$) ορίζουμε τις ακολουθίες,

$$\begin{aligned} a_k &= \frac{1}{2} \log k - \int_{k-1/2}^k \log x \, dx = \int_{k-1/2}^k \log \frac{k}{x} \, dx, \\ b_k &= \int_k^{k+1/2} \log x \, dx - \frac{1}{2} \log k = \int_k^{k+1/2} \log \frac{x}{k} \, dx. \end{aligned}$$

Από τον ορισμό των δύο ακολουθιών, η διαφορά τους ισούται με,

$$a_k - b_k = \int_{k-1/2}^{k+1/2} \log \frac{k}{x} \, dx = - \int_{k-1/2}^{k+1/2} \log x \, dx + \log k,$$

οπότε, αθροίζοντας τις διαφορές,

$$\sum_{k=1}^{n-1} (a_k - b_k) = \log((n-1)!) - \int_{1/2}^{n-1/2} \log x \, dx,$$

έχουμε,

$$\begin{aligned} \sum_{k=1}^{n-1} (a_k - b_k) + a_n &= \log((n-1)!) + \frac{1}{2} \log n - \int_{1/2}^{n-1/2} \log x \, dx - \int_{n-1/2}^n \log x \, dx \\ &= \log n! - \frac{1}{2} \log n - \int_{1/2}^n \log x \, dx. \end{aligned}$$

Τώρα προσθέτοντας και αφαιρώντας το ολοκλήρωμα $I(1/2)$, βρίσκουμε,

$$\begin{aligned} \sum_{k=1}^{n-1} (a_k - b_k) + a_n &= \log n! - \frac{1}{2} \log n - \int_0^n \log x \, dx + I(1/2) \\ &= \log n! - \left(n + \frac{1}{2}\right) \log n + n + I(1/2). \end{aligned}$$

Συνεπώς, για να αποδείξουμε τη ζητούμενη σχέση (7.12), αρκεί να δείξουμε πως το αριστερό μέρος της παραπάνω σχέσης, δηλαδή το $\sum_{k=1}^{n-1} (a_k - b_k) + a_n$, τείνει σε κάποιο (πεπερασμένο) όριο καθώς το $n \rightarrow \infty$.

Παρατηρούμε πως, μετά από μια απλή αλλαγή μεταβλητών, οι ακολουθίες $\{a_k\}$ και $\{b_k\}$ μπορούν να εκφραστούν ως,

$$\begin{aligned} a_k &= \int_0^{1/2} \log\left(\frac{1}{1-t/k}\right) dt, \\ b_k &= \int_0^{1/2} \log\left(1 + \frac{t}{k}\right) dt, \end{aligned}$$

απ' όπου εύκολα προκύπτει πως $a_n \rightarrow 0$ καθώς $n \rightarrow \infty$, και πως η διαφορά τους,

$$a_k - b_k = - \int_0^{1/2} \log\left(1 - \frac{t^2}{k^2}\right) dt = \int_0^{1/2} \log\left(1 + \frac{t^2}{k^2 - t^2}\right) dt \geq 0.$$

Άρα, για να ολοκληρώσουμε την απόδειξη, αρκεί να δείξουμε πως η σειρά $\sum_{k=1}^{\infty} (a_k - b_k)$ (των θετικών όρων $a_k - b_k$) συγκλίνει σε κάποιο πεπερασμένο άθροισμα.

Από την απλή ανισότητα (βλ. Λήμμα 7.3 πιο κάτω) $\log(1+x) \leq x$ έχουμε,

$$\log\left(1 + \frac{t^2}{k^2 - t^2}\right) \leq \frac{t^2}{k^2 - t^2},$$

και εφόσον η συνάρτηση $g(t) = \frac{t^2}{k^2 - t^2}$ είναι αύξουσα ως προς t (γιατί;), για τις τιμές του t στο σχετικό ολοκλήρωμα έχουμε $g(t) \leq g(1)$ και,

$$\log\left(1 + \frac{t^2}{k^2 - t^2}\right) \leq \frac{1}{k^2 - 1} \leq \frac{2}{k^2},$$

όπου η τελευταία στοιχειώδης ανισότητα ισχύει για κάθε $k \geq 2$. Τελικά, λοιπόν, βρίσκουμε πως,

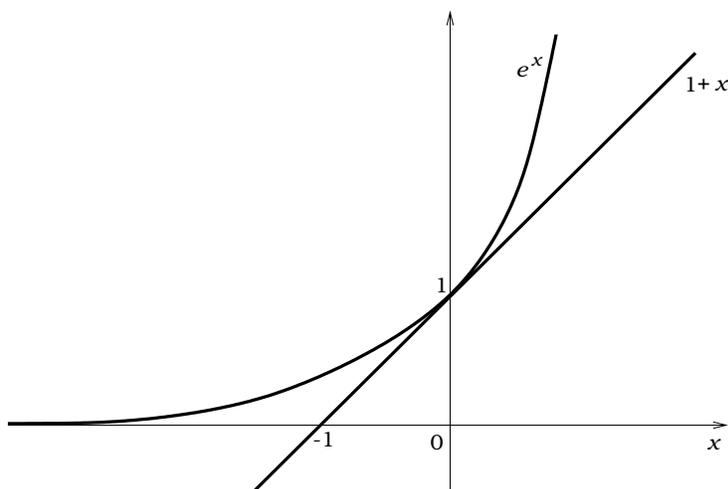
$$a_k - b_k = \int_0^{1/2} \log\left(\frac{t^2}{k^2 - t^2}\right) dt \leq \frac{1}{k^2},$$

και αφού η σειρά $\sum_{k=1}^{\infty} (1/k^2) < \infty$, τότε συγκλίνει και η ζητούμενη σειρά, ολοκληρώνοντας την απόδειξη. ■

Κλείνουμε αυτό το κεφάλαιο με μια στοιχειώδη ανισότητα την οποία χρησιμοποιήσαμε ήδη πιο πάνω και η οποία θα μας φανεί χρήσιμη και αργότερα.

Λήμμα 7.3 Για κάθε πραγματικό αριθμό $x > -1$ έχουμε $\log(1+x) \leq x$ ή, ισοδύναμα:

$$e^x \geq 1+x, \quad \text{για κάθε } x \in \mathbb{R}. \quad (7.13)$$



Σχήμα 7.1: Σχηματική αναπαράσταση της ανισότητας στο Λήμμα 7.3.

Απόδειξη: Έστω η συνάρτηση $q(x) = e^x - x - 1$. Αρκεί να δείξουμε πως $q(x) \geq 0$ για κάθε x , και εφόσον $q(0) = 0$ αρκεί να δείξουμε πως η παράγωγος $q'(x) \geq 0$ για $x \geq 0$ και $q'(x) \leq 0$ για $x \leq 0$. Πράγματι, έχουμε $q'(x) = e^x - 1$, το οποίο προφανώς ικανοποιεί τις ζητούμενες συνθήκες. Η ανισότητα (7.13) αναπαρίσταται γραφικά και στο Σχήμα 7.1. ■

7.5 Ασκήσεις

1. *Overbooking.* Ένα αεροπλάνο έχει 50 θέσεις, και έχουν γίνει κρατήσεις από 55 επιβάτες. Η πιθανότητα καθέννας από αυτούς να έρθει στο αεροδρόμιο είναι 90%, ανεξάρτητα από τους υπόλοιπους. Ποια είναι η πιθανότητα κάποιοι επιβάτες που θα φθάσουν στο αεροδρόμιο να είναι υπεράριθμοι;

2. *Δείκτριες τυχαίες μεταβλητές.*

(α') Έστω ένα οποιοδήποτε ενδεχόμενο A με πιθανότητα $p = \mathbb{P}(A)$, και έστω η «δείκτρια» Τ.Μ. X , που παίρνει την τιμή $X = 1$ όταν συμβαίνει το A , και $X = 0$ όταν δεν συμβαίνει. Ποια είναι η κατανομή του X ;

(β') Έστω μια οποιαδήποτε Τ.Μ. Y με συνάρτηση κατανομής $F(y)$. Για δεδομένο $x \in \mathbb{R}$, η δείκτρια συνάρτηση $h_x : \mathbb{R} \rightarrow \{0, 1\}$ ορίζεται ως,

$$h_x(y) = \begin{cases} 1, & \text{αν } y \leq x, \\ 0, & \text{αν } y > x. \end{cases}$$

Δείξτε πως $E[h_x(Y)] = F(x)$.

3. *★ Η ουρά της μέσης τιμής.* Έστω μια διακριτή Τ.Μ. Y η οποία παίρνει πάντα τιμές μεγαλύτερες ή ίσες του μηδενός και έχει πεπερασμένη μέση τιμή $E(Y)$. Δείξτε πως, καθώς το $x \rightarrow \infty$,

$$E(Y[1 - h_x(Y)]) \rightarrow 0,$$

όπου $h_x(y)$ είναι οι δείκτριες συναρτήσεις που ορίσαμε στην Άσκηση 2. Ερμηνεύστε διαισθητικά αυτό το αποτέλεσμα.

4. *Ελάχιστο δύο γεωμετρικών τυχαίων μεταβλητών.* Έστω X και Y δύο ανεξάρτητες τυχαίες μεταβλητές, με παραμέτρους p_1 και p_2 αντίστοιχα. Υπολογίστε την πυκνότητα της τυχαίας μεταβλητής $Z = \min(X, Y)$. Μπορείτε να ερμηνεύσετε διαισθητικά το αποτέλεσμα; *Υπόδειξη.* Ίσως σας βοηθήσει να υπολογίσετε πρώτα την πιθανότητα $\Pr(Z \geq k)$.

5. *Άθροισμα Poisson.* Αν οι τυχαίες μεταβλητές X και Y είναι ανεξάρτητες, με κατανομή $\text{Poisson}(\lambda)$ και $\text{Poisson}(\mu)$ αντίστοιχα, να δείξετε ότι η τυχαία μεταβλητή $X + Y$ έχει κατανομή $\text{Poisson}(\lambda + \mu)$. *Υπόδειξη.* Το διωνυμικό θεώρημα λέει πως $(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}$.

6. *Ταυτότητα Vandermonde.* Αποδείξτε την ταυτότητα Vandermonde,

$$\sum_{k=0}^m \binom{n_1}{k} \binom{n_2}{m-k} = \binom{n_1 + n_2}{m}, \quad (7.14)$$

όπου $n_1, n_2, m \in \mathbb{N}$. Χρησιμοποιήστε επιχειρήματα που βασίζονται αποκλειστικά στη συνδυαστική. Υπενθυμίζεται πως, κατά σύμβαση, θέτουμε $\binom{n}{0} = 1$ για κάθε n , και $\binom{n}{k} = 0$ για κάθε $k > n$.

7. Άθροισμα διωνυμικών T.M. Χρησιμοποιήστε την ταυτότητα Vandermonde της Άσκησης 6, προκειμένου να δείξετε ότι το άθροισμα δύο διωνυμικών T.M. X, Y με παραμέτρους n_1, p και n_2, p αντίστοιχα, ανεξάρτητων μεταξύ τους, έχει $\Delta\text{ιων}(n_1 + n_2, p)$ κατανομή.

Επίσης εξηγήστε πως θα μπορούσαμε να αποδείξουμε το ίδιο αποτέλεσμα χωρίς να χρειαστεί να κάνουμε υπολογισμούς.

8. Διάφορες κατανομές. Παρακάτω ορίζονται κάποιες τυχαίες μεταβλητές. Περιγράψτε την κατανομή και τις αντίστοιχες παραμέτρους καθεμιάς από αυτές.

(α') Ρίχνουμε διαδοχικές (ανεξάρτητες) ζαριές με δύο ζάρια, μέχρι την 1^η φορά που θα φέρουμε διπλή. Έστω Y = συνολικό πλήθος από ζαριές που ρίξαμε.

(β') Επιλέγουμε στην τύχη, με επανατοποθέτηση, 6 φύλλα από μια τράπουλα. Έστω X το πλήθος από κούπες που επιλέξαμε.

(γ') Όπως στο (β'), αλλά χωρίς επανατοποθέτηση.

(δ') Ρίχνουμε 20 φορές ένα νόμισμα με $\text{Pr}(\text{Κορώνα}) = 0.3$. Έστω Z = το πλήθος των φορών που φέραμε Γράμματα.

9. Ουρά στην τράπεζα. Σε μια τράπεζα, κάθε πεντάλεπτο υπάρχει πιθανότητα 5% να έρθει ένας νέος πελάτης. Υποθέτουμε πως οι αφίξεις πελατών σε διαφορετικά πεντάλεπτα είναι ανεξάρτητες, πως ποτέ δεν έρχονται δύο ή παραπάνω πελάτες σε ένα δεδομένο πεντάλεπτο, και ορίζουμε:

$$X = \text{πλήθος πελατών που έφτασαν τις πρώτες 2 ώρες,}$$

$$Y = \text{το πρώτο 5λεπτο κατά το οποίο έφτασε πελάτης.}$$

(α') Ποια είναι η κατανομή του X ;

(β') Βρείτε την πιθανότητα να έρθουν ακριβώς 3 πελάτες τις πρώτες δύο ώρες.

(γ') Ποια είναι η κατανομή του Y ;

(δ') Κατά μέσο όρο πόσα λεπτά θα περιμένουν οι υπάλληλοι μέχρι την άφιξη του πρώτου πελάτη;

(ε') Δεδομένου ότι δεν ήρθε κανείς τις πρώτες 2 ώρες, ποια είναι η πιθανότητα να μην έρθει κανείς και κατά την επόμενη μισή ώρα;

10. Επιθέσεις μηνυμάτων spam. Κάποιος διαφημιστής στέλνει ένα email στους 10 χιλιάδες λογαριασμούς του domain aueb.gr ζητώντας τον αριθμό της πιστωτικής κάρτας του παραλήπτη. Από προηγούμενες προσπάθειές του γνωρίζει ότι οι διαφορετικοί χρήστες ανταποκρίνονται ανεξάρτητα ο ένας από τον άλλον, και η πιθανότητα να απαντήσει κάποιος με τα ζητούμενα στοιχεία είναι 0.018%. Ποια είναι η πιθανότητα να του απαντήσουν τουλάχιστον τρεις χρήστες;

11. Τα «ν» του Βαρουφάκη. Ως γνωστόν, ο πρώην υπουργός Οικονομικών κύριος Γιάνης Βαρουφάκης αρέσκειται στο να γράφει το μικρό του όνομα με ένα «ν» αντί για δύο. Στα κείμενα που ανεβαίνουν σε ένα δημοσιογραφικό σάιτ, με πιθανότητα 0.3% ο διορθωτής κάνει λάθος και γράφει το όνομά του με δύο «ν», ανεξάρτητα από φορά σε φορά. Αν σε έναν μήνα το όνομα του x . Βαρουφάκη αναφέρθηκε 800 φορές:
- (α') Εκφράστε το συνολικό πλήθος από «ν» που χρησιμοποιήθηκαν για όλες τις φορές που γράφτηκε το όνομά του, ως το άθροισμα μιας σταθεράς και μιας Τ.Μ. με γνωστή κατανομή. Ποια είναι αυτή η κατανομή;
 - (β') Υπολογίστε μια προσέγγιση για την πιθανότητα το συνολικό πλήθος από «ν» που χρησιμοποιήθηκαν να είναι ίσο με 810.
12. Άσχετος φοιτητής. Σε ένα διαγώνισμα υπάρχουν 250 ερωτήσεις πολλαπλής επιλογής. Ο εξεταζόμενος έχει παρανοήσει εντελώς το αντικείμενο του μαθήματος, και η πιθανότητα να απαντήσει σωστά είναι (πολύ χειρότερη από το να απαντούσε τυχαία!) μόνο 0.5% για την κάθε ερώτηση, ανεξαρτήτως των υπολοίπων ερωτήσεων. Βρείτε μια προσέγγιση της πιθανότητας ο εξεταζόμενος να απαντήσει σωστά σε ακριβώς τρεις ερωτήσεις.
13. Δύο διαφορετικές δημοσκοπήσεις. Σε κάποιον πληθυσμό M ψηφοφόρων, $3M/10$ άτομα, δηλαδή το 30%, είναι ψηφοφόροι κάποιου συγκεκριμένου κόμματος. Για να προβλέψει το ποσοστό αυτού του κόμματος στις εκλογές, μια εταιρία δημοσκοπήσεων διαλέγει τυχαία, ομοιόμορφα και με επανατοποθέτηση N άτομα από τον πληθυσμό, και καταγράφει το πλήθος Y εκείνων που δήλωσαν ψηφοφόροι του. Ως πρόβλεψη του εκλογικού αποτελέσματος του κόμματος, δίνει το ποσοστό Y/N των ψηφοφόρων του στο τυχαίο δείγμα της.
- (α') Ποια είναι η κατανομή του Y ; Ποια είναι η μέση τιμή και η διασπορά της πρόβλεψης του ποσοστού Y/N ;
 - (β') Επαναλάβετε το προηγούμενο ερώτημα στην περίπτωση που η επιλογή γίνεται χωρίς επανατοποθέτηση. Συγκρίνετε με τα προηγούμενα αποτελέσματα. Ποια μέθοδος είναι πιο συμφέρουσα;
14. Βελάκια. Ρίχνουμε δέκα βελάκια σε ένα στόχο, με κλειστά μάτια. Οι διαδοχικές ρίψεις είναι ανεξάρτητες, και η πιθανότητα να πετύχουμε τον στόχο είναι μόνο 1%.
- (α') Ποια η πιθανότητα ακριβώς 2 βελάκια να πέτυχαν τον στόχο;
 - (β') Αν αντί για 10, ρίχναμε 140 βελάκια, βρείτε μια προσέγγιση της ίδιας πιθανότητας.
15. Γινόμενο Bernoulli. Έστω n ανεξάρτητες τυχαίες μεταβλητές, X_1, X_2, \dots, X_n , όπου η κάθε $X_i \sim \text{Bern}(p_i)$, και έστω Z το γινόμενό τους. Παρατηρήστε πως το γινόμενο έχει επίσης κατανομή Bernoulli και υπολογίστε την παράμετρό της:
- (α') Χρησιμοποιώντας το Θεώρημα 6.1, και (β') χωρίς να χρησιμοποιήσετε το Θεώρημα 6.1.

16. *Χρηματιστήριο.* Στο χρηματιστήριο, κάθε μέρα η τιμή μιας μετοχής αυξάνεται κατά 1 ευρώ με πιθανότητα 14%, αλλιώς μένει σταθερή, και οι αλλαγές αυτές είναι ανεξάρτητες από τη μία μέρα στην άλλη. Έστω ότι το πρωί της 1^{ης} μέρας η τιμή είναι 100 ευρώ, και έστω Y_1 η μεταβολή της τιμής κατά την 1^η μέρα, Y_2 η μεταβολή της τιμής κατά τη 2^η μέρα, και γενικά Y_i η μεταβολή της τιμής κατά τη μέρα i .

- (α') Ποια είναι η κατανομή των τυχαίων μεταβλητών Y_i ;
- (β') Έστω X η συνολική διαφορά της τιμής μετά από ένα μήνα (δηλαδή 30 ημέρες). Ποια είναι η κατανομή της τυχαίας μεταβλητής X ;
- (γ') Ποια είναι η πιθανότητα του ενδεχόμενου E ότι μετά από ένα μήνα η τιμή θα είναι μεταξύ 110 και 112 ευρώ (συμπεριλαμβανομένων);
- (δ') Αν Z είναι η τιμή μετά από 60 μέρες, να βρεθεί η μέση τιμή και η διασπορά της.
- (ε') Αν αντί για 1 ευρώ, η τιμή αυξανόταν κατά δύο ευρώ ή έμενε σταθερή (με τις ίδιες πιθανότητες), να βρεθεί η μέση τιμή και η διασπορά της τιμής της μετοχής μετά από δύο μήνες.

17. *XOR Bernoulli.* Έστω X και Y δύο ανεξάρτητες T.M. οι οποίες περιγράφουν δύο τυχαία bits στην εκτέλεση ενός προγράμματος. Έστω ότι $X \sim \text{Bern}(1/4)$, $Y \sim \text{Bern}(1/2)$, και έστω μια νέα T.M. $Z = X \text{ XOR } Y$.

- (α') Η Z είναι δυαδική T.M. άρα έχει κατανομή Bernoulli. Να βρεθεί η παράμετρος της.
- (β') Είναι η Z ανεξάρτητη από τη X ή όχι; Αποδείξτε την απάντησή σας.

18. * *Διασπορά υπεργεωμετρικής κατανομής.* Στην παρατήρηση που ακολουθεί το Παράδειγμα 7.8 είδαμε πως μια T.M. $Y \sim \text{Υπερ}(N, k, n)$ μπορεί να εκφραστεί ως το άθροισμα n T.M. X_1, X_2, \dots, X_n οι οποίες έχουν όλες κατανομή $\text{Bern}(k/N)$ αλλά δεν είναι ανεξάρτητες. Εδώ θα αποδείξουμε τον τύπο (7.3) για τη διασπορά της Y .

- (α') Δείξτε πως:

$$E(Y^2) = \sum_{i=1}^n \sum_{j=1}^n E(X_i X_j).$$

- (β') Εξηγήστε γιατί, με βάση τον συλλογισμό που χρησιμοποιήσαμε για τον υπολογισμό της μέσης τιμής, όλοι οι όροι στο παραπάνω άθροισμα ικανοποιούν:

$$\begin{aligned} E(X_i X_j) &= E(X_1 X_2), & \text{για κάθε } i \neq j, \\ \text{και } E(X_i^2) &= E(X_1^2) = E(X_1), & \text{για κάθε } i. \end{aligned}$$

- (γ') Υπολογίστε την παράμετρο της Bernoulli τυχαίας μεταβλητής $X_1 X_2$.

(δ') Χρησιμοποιώντας τα τρία πιο πάνω βήματα, δείξτε πως:

$$E(Y^2) = n \frac{k}{N} + n(n-1) \frac{k}{N} \frac{k-1}{N-1}.$$

(ε') Χρησιμοποιώντας το αποτέλεσμα του βήματος (δ'), αποδείξτε τον τύπο (7.3) για τη διασπορά της Y .

19. ★ Αναπαράσταση της e^x . Η τρίτη ιδιότητα του Λήμματος 7.2 λέει ότι, για οποιαδήποτε πραγματική ακολουθία $\{x_n\}$ που τείνει στο $x \in \mathbb{R}$ καθώς το $n \rightarrow \infty$, ισχύει η σχέση (7.5):

$$e^x = \lim_{n \rightarrow \infty} \left(1 + \frac{x_n}{n}\right)^n.$$

Στα τρία πιο κάτω βήματα θα δούμε την απόδειξή της.

(α') Χρησιμοποιώντας το ανάπτυγμα Taylor, δείξτε ότι για οποιοδήποτε $y > -1$, έχουμε,

$$\log(1+y) = y - \frac{y^2}{2(1+\zeta)^2},$$

για κάποιο ζ τέτοιο ώστε $|\zeta| \leq |y|$.

(β') Χρησιμοποιήστε το προηγούμενο σέλος για να δείξετε ότι υπάρχει κάποιο $n_0 \geq 1$ και μια φραγμένη ακολουθία $\{\xi_n\}$, τέτοια ώστε:

$$n \log \left(1 + \frac{x_n}{n}\right) = x_n - \frac{1}{2n} \xi_n, \quad \text{για } n \geq n_0.$$

(γ') Εξηγήστε πως το πιο πάνω αποτέλεσμα συνεπάγεται ότι ισχύει η (7.5).

Κεφάλαιο 8

Παραδείγματα πιθανοκρατικής ανάλυσης αλγορίθμων

Σ' αυτό το κεφάλαιο θα αναπτύξουμε κάποια πιο ρεαλιστικά παραδείγματα εφαρμογών των πιθανοτήτων στην πληροφορική. Σε μια πρώτη ανάγνωση του βιβλίου, ο αναγνώστης που θέλει απλά να εξοικειωθεί με τις βασικές ιδέες και τεχνικές των πιθανοτήτων, μπορεί ασφαλώς να το προσπεράσει.

Στις Ενότητες **8.1** και **8.2** θα εξετάσουμε δύο κλασικά προβλήματα της επιστήμης υπολογιστών, και θα δούμε πως για το καθένα από αυτά υπάρχει μια μέθοδος να λυθεί χρησιμοποιώντας έναν «randomized» (ή τυχαιοκρατικό) αλγόριθμο. Όπως αναφέραμε στο εισαγωγικό Κεφάλαιο **1**, υπάρχουν αρκετά προβλήματα τα οποία επιλύονται πολύ πιο αποτελεσματικά αν επιτρέψουμε στον αλγόριθμο που χρησιμοποιούμε να πραγματοποιεί ορισμένα τυχαία βήματα κατά την εκτέλεσή του. Με προσεκτικό σχεδιασμό, αυτή η μέθοδος μπορεί να μειώσει σημαντικά την πολυπλοκότητα του αλγορίθμου. Το τίμημα αυτής της επιτάχυνσης είναι πως, λόγω της τυχαιότητας, υπάρχει πάντα κάποια πιθανότητα το αποτέλεσμα να είναι εσφαλμένο. Και πάλι με προσεκτικό σχεδιασμό ωστόσο, μπορούμε να εγγυηθούμε ότι η πιθανότητα αυτή θα είναι αποδεκτά μικρή, όπως θα δούμε λεπτομερώς παρακάτω.

Κατόπιν, στις Ενότητες **8.3** και **8.4** θα δούμε δύο παραδείγματα πιθανοκρατικής ανάλυσης αλγορίθμων. Συγκεκριμένα, θα παρουσιάσουμε δύο σχετικά απλούς, γνωστούς αλγορίθμους, και θα παρατηρήσουμε πως η κλασική τους πολυπλοκότητα (δηλαδή η πολυπλοκότητά τους για τη χειρότερη περίπτωση των δεδομένων εισόδου) είναι σχετικά μεγάλη. Αντίθετα, η συμπεριφορά τους πάνω σε τυχαία δεδομένα είναι πολύ διαφορετική: Η πολυπλοκότητά τους είναι σημαντικά μικρότερη με πιθανότητα πολύ κοντά στο 100%. Αυτό το φαινόμενο – ένας αλγόριθμος να είναι σχεδόν πάντα πολύ αποτελεσματικότερος από ό,τι στη χειρότερη δυνατή περίπτωση – είναι αρκετά συχνό και παίζει ιδιαίτερα σημαντικό ρόλο σε πολλές πρακτικές εφαρμογές.

Είναι επίσης αξιοσημείωτο πως, παρότι τα αποτελέσματα που θα αποδείξουμε είναι ισχυρά και έχουν ουσιαστικό πρακτικό ενδιαφέρον στο πλαίσιο σύγχρονων προβλημάτων της πληροφορικής,

η μαθηματική τους ανάλυση είναι σχετική απλή και δεν απαιτεί τίποτα παραπάνω από κάποιες από τις απλούστερες τεχνικές των πιθανοτήτων τις οποίες έχουμε συναντήσει στα Κεφάλαια 2 έως 6.

Κάποιες ενδιαφέρουσες επεκτάσεις και γενικεύσεις των τεσσάρων αυτών παραδειγμάτων δίνονται στις ασκήσεις στο τέλος του κεφαλαίου.

8.1 Επαλήθευση ισότητας πολυωνύμων

Εδώ εξετάζουμε το πρόβλημα της επαλήθευσης ισότητας πολυωνύμων. Τα δεδομένα μας είναι δύο πολυώνυμα βαθμού n ,

$$\begin{aligned} f(x) &= a_0 + a_1x + a_2x^2 + \cdots + a_{n-1}x^{n-1} + x^n \\ g(x) &= (x - b_1)(x - b_2) \cdots (x - b_n), \end{aligned}$$

και το ζητούμενο είναι να προσδιορίσουμε αν είναι τα ίδια, δηλαδή αν $f(x) = g(x)$ για κάθε x , ή όχι. Αν και δεν θα επεκταθούμε περαιτέρω, απλά σημειώνουμε πως αυτό το πρόβλημα και οι διάφορες προσεγγίσεις που έχουν αναπτυχθεί για τη λύση του έχουν μεγάλο εύρος πρακτικών εφαρμογών.

Θα δούμε τρεις μεθόδους, εκ των οποίων η τρίτη είναι ένας απλός randomized αλγόριθμος, και θα εκτιμήσουμε την πολυπλοκότητά τους. Για να απλοποιήσουμε την ανάλυση, η πολυπλοκότητα θα υπολογιστεί βάσει του πλήθους των πολλαπλασιασμών που απαιτεί η κάθε μέθοδος, δηλαδή θα αγνοήσουμε τις προσθέσεις και τις πράξεις που απαιτούνται για να υψώσουμε κάποιο πραγματικό αριθμό x σε μια ακέραια δύναμη k (βλ. την Άσκηση 3 στο τέλος του κεφαλαίου για μια πιο λεπτομερή ανάλυση).

1. Σύγκριση συντελεστών. Η πιο προφανής μέθοδος είναι να αναπτύξουμε το γινόμενο στην περιγραφή του $g(x)$ ώστε να έρθει στην ίδια μορφή με το $f(x)$, και να συγκρίνουμε τους αντίστοιχους συντελεστές.

Ο πολλαπλασιασμός $(x - b_1)(x - b_2)$ των δύο πρώτων όρων απαιτεί έναν πολλαπλασιασμό για να βρούμε το b_1b_2 . Στη συνέχεια, ο πολλαπλασιασμός του αποτελέσματος,

$$x^2 - (b_1 + b_2)x + b_1b_2,$$

με τον επόμενο όρο $(x - b_3)$ δίνει το $x^3 - (b_1 + b_2 + b_3)x^2 + [b_1b_2 - b_3(b_1 + b_2)]x - b_1b_2b_3$ και απαιτεί 2 επιπλέον πολλαπλασιασμούς.

Γενικά, στο βήμα k πολλαπλασιάζουμε το αποτέλεσμα του προηγούμενου βήματος (που είναι ένα πολυώνυμο βαθμού k) με τον επόμενο όρο $(x - b_{k+1})$, πράγμα που απαιτεί k επιπλέον πολλαπλασιασμούς (για να βρούμε το γινόμενο του b_{k+1} με τον καθένα από τους k συντελεστές του προηγούμενου πολυωνύμου). Συνολικά, το πλήθος των πολλαπλασιασμών που απαιτούνται είναι,

$$1 + 2 + \cdots + (n - 1) = \sum_{k=1}^{n-1} k = \frac{(n-1)n}{2} = \frac{1}{2}n^2 - \frac{1}{2}n.$$

Άρα, για μεγάλα n , η πολυπλοκότητα της μεθόδου αυτής είναι $O(n^2)$, δηλαδή «της τάξης του n^2 » ή «τετραγωνική».¹

2. Σύγκριση ριζών. Η αναπαράσταση του $g(x)$ μάς λέει πως έχει n πραγματικές ρίζες, τις b_1, b_2, \dots, b_n . Εφόσον ένα άλλο πολυώνυμο βαθμού n θα είναι ίσο με το $g(x)$ αν και μόνο αν έχουν τις ίδιες ρίζες, μπορούμε να υπολογίσουμε τις τιμές $f(b_i)$ για κάθε $i = 1, 2, \dots, n$. Αν όλες είναι ίσες με το μηδέν, τότε τα δύο πολυώνυμα θα είναι τα ίδια, ενώ αν $f(b_i) \neq 0$ για κάποιο i , θα είναι διαφορετικά.

Ο υπολογισμός του κάθε $f(b_i)$ απαιτεί n πολλαπλασιασμούς, και η επανάληψη για όλα τα $i = 1, 2, \dots, n$ απαιτεί συνολικά n^2 πολλαπλασιασμούς. Και πάλι λοιπόν έχουμε έναν αλγόριθμο με πολυπλοκότητα της τάξης του n^2 .

3. Ένας randomized αλγόριθμος. Έστω τώρα πως επιλέγουμε έναν τυχαίο ακέραιο X μεταξύ του 1 και του $100n$. Δηλαδή το X είναι μια τυχαία μεταβλητή με σύνολο τιμών το $S = \{1, 2, \dots, 100n\}$ και πυκνότητα $P(x) = \frac{1}{100n}$ για κάθε $x \in S$. Υπολογίζουμε τις τιμές $f(X)$ και $g(X)$, και αν είναι ίσες δηλώνουμε πως τα δύο πολυώνυμα είναι τα ίδια, ενώ αν $f(X) \neq g(X)$ δηλώνουμε πως είναι διαφορετικά.

Πριν εξετάσουμε την πιθανότητα να δώσουμε λάθος αποτέλεσμα, παρατηρούμε ότι η πολυπλοκότητα αυτού του randomized αλγορίθμου είναι μόλις $2n$, δηλαδή γραμμική ως προς το n και άρα σημαντικά μικρότερη από εκείνη των δύο προηγούμενων μεθόδων.

Έστω τώρα $P_e = \Pr(\text{«είγοι»})$ η πιθανότητα σφάλματος του αλγορίθμου. Εξετάζουμε δύο περιπτώσεις. Αν τα $f(x)$ και $g(x)$ είναι ίσα, τότε θα έχουμε πάντοτε $f(X) = g(X)$ και ο αλγόριθμος θα δίνει τη σωστή απάντηση, άρα $P_e = 0$. Αν τα $f(x)$ και $g(x)$ είναι διαφορετικά, τότε η διαφορά τους $f(x) - g(x)$ είναι ένα πολυώνυμο βαθμού το πολύ $n - 1$. Κατά συνέπεια θα έχει το πολύ $n - 1$ πραγματικές ρίζες, τις οποίες συμβολίζουμε ως $\{\rho_1, \rho_2, \dots, \rho_k\}$ για κάποιον ακέραιο $k \leq n - 1$. Στην περίπτωση, λοιπόν, που τα $f(x)$ και $g(x)$ είναι διαφορετικά, μπορούμε να εκφράσουμε την πιθανότητα σφάλματος ως εξής:

$$\begin{aligned} P_e &= \Pr(f(X) = g(X)) \\ &= \Pr(f(X) - g(X) = 0) \\ &= \Pr(X \in \{\rho_1, \rho_2, \dots, \rho_k\}) \\ &= \sum_{i=1}^k \Pr(X = \rho_i). \end{aligned}$$

Αν κάποιο ρ_i είναι ακέραιος μεταξύ 1 και $100n$, τότε $\Pr(X = \rho_i) = \frac{1}{100n}$, και αν όχι τότε $\Pr(X = \rho_i) = 0$. Σε κάθε περίπτωση, $\Pr(X = \rho_i) \leq \frac{1}{100n}$, οπότε,

$$P_e \leq k \cdot \frac{1}{100n} \leq 1\%,$$

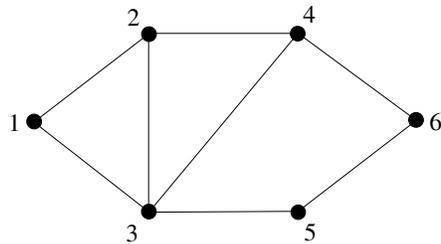
¹Εδώ χρησιμοποιήσαμε τον τύπο για το άθροισμα των όρων μιας αριθμητικής προόδου $\sum_{k=1}^n k = n(n+1)/2$. Η Άσκηση 2 στο τέλος του κεφαλαίου δίνει κάποιες λεπτομέρειες για την απόδειξή του.

όπου στο δεύτερο βήμα χρησιμοποιήσαμε το γεγονός ότι το k είναι το πολύ $n - 1$ και το προφανές φράγμα $n - 1 \leq n$.

Συνεπώς, με τη χρήση τυχειότητας στο σχεδιασμό του αλγορίθμου, πετύχαμε και μεγάλη μείωση της πολυπλοκότητας και μικρή πιθανότητα σφάλματος. Συγκεκριμένα, όταν το αποτέλεσμα του αλγορίθμου είναι « $f \neq g$ », είμαστε βέβαιοι ότι είναι σωστό. Ενώ, όταν το αποτέλεσμα είναι « $f \equiv g$ », ξέρουμε ότι η πιθανότητα να είναι πράγματι ίσα τα δύο πολυώνυμα είναι τουλάχιστον 99%.

8.2 Εύρεση ελαχιστιαίου cut set σε γράφους

Σκεφτείτε ένα δίκτυο n υπολογιστών όπου ο κάθε υπολογιστής είναι συνδεδεμένος με κάποιους, αλλά όχι απαραίτητα όλους, από τους υπόλοιπους. Ένα παράδειγμα ενός τέτοιου δικτύου $n = 6$ υπολογιστών δίνεται στο Σχήμα 8.1, όπου έχουμε συμβολίσει τον κάθε υπολογιστή i με τον αντίστοιχο κόμβο i , για $i = 1, 2, \dots, n$, και δύο υπολογιστές i, j είναι συνδεδεμένοι αν και μόνο αν οι αντίστοιχοι κόμβοι στον γράφο συνδέονται με την ακμή ij .



Σχήμα 8.1: Σχηματική αναπαράσταση του γράφου που περιγράφει ένα δίκτυο έξι υπολογιστών και τις αντίστοιχες μεταξύ τους συνδέσεις.

Δύο υπολογιστές i, j στο πιο πάνω δίκτυο μπορούν να επικοινωνήσουν αν είναι άμεσα συνδεδεμένοι, αν δηλαδή η αντίστοιχη ακμή ij περιέχεται στον γράφο, ή αν υπάρχει ένα μονοπάτι ακμών που να ξεκινάει από τον i και να καταλήγει στον j . Για παράδειγμα, οι κόμβοι 1 και 2 επικοινωνούν άμεσα, ενώ οι 2 και 5 επικοινωνούν μέσω του μονοπατιού (23, 35).

Σε πολλών ειδών πραγματικά επικοινωνιακά δίκτυα (φυσικά, πολύ μεγαλύτερα και πιο πολύπλοκα από το παραπάνω απλοϊκό παράδειγμα), μας ενδιαφέρει να μπορούν όλοι οι κόμβοι να επικοινωνήσουν μεταξύ τους, δηλαδή, στον γράφο που τα αναπαριστά να υπάρχει τουλάχιστον ένα μονοπάτι ακμών που να συνδέει οποιουδήποτε δύο κόμβους. Όταν αυτό συμβαίνει, λέμε ότι ο γράφος είναι *συνδεδεμένος* ή *συνεκτικός*.

Έστω λοιπόν ένας (μη κατευθυνόμενος) γράφος (V, E) , όπου το $V = \{1, 2, \dots, n\}$ είναι το σύνολο των $n = \#V$ κόμβων του και E είναι το σύνολο των ακμών του, και όπου η ακμή $e \in E$ που συνδέει τους κόμβους i και j συμβολίζεται ως $e = ij$. Ας υποθέσουμε ότι ο γράφος (V, E) είναι συνδεδεμένος. Αν, π.χ., αυτός ο γράφος περιγράφει ένα επικοινωνιακό δίκτυο, μας ενδιαφέρει να γνωρίζουμε πόσο ευαίσθητο είναι το δίκτυο σε δυσλειτουργίες ή βλάβες κάποιων

από τις συνδέσεις του. Για παράδειγμα, ποιο είναι το μικρότερο πλήθος συνδέσεων που, αν πάψουν να λειτουργούν, θα υπάρχουν κάποιοι κόμβοι που δε θα μπορούν να επικοινωνήσουν μεταξύ τους; Με άλλα λόγια, ποιο είναι το μέγεθος του μικρότερου υποσυνόλου $C \subset E$, τέτοιο ώστε, αν αφαιρέσουμε τις ακμές που περιέχονται στο C , ο νέος γράφος δεν θα είναι πια συνδεδεμένος;

Στα διακριτά μαθηματικά κάθε τέτοιο σύνολο $C \subset E$ ονομάζεται *cut set* (ή «σύνολο αποκοπτούσων ακμών»). Εάν επιπλέον ένα cut set C έχει την ιδιότητα να έχει μέγεθος μικρότερο ή ίσο με εκείνο οποιουδήποτε άλλου cut set, τότε ονομάζεται «ελαχιστιαίο cut set» ή πιο απλά *min-cut*. Για παράδειγμα, στο γράφο του Σχήματος 8.1 παρατηρούμε πως τα $C_1 = \{12, 13\}$, $C_2 = \{24, 34, 35\}$ και $C_3 = \{46, 56\}$ είναι cut sets και πως τα C_1 και C_3 είναι min-cuts.

Πέραν των εφαρμογών που αναφέραμε παραπάνω, το πρόβλημα της εύρεσης ενός min-cut σε ένα γράφο είναι και θεωρητικά ενδιαφέρον. Για παράδειγμα, δεν υπάρχει κάποιος προφανής τρόπος να βρούμε σχετικά εύκολα ένα min-cut, παρά μόνο να εξετάσουμε όλα τα δυνατά cut sets και να επιλέξουμε το μικρότερο. Αλλά το πλήθος των cut sets ενός γράφου με n κόμβους είναι $\Gamma_n = 2^{n-1} - 1$, το οποίο είναι προφανώς απαγορευτικά (εκθετικά) μεγάλο για ρεαλιστικές τιμές του n .

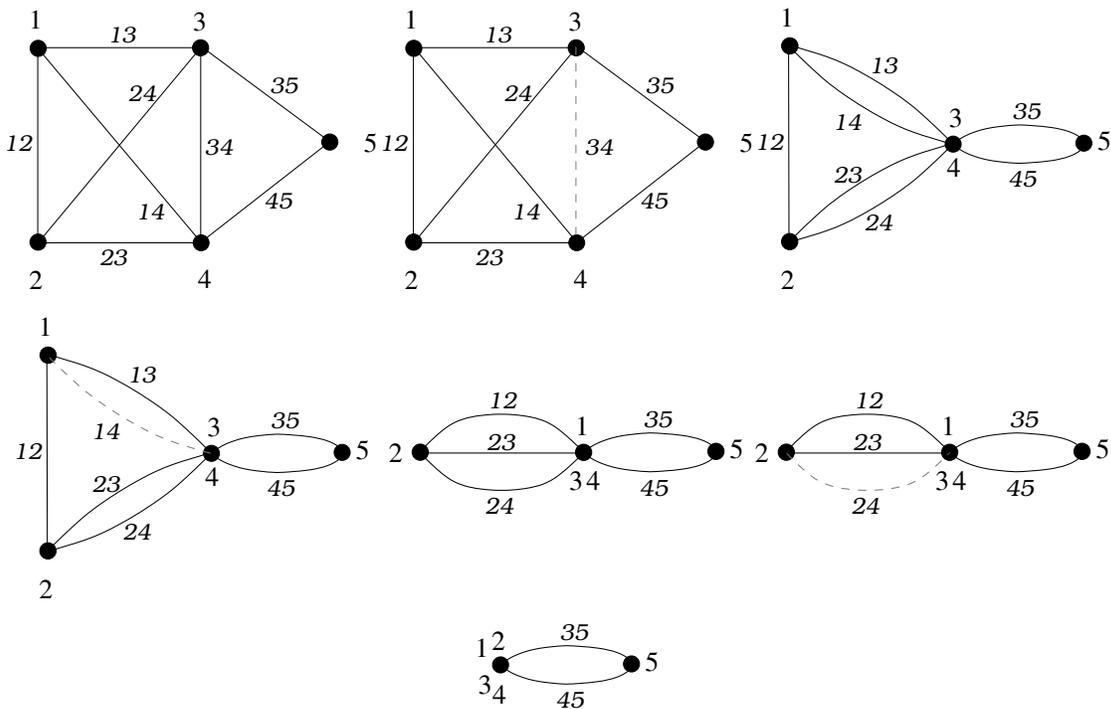
Για να δούμε ότι πράγματι υπάρχουν Γ_n υποψήφια cut sets, παρατηρούμε πως κάθε cut set αντιστοιχεί σε ένα διαχωρισμό των κόμβων V του γράφου σε δύο (ξένα) υποσύνολα K_1 και K_2 , τέτοια ώστε κανένας κόμβος στο K_1 να μην επικοινωνεί με οποιονδήποτε κόμβο στο K_2 και η ένωσή τους να μας δίνει όλους τους κόμβους, δηλαδή $V = K_1 \cup K_2$. Κάθε τέτοιος διαχωρισμός αντιστοιχεί ακριβώς σε ένα cut set, το οποίο αποτελείται από τις ακμές που συνδέουν κάποιο στοιχείο του K_1 με κάποιο στοιχείο του K_2 . Άρα, κάθε cut set μπορεί να περιγραφεί από την επιλογή ενός υποσυνόλου K_1 του V , όπου φυσικά υπάρχουν $2^n - 2$ υποψήφια υποσύνολα (μια που δεν συμπεριλαμβάνουμε το κενό σύνολο και ολόκληρο το V). Αλλά επειδή οι ρόλοι των K_1 και K_2 είναι συμμετρικοί, κάθε cut set μπορεί να περιγραφεί μέσω δύο διαχωρισμών, απλά αντιστρέφοντας τους ρόλους του K_1 και του K_2 . Συνεπώς, το πλήθος όλων των δυνατών διαχωρισμών, δηλαδή το πλήθος όλων των υποψήφιων cut sets, είναι πράγματι $\Gamma_n = (2^n - 2)/2 = 2^{n-1} - 1$.

Ο παρακάτω randomized αλγόριθμος μάς δίνει μια απλή και σχετικά χαμηλής πολυπλοκότητας μέθοδο για την εύρεση ενός min-cut σε έναν οποιονδήποτε γράφο, με μεγάλη πιθανότητα. Η κεντρική ιδέα του αλγορίθμου είναι η έννοια της *συνένωσης* δύο κόμβων.

Ο αλγόριθμος του Karger (1993) Έστω ένας μη κατευθυνόμενος, συνδεδεμένος γράφος (V, E) με $n = \#V$ κόμβους.

- Επιλέγουμε τυχαία και ομοιόμορφα μία από τις ακμές $e = ij$ του γράφου, και *συνενώνουμε* τους δύο κόμβους i και j σε έναν. Συμβολίζουμε τον νέο κόμβο i, j , αφαιρούμε την ακμή ij και διατηρούμε όλες τις υπόλοιπες.
- Π.χ., ξεκινώντας από τον γράφο στο πρώτο διάγραμμα του Σχήματος 8.2, αν επιλέξουμε την ακμή 34 όπως φαίνεται στο δεύτερο διάγραμμα, καταλήγουμε στον γράφο του τρίτου διαγράμματος.

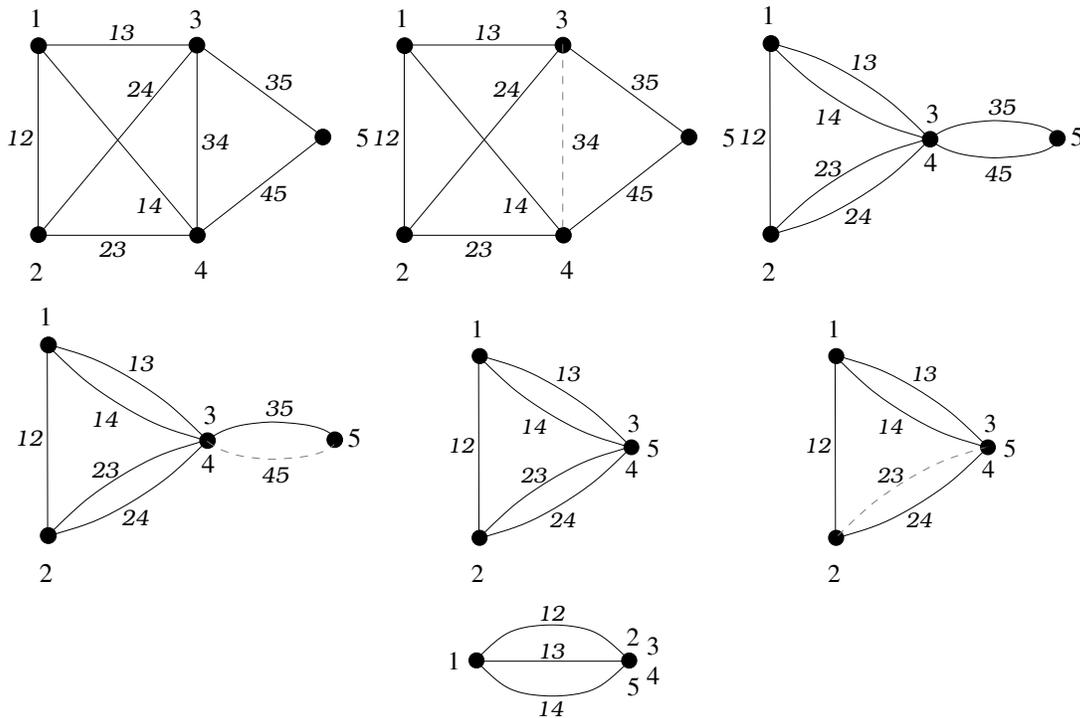
- Παρατηρούμε τώρα ότι ένας κόμβος του νέου γράφου θα συμβολίζεται με δύο αριθμούς αντί για έναν, και ότι κάποιοι κόμβοι μπορεί να συνδέονται με παραπάνω από μία ακμές.
- Συνεχίζουμε επαγωγικά: Επιλέγουμε τυχαία και ομοιόμορφα μία από τις υπάρχουσες ακμές του νέου γράφου και επαναλαμβάνουμε την ίδια διαδικασία: *Συνενώνουμε* τους δύο κόμβους που συνδέει η επιλεγμένη ακμή, συμβολίζουμε τον νέο κόμβο με το σύνολο των συμβόλων που περιέγραφαν τους δύο κόμβους, αφαιρούμε τις ακμές που τους συνέδεαν μεταξύ τους και διατηρούμε όλες τις υπόλοιπες.
- Π.χ., ξεκινώντας από το γράφο στο τρίτο διάγραμμα του Σχήματος 8.2, αν επιλέξουμε την ακμή 14 όπως φαίνεται στο τέταρτο διάγραμμα, καταλήγουμε στο γράφο του πέμπτου διαγράμματος.
- Η ίδια διαδικασία επαναλαμβάνεται για ακριβώς $(n - 2)$ βήματα, οπότε και θα έχουμε αναγκαστικά καταλήξει σε ένα γράφο με ακριβώς δύο κόμβους.



Σχήμα 8.2: Παράδειγμα της εκτέλεσης του αλγορίθμου του Karger σε έναν απλό γράφο. Σε κάθε βήμα, η ακμή που επιλέγεται συμβολίζεται με διακεκομμένη γκριζα γραμμή.

Παρατηρήσεις

1. Το αποτέλεσμα του αλγορίθμου του Karger είναι ο διαχωρισμός όλων των κόμβων του αρχικού γράφου σε δύο ομάδες που συνδέονται μεταξύ τους μέσω όλων των ακμών του αρχικού γράφου οι οποίες συνδέαν κάποιον κόμβο της πρώτης ομάδας με κάποιον της δεύτερης. Συνεπώς, το σύνολο των ακμών στο οποίο καταλήγει ο αλγόριθμος πάντοτε αντιστοιχεί σε ένα cut set. Για παράδειγμα, στο Σχήμα 8.3 καταλήγουμε στο cut set $C = \{35, 45\}$, το οποίο εύκολα βλέπουμε πως είναι και min-cut για αυτόν το γράφο.
2. Στο Σχήμα 8.3 βλέπουμε μια διαφορετική πιθανή έκβαση του αλγορίθμου του Karger, όπου καταλήγουμε σε ένα διαφορετικό cut set, το $C' = \{12, 13, 14\}$, το οποίο προφανώς δεν είναι min-cut. Άρα παρατηρούμε πως ο αλγόριθμος δεν μας δίνει πάντα ένα min-cut ως αποτέλεσμα.



Σχήμα 8.3: Ένα διαφορετικό παράδειγμα της εκτέλεσης του αλγορίθμου του Karger.

Το επόμενο θεώρημα μάς δίνει πολύ ακριβείς πληροφορίες για το πώς ακριβώς και πόσο αποτελεσματικά μπορεί να εφαρμοστεί ο πιο πάνω αλγόριθμος στην πράξη. [Θυμίζουμε πως με \log συμβολίζουμε τον φυσικό λογάριθμο με βάση το e , ο οποίος συχνά συμβολίζεται και ως \ln .]

Θεώρημα 8.1 Έστω ένας οποιοσδήποτε μη κατευθυνόμενος, συνδεδεμένος γράφος (V, E) με $n = \#V$ κόμβους.

1. Αν εφαρμόσουμε τον αλγόριθμο του Karger, η πιθανότητα P_s το cut set στο οποίο θα καταλήξουμε να είναι και *min-cut* για αυτόν το γράφο δεν μπορεί να είναι αυθαίρετα μικρή· συγκεκριμένα,

$$P_s = \Pr(\text{«success»}) \geq \frac{2}{n(n-1)}.$$

2. Αν εφαρμόσουμε τον αλγόριθμο του Karger σε $n(n-1) \log n$ ανεξάρτητες επαναλήψεις, και ως αποτέλεσμα πάρουμε το μικρότερο από όλα τα cut sets στα οποία θα καταλήξουμε, η πιθανότητα P_e το τελικό αποτέλεσμα να μην είναι *min-cut* είναι μικρή· συγκεκριμένα,

$$P_e = \Pr(\text{«error»}) \leq \frac{1}{n^2}.$$

Πριν την απόδειξη, παραθέτουμε κάποιες σημαντικές παρατηρήσεις σχετικά με το παραπάνω αποτέλεσμα και με το γενικό πρόβλημα της εύρεσης *min-cuts* σε γράφους.

Παρατηρήσεις

1. Ο πιο απλός τρόπος να κατασκευάσουμε ένα cut set είναι να επιλέξουμε όλες τις ακμές ενός συγκεκριμένου κόμβου· προφανώς, αν τις αφαιρέσουμε, τότε ο γράφος χωρίζεται σε δύο μέρη (τον συγκεκριμένο κόμβο και όλους τους υπόλοιπους) τα οποία δεν συνδέονται μεταξύ τους. Κατά συνέπεια, το πλήθος των στοιχείων που περιέχει ένα *min-cut* θα είναι πάντα μικρότερο ή ίσο του πλήθους των ακμών που συνδέονται με οποιονδήποτε κόμβο. Ή, αντίστροφα, αν υπάρχει κάποιο *min-cut* $C \subset E$ το οποίο περιέχει k στοιχεία, τότε ο κάθε κόμβος του γράφου θα έχει τουλάχιστον k ακμές.
2. Εφόσον κάθε επανάληψη του αλγορίθμου του Karger απαιτεί $(n-2)$ βήματα, ο αλγόριθμος που περιγράφεται στο δεύτερο μέρος του θεωρήματος απαιτεί $n(n-1)(\log n)(n-2)$ βήματα, και άρα η πολυπλοκότητα του είναι της τάξης $O(n^3 \log n)$. Επιπλέον, για πρακτικά ρεαλιστικές τιμές του n η πιθανότητα σφάλματος είναι αμελητέα· π.χ., για $n \geq 100$, η πιθανότητα σφάλματος είναι το πολύ 0.01%.
3. Υπάρχει μια τροποποίηση της μεθόδου που είδαμε στο θεώρημα, ο αλγόριθμος Karger-Stein, ο οποίος επιτρέπει την εύρεση ενός *min-cut* με μεγάλη πιθανότητα και με πολυπλοκότητα της τάξης $O(n^2(\log n)^r)$ για κάποιο r .

Από την άλλη μεριά, παρατηρούμε πως υπάρχουν γράφοι με $O(n^2)$ πλήθος ακμών (όπως για παράδειγμα ο πλήρης γράφος σε n κόμβους) και πως για να βρούμε κάποιο *min-cut* πρέπει να εξετάσουμε την κάθε ακμή τουλάχιστον μία φορά (για να διαπιστώσουμε αν ο γράφος που προκύπτει είναι συνδεδεμένος). Συνεπώς, η μικρότερη δυνατή πολυπλοκότητα οποιουδήποτε

ντετερμινιστικού αλγορίθμου για την εύρεση κάποιου min-cut είναι τουλάχιστον της τάξης των $O(n^2)$ βημάτων.

Ο καλύτερος γνωστός ως σήμερα ντετερμινιστικός αλγόριθμος γι' αυτό το πρόβλημα, απαιτεί $O(mn + n^2 \log n)$ βήματα, όπου $m = \#E$ είναι το πλήθος των ακμών του αρχικού γράφου.

4. Χρησιμοποιώντας κάποιους από τους υπολογισμούς που θα κάνουμε στην απόδειξη του θεωρήματος πιο κάτω, στην Άσκηση 5 στο τέλος του κεφαλαίου θα δείξετε το ακόλουθο ενδιαφέρον πόρισμα: Για οποιονδήποτε γράφο με n κόμβους, υπάρχουν το πολύ $\binom{n}{2}$ min-cuts.
5. Έχοντας δει το αποτέλεσμα του θεωρήματος, μια εύλογη ερώτηση είναι η εξής: Δεν θα ήταν ευκολότερο, αντί να χρησιμοποιήσουμε τον αλγόριθμο του Karger, απλά να επιλέξουμε στην τύχη ένα οποιοδήποτε cut set πολλές φορές και να πάρουμε ως αποτέλεσμα ενός υποψηφίου min-cut το μικρότερο από αυτά; Όπως θα δείξετε στην Άσκηση 6 στο τέλος του κεφαλαίου, αυτή η μέθοδος δεν είναι αποτελεσματική, γιατί η πιθανότητα ένα τυχαία επιλεγμένο cut set να είναι και min-cut είναι σημαντικά μικρότερη από την αντίστοιχη πιθανότητα P_s του αλγορίθμου του Karger.

Απόδειξη του Θεωρήματος 8.1: Όπως έχουμε ήδη παρατηρήσει, σε κάποιον γράφο (V, E) μπορούν να υπάρχουν περισσότερα από ένα min-cuts. Έστω $C \subset E$ ένα συγκεκριμένο από αυτά τα min-cuts, και έστω $k = \#C$ το πλήθος των στοιχείων του. Όπως αναφέραμε στην πρώτη από τις παραπάνω παρατηρήσεις, ο κάθε κόμβος του γράφου θα έχει τουλάχιστον k ακμές, άρα το συνολικό πλήθος $\#E$ των ακμών του θα είναι τουλάχιστον $\#E \geq kn/2$.

Κατ' αρχάς θα υπολογίσουμε ένα φράγμα για την πιθανότητα του ενδεχομένου:

$$S = \{\text{το αποτέλεσμα του αλγορίθμου είναι το } C\}.$$

Ορίζουμε για κάθε $i = 1, 2, \dots, n-2$ τα ενδεχόμενα,

$$S_i = \{\text{στο βήμα } i \text{ ο αλγόριθμος δεν επέλεξε κάποια ακμή από το } C\},$$

και παρατηρούμε πως το C θα είναι το αποτέλεσμα του αλγορίθμου αν και μόνο αν σε κανένα βήμα δεν επιλεγεί κάποια ακμή του, δηλαδή,

$$S = \bigcap_{i=1}^{n-2} S_i.$$

Οπότε, από τον ορισμό της δεσμευμένης πιθανότητας (και τη γενίκευσή του στην Άσκηση 4 του Κεφαλαίου 5) έχουμε:

$$\Pr(S) = \Pr(S_1) \prod_{i=2}^{n-2} \Pr(S_i | S_1 \cap S_2 \cap \dots \cap S_{i-1}). \quad (8.1)$$

Για την πρώτη πιθανότητα, εφόσον η επιλογή στο πρώτο βήμα γίνεται ομοιόμορφα ανάμεσα σε όλες τις ακμές του γράφου, έχουμε πως,

$$\Pr(S_1) = 1 - \Pr(S'_1) = 1 - \frac{\#C}{\#E} \geq 1 - \frac{k}{kn/2} = \frac{n-2}{n}, \quad (8.2)$$

όπου χρησιμοποιήσαμε την παραπάνω παρατήρηση πως $\#E \geq kn/2$.

Τώρα, δεδομένου ότι στα πρώτα $i-1$ βήματα δεν έχει επιλεχθεί καμία ακμή του C , έχουμε έναν νέο γράφο (V_i, E_i) ο οποίος περιέχει $(n-(i-1))$ κόμβους και επίσης ακόμα περιέχει το min-cut C , που έχει k στοιχεία. Άρα, ο νέος γράφος έχει τουλάχιστον $k(n-(i-1))/2$ ακμές, και για καθεμία από τις υπόλοιπες πιθανότητες στο γινόμενο της σχέσης (8.1), με τον ίδιο συλλογισμό όπως πριν βρίσκουμε,

$$\begin{aligned} \Pr(S_i | S_1 \cap S_2 \cap \dots \cap S_{i-1}) &= 1 - \Pr(S'_i | S_1 \cap S_2 \cap \dots \cap S_{i-1}) \\ &= 1 - \frac{\#C}{\#E'} \\ &\geq 1 - \frac{k}{k(n-i+1)/2} \\ &= \frac{n-i-1}{n-i+1}. \end{aligned} \quad (8.3)$$

Αντικαθιστώντας τα φράγματα των σχέσεων (8.2) και (8.3) στην (8.1), έχουμε,

$$\Pr(S) \geq \frac{n-2}{n} \times \frac{n-3}{n-1} \times \frac{n-4}{n-2} \times \dots \times \frac{3}{5} \times \frac{2}{4} \times \frac{1}{3},$$

όπου σχεδόν όλοι οι όροι απλοποιούνται (εκτός από τους δύο πρώτους παρονομαστές και τους δύο τελευταίους αριθμητές) ώστε,

$$\Pr(S) \geq \frac{2}{n(n-1)}.$$

Τέλος, παρατηρούμε πως το ενδεχόμενο S είναι υποσύνολο του ενδεχομένου «success», μια που έχουμε επιτυχία του αλγορίθμου αν καταλήξει στο C ή σε οποιοδήποτε άλλο min-cut. Οπότε, από τον δεύτερο κανόνα πιθανότητας, τελικά βρίσκουμε,

$$P_s = \Pr(\text{«success»}) \geq \Pr(C) \geq \frac{2}{n(n-1)},$$

κι έτσι έχουμε αποδείξει το πρώτο μέρος του θεωρήματος.

Για το δεύτερο μέρος, ορίζουμε για κάθε $j = 1, 2, \dots, R_n$ τα ενδεχόμενα,

$$F_j = \{\text{στην εκτέλεση } j \text{ ο αλγόριθμος δεν κατέληξε σε min-cut}\},$$

όπου $R_n = n(n-1) \log n$ είναι το πλήθος των επαναλήψεων, και παρατηρούμε πως, για κάθε j , από το πρώτο μέρος έχουμε,

$$\Pr(F_j) = 1 - \Pr(F'_j) = 1 - P_s \leq \left(1 - \frac{2}{n(n-1)}\right). \quad (8.4)$$

Συνεπώς, εφόσον έχουμε «error» αν και μόνο αν σε καθεμία από τις ανεξάρτητες επαναλήψεις j δεν καταλήξει ο αλγόριθμος σε min-cut, βρίσκουμε ότι,²

$$\begin{aligned} P_e = \Pr(\text{«error»}) &= \Pr\left(\bigcap_{j=1}^{R_n} F_j\right) \\ &\stackrel{(a)}{=} \prod_{j=1}^{R_n} \Pr(F_j) \\ &= (1 - P_s)^{R_n} \\ &\stackrel{(b)}{\leq} \left(1 - \frac{2}{n(n-1)}\right)^{n(n-1) \log n} \\ &\stackrel{(c)}{\leq} \left(e^{-\frac{2}{n(n-1)}}\right)^{n(n-1) \log n} \\ &= e^{-2 \log n} = \frac{1}{n^2}, \end{aligned}$$

το οποίο είναι και το ζητούμενο αποτέλεσμα. Στον πιο πάνω υπολογισμό, στο βήμα (a) χρησιμοποιήσαμε το γεγονός ότι οι επαναλήψεις είναι ανεξάρτητες μεταξύ τους, στο (b) εφαρμόσαμε το φράγμα της σχέσης (8.4), και στο (c) χρησιμοποιήσαμε την απλή ανισότητα του Λήμματος 7.3, με $x = -\frac{2}{n(n-1)}$. ■

8.3 String matching

Πολλοί αλγόριθμοι επεξεργασίας σήματος, συμπίεσης, κλπ., οι οποίοι δρουν σε ακολουθίες ψηφιακών δεδομένων, εκτελούν κάποιο είδος «string matching» (ή *ταίριασμα ακολουθιών*), δηλαδή ανιχνεύουν συγκεκριμένα μοτίβα στα δεδομένα τα οποία κατόπιν εκμεταλλεύονται για την πιο αποτελεσματική επεξεργασία τους. Σε αυτή την ενότητα θα εξετάσουμε ένα πολύ απλό τέτοιο παράδειγμα.

Έστω ότι ένας αλγόριθμος εκτελείται με δεδομένα εισόδου μια ακολουθία x_1, x_2, \dots, x_n από bits x_i , και ο χρόνος εκτέλεσής του ισούται με το μήκος της μεγαλύτερης ακολουθίας συνεχόμενων «1» μεταξύ των x_i .

²Στον υπολογισμό της πιθανότητας P_e εμμέσως υποθέτουμε ότι το πλήθος $R_n = n(n-1) \log n$ των επαναλήψεων είναι ακέραιος αριθμός. Άσκηση. Αν το R_n δεν είναι ακέραιος, δείξτε πως μπορούμε να επαναλάβουμε τα ίδια ακριβώς βήματα για $R_n = \lceil n(n-1) \log n \rceil$ και να καταλήξουμε στο ίδιο ακριβώς αποτέλεσμα.

Προφανώς η πολυπλοκότητα αυτού του αλγορίθμου ισούται με n , αφού στη χειρότερη περίπτωση θα έχουμε $x_1, x_2, \dots, x_n = 1, 1, \dots, 1$. Θα δείξουμε πως, αν τα δεδομένα εισόδου είναι τυχαία, δηλαδή προκύπτουν από μια ακολουθία ανεξάρτητων $\text{Bern}(1/2)$ τυχαίων μεταβλητών X_1, X_2, \dots, X_n , τότε, για μεγάλα n , η πολυπλοκότητα του αλγορίθμου δεν ξεπερνά το $2 \log_2 n$, με μεγάλη πιθανότητα. Συγκεκριμένα, ορίζοντας μια νέα τυχαία μεταβλητή,

$$L_n = \text{μέγιστο μήκος από συνεχόμενα «1» στα } X_1, X_2, \dots, X_n,$$

τότε,

$$\Pr(L_n \leq 2 \log_2 n) \geq 1 - \frac{1}{n}. \quad (8.5)$$

Πριν αποδείξουμε την (8.5) παρατηρούμε πως υπάρχει δραματικά μεγάλη διαφορά μεταξύ του n και του $(2 \log_2 n)$ για ρεαλιστικές τιμές του n . Για παράδειγμα, αν έχουμε $n = 10^6 = \text{ένα εκατομμύριο bits}$, δηλαδή 125 KB, τότε το $(2 \log_2 n)$ είναι λίγο μικρότερο από 40, και η πιθανότητα να έχουμε πολυπλοκότητα μεγαλύτερη από 40 είναι το πολύ $1/n = 10^{-6}$!

Θα αποδείξουμε την εξής σχέση, η οποία είναι ισοδύναμη (από τον κανόνα πιθανότητας #4 του Κεφαλαίου 3) με την (8.5):

$$\Pr(L_n > 2 \log_2 n) \leq \frac{1}{n}. \quad (8.6)$$

Θέτουμε $k = 2 \log_2 n$, και για κάθε $i = 1, 2, \dots, n - k + 1$, ορίζουμε τα ενδεχόμενα,

$$\begin{aligned} E_i &= \text{«}k \text{ συνεχόμενα «1» ξεκινώντας από τη θέση } i\text{»} \\ &= \{X_i = X_{i+1} = \dots = X_{i+k-1} = 1\}. \end{aligned}$$

Παρατηρούμε πως θα υπάρχει στα X_i μια ακολουθία από «1» με μήκος τουλάχιστον k αν και μόνο αν σε κάποια θέση i ξεκινάει μια ακολουθία από «1» μήκους k . Συνεπώς,

$$\Pr(L_n > 2 \log_2 n) = \Pr(L_n > k) = \Pr\left(\bigcup_{i=1}^{n-k+1} E_i\right).$$

Εφαρμόζοντας τώρα το φράγμα ένωσης του Λήμματος 5.2,

$$\Pr(L_n > 2 \log_2 n) \leq \sum_{i=1}^{n-k+1} \Pr(E_i),$$

όπου η πιθανότητα του κάθε ενδεχομένου E_i είναι,

$$\Pr(E_i) = \Pr(X_i = 1, X_{i+1} = 1, \dots, X_{i+k-1} = 1) = \left(\frac{1}{2}\right)^k,$$

άρα,

$$\Pr(L_n > 2 \log_2 n) \leq (n - k + 1)2^{-k}.$$

Τέλος, χρησιμοποιώντας το προφανές φράγμα $(n - k + 1) \leq n$, αντικαθιστώντας την τιμή του k και απλοποιώντας την πιο πάνω έκφραση, βρίσκουμε,

$$\Pr(L_n > 2 \log_2 n) \leq n2^{-2 \log_2 n} = n2^{\log_2(n^{-2})} = \frac{1}{n},$$

το οποίο είναι ακριβώς το ζητούμενο φράγμα (8.6).³

8.4 Γρήγορη ταξινόμηση δεδομένων

Το πρόβλημα ταξινόμησης είναι ένα από τα πιο θεμελιώδη προβλήματα της πληροφορικής. Δεδομένης μιας λίστας $\Lambda = [x_1, x_2, \dots, x_n]$, που περιέχει n στοιχεία x_i , το ζητούμενο είναι, με τον πιο αποτελεσματικό και γρήγορο τρόπο, να ταξινομηθεί η λίστα δηλαδή να αναδιαταχθούν τα στοιχεία της με βάση κάποιο συγκεκριμένο κριτήριο. Για παράδειγμα, αν τα x_i είναι πραγματικοί αριθμοί μπορεί να τους διατάξουμε από τον μικρότερο προς τον μεγαλύτερο, ή, αν είναι λέξεις ενός κειμένου, μπορούμε να τις ταξινομήσουμε αλφαβητικά. Η μορφή των δεδομένων και το συγκεκριμένο κριτήριο δεν έχουν σημασία, αρκεί να καθορίζεται σαφώς, για κάθε ζευγάρι στοιχείων x και y , ποιο είναι «μεγαλύτερο» και ποιο «μικρότερο», δηλαδή ποιο από τα δύο πρέπει να τοποθετηθεί νωρίτερα στη νέα λίστα. Για να αποφύγουμε τεχνικές λεπτομέρειες που δεν παρουσιάζουν κανένα ενδιαφέρον, από εδώ και στο εξής θα θεωρούμε πάντα πως τα στοιχεία x_1, x_2, \dots, x_n της λίστας Λ είναι όλα διαφορετικά μεταξύ τους.

Μια μέθοδος θεωρούμε πως είναι «αποτελεσματική και γρήγορη» όπως είπαμε πιο πάνω, αν το πλήθος από ζευγάρια στοιχείων x_i, x_j τα οποία απαιτεί να συγκριθούν είναι σχετικά μικρό. Για παράδειγμα, η προφανής, απλοϊκή μέθοδος ταξινόμησης είναι να συγκρίνουμε το κάθε x_i με όλα τα υπόλοιπα. Αυτό απαιτεί $n(n - 1)$ συγκρίσεις και άρα έχει πολυπλοκότητα της τάξης του n^2 .

Εδώ θα παρουσιάσουμε τον αλγόριθμο quicksort (ή «γρήγορη ταξινόμηση») και θα αναλύσουμε τη συμπεριφορά του. Αν και έχει και αυτός πολυπλοκότητα $O(n^2)$ στη χειρότερη περίπτωση, όπως θα δούμε αυτή η περίπτωση είναι μάλλον σπάνια. Κατά μέσο όρο, ο quicksort απαιτεί μόνο $O(n \log n)$ συγκρίσεις για να ταξινομήσει n στοιχεία, και στην πράξη έχει παρατηρηθεί πως συχνά είναι γρηγορότερος από άλλους αλγόριθμους με πολυπλοκότητα $O(n \log n)$. Για αυτόν το λόγο έχει υιοθετηθεί ευρύτατα ως η κύρια μέθοδος ταξινόμησης σε πολλά συστήματα υπολογιστών (π.χ. το λειτουργικό σύστημα Unix), και στις βιβλιοθήκες πολλών γλωσσών προγραμματισμού, όπως η C/C++ και η Java, στις οποίες είναι υλοποιημένος στη συνάρτηση qsort.

³Όπως και στην απόδειξη του Θεωρήματος 8.1, στον πιο πάνω υπολογισμό έχουμε εμμέσως υποθέσει ότι το $k = 2 \log_2 n$ είναι ακέραιος. Άσκηση. Αν το $2 \log_2 n$ δεν είναι ακέραιος αριθμός, δείξτε πως μπορούμε να επαναλάβουμε τα ίδια ακριβώς βήματα για $k = \lceil 2 \log_2 n \rceil$ και να καταλήξουμε στο ίδιο ακριβώς φράγμα.

Ο αλγόριθμος quicksort (Hoare, 1960) Ο quicksort ανήκει στην κατηγορία των «divide and conquer» (δηλαδή, «διαίρει και βασίλευε») αλγορίθμων, όπου το αρχικό πρόβλημα διασπάται σε μικρότερα προβλήματα και το καθένα από αυτά λύνεται ξεχωριστά.

Δεδομένης μιας λίστας $\Lambda = [x_1, x_2, \dots, x_n]$ που περιέχει n διαφορετικά στοιχεία, επαναλαμβάνουμε τα πιο κάτω βήματα αναδρομικά:

1. Αν η λίστα Λ που εξετάζουμε είναι κενή ή περιέχει μόνο ένα στοιχείο, την επιστρέφουμε ως έχει.
2. Αν η λίστα Λ που εξετάζουμε έχει δύο ή περισσότερα στοιχεία, επιλέγουμε ένα από αυτά, έστω x , ως «οδηγό», συγκρίνουμε όλα τα υπόλοιπα στοιχεία της με τον οδηγό, και κατασκευάζουμε δύο νέες λίστες:
 - Στη λίστα Λ_1 τοποθετούμε όλα τα στοιχεία που είναι μικρότερα του x .
 - Στη λίστα Λ_2 τοποθετούμε όλα τα στοιχεία που είναι μεγαλύτερα του x .
3. Για καθεμία από τις δύο λίστες Λ_1 και Λ_2 επαναλαμβάνουμε τα βήματα 1 και 2.
4. Επιστρέφουμε τη νέα λίστα με τη διάταξη Λ_1, x, Λ_2 .

Παρατηρήσεις

1. Η περιγραφή του αλγορίθμου αφήνει ανοιχτό το θέμα του πώς επιλέγουμε ένα στοιχείο της κάθε λίστας ως οδηγό. Μια και θέλουμε να σχεδιάσουμε έναν αλγόριθμο ο οποίος να έχει καλή συμπεριφορά κάτω από γενικές συνθήκες, δεν μπορούμε να θεωρήσουμε πως έχουμε κάποια εκ των προτέρων γνώση για τη δομή και τη φύση των δεδομένων. Άρα, εφόσον δεν μπορούμε να προϋποθέσουμε τίποτα για την αρχική διάταξη των x_i , λογικά φαίνεται αδιάφορο το πώς θα επιλέξουμε τους οδηγούς. Γι' αυτόν το λόγο, και για λόγους απλότητας στην περιγραφή και την υλοποίηση του αλγορίθμου, υιοθετούμε τη σύμβαση ότι, για κάθε λίστα, ως οδηγός θα επιλέγεται πάντα το πρώτο στοιχείο της.

Μια εξαίρεση σε αυτό το πρωτόκολλο αποτελεί η περίπτωση της randomized μορφής του quicksort. Στο Θεώρημα 8.3 στο τέλος αυτής της ενότητας θα δούμε πως, και στην περίπτωση που επιλέξουμε τους οδηγούς τυχαία, η συμπεριφορά του αλγορίθμου δεν αλλάζει ουσιαστικά.

2. Για να είναι απολύτως πλήρης η περιγραφή του αλγορίθμου υιοθετούμε και την επιπλέον σύμβαση ότι, μετά την επιλογή του οδηγού, η σχετική θέση που έχουν τα υπόλοιπα στοιχεία στις δύο νέες λίστες είναι η ίδια με την αρχική σχετική τους θέση. Έτσι, αν, για παράδειγμα, εκτελέσουμε ένα βήμα του quicksort στην λίστα $\Lambda = [2, 5, 1, 4, 3]$ και επιλέξουμε το πρώτο στοιχείο, δηλαδή το 2, ως οδηγό, τότε το αποτέλεσμα είναι η τριάδα: $[1], 2, [5, 4, 3]$.

3. Η εκτέλεση του αλγορίθμου σε μια απλή περίπτωση περιγράφεται αναλυτικά στο Παράδειγμα 8.1. Όπως αναφέραμε και στην αρχή, στη χειρότερη περίπτωση η πολυπλοκότητα του quicksort είναι της τάξης των n^2 συγκρίσεων· στο Παράδειγμα 8.2 θα δούμε ένα απλό σενάριο στο οποίο πράγματι συμβαίνει αυτό. Αλλά το βασικό μας αποτέλεσμα, όπως προκύπτει από τη λεπτομερή πιθανοκρατική ανάλυση του Θεωρήματος 8.2, είναι το εξής: Για τυχαία δεδομένα, το μέσο πλήθος συγκρίσεων που απαιτεί η εκτέλεση του quicksort είναι κατά προσέγγιση $2n \log n$, το οποίο είναι σημαντικά μικρότερο του n^2 για μεγάλα n .

Παράδειγμα 8.1 Έστω πως ξεκινάμε με τη λίστα,

$$\Lambda = [5, 3, 2, 4, 1, 6, 7].$$

Στο πρώτο βήμα, επιλέγοντας ως οδηγό το 5, δημιουργούμε τις δύο νέες λίστες $\Lambda_1 = [3, 2, 4, 1]$ και $\Lambda_2 = [6, 7]$ και καταλήγουμε στο αποτέλεσμα:

$$[3, 2, 4, 1], 5, [6, 7].$$

Στο επόμενο βήμα επαναλαμβάνουμε την ίδια διαδικασία, μία φορά για την καθεμία από τις δύο παραπάνω λίστες. Επιλέγοντας ως οδηγούς το 3 και το 6, αντίστοιχα, έχουμε το αποτέλεσμα:

$$[2, 1], 3, [4], 5, 6, [7].$$

Τέλος, παρατηρούμε ότι, από τις τέσσερις λίστες που προέκυψαν, μόνο η πρώτη περιέχει περισσότερα από ένα στοιχεία. Επιλέγοντας για αυτήν το 2 ως οδηγό, φτάνουμε στο τελικό αποτέλεσμα της ταξινομημένης λίστας:

$$[1], 2, 3, [4], 5, 6, [7].$$

Παράδειγμα 8.2 (χειρότερη περίπτωση) Έστω τώρα πως η αρχική μας λίστα είναι η $\Lambda = [n, n - 1, \dots, 2, 1]$. Επιλέγοντας το στοιχείο n ως οδηγό, μετά από $(n - 1)$ συγκρίσεις η πρώτη νέα λίστα που δημιουργούμε είναι η $\Lambda_1 = [n - 1, n - 2, \dots, 2, 1]$ και η δεύτερη είναι κενή. Έτσι καταλήγουμε στο αποτέλεσμα $[n - 1, n - 2, \dots, 2, 1], n$. Επαναλαμβάνοντας την ίδια διαδικασία, μετά από $(n - 2)$ συγκρίσεις έχουμε το αποτέλεσμα $[n - 2, n - 3, \dots, 2, 1], n - 1, n$, και συνεχίζοντας με αυτό τον τρόπο, μετά από $(n - 1)$ βήματα καταλήγουμε στην πλήρως ταξινομημένη λίστα $1, 2, \dots, n$. Συνολικά, το πλήθος των συγκρίσεων που πραγματοποιήσαμε είναι,

$$\sum_{k=1}^{n-1} (n - k) = \sum_{k=1}^{n-1} k = \frac{(n - 1)n}{2},$$

το οποίο είναι της τάξης του n^2 , και όπου χρησιμοποιήσαμε και πάλι τον τύπο για το άθροισμα μιας αριθμητικής προόδου όπως στην Ενότητα 8.1.

Θεώρημα 8.2 (Πιθανοκρατική ανάλυση του quicksort) Έστω πως ο quicksort εκτελείται σε μια λίστα n διαφορετικών στοιχείων $\Lambda = [x_1, x_2, \dots, x_n]$, όπου ως οδηγό επιλέγουμε πάντα το πρώτο στοιχείο της κάθε λίστας.

Αν τα δεδομένα είναι τυχαία, δηλαδή αν η αρχική διάταξη των x_i είναι επιλεγμένη ανάμεσα σε όλες τις $n!$ δυνατές διατάξεις τους τυχαία και ομοιόμορφα (δηλαδή καθεμία από τις $n!$ διατάξεις έχει πιθανότητα $1/n!$), τότε η μέση τιμή του πλήθους συγκρίσεων που απαιτεί η εκτέλεση του quicksort ισούται με:

$$2n \log n + O(n) \text{ συγκρίσεις.}$$

Απόδειξη: Έστω $M = [y_1, y_2, \dots, y_n]$ η ταξινομημένη μορφή της λίστας Λ . Για κάθε i, j τέτοια ώστε $1 \leq i < j \leq n$, ορίζουμε μια Τ.Μ. Z_{ij} ως:

$$Z_{ij} = \begin{cases} 1, & \text{αν το } y_i \text{ και το } y_j \text{ κάποια στιγμή συγκρίνονται,} \\ 0, & \text{αν όχι.} \end{cases}$$

Η κάθε Z_{ij} είναι μια Τ.Μ. με κατανομή $\text{Bern}(p_{ij})$, της οποίας η παράμετρος $p_{ij} = \Pr(Z_{ij} = 1)$ είναι ίση με την πιθανότητα το y_i και το y_j να συγκριθούν κάποια στιγμή κατά τη διάρκεια της εκτέλεσης του αλγορίθμου. Η σημαντική παρατήρηση εδώ είναι πως αυτό συμβαίνει *αν και μόνο αν*, κατά την εκτέλεση του quicksort, το πρώτο στοιχείο μεταξύ των $\{y_i, y_{i+1}, \dots, y_j\}$ το οποίο επιλέγεται ως οδηγός είναι το y_i ή το y_j .

Γιατί ισχύει αυτή η ισοδυναμία; Κατ' αρχάς παρατηρούμε δύο πράγματα. Πρώτον ότι, εφόσον πάντα καταλήγουμε σε λίστες με ένα ή κανένα στοιχείο, αποκλείεται καθ' όλη τη διάρκεια εκτέλεσης του αλγορίθμου να μην επιλεγεί κανένα από τα $\{y_i, y_{i+1}, \dots, y_j\}$ ως οδηγός. Δεύτερον, ότι τα στοιχεία μιας οποιασδήποτε λίστας μόνο με οδηγούς συγκρίνονται. Οπότε, αν κάποιο στοιχείο $y \in \{y_i, y_{i+1}, \dots, y_j\}$, διαφορετικό από τα y_i και y_j , επιλεγεί πρώτο ως οδηγός, τότε τα y_i και y_j σίγουρα θα καταλήξουν σε διαφορετικές υπο-λίστες και ποτέ δεν θα συγκριθούν μεταξύ τους. Τέλος, αν το πρώτο στοιχείο μεταξύ των $\{y_i, y_{i+1}, \dots, y_j\}$ που επιλεγεί ως οδηγός είναι, π.χ., το y_i , τότε σίγουρα ως εκείνη την ώρα τα y_i και y_j θα έχουν παραμείνει στην ίδια υπο-λίστα, και αναγκαστικά το y_j θα συγκριθεί με τον οδηγό y_i . Παρομοίως και για την περίπτωση που το πρώτο στοιχείο μεταξύ των $\{y_i, y_{i+1}, \dots, y_j\}$ που επιλεγεί ως οδηγός είναι το y_j .

Επιπλέον, λόγω της συμμετρίας του προβλήματος, αφού δηλαδή η αρχική διάταξη των δεδομένων είναι εντελώς τυχαία, η πιθανότητα να επιλεγεί οποιοδήποτε από τα στοιχεία $\{y_i, y_{i+1}, \dots, y_j\}$ πρώτο ως οδηγός, είναι η ίδια. Συνεπώς,

$$\begin{aligned} p_{ij} &= \Pr(\text{«πρώτο από τα } \{y_i, y_{i+1}, \dots, y_j\} \text{ θα επιλεγεί ως οδηγός το } y_i \text{ ή το } y_j\text{»}) \\ &= \frac{2}{j - i + 1}, \end{aligned}$$

και τώρα είμαστε έτοιμοι να υπολογίσουμε τη ζητούμενη ποσότητα.

Έστω Σ_n η Τ.Μ. που περιγράφει το συνολικό πλήθος συγκρίσεων που απαιτεί η εκτέλεση του quicksort. Τότε:

$$E(\Sigma_n) = E\left(\sum_{1 \leq i < j \leq n} Z_{ij}\right) = \sum_{1 \leq i < j \leq n} E(Z_{ij}) = \sum_{i=1}^{n-1} \sum_{j=i+1}^n p_{ij},$$

όπου χρησιμοποιήσαμε την πρώτη ιδιότητα της μέσης τιμής από το Θεώρημα 6.1, και το γεγονός ότι η μέση τιμή μιας $\text{Bern}(p)$ Τ.Μ. ισούται με p . Συνεπώς,

$$\begin{aligned} E(\Sigma_n) &= \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{2}{j-i+1} \\ &\stackrel{(a)}{=} \sum_{i=1}^{n-1} \sum_{k=2}^{n-i+1} \frac{2}{k} \\ &\stackrel{(b)}{=} \sum_{k=2}^n \frac{2}{k} (n+1-k) \\ &= \sum_{k=2}^n \frac{2(n+1)}{k} - \sum_{k=2}^n 2, \end{aligned}$$

όπου στο βήμα (a) κάναμε την αλλαγή μεταβλητής του πρώτου αθροίσματος από i σε $k = j - i + 1$, και στο βήμα (b) παρατηρήσαμε πως οι όροι του διπλού αθροίσματος δεν εξαρτώνται απ' το i , και πως κάθε όρος $2/k$ στο διπλό άθροισμα εμφανίζεται $(n+1-k)$ φορές. Τώρα, ορίζοντας ως S_n το άθροισμα των n πρώτων όρων της αρμονικής σειράς όπως στο Λήμμα 8.1 πιο κάτω, έχουμε,

$$E(\Sigma_n) = 2(n+1) \sum_{k=2}^n \frac{1}{k} - 2(n-1) = 2(n+1)(S_n - 1) - 2(n-1) = 2(n+1)S_n - 4n + 4,$$

και χρησιμοποιώντας το Λήμμα 8.1 παίρνουμε το ζητούμενο αποτέλεσμα ότι, πράγματι, το μέσο πλήθος συγκρίσεων $E(\Sigma_n) = 2n \log n + O(n)$. ■

Λήμμα 8.1 Έστω S_n το άθροισμα των n πρώτων όρων της αρμονικής σειράς,

$$S_n = \sum_{k=1}^n \frac{1}{k}.$$

Για κάθε $n \geq 1$, η τιμή του S_n είναι πολύ κοντά στην τιμή του $\log n$:

$$\log(n+1) \leq S_n \leq \log n + 1.$$

Συνεπώς, για μεγάλα n , έχουμε $S_n = \log n + O(1)$.

Απόδειξη: Η περίπτωση $n = 1$ είναι τετριμμένη, αφού,

$$\log(1 + 1) = \log(2) = 0.6931 \dots \leq 1 = 1/1 = S_1 = 0 + 1 = \log 1 + 1,$$

οπότε θεωρούμε πως $n \geq 2$. Χρησιμοποιώντας τον συνήθη συμβολισμό $\lceil x \rceil$ (αντίστοιχα $\lfloor x \rfloor$) για τον μικρότερο ακέραιο ο οποίος είναι $\geq x$ (αντίστοιχα, για τον μεγαλύτερο ακέραιο ο οποίος είναι $\leq x$), έχουμε,

$$\log n = \int_1^n \frac{1}{x} dx \leq \int_1^n \frac{1}{\lfloor x \rfloor} dx = \sum_{k=1}^{n-1} \int_k^{k+1} \frac{1}{\lfloor x \rfloor} dx = \sum_{k=1}^{n-1} \frac{1}{k} = S_{n-1}.$$

Συνεπώς, $S_{n-1} \geq \log n$ ή, ισοδύναμα, $S_n \geq \log(n+1)$. Παρομοίως, για το άνω φράγμα βρίσκουμε,

$$\log n = \int_1^n \frac{1}{x} dx \geq \int_1^n \frac{1}{\lceil x \rceil} dx = \sum_{k=2}^n \int_{k-1}^k \frac{1}{\lceil x \rceil} dx = \sum_{k=2}^n \frac{1}{k} = S_n - 1,$$

το οποίο μας δίνει το ζητούμενο $S_n \leq \log n + 1$. ■

Κλείνοντας, αναφέρουμε πως υπάρχει και μια σχετικά προφανής randomized μορφή του quicksort. Όπως περιγράψαμε στο τελευταίο μας αποτέλεσμα (το οποίο θα αποδείξετε στην Άσκηση 12), κατά μέσο όρο ο randomized quicksort έχει και αυτός πολυπλοκότητα της τάξης των $2n \log n + O(n)$ συγκρίσεων.

Θεώρημα 8.3 (Ανάλυση του randomized quicksort) Έστω πως ο quicksort εκτελείται σε μια αυθαίρετη λίστα n διαφορετικών στοιχείων $\Lambda = [x_1, x_2, \dots, x_n]$, όπου σε κάθε υπο-λίστα ως οδηγό επιλέγουμε τυχαία και ομοιόμορφα ένα από τα στοιχεία της.

Τότε, και πάλι η μέση τιμή του πλήθους συγκρίσεων που απαιτεί η εκτέλεση του quicksort ισούται με:

$$2n \log n + O(n) \text{ συγκρίσεις.}$$

8.5 Ασκήσεις

1. Ένας άλλος *randomized* αλγόριθμος. Έστω πως έχουμε αποθηκευμένες σε έναν πίνακα τις τιμές κάποιας συνάρτησης $F : \{0, 1, \dots, n-1\} \rightarrow \{1, 2, \dots, m\}$, αλλά ξέρουμε πως το $1/5$ των τιμών που περιέχει ο πίνακας είναι λάθος. Επίσης όμως ξέρουμε ότι η $F(x)$ έχει την εξής ιδιότητα: $F(x+y \bmod n) = F(x)F(y)$, για κάθε x και y . Περιγράψτε έναν απλό *randomized* αλγόριθμο ο οποίος να υπολογίζει την τιμή $F(x)$ και να μπορεί να μας εγγυηθεί πως η πιθανότητα η τιμή που υπολογίζει να είναι σωστή, είναι τουλάχιστον 50%. Ο αλγόριθμός σας πρέπει να λειτουργεί με αυτό τον τρόπο για κάθε x , και χωρίς να ξέρουμε ποιες από τις τιμές του πίνακα είναι λάθος.
2. *Άθροισματα*. Έστω ένας άρτιος ακέραιος $n \geq 2$. Το άθροισμα των ακεραίων από το 1 έως το n μπορεί να γραφτεί ως,

$$1 + 2 + \dots + \left(\frac{n}{2} - 1\right) + \frac{n}{2} \\ + n + (n-1) + \dots + \left(\frac{n}{2} + 2\right) + \left(\frac{n}{2} + 1\right),$$

και παρατηρούμε (όπως παρατήρησε και ο Gauss στην τρίτη δημοτικού) ότι το άθροισμα κάθε στήλης είναι $(n+1)$. Εφόσον υπάρχουν $n/2$ στήλες, το συνολικό άθροισμα $\sum_{k=1}^n k$ πρέπει λοιπόν να ισούται με $(n/2)(n+1)$.

- (α') Δείξτε το ίδιο αποτέλεσμα για περιττά n .
 (β') Δείξτε πως για κάθε $n \geq 1$ και $1 \leq m \leq n-1$,

$$\sum_{k=m+1}^n (2k) = (n+m+1)(n-m).$$

3. *Μέτρηση πράξεων*. Στην Ενότητα 8.1 είδαμε τρεις μεθόδους επίλυσης του προβλήματος επαλήθευσης ισότητας πολυωνύμων, εκ των οποίων η τρίτη μέθοδος ήταν ένας απλός *randomized* αλγόριθμος. Μετρήσαμε και για τις τρεις μεθόδους το πλήθος των πολλαπλασιασμών που απαιτούνταν. Εδώ μας ενδιαφέρει το συνολικό πλήθος των «πράξεων» που απαιτεί η καθεμία από τις τρεις μεθόδους, όπου αγνοούμε τις προσθέσεις όπως πριν, κάθε πολλαπλασιασμός μετράει για μία πράξη, και για να υψώσουμε έναν αριθμό a σε κάποιον εκθέτη k χρειάζονται $(\log k)$ πράξεις. Για καθεμία από τις τρεις μεθόδους, υπολογίστε το πλήθος των πράξεων που χρειάζονται, όταν ο βαθμός n των πολυωνύμων είναι μεγάλος. Συγκρίνετε τις τρεις μεθόδους μεταξύ τους.

Σημειώσεις. (1.) Μπορείτε να χρησιμοποιήσετε τον τύπο του Stirling, όπως στο Λήμμα 7.1: Για μεγάλα N , έχουμε την προσέγγιση, $N! \approx C N^N \sqrt{N} e^{-N}$, για κάποια σταθερά $C > 0$. (2.) Στην πράξη, η πρόσθεση είναι πράγματι σημαντικά πιο απλή από τον πολλαπλασιασμό, και ο πιο συνηθισμένος αλγόριθμος για τον υπολογισμό της ύψωσης στη δύναμη k πράγματι απαιτεί κατά προσέγγιση $(\log k)$ πολλαπλασιασμούς.

4. *Μικρότερη πιθανότητα σφάλματος.* Στην Ενότητα 8.1 δείξαμε ότι η πιθανότητα σφάλματος ενός απλού randomized αλγορίθμου για το πρόβλημα επαλήθευσης ισότητας δύο πολυωνύμων $f(x)$ και $g(x)$, είναι το πολύ 0.01. Έστω τώρα ότι αυτό το ποσοστό δεν μας είναι ικανοποιητικό και θέλουμε να μειώσουμε την πιθανότητα σφάλματος σε κάποιο δεδομένο $p < 0.01$. Περιγράψτε δύο απλούς τρόπους με τους οποίους μπορούμε να τροποποιήσουμε τον αλγόριθμο, οι οποίοι θα έχουν πιθανότητα σφάλματος το πολύ p .
5. *Πόσα min-cuts υπάρχουν;* Έστω ότι για κάποιον συγκεκριμένο (μη κατευθυνόμενο, συνδεδεμένο) γράφο με n κόμβους, υπάρχουν ακριβώς M min-cuts, τα C_1, C_2, \dots, C_M . Στην απόδειξη του Θεωρήματος 8.1 είδαμε πως, για κάθε ένα από αυτά, η πιθανότητα του ενδεχομένου $S(j) = \{\text{το αποτέλεσμα του αλγορίθμου είναι το } C_j\}$ είναι τουλάχιστον $\Pr(S(j)) \geq \frac{2}{n(n-1)}$, για κάθε $j = 1, 2, \dots, M$.
- Χρησιμοποιώντας αυτό το αποτέλεσμα, αποδείξτε πως το πλήθος M των min-cuts δεν μπορεί να είναι μεγαλύτερο από $\binom{n}{2}$.
6. *★ Τυχαία cut sets.* Εδώ θα απαντήσετε στην ερώτηση που θέσαμε στην Παρατήρηση 5 της Ενότητας 8.2.

Δεδομένου ενός (μη κατευθυνόμενου, συνδεδεμένου) γράφου (V, E) με $n = \#V$ κόμβους, επιλέγουμε ένα τυχαίο υποσύνολο $C \subset E$ των ακμών του ως εξής. Κατ' αρχάς επιλέγουμε ένα τυχαίο σύνολο κόμβων $K_1 \subset V$, όπου κάθε κόμβος i συμπεριλαμβάνεται στο K_1 με πιθανότητα $1/2$, ανεξαρτήτως των υπόλοιπων κόμβων (αν καταλήξουμε με το K_1 να είναι το κενό σύνολο ή ολόκληρο το V , το απορρίπτουμε και επαναλαμβάνουμε τη διαδικασία). Δεδομένου του K_1 , ορίζουμε το σύνολο $C \subset E$ ως το σύνολο που περιέχει όλες τις ακμές $e \in E$ οι οποίες συνδέουν κάποιον κόμβο του K_1 με κάποιον κόμβο εκτός του K_1 .

- (α) Εξηγήστε γιατί το σύνολο C είναι πάντα cut set.
- (β) Αποδείξτε πως το C είναι ομοιόμορφα κατανομημένο, δηλαδή πως η πιθανότητα το C να ισούται με οποιοδήποτε cut set είναι πάντα η ίδια.
- (γ) Αποδείξτε πως η πιθανότητα $P_{s'}$ το τυχαίο cut set C να είναι min-cut είναι πάντα εκθετικά μικρή, και συγκεκριμένα ότι,

$$P_{s'} \leq \frac{n(n-1)}{2(2^{n-1}-1)}.$$

[Υπόδειξη. Ίσως σας φανούν χρήσιμα το αποτέλεσμα της Άσκησης 5, και η παρατήρηση πριν από την περιγραφή του αλγορίθμου του Karger πως το πλήθος όλων των cut sets του γράφου ισούται με Γ_n .]

Άρα, πράγματι αυτή η μέθοδος είναι πολύ λιγότερο αποτελεσματική από τον αλγόριθμο του Karger, μια που η πιθανότητα $P_{s'}$ να επιλέξουμε ένα min-cut είναι εκθετικά μικρή, ενώ η πιθανότητα P_s του Θεωρήματος 8.1 είναι τουλάχιστον της τάξης του $1/n^2$.

7. *Φράγμα τομής.* Εδώ θα δείξετε πως, αν καθένα από τα ενδεχόμενα A_1, A_2, \dots, A_n έχει πιθανότητα κοντά στο 100% να συμβεί, τότε και η πιθανότητα του να συμβούν όλα μαζί είναι κοντά στο 100%. Χρησιμοποιήστε το φράγμα ένωσης του Κεφαλαίου 5 για να αποδείξετε πως:

$$\Pr\left(\bigcap_{i=1}^n A_i\right) \geq 1 + \sum_{i=1}^n \Pr(A_i) - n.$$

Άρα, αν $\Pr(A_i) \approx 1$ για όλα τα i , και η πιθανότητα της τομής τους θα είναι ≈ 1 .

8. * *Μια γενικότερη πιθανοκρατική ανάλυση.* Στην Ενότητα 8.3 αναλύσαμε το κόστος εκτέλεσης L_n ενός αλγορίθμου που δρα σε μια ακολουθία n τυχαίων bits, X_1, X_2, \dots, X_n , όπου το L_n ισούται με το μήκος της μακρύτερης ακολουθίας συνεχόμενων «1» ανάμεσα στα X_i . Υποθέτοντας ότι τα X_i είναι ανεξάρτητες Bern(1/2) τυχαίες μεταβλητές, αποδείξαμε ότι η πιθανότητα το L_n να είναι μεγαλύτερο από $2 \log_2 n$ τείνει στο μηδέν.

Υποθέτοντας τώρα ότι τα X_i είναι n ανεξάρτητες Bern(p) τυχαίες μεταβλητές (για κάποιο p μεταξύ 0 και 1), αποδείξτε πως, για οποιοδήποτε $\epsilon > 0$, η πιθανότητα το L_n να είναι μεγαλύτερο από,

$$k = \left(\frac{1 + \epsilon}{\log_2(1/p)}\right) \log_2 n,$$

τείνει στο μηδέν καθώς το $n \rightarrow \infty$. Δικαιολογήστε όλα τα βήματα στον υπολογισμό σας. Πώς μεταβάλλεται ο συντελεστής του $(\log_2 n)$ για διαφορετικές τιμές του p ; Σχολιάστε.

9. *Προσομοίωση.* Στην προηγούμενη άσκηση δείξατε ότι, για κάθε $\epsilon > 0$,

$$\frac{L_n}{\log_2 n} \leq \frac{1 + \epsilon}{\log_2(1/p)}, \quad \text{με μεγάλη πιθανότητα,}$$

για μεγάλα n . Εδώ θα εξετάσουμε μέσω προσομοίωσης αν αυτό το φράγμα είναι ακριβές καθώς το $n \rightarrow \infty$, δηλαδή αν $\frac{L_n}{\log_2 n} \rightarrow \frac{1}{\log_2(1/p)}$. Χρησιμοποιήστε το scilab, το matlab ή όποιο άλλο προγραμματιστικό περιβάλλον σάς βολεύει για τα παρακάτω ερωτήματα.

- (α') Επιλέξτε ένα αυθαίρετο p στο διάστημα (0.4, 0.6). Αυτή η τιμή του p θα μείνει σταθερή σε όλα τα υποερωτήματα.
- (β') Γράψτε ένα πρόγραμμα που να προσομοιώνει τις τιμές $M = 30000$ ανεξάρτητων Bern(p) τυχαίων μεταβλητών X_1, X_2, \dots, X_M , και να υπολογίζει την τιμή του L_n για $n = 2000, 4000, \dots, 30000$.
- (γ') Κάντε ένα γράφημα που να παρουσιάζει την τιμή του $L_n/(\log_2 n)$ ως συνάρτηση του n για τις πιο πάνω τιμές του n . Καθώς το n μεγαλώνει, παρατηρείτε να συγκλίνει η τιμή του $\frac{L_n}{\log_2 n}$ στην τιμή $\frac{1}{\log_2(1/p)}$ ή όχι;

10. ★ *Μεγάλοι χρόνοι αναμονής.* Έστω μια ακολουθία τυχαίων bits X_1, X_2, \dots δηλαδή τα X_i είναι ανεξάρτητες $\text{Bern}(1/2)$ Τ.Μ. Για κάθε $N \geq 1$ ορίζουμε μια νέα Τ.Μ., την W_N , η οποία περιγράφει την πρώτη χρονική στιγμή κατά την οποία εμφανίζεται το μοτίβο $0 \dots 01 \dots 1$ μήκους $2N$ (δηλαδή πρώτα N «0» και μετά N «1») στα τυχαία bits. Συγκεκριμένα, το W_N είναι η πρώτη θέση από την οποία ξεκινά και εμφανίζεται αυτό το μοτίβο. Π.χ. αν $N = 3$ και,

$$X_1, X_2, \dots = 00100100111110 \underbrace{000111}_{\text{μοτίβο}} 0100101100 \dots,$$

τότε $W_N = 14$. Αποδείξτε πως, για μεγάλα N , ο χρόνος αναμονής W_N θα είναι εκθετικά μεγάλος με πιθανότητα πολύ κοντά στο 1. Συγκεκριμένα, δείξτε ότι,

$$\Pr(W_N > 2^N) \geq 1 - 2^{-N}.$$

11. *Τυχαίες διατάξεις.* Υπάρχουν $n!$ διαφορετικές διατάξεις των στοιχείων του συνόλου $\{1, 2, \dots, n\}$. Έστω $\{Z_1, Z_2, \dots, Z_n\}$ μια τυχαία διάταξη, επιλεγμένη ομοιόμορφα από όλες τις $n!$ δυνατές διατάξεις, δηλαδή, με όλες τις δυνατές $n!$ επιλογές να έχουν την ίδια πιθανότητα.

- (α') Εξηγήστε γιατί, όταν ταξινομούμε τη λίστα $\{Z_1, Z_2, \dots, Z_n\}$, το κάθε στοιχείο Z_i μετακινείται κατά $|Z_i - i|$ θέσεις από την αρχική του θέση μέχρι τη θέση του στην ταξινομημένη λίστα.
- (β') Εξηγήστε γιατί η κάθε Τ.Μ. Z_i έχει ομοιόμορφη κατανομή στο $S = \{1, 2, \dots, n\}$, δηλαδή έχει σύνολο τιμών το S και πυκνότητα $P(j) = \Pr(Z_i = j) = 1/n$, για κάθε i, j .
- (γ') Υπολογίστε τη μέση τιμή της συνολικής απόστασης που τα στοιχεία της λίστας μετακινούνται μέχρι την τελική τους ταξινόμηση, δηλαδή την:

$$E \left(\sum_{i=1}^n |Z_i - i| \right).$$

Υπόδειξη. Ίσως χρειαστείτε τον τύπο για το άθροισμα μιας αριθμητικής προόδου όπως στην Άσκηση 2 παραπάνω, και τον αντίστοιχο τύπο για το άθροισμα διαδοχικών τετραγώνων:

$$\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}. \quad (8.7)$$

12. *Ο randomized quicksort.* Εξηγήστε γιατί η απόδειξη του Θεωρήματος 8.3 είναι μια πολύ απλή τροποποίηση της απόδειξης του Θεωρήματος 8.2.

Κεφάλαιο 9

Ανισότητες, από κοινού κατανομή, Νόμος των Μεγάλων Αριθμών

9.1 Ανισότητες Markov και Chebychev

Ξεκινάμε αυτό το κεφάλαιο με δύο σημαντικά αποτελέσματα τα οποία, πέραν της μεγάλης χρησιμότητάς τους, δίνουν και μια πιο σαφή χρηστική σημασία στην έννοια της μέσης τιμής και της διασποράς.

Το πρώτο λέει πως, αν μια τυχαία μεταβλητή έχει μικρή μέση τιμή, τότε δεν μπορεί να παίρνει μεγάλες τιμές με μεγάλη πιθανότητα:

Θεώρημα 9.1 (Ανισότητα του Markov) Έστω μια διακριτή Τ.Μ. X που παίρνει πάντα τιμές μεγαλύτερες ή ίσες του μηδενός και έχει μέση τιμή $\mu = E(X)$. Τότε:

$$\Pr(X \geq c) \leq \frac{\mu}{c}, \quad \text{για οποιαδήποτε σταθερά } c > 0.$$

Απόδειξη: Έστω S το σύνολο τιμών της X . Ξεκινώντας από τον ορισμό της μέσης τιμής, έχουμε,

$$\mu = \sum_{x \in S} x P(x) = \sum_{x \in S: x \geq c} x P(x) + \sum_{x \in S: x < c} x P(x),$$

όπου χωρίσαμε το άθροισμα σε δύο μέρη, αυτό που αντιστοιχεί σε τιμές $x \in S$ μικρότερες του c , και τις τιμές $x \geq c$. Εφόσον όλες οι τιμές $x \in S$ της X είναι μεγαλύτερες ή ίσες του μηδενός, το παραπάνω τελευταίο άθροισμα είναι κι αυτό μεγαλύτερο ή ίσο του μηδενός, συνεπώς,

$$\mu \geq \sum_{x \in S: x \geq c} x P(x) \geq \sum_{x \in S: x \geq c} c P(x) = c \sum_{x \in S: x \geq c} P(x),$$

όπου χρησιμοποιήσαμε το γεγονός ότι στο παραπάνω άθροισμα όλα τα x είναι μεγαλύτερα ή ίσα του c . Παρατηρώντας, τέλος, πως από την τρίτη βασική ιδιότητα της πυκνότητας (βλ. Κεφάλαιο 6), το παραπάνω τελευταίο άθροισμα ισούται με $\Pr(X \geq c)$, έχουμε,

$$\mu \geq c \Pr(X \geq c),$$

που είναι η ζητούμενη ανισότητα. ■

Παράδειγμα 9.1 Όπως θα δούμε πιο κάτω – ειδικά στην απόδειξη του Νόμου των Μεγάλων Αριθμών – το φράγμα που δίνει η ανισότητα Markov συχνά είναι εξαιρετικά χρήσιμο. Αλλά σε κάποιες περιπτώσεις, όπως σε αυτό το παράδειγμα, μπορεί να είναι και σχετικά ασθενές.

Έστω μια Τ.Μ. X με κατανομή Γεωμ($1/5$), οπότε η X έχει μέση τιμή $E(X) = \frac{1}{1/5} = 5$. Από την ανισότητα Markov έχουμε πως,

$$\Pr(X \geq 15) \leq 5/15 = 0.3333 \dots ,$$

ενώ στην πραγματικότητα (χρησιμοποιώντας την πρώτη ιδιότητα της γεωμετρικής κατανομής από το Θεώρημα 7.1) η πιο πάνω πιθανότητα είναι,

$$\Pr(X \geq 15) = \Pr(X > 14) = \left(1 - \frac{1}{5}\right)^{14} \approx 0.044,$$

η οποία είναι σημαντικά μικρότερη.

Αν και, πράγματι, το πιο πάνω φράγμα είναι αρκετά ασθενές, η ισχύς του έγκειται κυρίως στο γεγονός ότι δεν χρησιμοποιεί την πληροφορία πως η X έχει γεωμετρική κατανομή, και ισχύει για οποιαδήποτε τυχαία μεταβλητή X με $E(X) = 5$.

Το επόμενο αποτέλεσμα (το οποίο προκύπτει από μια σχετικά απλή εφαρμογή της ανισότητας του Markov) λέει πως, αν μια τυχαία μεταβλητή έχει μικρή διασπορά, τότε δεν μπορεί να έχει μεγάλες διακυμάνσεις, υπό την έννοια ότι δεν μπορεί να παίρνει τιμές μακριά από τη μέση τιμή της με μεγάλη πιθανότητα:

Θεώρημα 9.2 (Ανισότητα του Chebychev) Έστω μια διακριτή Τ.Μ. X με μέση τιμή $\mu = E(X)$ και διασπορά $\sigma^2 = \text{Var}(X)$. Τότε:

$$\Pr(|X - \mu| \geq c) \leq \frac{\sigma^2}{c^2}, \quad \text{για οποιαδήποτε σταθερά } c > 0.$$

Απόδειξη: Έστω μια νέα τυχαία μεταβλητή $Y = (X - \mu)^2$ η οποία, εξ ορισμού, έχει πάντα $Y \geq 0$. Από τον ορισμό της διασποράς, παρατηρούμε πως $E(Y) = E[(X - \mu)^2] = \text{Var}(X)$. Εφαρμόζοντας την ανισότητα του Markov για την Τ.Μ. Y βρίσκουμε,

$$\Pr(|X - \mu| \geq c) = \Pr((X - \mu)^2 \geq c^2) = \Pr(Y \geq c^2) \leq \frac{E(Y)}{c^2} = \frac{\sigma^2}{c^2}.$$

■

Παράδειγμα 9.2 Έστω μια Τ.Μ. X με κατανομή Υπερ(500, 150, 15). Από τις γνωστές ιδιότητες της υπεργεωμετρικής κατανομής στο Κεφάλαιο 7, η X έχει μέση τιμή και διασπορά,

$$\begin{aligned}\mu &= \frac{150 \times 15}{500} = 4.5, \\ \sigma^2 &= \frac{150 \times 15 \times (500 - 150) \times (500 - 15)}{500^2 \times (500 - 1)} \approx 3.06162.\end{aligned}$$

Έστω τώρα πως θέλουμε να υπολογίσουμε την $\Pr(|X - 4.5| \geq 3)$, δηλαδή την πιθανότητα το X να ισούται με 0 ή με 1 ή με 8 ή με 9 κ.ο.κ. Εφόσον γνωρίζουμε την πυκνότητα της X , μπορούμε να υπολογίσουμε αυτή την πιθανότητα ως $P(0) + P(1) + P(8) + P(9) + \dots + P(15)$. Αλλά λόγω της πολυπλοκότητας του τύπου της πυκνότητας $P(x)$, και επιπλέον επειδή απαιτεί τον υπολογισμό παραγοντικών $k!$ για μεγάλα k – πράγμα το οποίο, όπως παρατηρήσαμε στο Κεφάλαιο 7, μπορεί να οδηγήσει σε σημαντικά αριθμητικά σφάλματα – είναι πολύ πιο εύκολο να χρησιμοποιήσουμε την ανισότητα Chebychev για να βρούμε ένα φράγμα για τη ζητούμενη πιθανότητα:

$$\Pr(|X - 4.5| \geq 3) = \Pr(|X - \mu| \geq 3) \leq \frac{\sigma^2}{3^2} \approx 0.34018.$$

Παράδειγμα 9.3 Από εμπειρικές μετρήσεις γνωρίζουμε πως ο χρόνος απόκλισης X (σε δευτερόλεπτα) ενός συστήματος έχει μέση τιμή $\mu = E(X) = 0$ και τυπική απόκλιση $\sigma = 6$ δευτερόλεπτα. Θέλουμε να βρούμε μια τιμή θ τέτοια ώστε να μπορούμε να εγγυηθούμε ότι, με πιθανότητα τουλάχιστον 99%, η χρονική απόκλιση του συστήματος δεν θα ξεπερνά τα $\pm\theta$ δευτερόλεπτα. Δηλαδή, θέλουμε $\Pr(|X| < \theta) \geq 0.99$, ή, ισοδύναμα,

$$\Pr(|X| \geq \theta) \leq 0.01. \quad (9.1)$$

Αλλά από την ανισότητα Chebychev έχουμε πως,

$$\Pr(|X| \geq \theta) = \Pr(|X - \mu| \geq \theta) \leq \frac{\sigma^2}{\theta^2} = \frac{6^2}{\theta^2}.$$

Συνεπώς, για να ισχύει το ζητούμενο (9.1), αρκεί να διαλέξουμε το θ έτσι ώστε να έχουμε $36/\theta^2 \leq 0.01$, για το οποίο αρκεί να διαλέξουμε $\theta = 60$ δευτερόλεπτα. Με άλλα λόγια, μπορούμε να εγγυηθούμε πως, με πιθανότητα τουλάχιστον 99%, ο χρόνος απόκλισης του συστήματος δεν θα ξεπερνά το ένα λεπτό.

9.2 Από κοινού κατανομή και συνδιακύμανση

Όταν δύο Τ.Μ. X, Y με αντίστοιχες πυκνότητες $P_X(x)$ και $P_Y(y)$ είναι ανεξάρτητες, η πιθανότητα να πάρουν ένα ζευγάρι τιμών (x, y) μπορεί να υπολογιστεί απλά ως το γινόμενο,

$$\Pr(X = x, Y = y) = P_X(x)P_Y(y),$$

όπως μας λέει ο ορισμός της ανεξαρτησίας τυχαίων μεταβλητών στο Κεφάλαιο 6. Για ζευγάρια Τ.Μ. που ενδεχομένως δεν είναι ανεξάρτητες, οι αντίστοιχες πιθανότητες περιγράφονται από την από κοινού πυκνότητά τους, την οποία θα ορίσουμε αμέσως μετά το παρακάτω παράδειγμα.

Παράδειγμα 9.4 Δύο διαφορετικά προγράμματα κατανέμονται τυχαία σε τρεις υπολογιστές. Έστω X_i = «πλήθος προγραμμάτων που κατέληξαν στον υπολογιστή i », για κάθε $i = 1, 2, 3$, οπότε $X_1 + X_2 + X_3 = 2$. Προφανώς αυτές οι Τ.Μ. δεν είναι ανεξάρτητες – διαισθητικά τουλάχιστον: Αν ξέρουμε, π.χ., πως $X_1 = 2$, τότε και τα δύο προγράμματα κατέληξαν στον πρώτο υπολογιστή, συνεπώς $X_2 = X_3 = 0$.

Ας εξετάσουμε τώρα τις δύο μεταβλητές X_1, X_2 . Εφόσον υποθέτουμε πως η κατανομή έγινε τυχαία, το κάθε πρόγραμμα έχει, ανεξάρτητα από το άλλο, πιθανότητα $1/3$ να πάει στον κάθε έναν από τους τρεις υπολογιστές. Οπότε, για παράδειγμα η πιθανότητα να έχουμε $X_1 = 0$ και $X_2 = 0$, είναι η πιθανότητα του $X_3 = 2$, δηλαδή και τα δύο προγράμματα να πήγαν στον τρίτο υπολογιστή:

$$\Pr(X_1 = 0, X_2 = 0) = \Pr(X_3 = 2) = \frac{1}{3} \times \frac{1}{3} = \frac{1}{9}.$$

Παρομοίως, η πιθανότητα να έχουμε $X_1 = 0$ και $X_2 = 1$, είναι η πιθανότητα το πρώτο πρόγραμμα να πήγε στον υπολογιστή 2 και το δεύτερο στον 3, ή αντίστροφα, δηλαδή,

$$\Pr(X_1 = 0, X_2 = 1) = \frac{1}{3} \times \frac{1}{3} + \frac{1}{3} \times \frac{1}{3} = \frac{2}{9}.$$

Όλες οι πιθανότητες της μορφής $\Pr(X_1 = x_1, X_2 = x_2)$ μπορούν να υπολογιστούν με τον ίδιο τρόπο, όπως συνοψίζεται στον πιο κάτω πίνακα:

X_1	0	1	2
X_2			
0	1/9	2/9	1/9
1	2/9	2/9	0
2	1/9	0	0

Ορισμός 9.1 1. Η από κοινού πυκνότητα (ή διακριτή συνάρτηση από κοινού πυκνότητας πιθανότητας) δύο διακριτών Τ.Μ. X, Y με αντίστοιχα σύνολα τιμών S_X, S_Y είναι η συνάρτηση $P : S_X \times S_Y \rightarrow [0, 1]$ που ορίζεται ως,

$$P(x, y) = \Pr(X = x, Y = y), \quad \text{για κάθε } x \in S_X, y \in S_Y.$$

2. Η από κοινού πυκνότητα (ή διακριτή συνάρτηση από κοινού πυκνότητας πιθανότητας) ενός πεπερασμένου πλήθους διακριτών Τ.Μ. X_1, X_2, \dots, X_N όπου η κάθε X_i έχει σύνολο τιμών S_{X_i} είναι η συνάρτηση,

$$P(x_1, x_2, \dots, x_N) = \Pr(X_1 = x_1, X_2 = x_2, \dots, X_N = x_N),$$

για κάθε N -άδα τιμών $x_1 \in S_{X_1}, x_2 \in S_{X_2}, \dots, x_N \in S_{X_N}$.

Παράδειγμα 9.5 Για τις Τ.Μ. X_1, X_2 του Παραδείγματος 9.4, οι τιμές της από κοινού πυκνότητας $P(x_1, x_2)$ συνοψίζονται στον εκεί πίνακα. Έστω, τώρα, πως θέλουμε να υπολογίσουμε από αυτές τις τιμές την πιθανότητα $\Pr(X_1 = 1)$. Χρησιμοποιώντας τον κανόνα συνολικής πιθανότητας έχουμε,

$$\begin{aligned}\Pr(X_1 = 1) &= \Pr(X_1 = 1, X_2 = 0) + \Pr(X_1 = 1, X_2 = 1) + \Pr(X_1 = 1, X_2 = 2) \\ &= P(1, 0) + P(1, 1) + P(1, 2) \\ &= \frac{2}{9} + \frac{2}{9} + 0 = \frac{4}{9}.\end{aligned}$$

Άρα, για να βρούμε την $\Pr(X_1 = 1)$, αθροίσαμε όλες τις τιμές του πίνακα που αντιστοιχούσαν στην στήλη $X_1 = 1$:

X_1	0	1	2	
X_2				
0	1/9	2/9	1/9	4/9
1	2/9	2/9	0	4/9
2	1/9	0	0	1/9
	4/9	4/9	1/9	

Με παρόμοιο τρόπο, αθροίζοντας τις τιμές στις στήλες $X_1 = 0$ και $X_1 = 2$ βρίσκουμε τις $\Pr(X_1 = 0)$ και $\Pr(X_1 = 2)$, και αντίστοιχα, αθροίζοντας τις τιμές που βρίσκονται στην κάθε γραμμή $X_2 = x$ του πίνακα, βρίσκουμε τις πιθανότητες $\Pr(X_2 = x)$.

Λόγω του ότι οι τιμές των πιθανοτήτων $\Pr(X_i = x)$ που αφορούν μόνο μία από τις δύο Τ.Μ. μπορούν να γραφτούν, όπως πιο πάνω, στο περιθώριο του πίνακα της από κοινού πυκνότητας, οι επιμέρους πυκνότητές τους συχνά ονομάζονται *περιθώριες πυκνότητες*. Πριν δώσουμε τον ακριβή τους ορισμό, παρατηρούμε πως είναι τώρα εύκολο να διαπιστώσουμε και μαθηματικά ότι οι X_1, X_2 δεν είναι ανεξάρτητες. Με τη βοήθεια του πίνακα βλέπουμε πως:

$$\Pr(X_1 = 2) \Pr(X_2 = 2) = \frac{1}{9} \cdot \frac{1}{9} \neq 0 = \Pr(X_1 = 2, X_2 = 2).$$

Ορισμός 9.2 Η *περιθώρια πυκνότητα* της X_i είναι η πυκνότητα $P_{X_i}(x_i)$ μίας από N διακριτές τυχαίες μεταβλητές X_1, X_2, \dots, X_N , οι οποίες έχουν από κοινού πυκνότητα $P(x_1, \dots, x_n)$.

Ιδιότητες της από κοινού πυκνότητας. Έστω δύο διακριτές Τ.Μ. X, Y με σύνολα τιμών S_X, S_Y και πυκνότητες $P_X(x)$ και $P_Y(y)$, αντίστοιχα. Η από κοινού πυκνότητά τους $P(x, y)$ έχει τις εξής ιδιότητες:

- $\sum_{x \in S_X, y \in S_Y} P(x, y) = 1.$

$$2. P_X(x) = \sum_{y \in S_Y} P(x, y), \quad \text{και} \quad P_Y(y) = \sum_{x \in S_X} P(x, y).$$

3. Οι X, Y είναι ανεξάρτητες αν και μόνο αν,

$$P(x, y) = P_X(x)P_Y(y), \quad \text{για όλα τα } x \in S_X, y \in S_Y.$$

Στην Άσκηση 3 στο τέλος του κεφαλαίου σας ζητείται να διατυπώσετε και να αποδείξετε τις φυσικές γενικεύσεις των πιο πάνω ιδιοτήτων για ένα οποιοδήποτε πεπερασμένο πλήθος N τυχαίων μεταβλητών X_1, X_2, \dots, X_N .

Απόδειξη: Εφόσον για διαφορετικά ζεύγη τιμών (x, y) τα ενδεχόμενα $\{X = x, Y = y\}$ είναι ξένα μεταξύ τους, και προφανώς η ένωσή τους καλύπτει όλο το Ω , από τον κανόνα συνολικής πιθανότητας έχουμε,

$$\begin{aligned} 1 &= \Pr(\Omega) \\ &= \Pr\left(\bigcup_{x \in S_X, y \in S_Y} \{X = x, Y = y\}\right) \\ &= \sum_{x \in S_X, y \in S_Y} \Pr(X = x, Y = y) \\ &= \sum_{x \in S_X, y \in S_Y} P(x, y), \end{aligned}$$

που αποδεικνύει την πρώτη ιδιότητα. Παρότι η δεύτερη ιδιότητα έχει ήδη αποδειχθεί στο Λήμμα 6.1 του Κεφαλαίου 6, θυμίζουμε πως, για κάθε x , τα ενδεχόμενα $\{X = x, Y = y\}$ είναι ξένα μεταξύ τους και η ένωσή τους για διαφορετικά y ισούται με το ενδεχόμενο $\{X = x\}$. Άρα, πάλι από τον κανόνα συνολικής πιθανότητας, έχουμε:

$$\begin{aligned} P_X(x) &= \Pr(X = x) \\ &= \Pr\left(\bigcup_{y \in S_Y} \{X = x, Y = y\}\right) \\ &= \sum_{y \in S_Y} \Pr(X = x, Y = y) \\ &= \sum_{y \in S_Y} P(x, y). \end{aligned}$$

Η αντίστοιχη σχέση για την $P_Y(y)$ αποδεικνύεται με τον ίδιο τρόπο. Τέλος, η τρίτη ιδιότητα είναι απλώς αναδιατυπωμένος ο ορισμός της ανεξαρτησίας διακριτών τυχαίων μεταβλητών. ■

Παράδειγμα 9.6 Μερικές φορές το πλήθος των τιμών που παίρνει μια Τ.Μ. Y είναι απαγορευτικά μεγάλο για να αναπαραστήσουμε σε πίνακα την από κοινού πυκνότητά της με μια άλλη Τ.Μ. X . Έστω, για παράδειγμα, πως η X ορίζεται ως,

$$X = \begin{cases} 2, & \text{με πιθανότητα } 1/3, \\ 3, & \text{με πιθανότητα } 1/3, \\ 4, & \text{με πιθανότητα } 1/3, \end{cases}$$

και πως, δεδομένου ότι η X παίρνει την τιμή x , η Y έχει Γεωμ($1/x$) κατανομή. Συνεπώς, τα αντίστοιχα σύνολα τιμών είναι $S_X = \{2, 3, 4\}$ και $S_Y = \mathbb{N} = \{1, 2, \dots\}$, όπου το S_Y προφανώς είναι άπειρο.

Η από κοινού πυκνότητα των X, Y μπορεί εύκολα να υπολογιστεί από τον ορισμό της δεσμευμένης πιθανότητας: Για $m \in S_X$ και $n \in S_Y$,

$$\begin{aligned} P(m, n) &= \Pr(X = m, Y = n) \\ &= \Pr(X = m) \Pr(Y = n | X = m) \\ &= \frac{1}{3} \left(1 - \frac{1}{m}\right)^{n-1} \frac{1}{m}, \end{aligned}$$

όπου χρησιμοποιήσαμε τον τύπο της πυκνότητας της Γεωμ($1/m$) κατανομής.

Η περιθώρια πυκνότητα της X μας είναι εξ ορισμού γνωστή και, από τη δεύτερη ιδιότητα της από κοινού πυκνότητας, μπορούμε εύκολα να υπολογίσουμε και τις τιμές της περιθώριας πυκνότητας της Y . Για παράδειγμα,

$$\begin{aligned} P_Y(1) &= P(2, 1) + P(3, 1) + P(4, 1) \\ &= \frac{1}{3} \left(1 - \frac{1}{2}\right)^{1-1} \frac{1}{2} + \frac{1}{3} \left(1 - \frac{1}{3}\right)^{1-1} \frac{1}{3} + \frac{1}{3} \left(1 - \frac{1}{4}\right)^{1-1} \frac{1}{4} = \frac{13}{36}. \end{aligned}$$

Από τον ορισμό της X γνωρίζουμε πως η πιθανότητα $\Pr(X = 4) = 1/3$. Αλλά δεδομένου ότι το $Y = 1$, ποια θα ήταν η πιθανότητα να έχουμε $X = 4$; Το $X = 4$ αντιστοιχεί στο να έχει το Y γεωμετρική κατανομή με παράμετρο $1/4$, δηλαδή με πιθανότητα επιτυχίας που είναι η μικρότερη από τις τρεις δυνατές περιπτώσεις ($1/2$, $1/3$ ή $1/4$). Άρα θεωρούμε ότι είναι μάλλον απίθανο να έχουμε $X = 4$ δεδομένου ότι είχαμε $Y = 1$, δηλαδή επιτυχία από το πρώτο κίολας πείραμα. Με άλλα λόγια, περιμένουμε πως η δεσμευμένη πιθανότητα $\Pr(X = 4 | Y = 1)$ θα είναι μικρότερη από την αρχική $\Pr(X = 4) = 1/3$. Πράγματι, από τον ορισμό της δεσμευμένης πιθανότητας βρίσκουμε,

$$\Pr(X = 4 | Y = 1) = \frac{P(4, 1)}{P_Y(1)} = \frac{\frac{1}{3} \left(1 - \frac{1}{4}\right)^{1-1} \frac{1}{4}}{13/36} = \frac{3}{13} < \frac{1}{3},$$

όπως αναμενόταν.

Παράδειγμα 9.7 Έστω δύο διακριτές Τ.Μ. X, Y με από κοινού πυκνότητα όπως στον πιο κάτω πίνακα:

X	0	1	2	
Y				
0	1/6	0	1/4	5/12
1	0	1/6	0	1/6
2	1/4	0	1/6	1/9
	5/12	1/6	5/12	

Όπως και στο Παράδειγμα 9.5, οι περιθώριες πυκνότητες $P_X(x)$ και $P_Y(y)$ εύκολα υπολογίζονται στον πίνακα βάσει της δεύτερης ιδιότητας της από κοινού πυκνότητας. Και όπως και νωρίτερα, διαπιστώνουμε πως και εδώ οι X, Y δεν είναι ανεξάρτητες, εφόσον:

$$P(0, 1) = 0 \neq \frac{5}{12} \cdot \frac{1}{6} = P_X(0) \cdot P_Y(1).$$

Η συνδιακύμανση μεταξύ δύο Τ.Μ. είναι μια ποσοτική έκφραση του «πόσο μη ανεξάρτητες» είναι:

Ορισμός 9.3 Η συνδιακύμανση $\text{Cov}(X, Y)$ μεταξύ δύο τυχαίων μεταβλητών X, Y ορίζεται ως:

$$\text{Cov}(X, Y) = E[(X - E(X))(Y - E(Y))].$$

Παράδειγμα 9.8 Για τις Τ.Μ. X_1, X_2 του Παραδείγματος 9.5 έχουμε,

$$E(X_1) = 0 \cdot \frac{4}{9} + 1 \cdot \frac{4}{9} + 2 \cdot \frac{1}{9} = \frac{2}{3},$$

και παρομοίως $E(X_2) = 2/3$. Από τον ορισμό της συνδιακύμανσης και την από κοινού πυκνότητα των X_1, X_2 , εύκολα υπολογίζουμε,

$$\begin{aligned} \text{Cov}(X_1, X_2) &= \frac{1}{9} \left(0 - \frac{2}{3}\right) \left(0 - \frac{2}{3}\right) + \frac{2}{9} \left(1 - \frac{2}{3}\right) \left(0 - \frac{2}{3}\right) + \frac{1}{9} \left(2 - \frac{2}{3}\right) \left(0 - \frac{2}{3}\right) \\ &\quad + \frac{2}{9} \left(0 - \frac{2}{3}\right) \left(1 - \frac{2}{3}\right) + \frac{2}{9} \left(1 - \frac{2}{3}\right) \left(1 - \frac{2}{3}\right) + \frac{1}{9} \left(0 - \frac{2}{3}\right) \left(2 - \frac{2}{3}\right) \\ &= \frac{2}{9}, \end{aligned}$$

όπου στον υπολογισμό παραλείψαμε τα τρία ζεύγη τιμών (x_1, x_2) με μηδενική πιθανότητα.

Παράδειγμα 9.9 Για μια οποιαδήποτε Τ.Μ. X , μπορούμε να υπολογίσουμε τη συνδιακύμανση της X με τον εαυτό της. Από τους ορισμούς βρίσκουμε,

$$\text{Cov}(X, X) = E[(X - E(X))^2] = \text{Var}(X) > 0,$$

άρα η συνδιακύμανση μιας οποιασδήποτε Τ.Μ. X με τον εαυτό της είναι πάντα θετική. [Εκτός, βέβαια, από την τετριμμένη περίπτωση που η X έχει μηδενική διασπορά, δηλαδή ισούται με κάποια

σταθερά με πιθανότητα 1, οπότε η διασπορά της είναι μηδέν.] Παρομοίως, η συνδιακύμανση μεταξύ της X και της $-X$ ισούται με,

$$\begin{aligned}\text{Cov}(X, -X) &= E[(X - E(X))(-X - E(-X))] \\ &= E[-(X - E(X))^2] \\ &= -\text{Var}(X),\end{aligned}$$

η οποία είναι πάντοτε αρνητική.

Παρατήρηση: Μια πρώτη διαισθητική ερμηνεία της συνδιακύμανσης είναι πως, όταν έχουμε $\text{Cov}(X, Y) > 0$, τότε οι δύο Τ.Μ. X, Y τείνουν να παίρνουν και οι δύο μεγάλες τιμές, ή και οι δύο μικρές τιμές. Αντίστοιχα, όταν $\text{Cov}(X, Y) < 0$, τότε όταν η μία Τ.Μ. παίρνει μεγάλες τιμές η άλλη τείνει να παίρνει μικρές τιμές.

Αυτό το φαινόμενο στην ακραία του μορφή φαίνεται στο Παράδειγμα 9.9 πιο πάνω, όπου είδαμε πως, για οποιαδήποτε Τ.Μ. X , η συνδιακύμανση $\text{Cov}(X, X)$ είναι πάντοτε θετική, ενώ η συνδιακύμανση $\text{Cov}(X, -X)$ είναι πάντοτε αρνητική.

Ιδιότητες της συνδιακύμανσης. Έστω δύο διακριτές Τ.Μ. X, Y .

1. Η συνδιακύμανση $\text{Cov}(X, Y)$ εναλλακτικά ισούται με:

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y). \quad (9.2)$$

2. Αν οι X, Y είναι ανεξάρτητες, τότε η συνδιακύμανσή τους ισούται με μηδέν, δηλαδή, $\text{Cov}(X, Y) = 0$.
3. Αν η συνδιακύμανσή $\text{Cov}(X, Y)$ ισούται με μηδέν, αυτό δεν συνεπάγεται απαραίτητα πως οι X, Y είναι ανεξάρτητες.
4. Πάντοτε έχουμε:

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y).$$

5. Αν οι X, Y είναι ανεξάρτητες, τότε:

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y).$$

Παρατηρήσεις

1. Την Ιδιότητα 5 την έχουμε ήδη συναντήσει στο Θεώρημα 6.1. Εδώ θα δούμε πώς προκύπτει, εναλλακτικά, ως ειδική περίπτωση της παραπάνω Ιδιότητας 4.

2. Αν και ίσως δεν συνδέεται άμεσα με την ύλη που έχουμε συναντήσει ως τώρα σε αυτό το κεφάλαιο, παραθέτουμε εδώ μία σημαντική ισοδύναμη συνθήκη για την ανεξαρτησία δύο Τ.Μ., η οποία θα μας φανεί πολύ χρήσιμη στη μετέπειτα ανάπτυξη των αποτελεσμάτων μας. Η απόδειξή της δίνεται στην Άσκηση 6 στο τέλος του κεφαλαίου.

Δύο διακριτές Τ.Μ. X, Y είναι ανεξάρτητες αν και μόνο αν, για οποιεσδήποτε συναρτήσεις $f(x)$ και $g(y)$, οι νέες Τ.Μ. $f(X)$ και $g(Y)$ είναι ανεξάρτητες.

Απόδειξη: Για την πρώτη ιδιότητα, από τον ορισμό της συνδιακύμανσης και την πρώτη ιδιότητα του Θεωρήματος 6.1, έχουμε:

$$\begin{aligned} \text{Cov}(X, Y) &= E[(X - E(X))(Y - E(Y))] \\ &= E[XY - E(X)Y - XE(Y) + E(X)E(Y)] \\ &= E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y). \end{aligned}$$

Ιδιότητα 2: Αν οι X, Y είναι ανεξάρτητες, τότε από την Ιδιότητα 3 του Θεωρήματος 6.1 έχουμε $E(XY) = E(X)E(Y)$, και αντικαθιστώντας στην (9.2) η συνδιακύμανση πράγματι ισούται με μηδέν,

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0.$$

Για την τρίτη ιδιότητα αρκεί να δώσουμε ένα παράδειγμα δύο Τ.Μ. με μηδενική συνδιακύμανση, οι οποίες δεν είναι ανεξάρτητες: Έστω δύο διακριτές Τ.Μ. X, Y με από κοινού πυκνότητα όπως στον παρακάτω πίνακα:

X	-1	0	1	
Y	-1	0	1	
-1	0	1/4	0	1/4
0	1/4	0	1/4	1/2
1	0	1/4	0	1/4
	1/4	1/2	1/4	

Οι περιθώριες πυκνότητες $P_X(x), P_Y(y)$ έχουν επίσης υπολογιστεί στον πίνακα βάσει της Ιδιότητας 2 της από κοινού πυκνότητας. Οι X, Y δεν είναι ανεξάρτητες, αφού:

$$P(0, 0) = 0 \neq \frac{1}{2} \cdot \frac{1}{2} = P_X(0) \cdot P_Y(0).$$

Για τη συνδιακύμανσή τους, πρώτα υπολογίζουμε,

$$E(X) = (-1) \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} = 0,$$

και παρομοίως $E(Y) = 0$ αφού οι X και Y έχουν την ίδια πυκνότητα. Επιπλέον παρατηρούμε πως,

$$E(XY) = \sum_{x=-1,0,1} \sum_{y=-1,0,1} P(x,y) \cdot x \cdot y = 0,$$

διότι όλοι οι όροι του πιο πάνω αθροίσματος είναι μηδενικοί. [Ή $x = 0$ ή $y = 0$ ή $P(x,y) = 0$, για όλα τα δυνατά ζεύγη τιμών (x,y) .] Άρα,

$$\text{Cov}(X,Y) = E(XY) - E(X)E(Y) = 0 - 0 \cdot 0 = 0.$$

[Παρατηρήστε πως αυτό το παράδειγμα είναι το ίδιο με εκείνο που δόθηκε για τη λύση της Άσκησης 5 του Κεφαλαίου 6.]

Για την τέταρτη ιδιότητα, ξεκινώντας από τον ορισμό της διασποράς και εφαρμόζοντας την Ιδιότητα 1 του Θεωρήματος 6.1, έχουμε,

$$\begin{aligned} \text{Var}(X+Y) &= E\left[\left((X+Y) - E(X+Y)\right)^2\right] \\ &= E\left[\left((X - E(X)) + (Y - E(Y))\right)^2\right] \\ &= E[(X - E(X))^2] + E[(Y - E(Y))^2] + 2E[(X - E(X))(Y - E(Y))] \\ &= \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X,Y). \end{aligned}$$

Τέλος, η πέμπτη ιδιότητα είναι άμεση συνέπεια του συνδυασμού των Ιδιοτήτων 4 και 2. ■

Παράδειγμα 9.10 Έστω τρεις ανεξάρτητες Τ.Μ. X_1, X_2, X_3 που έχουν όλες την ίδια πυκνότητα, και, κατά συνέπεια, την ίδια μέση τιμή $\mu = E(X_i)$. Ορίζουμε δύο νέες Τ.Μ., τις $Y = X_1 + X_2$ και $Z = X_1 + X_3$. Εφόσον και οι δύο εξαρτώνται από την X_1 περιμένουμε πως δεν θα είναι ανεξάρτητες, και πως πιθανότατα η συνδιακύμανσή τους δεν θα είναι μηδενική. Χρησιμοποιώντας την εναλλακτική έκφραση για τη συνδιακύμανση στη σχέση (9.2), έχουμε,

$$\begin{aligned} \text{Cov}(Y,Z) &= E(YZ) - E(Y)E(Z) \\ &= E((X_1 + X_2)(X_1 + X_3)) - E(X_1 + X_2)E(X_1 + X_3) \\ &= E(X_1^2 + X_1X_3 + X_1X_2 + X_2X_3) - E(X_1 + X_2)E(X_1 + X_3) \\ &= E(X_1^2) + E(X_1X_3) + E(X_1X_2) + E(X_2X_3) \\ &\quad - (E(X_1))^2 - E(X_1)E(X_3) - E(X_1)E(X_2) - E(X_2)E(X_3), \end{aligned}$$

όπου στο τελευταίο βήμα χρησιμοποιήσαμε την Ιδιότητα 1 του Θεωρήματος 6.1. Από το ίδιο θεώρημα επίσης γνωρίζουμε πως, λόγω της ανεξαρτησίας των X_i , έχουμε $E(X_iX_j) = E(X_i)E(X_j)$. Συνεπώς, οι παραπάνω έξι όροι εκτός του πρώτου και του τέταρτου απλοποιούνται, διότι είναι όλοι ίσοι με $(E(X_i))^2 = \mu^2$, και χρησιμοποιώντας την εναλλακτική έκφραση της διασποράς βρίσκουμε πως $\text{Cov}(Y,Z) = \text{Var}(X_1)$.

Άρα, η συνδιακύμανση μεταξύ Y και Z όχι μόνο είναι μη μηδενική (όπως περιμέναμε), αλλά επιπλέον ισούται ακριβώς με τη «διακύμανση» (δηλαδή τη διασπορά) της Τ.Μ. X_1 η οποία συνδέει τις Y και Z .

Παράδειγμα 9.11 Παρομοίως με το προηγούμενο παράδειγμα, έστω ότι από τρεις ανεξάρτητες τυχαίες μεταβλητές $X_1 \sim \text{Poisson}(\lambda)$, $X_2 \sim \text{Poisson}(\mu)$ και $X_3 \sim \text{Poisson}(\nu)$, ορίζουμε δύο νέες Τ.Μ., την $Y = X_1 + X_2$ και την $Z = X_2 + X_3$. Όπως είδαμε στην Άσκηση 5 του Κεφαλαίου 7, το άθροισμα δύο ανεξάρτητων Τ.Μ. με κατανομή Poisson έχει κι αυτό κατανομή Poisson, και η παράμετρος του ισούται με το άθροισμα των δύο αρχικών παραμέτρων. Άρα εδώ έχουμε $Y \sim \text{Poisson}(\lambda + \mu)$ και $Z \sim \text{Poisson}(\mu + \nu)$.

Ακολουθώντας τα ίδια βήματα όπως στο προηγούμενο παράδειγμα, υπολογίζουμε τη συνδιακύμανση μεταξύ Y και Z ,

$$\begin{aligned} \text{Cov}(Y, Z) &= E(YZ) - E(Y)E(Z) \\ &= E((X_1 + X_2)(X_2 + X_3)) - E(X_1 + X_2)E(X_2 + X_3) \\ &= E(X_1)E(X_2) + E(X_1)E(X_3) + E(X_2^2) + E(X_2)E(X_3) \\ &\quad - E(X_1)E(X_2) - E(X_1)E(X_3) - (E(X_2))^2 - E(X_2)E(X_3) \\ &= E(X_2^2) - (E(X_2))^2 \\ &= \text{Var}(X_2) = \mu. \end{aligned}$$

Παρατηρούμε πως, παρότι οι τρεις Τ.Μ. X_1, X_2, X_3 εδώ, σε αντίθεση με το προηγούμενο παράδειγμα, δεν έχουν ακριβώς την ίδια κατανομή (γιατί έχουν διαφορετικές παραμέτρους), καταλήγουμε στο ίδιο αποτέλεσμα: Η συνδιακύμανση μεταξύ Y και Z είναι ίση με τη διασπορά της Τ.Μ. X_2 η οποία συνδέει τις Y και Z .

9.3 Ο Νόμος των Μεγάλων Αριθμών

Σε αυτό το σημείο έχουμε πλέον αναπτύξει αρκετά μαθηματικά εργαλεία ώστε να είμαστε σε θέση να διατυπώσουμε και να αποδείξουμε ίσως το πιο θεμελιώδες αποτέλεσμα της θεωρίας των πιθανοτήτων, το *Νόμο των Μεγάλων Αριθμών* (ή Ν.Μ.Α., για συντομία). Χωρίς να μπορούμε ακόμα σε μαθηματικές λεπτομέρειες, ο Ν.Μ.Α. λέει το εξής:

Ν.Μ.Α.: Έστω πως έχουμε ένα μεγάλο πλήθος N από ανεξάρτητες τυχαίες μεταβλητές X_1, X_2, \dots, X_N , οι οποίες έχουν όλες την ίδια «κατανομή», δηλαδή την ίδια πυκνότητα $P(x)$ και, κατά συνέπεια, και την ίδια μέση τιμή $\mu = E(X_i)$. Τότε, με μεγάλη πιθανότητα, ο εμπειρικός μέσος όρος τους θα ισούται με τη μέση τιμή τους, δηλαδή:

$$\frac{1}{N} \sum_{i=1}^N X_i \approx \mu, \quad \text{με πιθανότητα} \approx 1.$$

Πριν δώσουμε την ακριβή μαθηματική διατύπωση του Ν.Μ.Α., ας δούμε μερικά απλά αλλά πολύ σημαντικά παραδείγματα.

Παράδειγμα 9.12 (Τι θα πει «πιθανότητα»;) Έστω πως μας ενδιαφέρει η πιθανότητα $p = \mathbb{P}(A)$ του να συμβεί ένα συγκεκριμένο ενδεχόμενο A . Λόγου χάρη, το A θα μπορούσε να αντιστοιχεί στο «η θεραπεία του ασθενή ήταν επιτυχής» ή «το δίκτυο παρουσίασε σφάλμα» ή «η εκτέλεση του αλγορίθμου ολοκληρώθηκε κανονικά» ή «ανέβηκε σήμερα στο χρηματιστήριο η τιμή της μετοχής τάδε» κλπ. Ένας τρόπος για να δώσουμε ένα φυσικό νόημα στην πιθανότητα p είναι να φανταστούμε πως επαναλαμβάνεται το ίδιο ακριβώς πείραμα πάρα πολλές, ανεξάρτητες φορές, έτσι ώστε, τουλάχιστον στο διαισθητικό επίπεδο, η πιθανότητα $p = \mathbb{P}(A)$ να αντιστοιχεί στο ποσοστό των φορών, μακροπρόθεσμα, που συμβαίνει το A . Ο Ν.Μ.Α. είναι ακριβώς το αποτέλεσμα εκείνο που τεκμηριώνει αυτόν τον συλλογισμό μαθηματικά.

Συγκεκριμένα, το πιο πάνω φανταστικό σενάριο μπορεί να περιγραφεί ως εξής. Έστω πως έχουμε ένα μεγάλο πλήθος N ανεξάρτητων επαναλήψεων του ίδιου πειράματος, και ας ορίσουμε, για την κάθε επανάληψη i , μια Τ.Μ. X_i που να ισούται με 1 αν το A συνέβη τη φορά i , αλλιώς να ισούται με 0. Τότε,

$$\text{«ποσοστό των φορών που συνέβη το } A\text{»} = \frac{1}{N} \sum_{i=1}^N X_i,$$

και ο Ν.Μ.Α. μας λέει πως, με πιθανότητα ≈ 1 , αυτό θα ισούται με τη μέση τιμή των X_i . Εφόσον τα X_i είναι Bernoulli τυχαίες μεταβλητές, η μέση τιμή $E(X_i)$ είναι απλά η πιθανότητα p του να έχουμε $X_i = 1$, δηλαδή του να συμβεί το A . Με άλλα λόγια:

Η πιθανότητα p του να συμβεί ένα οποιοδήποτε ενδεχόμενο A , ισούται με τη συχνότητα με την οποία θα συμβεί, μακροπρόθεσμα, το A , σε πολλές ανεξάρτητες επαναλήψεις του ίδιου πειράματος.

Παράδειγμα 9.13 (Γιατί γίνονται δημοσκοπήσεις;) Έστω πως σε έναν πληθυσμό K ατόμων οι $K/4$ είναι ψηφοφόροι κάποιου κόμματος, δηλαδή το κόμμα αυτό στις εκλογές θα πάρει ποσοστό 25%. Μια δημοσκόπηση επιλέγει στην τύχη ένα σχετικά μεγάλο πλήθος ατόμων N (με επανατοποθέτηση) και τα ρωτά αν θα ψηφίσουν το κόμμα αυτό. [Σε τυπικές δημοσκοπήσεις έχουμε N μεταξύ 1000 και 5000 ερωτηθέντων σε έναν πληθυσμό K μεταξύ 500 χιλιάδων και πολλών εκατομμυρίων ατόμων.]

Έστω X_i μια Τ.Μ. που ισούται με 1 αν το άτομο i που ρωτάται είναι ψηφοφόρος του κόμματος, και 0 αν όχι, οπότε οι X_i είναι ανεξάρτητες Τ.Μ. με κατανομή Bern(0.25). Εδώ ο Ν.Μ.Α. μας λέει πως, αν το μέγεθος N του δείγματος είναι αρκετά μεγάλο, τότε, με πιθανότητα ≈ 1 :

$$\text{«ποσοστό ερωτηθέντων που είναι ψηφοφόροι του κόμματος»} = \frac{1}{N} \sum_{i=1}^N X_i \approx E(X_i) = 25\%.$$

Άρα, από ένα αρκετά μεγάλο (τυχαίο) δείγμα, μπορούμε να εκτιμήσουμε το πραγματικό ποσοστό ψήφων που αυτό το κόμμα θα πάρει στις εκλογές.

Παράδειγμα 9.14 (Εκτίμηση μέσω δειγματοληψίας) Έστω πως θέλουμε να υπολογίσουμε το μέσο ύψος \bar{y} (σε εκατοστά) ενός κοριτσιού ηλικίας 18 μηνών, μεταξύ των M κοριτσιών ενός πληθυσμού αυτής της ηλικίας. Αν y_k είναι το ύψος του κοριτσιού $k = 1, 2, \dots, M$, θέλουμε να εκτιμήσουμε το,

$$\text{«μέσο ύψος»} = \bar{y} = \frac{1}{M} \sum_{i=1}^M y_k.$$

Αντί να εξετάσουμε ολόκληρο τον πληθυσμό, επιλέγουμε τυχαία, με επανατοποθέτηση, N μέλη του πληθυσμού, όπου το μέγεθος N του δείγματός μας είναι μεν σχετικά μεγάλο, αλλά είναι σημαντικά μικρότερο του συνολικού μεγέθους M του πληθυσμού (όπως και στο προηγούμενο παράδειγμα).

Ορίζουμε τις εξής Τ.Μ.: Έστω $X_j = \text{«ύψος του κοριτσιού } j\text{»}$ που επιλέξαμε. Τότε τα X_j είναι ανεξάρτητα, και το κάθε X_j έχει σύνολο τιμών $S = \{y_1, y_2, \dots, y_M\}$ και πυκνότητα,

$$P(y_k) = \Pr(X_j = y_k) = \Pr(\text{«επιλέξαμε το κορίτσι } k\text{»}) = \frac{1}{M}.$$

Άρα τα X_j έχουν μέση τιμή,

$$E(X_j) = \sum_{k=1}^M y_k \cdot P(y_k) = \sum_{k=1}^M y_k \cdot \frac{1}{M} = \frac{1}{M} \sum_{k=1}^M y_k = \bar{y}.$$

Αν εκτιμήσουμε λοιπόν το \bar{y} ως τον μέσο όρο από τα ύψη των κοριτσιών που επιλέξαμε, ο Ν.Μ.Α. είναι αυτός που μας διαβεβαιώνει πως, αν το μέγεθος N του δείγματός μας είναι αρκετά μεγάλο, τότε η εκτίμησή μας θα είναι ακριβής με πιθανότητα κοντά στο 100%:

$$\text{«μέσο ύψος δειγμάτων»} = \frac{1}{N} \sum_{i=1}^N X_i \approx E(X_i) = \bar{y}, \quad \text{με πιθανότητα } \approx 1.$$

Για να διατυπώσουμε τον Ν.Μ.Α. μαθηματικά με ακρίβεια, θα χρειαστούμε την έννοια της «σύγκλισης κατά πιθανότητα». Όταν η ακολουθία που μας ενδιαφέρει αποτελείται από τυχαίες μεταβλητές X_1, X_2, X_3, \dots , και όχι απλά από πραγματικούς αριθμούς, πρέπει να ορίσουμε εκ νέου τι εννοούμε λέγοντας ότι «η ακολουθία συγκλίνει».

Ορισμός 9.4 (Σύγκλιση κατά πιθανότητα) Έστω μια ακολουθία τυχαίων μεταβλητών $\{X_n\} = \{X_1, X_2, \dots\}$, μια σταθερά c , και μια άλλη Τ.Μ. X .

1. Η ακολουθία τυχαίων μεταβλητών $\{X_n\}$ συγκλίνει κατά πιθανότητα στη σταθερά c , αν για κάθε $\epsilon > 0$:

$$\Pr(|X_n - c| < \epsilon) \rightarrow 1, \quad \text{καθώς το } n \rightarrow \infty.$$

2. Γενικότερα, η ακολουθία τυχαίων μεταβλητών $\{X_n\}$ συγκλίνει κατά πιθανότητα στην Τ.Μ. X , αν για κάθε $\epsilon > 0$:

$$\Pr(|X_n - X| < \epsilon) \rightarrow 1, \quad \text{καθώς το } n \rightarrow \infty.$$

Θεώρημα 9.3 (Ο Νόμος των Μεγάλων Αριθμών) Έστω μια ακολουθία από ανεξάρτητες διακριτές τυχαίες μεταβλητές X_1, X_2, \dots , όπου έχουν όλες την ίδια κατανομή, δηλαδή την ίδια πυκνότητα, και κατά συνέπεια την ίδια μέση τιμή $\mu = E(X_i)$ και την ίδια διασπορά $\sigma^2 = \text{Var}(X_i) < \infty$. Τότε:

1. [Διαισθητικά] Για μεγάλα N , ο εμπειρικός μέσος όρος των X_1, X_2, \dots, X_N ,

$$\bar{X}_N = \frac{1}{N} \sum_{i=1}^N X_i \approx \mu, \quad \text{με μεγάλη πιθανότητα.}$$

2. [Μαθηματικά] Καθώς το $N \rightarrow \infty$ ο εμπειρικός μέσος όρος \bar{X}_N τείνει στη μέση τιμή μ κατά πιθανότητα, δηλαδή, για κάθε $\epsilon > 0$:

$$\Pr(|\bar{X}_N - \mu| < \epsilon) \rightarrow 1, \quad \text{καθώς το } N \rightarrow \infty.$$

Απόδειξη: Η απόδειξη είναι μια απλή εφαρμογή της ανισότητας Chebychev. Χρησιμοποιώντας την πρώτη ιδιότητα του Θεωρήματος 6.1, η μέση τιμή της Τ.Μ. \bar{X}_N είναι,

$$E(\bar{X}_N) = E\left(\frac{1}{N} \sum_{i=1}^N X_i\right) = \frac{1}{N} \sum_{i=1}^N E(X_i) = \frac{1}{N} \cdot N \cdot \mu = \mu.$$

Παρομοίως, χρησιμοποιώντας τις Ιδιότητες 2 και 4 του Θεωρήματος 6.1, και το γεγονός ότι τα X_i είναι ανεξάρτητα, η διασπορά της Τ.Μ. \bar{X}_N είναι,

$$\begin{aligned} \text{Var}(\bar{X}_N) &= \text{Var}\left(\frac{1}{N} \sum_{i=1}^N X_i\right) \\ &= \frac{1}{N^2} \text{Var}\left(\sum_{i=1}^N X_i\right) \\ &= \frac{1}{N^2} \sum_{i=1}^N \text{Var}(X_i) \\ &= \frac{1}{N^2} \cdot N \cdot \sigma^2 \\ &= \frac{\sigma^2}{N}. \end{aligned}$$

Τέλος, δεδομένου $\epsilon > 0$, από την ανισότητα Chebychev έχουμε,

$$\Pr(|\bar{X}_N - \mu| \geq \epsilon) = \Pr(|\bar{X}_N - E(\bar{X}_N)| \geq \epsilon) \leq \frac{\text{Var}(\bar{X}_N)}{\epsilon^2} = \frac{\sigma^2}{N\epsilon^2}.$$

Άρα,

$$\Pr(|\bar{X}_N - \mu| < \epsilon) = 1 - \Pr(|\bar{X}_N - \mu| \geq \epsilon) \geq 1 - \frac{\sigma^2}{N\epsilon^2},$$

το οποίο προφανώς τείνει στο 1, καθώς το $N \rightarrow \infty$. ■

Παρατηρήσεις

1. Το αποτέλεσμα του Θεωρήματος 9.3 συχνά αναφέρεται στη βιβλιογραφία ως ο «Ασθενής» Νόμος των Μεγάλων Αριθμών. Ο λόγος είναι ότι υπάρχει ένα πράγματι ισχυρότερο αποτέλεσμα, ο λεγόμενος «Ισχυρός» Νόμος των Μεγάλων Αριθμών, που μας λέει πως, κάτω από τις ίδιες συνθήκες, ο εμπειρικός μέσος όρος \bar{X}_N συγκλίνει στη μέση τιμή μ με πιθανότητα 1, όχι μόνο κατά πιθανότητα.

Αυτή η μορφή σύγκλισης λέει το εξής: Ας εξετάσουμε τις Τ.Μ. \bar{X}_N στην πιο βασική τους μορφή, δηλαδή ως συναρτήσεις $\bar{X}_N : \Omega \rightarrow \mathbb{R}$ σε κάποιο χώρο πιθανότητας Ω . Για οποιοδήποτε συγκεκριμένο $\omega \in \Omega$, η ακολουθία $\{\bar{X}_N(\omega) : N \geq 1\}$ είναι απλά μια ακολουθία πραγματικών αριθμών. Μπορούμε λοιπόν να εξετάσουμε το ενδεχόμενο,

$$\Sigma = \left\{ \omega \in \Omega : \lim_{N \rightarrow \infty} \bar{X}_N(\omega) = \mu \right\} = \left\{ \lim_{N \rightarrow \infty} \bar{X}_N = \mu \right\}.$$

Λέμε ότι τα \bar{X}_N συγκλίνουν στο μ με πιθανότητα 1, αν το ενδεχόμενο Σ έχει πιθανότητα $\Pr(\Sigma) = 1$. Δεν είναι ιδιαίτερα δύσκολο ναδειχθεί (βλ. Άσκηση 14 στο τέλος του κεφαλαίου) ότι η σύγκλιση με πιθανότητα 1 συνεπάγεται και τη σύγκλιση κατά πιθανότητα, πράγμα που εξηγεί τις σχετικές ονομασίες των δύο αυτών αποτελεσμάτων.

2. Αν και στη διατύπωση του Ν.Μ.Α. στο Θεώρημα 9.3 υποθέσαμε πως οι Τ.Μ. X_i έχουν πεπερασμένη διασπορά $\sigma^2 = \text{Var}(X_i) < \infty$, σημειώνουμε πως αυτή η συνθήκη δεν είναι απαραίτητη και, αν ήμασταν διατεθειμένοι να υποστούμε μια κάπως πιο σύνθετη και μακροσκελή απόδειξη, θα μπορούσαμε να την αντικαταστήσουμε με την ασθενέστερη υπόθεση ότι η μέση τιμή $\mu_0 = E(|X_i|)$ είναι πεπερασμένη.
3. Κλείνουμε με μια ακόμη παρατήρηση η οποία, αν και τεχνικής φύσεως, όπως θα δούμε σε επόμενα κεφάλαια είναι θεμελιώδους σημασίας για τις περισσότερες στατιστικές εφαρμογές των πιθανοτήτων.

Δύο βασικά και μάλλον προφανή ερωτήματα που γεννιούνται από τον Ν.Μ.Α. είναι τα εξής:

- (α') Πόσο μεγάλο πρέπει να είναι το πλήθος N των δειγμάτων, ώστε να έχουμε κάποια σχετική βεβαιότητα πως ο εμπειρικός μέσος όρος \bar{X}_N θα είναι αρκετά κοντά στη μέση τιμή μ ;

(β') Αντίστοιχα, δεδομένου του πλήθους N , πόσο μικρή είναι η πιθανότητα το \bar{X}_N να απέχει κατά πολύ από το μ ;

Αν εξετάσουμε προσεκτικά την παραπάνω απόδειξη θα δούμε πως, πέρα από το ασυμπτωτικό αποτέλεσμα του θεωρήματος (το οποίο ισχύει μόνο καθώς το $N \rightarrow \infty$), περιέχει και κάποιες πρώτες απαντήσεις στα πιο πάνω ερωτήματα. Συγκεκριμένα, δείξαμε πως η μέση τιμή και η διασπορά του \bar{X}_N είναι, αντίστοιχα,

$$E(\bar{X}_N) = \mu \quad \text{και} \quad \text{Var}(\bar{X}_N) = \frac{\sigma^2}{N}, \quad \text{για κάθε } N.$$

Επιπλέον βρήκαμε ένα ακριβές ποσοτικό φράγμα για την πιθανότητα ο εμπειρικός μέσος όρος \bar{X}_N να απέχει από τη μέση τιμή μ κατά τουλάχιστον ϵ :

$$\Pr(|\bar{X}_N - \mu| \geq \epsilon) \leq \frac{\sigma^2}{N\epsilon^2}. \quad (9.3)$$

Αν και μαθηματικά σωστό, το πιο πάνω φράγμα στις περισσότερες περιπτώσεις δεν είναι αρκετά ακριβές ώστε να είναι χρήσιμο στην πράξη. Με άλλα λόγια, η ζητούμενη πιθανότητα στη σχέση (9.3) είναι συνήθως σημαντικά μικρότερη από το φράγμα $\sigma^2/(N\epsilon^2)$.

Το δεύτερο θεμελιώδες αποτέλεσμα των πιθανοτήτων, το *Κεντρικό Οριακό Θεώρημα* (ή Κ.Ο.Θ.), το οποίο αποτελεί το αντικείμενο του Κεφαλαίου 12, είναι αυτό που μας δίνει μια ακριβή προσέγγιση της πιθανότητας $\Pr(|\bar{X}_N - \mu| \geq \epsilon)$. Συγκεκριμένα, το Κ.Ο.Θ. λέει πως, κάτω από ορισμένες συνθήκες, η κατανομή του εμπειρικού μέσου όρου \bar{X}_N μπορεί να προσεγγιστεί με μεγάλη ακρίβεια από την *κανονική κατανομή*.

Η κανονική (ή γκαουσιανή) κατανομή αντιστοιχεί σε μια κατηγορία τυχαίων μεταβλητών οι οποίες έχουν ως σύνολο τιμών ολόκληρο το \mathbb{R} . Συνεπώς, πριν μπορέσουμε να διατυπώσουμε με ακρίβεια το Κ.Ο.Θ., θα πρέπει να μελετήσουμε τις *συνεχείς τυχαίες μεταβλητές*. Αυτό είναι το αντικείμενο του Κεφαλαίου 10 που ακολουθεί.

9.4 Ασκήσεις

1. *Αφρικανικό χελιδόσι.* Ένας βιολόγος ισχυρίζεται το εξής: «Το μέσο βάρος του αφρικανικού χελιδονιού είναι 100 γραμμάρια, ενώ το 60% των αφρικανικών χελιδονιών έχει βάρος άνω των 200 γραμμαρίων». Είναι δυνατόν αυτός ο ισχυρισμός να ευσταθεί; Εξηγήστε.
2. *Απόσταση Τ.Μ. από τη μέση τιμή της.* Για μια τυχαία μεταβλητή X με μέση τιμή μ και άγνωστη κατανομή, κάποιος στατιστικολόγος ισχυρίζεται ότι η πιθανότητα η τιμή της X να απέχει από τη μέση τιμή της κατά παραπάνω από 3μ είναι μικρή. Υποστηρίζει αυτό το συμπέρασμα διότι γνωρίζει πως, στη συγκεκριμένη περίπτωση, η διασπορά της X είναι αρκετά μικρότερη από τη μέση τιμή στο τετράγωνο (δηλαδή από το μ^2). Έχει δίκιο στον ισχυρισμό του ή όχι; Αποδείξτε την απάντησή σας.
3. *Ιδιότητες από κοινού πυκνότητας.* Αν, αντί για δύο διακριτές Τ.Μ. X, Y , είχαμε ένα πεπερασμένο πλήθος X_1, X_2, \dots, X_N , διατυπώστε και αποδείξτε τις φυσικές γενικεύσεις των τριών ιδιοτήτων της από κοινού πυκνότητας που ακολουθούν τον Ορισμό 9.2.
4. *Χρόνος εκτέλεσης αλγορίθμου.* Ένας αλγόριθμος έχει ως δεδομένα εισόδου έναν ακέραιο αριθμό Y και μια σειρά από n bits, X_1, X_2, \dots, X_n . Ο χρόνος εκτέλεσης του αλγορίθμου είναι:

$$T = Y + 2 \sum_{i=1}^n X_i + \left(\sum_{i=1}^n X_i \right)^2 \quad \text{δευτερόλεπτα.}$$

Αν υποθέσουμε ότι τα Y, X_1, X_2, \dots, X_n είναι ανεξάρτητες τυχαίες μεταβλητές όπου τα $X_i \sim \text{Bern}(1/4)$ και το Y παίρνει τις τιμές 0, 2, 5 και 9 με πιθανότητα $1/4$ για την καθεμία:

- (α) Να υπολογίσετε τον μέσο χρόνο εκτέλεσης.
- (β) Να δείξετε ότι η πιθανότητα ο χρόνος εκτέλεσης να ξεπερνά τα n^2 δευτερόλεπτα είναι το πολύ,

$$\frac{1}{16} + \frac{11}{16n} + \frac{4}{n^2}.$$

[Συνεπώς, για μεγάλα n αυτή η πιθανότητα είναι το πολύ $\approx \frac{1}{16} = 6.25\%$.]

- (γ) Υπολογίστε τη συνδιακύμανση μεταξύ των Τ.Μ. T και Y . Πώς εξαρτάται το αποτέλεσμα από το n ; Εξηγήστε.

5. *Μέγιστο και ελάχιστο δύο ζαριών.* Ρίχνουμε ένα ζάρι δύο φορές. Έστω X_1 και X_2 τα αποτελέσματα των δύο ρίψεων.
 - (α) Ποια είναι η από κοινού πυκνότητα του ζεύγους X_1, X_2 , υποθέτοντας ότι οι ρίψεις είναι ανεξάρτητες και το ζάρι είναι δίκαιο;
 - (β) Έστω $X = \min(X_1, X_2)$, $Y = \max(X_1, X_2)$. Βρείτε την από κοινού πυκνότητα των X, Y .

(γ') Βρείτε τις περιθώριες πυκνότητες των X, Y . Συγκρίνετε όλα τα πιο πάνω αποτελέσματα με αυτά της Άσκησης 7 του Κεφαλαίου 6.

6. *Συναρτήσεις τυχαίων μεταβλητών.* Έστω X, Y δύο διακριτές τυχαίες μεταβλητές με σύνολο τιμών $\mathbb{Z} = \{\dots, -1, 0, 1, 2, \dots\}$.

(α') Δείξτε πως, αν τα X, Y είναι ανεξάρτητα, τότε και τα X^2 και Y^2 είναι ανεξάρτητα.

(β') Βρείτε μια από κοινού πυκνότητα για την οποία τα X, Y δεν είναι ανεξάρτητα, αλλά τα X^2, Y^2 είναι.

(γ') Γενικεύοντας το πρώτο ερώτημα, εδώ θα αποδείξετε μια σημαντική παρατήρηση που κάναμε στην Ενότητα 9.2: Έστω οποιεσδήποτε δύο διακριτές Τ.Μ. X, Y με σύνολα τιμών S_X, S_Y . Δείξτε πως οι X, Y είναι ανεξάρτητες αν και μόνο αν, για οποιεσδήποτε συναρτήσεις $f : S_X \rightarrow S'_X$ και $g : S_Y \rightarrow S'_Y$, οι $f(X), g(Y)$ είναι ανεξάρτητες.

Γιατί αυτό το αποτέλεσμα δεν αντικρούει το προηγούμενο ερώτημα;

[Υπόδειξη. Ίσως σας φανεί χρήσιμο να δείτε τη λύση της Άσκησης 1 του Κεφαλαίου 6.]

7. * *Το φράγμα του Chernoff.* Έστω πως, για μια διακριτή Τ.Μ. X , θέλουμε να βρούμε ένα ακριβέστερο φράγμα από αυτό της ανισότητας Markov για την πιθανότητα η X να παίρνει τιμές μεγαλύτερες από κάποια σταθερά $c \in \mathbb{R}$.

(α') Αποδείξτε πως, για οποιοδήποτε $\lambda > 0$, έχουμε το φράγμα:

$$\Pr(X \geq c) \leq e^{-\lambda c} E(e^{\lambda X}). \quad (9.4)$$

Παρατηρήστε πως ενώ το αριστερό μέρος δεν εξαρτάται από το λ , το δεξί μέρος ισχύει για κάθε θετικό λ .

(β') Έστω τώρα πως η X έχει κατανομή Διων($n, 1/2$). Χρησιμοποιώντας τη σχέση (9.4), δείξτε πως, για οποιοδήποτε $c \geq 0$, έχουμε:

$$\Pr(X \geq nc) \leq \left\{ \frac{1}{2} e^{-\lambda c} [1 + e^\lambda] \right\}^n. \quad (9.5)$$

(γ') Έστω πάλι πως $X \sim \text{Διων}(n, 1/2)$. Για $n = 10$ και $c = 0.9$:

- i. Υπολογίστε την τιμή του λ που μας δίνει το καλύτερο (δηλαδή το μικρότερο) δυνατό φράγμα στη σχέση (9.5), καθώς και το αντίστοιχο φράγμα γι' αυτήν τη βέλτιστη τιμή του λ .
- ii. Υπολογίστε το αντίστοιχο φράγμα που μας δίνει η ανισότητα του Markov. Παρατηρήστε ότι είναι ασθενέστερο.
- iii. Υπολογίστε ακριβώς την πιθανότητα $\Pr(X \geq nc)$ και συγκρίνετε την πραγματική τιμή της με τα δύο παραπάνω φράγματα.

8. Συμπλήρωση πίνακα. Οι διακριτές τυχαίες μεταβλητές X και Y έχουν από κοινού κατανομή που δίνεται από τον παρακάτω πίνακα:

X	-5	-2	2	5	
Y					$P_Y(y)$
0	0.01	0.01	0.01		0.19
1	0.05		0.05	0	
2	0.07	0.07		0.07	
3	0.15		0	0.15	
$P_X(x)$		0.26	0.15		

- (α') Συμπληρώστε τις τιμές που λείπουν.
 (β') Είναι οι X, Y ανεξάρτητες ή όχι; Αποδείξτε την απάντησή σας.
 (γ') Υπολογίστε την $\Pr(Y = 1|X = 2)$.
9. *Ενεργοποιημένες συνδέσεις.* Το πλήθος των ενεργοποιημένων συνδέσεων σε ένα δίκτυο είναι X , όπου η X έχει κάποια άγνωστη κατανομή με μέση τιμή 2000 και τυπική απόκλιση 500. Αν κάποια στιγμή το πλήθος των συνδέσεων αποκλίνει από το μέσο πλήθος κατά περισσότερο από $\pm s$, τότε το δίκτυο «πέφτει». Βρείτε μια τιμή για το s ώστε το ενδεχόμενο να πέσει το δίκτυο να έχει πιθανότητα το πολύ 1%.
10. *Συντελεστής συσχέτισης.* Ο συντελεστής συσχέτισης μεταξύ δύο τυχαίων μεταβλητών X, Y , ορίζεται ως:

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{[\text{Var}(X)\text{Var}(Y)]^{\frac{1}{2}}}$$

Να δείξετε ότι, αν $Y = aX + b$, όπου η σταθερά $a \neq 0$, τότε ο συντελεστής συσχέτισης $\rho_{X,Y}$ των X, Y , ισούται με $+1$ ή με -1 . Πότε ισούται με $+1$ και πότε με -1 ;

11. *Ασυσχέτιστες αλλά όχι ανεξάρτητες Τ.Μ.* Εδώ θα δούμε ένα παράδειγμα, διαφορετικό από εκείνο της απόδειξης της τρίτης ιδιότητας της συνδιακύμανσης, δύο τυχαίων μεταβλητών που έχουν μηδενική συνδιακύμανση αλλά δεν είναι ανεξάρτητες.
 Έστω X, Y , ανεξάρτητες τυχαίες μεταβλητές Bernoulli με παράμετρο $p = \frac{1}{2}$. Να δείξετε ότι οι τυχαίες μεταβλητές $V = X + Y$ και $W = |X - Y|$ δεν είναι ανεξάρτητες, αλλά έχουν μηδενική συνδιακύμανση.

12. *Σύγκλιση κατά πιθανότητα.*

(α') Έστω μια ακολουθία $\{X_n\} = \{X_1, X_2, \dots\}$ ανεξάρτητων Τ.Μ. όπου η κάθε X_n έχει κατανομή $\text{Bern}(1/n)$. Δείξτε πως οι $\{X_n\}$ συγκλίνουν στο μηδέν κατά πιθανότητα, καθώς το $n \rightarrow \infty$.

- (β') Εξηγήστε γιατί το αποτέλεσμα του παραπάνω ερωτήματος εξακολουθεί να ισχύει ακόμα και αν οι Τ.Μ. $\{X_n\}$ δεν είναι ανεξάρτητες.
- (γ') Μπορούμε με το ίδιο σκεπτικό να συμπεράνουμε ότι και ο Ν.Μ.Α. ισχύει αν οι Τ.Μ. $\{X_n\}$ δεν είναι ανεξάρτητες; Αν ναι, αποδείξτε το. Αν όχι, εξηγήστε γιατί.

13. Εκτίμηση με θόρυβο.

- (α') Έστω πως θέλουμε να εκτιμήσουμε τη μέση τιμή $\mu = E(X)$ κάποιας Τ.Μ. X , χρησιμοποιώντας τις τιμές N ανεξάρτητων Τ.Μ. X_1, X_2, \dots, X_N όλες με την ίδια κατανομή με την X . Αλλά αντί για τα X_i τα ίδια, το μόνο που έχουμε διαθέσιμο είναι τα δείγματα $Y_i = X_i + Z_i$, για $i = 1, 2, \dots, N$, όπου το κάθε Y_i είναι μια τυχαία μέτρηση του αντίστοιχου X_i συν κάποιο τυχαίο «θόρυβο» Z_i . Υποθέτουμε ότι τα Z_i είναι ανεξάρτητα από τα X_i , και επίσης ότι είναι ανεξάρτητα μεταξύ τους και όλα έχουν την ίδια κατανομή, με κάποια γνωστή μέση τιμή $\nu = E(Z_i)$.

Περιγράψτε μια μέθοδο εκτίμησης της μέσης τιμής μ , και αποδείξτε ότι συγκλίνει στο ζητούμενο μ κατά πιθανότητα, καθώς το πλήθος των δειγμάτων $N \rightarrow \infty$.

- (β') Έστω τώρα πως θέλουμε να εκτιμήσουμε όχι κάποια σταθερά όπως πιο πάνω, αλλά την τυχαία τιμή κάποιας Τ.Μ. X , χρησιμοποιώντας τις τιμές N δειγμάτων $Y_i' = X + Z_i$, για $i = 1, 2, \dots, N$, όπου τα Z_i είναι ανεξάρτητα από το X , είναι ανεξάρτητα μεταξύ τους και όλα έχουν την ίδια κατανομή, με γνωστή μέση τιμή $\nu = E(Z_i)$.

Περιγράψτε μια μέθοδο εκτίμησης της X και αποδείξτε ότι συγκλίνει στην X κατά πιθανότητα, καθώς $N \rightarrow \infty$.

14. ★ Σύγκλιση με πιθανότητα 1. Έστω μια ακολουθία Τ.Μ. $\{X_n\}$ σε κάποιο χώρο πιθανότητας Ω με αντίστοιχο μέτρο πιθανότητας \mathbb{P} . Υποθέτουμε ότι οι $\{X_n\}$ τείνουν με πιθανότητα 1 σε κάποια Τ.Μ. X η οποία είναι επίσης ορισμένη στον Ω , δηλαδή ότι το ενδεχόμενο,

$$\Sigma = \left\{ \omega \in \Omega : \lim_{n \rightarrow \infty} \bar{X}_n(\omega) = X(\omega) \right\} = \left\{ \lim_{n \rightarrow \infty} \bar{X}_n = X \right\},$$

έχει πιθανότητα $\mathbb{P}(\Sigma) = 1$. Ακολουθώντας τα πιο κάτω βήματα θα δείξουμε ότι οι $\{X_n\}$ τείνουν στη X και κατά πιθανότητα. Έστω $\epsilon > 0$ αυθαίρετο.

- (α') Για κάθε ακέραιο $m \geq 1$, ορίζουμε τα ενδεχόμενα,

$$\begin{aligned} T_m &= \{ \omega \in \Omega : |X_n(\omega) - X(\omega)| < \epsilon \text{ για κάθε } n \geq m \} \\ &= \{ |X_n - X| < \epsilon \text{ για κάθε } n \geq m \}. \end{aligned}$$

Εξηγήστε γιατί η ακολουθία των ενδεχομένων $\{T_n\}$ είναι αύξουσα, δηλαδή $T_1 \subset T_2 \subset T_3 \subset \dots$, και γιατί έχουμε,

$$\Sigma \subset \bigcup_{m=1}^{\infty} T_m. \quad (9.6)$$

- (β') Χρησιμοποιώντας το αποτέλεσμα της Άσκησης 9 του Κεφαλαίου 3, αποδείξτε πως $\mathbb{P}(T_m) \rightarrow 1$ καθώς το $m \rightarrow \infty$.
- (γ') Εξηγήστε πώς τα πιο πάνω συνεπάγονται ότι οι $\{X_n\}$ τείνουν στη X κατά πιθανότητα.

Κεφάλαιο 10

Συνεχείς τυχαίες μεταβλητές

Σε αυτό το κεφάλαιο θα εξετάσουμε τις ιδιότητες που έχουν οι συνεχείς τυχαίες μεταβλητές. Εκείνες οι Τ.Μ. X , δηλαδή, των οποίων το σύνολο τιμών δεν είναι διακριτό, όπως για παράδειγμα αν έχουμε $S_X = [0, 1]$, $S_X = \mathbb{R}$, $S_X = [-5, -1] \cup (2, 3) \cup [7, \infty)$, κλπ.¹

Δύο βασικοί λόγοι κινούν το ενδιαφέρον μας για τις συνεχείς Τ.Μ. Ο ένας είναι προφανής: Πολλές ποσότητες που είναι σημαντικές στην πράξη, είναι από τη φύση τους συνεχείς – π.χ., ο χρόνος που διαρκεί η εκτέλεση ενός αλγορίθμου, η θερμοκρασία ενός επεξεργαστή, η απόσταση μεταξύ ενός κινητού τηλεφώνου και της κοντινότερης κεραίας με την οποία επικοινωνεί κ.ο.κ.

Ο δεύτερος λόγος είναι πιο λεπτός και σχετίζεται με το Νόμο των Μεγάλων Αριθμών, που είδαμε στον προηγούμενο κεφάλαιο. Ο Ν.Μ.Α. περιγράφει τη συμπεριφορά του εμπειρικού μέσου όρου,

$$\bar{X}_N = \frac{1}{N} \sum_{i=1}^N X_i,$$

όταν οι Τ.Μ. X_1, X_2, \dots, X_N είναι ανεξάρτητες και έχουν όλες την ίδια κατανομή. Μία από τις ισοδύναμες διατυπώσεις του Θεωρήματος 9.3 είναι πως, αν το πλήθος N των X_i είναι μεγάλο, τότε η πιθανότητα ο εμπειρικός μέσος όρος τους \bar{X}_N να απέχει σημαντικά από τη μέση τιμή τους μ είναι μικρή. Συγκεκριμένα, για κάθε $\epsilon > 0$, έχουμε,

$$\Pr(|\bar{X}_N - \mu| \geq \epsilon) \rightarrow 0, \quad \text{καθώς το } N \rightarrow \infty. \quad (10.1)$$

Για να χρησιμοποιήσουμε αυτό το αποτέλεσμα στην πράξη, όπως έχουμε ήδη αναφέρει, μια προφανής βασική ερώτηση που γεννιέται είναι «πόσο μικρή είναι η πιθανότητα απόκλισης στη σχέση (10.1);». Όπως θα δούμε λεπτομερώς στο Κεφάλαιο 12, η απάντηση (την οποία μας δίνει το Κεντρικό Οριακό Θεώρημα) είναι πως, κάτω από αρκετά γενικές συνθήκες, η κατανομή

¹Ο ακριβής ορισμός του όρου «συνεχής τυχαία μεταβλητή» δίνεται στην επόμενη ενότητα. Όπως και στο Κεφάλαιο 6, έτσι και εδώ ορισμένα από τα αποτελέσματα που θα δούμε απαιτούν κάποιες επιπλέον τεχνικές υποθέσεις, οι οποίες όμως δεν επηρεάζουν την ουσία τους. Κάποιες περαιτέρω λεπτομέρειες δίνονται στην Ενότητα 10.3.

του εμπειρικού μέσου \bar{X}_N μπορεί να προσεγγιστεί μέσω της κατανομής μιας συνεχούς τυχαίας μεταβλητής κατανομής, συγκεκριμένα μιας Τ.Μ. που ακολουθεί τη λεγόμενη κανονική κατανομή.

Πριν αναπτύξουμε τη θεωρία που απαιτείται για να μελετήσουμε τις συνεχείς Τ.Μ. συστηματικά, ας δούμε ένα απλό παράδειγμα.

Παράδειγμα 10.1 Έστω πως ο χρόνος X (σε δευτερόλεπτα) που απαιτείται για την εκκίνηση της λειτουργίας ενός δικτύου είναι πάντοτε μεταξύ 10 και 20 δευτερολέπτων και, κατά τα άλλα, είναι «εντελώς τυχαίος». Για να περιγράψουμε το χρόνο X ως μια Τ.Μ., θα θέλαμε να έχει σύνολο τιμών $S_X = [10, 20]$ και η πιθανότητα του να πάρει οποιαδήποτε τιμή σε αυτό το διάστημα να είναι κατά κάποιον τρόπο ομοιόμορφη. Π.χ., θα θέλαμε η πιθανότητα το X να είναι μεταξύ 10 και 11 δευτερολέπτων να είναι η ίδια με την πιθανότητα να έχουμε $19 \leq X \leq 20$. Επιπλέον, αφού ο χρόνος X είναι «εντελώς τυχαίος», θα περιμέναμε λογικά πως θα ικανοποιεί,

$$\Pr(10 \leq X \leq 15) = \Pr(15 \leq X \leq 20) = 1/2, \quad (10.2)$$

δηλαδή πως θα είναι το ίδιο πιθανό η εκκίνηση να γίνει τα πρώτα 5 ή τα τελευταία 5 δευτερόλεπτα.

Στην επόμενη ενότητα θα περιγράψουμε τον συστηματικό τρόπο με τον οποίο μπορούν να οριστούν τέτοιου είδους Τ.Μ., και στην Ενότητα 11.1 του επόμενου κεφαλαίου θα δούμε πως η Τ.Μ. X αυτού του παραδείγματος ανήκει σε μια συνηθισμένη και χρήσιμη οικογένεια τυχαίων μεταβλητών, αυτών που έχουν ομοιόμορφη κατανομή.

10.1 Συνεχείς Τ.Μ. και συνεχής πυκνότητα

Ορισμός 10.1 Μια τυχαία μεταβλητή X είναι συνεχής με πυκνότητα $f(x)$, αν η Τ.Μ. X και η συνάρτηση $f(x)$ ικανοποιούν τις εξής ιδιότητες:

1. Η πυκνότητα $f(x)$ είναι μια συνάρτηση $f : \mathbb{R} \rightarrow [0, \infty)$ τέτοια ώστε,

$$\int_{-\infty}^{+\infty} f(x) dx = 1.$$

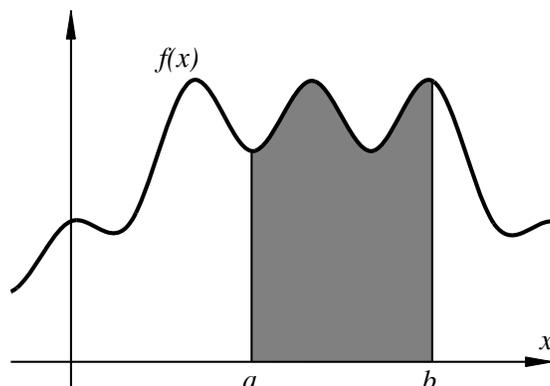
2. Το σύνολο τιμών S_X της X μπορεί να εκφραστεί ως ένωση ενός πεπερασμένου πλήθους (μη τετριμμένων) διαστημάτων πραγματικών αριθμών, και επιπλέον το S_X αποτελείται από εκείνα τα $x \in \mathbb{R}$ για τα οποία η πυκνότητα $f(x)$ δεν είναι μηδενική, δηλαδή:

$$S_X = \{x \in \mathbb{R} : f(x) > 0\}.$$

3. Για οποιαδήποτε $a \leq b$, η πιθανότητα η Τ.Μ. X να πάρει κάποια τιμή στο διάστημα $[a, b]$ μπορεί να εκφραστεί ως προς την πυκνότητα $f(x)$ μέσω της σχέσης,

$$\Pr(a \leq X \leq b) = \int_a^b f(x) dx, \quad (10.3)$$

όπως αναπαρίσταται στο Σχήμα 10.1.



Σχήμα 10.1: Γραφική αναπαράσταση του υπολογισμού της πιθανότητας $\Pr(a \leq X \leq b)$ για μια συνεχή τυχαία μεταβλητή X μέσω της πυκνότητάς $f(x)$. Παρατηρούμε πως η πιθανότητα $\Pr(a \leq X \leq b)$ είναι ίση με το εμβαδόν μεταξύ του καμπύλης $y = f(x)$ και του άξονα x , ανάμεσα στα σημεία a και b .

Παρατηρήσεις

1. Ο λόγος για τον οποίο στον ορισμό απαιτούμε τα διαστήματα που απαρτίζουν το σύνολο τιμών της X να είναι μη τετριμμένα, δηλαδή να μην είναι της μορφής $[a, a]$, είναι διότι τέτοια διαστήματα αποτελούνται από μόνο ένα στοιχείο, $[a, a] = \{a\}$. Αν η X είχε, για παράδειγμα, σύνολο τιμών το,

$$S_X = [0, 0] \cup [1, 1] \cup [2, 2] = \{0, 1, 2\},$$

τότε προφανώς θα ήταν διακριτή και όχι συνεχής.

2. Θυμίζουμε πως, από τον Ορισμό 6.3 του Κεφαλαίου 6, η συνάρτηση κατανομής $F(x)$ μιας οποιασδήποτε τυχαίας μεταβλητής X είναι η $F(x) = \Pr(X \leq x)$, για $x \in \mathbb{R}$. Αν, τώρα, η X είναι συνεχής με πυκνότητα $f(x)$, τότε η συνάρτηση κατανομής $F(x)$ υπολογίζεται εύκολα ως,

$$F(x) = \Pr(X \leq x) = \int_{-\infty}^x f(y) dy, \quad x \in \mathbb{R}, \quad (10.4)$$

βάσει της δεύτερης βασικής ιδιότητας που διατυπώνεται αμέσως μετά, στη σχέση (10.6). Περαιτέρω ιδιότητες της συνάρτησης κατανομής μιας συνεχούς Τ.Μ. δίνονται μετά το Παράδειγμα 10.2.

3. Αν και η τιμή $f(x)$ της πυκνότητας μιας συνεχούς Τ.Μ. X δεν αντιστοιχεί ακριβώς σε κάποια πιθανότητα, παρατηρούμε ότι η X είναι πιο πιθανό να πάρει τιμές κοντά σε κάποιο

x_0 όπου η τιμή της $f(x)$ είναι σχετικά μεγάλη. Για παράδειγμα, αν η $f(x)$ είναι συνεχής στο x_0 , τότε για μικρά δ κατά προσέγγιση έχουμε,

$$\Pr(x_0 - \delta/2 \leq X \leq x_0 + \delta/2) = \int_{x_0 - \delta/2}^{x_0 + \delta/2} f(x) dx \approx \delta f(x_0).$$

Άρα είναι μεγαλύτερη η πιθανότητα το X να πάρει τιμές κοντά στο x_0 αν η τιμή της πυκνότητας $f(x_0)$ είναι μεγάλη, ενώ είναι λιγότερο πιθανό να πάρει τιμές κοντά σε κάποιο x_0 όπου η $f(x_0)$ είναι κοντά στο μηδέν.

Συνεχείς Τ.Μ.: Βασικές ιδιότητες. Για μια οποιαδήποτε συνεχή Τ.Μ. X με πυκνότητα $f(x)$ έχουμε:

1. Για κάθε τιμή $a \in \mathbb{R}$, η πιθανότητα η X να ισούται ακριβώς με το a είναι μηδενική:

$$\Pr(X = a) = 0, \quad \text{για οποιοδήποτε } a \in \mathbb{R}.$$

2. Για κάθε $a < b$,

$$\Pr(a \leq X \leq b) = \Pr(a < X \leq b) = \Pr(a \leq X < b) = \Pr(a < X < b),$$

και όλες οι πιο πάνω πιθανότητες είναι ίσες με $\int_a^b f(x) dx$.

3. Η σχέση (10.3) ισχύει ακόμη και στην περίπτωση που το a ή το b ή και τα δύο παίρνουν άπειρες τιμές. Δηλαδή, για κάθε a και b ,

$$\Pr(X \geq a) = \Pr(a \leq X < \infty) = \int_a^{\infty} f(x) dx \quad (10.5)$$

$$\Pr(X \leq b) = \Pr(-\infty < X \leq b) = \int_{-\infty}^b f(x) dx, \quad (10.6)$$

και, προφανώς,

$$\Pr(-\infty < X < \infty) = \int_{-\infty}^{\infty} f(x) dx = 1.$$

Απόδειξη: Η πρώτη ιδιότητα είναι άμεση συνέπεια του Ορισμού 10.1: Εφόσον το ολοκλήρωμα οποιασδήποτε συνάρτησης από το a ως το a είναι ίσο με μηδέν (σχηματικά, το αντίστοιχο εμβαδόν είναι προφανώς μηδενικό), έχουμε:

$$\Pr(X = a) = \Pr(a \leq X \leq a) = \int_a^a f(x) dx = 0.$$

Για το πρώτο σκέλος της δεύτερης ιδιότητας, εξετάζοντας τα ξένα ενδεχόμενα $\{X = a\}$ και $\{a < X \leq b\}$, βρίσκουμε πως,

$$\Pr(a \leq X \leq b) = \Pr(\{X = a\} \cup \{a < X \leq b\}) = \Pr(X = a) + \Pr(a < X \leq b),$$

και χρησιμοποιώντας την πρώτη ιδιότητα έχουμε, $\Pr(a \leq X \leq b) = \Pr(a < X \leq b)$. Παρομοίως αποδεικνύονται και οι ισότητες των άλλων δύο περιπτώσεων.

* Για την τρίτη ιδιότητα θα χρησιμοποιήσουμε το αποτέλεσμα της Άσκησης 9 του Κεφαλαίου 3, όπου αποδεικνύεται η συνέχεια του μέτρου πιθανότητας. Ορίζουμε την ακολουθία ενδεχομένων,

$$A_n = \{a \leq X \leq a + n\}, \quad \text{για κάθε } n \geq 1,$$

και παρατηρούμε πως $A_n \subset A_{n+1}$ και πως το όριό τους,

$$\lim_{n \rightarrow \infty} A_n = \bigcup_{n=1}^{\infty} A_n = \{a \leq X < \infty\} = \{X \geq a\}.$$

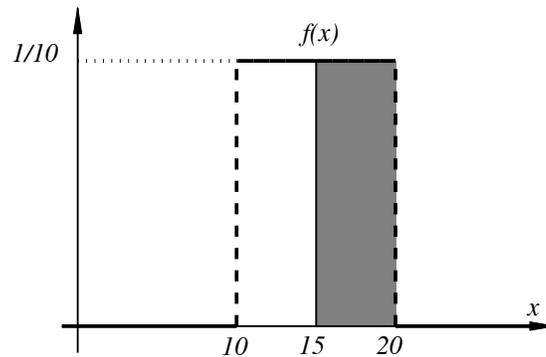
Συνεπώς, η συνέχεια του μέτρου πιθανότητας συνεπάγεται το ότι η πιθανότητα,

$$\begin{aligned} \Pr(X \geq a) &= \Pr\left(\lim_{n \rightarrow \infty} A_n\right) \\ &= \lim_{n \rightarrow \infty} \Pr(A_n) \\ &= \lim_{n \rightarrow \infty} \Pr(a \leq X \leq a + n) \\ &= \lim_{n \rightarrow \infty} \int_a^{a+n} f(x) dx \\ &= \int_a^{\infty} f(x) dx, \end{aligned}$$

που αποδεικνύει τη σχέση (10.5). Παρομοίως αποδεικνύονται και τα αποτελέσματα των άλλων δύο περιπτώσεων. ■

Παράδειγμα 10.2 Συνεχίζοντας το Παράδειγμα 10.1, θα ορίσουμε την πυκνότητα της Τ.Μ. X που εξετάσαμε εκεί. Εφόσον η X έχει προφανώς σύνολο τιμών $S_X = [10, 20]$, θέτουμε την πυκνότητά της $f(x)$ ίση με μηδέν για $x \notin [10, 20]$. Επιπλέον, για να παίρνει η X «ομοιόμορφα τυχαίες τιμές» στο διάστημα αυτό, λογικά ορίζουμε την πυκνότητά $f(x)$ ως μια σταθερά $c > 0$, για $x \in [10, 20]$:

$$f(x) = \begin{cases} c, & \text{για } x \in [10, 20], \\ 0, & \text{για } x \notin [10, 20]. \end{cases}$$



Σχήμα 10.2: Γραφική αναπαράσταση της πυκνότητας $f(x)$ της Τ.Μ. X στο Παράδειγμα 10.2, και του υπολογισμού του εμβαδού το οποίο ισούται με την πιθανότητα $\Pr(15 \leq X \leq 20)$.

Για να υπολογίσουμε την τιμή της σταθεράς c παρατηρούμε πως, από την πρώτη παραπάνω ιδιότητα, το ολοκλήρωμα της $f(x)$ για όλα τα $x \in \mathbb{R}$ πρέπει να ισούται με 1, άρα,

$$1 = \int_{-\infty}^{\infty} f(x) dx = \int_{10}^{20} c dx = [cx]_{10}^{20} = 20c - 10c = 10c,$$

συνεπώς το $c = 1/10$, βλ. Σχήμα 10.2.

Για να ελέγξουμε αν αυτή η Τ.Μ. πράγματι έχει ιδιότητες που ανταποκρίνονται στις απαιτήσεις μας ως προς την ποσότητα την οποία θέλουμε να περιγράψουμε, υπολογίζουμε την πιθανότητα,

$$\Pr(15 \leq X \leq 20) = \int_{15}^{20} f(x) dx = \int_{15}^{20} \frac{1}{10} dx = \frac{20 - 15}{10} = 1/2,$$

βλ. Σχήμα 10.2, και παρομοίως βρίσκουμε πως και η πιθανότητα $\Pr(10 \leq X \leq 15)$ ισούται με $1/2$. Συνεπώς η σχέση (10.2), την οποία διατυπώσαμε διαισθητικά, επαληθεύεται και μαθηματικά.

Ιδιότητες της συνάρτησης κατανομής. Για μια οποιαδήποτε συνεχή Τ.Μ. X με πυκνότητα $f(x)$, όπως είδαμε στη σχέση (10.4), η συνάρτηση κατανομής μπορεί να εκφραστεί ως $F(x) = \Pr(X \leq x) = \int_{-\infty}^x f(y) dy$. Επιπλέον:

1. Από το θεμελιώδες θεώρημα του διαφορικού λογισμού, αμέσως προκύπτει πως ισχύει και η αντίστροφη σχέση της (10.4), δηλαδή,

$$f(x) = F'(x) = \frac{d}{dx} F(x), \quad (10.7)$$

για όλα τα x για τα οποία υπάρχει η παράγωγος $F'(x)$.

2. Όταν γνωρίζουμε τη συνάρτηση κατανομής $F(x)$, τότε όλες οι πιθανότητες της μορφής $\Pr(a \leq X \leq b)$ μπορούν να υπολογιστούν απευθείας από τη σχέση,

$$\Pr(a \leq X \leq b) = F(b) - F(a). \quad (10.8)$$

Η (10.8) είναι άμεσο επακόλουθο της (10.4) και του Ορισμού 10.1, αλλά εύκολα αποδεικνύεται και ευθέως: Εφόσον τα ενδεχόμενα $\{X \leq a\}$ και $\{a < X \leq b\}$ είναι ξένα,

$$\begin{aligned} F(b) = \Pr(X \leq b) &= \Pr(\{X \leq a\} \cup \{a < X \leq b\}) \\ &= \Pr(X \leq a) + \Pr(a < X \leq b) = F(a) + \Pr(a \leq X \leq b), \end{aligned}$$

από όπου προκύπτει η (10.8).

3. Υπενθυμίζουμε (βλ. Κεφάλαιο 6) πως η συνάρτηση κατανομής $F(x)$ μιας οποιασδήποτε (διακριτής ή συνεχούς) τυχαίας μεταβλητής ικανοποιεί:

$$\begin{aligned} \lim_{x \rightarrow -\infty} F(x) &= 0 \\ \text{και} \quad \lim_{x \rightarrow +\infty} F(x) &= 1. \end{aligned}$$

Παράδειγμα 10.3 Για την Τ.Μ. X του Παραδείγματος 10.2 μπορούμε να υπολογίσουμε τη συνάρτηση κατανομής $F(x)$ ως εξής: Κατ' αρχάς, εφόσον η X παίρνει τιμές μόνο στο διάστημα $[10, 20]$, είναι αδύνατον να έχουμε $X \leq x$ όταν το x είναι μικρότερο από 10, συνεπώς έχουμε $F(x) = \Pr(X \leq x) = 0$ για $x < 10$. Παρομοίως, αν το x είναι μεγαλύτερο του 20, τότε $F(x) = \Pr(X \leq x) = 1$. Τέλος, για $x \in [10, 20]$, βρίσκουμε,

$$F(x) = \Pr(X \leq x) = \int_{-\infty}^x f(y) dy = \int_{10}^x \frac{1}{10} dy = \frac{x-10}{10},$$

οπότε, συνοψίζοντας:

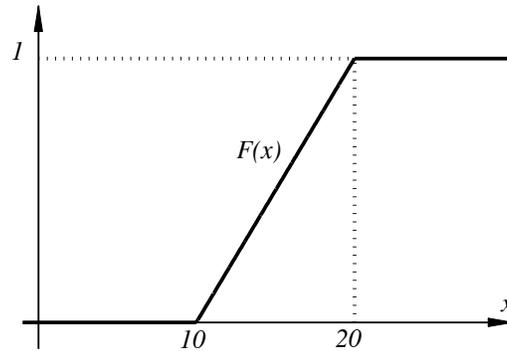
$$F(x) = \begin{cases} 0, & \text{για } x < 10, \\ \frac{x-10}{10}, & \text{για } x \in [10, 20], \\ 1, & \text{για } x > 20. \end{cases}$$

Η γραφική της αναπαράσταση δίνεται στο Σχήμα 10.3.

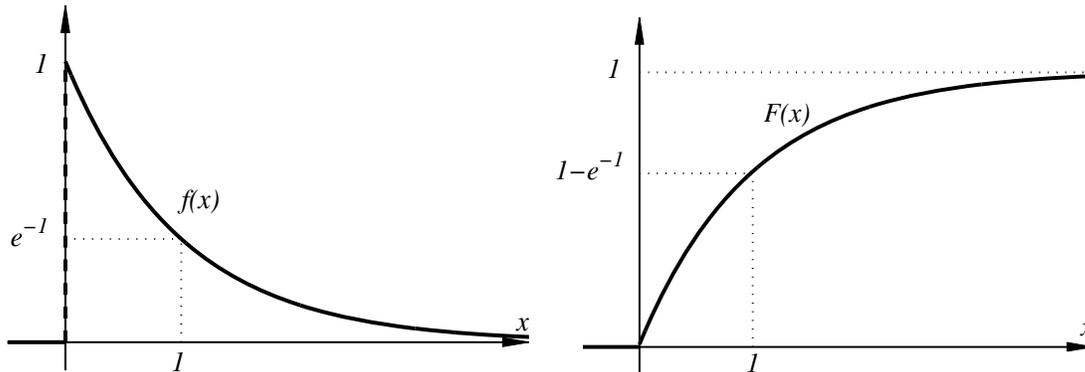
Παράδειγμα 10.4 Έστω πως η διάρκεια ζωής X , σε χρόνια, μιας οθόνης υπολογιστή είναι μια συνεχής Τ.Μ. με σύνολο τιμών $S = [0, \infty)$ και πυκνότητα,

$$f(x) = \begin{cases} 0, & \text{για } x < 0, \\ e^{-x}, & \text{για } x \geq 0. \end{cases}$$

Η γραφική της αναπαράσταση δίνεται στο Σχήμα 10.4.



Σχήμα 10.3: Γραφική αναπαράσταση της συνάρτησης κατανομής $F(x)$ της Τ.Μ. X στο Παράδειγμα 10.3.



Σχήμα 10.4: Γραφική αναπαράσταση της πυκνότητας $f(x)$ (αριστερά), και της συνάρτησης κατανομής $F(x)$ (δεξιά) της Τ.Μ. X στο Παράδειγμα 10.4.

Συνεπώς, η συνάρτηση κατανομής $F(x)$ της X θα ισούται με $F(x) = \Pr(X \leq x) = 0$ για $x < 0$, ενώ για $x \geq 0$,

$$F(x) = \Pr(X \leq x) = \int_{-\infty}^x f(y) dy = \int_0^x e^{-y} dy = [-e^{-y}]_0^x = 1 - e^{-x}.$$

Η γραφική της αναπαράσταση δίνεται στο Σχήμα 10.4.

Όπως παρατηρήσαμε πιο πάνω, το ότι γνωρίζουμε τη συνάρτηση κατανομής διευκολύνει σημαντικά τον υπολογισμό πιθανοτήτων σχετικά με τη X . Για παράδειγμα, η $\Pr(X > 3)$ μπορεί να υπολογιστεί απευθείας από την πυκνότητα ως,

$$\Pr(X > 3) = \Pr(3 < X < \infty) = \int_3^{\infty} f(x) dx = \int_3^{\infty} e^{-x} dx = [-e^{-x}]_3^{\infty} = e^{-3},$$

ή, εναλλακτικά (και ευκολότερα), μέσω της συνάρτησης κατανομής,

$$\Pr(X > 3) = 1 - \Pr(X \leq 3) = 1 - F(3) = 1 - (1 - e^{-3}) = e^{-3} \approx 0.0498 = 4.98\%.$$

Παρομοίως μπορούμε να υπολογίσουμε και δεσμευμένες πιθανότητες για τη X , όπως π.χ.,

$$\begin{aligned} \Pr(X < 4 | X > 3) &= \frac{\Pr(3 < X < 4)}{\Pr(X > 3)} \\ &\stackrel{(a)}{=} \frac{\Pr(3 \leq X \leq 4)}{1 - \Pr(X \leq 3)} \\ &\stackrel{(b)}{=} \frac{F(4) - F(3)}{1 - F(3)} \\ &= \frac{1 - e^{-4} - (1 - e^{-3})}{1 - (1 - e^{-3})} = 1 - e^{-1} \approx 0.6321, \end{aligned}$$

όπου στο βήμα (a) χρησιμοποιήσαμε τη δεύτερη βασική ιδιότητα της πυκνότητας, και στο βήμα (b) εφαρμόσαμε τη σχέση (10.8).

10.2 Μέση τιμή και διασπορά

Η μέση τιμή και η διασπορά για μια συνεχή Τ.Μ. X ορίζονται κατά τρόπο ανάλογο με εκείνον που είδαμε στην περίπτωση διακριτών τυχαίων μεταβλητών. Το ρόλο της (διακριτής) πυκνότητας $P(x)$ εδώ παίζει η (συνεχής) πυκνότητα $f(x)$, και τα αθροίσματα αντικαθίστανται από τα αντίστοιχα ολοκληρώματα.

Ορισμός 10.2 Η μέση τιμή (ή αναμενόμενη τιμή, ή προσδοκώμενη τιμή) μιας συνεχούς Τ.Μ. X με σύνολο τιμών S και πυκνότητα $f(x)$, ορίζεται ως:

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x) dx. \quad (10.9)$$

Γενικότερα, για οποιαδήποτε συνάρτηση $g : S \rightarrow \mathbb{R}$, η μέση τιμή της νέας Τ.Μ. $g(X)$ ορίζεται ως:

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx. \quad (10.10)$$

Ορισμός 10.3 Για μια συνεχή Τ.Μ. X με μέση τιμή μ , η διασπορά της X ορίζεται, ακριβώς όπως και στη διακριτή περίπτωση, ως,

$$\sigma^2 = \text{Var}(X) = E[(X - \mu)^2], \quad (10.11)$$

και η τυπική απόκλιση της X είναι:

$$\sigma = \sqrt{\text{Var}(X)}.$$

Παρατηρήσεις

1. Όπως και για τις διακριτές Τ.Μ., διαισθητικά η μέση τιμή μ μας λέει πως η X τείνει να κυμαίνεται γύρω από την τιμή μ . Παρομοίως, η διασπορά σ^2 της X είναι η μέση τετραγωνική απόκλιση της X από το μ , δηλαδή η μέση τιμή του τετραγώνου της απόστασης της τυχαίας τιμής X από τη μέση τιμή της.
2. Οι δύο παραπάνω ορισμοί (10.9) και (10.10) της μέσης τιμής μιας συνεχούς Τ.Μ. είναι ακριβώς ανάλογοι των αντίστοιχων ορισμών (6.4) και (6.5) που είδαμε στο Κεφάλαιο 6 για διακριτές Τ.Μ. Επιπλέον, εφόσον κάθε συνάρτηση μιας διακριτής Τ.Μ. είναι αναγκαστικά κι αυτή μια διακριτή Τ.Μ., στην Άσκηση 1 του Κεφαλαίου 6 δείξαμε πως ο δεύτερος ορισμός είναι στην πραγματικότητα συνέπεια του πρώτου.

Στη συνεχή περίπτωση, η αντιστοιχία είναι τεχνικά πιο σύνθετη, αλλά η ουσιαστική σχέση παραμένει η ίδια: Ο δεύτερος ορισμός (10.10) είναι και πάλι συνέπεια του πρώτου (10.9), αλλά η απόδειξη είναι αρκετά πιο πολύπλοκη και βασίζεται σε τεχνικά αποτελέσματα του τομέα της πραγματικής ανάλυσης, τα οποία αφενός ξεπερνούν τους στόχους του παρόντος βιβλίου, και αφετέρου δεν σχετίζονται άμεσα με τις ιδέες και τις τεχνικές της θεωρίας πιθανοτήτων. Μια πρώτη ένδειξη της μεγαλύτερης πολυπλοκότητας που προκύπτει στην περίπτωση ακόμα και μιας απλής συνάρτησης $g(X)$ μιας συνεχούς Τ.Μ. X θα δούμε στην Άσκηση 4 στο τέλος του κεφαλαίου, όπου θα εξετάσουμε ένα παράδειγμα στο οποίο η $g(X)$ είναι διακριτή Τ.Μ. παρότι η X είναι συνεχής.

Κατά συνέπεια, χάριν ευκολίας, επιλέγουμε να δεχτούμε τον τύπο (10.10) για τη μέση τιμή της συνάρτησης μιας συνεχούς Τ.Μ. ως δεδομένο και παραλείπουμε την απόδειξή του.

3. Συνδυάζοντας, όπως και στη διακριτή περίπτωση, τον ορισμό της διασποράς (10.11) με τον γενικό ορισμό της μέσης τιμής (10.10) προκύπτει ότι, για κάθε Τ.Μ. X με πυκνότητα $f(x)$, η διασπορά ισούται με:

$$\sigma^2 = \text{Var}(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx. \quad (10.12)$$

4. Και πάλι όπως στη διακριτή περίπτωση, η διασπορά μπορεί εναλλακτικά να εκφραστεί ως:

$$\sigma^2 = \text{Var}(X) = E(X^2) - \mu^2. \quad (10.13)$$

Ξεκινώντας από τη σχέση (10.12), έχουμε:

$$\begin{aligned}\text{Var}(X) &= \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx \\ &= \int_{-\infty}^{\infty} (x^2 - 2x\mu + \mu^2) f(x) dx \\ &= \int_{-\infty}^{\infty} x^2 f(x) dx - 2\mu \int_{-\infty}^{\infty} x f(x) dx + \mu^2 \int_{-\infty}^{\infty} f(x) dx.\end{aligned}$$

Το πρώτο ολοκλήρωμα πιο πάνω ισούται με $E(X^2)$ από την (10.10), το δεύτερο ολοκλήρωμα ισούται με $E(X) = \mu$ εξ ορισμού, και το τρίτο ολοκλήρωμα ισούται με 1 από την αντίστοιχη ιδιότητα της πυκνότητας. Συνεπώς, $\text{Var}(X) = E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2$, και άρα έχουμε αποδείξει τη σχέση (10.13).

Παράδειγμα 10.5 Για την Τ.Μ. X του Παραδείγματος 10.4, η μέση τιμή της υπολογίζεται εύκολα ολοκληρώνοντας κατά παράγοντες:

$$\begin{aligned}E(X) &= \int_{-\infty}^{\infty} x f(x) dx \\ &= \int_0^{\infty} x e^{-x} dx \\ &\stackrel{(a)}{=} [-x e^{-x}]_0^{\infty} + \int_0^{\infty} e^{-x} dx \\ &= 0 + [-e^{-x}]_0^{\infty} \\ &= 1,\end{aligned}$$

όπου στο βήμα (a) θέσαμε $u = x$ και $dv = e^{-x} dx$, έτσι ώστε $du = dx$ και $v = -e^{-x}$. Άρα, ο μέσος όρος ζωής μιας οθόνης είναι ένας χρόνος.

Με παρόμοιο τρόπο υπολογίζουμε και τη μέση τιμή του X^2 ,

$$\begin{aligned}E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x) dx \\ &= \int_0^{\infty} x^2 e^{-x} dx \\ &\stackrel{(b)}{=} [-x^2 e^{-x}]_0^{\infty} + \int_0^{\infty} 2x e^{-x} dx \\ &= 0 + 2E(X) \\ &= 2,\end{aligned}$$

όπου στο βήμα (b) θέσαμε $u = x^2$ και $dv = e^{-x} dx$, έτσι ώστε $du = 2x dx$ και $v = -e^{-x}$. Συνεπώς, από την εναλλακτική έκφραση για τη διασπορά (10.13), βρίσκουμε πως η διασπορά του X ισούται με:

$$\text{Var}(X) = E(X^2) - [E(X)]^2 = 2 - 1^2 = 1.$$

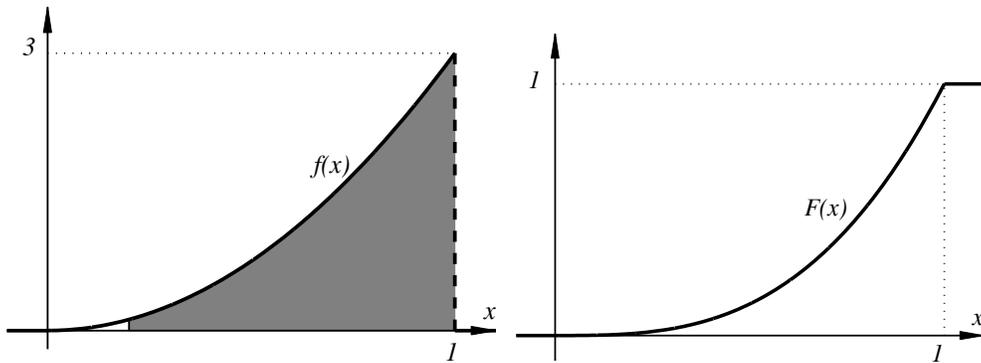
Παράδειγμα 10.6 Έστω μια συνεχής Τ.Μ. X με πυκνότητα,

$$f(x) = \begin{cases} cx^2, & \text{για } x \in [0, 1], \\ 0, & \text{για } x \notin [0, 1]. \end{cases}$$

Όπως στο Παράδειγμα 10.2, η τιμή της c μπορεί να υπολογιστεί από τον ορισμό της πυκνότητας, ο οποίος μας λέει πως το ολοκλήρωμα της $f(x)$ για όλα τα $x \in \mathbb{R}$ πρέπει να ισούται με 1,

$$1 = \int_{-\infty}^{\infty} f(x) dx = \int_0^1 cx^2 dx = \left[\frac{cx^3}{3} \right]_0^1 = \frac{c}{3},$$

και συνεπώς έχουμε $c = 3$, βλ. Σχήμα 10.5.



Σχήμα 10.5: Παράδειγμα 10.6. Αριστερά: Γραφική αναπαράσταση της πυκνότητας $f(x)$ της Τ.Μ. X και του υπολογισμού του εμβαδού το οποίο ισούται με την πιθανότητα $\Pr(X \geq 0.2)$. Δεξιά: Η συνάρτηση κατανομής $F(x)$ της X .

Έχοντας την πυκνότητα, μπορούμε να υπολογίσουμε πιθανότητες για τις τιμές της X , όπως, π.χ.,

$$\Pr(X \geq 0.2) = \Pr(0.2 \leq X \leq 1) = \int_{0.2}^1 3x^2 dx = [x^3]_{0.2}^1 = 0.992.$$

Ένας παρόμοιος υπολογισμός μάς δίνει και τη συνάρτηση κατανομής της X . Προφανώς έχουμε $F(x) = \Pr(X \leq x) = 0$ για $x < 0$, και $F(x) = \Pr(X \leq x) = 1$ όταν $x > 1$, ενώ για $x \in [0, 1]$,

$$F(x) = \Pr(X \leq x) = \int_0^x 3y^2 dy = x^3,$$

οπότε,

$$F(x) = \begin{cases} 0, & \text{για } x < 0, \\ x^3, & \text{για } x \in [0, 1], \\ 1, & \text{για } x > 1, \end{cases}$$

βλ. Σχήμα 10.5. Η πιθανότητα που υπολογίσαμε πιο πάνω μπορεί εναλλακτικά να υπολογιστεί και απευθείας από την $F(x)$ ως,

$$\Pr(X \geq 0.2) = 1 - \Pr(X < 0.2) = 1 - \Pr(X \leq 0.2) = 1 - F(0.2) = 1 - (0.2)^3 = 0.992.$$

Επίσης από την πυκνότητα μπορούμε να υπολογίσουμε τη μέση τιμή $\mu = E(X)$ της X και την τιμή του $E(X^2)$,

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_0^1 3x^3 dx = \left[\frac{3x^4}{4} \right]_0^1 = \frac{3}{4},$$

και

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx = \int_0^1 3x^4 dx = \left[\frac{3x^5}{5} \right]_0^1 = \frac{3}{5},$$

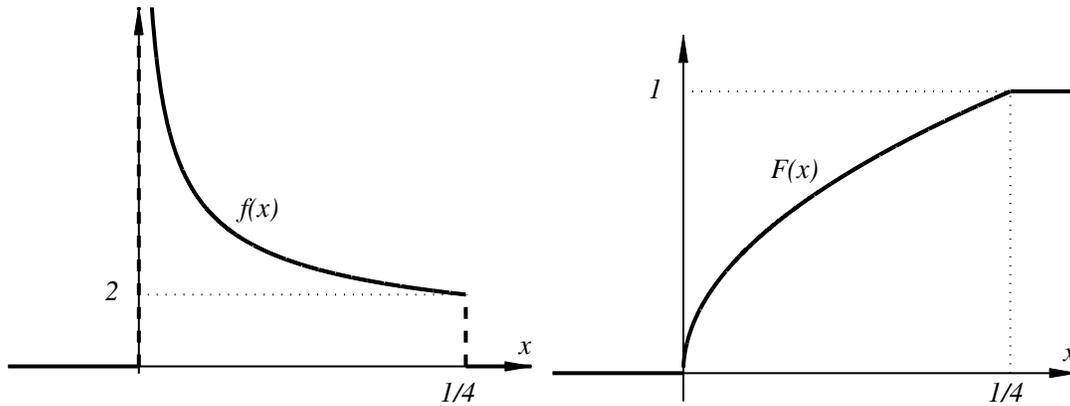
οπότε η διασπορά της X είναι:

$$\sigma^2 = \text{Var}(X) = E(X^2) - \mu^2 = \frac{3}{5} - \left(\frac{3}{4}\right)^2 = \frac{3}{80} = 0.0375.$$

Παράδειγμα 10.7 Έστω πως η συνεχής Τ.Μ. Y έχει συνάρτηση κατανομής,

$$F(y) = \begin{cases} 0, & \text{για } y < 0, \\ 2\sqrt{y}, & \text{για } 0 \leq y \leq 1/4, \\ 1, & \text{για } y > 1/4, \end{cases}$$

βλ. Σχήμα 10.6. Ποιο είναι το σύνολο τιμών της;



Σχήμα 10.6: Γραφική αναπαράσταση της συνάρτησης κατανομής $F(y)$ και της πυκνότητας $f(y)$ της Τ.Μ. Y στο Παράδειγμα 10.7.

Για να απαντήσουμε, κατ' αρχάς υπολογίζουμε την πυκνότητα $f(y)$ της Y μέσω της σχέσης (10.7),

$$f(y) = F'(y) = \begin{cases} 0, & \text{για } y < 0, \\ \frac{1}{\sqrt{y}}, & \text{για } 0 \leq y \leq 1/4, \\ 0, & \text{για } y > 1/4, \end{cases}$$

βλ. Σχήμα 10.6. Συνεπώς, $S_Y = [0, 1/4]$.

Χρησιμοποιώντας τη συνάρτηση κατανομής, μπορούμε τώρα να υπολογίσουμε πιθανότητες για τη X , όπως λ.χ. την,

$$\begin{aligned} \Pr\left(\frac{1}{16} < X < \frac{1}{9}\right) &= \Pr\left(\frac{1}{16} < X \leq \frac{1}{9}\right) \\ &= \Pr\left(X \leq \frac{1}{9}\right) - \Pr\left(X \leq \frac{1}{16}\right) \\ &= F(1/9) - F(1/16), \end{aligned}$$

το οποίο ισούται με $2\sqrt{1/9} - 2\sqrt{1/16} = 1/6$, ή, παρομοίως, δεσμευμένες πιθανότητες όπως η $\Pr(Y > 1/9 | Y > 1/16)$:

$$\begin{aligned} \Pr\left(Y > \frac{1}{9} \mid Y > \frac{1}{16}\right) &= \frac{\Pr(Y > 1/9 \text{ και } Y > 1/16)}{\Pr(Y > 1/16)} \\ &= \frac{\Pr(Y > 1/9)}{\Pr(Y > 1/16)} \\ &= \frac{1 - \Pr(Y \leq 1/9)}{1 - \Pr(Y \leq 1/16)} \\ &= \frac{1 - F(1/9)}{1 - F(1/16)} \\ &= \frac{1 - 2\sqrt{1/9}}{1 - 2\sqrt{1/16}} = 2/3. \end{aligned}$$

10.3 * Μετρησιμότητα και άπειρες τιμές

10.3.1 Ορισμός μιας συνεχούς Τ.Μ.

Όπως αναφέραμε στην Ενότητα 6.3 του Κεφαλαίου 6, όταν ο χώρος πιθανότητας Ω ή το σύνολο τιμών μιας Τ.Μ. είναι ένα άπειρο και μη αριθμησιμο σύνολο, τότε απαιτείται ιδιαίτερη προσοχή στο πώς ορίζονται οι Τ.Μ. που χρησιμοποιούμε έτσι ώστε να μην μας οδηγούν σε «μη μετρήσιμα» υποσύνολα του Ω . Σε αυτό και τα επόμενα κεφάλαια, ο τρόπος με τον οποίο αποφεύγουμε τέτοιες ενδεχόμενες παθολογίες στην περίπτωση που θέλουμε να εξετάσουμε μια συνεχή Τ.Μ. X (της οποίας πάντοτε το σύνολο τιμών είναι, εξ ορισμού, άπειρο και μη αριθμησιμο), είναι εισάγοντας τους εξής περιορισμούς:

(1.) Τα μόνα σύνολα τιμών που επιτρέπουμε για μια συνεχή Τ.Μ. είναι εκείνα που μπορούν να εκφραστούν ως ένωση ενός πεπερασμένου πλήθους διαστημάτων στο \mathbb{R} .

(2.) Μια συνεχής τυχαία μεταβλητή X ορίζεται πάντοτε σε συνδυασμό με την πυκνότητά της $f(x)$, της οποίας το ολοκλήρωμα ορίζεται σε οποιοδήποτε διάστημα $[a, b]$.

(3.) Τα μόνα ενδεχόμενα των οποίων υπολογίζουμε τις πιθανότητες (ή τις δεσμευμένες πιθανότητες) είναι της μορφής $\{X \in A\}$ για σύνολα A τα οποία μπορούν να εκφραστούν ως ένωση ενός πεπερασμένου πλήθους διαστημάτων στο \mathbb{R} .

Αν και, από αυστηρά μαθηματική σκοπιά, αυτοί οι περιορισμοί δεν είναι απαραίτητοι (υπάρχουν και πιο γενικές συνθήκες κάτω από τις οποίες μπορεί να οριστεί η έννοια της συνεχούς τυχαίας μεταβλητής), οι συνθήκες (1.)–(3.) αφενός είναι αρκετά γενικές ώστε να συμπεριλαμβάνουν όλες τις σημαντικές για εμάς εφαρμογές, και αφετέρου μας επιτρέπουν, δεδομένης οποιασδήποτε πυκνότητας $f(x)$, όπως στον Ορισμό 10.1, να κατασκευάσουμε ένα χώρο πιθανότητας Ω και μια συνάρτηση $X : \Omega \rightarrow \mathbb{R}$ έτσι ώστε η X να είναι μια (μετρήσιμη) συνεχής Τ.Μ. με πυκνότητα $f(x)$. Η κατασκευή αυτή είναι η εξής:

Έστω πως έχουμε μια συνάρτηση $f(x)$ που ικανοποιεί τις συνθήκες του Ορισμού 10.1. Θέτουμε $\Omega = \mathbb{R}$, και ορίζουμε τη συνάρτηση $X : \mathbb{R} \rightarrow \mathbb{R}$ ως $X(\omega) = \omega$. Επιπλέον, έστω $A \subset \mathbb{R}$ ένα οποιοδήποτε σύνολο που μπορεί να εκφραστεί ως ένωση ενός πεπερασμένου πλήθους N διαστημάτων (που δεν αλληλοεπικαλύπτονται), και ας πούμε πως τα άκρα αυτών των διαστημάτων είναι $a_i \leq b_i$ για $i = 1, 2, \dots, N$. Ορίζουμε ένα μέτρο πιθανότητας \mathbb{P} τέτοιο ώστε, για κάθε σύνολο A αυτής της μορφής,

$$\mathbb{P}(A) = \int_{a_1}^{b_1} f(x) dx + \int_{a_2}^{b_2} f(x) dx + \dots + \int_{a_n}^{b_n} f(x) dx. \quad (10.14)$$

Ένα από τα μεγάλα επιτεύγματα της θεωρίας μέτρου είναι η απόδειξη του ακόλουθου αποτελέσματος: Κάτω από τις πιο πάνω υποθέσεις, υπάρχει μια οικογένεια \mathcal{F} υποσυνόλων του Ω με τις πιο κάτω ιδιότητες:

- Το κενό σύνολο και το Ω ανήκουν στην \mathcal{F} : $\emptyset \in \mathcal{F}$, $\Omega \in \mathcal{F}$.
- Όλα τα υποσύνολα του \mathbb{R} τα οποία μπορούν να εκφραστούν ως ένωση ενός πεπερασμένου πλήθους διαστημάτων ανήκουν στην \mathcal{F} .
- Υπάρχει ένα μέτρο πιθανότητας \mathbb{P} το οποίο ορίζεται για κάθε στοιχείο $A \in \mathcal{F}$, και το οποίο ικανοποιεί τη σχέση (10.14).

Τώρα μπορούμε εύκολα να ελέγξουμε πως η συνάρτηση X που έχουμε ορίσει είναι «μια συνεχής τυχαία μεταβλητή με πυκνότητα $f(x)$ », υπό την έννοια του Ορισμού 10.1, δηλαδή, να επιβεβαιώσουμε πως ισχύει η τρίτη συνθήκη του ορισμού αυτού: Για $a \leq b$,

$$\Pr(a \leq X \leq b) = \Pr(X \in [a, b]) = \Pr(\{\omega : X(\omega) \in [a, b]\}),$$

και εφόσον $X(\omega) = \omega$, πράγματι προκύπτει πως ισχύει η ζητούμενη σχέση,

$$\Pr(a \leq X \leq b) = \Pr(\{\omega : \omega \in [a, b]\}) = \mathbb{P}([a, b]) = \int_a^b f(x) dx,$$

όπου στο τελευταίο βήμα εφαρμόσαμε τη (10.14) στην ειδική περίπτωση $N = 1$ ενός μόνο διαστήματος.

10.3.2 Ορισμοί $E(X)$, $\text{Var}(X)$, και επιπλέον συνθήκες

Για μια οποιαδήποτε συνεχή Τ.Μ. X με πυκνότητα $f(x)$, η μέση τιμή $E(X)$ δίνεται από ένα καταχρηστικό ολοκλήρωμα, ένα ολοκλήρωμα, δηλαδή, του οποίου τα όρια είναι $\pm\infty$. Όπως γνωρίζουμε από τα βασικά αποτελέσματα του διαφορικού λογισμού, τέτοια ολοκληρώματα μπορεί να παίρνουν άπειρες τιμές, ή ακόμα και να μην ορίζονται. (Ένα παράδειγμα μιας συνεχούς Τ.Μ. με άπειρη μέση τιμή δίνεται στην Άσκηση 5 στο τέλος του κεφαλαίου.)

Για να αποφύγουμε τεχνικές λεπτομέρειες που ξεφεύγουν από τα ζητούμενα του παρόντος βιβλίου, υιοθετούμε τις πιο κάτω συμβατικές υποθέσεις (ανάλογες εκείνων του Κεφαλαίου 6 για διακριτές Τ.Μ.), οι οποίες θα παραμείνουν εν ισχύ σε όλα τα υπόλοιπα κεφάλαια.

Συμβάσεις

- Πάντοτε, όταν λέμε πως «η συνεχής Τ.Μ. X έχει μέση τιμή $\mu = E(X)$ », εμμέσως υποθέτουμε ότι το ολοκλήρωμα που δίνει την $E(X)$ ορίζεται και ότι η τιμή του είναι πεπερασμένη.
- Πάντοτε, όταν λέμε ότι «η συνεχής Τ.Μ. X έχει διασπορά $\sigma^2 = \text{Var}(X)$ », εμμέσως υποθέτουμε ότι η μέση τιμή $\mu = E(X)$ ορίζεται και είναι πεπερασμένη και πως το ολοκλήρωμα που δίνει την $\text{Var}(X)$ ορίζεται και δίνει πεπερασμένο αποτέλεσμα.
- Όποτε διατυπώνεται μια ιδιότητα για τη μέση τιμή (ή τη διασπορά), εμμέσως υποθέτουμε πως η αντίστοιχη μέση τιμή (αντίστοιχα, διασπορά) ορίζεται και είναι πεπερασμένη.

10.4 Ασκήσεις

1. *Κυκλικός δίσκος.* Έστω ότι επιλέγεται εντελώς τυχαία ένα σημείο του κυκλικού δίσκου Ω με κέντρο την αρχή των αξόνων και ακτίνα 1, δηλαδή χωρίς κάποια προτίμηση σε κάποια περιοχή του δίσκου. Έστω Z η απόσταση του σημείου από την αρχή των αξόνων.

Βρείτε το σύνολο τιμών της Τ.Μ. Z , και υπολογίστε τη συνάρτηση κατανομής και την πυκνότητά της.

2. *Προσδιορισμός παραμέτρων.* Η συνάρτηση κατανομής μιας συνεχούς Τ.Μ. X δίνεται από τον τύπο:

$$F(x) = \begin{cases} 0, & x < 4, \\ Ax + B - \frac{4}{x}, & x \geq 4. \end{cases}$$

(α') Να βρεθούν οι τιμές των A και B .

(β') Να υπολογιστεί η πυκνότητα $f(x)$ της X και να γίνει η γραφική της παράσταση.

(γ') Βρείτε την πιθανότητα το X να είναι μικρότερο του 5, δεδομένου ότι είναι μικρότερο του 6.

3. *Μια απλή πυκνότητα.* Έστω μια συνεχής Τ.Μ. X με πυκνότητα:

$$f(x) = \begin{cases} 0, & x < 0, \\ cx, & 0 \leq x \leq 1, \\ c, & 1 \leq x \leq 2, \\ 0, & x > 2. \end{cases}$$

(α') Σχεδιάστε το γράφημα της $f(x)$ και βρείτε την τιμή της σταθεράς c .

(β') Βρείτε την πιθανότητα το X να είναι μεγαλύτερο από 1.5 ή μικρότερο από 0.5.

(γ') Υπολογίστε τη μέση τιμή της X .

4. *Μέση τιμή συναρτήσεων Τ.Μ.* Στον Ορισμό 10.2, η μέση τιμή $E[g(X)]$ μιας συνάρτησης $g(x)$ της Τ.Μ. X ορίστηκε μέσω του τύπου (10.10). Αλλά η $Y = g(X)$ είναι και η ίδια μια τυχαία μεταβλητή, με τη δική της (ενδεχομένως συνεχή ή διακριτή) πυκνότητα, οπότε, αν εφαρμόσουμε τον τύπο του αντίστοιχου ορισμού (6.4) ή (10.9) για τη μέση τιμή της Y , το αποτέλεσμα θα ισούται με εκείνο του γενικότερου τύπου (10.10) για την $E[g(X)]$. Εδώ θα αποδείξετε μια ειδική περίπτωση αυτής της γενικής ιδιότητας.

Έστω μια Τ.Μ. X με πυκνότητα όπως εκείνη του Παραδείγματος 10.4, και έστω η συνάρτηση $g : [0, \infty) \rightarrow \{-1, 0, 1\}$,

$$g(x) = \begin{cases} -1, & \text{για } 0 \leq x < 1, \\ 0, & \text{για } 1 \leq x \leq 5, \\ 1, & \text{για } x > 5. \end{cases}$$

- (α') Υπολογίστε τη μέση τιμή της $g(X)$ μέσω του τύπου (10.10).
 (β') Ορίστε τη διακριτή Τ.Μ. $Y = g(X)$, βρείτε το σύνολο τιμών και την πυκνότητά της και υπολογίστε τη μέση τιμή της μέσω του ορισμού (6.4) της μέσης τιμής μιας διακριτής Τ.Μ.

Συγκρίνετε το πιο πάνω αποτέλεσμα με αυτό της Άσκησης 1 του Κεφαλαίου 6.

5. Άπειρη μέση τιμή. Έστω μια συνεχής Τ.Μ. X με σύνολο τιμών το $S = [2, \infty)$ και πυκνότητα $f(x) = C/x^2$ για $x \geq 2$.

- (α') Υπολογίστε την τιμή της σταθεράς C .
 (β') Αποδείξτε πως η μέση τιμή $E(X) = +\infty$.

6. Κατανομή Βήτα. Μια συνεχής τυχαία μεταβλητή X έχει πυκνότητα:

$$f(x) = \begin{cases} cx(1-x), & 0 \leq x \leq 1, \\ 0, & x \notin [0, 1]. \end{cases}$$

- (α') Υπολογίστε την τιμή της σταθεράς c .
 (β') Υπολογίστε την πιθανότητα $\Pr(\frac{1}{2} \leq X \leq \frac{3}{4})$.
 (γ') Βρείτε τη συνάρτηση κατανομής $F(x)$.

7. Υποψήφιες πυκνότητες. Για καθεμία από τις παρακάτω συναρτήσεις $f(x)$ σχεδιάστε το γράφημά της και αποφασίστε αν είναι πυκνότητα ή όχι. Δικαιολογήστε τις απαντήσεις σας. Για όσες είναι πυκνότητες, προσδιορίστε τη μέση τιμή, τη διασπορά, και την πιθανότητα να είναι η αντίστοιχη Τ.Μ. X μεγαλύτερη από τη μέση τιμή της.

(α')

$$f(x) = \begin{cases} \frac{3}{4}(1-x^2), & x \in [-1, 1], \\ 0, & x \notin [-1, 1]. \end{cases}$$

(β')

$$f(x) = \begin{cases} \frac{3}{4}(1-x), & x \in [0, 1], \\ 0, & x \notin [0, 1]. \end{cases}$$

(γ')

$$f(x) = \begin{cases} \frac{3}{4}(1-x), & x \in [-\frac{2}{3}, \frac{2}{3}], \\ 0, & x \notin [-\frac{2}{3}, \frac{2}{3}]. \end{cases}$$

(δ')

$$f(x) = \begin{cases} (\frac{1}{2}-x), & x \in [-1, 1], \\ 0, & x \notin [-1, 1]. \end{cases}$$

8. Μια απλή πυκνότητα στο \mathbb{R} . Έστω πως μια συνεχής Τ.Μ. X έχει πυκνότητα,

$$f(x) = ce^{-4|x|}, \quad x \in \mathbb{R},$$

όπου η c είναι μια άγνωστη θετική παράμετρος.

- (α') Ποια είναι η τιμή της c ;
 (β') Πόση είναι η μέση τιμή $E(X)$;
 (γ') Πόση είναι η διασπορά $\text{Var}(X)$;
 (δ') Υπολογίστε την πιθανότητα $\Pr(|X| > 1/2)$.
9. *Βαρουφάκης εναντίον Merkel.* Πριν από κάθε συνεδρίαση του Eurogroup την άνοιξη του 2015, ο (τότε) υπουργός Οικονομικών Γιάνης Βαρουφάκης βρισκόταν σε μία από δύο πιθανές ψυχικές διαθέσεις. Με πιθανότητα $1/2$ ήταν σοβαρός, οπότε συζητούσε κανονικά με την καγκελάριο Merkel για ένα τυχαίο διάστημα X μέχρι τρεις ώρες. Αντίθετα, με πιθανότητα $1/2$ ήταν σε διάθεση χαβαλέ, οπότε τρόλαρε την καγκελάριο Merkel για το πολύ ένα τέταρτο και διαλυόταν νωρίς το Eurogroup. Έστω Y η τυχαία διάρκεια του Eurogroup στη δεύτερη περίπτωση.
- Αν η Τ.Μ. X έχει πυκνότητα $g(x) = 1/3$ για $0 \leq x \leq 3$ και $g(x) = 0$ για $x \notin [0, 3]$, ενώ η Y έχει την πυκνότητα $f(y)$ του Παραδείγματος 10.7, υπολογίστε την πιθανότητα ένα τυχαίο Eurogroup να διήρκεσε το πολύ 10 λεπτά.
10. *Συνεχείς και διακριτές τυχαίες μεταβλητές.* Έστω μια διακριτή Τ.Μ. X , η οποία παίρνει τις τιμές 1, 2 και 3, με πιθανότητα $1/3$ για την καθεμία. Δεδομένου ότι $X = a$, ορίζουμε δύο συνεχείς Τ.Μ. Y_1 και Y_2 , οι οποίες είναι ανεξάρτητες και έχουν την ίδια πυκνότητα $f_a(y)$, όπου $f_a(y) = 1/a$ για $y \in [0, a]$, και $f_a(y) = 0$ για $y \notin [0, a]$.
- (α') Υπολογίστε την πιθανότητα του ενδεχομένου $\{Y_1 \leq 1/2\}$.
 (β') Υπολογίστε την πυκνότητα $f(y)$ του Y_1 . Βεβαιωθείτε πως η $f(y)$ που βρήκατε πράγματι είναι η πυκνότητα του Y_1 .
 (γ') Υπολογίστε τη (δεσμευμένη) πιθανότητα του ενδεχομένου $\{Y_1 \leq 1/2 \text{ και } Y_2 \leq 1/2\}$, δεδομένου ότι $X = 1$.
 (δ') Αντίστροφα, υπολογίστε τη δεσμευμένη πιθανότητα του ενδεχομένου $\{X = 1\}$, δεδομένου ότι $Y_1 \leq 1/2$ και $Y_2 \leq 1/2$.
11. *Απόσταση χ^2 .* Στην Άσκηση 18 του Κεφαλαίου 6 ορίσαμε την χ^2 -απόσταση μεταξύ δύο διακριτών πυκνοτήτων. Για δύο συνεχείς πυκνότητες $f(x), g(x)$ με κοινό σύνολο τιμών το $S \subseteq \mathbb{R}$ παρομοίως ορίζουμε την χ^2 -απόσταση της $g(x)$ από την $f(x)$ ως:

$$d_{\chi^2}(f, g) = \int_S \frac{(f(x) - g(x))^2}{g(x)} dx.$$

(α') Δείξτε ότι η $d_{\chi^2}(f, g)$ μπορεί εναλλακτικά να εκφραστεί ως:

$$d_{\chi^2}(f, g) = \int_S \frac{f(x)^2}{g(x)} dx - 1.$$

(β') Διατυπώστε και αποδείξτε το αντίστοιχο αποτέλεσμα στην περίπτωση διακριτών πυκνοτήτων.

Κεφάλαιο 11

Συνεχείς κατανομές και ο Ν.Μ.Α.

Στο προηγούμενο κεφάλαιο ορίσαμε την έννοια της *συνεχούς τυχαιάς μεταβλητής*, και είδαμε τις βασικές της ιδιότητες. Εδώ θα περιγράψουμε κάποιους ιδιαίτερους τύπους τυχαιών μεταβλητών, οι οποίες είναι χρήσιμες στην πράξη και εμφανίζονται συχνά σε βασικά προβλήματα των πιθανοτήτων. Συγκεκριμένα, θα ορίσουμε δύο οικογένειες κατανομών για συνεχείς Τ.Μ., και θα αποδείξουμε ορισμένες ιδιότητές τους. Τέλος, στις Ενότητες 11.2 και 11.3 θα διατυπώσουμε τις φυσικές γενικεύσεις στη συνεχή περίπτωση κάποιων σημαντικών αποτελεσμάτων που είδαμε σε προηγούμενα κεφάλαια.

11.1 Ομοιόμορφη και εκθετική κατανομή

Η πιο απλή περίπτωση μιας συνεχούς Τ.Μ. είναι εκείνη που παίρνει «ομοιόμορφα τυχαιές» τιμές σε κάποιο διάστημα $[a, b]$ στο \mathbb{R} .

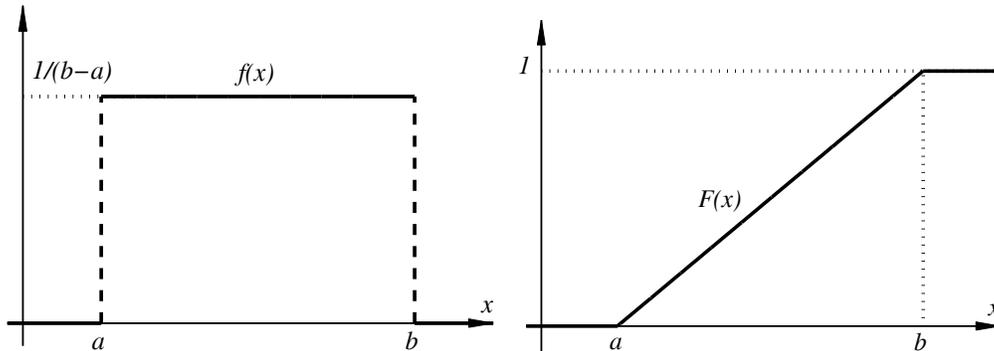
Ορισμός 11.1 Μια συνεχής Τ.Μ. X έχει ομοιόμορφη κατανομή στο διάστημα $[a, b]$, για κάποια $a < b$, αν έχει σύνολο τιμών το $S = [a, b]$ και πυκνότητα,

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{για } x \in [a, b], \\ 0, & \text{για } x \notin [a, b], \end{cases}$$

βλ. Σχήμα 11.1. Για συντομία, αυτό συμβολίζεται: $X \sim U[a, b]$.

Παρατηρήσεις

1. Μια Τ.Μ. X με ομοιόμορφη κατανομή στο $[a, b]$ περιγράφει μια ποσότητα που παίρνει «εντελώς τυχαιές» τιμές σε αυτό το διάστημα. Π.χ., η Τ.Μ. X στο Παράδειγμα 10.2 του Κεφαλαίου 10 έχει $X \sim U[10, 20]$.



Σχήμα 11.1: Γραφική αναπαράσταση της πυκνότητας $f(x)$ (αριστερά) και της συνάρτησης κατανομής $F(x)$ (δεξιά) μιας Τ.Μ. X με ομοιόμορφη κατανομή στο διάστημα $[a, b]$.

2. Η συνάρτηση κατανομής $F(x)$ μιας Τ.Μ. $X \sim U[a, b]$ υπολογίζεται εύκολα από τη σχέση (10.4). Προφανώς έχουμε,

$$F(x) = \Pr(X \leq x) = 0, \quad \text{για } x < a,$$

$$\text{και } F(x) = \Pr(X \leq x) = 1, \quad \text{για } x > b,$$

ενώ για $x \in [a, b]$ βρίσκουμε,

$$F(x) = \Pr(X \leq x) = \int_{-\infty}^x f(y) dy = \int_a^x \frac{1}{b-a} dy = \left[\frac{y}{b-a} \right]_a^x = \frac{x-a}{b-a}.$$

Η γραφική της αναπαράσταση δίνεται στο Σχήμα 11.1

3. Η μέση τιμή μιας Τ.Μ. $X \sim U[a, b]$ επίσης υπολογίζεται εύκολα:

$$\begin{aligned} \mu = E(X) &= \int_{-\infty}^{\infty} x f(x) dx = \int_a^b \frac{x}{b-a} dx = \left[\frac{x^2}{2(b-a)} \right]_a^b \\ &= \frac{b^2 - a^2}{2(b-a)} = \frac{(b-a)(b+a)}{2(b-a)} = \frac{a+b}{2}. \end{aligned}$$

Παρομοίως υπολογίζουμε και τη μέση τιμή της X^2 ,

$$\begin{aligned} E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x) dx = \int_a^b \frac{x^2}{b-a} dx = \left[\frac{x^3}{3(b-a)} \right]_a^b = \frac{b^3 - a^3}{3(b-a)} \\ &= \frac{(a^2 + ab + b^2)(b-a)}{3(b-a)} = \frac{a^2 + ab + b^2}{3}, \end{aligned}$$

οπότε, η διασπορά της X από την έκφραση (10.13) ισούται με,

$$\begin{aligned}\text{Var}(X) &= E(X^2) - \mu^2 = \frac{a^2 + ab + b^2}{3} - \left(\frac{a+b}{2}\right)^2 \\ &= \frac{4a^2 + 4ab + 4b^2 - 3a^2 - 6ab - 3b^2}{12} = \frac{(b-a)^2}{12}.\end{aligned}$$

Συνοψίζοντας, μια Τ.Μ. $X \sim U[a, b]$ έχει μέση τιμή που ισούται με το μέσο $(a+b)/2$ του $[a, b]$ και διασπορά $(b-a)^2/12$ η οποία εξαρτάται μόνο από το μήκος $(b-a)$ του διαστήματος όπου η X παίρνει τιμές, και όχι από τη θέση του:

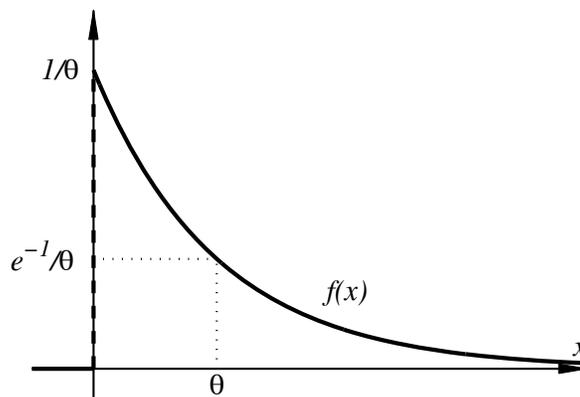
$$E(X) = \frac{a+b}{2} \quad \text{και} \quad \text{Var}(X) = \frac{(b-a)^2}{12}.$$

Μια άλλη σημαντική οικογένεια τυχαίων μεταβλητών είναι εκείνες που έχουν *εκθετική κατανομή*. Μια ειδική περίπτωση αυτής της κατανομής συναντήσαμε ήδη στο Παράδειγμα 10.4 του προηγούμενου κεφαλαίου.

Ορισμός 11.2 Μια συνεχής Τ.Μ. X έχει *εκθετική κατανομή* με παράμετρο $\theta > 0$, αν έχει σύνολο τιμών το $S = [0, \infty)$ και πυκνότητα,

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-x/\theta}, & \text{για } x \geq 0, \\ 0, & \text{για } x < 0, \end{cases}$$

βλ. Σχήμα 11.2. Για συντομία, αυτό συμβολίζεται: $X \sim \text{Eκ}\theta(\theta)$.



Σχήμα 11.2: Γραφική αναπαράσταση της πυκνότητας $f(x)$ μιας Τ.Μ. X με κατανομή $\text{Eκ}\theta(\theta)$.

Παρατήρηση: Από κάποιες απόψεις, η εκθετική κατανομή είναι ανάλογη στη συνεχή περίπτωση της γεωμετρικής κατανομής την οποία συναντήσαμε στο Κεφάλαιο 7 και συχνά παρομοίως χρησιμοποιείται για να περιγράψει το χρόνο που μεσολαβεί μέχρι να συμβεί κάποιο τυχαίο γεγονός. Λόγου χάρη, η Τ.Μ. X στο Παράδειγμα 10.4 του προηγούμενου κεφαλαίου, η οποία είχε $X \sim \text{Eκ}\theta(1)$, περιέγραφε το χρόνο μέχρι να εμφανιστεί η πρώτη βλάβη μιας οθόνης υπολογιστή.

Ένας λόγος για τον οποίο θεωρούμε την εκθετική και τη γεωμετρική συγγενείς κατανομές είναι διότι μοιράζονται αρκετές από τις θεμελιώδεις ιδιότητές τους, όπως βλέπουμε συγκρίνοντας τα αποτελέσματα του Θεωρήματος 7.1 με εκείνα του Θεωρήματος 11.1 που ακολουθεί. Ένας πιο φορμαλιστικός τρόπος για να διαπιστώσουμε την ομοιότητά τους είναι το γεγονός πως η πυκνότητα $f(x)$ μιας $\text{Eκ}\theta(\theta)$ τυχαίας μεταβλητής είναι μαθηματικά πανομοιότυπη με την πυκνότητα $P(k)$ μιας Τ.Μ. με Γεωμ(p) κατανομή.

Συγκεκριμένα, στη συνεχή περίπτωση, για $x > 0$ η πυκνότητα $f(x)$ είναι της μορφής,

$$f(x) = \frac{1}{\theta} e^{-x/\theta} = C \gamma^x, \quad x > 0, \quad (11.1)$$

όπου ορίσαμε τις σταθερές $C = 1/\theta$ και $\gamma = e^{-1/\theta}$. Αντίστοιχα, στη διακριτή περίπτωση, για $k \geq 1$ η πυκνότητα $P(k)$ μπορεί να εκφραστεί ως,

$$P(k) = p(1-p)^{k-1} = \frac{p}{1-p} \cdot (1-p)^k = C \gamma^k, \quad k \geq 1, \quad (11.2)$$

όπου τώρα έχουμε τις σταθερές $C = p/(1-p)$ και $\gamma = (1-p)$. Η ομοιότητα μεταξύ των δύο εκφράσεων στις σχέσεις (11.1) και (11.2) είναι εμφανής.

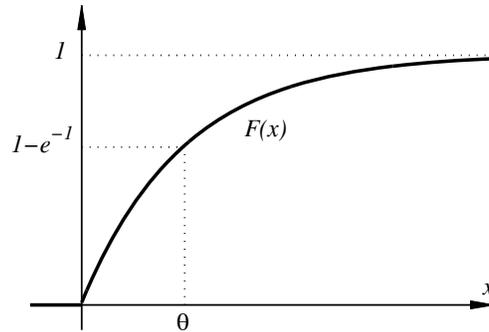
Θεώρημα 11.1 (Ιδιότητες της εκθετικής κατανομής)

Έστω $X \sim \text{Eκ}\theta(\theta)$. Η X έχει τις εξής ιδιότητες:

1. Μέση τιμή: $E(X) = \theta$.
2. Διασπορά: $\text{Var}(X) = \theta^2$.
3. Συνάρτηση κατανομής: $F(x) = 0$ για $x < 0$ και $F(x) = 1 - e^{-x/\theta}$ για $x \geq 0$. Η γραφική της αναπαράσταση δίνεται στο Σχήμα 11.3.
4. Ιδιότητα έλλειψης μνήμης: Για κάθε $a, b > 0$:

$$\Pr(X \geq a + b | X \geq a) = \Pr(X \geq b).$$

Η πιθανότητα $\Pr(X \geq a + b | X \geq a)$ είναι ανεξάρτητη του a , και ίση με εκείνη που αντιστοιχεί στο $b = 0$, δηλαδή, $\Pr(X \geq b)$.



Σχήμα 11.3: Γραφική αναπαράσταση της συνάρτησης κατανομής $F(x)$ μιας Τ.Μ. X με κατανομή $\text{Exp}(\theta)$.

Απόδειξη: Από τον ορισμό της μέσης τιμής μιας συνεχούς Τ.Μ. και την πυκνότητα της εκθετικής κατανομής, βρίσκουμε,

$$\begin{aligned}
 E(X) &= \int_{-\infty}^{\infty} x f(x) dx \\
 &= \int_0^{\infty} \frac{x}{\theta} e^{-x/\theta} dx \\
 &\stackrel{(a)}{=} \int_0^{\infty} y e^{-y} \theta dy \\
 &\stackrel{(b)}{=} [-\theta y e^{-y}]_0^{\infty} + \theta \int_0^{\infty} e^{-y} dy \\
 &= 0 + [-\theta e^{-y}]_0^{\infty} = \theta,
 \end{aligned}$$

όπου στο βήμα (a) κάναμε την αντικατάσταση $y = x/\theta$, και στο βήμα (b) ολοκληρώσαμε κατά παράγοντες, θέτοντας $u = y$ και $dv = e^{-y} dy$, έτσι ώστε $du = dy$ και $v = -e^{-y}$. Άρα έχουμε $\mu = E(X) = \theta$, αποδεικνύοντας την πρώτη ιδιότητα.

Για τη δεύτερη ιδιότητα, παρομοίως υπολογίζουμε τη μέση τιμή της X^2 ,

$$\begin{aligned}
 E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x) dx \\
 &= \int_0^{\infty} \frac{x^2}{\theta} e^{-x/\theta} dx \\
 &\stackrel{(c)}{=} \int_0^{\infty} \theta y^2 e^{-y} \theta dy \\
 &\stackrel{(d)}{=} [-\theta^2 y^2 e^{-y}]_0^{\infty} + 2\theta^2 \int_0^{\infty} y e^{-y} dy,
 \end{aligned}$$

όπου στο βήμα (c) κάναμε πάλι την αντικατάσταση $y = x/\theta$, και στο βήμα (d) ολοκληρώσαμε κατά παράγοντες, θέτοντας $u = y^2$ και $dv = e^{-y}dy$, έτσι ώστε $du = 2ydy$ και $v = -e^{-y}$. Τώρα παρατηρούμε πως ο πρώτος όρος στην τελευταία παραπάνω έκφραση είναι μηδενικός και πως το τελευταίο ολοκλήρωμα ισούται με τη μέση τιμή μιας Τ.Μ. με κατανομή $\text{Ex}\theta(1)$, άρα, από την πρώτη ιδιότητα, είναι ίσο με 1. Συνεπώς έχουμε $E(X^2) = 0 + 2\theta^2 \cdot 1 = 2\theta^2$ και η διασπορά της X είναι,

$$\text{Var}(X) = E(X^2) - \mu^2 = 2\theta^2 - \theta^2 = \theta^2,$$

οπότε έχουμε αποδείξει και τη δεύτερη ιδιότητα.

Για την Ιδιότητα 3 παρατηρούμε πως, εφόσον το σύνολο τιμών της X είναι το $[0, \infty)$, προφανώς για $x < 0$ έχουμε, $F(x) = \Pr(X \leq x) = 0$. Για $x \geq 0$, βρίσκουμε,

$$F(x) = \Pr(X \leq x) = \int_{-\infty}^x f(y) dy = \int_0^x \frac{1}{\theta} e^{-y/\theta} dy = \left[-e^{-y/\theta} \right]_0^x = 1 - e^{-x/\theta}.$$

Τέλος, για την Ιδιότητα 4, χρησιμοποιούμε τον ορισμό της δεσμευμένης πιθανότητας σε συνδυασμό με την Ιδιότητα 3:

$$\begin{aligned} \Pr(X \geq a + b | X \geq a) &= \frac{\Pr(X \geq a + b \text{ και } X \geq a)}{\Pr(X \geq a)} \\ &= \frac{\Pr(X \geq a + b)}{\Pr(X \geq a)} \\ &= \frac{1 - \Pr(X \leq a + b)}{1 - \Pr(X \leq a)} \\ &= \frac{1 - F(a + b)}{1 - F(a)} \\ &= \frac{e^{-(a+b)/\theta}}{e^{-a/\theta}} = e^{-b/\theta} = 1 - F(b) = \Pr(X \geq b). \end{aligned}$$

■

Παράδειγμα 11.1 Έστω πως ο χρόνος X , σε μήνες, μέχρι την πρώτη φορά που ένας σκληρός δίσκος θα παρουσιάσει κάποιο σφάλμα, έχει εκθετική κατανομή με μέσο όρο 30 μήνες. Συνεπώς, από την πρώτη ιδιότητα του Θεωρήματος 11.1, η $X \sim \text{Ex}\theta(30)$.

Η πιθανότητα το πρώτο σφάλμα να εμφανιστεί μετά τους πρώτους 30 μήνες, είναι,

$$\Pr(X > 30) = 1 - \Pr(X \leq 30) = 1 - F(30) = 1 - [1 - e^{-30/30}] = e^{-1} \approx 0.3679.$$

Και η πιθανότητα ο δίσκος να μην παρουσιάσει σφάλμα για τους επόμενους 30 μήνες, δεδομένου ότι ήδη λειτουργεί 30 μήνες χωρίς πρόβλημα, είναι και πάλι,

$$\Pr(X > 60 | X > 30) = \Pr(X > 30) = e^{-1} \approx 0.3679,$$

όπου χρησιμοποιήσαμε την ιδιότητα έλλειψης μνήμης. Τέλος, η πιθανότητα το πρώτο σφάλμα να εμφανιστεί μετά τον δέκατο μήνα αλλά πριν τον εικοστό είναι,

$$\begin{aligned}\Pr(10 < X < 20) &= \Pr(10 \leq X \leq 20) = F(20) - F(10) \\ &= 1 - e^{-20/30} - [1 - e^{-10/30}] \approx 0.2031,\end{aligned}$$

όπου εδώ χρησιμοποιήσαμε τη γενική ιδιότητα (10.8).

11.2 Μετασχηματισμοί

Σε πολλές περιπτώσεις, όπως έχουμε δει σε παραδείγματα προηγούμενων κεφαλαίων, μπορεί να μας ενδιαφέρει περισσότερο κάποια συνάρτηση μιας τυχαίας μεταβλητής X παρά η ίδια η τιμή της.

Παράδειγμα 11.2 Έστω μια συνεχής Τ.Μ. X με πυκνότητα $f(x)$, μέση τιμή $\mu = E(X)$ και διασπορά $\sigma^2 = \text{Var}(X)$. Ποια είναι η μέση τιμή και η διασπορά της νέας Τ.Μ. $Y = aX + b$, για κάποιες σταθερές $a, b \in \mathbb{R}$;

Από τον γενικό ορισμό της μέσης τιμής, έχουμε,

$$E(aX + b) = \int_{-\infty}^{\infty} (ax + b) f(x) dx = a \int_{-\infty}^{\infty} x f(x) dx + b \int_{-\infty}^{\infty} f(x) dx = aE(X) + b.$$

Παρομοίως για τη διασπορά της Y βρίσκουμε,

$$\begin{aligned}\text{Var}(aX + b) &= E\left\{[(aX + b) - E(aX + b)]^2\right\} \\ &= \int_{-\infty}^{\infty} [ax + b - (a\mu + b)]^2 f(x) dx \\ &= a^2 \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx \\ &= a^2 \text{Var}(X).\end{aligned}$$

Τα δύο παραπάνω αποτελέσματα, μαζί με ένα ακόμα συνοψίζονται στο θεώρημα που ακολουθεί.

Θεώρημα 11.2 (Γραμμικός μετασχηματισμός)

Έστω μια συνεχής Τ.Μ. X με πυκνότητα $f(x)$, μέση τιμή $\mu = E(X)$ και διασπορά $\sigma^2 = \text{Var}(X)$. Για $a, b \in \mathbb{R}$, ορίζουμε τη νέα Τ.Μ. $Y = aX + b$.

1. Η Y έχει μέση τιμή $E(Y) = E(aX + b) = aE(X) + b$.
2. Η Y έχει διασπορά $\text{Var}(Y) = a^2 \text{Var}(X)$.

3. Αν η σταθερά a είναι θετική, τότε η πυκνότητα $g(y)$ της Y δίνεται από τη σχέση:

$$g(y) = \frac{1}{a} f\left(\frac{y-b}{a}\right), \quad \text{για κάθε } y.$$

Η περίπτωση $a < 0$ εξετάζεται στην Άσκηση 6 στο τέλος του κεφαλαίου.

Απόδειξη: Οι δύο πρώτες ιδιότητες έχουν ήδη αποδειχθεί στο Παράδειγμα 11.2. Για την τρίτη ιδιότητα εξετάζουμε αρχικά τις συναρτήσεις κατανομής, έστω $F(x)$ και $G(y)$ των τυχαίων μεταβλητών X, Y , αντίστοιχα. Για την $G(y)$ έχουμε,

$$G(y) = \Pr(Y \leq y) = \Pr(aX + b \leq y) = \Pr\left(X \leq \frac{y-b}{a}\right) = F\left(\frac{y-b}{a}\right).$$

Άρα, από τη σχέση (10.7), μπορούμε εύκολα να υπολογίσουμε την πυκνότητα της Y ως την παράγωγο της συνάρτησης κατανομής της,

$$g(y) = G'(y) = \frac{d}{dy} F\left(\frac{y-b}{a}\right) = \frac{1}{a} F'\left(\frac{y-b}{a}\right) = \frac{1}{a} f\left(\frac{y-b}{a}\right),$$

που είναι ακριβώς η ζητούμενη σχέση. ■

Εφαρμόζοντας το παραπάνω αποτέλεσμα, έχουμε το εξής σημαντικό πόρισμα, το οποίο θα αποδείξετε στην Άσκηση 7 στο τέλος του κεφαλαίου.

Πόρισμα 11.1 Για μια οποιαδήποτε (συνεχή ή διακριτή) τυχαία μεταβλητή X με μέση τιμή μ και διασπορά σ^2 , η κανονικοποιημένη τυχαία μεταβλητή,

$$Y = \frac{X - \mu}{\sigma}, \quad \text{έχει } E(Y) = 0 \text{ και } \text{Var}(Y) = 1.$$

Όπως είδαμε στην απόδειξη της τρίτης ιδιότητας του Θεωρήματος 11.2, για να υπολογίσουμε την πυκνότητα μιας συνάρτησης μιας συνεχούς Τ.Μ. συχνά είναι απλούστερο να εξετάσουμε πρώτα τη συνάρτηση κατανομής της. Ένα λίγο πιο πολύπλοκο παράδειγμα είναι το ακόλουθο.

Παράδειγμα 11.3 Έστω μια Τ.Μ. X με κατανομή $\text{Exp}(1)$, και έστω μια νέα Τ.Μ. $Y = X^2$. Εφόσον η X έχει σύνολο τιμών το $[0, \infty)$, και η Y θα έχει προφανώς το ίδιο σύνολο τιμών, άρα η πυκνότητα $g(y)$ της Y θα ισούται με 0 για $y \leq 0$.

Έστω, τώρα, $f(x)$ και $F(x)$ η πυκνότητα και η συνάρτηση κατανομής της X , αντίστοιχα. Για να υπολογίσουμε την πυκνότητα $g(y)$ της Y για $y > 0$, πρώτα εξετάζουμε τη συνάρτηση κατανομής της, έστω $G(y)$:

$$G(y) = \Pr(Y \leq y) = \Pr(X^2 \leq y) = \Pr(X \leq \sqrt{y}) = F(\sqrt{y}),$$

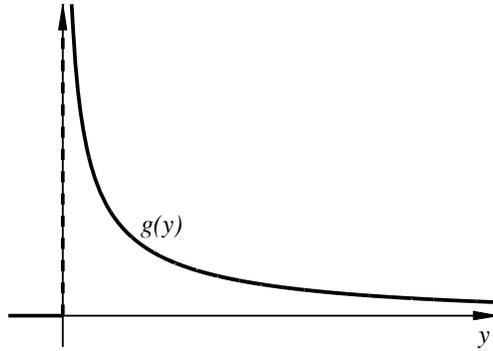
οπότε η πυκνότητα της Y για $y > 0$ προκύπτει από τη σχέση (10.7) ως,

$$g(y) = G'(y) = \frac{d}{dy} F(\sqrt{y}) = \frac{d}{dy} (\sqrt{y}) F'(\sqrt{y}) = \frac{1}{2\sqrt{y}} f(\sqrt{y}).$$

Και εφόσον γνωρίζουμε, εξ' ορισμού, πως για $x \geq 0$ η $\text{Exp}(1)$ πυκνότητα ισούται με $f(x) = e^{-x}$, αντικαθιστώντας,

$$g(y) = \frac{1}{2\sqrt{y}} e^{-\sqrt{y}}, \quad \text{για } y > 0.$$

Η γραφική της αναπαράσταση δίνεται στο Σχήμα 11.4.



Σχήμα 11.4: Γραφική αναπαράσταση της πυκνότητας $g(y)$ της Τ.Μ. Y στο Παράδειγμα 11.3.

11.3 Ανεξαρτησία, ανισότητες και ο Ν.Μ.Α.

Όπως είδαμε στο Κεφάλαιο 9, η μέση τιμή και η διασπορά ικανοποιούν δύο σημαντικές ανισότητες, εκείνες του Markov και του Chebyshev, οι οποίες, πέραν της μεγάλης χρησιμότητάς τους σε εφαρμογές, δίνουν και μια σαφή χρηστική σημασία στις έννοιες αυτές. Ξεκινάμε αυτή την ενότητα με τις γενικεύσεις αυτών των δύο σημαντικών αποτελεσμάτων για συνεχείς τυχαίες μεταβλητές.

Η ανισότητα του Markov λέει πως, αν μια τυχαία μεταβλητή έχει μικρή μέση τιμή, τότε δεν μπορεί να παίρνει μεγάλες τιμές με μεγάλη πιθανότητα. Η απόδειξη είναι παρόμοια με εκείνη που είδαμε στην περίπτωση διακριτών τυχαίων μεταβλητών (βλ. Άσκηση 11 στο τέλος του κεφαλαίου).

Θεώρημα 11.3 (Ανισότητα του Markov) Έστω μια συνεχής Τ.Μ. X που παίρνει πάντα τιμές μεγαλύτερες ή ίσες του μηδενός και έχει μέση τιμή $\mu = E(X)$. Τότε:

$$\Pr(X \geq c) \leq \frac{\mu}{c}, \quad \text{για οποιαδήποτε σταθερά } c > 0.$$

Συνεχίζοντας όπως στο Κεφάλαιο 9, μια απλή εφαρμογή της ανισότητας του Markov μάς δίνει την ανισότητα του Chebychev, η οποία λέει το εξής: Αν μια τυχαία μεταβλητή έχει «μικρή» διασπορά, τότε δεν μπορεί να παίρνει τιμές μακριά από τη μέση τιμή της με μεγάλη πιθανότητα. Η απόδειξη παραλείπεται μια και είναι ακριβώς η ίδια με εκείνη του Θεωρήματος 9.2.

Θεώρημα 11.4 (Ανισότητα του Chebychev) Έστω μια συνεχής Τ.Μ. X με μέση τιμή $\mu = E(X)$ και διασπορά $\sigma^2 = \text{Var}(X)$. Τότε:

$$\Pr(|X - \mu| \geq c) \leq \frac{\sigma^2}{c^2}, \quad \text{για οποιαδήποτε σταθερά } c > 0.$$

Το τελευταίο μεγάλο αποτέλεσμα αυτού του κεφαλαίου είναι ο Νόμος των Μεγάλων Αριθμών (Ν.Μ.Α.) για συνεχείς Τ.Μ., στο Θεώρημα 11.6. Για τη διατύπωσή του χρειαζόμαστε τη γενίκευση της έννοιας της ανεξαρτησίας για συνεχείς τυχαίες μεταβλητές.

Ορισμός 11.3 1. Δύο συνεχείς τυχαίες μεταβλητές X, Y είναι **ανεξάρτητες** αν, για κάθε $a \leq b, a' \leq b'$, τα ενδεχόμενα $\{a \leq X \leq b\}$ και $\{a' \leq Y \leq b'\}$ είναι ανεξάρτητα. Ισοδύναμα, οι X, Y είναι ανεξάρτητες αν και μόνο αν,

$$\Pr(a \leq X \leq b, a' \leq Y \leq b') = \Pr(a \leq X \leq b) \Pr(a' \leq Y \leq b'),$$

για κάθε $a \leq b, a' \leq b'$.

2. Οι συνεχείς Τ.Μ. X_1, X_2, \dots, X_N είναι **ανεξάρτητες** αν,

$$\begin{aligned} \Pr(a_1 \leq X_1 \leq b_1, a_2 \leq X_2 \leq b_2, \dots, a_N \leq X_N \leq b_N) \\ = \Pr(a_1 \leq X_1 \leq b_1) \Pr(a_2 \leq X_2 \leq b_2) \cdots \Pr(a_N \leq X_N \leq b_N), \end{aligned}$$

για κάθε N -άδα ζευγαριών τιμών $a_1 \leq b_1, a_2 \leq b_2, \dots, a_N \leq b_N$.

3. Οι συνεχείς Τ.Μ. X_i σε μια άπειρη ακολουθία X_1, X_2, \dots είναι **ανεξάρτητες** αν οι Τ.Μ. X_1, X_2, \dots, X_N είναι ανεξάρτητες για κάθε $N \geq 1$.

Όπως και στο Κεφάλαιο 9, η απόδειξη του Ν.Μ.Α. στο Θεώρημα 11.6 βασίζεται στον υπολογισμό της μέσης τιμής και της διασποράς ενός αθροίσματος ανεξάρτητων τυχαίων μεταβλητών. Γ' αυτόν τον υπολογισμό θα χρειαστούμε τις παρακάτω ιδιότητες, τις οποίες θα αποδείξουμε στο Κεφάλαιο 15, βλ. Πρόταση 15.1 και Θεώρημα 15.1.

Θεώρημα 11.5 (Μέση τιμή και διασπορά αθροίσματος)

1. Για οποιοσδήποτε συνεχείς Τ.Μ. X, Y και σταθερές $a, b \in \mathbb{R}$:

$$E(aX + bY) = aE(X) + bE(Y).$$

2. Αν οι Τ.Μ. X, Y είναι ανεξάρτητες, τότε:

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y).$$

Παρατήρηση: Όπως και στη διακριτή περίπτωση, και οι δύο ιδιότητες του Θεωρήματος 11.5 εύκολα επεκτείνονται και για οποιοδήποτε (πεπερασμένο) πλήθος N τυχαίων μεταβλητών. Με άλλα λόγια, για οποιοδήποτε (συνεχείς ή διακριτές) τυχαίες μεταβλητές X_1, X_2, \dots, X_N , και οποιοδήποτε σταθερές a_1, a_2, \dots, a_N ,

$$E\left(\sum_{i=1}^N a_i X_i\right) = \sum_{i=1}^N a_i E(X_i). \quad (11.3)$$

Και αν, επιπλέον, οι X_i είναι ανεξάρτητες, τότε,

$$\text{Var}\left(\sum_{i=1}^N X_i\right) = \sum_{i=1}^N \text{Var}(X_i). \quad (11.4)$$

Πριν δούμε το κεντρικό αποτέλεσμα του κεφαλαίου, τον Ν.Μ.Α. για συνεχείς Τ.Μ., θα αποδείξουμε μια ανισότητα που αποτελεί ένα εξαιρετικά σημαντικό εργαλείο στις πιθανότητες. Ξεκινάμε παρατηρώντας πως, από το προφανές γεγονός ότι η διασπορά μιας Τ.Μ. είναι πάντα μεγαλύτερη ή ίση του μηδενός, έχουμε ότι,

$$0 \leq \text{Var}(X) = E(X^2) - [E(X)]^2,$$

από το οποίο προκύπτει πως,

$$[E(X)]^2 \leq E(X^2),$$

για οποιαδήποτε Τ.Μ. X . Η ανισότητα Cauchy-Schwarz, την οποία αποδεικνύουμε αμέσως μετά, μπορεί να θεωρηθεί μια γενίκευση αυτού του αποτελέσματος για την περίπτωση του γινομένου δύο τυχαίων μεταβλητών.

Πρόταση 11.1 (Ανισότητα Cauchy-Schwarz) Για δύο οποιοδήποτε (συνεχείς ή διακριτές) Τ.Μ. X, Y , έχουμε:

$$[E(XY)]^2 \leq E(X^2)E(Y^2).$$

Απόδειξη: Η απόδειξη είναι πολύ χαριτωμένη και εντυπωσιακά απλή. Για οποιοδήποτε $z \in \mathbb{R}$, η Τ.Μ. $(X + zY)^2$ παίρνει πάντα τιμές μεγαλύτερες ή ίσες του μηδενός, άρα και η μέση τιμή της θα είναι $E[(X + zY)^2] \geq 0$. Οπότε, αναπτύσσοντας το τετράγωνο και χρησιμοποιώντας την πρώτη ιδιότητα της μέσης τιμής από τα Θεωρήματα 6.1 και 11.2,

$$0 \leq E[(X + zY)^2] = E[z^2Y^2 + 2zXY + X^2] = z^2E(Y^2) + 2zE(XY) + E(X^2).$$

Έχουμε λοιπόν ένα τριώνυμο της μορφής $az^2 + bz + c$, το οποίο είναι μεγαλύτερο ή ίσο του μηδενός για όλες τις τιμές του $z \in \mathbb{R}$, δηλαδή ή δεν έχει καμία ρίζα ή έχει μία διπλή. Συνεπώς η διακρίνουσά του $\Delta = b^2 - 4ac$ πρέπει να είναι $\Delta \leq 0$, δηλαδή πρέπει να έχουμε,

$$[2E(XY)]^2 - 4E(X^2)E(Y^2) \leq 0,$$

η οποία είναι ακριβώς η ζητούμενη ανισότητα. ■

Κλείνουμε αυτό το κεφάλαιο με τη διατύπωση του Ν.Μ.Α. για συνεχείς Τ.Μ. Η απόδειξή του είναι ακριβώς ίδια με εκείνη του Θεωρήματος 9.3 στη διακριτή περίπτωση, και γι' αυτόν τον λόγο παραλείπεται.

Θεώρημα 11.6 (Νόμος των Μεγάλων Αριθμών) Έστω μια ακολουθία από ανεξάρτητες συνεχείς τυχαίες μεταβλητές X_1, X_2, \dots που έχουν όλες την ίδια κατανομή, δηλαδή την ίδια πυκνότητα $f(x)$, και κατά συνέπεια την ίδια μέση τιμή $\mu = E(X_i)$ και την ίδια διασπορά $\sigma^2 = \text{Var}(X_i) < \infty$. Τότε:

1. [Διαισθητικά] Για μεγάλα N , ο εμπειρικός μέσος όρος των X_1, X_2, \dots, X_N ,

$$\bar{X}_N = \frac{1}{N} \sum_{i=1}^N X_i \approx \mu, \quad \text{με μεγάλη πιθανότητα.}$$

2. [Μαθηματικά] Καθώς το $N \rightarrow \infty$ ο εμπειρικός μέσος όρος \bar{X}_N τείνει στη μέση τιμή μ κατά πιθανότητα, δηλαδή: Για κάθε $\epsilon > 0$, έχουμε,

$$\Pr(|\bar{X}_N - \mu| < \epsilon) \rightarrow 1, \quad \text{καθώς το } N \rightarrow \infty.$$

Τέλος, όπως αναφέραμε και στο Κεφάλαιο 9, σημειώνουμε πως, παρότι στην παραπάνω διατύπωση του Ν.Μ.Α. στο Θεώρημα 11.6 υποθέσαμε πως οι συνεχείς Τ.Μ. X_i έχουν πεπερασμένη διασπορά $\sigma^2 = \text{Var}(X_i) < \infty$, αυτή η υπόθεση δεν είναι απαραίτητη και μπορεί να αντικατασταθεί από την ασθενέστερη υπόθεση ότι η μέση τιμή $\mu_0 = E(|X_i|)$ είναι πεπερασμένη. Βέβαια, σε αυτήν την περίπτωση απαιτείται μια διαφορετική απόδειξη, αρκετά πιο μακροσκελής και απαιτητική από μαθηματική άποψη.

11.4 Ασκήσεις

1. Από την ομοιόμορφη στην εκθετική κατανομή. Έστω μια Τ.Μ. U ομοιόμορφα κατανομημένη στο $[0, 1]$. Αν $\lambda > 0$, δείξτε πως η νέα Τ.Μ. $Y = -\lambda \log U$ έχει κατανομή $\text{Eκθ}(\lambda)$.
2. Χρόνος μετάδοσης. Ο χρόνος T που μεσολαβεί από την αποστολή μέχρι την παράδοση ενός email μεγέθους X KB από ένα διακομιστή σε έναν άλλο είναι $T = X(X+1)/4$ λεπτά. Αν το X έχει εκθετική κατανομή με μέση τιμή 5 KB, να βρεθεί:
 - (α') Η μέση τιμή του χρόνου T .
 - (β') Η πιθανότητα ο χρόνος εκτέλεσης να ξεπεράσει τα 10 λεπτά.
3. Μηνυματάκια! Η διάρκεια αποστολής ενός SMS σε δευτερόλεπτα έχει ομοιόμορφη κατανομή στο διάστημα $[1, 3]$, και η διάρκεια αποστολής ενός MMS έχει εκθετική κατανομή με μέση τιμή τα 8 δευτερόλεπτα. Στέλνουμε, σε τρεις ανεξάρτητες αποστολές, 2 SMS και ένα MMS.
 - (α') Ποια είναι η μέση τιμή της συνολικής διάρκειας αποστολής;
 - (β') Ποια είναι η πιθανότητα και τα 2 SMS να έχουν διάρκεια πάνω από 2 δευτερόλεπτα το καθένα;
 - (γ') Ποια είναι η πιθανότητα το MMS να έχει διάρκεια μεγαλύτερη από τη μέση τιμή της συνολικής διάρκειας των 2 SMS;
 - (δ') Δεδομένου ότι το MMS έχει διάρκεια πάνω από 10 δευτερόλεπτα, ποια η πιθανότητα να διαρκέσει μεταξύ 10 και 20 δευτερόλεπτα;
4. Ελάχιστο και μέγιστο δύο Τ.Μ. Έστω οι ανεξάρτητες Τ.Μ. X, Y με συναρτήσεις κατανομής, αντίστοιχα, $F_X(x), F_Y(y)$. Έστω επίσης οι τυχαίες μεταβλητές $W = \max\{X, Y\}$ και $V = \min\{X, Y\}$.
 - (α') Να υπολογίσετε τις συναρτήσεις κατανομής $F_W(w)$ και $F_V(v)$ των W και V , συναρτήσει των $F_X(x), F_Y(y)$.
 - (β') Αν η $X \sim \text{Eκθ}(\theta)$ και η $Y \sim \text{Eκθ}(\phi)$, ποια είναι η κατανομή του $V = \min\{X, Y\}$; Συγκρίνετε το αποτέλεσμά σας με αυτό της Άσκησης 4 του Κεφαλαίου 7.
5. Ελάχιστο N Τ.Μ. Έστω N ανεξάρτητες Τ.Μ. X_1, X_2, \dots, X_N , όλες με την ίδια συνάρτηση κατανομής $F_X(x)$, και έστω $V = \min\{X_1, X_2, \dots, X_N\}$.
 - (α') Βρείτε τη συνάρτηση κατανομής $F_V(v)$ της V συναρτήσει της $F_X(x)$.
 - (β') Στην ειδική περίπτωση που όλες οι X_i έχουν κατανομή $\text{Eκθ}(\theta)$, ποια είναι η κατανομή της V ;

6. *Γραμμικός μετασχηματισμός.* Διατυπώστε και αποδείξτε την Ιδιότητα 3 του Θεωρήματος 11.2 στην περίπτωση που η σταθερά a είναι αρνητική.
7. *Κανονικοποίηση.* Αποδείξτε το Πρόγραμμα 11.1: (α') για την περίπτωση που η Τ.Μ. X είναι διακριτή, και (β') για την περίπτωση που είναι συνεχής.
8. *Το φράγμα του Chernoff για συνεχείς Τ.Μ.* Όπως στην Άσκηση 7 του Κεφαλαίου 9, εδώ θα δούμε ένα φράγμα για την πιθανότητα μια συνεχής Τ.Μ. X να παίρνει τιμές μεγαλύτερες από κάποια σταθερά $c \in \mathbb{R}$, ακριβέστερο απ' αυτό που μας δίνει η ανισότητα του Markov.

(α') Αποδείξτε πως, για οποιοδήποτε $\lambda > 0$, έχουμε το φράγμα:

$$\Pr(X \geq c) \leq e^{-\lambda c} E(e^{\lambda X}). \quad (11.5)$$

Παρατηρήστε πως, ενώ το αριστερό μέρος δεν εξαρτάται από το λ , το δεξί μέρος ισχύει για κάθε λ .

(β') Έστω τώρα πως η X έχει κατανομή $\text{Ex}(\theta)$. Υπολογίστε τη μέση τιμή $E(e^{\lambda X})$, όπου η λ είναι μια σταθερά με $0 < \lambda < 1/\theta$, και χρησιμοποιώντας τη σχέση (11.5) δείξτε πως, για οποιοδήποτε $c \geq 0$ και $0 < \lambda < 1/\theta$, έχουμε:

$$\Pr(X \geq c) \leq \frac{e^{-\lambda c}}{1 - \lambda\theta}. \quad (11.6)$$

(γ') Έστω πάλι πως $X \sim \text{Ex}(\theta)$. Για $\theta = 1$ και $c = 10$:

- i. Υπολογίστε την τιμή του $\lambda \in (0, 1)$ που μας δίνει το καλύτερο (δηλαδή το μικρότερο) δυνατό φράγμα στη σχέση (11.6), καθώς και το αντίστοιχο φράγμα γι' αυτήν τη βέλτιστη τιμή του λ .
- ii. Υπολογίστε το αντίστοιχο φράγμα που μας δίνει η ανισότητα του Markov. Παρατηρήστε ότι είναι ασθενέστερο.
- iii. Υπολογίστε ακριβώς την πιθανότητα $\Pr(X \geq c)$ και συγκρίνετε την πραγματική τιμή της με τα δύο παραπάνω φράγματα.

9. *Γραμμικοί συνδυασμοί.* Οι τυχαίες μεταβλητές X και Y είναι ανεξάρτητες και έχουν εκθετική κατανομή με μέση τιμή 2. Έστω δύο νέες τυχαίες μεταβλητές $A = 2X + Y$ και $B = 2X - Y$. Να βρεθούν τα ακόλουθα: $E(A)$, $E(B)$, $\text{Var}(X)$, $\text{Var}(Y)$, $E(AB)$.
10. *Συνάρτηση μιας εκθετικής Τ.Μ.* Το κόστος συντήρησης ενός δικτύου κινητής τηλεφωνίας με συνολική έκταση $X \text{ km}^2$ είναι $Y = X^3 + 5$ χιλιάδες ευρώ. Αν η Τ.Μ. X έχει εκθετική κατανομή με παράμετρο 20, να βρεθεί η πυκνότητα του Y .
11. *Ανισότητα Markov.* Να αποδείξετε την ανισότητα Markov του Θεωρήματος 11.3 στην περίπτωση που το σύνολο τιμών S της τυχαίας μεταβλητής X είναι το $[0, \infty)$.

[Παρατήρηση. Η υπόθεση $S_X = [0, \infty)$ γίνεται απλώς για να σας διευκολύνει. Η απόδειξη της γενικής περίπτωσης όπου το S_X είναι κάποιο υποσύνολο του $[0, \infty)$ είναι ακριβώς η ίδια, απλά απαιτεί μεγαλύτερη προσοχή στο συμβολισμό. Υπόδειξη. Ακολουθήστε τα ίδια βήματα όπως στην απόδειξη της ανισότητας Markov για διακριτές Τ.Μ. στο Θεώρημα 9.1 του Κεφαλαίου 9.]

12. *Ανισότητα Chebychev.* Για μια διακριτή Τ.Μ. αποδείξαμε την ανισότητα Chebychev στο Θεώρημα 9.2. Εξηγήστε με ποιο τρόπο τεκμηριώνονται τα βήματα της απόδειξης στην περίπτωση που η Τ.Μ. είναι συνεχής.
13. *Πόσο ακριβής είναι η ανισότητα Chebychev;* Συγκρίνετε το φράγμα που δίνει η ανισότητα Chebychev για την πιθανότητα του ενδεχόμενου $\{|X - \mu| \geq c\}$ με την ακριβή πιθανότητα, στην περίπτωση που η X είναι ομοιόμορφα κατανομημένη στο διάστημα $[-k, k]$. Σχολιάστε πόσο ακριβές είναι το φράγμα, δηλαδή πόσο κοντά είναι στην πραγματική τιμή της πιθανότητας.
14. *Απόδειξη του N.M.A.* Για μια ακολουθία διακριτών Τ.Μ. αποδείξαμε τον N.M.A. στο Θεώρημα 9.3. Εξηγήστε με ποιόν τρόπο τεκμηριώνονται τα βήματα της απόδειξης στην περίπτωση που οι Τ.Μ. είναι συνεχείς.
15. *Συνεχείς και διακριτές Τ.Μ.* Θυμηθείτε τις συνεχείς Τ.Μ. Y_1 και Y_2 τις οποίες ορίσαμε στην Άσκηση 10 του προηγούμενου κεφαλαίου. Είναι ή όχι οι Y_1, Y_2 ανεξάρτητες; Αποδείξτε την απάντησή σας και εξηγήστε τη διαισθητικά.
16. *Άλλος ένας γραμμικός μετασχηματισμός.* Έστω μια Τ.Μ. X με ομοιόμορφη κατανομή στο $[0, 1]$, και έστω η νέα Τ.Μ. $Y = 10X + 5$. Ποια είναι η συνάρτηση κατανομής της Y ;
17. *Συνδιακύμανση και συσχέτιση.* Έστω δύο οποιεσδήποτε διακριτές Τ.Μ. X και Y .

(α') Αποδείξτε πως για τη συνδιακύμανσή τους έχουμε:

$$\text{Cov}(X, Y) \leq \sqrt{\text{Var}(X)\text{Var}(Y)}.$$

(β') Αποδείξτε πως ο συντελεστής συσχέτισης $\rho_{X,Y}$ που ορίσαμε στην Άσκηση 10 του Κεφαλαίου 9 πάντα ικανοποιεί $|\rho_{X,Y}| \leq 1$, καθώς και ότι υπάρχουν περιπτώσεις στις οποίες πράγματι ισούται με τις ακραίες τιμές ± 1 .

Κεφάλαιο 12

Το Κεντρικό Οριακό Θεώρημα

Όπως περιγράψαμε λεπτομερώς στο Κεφάλαιο 9, ο Νόμος των Μεγάλων Αριθμών είναι ένα από τα πιο θεμελιώδη αποτελέσματα της θεωρίας των πιθανοτήτων (βλ. τα Θεωρήματα 9.3 και 11.6 για τη διατύπωση του Ν.Μ.Α. για διακριτές και συνεχείς τυχαίες μεταβλητές αντίστοιχα). Συγκεκριμένα, αν X_1, X_2, \dots είναι μια ακολουθία από ανεξάρτητες (συνεχείς ή διακριτές) Τ.Μ. που έχουν όλες την ίδια κατανομή, τότε ο Ν.Μ.Α. μας λέει πως, για μεγάλα N , ο εμπειρικός μέσος όρος \bar{X}_N των πρώτων N απ' τα X_i θα ισούται κατά προσέγγιση με την μέση τιμή τους μ , με μεγάλη πιθανότητα.

Για παράδειγμα, έστω πως έχουμε N ανεξάρτητα, τυχαία δείγματα X_1, X_2, \dots, X_N , όπου το κάθε X_i αντιστοιχεί στο αποτέλεσμα της μέτρησης της θερμοκρασίας μιας μεγάλης δεξαμενής. Αν θέλουμε να εκτιμήσουμε την (άγνωστη σε μας) μέση θερμοκρασία, έστω μ , του υγρού που περιέχει η δεξαμενή, τότε ο Ν.Μ.Α. μας διαβεβαιώνει πως ο εμπειρικός μέσος όρος των δειγμάτων X_i ,

$$\bar{X}_N = \frac{1}{N} \sum_{i=1}^n X_i,$$

θα μας δώσει μια καλή εκτίμηση για τη μ , εφόσον το πλήθος N των δειγμάτων είναι αρκετά μεγάλο. Φυσικά για να είναι πρακτικά χρήσιμο αυτό το αποτέλεσμα είναι απαραίτητο να δώσουμε μια ακριβέστερη ποσοτική περιγραφή του τι θα πει «καλή εκτίμηση» και πόσο «αρκετά μεγάλο» πρέπει να είναι το πλήθος N των δειγμάτων. Αυτό είναι το αντικείμενο του παρόντος κεφαλαίου.

Από μαθηματική άποψη, ο Ν.Μ.Α. μας λέει πως ο εμπειρικός μέσος όρος \bar{X}_N τείνει στη μέση τιμή μ , καθώς το N τείνει στο άπειρο. Το Κεντρικό Οριακό Θεώρημα (Κ.Ο.Θ.) που θα δούμε στην Ενότητα 12.2 είναι μια ακριβέστερη μορφή αυτού του αποτελέσματος: Λέει πως, για μεγάλα N , η κατανομή του εμπειρικού μέσου όρου \bar{X}_N μπορεί να προσεγγιστεί από τη λεγόμενη «κανονική κατανομή με παραμέτρους μ και σ^2/N », όπου σ^2 είναι η διασπορά των X_i .

Συνεπώς, γνωρίζοντας (κατά προσέγγιση) την κατανομή του \bar{X}_N , μπορούμε να υπολογίσουμε την πιθανότητα το \bar{X}_N να είναι σχετικά κοντά ή μακριά από την τιμή μ στην οποία τείνει. Λόγου χάρη, στο πιο πάνω παράδειγμα, ποια είναι η πιθανότητα η εκτίμησή μας \bar{X}_N για τη μέση θερμοκρασία μ , να απέχει από την πραγματική τιμή κατά το πολύ έναν βαθμό; Αν ορίσουμε μια Τ.Μ. Z

με την κατανομή που μας δίνει το Κ.Ο.Θ., θα έχουμε την προσέγγιση,

$$\Pr(|\bar{X}_N - \mu| \leq 1) \approx \Pr(|Z - \mu| \leq 1) = \Pr(\mu - 1 \leq Z \leq \mu + 1),$$

όπου η τελευταία πιθανότητα μπορεί εύκολα να υπολογιστεί, αφού η Τ.Μ. Z έχει γνωστή κατανομή.

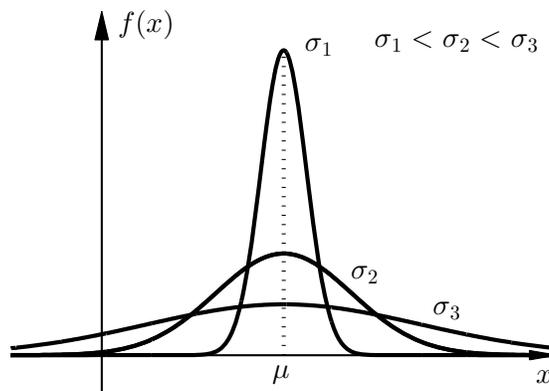
Στην Ενότητα 12.1 θα ορίσουμε την οικογένεια των τυχαίων μεταβλητών με κανονική κατανομή και θα εξετάσουμε τις βασικές τους ιδιότητες. Στη συνέχεια, στην Ενότητα 12.2 θα διατυπώσουμε το Κ.Ο.Θ. και στην 12.3 θα δούμε αρκετά παραδείγματα του πώς χρησιμοποιείται στην πράξη. Στο Κεφάλαιο 13 θα δούμε κάποια θεωρητικά αποτελέσματα που σχετίζονται με το Κ.Ο.Θ., συμπεριλαμβανομένων των βασικών στοιχείων της απόδειξής του, και τέλος στο Κεφάλαιο 14 θα παρουσιάσουμε τις κεντρικές ιδέες του τρόπου με τον οποίο το Κ.Ο.Θ. εφαρμόζεται σε κάποια από τα θεμελιώδη προβλήματα της στατιστικής.

12.1 Κανονική κατανομή

Ορισμός 12.1 Μια συνεχής Τ.Μ. X έχει κανονική κατανομή (ή γκαουσιανή κατανομή) με παραμέτρους $\mu \in \mathbb{R}$ και $\sigma^2 > 0$, αν έχει σύνολο τιμών το $S = \mathbb{R}$ και πυκνότητα,

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}, \quad \text{για } x \in \mathbb{R}, \quad (12.1)$$

βλ. Σχήμα 12.1. Για συντομία, αυτό συμβολίζεται: $X \sim N(\mu, \sigma^2)$.



Σχήμα 12.1: Γραφική αναπαράσταση της πυκνότητας $f(x)$ μιας Τ.Μ. X με κατανομή $N(\mu, \sigma^2)$, για τρεις διαφορετικές τιμές της παραμέτρου σ^2 . Η πυκνότητα έχει την κορυφή της στο σημείο $x = \mu$, και είναι συμμετρική γύρω από τον άξονα $x = \mu$.

Παρατηρήσεις

1. Αν και, εκ πρώτης όψεως, η πολύπλοκη μορφή της πυκνότητας (12.1) ίσως μας ξενίζει, είναι σημαντικό να σημειώσουμε πως η κανονική κατανομή αντλεί τη σημασία της από το Κεντρικό Οριακό Θεώρημα, το οποίο μας λέει πως η κανονική κατανομή είναι, κατά κάποιον τρόπο, αναπόφευκτη! Όταν εξετάζουμε τον εμπειρικό μέσο όρο \bar{X}_N ενός σχετικά μεγάλου πλήθους τυχαίων μεταβλητών X_i , η κατανομή του \bar{X}_N τείνει πάντοτε στην κανονική κατανομή, ανεξάρτητα του ποια κατανομή έχουν τα μεμονωμένα X_i .
2. Για να βεβαιωθούμε πως ο παραπάνω ορισμός είναι μαθηματικά ορθός σύμφωνα με τον Ορισμό 10.1, θα έπρεπε να ελέγξουμε ότι το ολοκλήρωμα (από το $-\infty$ ως το $+\infty$) της πυκνότητας $f(x)$ στη σχέση (12.1) ισούται με 1, για κάθε ζευγάρι τιμών των παραμέτρων μ και σ^2 . Αυτό φυσικά πράγματι ισχύει, αλλά ο σχετικός υπολογισμός, τον οποίο δίνουμε στην Ενότητα 13.1 του επόμενου κεφαλαίου, είναι κάπως τεχνικός και μακροσκελής.
3. Έστω πως, για μια Τ.Μ. X με κατανομή $X \sim N(\mu, \sigma^2)$, θέλουμε να υπολογίσουμε την πιθανότητα $\Pr(a \leq X \leq b)$. Αν και μπορεί να εκφραστεί ως το ολοκλήρωμα της πυκνότητας $f(x)$ στη σχέση (12.1),

$$\Pr(a \leq X \leq b) = \int_a^b \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} dx,$$

αυτό το ολοκλήρωμα δεν μπορεί να εκφραστεί σε κλειστή μορφή. Γί' αυτόν τον λόγο, θα δείξουμε στις πιο κάτω ιδιότητες πως οποιαδήποτε πιθανότητα που αφορά μια $N(\mu, \sigma^2)$ κατανομή μπορεί να εκφραστεί ως μια αντίστοιχη πιθανότητα για τη λεγόμενη «τυπική κανονική κατανομή» $N(0, 1)$. Βλ. τη «γενική μέθοδο» που δίνεται μετά το Θεώρημα 12.1 και τα παραδείγματα που ακολουθούν.

Θεώρημα 12.1 (Ιδιότητες της κανονικής κατανομής)

Έστω $X \sim N(\mu, \sigma^2)$. Η X έχει τις εξής ιδιότητες:

1. **Μέση τιμή:** $E(X) = \mu$.
2. **Διασπορά:** $\text{Var}(X) = \sigma^2$.

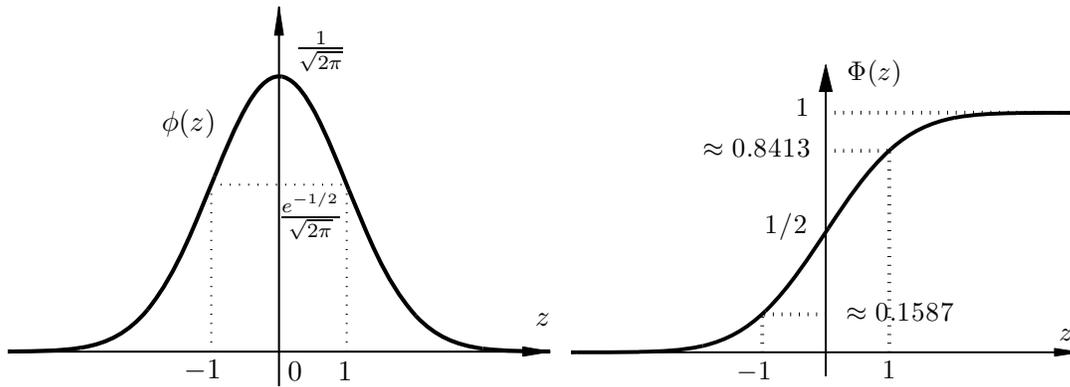
Το γεγονός πως οι παράμετροι μ και σ^2 αντιστοιχούν στη μέση τιμή και τη διασπορά της κατανομής $N(\mu, \sigma^2)$ αναπαρίσταται και γραφικά στο Σχήμα 12.1.

3. **Κανονικοποίηση:** Η Τ.Μ. $Z = \frac{X-\mu}{\sigma}$ έχει κατανομή $Z \sim N(0, 1)$.
4. **Γραμμικός συνδυασμός:** Αν οι τυχαίες μεταβλητές $X \sim N(\mu, \sigma^2)$ και $Y \sim N(\nu, \tau^2)$ είναι ανεξάρτητες, τότε η Τ.Μ. $aX + bY$ έχει κατανομή $N(a\mu + b\nu, a^2\sigma^2 + b^2\tau^2)$.

5. Συμβολισμός: Αν η X έχει $\mu = 0$ και $\sigma^2 = 1$, δηλαδή $X \sim N(0, 1)$, τότε λέμε πως έχει τυπική κανονική κατανομή και συμβολίζουμε την πυκνότητα της και τη συνάρτηση κατανομής της αντίστοιχα ως,

$$\begin{aligned}\phi(x) &= \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, & x \in \mathbb{R}, \\ \Phi(z) &= \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx, & z \in \mathbb{R}.\end{aligned}\quad (12.2)$$

Η γραφική αναπαράσταση της πυκνότητας $\phi(x)$ και της συνάρτησης κατανομής $\Phi(z)$ της τυπικής κανονικής κατανομής $N(0, 1)$ δίνονται στο Σχήμα 12.2.



Σχήμα 12.2: Γραφική αναπαράσταση της πυκνότητας $\phi(z)$ και της συνάρτησης κατανομής $\Phi(z)$ της τυπικής κανονικής κατανομής $N(0, 1)$.

Απόδειξη: Η μέση τιμή της X υπολογίζεται απλά μέσω του ορισμού. Αρχικά έχουμε,

$$\begin{aligned}E(X) &= \int_{-\infty}^{\infty} x f(x) dx \\ &= \int_{-\infty}^{\infty} (x - \mu + \mu) f(x) dx \\ &= \int_{-\infty}^{\infty} (x - \mu) f(x) dx + \mu \int_{-\infty}^{\infty} f(x) dx \\ &= \sigma^2 \int_{-\infty}^{\infty} \frac{(x - \mu)}{\sigma^2} f(x) dx + \mu,\end{aligned}$$

αλλά παρατηρούμε πως η συνάρτηση του τελευταίου ολοκληρώματος σχετίζεται με την παράγωγο

της πυκνότητας,

$$f'(x) = \frac{d}{dx} \left[-\frac{(x-\mu)^2}{2\sigma^2} \right] \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} = -\frac{(x-\mu)}{\sigma^2} f(x),$$

οπότε αντικαθιστώντας βρίσκουμε,

$$\begin{aligned} E(X) &= \mu + \sigma^2 \int_{-\infty}^{\infty} (-f'(x)) dx \\ &= \mu - \sigma^2 [f(x)]_{-\infty}^{\infty} = \mu, \end{aligned}$$

αποδεικνύοντας την πρώτη ιδιότητα. Για τη δεύτερη ιδιότητα, χρησιμοποιώντας και πάλι την έκφραση για την παράγωγο της $f(x)$,

$$\begin{aligned} \text{Var}(X) &= E[(X - \mu)^2] \\ &= \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx \\ &= \sigma^2 \int_{-\infty}^{\infty} (x - \mu) \frac{(x - \mu)}{\sigma^2} f(x) dx \\ &= \sigma^2 \int_{-\infty}^{\infty} (x - \mu)(-f'(x)) dx \\ &\stackrel{(a)}{=} [-\sigma^2(x - \mu)f(x)]_{-\infty}^{\infty} + \sigma^2 \int_{-\infty}^{\infty} f'(x) dx \\ &= 0 + \sigma^2 \cdot 1 = \sigma^2, \end{aligned}$$

όπου στο βήμα (a) ολοκληρώσαμε κατά παράγοντες, θέτοντας $u = x - \mu$ και $dv = f'(x)dx$, έτσι ώστε $du = dx$ και $v = f(x)$.

Για την τρίτη ιδιότητα, έστω $f(x)$ η πυκνότητα της X όπως στην (12.1). Εφόσον $Z = \frac{1}{\sigma}X - \frac{\mu}{\sigma}$, τότε από την τρίτη ιδιότητα του Θεωρήματος 11.2 με $a = 1/\sigma$ και $b = -\mu/\sigma$, η Z έχει πυκνότητα,

$$\frac{1}{1/\sigma} f\left(\frac{x - (-\mu/\sigma)}{1/\sigma}\right) = \sigma f(\sigma x + \mu) = \sigma \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(\sigma x + \mu)^2/2\sigma^2} = \frac{1}{\sqrt{2\pi}} e^{-x^2/2},$$

η οποία είναι πράγματι η πυκνότητα της $N(0, 1)$ κατανομής.

Η τέταρτη ιδιότητα θα αποδειχθεί στο Πρόρισμα 15.1 του Κεφαλαίου 15. Για την ειδική περίπτωση όπου η Y είναι απλά μια σταθερά, δείτε την Άσκηση 7 στο τέλος του κεφαλαίου. ■

Γενική μέθοδος υπολογισμού $N(\mu, \sigma^2)$ πιθανοτήτων. Εφόσον, όπως σημειώσαμε πιο πάνω, ολοκληρώματα της μορφής $\int_a^b f(x) dx$ για την κανονική πυκνότητα (12.1) δεν μπορούν να υπολογιστούν σε κλειστή μορφή, θα χρησιμοποιήσουμε την Ιδιότητα 3 του Θεωρήματος 12.1 ως εξής. Έστω πως θέλουμε να υπολογίσουμε κάποια πιθανότητα της μορφής $\Pr(a \leq X \leq b)$ για μια Γ.Μ. $X \sim N(\mu, \sigma^2)$:

1. Ορίζουμε μια νέα Τ.Μ. $Z \sim N(0, 1)$ και εκφράζουμε τη ζητούμενη πιθανότητα ως,

$$\Pr(a \leq X \leq b) = \Pr\left(\frac{a - \mu}{\sigma} \leq \frac{X - \mu}{\sigma} \leq \frac{b - \mu}{\sigma}\right) = \Pr\left(\frac{a - \mu}{\sigma} \leq Z \leq \frac{b - \mu}{\sigma}\right).$$

Π.χ., αν η ζητούμενη πιθανότητα ήταν $\Pr(0 \leq X \leq 1.5)$ για την $X \sim N(1, 3)$, τότε,

$$\Pr(0 \leq X \leq 1) = \Pr\left(\frac{0 - 1}{\sqrt{3}} \leq \frac{X - \mu}{\sigma} \leq \frac{1.5 - 1}{\sqrt{3}}\right) \approx \Pr(-0.5774 \leq Z \leq 0.2887).$$

2. Εφόσον η Z έχει συνάρτηση κατανομής $\Phi(z)$ όπως στην (12.2), χρησιμοποιώντας τη γενική ιδιότητα « $\Pr(a \leq X \leq b) = F(b) - F(a)$ » που είδαμε στη (10.8), έχουμε,

$$\Pr(a \leq X \leq b) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right),$$

οπότε στο συγκεκριμένο πιο πάνω παράδειγμα,

$$\Pr(0 \leq X \leq 1) \approx \Phi(0.2887) - \Phi(-0.5774).$$

3. Τέλος, αντικαθιστούμε τις τιμές $\Phi(a)$ και $\Phi(b)$ από τους πίνακες τιμών της τυπικής κανονικής συνάρτησης κατανομής $\Phi(z)$, που δίνονται στο τέλος του κεφαλαίου. Η κάθε γραμμή περιέχει το ακέραιο μέρος και το πρώτο δεκαδικό ψηφίο του z και η κάθε στήλη το δεύτερο δεκαδικό του ψηφίο.

Παραδείγματος χάρι, για να βρούμε την τιμή $\Phi(-0.5774) \approx \Phi(-0.58)$, πάμε στον Πίνακα 12.1 που περιέχει τις τιμές της $\Phi(z)$ για αρνητικά z , και κοιτάζουμε στην 33^η γραμμή, που αντιστοιχεί στα $z = -0.5 \dots$, και στην ένατη στήλη, που περιέχει την τιμή της $\Phi(-0.58) \approx 0.2810$. Παρομοίως, από την τρίτη γραμμή και τελευταία στήλη του Πίνακα 12.2 βρίσκουμε την τιμή της $\Phi(0.2887) \approx \Phi(0.29) \approx 0.6164$. Οπότε τελικά έχουμε,

$$\Pr(0 \leq X \leq 1) \approx \Phi(0.29) - \Phi(-0.58) \approx 0.61640 - 0.2810 = 0.3354.$$

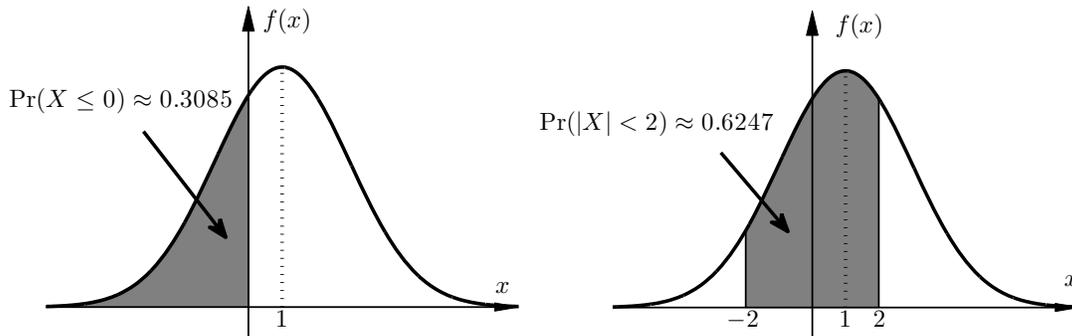
Παράδειγμα 12.1 Έστω πως το ρολόι (clock) ενός επεξεργαστή αποκλίνει κατά X δευτερόλεπτα μετά από ένα χρόνο συνεχούς λειτουργίας, όπου $X \sim N(1, 4)$. Ποια η πιθανότητα μετά από ένα χρόνο:

1. το ρολόι να πηγαίνει «πίσω», δηλαδή να έχει αρνητική απόκλιση;
2. η απόκλιση να είναι μεγαλύτερη από +3.5 δευτερόλεπτα;
3. η απόκλιση να είναι μικρότερη από ± 2 δευτερόλεπτα;

Χρησιμοποιώντας τη γενική παραπάνω μέθοδο με $\mu = 1$ και $\sigma = \sqrt{4} = 2$ και ορίζοντας μια τυπική κανονική Τ.Μ. $Z \sim N(0, 1)$, για το πρώτο ερώτημα εύκολα βρίσκουμε,

$$\Pr(X \leq 0) = \Pr\left(\frac{X - 1}{2} \leq \frac{0 - 1}{2}\right) = \Pr(Z \leq -0.5) = \Phi(-0.5),$$

και, αντικαθιστώντας την $\Phi(-0.5) \approx 0.3085$ από τον Πίνακα 12.1 έχουμε $\Pr(X \leq 0) \approx 0.3085$. βλ. Σχήμα 12.3.



Σχήμα 12.3: Γραφική αναπαράσταση του υπολογισμού των πιθανοτήτων στο πρώτο και το τρίτο ερώτημα του Παραδείγματος 12.1, για την Τ.Μ. $X \sim N(1, 4)$ με πυκνότητα $f(x)$.

Παρομοίως, για το δεύτερο ερώτημα,

$$\begin{aligned} \Pr(X > 3.5) &= \Pr\left(\frac{X - 1}{2} > \frac{3.5 - 1}{2}\right) \\ &= \Pr(Z > 1.25) \\ &= 1 - \Pr(Z \leq 1.25) \\ &= 1 - \Phi(1.25) \\ &\approx 1 - 0.8944 \\ &= 0.1056, \end{aligned}$$

όπου αντικαταστήσαμε την τιμή της $\Phi(1.25) \approx 0.8944$ από τον Πίνακα 12.2.

Τέλος, για το τρίτο ερώτημα, χρησιμοποιώντας και πάλι τους Πίνακες 12.1 και 12.2,

$$\begin{aligned} \Pr(|X| < 2) &= \Pr(-2 < X < 2) \\ &= \Pr\left(\frac{-2 - 1}{2} < \frac{X - 1}{2} < \frac{2 - 1}{2}\right) \\ &= \Pr(-1.5 < Z < 0.5) \\ &= \Phi(0.5) - \Phi(-1.5) \\ &\approx 0.6915 - 0.0668 \\ &= 0.6247, \end{aligned}$$

όπου αντικαταστήσαμε τις τιμές $\Phi(0.5) \approx 0.6915$ και $\Phi(-1.5) \approx 0.0668$ · βλ. Σχήμα 12.3.

Παράδειγμα 12.2 Έστω πως η ένταση X του ηλεκτρομαγνητικού θορύβου σε ένα κύκλωμα έχει κατανομή $X \sim N(-2, 3)$. Ποια η πιθανότητα το X να μην ξεπερνάει το 1 σε απόλυτη τιμή;

Όπως και στο προηγούμενο παράδειγμα, εφαρμόζοντας τη γενική παραπάνω μέθοδο με $\mu = -2$ και $\sigma = \sqrt{3}$ και ορίζοντας μια τυπική κανονική Τ.Μ. $Z \sim N(0, 1)$, βρίσκουμε,

$$\begin{aligned} \Pr(|X| \leq 1) &= \Pr(-1 \leq X \leq 1) \\ &= \Pr\left(\frac{-1 - (-2)}{\sqrt{3}} \leq \frac{X - (-2)}{\sqrt{3}} \leq \frac{1 - (-2)}{\sqrt{3}}\right) \\ &= \Pr(0.58 \leq Z \leq 1.73) \\ &= \Phi(1.73) - \Phi(0.58) \\ &\approx 0.9582 - 0.7190 = 0.2392, \end{aligned}$$

όπου αντικαταστήσαμε τις τιμές $\Phi(1.73) \approx 0.9582$ και $\Phi(0.58) \approx 0.7190$ από τον Πίνακα 12.2.

12.2 Το Κεντρικό Οριακό Θεώρημα

Όπως αναφέραμε στην αρχή του κεφαλαίου, πέραν του ότι ο εμπειρικός μέσος όρος,

$$\bar{X}_N = \frac{1}{N} \sum_{i=1}^N X_i,$$

N ανεξάρτητων Τ.Μ. X_i τείνει στη μέση τιμή τους μ σύμφωνα με τον Ν.Μ.Α, για να εφαρμόσουμε αυτό το αποτέλεσμα στην πράξη είναι απαραίτητο να γνωρίζουμε και ποσοτικά πόσο ακριβής είναι η προσέγγιση « $\bar{X}_N \approx \mu$, με μεγάλη πιθανότητα».

Το Κεντρικό Οριακό Θεώρημα (Κ.Ο.Θ.) απαντά σε αυτό ακριβώς το ερώτημα: Μας λέει πως ο εμπειρικός μέσος όρος \bar{X}_N έχει κατά προσέγγιση κανονική $N(\mu, \sigma^2/N)$ κατανομή. Ισοδύναμα, από την τρίτη ιδιότητα του Θεωρήματος 12.1, η Τ.Μ.,

$$\frac{\bar{X}_N - \mu}{\sqrt{\sigma^2/N}} = \frac{\sqrt{N}}{\sigma} \left[\frac{1}{N} \sum_{i=1}^N X_i - \mu \right] = \frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N (X_i - \mu), \quad (12.3)$$

έχει κατά προσέγγιση κανονική $N(0, 1)$ κατανομή.

Επιπλέον, το Κ.Ο.Θ. δεν είναι μόνο στο διαισθητικό επίπεδο μια ακριβέστερη μορφή του Ν.Μ.Α., αλλά όπως θα δούμε στην Ενότητα 13.2, μπορούμε και να αποδείξουμε τον Ν.Μ.Α. από το Κ.Ο.Θ.

Για να διατυπώσουμε το Κ.Ο.Θ. με μαθηματική ακρίβεια, θα χρειαστούμε μια άλλη έννοια για τη σύγκλιση μιας ακολουθίας τυχαίων μεταβλητών X_1, X_2, X_3, \dots , διαφορετική από τη σύγκλιση κατά πιθανότητα, την οποία συναντήσαμε στο Κεφάλαιο 9.

Ορισμός 12.2 Έστω μια ακολουθία Τ.Μ. $\{X_n\} = \{X_1, X_2, \dots\}$ και έστω μια άλλη Τ.Μ. X με συνάρτηση κατανομής $F(x)$. Η ακολουθία $\{X_n\}$ **συγκλίνει κατά κατανομή** στη X αν,

$$\Pr(X_n \leq x) \rightarrow F(x), \quad \text{καθώς το } n \rightarrow \infty,$$

για όλα τα x στα οποία η $F(x)$ είναι συνεχής.

Θεώρημα 12.2 (Κεντρικό Οριακό Θεώρημα) Έστω μια ακολουθία $\{X_n\}$ από ανεξάρτητες (συνεχείς ή διακριτές) τυχαίες μεταβλητές X_1, X_2, \dots , οι οποίες έχουν όλες την ίδια κατανομή και κατά συνέπεια την ίδια μέση τιμή $\mu = E(X_i)$ και την ίδια διασπορά $\sigma^2 = \text{Var}(X_i) < \infty$. Τότε:

[Διαισθητικά-1] Για μεγάλα N , ο εμπειρικός μέσος όρος,

$$\bar{X}_N = \frac{1}{N} \sum_{i=1}^N X_i \quad \text{έχει κατά προσέγγιση } N\left(\mu, \frac{\sigma^2}{N}\right) \text{ κατανομή.}$$

[Διαισθητικά-2] Για μεγάλα N , το κανονικοποιημένο άθροισμα,

$$\bar{S}_N = \frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N (X_i - \mu) \quad \text{έχει κατά προσέγγιση } N(0, 1) \text{ κατανομή.}$$

[Μαθηματικά] Καθώς το $N \rightarrow \infty$, το κανονικοποιημένο άθροισμα \bar{S}_N συγκλίνει κατά κατανομή στην τυπική κανονική κατανομή, δηλαδή: Για κάθε $x \in \mathbb{R}$, έχουμε,

$$\Pr(\bar{S}_N \leq x) = \Pr\left(\frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N (X_i - \mu) \leq x\right) \rightarrow \Phi(x), \quad \text{καθώς το } N \rightarrow \infty, \quad (12.4)$$

όπου $\Phi(x)$ είναι η συνάρτηση κατανομής της $N(0, 1)$ κατανομής.

Η πλήρης απόδειξη του Κ.Ο.Θ. είναι αρκετά τεχνική, και μάλλον ξεφεύγει από τους στόχους του παρόντος βιβλίου. Παρ' όλα αυτά, στο Κεφάλαιο 13 θα δώσουμε μια σειρά από σχετικά αποτελέσματα, συμπεριλαμβανομένης και μιας ολοκληρωμένης απόδειξης του Κ.Ο.Θ. Προς το παρόν σημειώνουμε μόνο πως, σε αντίθεση με τις υποθέσεις του Ν.Μ.Α. στα Θεωρήματα 9.3 και 11.6, εδώ η υπόθεση πως η διασπορά των X_i είναι πεπερασμένη είναι απαραίτητη και δεν μπορεί να παραλειφθεί.

Όπως θα δούμε στα παραδείγματα που ακολουθούν, η βασική χρήση του Κ.Ο.Θ. είναι μέσω της προσέγγισης,

$$\Pr(a \leq \bar{S}_N \leq b) = \Pr\left(a \leq \frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N (X_i - \mu) \leq b\right) \approx \Pr(a \leq Z \leq b) = \Phi(b) - \Phi(a) \quad (12.5)$$

όπου, προφανώς, η Τ.Μ. Z έχει τυπική κανονική κατανομή. Κάποιες πιο ρεαλιστικές και σύγχρονες εφαρμογές του Κ.Ο.Θ. θα εξετάσουμε λεπτομερώς στο Κεφάλαιο 14.

12.3 Παραδείγματα και απλές εφαρμογές

Παράδειγμα 12.3 Έστω πως X_i είναι η θερμοκρασία ενός επεξεργαστή την ημέρα i , όπου θεωρούμε πως τα X_i είναι ανεξάρτητες Τ.Μ. με κατανομή $X_i \sim U[10, 50]$. Ποια η πιθανότητα η μέση θερμοκρασία του επεξεργαστή σε μια περίοδο 3 μηνών (δηλαδή 90 ημερών) να ξεπερνά τους 31 βαθμούς;

Για να εφαρμόσουμε το Κ.Ο.Θ. αρχικά υπολογίζουμε τη μέση τιμή και τη διασπορά των X_i οι οποίες, από τις ιδιότητες της ομοιόμορφης κατανομής, είναι,

$$\mu = E(X_i) = \frac{10 + 50}{2} = 30, \quad \sigma^2 = \text{Var}(X_i) = \frac{(50 - 10)^2}{12} = \frac{400}{3} = 133.3333 \dots,$$

αντίστοιχα. Άρα, η ζητούμενη πιθανότητα μπορεί να εκφραστεί ως προς το κανονικοποιημένο άθροισμα του Κ.Ο.Θ. ως εξής,

$$\Pr\left(\frac{1}{N} \sum_{i=1}^N X_i > 31\right) = \Pr\left(\frac{1}{N} \sum_{i=1}^N (X_i - \mu) > 31 - 30\right) = \Pr\left(\frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N (X_i - \mu) > \frac{\sqrt{N}}{\sigma}\right),$$

με $\sqrt{N}/\sigma = \sqrt{90}/\sqrt{400/3} \approx 0.8216$. Άρα, από το Κ.Ο.Θ., όπως ακριβώς στην περίπτωση της προσέγγισης που δώσαμε στη σχέση (12.5), η ζητούμενη πιθανότητα μπορεί να προσεγγιστεί ως,

$$\Pr\left(\frac{1}{N} \sum_{i=1}^N X_i > 31\right) \approx \Pr(Z > 0.82) = 1 - \Pr(Z \leq 0.82) = 1 - \Phi(0.82),$$

όπου η Z έχει κατανομή $N(0, 1)$. Αντικαθιστώντας την τιμή $\Phi(0.82) \approx 0.7939$ από τον Πίνακα 12.2, τελικά βρίσκουμε πως η ζητούμενη πιθανότητα είναι $\approx 1 - 0.7939 = 0.2061$.

Παράδειγμα 12.4 Από εμπειρικές μετρήσεις γνωρίζουμε πως η διάρκεια εκτέλεσης ενός αλγορίθμου είναι κατά μέσο όρο 17.5 δευτερόλεπτα, με τυπική απόκλιση ± 4 δευτερόλεπτα. Ποια είναι η πιθανότητα, σε 400 διαδοχικές, ανεξάρτητες χρήσεις του αλγορίθμου, ο συνολικός χρόνος εκτέλεσης να μην ξεπερνά τις δύο ώρες;

Έστω X_i ο χρόνος εκτέλεσης του αλγορίθμου τη φορά i για $i = 1, 2, \dots, N = 400$. Από τα δεδομένα του προβλήματος έχουμε πως τα X_i είναι ανεξάρτητες Τ.Μ. με μέση τιμή $\mu = 17.5$ δευτερόλεπτα και τυπική απόκλιση $\sigma = 4$ δευτερόλεπτα. Συνεπώς, η ζητούμενη πιθανότητα μπορεί να εκφραστεί ως προς το κανονικοποιημένο άθροισμα του Κ.Ο.Θ.,

$$\begin{aligned} \Pr\left(\sum_{i=1}^N X_i \leq 2 \times 60 \times 60\right) &= \Pr\left(\sum_{i=1}^N (X_i - \mu) \leq 7200 - 400 \times 17.5\right) \\ &= \Pr\left(\frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N (X_i - \mu) \leq \frac{200}{4 \times \sqrt{400}}\right) \\ &\approx \Pr(Z \leq 2.5), \end{aligned}$$

όπου $Z \sim N(0, 1)$. Και πάλι εφαρμόσαμε το Κ.Ο.Θ. όπως στην προσέγγιση (12.5). Άρα η ζητούμενη πιθανότητα είναι,

$$\Pr\left(\sum_{i=1}^N X_i \leq 7200\right) \approx \Pr(Z \leq 2.5) = \Phi(2.5) \approx 0.9938,$$

με την τιμή $\Phi(2.5) \approx 0.9938$ να έχει αντικατασταθεί από τον Πίνακα 12.2.

Παρατηρήσεις

- Μια καίρια παρατήρηση για το τελευταίο από τα παραπάνω παραδείγματα, είναι πως υπολογίσαμε την πιθανότητα ο μέσος χρόνος εκτέλεσης σε πολλές επαναλήψεις να μην ξεπερνά τις δύο ώρες, χωρίς να γνωρίζουμε την κατανομή του χρόνου μίας εκτέλεσης. Αυτό το φαινόμενο είναι συχνό και πολύ σημαντικό: Το Κ.Ο.Θ. μας επιτρέπει να προσεγγίσουμε την κατανομή του εμπειρικού μέσου όρου ενός μεγάλου πλήθους δειγμάτων, ακόμα κι αν δεν γνωρίζουμε την κατανομή των ίδιων των δειγμάτων. Υπό αυτή την έννοια, η κανονική κατανομή είναι «καθολική», μια και προκύπτει φυσικά και κατά κάποιον τρόπο αναπόφευκτα για τυχαία δείγματα οποιασδήποτε προέλευσης, από την αστρονομία μέχρι την ιατρική.
- Και στα δύο παραπάνω παραδείγματα οι Τ.Μ. X_i ήταν συνεχείς και χρησιμοποιήσαμε το Κ.Ο.Θ. για να προσεγγίσουμε τη (συνεχή) κατανομή του αθροίσματός τους μέσω της (επίσης συνεχούς) κανονικής κατανομής. Όταν οι Τ.Μ. X_i είναι διακριτές, μπορούμε και πάλι να χρησιμοποιήσουμε το Κ.Ο.Θ., αλλά εφόσον προσεγγίζουμε μια διακριτή κατανομή (αυτή του αθροίσματος των X_i ή του εμπειρικού μέσου όρου τους) με μια συνεχή κατανομή (την κανονική), η εφαρμογή του Κ.Ο.Θ. απαιτεί λίγη περισσότερη προσοχή, όπως θα δούμε στα επόμενα παραδείγματα.

Παράδειγμα 12.5 Έστω πως στρίβουμε N ανεξάρτητες φορές ένα νόμισμα με πιθανότητα να φέρει Κορώνα ίση με κάποιο γνωστό $p \in (0, 1)$. Ορίζοντας τις ανεξάρτητες Τ.Μ. X_1, X_2, \dots, X_N ,

όπου η κάθε $X_i \sim \text{Bern}(p)$ περιγράφει το αποτέλεσμα της ρίψης τη φορά i (δηλαδή $X_i = 1$ αν φέραμε Κορώνα, και $X_i = 0$ αν φέραμε Γράμματα), από τον Ν.Μ.Α. έχουμε πως, αν το πλήθος N των ρίψεων είναι μεγάλο, τότε:

$$\text{ποσοστό από } K \text{ στις } N \text{ ρίψεις} = \bar{X}_N = \frac{1}{N} \sum_{i=1}^N X_i \approx \mu, \quad \text{με μεγάλη πιθανότητα.}$$

Το Κ.Ο.Θ. μας επιτρέπει να ποσοτικοποιήσουμε το παραπάνω αποτέλεσμα. Συγκεκριμένα, για κάθε ζευγάρι ακεραίων $m \leq M$, μπορούμε να προσεγγίσουμε την πιθανότητα το συνολικό πλήθος από K να είναι μεταξύ m και M ,

$$\Pr\left(m \leq \sum_{i=1}^N X_i \leq M\right) = \Pr\left(m - \frac{1}{2} \leq \sum_{i=1}^N X_i \leq M + \frac{1}{2}\right).$$

Εφόσον τα m, M και η Τ.Μ. $\sum_{i=1}^N X_i$ παίρνουν όλα μόνο ακεραίες τιμές, οι δύο πιθανότητες είναι ίσες. Ο λόγος για τον οποίο προσθέτουμε ή αφαιρούμε $1/2$ από τα m, M είναι διότι, εφόσον θα προσεγγίσουμε την κατανομή του αθροίσματος αυτού με μια συνεχή κατανομή (την κανονική), λογικό είναι να μοιράσουμε την απόσταση ανάμεσα στις διαδοχικές τιμές $m - 1$ και m , και αντίστοιχα για τις M και $M + 1$.

Σημειώνοντας ότι τα X_i έχουν $\mu = E(X_i) = p$ και $\sigma^2 = \text{Var}(X_i) = p(1 - p)$, μπορούμε τώρα να εφαρμόσουμε το Κ.Ο.Θ. ως εξής,

$$\begin{aligned} & \Pr\left(m \leq \sum_{i=1}^N X_i \leq M\right) \\ &= \Pr\left(m - \frac{1}{2} \leq \sum_{i=1}^N X_i \leq M + \frac{1}{2}\right) \\ &= \Pr\left(m - Np - \frac{1}{2} \leq \sum_{i=1}^N (X_i - p) \leq M - Np + \frac{1}{2}\right) \\ &= \Pr\left(\frac{m - Np - 1/2}{\sqrt{p(1-p)N}} \leq \frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N (X_i - p) \leq \frac{M - Np + 1/2}{\sqrt{p(1-p)N}}\right), \end{aligned}$$

όπου, χρησιμοποιώντας την προσέγγιση που δώσαμε στη σχέση (12.5), η ζητούμενη πιθανότητα μπορεί να προσεγγιστεί ως,

$$\begin{aligned} \Pr\left(m \leq \sum_{i=1}^N X_i \leq M\right) &\approx \Pr\left(\frac{m - Np - 1/2}{\sqrt{p(1-p)N}} \leq Z \leq \frac{M - Np + 1/2}{\sqrt{p(1-p)N}}\right) \\ &= \Phi\left(\frac{M - Np + 1/2}{\sqrt{p(1-p)N}}\right) - \Phi\left(\frac{m - Np - 1/2}{\sqrt{p(1-p)N}}\right). \end{aligned}$$

Παρατηρώντας πως η Τ.Μ. την οποία εξετάσαμε, $\sum_{i=1}^N X_i$ έχει κατανομή Διων(N, p), έχουμε μόλις αποδείξει το ακόλουθο πόρισμα, το οποίο συμπληρώνει το αποτέλεσμα του Πορίσματος 7.1 στο Κεφάλαιο 7, την προσέγγιση Poisson στη διωνυμική κατανομή.

Πόρισμα 12.1 (Κανονική προσέγγιση στη διωνυμική κατανομή)

Αν μια τυχαία μεταβλητή Y έχει κατανομή Διων(N, p) με παραμέτρους τέτοιες ώστε:

- το N είναι αρκετά «μεγάλο», δηλαδή, $N \geq 80$,
- το p δεν είναι τόσο μικρό ώστε το γινόμενο Np να είναι της τάξεως του 1,

τότε η κατανομή της Y μπορεί να προσεγγιστεί από την κανονική κατανομή υπό την έννοια ότι, για κάθε ζευγάρι ακεραίων τιμών $m \leq M$,

$$\Pr(m \leq Y \leq M) \approx \Phi\left(\frac{M - Np + 1/2}{\sqrt{p(1-p)N}}\right) - \Phi\left(\frac{m - Np - 1/2}{\sqrt{p(1-p)N}}\right).$$

Στη συνέχεια, θα ξαναδούμε τα τρία θεμελιώδη παραδείγματα του Ν.Μ.Α. από την Ενότητα 9.3 του Κεφαλαίου 9, και θα τα επανεξετάσουμε κάτω από το πολύ πιο ακριβές πρίσμα του Κ.Ο.Θ.

Παράδειγμα 12.6 (Τι θα πει «πιθανότητα»;) Επιστρέφοντας στο Παράδειγμα 9.12 του Κεφαλαίου 9, όπου μας ενδιαφέρει η πιθανότητα $p = \mathbb{P}(A)$ του να συμβεί ένα συγκεκριμένο ενδεχόμενο A , θυμίζουμε πως ο Ν.Μ.Α. μάς εγγυάται ότι, αν επαναλάβουμε το ίδιο ακριβώς πείραμα πάρα πολλές, ανεξάρτητες φορές, σε διασθητικό επίπεδο: «η πιθανότητα $p = \mathbb{P}(A)$ είναι το ποσοστό των φορών, μακροπρόθεσμα, που συμβαίνει το A ».

Μαθηματικά, αυτό το σενάριο περιγράφεται ακριβώς όπως στο Παράδειγμα 12.5: Σε N επαναλήψεις του πειράματος, τις οποίες περιγράφουμε με την ακολουθία των ανεξάρτητων Τ.Μ. $X_i \sim \text{Bern}(p)$, για $i = 1, 2, \dots, N$, η συχνότητα με την οποία σημειώθηκε επιτυχία στο πείραμα, δηλαδή ο εμπειρικός μέσος όρος \bar{X}_N των X_i , θα συγκλίνει κατά πιθανότητα στο p , σύμφωνα με τον Ν.Μ.Α. Το Κ.Ο.Θ. μάς δίνει τη δυνατότητα να υπολογίσουμε με μεγαλύτερη ακρίβεια τον τρόπο με τον οποίο το \bar{X}_N συγκλίνει στο p . Στο διασθητικό επίπεδο, μας λέει πως, για μεγάλα N , η κατανομή του \bar{X}_N είναι κατά προσέγγιση $N(p, p(1-p)/N)$.

Για να διατυπώσουμε αυτή την προσέγγιση με μεγαλύτερη μαθηματική ακρίβεια, ορίζουμε το συνολικό πλήθος των επιτυχιών, $Y = N\bar{X}_N = \sum_{i=1}^N X_i$. Τότε η Y έχει κατανομή Διων(N, p) και μπορούμε να χρησιμοποιήσουμε τον υπολογισμό του Παραδείγματος 12.5, όπως διατυπώνεται στο Πόρισμα 12.1. Συγκεκριμένα, για οποιουσδήποτε δύο ακεραίους $m < M$, η πιθανότητα η συχνότητα \bar{X}_N των επιτυχιών να είναι μεταξύ m/N και M/N μπορεί να προσεγγιστεί ως:

$$\Pr\left(\frac{m}{N} \leq \bar{X}_N \leq \frac{M}{N}\right) = \Pr(m \leq Y \leq M) \approx \Phi\left(\frac{M - Np + 1/2}{\sqrt{p(1-p)N}}\right) - \Phi\left(\frac{m - Np - 1/2}{\sqrt{p(1-p)N}}\right).$$

Ας δούμε ένα αριθμητικό παράδειγμα αυτού του υπολογισμού: Σε $N = 150$ ανεξάρτητες επαναλήψεις ενός πειράματος με πιθανότητα επιτυχίας $p = 1/2$, ποια είναι η πιθανότητα το ποσοστό των επιτυχιών να είναι μεταξύ 60% και 70%; Ορίζοντας και πάλι την Τ.Μ. $Y = N\bar{X}_N \sim \text{Διων}(150, 1/2)$, η ζητούμενη πιθανότητα είναι,

$$\Pr(0.6 \leq \bar{X}_N \leq 0.7) = \Pr(90 \leq Y \leq 105),$$

η οποία προσεγγίζεται ως,

$$\begin{aligned} \Pr(90 \leq Y \leq 105) &\approx \Phi\left(\frac{105 - 75 + 1/2}{\sqrt{150/4}}\right) - \Phi\left(\frac{90 - 75 - 1/2}{\sqrt{150/4}}\right) \\ &\approx \Phi(4.98) - \Phi(2.37) \approx 1 - 0.9911 = 0.0089, \end{aligned}$$

όπου αντικαταστήσαμε την τιμή $\Phi(2.37) \approx 0.9911$ από τον Πίνακα 12.2, και χρησιμοποιήσαμε το γεγονός πως $\Phi(z) \approx 1$ για κάθε $z > 3.9$, το οποίο επίσης προκύπτει από τον Πίνακα 12.2.

Παράδειγμα 12.7 (Γιατί γίνονται δημοσκοπήσεις;) Όπως στο Παράδειγμα 9.13 του Κεφαλαίου 9, εξετάζουμε έναν πληθυσμό K ατόμων, το 25% των οποίων (δηλαδή $K/4$ άτομα) είναι ψηφοφόροι κάποιου κόμματος. Για τη διεξαγωγή μιας δημοσκόπησης επιλέγονται τυχαία N άτομα (με επανατοποθέτηση), και ρωτώνται αν θα ψηφίσουν αυτό το κόμμα. Όπως πριν, ορίζοντας τις ανεξάρτητες Τ.Μ. $X_i \sim \text{Bern}(0.25)$ για $i = 1, 2, \dots, N$ (όπου $X_i = 1$ αν το άτομο i είναι ψηφοφόρος του κόμματος, και 0 αν όχι), ο Ν.Μ.Α. μας λέει πως, αν το μέγεθος N του δείγματος είναι αρκετά μεγάλο, τότε το αποτέλεσμα της δημοσκόπησης, δηλαδή ο εμπειρικός μέσος όρος \bar{X}_N , θα είναι κοντά στο 25% με μεγάλη πιθανότητα.

Εστω τώρα πως έχουμε ένα δείγμα $N = 1000$ ατόμων, πράγμα συνηθισμένο σε πραγματικές πολιτικές δημοσκοπήσεις. Ποια είναι η πιθανότητα το αποτέλεσμα μας να πέσει έξω κατά περισσότερο από 2%; Δηλαδή, ποια είναι η,

$$\Pr(\bar{X}_N > 27\% \text{ ή } \bar{X}_N < 23\%) = 1 - \Pr(0.23 \leq \bar{X}_N \leq 0.27);$$

Ορίζοντας και πάλι την Τ.Μ.,

$$Y = \sum_{i=1}^N X_i = N\bar{X}_N,$$

η οποία έχει $\text{Διων}(N, 0.25)$ κατανομή, η πιθανότητα που μας ενδιαφέρει μπορεί να εκφραστεί ως $1 - \Pr(230 \leq Y \leq 270)$, και από το Πρόρισμα 12.1 με $N = 1000$, $p = 0.25$, $m = 230$ και $M = 270$ έχουμε την προσέγγιση,

$$\begin{aligned} \Pr(\text{«λάθος} > 2\% \text{»}) &= 1 - \Pr(230 \leq Y \leq 270) \\ &\approx 1 - \Phi\left(\frac{270 - 250 + 1/2}{\sqrt{0.25(1 - 0.25) \cdot 1000}}\right) + \Phi\left(\frac{230 - 250 - 1/2}{\sqrt{0.25(1 - 0.25) \cdot 1000}}\right) \\ &\approx 1 - \Phi(1.5) + \Phi(-1.5) \\ &\approx 1 - 0.9332 + 0.0668 = 0.1336. \end{aligned}$$

Συνεπώς, με ένα τυχαίο δείγμα 1000 ατόμων, η πιθανότητα η δημοσκόπηση να πέσει έξω κατά περισσότερο από 2% είναι μόλις 13%. Παρατηρήστε ότι το αποτέλεσμα αυτό δεν εξαρτάται από το μέγεθος K του πληθυσμού, αλλά μόνο από το μέγεθος N του δείγματος και το ποσοστό p των ψηφοφόρων. Στο τελευταίο παράδειγμα αυτής της ενότητας θα δούμε μια ρεαλιστική περίπτωση δημοσκόπησης, με πραγματικά δεδομένα από το δημοψήφισμα που έγινε στην Ελλάδα τον Ιούλιο του 2015.

Παράδειγμα 12.8 (Εκτίμηση μέσω δειγματοληψίας) Συνεχίζοντας όπως στο Παράδειγμα 9.14 του Κεφαλαίου 9, έστω πως θέλουμε να υπολογίσουμε το μέσο ύψος $\bar{y} = \frac{1}{M} \sum_{i=1}^M y_k$ ενός κοριτσιού, μεταξύ των M κοριτσιών ενός πληθυσμού, όπου y_k είναι το ύψος του κοριτσιού $k = 1, 2, \dots, M$. Αντί να εξετάσουμε ολόκληρο τον πληθυσμό, επιλέγουμε τυχαία, με επαναποδοθέτηση, N μέλη του πληθυσμού, όπου το μέγεθος N του δείγματός είναι σχετικά μεγάλο αλλά σημαντικά μικρότερο του συνολικού μεγέθους M του πληθυσμού.

Όπως και πριν, ορίζουμε τις ανεξάρτητες Τ.Μ. $X_j = \text{«ύψος του κοριτσιού } j \text{ που επιλέξαμε»}$, για $j = 1, 2, \dots, N$, οι οποίες έχουν σύνολο τιμών το $S = \{y_1, y_2, \dots, y_M\}$, και πυκνότητα $P(y_k) = \Pr(X_j = y_k) = \frac{1}{M}$, για κάθε j και κάθε k . Όπως υπολογίσαμε στο Παράδειγμα 9.14, τα X_j έχουν μέση τιμή, $E(X_j) = \bar{y}$, και παρομοίως υπολογίζουμε πως η διασπορά τους ισούται με,

$$\bar{\sigma}^2 = E[(X_j - \bar{y})^2] = \sum_{k=1}^M (y_k - \bar{y})^2 \cdot P(y_k) = \frac{1}{M} \sum_{k=1}^M (y_k - \bar{y})^2.$$

Αν εκτιμήσουμε λοιπόν το \bar{y} μέσω του εμπειρικού μέσου όρου \bar{X}_N από τα ύψη των κοριτσιών που επιλέξαμε, ο Ν.Μ.Α. μάς διαβεβαιώνει πως, για μεγάλα μεγέθη δειγμάτων N , θα έχουμε μια καλή εκτίμηση, δηλαδή $\bar{X}_n \approx \bar{y}$, με μεγάλη πιθανότητα. Επιπλέον, το Κ.Ο.Θ. μας δίνει την ακριβέστερη πληροφορία πως η κατανομή της εκτίμησής μας \bar{X}_n θα είναι κατά προσέγγιση κανονική, με μέση τιμή τη ζητούμενη ποσότητα \bar{y} και διασπορά ίση με $\bar{\sigma}^2/N$.

Κλείνουμε αυτή την ενότητα με τρία ακόμα απλά παραδείγματα εφαρμογής του Κ.Ο.Θ.

Παράδειγμα 12.9 Έστω πως η διάρκεια κάθε τηλεφωνικής κλήσης σε ένα δίκτυο έχει εκθετική κατανομή με μέση διάρκεια 85 δευτερόλεπτα, και έστω πως οι διάρκειες διαδοχικών κλήσεων είναι ανεξάρτητες. Ποια είναι η πιθανότητα η μέση διάρκεια μιας κλήσης ανάμεσα στις $N = 250$ που έγιναν μέσα σε μία μέρα να είναι μεγαλύτερη από 90 δευτερόλεπτα;

Ορίζουμε τις ανεξάρτητες Τ.Μ. X_1, X_2, \dots, X_N , όπου η κάθε $X_i \sim \text{Exp}(85)$. Από τις ιδιότητες της εκθετικής κατανομής ξέρουμε πως οι X_i έχουν μέση τιμή $\mu = E(X_i) = 85$ και τυπική απόκλιση $\sigma = 85$. Οπότε η ζητούμενη πιθανότητα $\Pr(\bar{X}_N > 90)$, όπου \bar{X}_N είναι ο εμπειρικός μέσος όρος των X_i , μπορεί να εκφραστεί ως,

$$\Pr\left(\frac{1}{N} \sum_{i=1}^N X_i > 90\right) = \Pr\left(\frac{1}{N} \sum_{i=1}^N (X_i - \mu) > 90 - 85\right) = \Pr\left(\frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N (X_i - \mu) > \frac{5\sqrt{250}}{85}\right),$$

και, εφαρμόζοντας το Κ.Ο.Θ., έχουμε,

$$\begin{aligned}\Pr(\bar{X}_N > 90) &\approx \Pr(\bar{S}_N > 0.93) \\ &= 1 - \Pr(\bar{S}_N \leq 0.93) \\ &\approx 1 - \Phi(0.93) \approx 1 - 0.8238 = 0.1762,\end{aligned}$$

όπου αντικαταστήσαμε την τιμή $\Phi(0.93) \approx 0.8238$ από τον Πίνακα 12.2.

Παράδειγμα 12.10 Κάθε φορά που ένας προγραμματιστής γράφει ένα καινούργιο κομμάτι κώδικα σε κάποιο πρόγραμμα, διαπιστώνει πως έχει κάνει X λάθη (δηλαδή βρίσκει X bugs στον κώδικά του), όπου το X είναι μια Τ.Μ. με σύνολο τιμών το $\{0, 1, 2, 3\}$ και πυκνότητα:

$$P(0) = \frac{1}{8}, \quad P(1) = \frac{1}{8}, \quad P(2) = \frac{3}{8}, \quad P(3) = \frac{3}{8}.$$

Επιπλέον, τα λάθη που κάνει σε διαφορετικά μέρη του κώδικα είναι ανεξάρτητα μεταξύ τους. Ποια είναι η πιθανότητα σε 100 νέα μέρη του κώδικα να κάνει συνολικά τουλάχιστον 210 λάθη;

Έστω $N = 100$ ανεξάρτητες Τ.Μ. X_1, X_2, \dots, X_N , όπου η κάθε X_i έχει την ίδια πυκνότητα $P(x)$ όπως πιο πάνω και συμβολίζει το πλήθος των λαθών στο μέρος i του κώδικα. Αν ορίσουμε μια νέα διακριτή Τ.Μ. Y ως το άθροισμα όλων των X_i , τότε η πιθανότητα που μας ενδιαφέρει είναι η $\Pr(Y \geq 210)$ και, για να την προσεγγίσουμε μέσω του Κ.Ο.Θ., χρειαζόμαστε τη μέση τιμή και τη διασπορά των X_i .

Από τον ορισμό των $E(X)$ και $\text{Var}(X)$ και την εναλλακτική έκφραση της διασποράς μιας διακριτής Τ.Μ., εύκολα υπολογίζουμε:

$$\begin{aligned}\mu = E(X_i) &= 0 \times \frac{1}{8} + 1 \times \frac{1}{8} + 2 \times \frac{3}{8} + 3 \times \frac{3}{8} = 2, \\ E(X_i^2) &= 0^2 \times \frac{1}{8} + 1^2 \times \frac{1}{8} + 2^2 \times \frac{3}{8} + 3^2 \times \frac{3}{8} = 5, \\ \text{και } \sigma^2 = \text{Var}(X_i) &= E(X_i^2) - (E(X_i))^2 = 1.\end{aligned}$$

Συνεπώς, η πιθανότητα που μας ενδιαφέρει μπορεί να εκφραστεί ως,

$$\Pr(Y \geq 210) = \Pr\left(\sum_{i=1}^{100} X_i \geq 210\right) = \Pr\left(\sum_{i=1}^{100} (X_i - \mu) \geq 210 - 100 \cdot 2 - \frac{1}{2}\right).$$

Παρατηρήστε πως, εφόσον η Τ.Μ. Y είναι διακριτή αλλά θα προσεγγίσουμε την κατανομή της με την κανονική η οποία είναι συνεχής, έχουμε αφαιρέσει και πάλι $1/2$ από την ποσότητα στο δεξί μέρος της τελευταίας από τις παραπάνω πιθανότητες με βάση το ίδιο σκεπτικό όπως στο Παράδειγμα 12.5.

Εφαρμόζοντας τώρα το Κ.Ο.Θ., βρίσκουμε πως η ζητούμενη πιθανότητα,

$$\begin{aligned} \Pr(Y \geq 210) &= \Pr\left(\sum_{i=1}^{100}(X_i - \mu) \geq 9.5\right) \\ &= \Pr\left(\frac{1}{\sigma\sqrt{N}} \sum_{i=1}^{100}(X_i - \mu) \geq \frac{9.5}{10}\right) \\ &= 1 - \Pr(\bar{S}_N \leq 0.95) \\ &\approx 1 - \Phi(0.95) \approx 0.1711, \end{aligned}$$

όπου αντικαταστήσαμε την τιμή $\Phi(0.95) \approx 0.8289$ από τον Πίνακα 12.2.

Αξίζει να σημειώσουμε πως, με τη βοήθεια του υπολογιστή, η ζητούμενη πιθανότητα μπορεί να υπολογιστεί ακριβώς και ισούται με 0.17150... Συνεπώς, η προσέγγιση που μας έδωσε το Κ.Ο.Θ. είναι πολύ ικανοποιητική. Επιπλέον, αν είχαμε κάνει τον πιο πάνω υπολογισμό μέσω του Κ.Ο.Θ. χωρίς να αφαιρέσουμε το επιπλέον 1/2, το αποτέλεσμα θα ήταν ≈ 0.1587 , το οποίο είναι σημαντικά χειρότερο από αυτό που υπολογίσαμε.

Παράδειγμα 12.11 (Δημοψήφισμα 2015: Ποιος δουλεύει ποιον;) Στις 5 Ιουλίου του 2015 πραγματοποιήθηκε στην Ελλάδα δημοψήφισμα σε σχέση με μια ενδεχόμενη συμφωνία δανεισμού της χώρας από την Ευρωπαϊκή Ένωση, το Διεθνές Νομισματικό Ταμείο και την Κεντρική Ευρωπαϊκή Τράπεζα. Στα αποτελέσματα σχετικής δημοσκόπησης που έκανε μία από τις μεγαλύτερες και πιο έγκυρες εταιρίες δημοσκοπήσεων, ανακοινώθηκε η πρόβλεψη πως η επιλογή του «OXI στη συμφωνία» θα έπαιρνε ποσοστό 49,8%, με «στατιστικό σφάλμα» $\pm 2.7\%$. Επιπλέον, όπως ανακοίνωσε η ίδια η εταιρία, η έρευνα έγινε σε δείγμα 1325 ατόμων κατά το διάστημα 2-3 Ιουλίου.

Το εκλογικό αποτέλεσμα τους διέψευσε οικτρά: Η επιλογή του OXI πήρε 61.31%. Εδώ θα εξετάσουμε πόσο πιστευτό είναι αυτό το αποτέλεσμα. Συγκεκριμένα, με δεδομένα τα πραγματικά αποτελέσματα, θα υπολογίσουμε πόσο πιθανό είναι η δημοσκόπηση να είχε πράγματι οδηγήσει στην πρόβλεψη που δόθηκε από την εταιρία στα MME. Για να είμαστε όσο πιο επιεικείς γίνεται, ας θεωρήσουμε πως η πρόβλεψη της εταιρίας ήταν πως το OXI θα έπαιρνε 49,8% + 2.7%, δηλαδή 52.5%, και πως η επιλογή του δείγματος έγινε με επανατοποθέτηση¹, οπότε βρισκόμαστε ακριβώς στο σενάριο του Παραδείγματος 12.7, με $N = 1325$ δείγματα σε έναν πληθυσμό $K = 9840525$ ψηφοφόρων, όπου το ποσοστό που θέλουμε να προβλέψουμε είναι το $p = 0.6131$.

1. Πόσο πιστευτή είναι η πρόβλεψη των δημοσκόπων βάσει του Κ.Ο.Θ.; Πρώτα θα υπολογίσουμε, με τα παραπάνω δεδομένα, την πιθανότητα να πέσει τόσο έξω η πρόβλεψη. Ποια είναι η πιθανότητα η δημοσκόπηση να δώσει αποτέλεσμα 52.5% ή μικρότερο;

¹Στην πράξη δεν έχει σχεδόν καμία διαφορά η επιλογή με ή χωρίς επανατοποθέτηση, μια που εύκολα υπολογίζεται (Άσκηση. Υπολογίστε το.) πως αν η επιλογή $N = 1325$ ατόμων μεταξύ των $K = 9840525$ εγγεγραμμένων ψηφοφόρων γίνει με επανατοποθέτηση, η πιθανότητα να μην επιλεγεί κανείς δύο φορές – δηλαδή το αποτέλεσμα να είναι το ίδιο με το να μην είχαμε επανατοποθέτηση – είναι μεγαλύτερη από 0.999.

Ακολουθώντας ακριβώς τα ίδια βήματα όπως στο Παράδειγμα 12.7, η ζητούμενη πιθανότητα είναι η $\Pr(\bar{X}_N \leq 0.525)$. Κατ' αρχάς παρατηρούμε πως εδώ,

$$\Pr(\bar{X}_N \leq 0.525) = \Pr(N\bar{X}_N \leq 0.525 \times 1325) = \Pr(Y \leq 695.625),$$

άρα δεν είναι απαραίτητη η διόρθωση του «ακέραιου αποτελέσματος» προσθέτοντας ή αφαιρώντας $1/2$. Οπότε, προσεγγίζουμε κανονικά τη ζητούμενη πιθανότητα μέσω του Κ.Ο.Θ.,

$$\begin{aligned} \Pr(\bar{X}_N \leq 0.525) &= \Pr\left(\frac{1}{N} \sum_{i=1}^N (X_i - \mu) \leq 0.525 - 0.6131\right) \\ &= \Pr\left(\frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N (X_i - \mu) \leq -\frac{0.0881 \cdot \sqrt{1325}}{\sqrt{0.6131 \cdot (1 - 0.6131)}}\right) \\ &= \Pr(\bar{S}_N \leq -6.7449948) \\ &\approx \Phi(-6.7449948) \approx 7.65 \times 10^{-12}, \end{aligned}$$

όπου η εξαιρετικά μικρή αυτή τιμή δεν συμπεριλαμβάνεται στους πίνακές μας, αλλά εύκολα μπορεί να βρεθεί με τη χρήση υπολογιστή.

2. *Ακριβώς πόσο πιστευτή είναι η πρόβλεψη των δημοσκοπών;* Η πιθανότητα που βρήκαμε πιο πάνω, 7.65×10^{-12} , είναι αστρονομικά μικρή. Για να έχουμε ένα μέτρο σύγκρισης, είναι παρόμοια με την πιθανότητα να κρύψουμε έναν συγκεκριμένο κόκκο άμμου σε μια παραλία και κάποιος που διαλέγει στην τύχη έναν κόκκο άμμου από ολόκληρη την παραλία να τον πετύχει! Αλλά είναι σημαντικό να θυμόμαστε πως ο υπολογισμός που μας οδήγησε σε αυτή την πιθανότητα ήταν προσεγγιστικός. Για τόσο μικρές τιμές, το σφάλμα που αναπόφευκτα εισάγεται από την προσέγγιση μπορεί να επηρεάζει σημαντικά το αποτέλεσμα.

Επανεξετάζουμε λοιπόν την πιθανότητα του να πέσει τόσο έξω η πρόβλεψη της εταιρίας, η οποία ήδη είδαμε πως μπορεί να εκφραστεί ως, $\Pr(Y \leq 695.625)$, όπου η Τ.Μ. Y έχει κατανομή $Y \sim \text{Διων}(N, p)$. Άρα, μπορούμε με τη χρήση του υπολογιστή να την υπολογίσουμε ακριβώς, χωρίς καμία προσέγγιση:

$$\begin{aligned} \Pr(Y \leq 695.625) &= \Pr(Y \leq 695) \\ &= \sum_{j=0}^{695} \Pr(Y = j) \\ &= \sum_{j=0}^{695} \binom{N}{j} p^j (1-p)^{N-j} \\ &\approx 3.501 \times 10^{-11}. \end{aligned}$$

Συμπεραίνουμε λοιπόν πως, αν η μεθοδολογία που χρησιμοποιήθηκε είναι αυτή η οποία δηλώνεται, με βάση τα πραγματικά δεδομένα στην καλύτερη περίπτωση τα αποτελέσματα που παρουσιάστηκαν από την εταιρία είναι μαθηματικώς ανεξήγητα.

3. *Χωρίς επανατοποθέτηση*; Τέλος, εξετάζουμε το ενδεχόμενο του να έχει επηρεαστεί σημαντικά ο υπολογισμός μας από την υπόθεση ότι η επιλογή των δειγμάτων γινόταν με επανατοποθέτηση. Στην περίπτωση που η επιλογή έγινε χωρίς επανατοποθέτηση, τότε η ζητούμενη πιθανότητα είναι και πάλι η $\Pr(Y \leq 695.625)$, αλλά τώρα η Τ.Μ. Y έχει κατανομή $Y \sim \text{Υπερ}(K, k, N)$, όπου $K = 9840525$ είναι το συνολικό πλήθος των ψηφοφόρων, $k = Kp \approx 6033226$ είναι το πλήθος εκείνων που σκόπευαν πράγματι να ψηφίσουν ΟΧΙ και $N = 1325$ είναι το μέγεθος του δείγματος.

Όπως πιο πάνω, μπορούμε να υπολογίσουμε,

$$\Pr(Y \leq 695.625) = \Pr(Y \leq 695) = \sum_{j=0}^{695} \Pr(j = k),$$

το οποίο, αντικαθιστώντας τον τύπο της πυκνότητας της υπεργεωμετρικής κατανομής, και με τη χρήση υπολογιστή, μας δίνει

$$\Pr(Y \leq 695.625) = \sum_{j=0}^{695} \frac{\binom{k}{j} \binom{K-k}{N-j}}{\binom{K}{N}} \approx 5.04 \times 10^{-11}.$$

Άρα και εδώ συμπεραίνουμε πως, αν η μεθοδολογία της εταιρίας είναι πράγματι αυτή την οποία περιγράφει, είναι μαθηματικά αδικοιολόγητα τα αποτελέσματα της δημοσκόπησης.

4. *Η δημοσκόπηση είναι εύκολη δουλειά*. Εδώ θα δούμε πως, παρά τις ατελείωτες αναλύσεις «ειδικών» στα ΜΜΕ, είναι μάλλον απλή υπόθεση η σχετικά ασφαλής πρόβλεψη του αποτελέσματος ενός δημοψηφίσματος: Έστω πως θέλαμε να προβλέψουμε το πραγματικό ποσοστό $p = 0.6131$, ακολουθώντας την απλούστερη δυνατή μεθοδολογία, επιλέγοντας, λ.χ., μόνο $N = 600$ ανεξάρτητα δείγματα από τον πληθυσμό των ψηφοφόρων (λιγότερα από τα μισά από όσα συμμετείχαν στην έρευνα της εταιρίας πιο πάνω), με επανατοποθέτηση. Τότε, η πιθανότητα το αποτέλεσμα \bar{X}_N της πρόβλεψής μας να απείχε από την πραγματικότητα περισσότερο από $\pm 3\%$ (όπου εδώ είμαστε πιο αυστηροί γιατί μετράμε το ενδεχόμενο λάθος και προς τα «πάνω» και προς τα «κάτω»), μπορεί εύκολα και με ακρίβεια να προσεγγιστεί μέσω του Κ.Ο.Θ. όπως και πριν. Αρχικά γράφουμε,

$$\begin{aligned} \Pr(|\bar{X}_N - p| > 0.03) &= \Pr\left(\frac{1}{N} \sum_{i=1}^N (X_i - p) > 0.03\right) + \Pr\left(\frac{1}{N} \sum_{i=1}^N (X_i - p) < -0.03\right) \\ &= \Pr\left(\frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N (X_i - p) > \frac{0.03 \times \sqrt{600}}{\sqrt{0.6131 \times (1 - 0.6131)}}\right) \\ &\quad + \Pr\left(\frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N (X_i - p) < -\frac{0.03 \times \sqrt{600}}{\sqrt{0.6131 \times (1 - 0.6131)}}\right), \end{aligned}$$

και από το Κ.Ο.Θ. βρίσκουμε,

$$\begin{aligned}\Pr(|\bar{X}_N - p| > 0.03) &\approx 1 - \Pr(\bar{S}_N < 1.509) + \Pr(\bar{S}_N < -1.509) \\ &\approx 1 - \Phi(1.51) + \Phi(-1.51) \\ &\approx 13.1\%.\end{aligned}$$

Άρα, με αυτή την πολύ απλοϊκή μεθοδολογία, και με πολύ λιγότερα δείγματα, είμαστε βέβαιοι με πιθανότητα $\approx 87\%$ ότι θα πετύχουμε το σωστό εκλογικό αποτέλεσμα με σφάλμα μικρότερο από $\pm 3\%$.

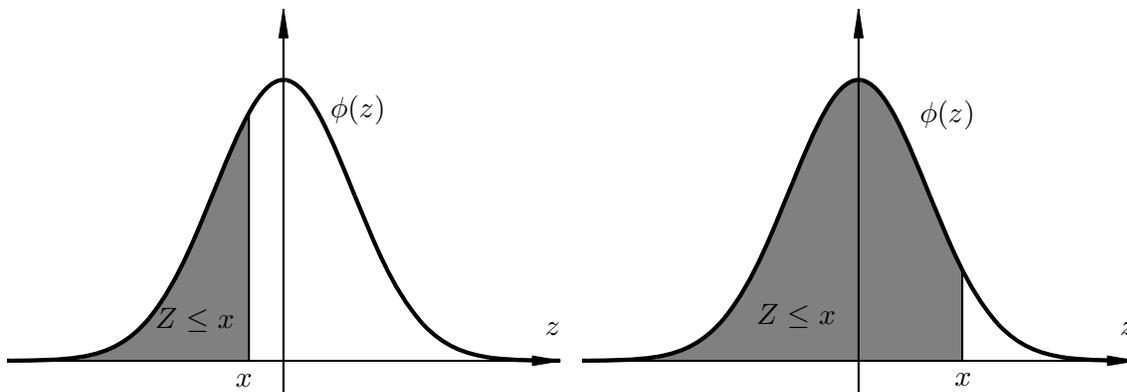
12.4 Πίνακες τιμών της τυπικής κανονικής κατανομής

Ο Πίνακας 12.1 περιέχει τιμές της συνάρτησης κατανομής,

$$\Phi(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du = \Pr(Z \leq z),$$

της τυπικής κανονικής κατανομής $Z \sim N(0, 1)$ για αρνητικές τιμές του z .

Παρομοίως, ο Πίνακας 12.2 περιέχει τιμές της συνάρτησης κατανομής $\Phi(z)$ της τυπικής κανονικής κατανομής για θετικές τιμές του z . Οι σχετικοί υπολογισμοί αναπαρίστανται γραφικά στο Σχήμα 12.4.



Σχήμα 12.4: Γραφική αναπαράσταση των τιμών της συνάρτησης κατανομής $\Phi(z)$ της τυπικής κανονικής κατανομής $N(0, 1)$ για αρνητικές τιμές του z (αριστερά) και για θετικές τιμές του z (δεξιά).

Είναι, φυσικά, προφανές πως ο δεύτερος πίνακας είναι περιττός: Λόγω της συμμετρίας της τυπικής κανονικής πυκνότητας $\phi(x)$, για οποιαδήποτε θετική τιμή του z η $\Phi(z)$ μπορεί εύκολα να υπολογιστεί ως $1 - \Phi(-z)$.

Πίνακας 12.1: Τιμές της τυπικής κανονικής συνάρτησης κατανομής $\Phi(z)$ για αρνητικά z .

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
-3.8	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
-3.7	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
-3.6	0.0002	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
-3.5	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

12.5 Ασκήσεις

1. *Τυχαία επιλογή Τ.Μ.* Η Τ.Μ. X έχει ομοιόμορφη κατανομή στο διάστημα $[3, 5]$, και η Τ.Μ. Y έχει κανονική κατανομή με μέση τιμή 1 και διασπορά 10. Ρίχνουμε ένα δίκαιο νόμισμα, και αν έρθει Κορώνα καταγράφουμε την τιμή του X , ενώ αν έρθει Γράμματα καταγράφουμε την τιμή του Y . Αν υποθέσουμε ότι η τιμή που καταγράψαμε είναι μεγαλύτερη του 4, ποια είναι η πιθανότητα να είχαμε επιλέξει τη X ;
2. *Ύψη μαθητών.* Το ύψος των μαθητών ενός σχολείου ακολουθεί κανονική κατανομή με μέση τιμή $m_A = 1.70$ m για τα αγόρια και $m_K = 1.60$ m για τα κορίτσια, και με διασπορά $\sigma^2 = (0.1 \text{ m})^2$ κοινή για αγόρια και κορίτσια. Αν υποθέσουμε ότι το πλήθος των αγοριών και των κοριτσιών σε αυτό το σχολείο είναι ακριβώς ίδιο, υπολογίστε την πιθανότητα το ύψος ενός μαθητή που επιλέγεται στην τύχη να ξεπερνά το 1.80 m.
3. *Αριθμητικοί υπολογισμοί με την κανονική κατανομή.* Έστω X μια τυχαία μεταβλητή με κανονική κατανομή $N(\mu, \sigma^2)$.
 - (α') Αν $\mu = 0$ και $\sigma^2 = 1$, να βρεθούν οι $\Pr(X < 1.22)$ και $\Pr(X > -1.22)$.
 - (β') Αν $\mu = 1$ και $\sigma^2 = 1$, να βρεθούν οι $\Pr(X > 2.7)$ και $\Pr(X < -4.7 \text{ ή } X > 2.7)$.
 - (γ') Αν $\mu = 1$ και $\sigma^2 = 1$, να βρεθεί η $\Pr(X > 2.1 \text{ ή } -1 < X < 1)$.
 - (δ') Αν $\mu = -1$ και $\sigma^2 = 4$, να βρεθούν οι $\Pr(X > 3)$ και $\Pr(X > 3 | X > 2)$.
 - (ε') Αν $\mu = -2$ και $\sigma^2 = 3.5$, να βρεθεί η μέση τιμή των $Y = 1 - X^2$ και $V = (X + 2)^{15}$.
 - (ς') Αν $\mu = 2$, να βρεθεί η τιμή της διασποράς για την οποία $\Pr(X < 0) \approx 1/3$.
4. *Τετράγωνο κανονικής Τ.Μ.* Έστω μια Τ.Μ. $X \sim N(0, \sigma^2)$ και μια σταθερά $a > 0$. Να βρεθούν η συνάρτηση κατανομής $G(y)$ και η πυκνότητα $g(y)$ της $Y = aX^2$.
5. *Η κατανομή χ^2 .* Έστω Z_1, Z_2, \dots, Z_k ανεξάρτητες Τ.Μ. με κάθε $Z_i \sim N(0, 1)$, και έστω η Τ.Μ. $Y = Z_1^2 + Z_2^2 + \dots + Z_k^2$. Η κατανομή της Y ονομάζεται κατανομή χ^2 με k βαθμούς ελευθερίας, και συμβολίζεται, $Y \sim \chi^2(k)$.
 - (α') Υπολογίστε την $E(Z_1^4)$.
 - (β') Υπολογίστε τη μέση τιμή της κατανομής $\chi^2(k)$.
 - (γ') Βρείτε τη διασπορά, την πυκνότητα και τη συνάρτηση κατανομής της $\chi^2(k)$ για $k = 1$.
6. *Λογαριθμοκανονική κατανομή.* Να βρεθούν η πυκνότητα $g(y)$ και η συνάρτηση κατανομής $G(y)$ της $Y = e^X$ συναρτήσει της πυκνότητας $f(x)$ και της συνάρτησης κατανομής $F(x)$ της συνεχούς Τ.Μ. X . Προσδιορίστε τις $g(y)$, $G(y)$ στην ειδική περίπτωση που η X έχει κανονική κατανομή $N(\mu, \sigma^2)$.
7. *Γραμμικός μετασχηματισμός.* Αποδείξτε πως, αν η $X \sim N(\mu, \sigma^2)$, τότε η $Y = aX + b$ έχει κατανομή $N(a\mu + b, a^2\sigma^2)$.

8. *Όλα στο 13!* Σε ένα καζίνο, αν ποντάρουμε στη ρουλέτα 1 ευρώ σε κάποιο νούμερο από το 1 ως το 36 και έρθει αυτό το νούμερο (πράγμα που συμβαίνει με πιθανότητα 1 στις 37) παίρνουμε 36 ευρώ (το αρχικό μας στοίχημα συν 35 ευρώ). Αν έρθει κάποιο άλλο νούμερο, χάνουμε το 1 ευρώ. Ένας παίκτης ξεκινάει με 120 ευρώ και τα ποντάρει όλα μαζί στο «13», κι ένας άλλος ξεκινάει με 120 ευρώ και ποντάρει 1 ευρώ στο «13» σε 120 διαδοχικά (ανεξάρτητα) παιχνίδια της ρουλέτας.
- (α') Δείξτε ότι και στις δύο περιπτώσεις, τα αναμενόμενα χρήματα που θα έχουν απομείνει στους παίκτες μετά το παιχνίδι είναι τα ίδια.
- (β') Και για τις δύο περιπτώσεις, υπολογίστε τη διασπορά των χρημάτων των παικτών. Σχολιάστε.
- (γ') Και για τις δύο περιπτώσεις, υπολογίστε την πιθανότητα ο παίκτης στο τέλος να έχει παραπάνω από 100 ευρώ. Σχολιάστε.
9. *Κόκκινο-μαύρο.* Ένας συντηρητικός παίκτης της ρουλέτας ποντάρει σε κάθε γύρο ένα ευρώ στο κόκκινο. Αν έρθει κόκκινο, πράγμα το οποίο συμβαίνει με πιθανότητα $16/33$, κερδίζει ένα ευρώ, ενώ αν έρθει μαύρο ή «0», με πιθανότητα $17/33$, χάνει το ευρώ που είχε ποντάρει. Υπολογίστε μια καλή προσέγγιση για την πιθανότητα μετά από 150 γύρους να μην έχει βγει χαμένος.
10. *Online game.* Σε ένα online παιχνίδι, όποτε ένα τανκ εκτελεί μια βολή, η βολή: (i) είναι άστοχη με πιθανότητα 30%, οπότε και προκαλείται στον αντίπαλο ζημιά 0 μονάδων ζωής, ή (ii) είναι απλώς επιτυχημένη με πιθανότητα 60%, οπότε προκαλείται στον αντίπαλο ζημιά 20 μονάδων ζωής, ή (iii) είναι εξαιρετικά επιτυχημένη (critical hit) με πιθανότητα 10%, οπότε προκαλείται ζημιά 40 μονάδων.
- (α') Έστω X η Τ.Μ. που περιγράφει τις μονάδες ζωής της ζημιάς που προκαλείται από μία βολή. Να υπολογίσετε τη μέση τιμή και τη διασπορά της.
- (β') Ποια είναι η πιθανότητα, μετά από 100 βολές, το άθροισμα όλων των ζημιών να υπερβαίνει τις 1700 μονάδες;
11. *Χαμένος χρόνος.* Κάθε πρωί, ο χρόνος που περιμένω μέχρι να έρθει το λεωφορείο για να πάω στο γραφείο μου έχει μέση τιμή 20 λεπτά και είναι εκθετικά κατανομημένος. Στις 259 εργάσιμες μέρες που έχει ένας χρόνος, ποια είναι η πιθανότητα να έχω χάσει πάνω από τέσσερις ολόκληρες μέρες της ζωής μου περιμένοντας το λεωφορείο;
12. *Ζυγοβίστι Αρκαδίας.* Η ετήσια βροχόπτωση στο Ζυγοβίστι Αρκαδίας περιγράφεται από μια συνεχή Τ.Μ. X (σε εκατοστά), με πυκνότητα:

$$f(x) = \begin{cases} \frac{1}{100}, & x \in [30, 130], \\ 0, & x \notin [30, 130]. \end{cases}$$

- (α') Κατά μέσο όρο, πόσα εκατοστά βροχής πέφτουν κάθε χρόνο στο Ζυγοβίσι; Ποια είναι η διασπορά τους;
- (β') Έχει παρατηρηθεί ότι η ποσότητα των καρυδιών που παράγει κάθε χρόνο το Ζυγοβίσι, σε τόνους, περιγράφεται από την Τ.Μ. $Y = 20X + 500$. Κατά μέσο όρο, πόσους τόνους καρύδια παράγει το Ζυγοβίσι κάθε χρόνο;
- (γ') Για να πετύχει η δενδροφύτευση σπάνιων κέδρων την οποία αποφάσισε η κοινότητα, απαιτείται για τα επόμενα 80 χρόνια να πέσουν λιγότερα από 7000 εκατοστά βροχής. Ποια είναι η πιθανότητα να επιτύχει η δενδροφύτευση;

13. *Αποσυγχρονισμός.* Έστω πως βλέπω ένα streaming βίντεο στον υπολογιστή μου, και παρατηρώ ότι, κάθε φορά που αποφασίζω να το πάω πιο μπροστά ή πιο πίσω, αποσυγχρονίζεται ο ήχος με την εικόνα κατά X δέκατα του δευτερολέπτου (ανεξάρτητα από τη μία φορά στην άλλη), όπου η X έχει την πυκνότητα $f(x)$ που είδαμε στην Άσκηση 8 του Κεφαλαίου 10.

Αν μετακινήσω το σημείο που βλέπω στο βίντεο 200 φορές, ποια είναι η πιθανότητα να έχει αποσυγχρονιστεί συνολικά ο ήχος κατά περισσότερα από ± 1 δευτερόλεπτο;

14. *Τηλεφωνικές κλήσεις.* Όπως στο Παράδειγμα 12.9, ας υποθέσουμε πως η διάρκεια κάθε κλήσης (σε δευτερόλεπτα) σε ένα δίκτυο έχει κατανομή $\text{Exp}(85)$ και οι διάρκειες διαδοχικών κλήσεων είναι ανεξάρτητες. Μια κλήση θεωρείται «σύντομη» αν διαρκέσει λιγότερο από ένα λεπτό. Ποια είναι η πιθανότητα, το πλήθος των σύντομων κλήσεων ανάμεσα στις 250 που έγιναν σε μία μέρα να είναι μικρότερο από 120;

15. *Εκτίμηση με θόρυβο.* Έστω πως θέλουμε να εκτιμήσουμε την άγνωστη μέση τιμή μ κάποιας Τ.Μ. X , χρησιμοποιώντας τις τιμές N ανεξάρτητων Τ.Μ. X_1, X_2, \dots, X_N , όλες με την ίδια κατανομή με τη X . Έστω επίσης πως γνωρίζουμε ότι οι X_i έχουν διασπορά $\sigma^2 = \text{Var}(X_i) = 1$.

(α') Αν για την εκτίμηση του μ χρησιμοποιήσουμε τον εμπειρικό μέσο όρο \bar{X}_N , υπολογίστε κατά προσέγγιση πόσο μεγάλο πρέπει να είναι το μέγεθος N του δείγματός μας, ώστε με πιθανότητα τουλάχιστον 95% η εκτίμηση να μην απέχει από την πραγματική τιμή κατά περισσότερο από 0.2, δηλαδή να έχουμε $|\bar{X}_N - \mu| \leq 0.2$.

(β') Έστω τώρα ότι (όπως στην Άσκηση 13 του Κεφαλαίου 9) αντί για τα X_i τα ίδια, το μόνο που έχουμε διαθέσιμο είναι τα δείγματα $Y_i = X_i + Z_i$, για $i = 1, 2, \dots, N$, όπου το κάθε Y_i είναι μια τυχαία μέτρηση του αντίστοιχου X_i συν κάποιο τυχαίο «θόρυβο» Z_i . Υποθέτουμε ότι τα Z_i είναι ανεξάρτητα από τα X_i , ότι είναι ανεξάρτητα μεταξύ τους, και ότι όλα έχουν την ίδια κατανομή, με κάποια γνωστή μέση τιμή $\nu = E(Z_i)$ και γνωστή διασπορά $\tau^2 = \text{Var}(Z_i) = 3$.

Περιγράψτε μια μέθοδο εκτίμησης της μέσης τιμής μ και πάλι υπολογίστε κατά προσέγγιση πόσο μεγάλο πρέπει να είναι το N , ώστε με πιθανότητα τουλάχιστον 95% η εκτίμηση να μην απέχει από την πραγματική τιμή κατά περισσότερο από 0.2.

Κεφάλαιο 13

★ Κ.Ο.Θ.: Λίγη θεωρία και αποδείξεις

Σε αυτό το κεφάλαιο θα δούμε τέσσερις αποδείξεις αποτελεσμάτων που σχετίζονται με την κανονική κατανομή και το Κ.Ο.Θ., οι οποίες είναι αρκετά πιο απαιτητικές, από μαθηματική άποψη, από τις περισσότερες αποδείξεις που έχουμε συναντήσει ως τώρα. Στην Ενότητα 13.1 θα δώσουμε την απόδειξη, την οποία παραλείψαμε νωρίτερα, ενός απλού αποτελέσματος από την Ενότητα 12.1: Το ολοκλήρωμα της πυκνότητας οποιασδήποτε κανονικής Τ.Μ. ισούται με ένα. Κατόπιν, στην Ενότητα 13.2 θα δούμε πως το Κ.Ο.Θ., κάτω από πολύ γενικές συνθήκες, μπορεί να χρησιμοποιηθεί για να αποδειχθεί ο Ν.Μ.Α. Κατά συνέπεια είναι, ως μαθηματικό αποτέλεσμα, αυστηρά ισχυρότερο από τον Ν.Μ.Α.

Όπως αναφέραμε στο προηγούμενο κεφάλαιο, η απόδειξη του Κ.Ο.Θ. είναι τεχνικής φύσεως και ξεπερνά τους στόχους του παρόντος βιβλίου. Παρ' όλα αυτά, στις Ενότητες 13.3 και 13.4 θα δούμε κάποιες ενδιαφέρουσες πτυχές αυτού του θέματος. Συγκεκριμένα, στην Ενότητα 13.3 θα αποδείξουμε μια ισχυρή μορφή του Κ.Ο.Θ., το λεγόμενο «θεώρημα de Moivre-Laplace», για την απλή περίπτωση που οι Τ.Μ. $\{X_n\}$ έχουν κατανομή Bernoulli. Η απόδειξη αυτή (όπως κι εκείνη της επόμενης ενότητας) είναι κάπως μακροσκελής, και η αλήθεια είναι πως υπάρχουν πολύ πιο σύντομες αποδείξεις του Κ.Ο.Θ. κάτω από πιο γενικές συνθήκες. Αυτές όμως απαιτούν τη χρήση εργαλείων της ανάλυσης Fourier ή άλλων αρκετά πιο προχωρημένων μαθηματικών περιοχών.

Τέλος, στην Ενότητα 13.4, θα δώσουμε μια όμορφη απόδειξη, εκείνη του Lindeberg, για τη γενική περίπτωση του Κ.Ο.Θ. όπως διατυπώθηκε στο Κεφάλαιο 12. Σε αντίθεση με την απόδειξη του θεωρήματος de Moivre-Laplace, η οποία βασίζεται σε μια σειρά από στοιχειώδεις υπολογισμούς, η κεντρική ιδέα της απόδειξης του Lindeberg είναι μια έξυπνη χρήση κάποιων πολύ απλών αποτελεσμάτων του διαφορικού λογισμού.

13.1 Το γκαουσιανό ολοκλήρωμα

Εδώ θα δείξουμε πως η πυκνότητα,

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}, \quad x \in \mathbb{R},$$

μιας Γ.Μ. με κατανομή $N(\mu, \sigma^2)$ έχει πράγματι ολοκλήρωμα ίσο με ένα, δηλαδή,

$$J = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} dx = 1. \quad (13.1)$$

Ξεκινάμε με την αντικατάσταση $y = (x - \mu)/\sqrt{2\sigma^2}$, οπότε $dx = \sqrt{2}\sigma dy$, και,

$$J = \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} e^{-x^2} dx.$$

Άρα, για να δείξουμε το ζητούμενο αποτέλεσμα της σχέσης (13.1) αρκεί να αποδείξουμε πως το πιο κάτω ολοκλήρωμα I ισούται με $\sqrt{\pi}$:

$$I = \int_{-\infty}^{\infty} e^{-y^2} dy = \sqrt{\pi}. \quad (13.2)$$

Εφόσον η e^{-y^2} είναι άρτια συνάρτηση του y , έχουμε,

$$I = 2 \int_0^{\infty} e^{-y^2} dy,$$

και παίρνοντας το τετράγωνο αυτής της σχέσης,

$$I^2 = \left(2 \int_0^{\infty} e^{-y^2} dy \right) \left(2 \int_0^{\infty} e^{-z^2} dz \right),$$

δηλαδή,

$$\frac{1}{4} I^2 = \int_0^{\infty} \left(\int_0^{\infty} e^{-(y^2+z^2)} dy \right) dz.$$

Τώρα, για κάθε συγκεκριμένο z , στο εσωτερικό ολοκλήρωμα κάνουμε την αντικατάσταση $y = zs$, οπότε $dy = z ds$ και,

$$\frac{1}{4} I^2 = \int_0^{\infty} \left(\int_0^{\infty} e^{-z^2(1+s^2)} z ds \right) dz = \int_0^{\infty} \left(\int_0^{\infty} e^{-z^2(1+s^2)} z dz \right) ds,$$

όπου στο τελευταίο βήμα αλλάξαμε τη σειρά των δύο ολοκληρωμάτων (βλ. θεώρημα του Fubini στο Κεφάλαιο 15). Παρατηρούμε πως,

$$\frac{d}{dz} e^{-z^2(1+s^2)} = -2z(1+s^2)e^{-z^2(1+s^2)},$$

άρα το εσωτερικό ολοκλήρωμα παραπάνω μπορεί να επιλυθεί ακριβώς,

$$\frac{1}{4}I^2 = \int_0^\infty \left[-\frac{1}{2(1+s^2)} e^{-z^2(1+s^2)} \right]_0^\infty ds = \frac{1}{2} \int_0^\infty \frac{1}{1+s^2} ds.$$

Τέλος, θυμίζουμε πως η παράγωγος της συνάρτησης $\arctan x$ είναι η $\frac{1}{1+x^2}$, οπότε καταλήγουμε στη σχέση,

$$\frac{1}{4}I^2 = \frac{1}{2} \left[\arctan s \right]_0^\infty = \frac{1}{2} \left[\frac{\pi}{2} - 0 \right] = \frac{\pi}{4},$$

η οποία μας δίνει το ζητούμενο αποτέλεσμα, $I = \sqrt{\pi}$, αποδεικνύοντας την (13.2) και, κατά συνέπεια, το αρχικό μας ζητούμενο αποτέλεσμα (13.1).

13.2 Κ.Ο.Θ. \Rightarrow Ν.Μ.Α.: Απόδειξη

Έστω οι ανεξάρτητες (συνεχείς ή διακριτές) Τ.Μ. X_1, X_2, \dots , οι οποίες έχουν όλες την ίδια κατανομή, και άρα έχουν την ίδια μέση τιμή $\mu = E(X_i)$ και την ίδια διασπορά $\sigma^2 = \text{Var}(X_i)$.

Ο Ν.Μ.Α. μάς λέει πως, για μεγάλα N , ο εμπειρικός τους μέσος όρος,

$$\bar{X}_N = \frac{1}{N} \sum_{i=1}^N X_i,$$

θα είναι κοντά στη μέση τιμή τους μ , με μεγάλη πιθανότητα. Επιπλέον, όπως είδαμε στο προηγούμενο κεφάλαιο, το Κ.Ο.Θ. μάς δίνει μια πιο λεπτομερή ποσοτική περιγραφή του \bar{X}_N : Για μεγάλα N , ο εμπειρικός μέσος όρος \bar{X}_N έχει κατά προσέγγιση $N(\mu, \sigma^2/N)$ κατανομή.

Σε αυτή την ενότητα θα δείξουμε πως το Κ.Ο.Θ. είναι πράγματι πιο ισχυρό αποτέλεσμα, με την αυστηρά μαθηματική έννοια, από τον Ν.Μ.Α. Είναι μάλιστα αξιοσημείωτο πως για να αποδείξουμε τον Ν.Μ.Α. δεν χρειάζεται να κάνουμε οποιαδήποτε άλλη υπόθεση για τις Τ.Μ. $\{X_n\}$ (λόγου χάρη, δεν είναι απαραίτητο να θεωρήσουμε καν ότι είναι ανεξάρτητες) πέραν του ότι ικανοποιούν το Κ.Ο.Θ.

Πρόταση 13.1 Έστω μια οποιαδήποτε ακολουθία $\{X_n\}$ τυχαίων μεταβλητών, και σταθερές $\mu \in \mathbb{R}, \sigma > 0$, για τις οποίες ισχύει το Κ.Ο.Θ. όπως διατυπώνεται στη σχέση (12.4) του Θεωρήματος 12.2.

Τότε, για τις $\{X_n\}$ ισχύει και ο Ν.Μ.Α.: Ο εμπειρικός τους μέσος όρος \bar{X}_N τείνει στη σταθερά μ κατά πιθανότητα καθώς το $N \rightarrow \infty$.

Υπενθύμιση: \limsup και \liminf . Ως γνωστόν, μια ακολουθία $\{a_n\}$ πραγματικών αριθμών δεν είναι απαραίτητο να έχει κάποιο όριο καθώς το $n \rightarrow \infty$. Αλλά μπορούμε πάντα να ορίσουμε,

$$\begin{aligned} \bar{a} &= \limsup_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \sup_{k \geq n} a_k \\ \text{και } \underline{a} &= \liminf_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \inf_{k \geq n} a_k, \end{aligned}$$

όπου τα δύο παραπάνω όρια πάντοτε υπάρχουν και ικανοποιούν $\bar{a} \geq a$ (γιατί;). Στην απόδειξη που ακολουθεί και σε κάποια άλλα σημεία του κεφαλαίου θα χρειαστούμε αυτές τις έννοιες, όπως και την απλή παρατήρηση ότι το όριο της $\{a_n\}$ υπάρχει και ισούται με a , αν και μόνο αν έχουμε $\bar{a} = a = a$.

Απόδειξη της Πρότασης 13.1: Έστω αυθαίρετο $\epsilon > 0$. Για τον Ν.Μ.Α. πρέπει να αποδείξουμε πως,

$$\Pr(|\bar{X}_N - \mu| < \epsilon) \rightarrow 1, \quad \text{καθώς το } N \rightarrow \infty,$$

για το οποίο αρκεί να δείξουμε πως για κάθε $\delta > 0$,

$$\liminf_{N \rightarrow \infty} \Pr(|\bar{X}_N - \mu| < \epsilon) \geq 1 - \delta. \quad (13.3)$$

Έστω λοιπόν $\delta > 0$ αυθαίρετο. Εφόσον η τυπική κανονική συνάρτηση κατανομής $\Phi(z) \rightarrow 1$ καθώς το $z \rightarrow \infty$, και $\Phi(z) \rightarrow 0$ καθώς το $z \rightarrow -\infty$, μπορούμε να διαλέξουμε ένα $K > 0$ αρκετά μεγάλο ώστε να έχουμε,

$$\Phi(K) - \Phi(-2K) \geq 1 - \delta. \quad (13.4)$$

Παρατηρούμε πως η πιθανότητα $\Pr(|\bar{X}_N - \mu| < \epsilon)$ που μας ενδιαφέρει ισούται με,

$$\begin{aligned} \Pr\left(\left|\frac{1}{N} \sum_{i=1}^N (X_i - \mu)\right| < \epsilon\right) &= \Pr\left(\left|\frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N (X_i - \mu)\right| < \frac{\epsilon\sqrt{N}}{\sigma}\right) \\ &= \Pr\left(|\bar{S}_N| < \epsilon\sqrt{N}/\sigma\right) \\ &= \Pr\left(\bar{S}_N < \epsilon\sqrt{N}/\sigma\right) - \Pr\left(\bar{S}_N \leq -\epsilon\sqrt{N}/\sigma\right) \\ &\geq \Pr\left(\bar{S}_N \leq \epsilon\sqrt{N}/2\sigma\right) - \Pr\left(\bar{S}_N \leq -\epsilon\sqrt{N}/\sigma\right), \end{aligned} \quad (13.5)$$

όπου \bar{S}_N είναι το κανονικοποιημένο άθροισμα όπως στο Κ.Ο.Θ. Από τον δεύτερο κανόνα πιθανότητας, ξέρουμε πως η πρώτη πιθανότητα στην έκφραση (13.5) είναι αύξουσα συνάρτηση του N και η δεύτερη φθίνουσα. Άρα, για κάθε $N > (2\sigma K/\epsilon)^2$, έχουμε,

$$\Pr(|\bar{X}_N - \mu| < \epsilon) \geq \Pr(\bar{S}_N \leq K) - \Pr(\bar{S}_N \leq -2K),$$

και παίρνοντας το όριο $N \rightarrow \infty$, από το Κ.Ο.Θ. συμπεραίνουμε πως,

$$\liminf_{N \rightarrow \infty} \Pr(|\bar{X}_N - \mu| < \epsilon) \geq \Phi(K) - \Phi(-2K) \geq 1 - \delta,$$

όπου η δεύτερη ανισότητα προκύπτει από τον τρόπο που επιλέξαμε το K στη σχέση (13.4). Άρα έχουμε αποδείξει πως, για κάθε $\delta > 0$ ισχύει η (13.3), αποδεικνύοντας έτσι τον Ν.Μ.Α. ■

13.3 Το θεώρημα de Moivre-Laplace

Σκοπός αυτής της παραγράφου είναι η παρουσίαση της απόδειξης μιας ισχυρής μορφής του Κ.Ο.Θ. στην περίπτωση που οι X_i είναι $\text{Bern}(p)$ τυχαίες μεταβλητές. Σε αυτή την ειδική περίπτωση το Κ.Ο.Θ. ονομάζεται «θεώρημα de Moivre-Laplace»¹ και μπορεί να αποδειχθεί με στοιχειώδεις υπολογισμούς. Όπως θα δούμε, το θεώρημα de Moivre-Laplace μάς δίνει μια ακριβή προσέγγιση για την πυκνότητα του εμπειρικού μέσου όρου \bar{X}_N των X_i ή, ισοδύναμα, για την πυκνότητα του αθροίσματός τους $S_n = N\bar{X}_N = X_1 + X_2 + \dots + X_N$.

Συμβολισμός $a_n \sim b_n$ και $a_n = o(b_n)$. Θυμίζουμε πως δύο ακολουθίες πραγματικών αριθμών $\{a_n\}$ και $\{b_n\}$ λέμε ότι είναι *ασυμπτωτικά ισοδύναμες*, το οποίο συμβολίζεται με $a_n \sim b_n$, όταν ο λόγος τους τείνει στο 1, $\lim_{n \rightarrow \infty} a_n/b_n = 1$. Επίσης, ο συμβολισμός $a_n = o(b_n)$ σημαίνει πως τα a_n τείνουν να είναι σημαντικά μικρότερα από τα b_n , δηλαδή, $\lim_{n \rightarrow \infty} a_n/b_n = 0$. Για παράδειγμα, αν $a_n = o(n)$, τότε η ακολουθία $\{a_n/n\}$ τείνει στο μηδέν καθώς το $n \rightarrow \infty$, ενώ αν $a_n = o(1)$ τότε $a_n \rightarrow 0$.

Θεώρημα 13.1 (Θεώρημα de Moivre-Laplace) Έστω μια ακολουθία $\{X_n\}$ από ανεξάρτητες τυχαίες μεταβλητές X_1, X_2, \dots που έχουν όλες κατανομή $\text{Bern}(p)$, και έστω S_n το άθροισμα τους $S_n = X_1 + X_2 + \dots + X_n$.

Για κάθε $y \in \mathbb{R}$, καθώς $n \rightarrow \infty$ έχουμε την εξής προσέγγιση για την πυκνότητα του S_n :

$$\Pr \left(S_n = \left\lfloor np + y\sqrt{np(1-p)} \right\rfloor \right) \sim \frac{1}{\sqrt{2\pi np(1-p)}} \exp(-y^2/2). \quad (13.6)$$

Παρατηρήσεις

1. Αν και το αποτέλεσμα του θεωρήματος έχει διαφορετική μορφή από εκείνη της γενικής περίπτωσης του Κ.Ο.Θ. στο Κεφάλαιο 12, η οποία αφορούσε το κανονικοποιημένο άθροισμα,

$$\bar{S}_n = \frac{1}{\sqrt{np(1-p)}} \sum_{i=1}^n (X_i - p),$$

σημειώνουμε πως η έκφραση (13.6) είναι ισχυρότερη από τη σύγκλιση των αντίστοιχων συναρτήσεων κατανομής όπως στο Θεώρημα 12.2. Συγκεκριμένα, ένας απλός (αλλά αρκετά τεχνικός και άνευ ιδιαίτερου ενδιαφέροντος) υπολογισμός βασισμένος στις απλές ιδιότητες του ορισμού του ολοκληρώματος Riemann, αποδεικνύει πως πράγματι η (13.6) συνεπάγεται ότι η συνάρτηση κατανομής των S_n τείνει στην $\Phi(x)$.

¹Το θεώρημα αυτό πρωτοδημοσιεύτηκε το 1738 από τον Abraham de Moivre στη δεύτερη έκδοση του βιβλίου του, *The Doctrine of Chances*. Η πρώτη έκδοσή του, το 1718, ήταν το πρώτο βιβλίο που γράφτηκε ποτέ με αντικείμενο εξ ολοκλήρου τις πιθανότητες. Ο de Moivre ήταν Γάλλος Ουγενότος, και έγραψε το βιβλίο στα αγγλικά ενώ ζούσε αυτοεξόριστος στην Αγγλία, προκειμένου να γλιτώσει τις διώξεις στις οποίες υπόκεινταν οι Ουγενότοι στη Γαλλία.

Το βασικό περίγραμμα αυτού του υπολογισμού έχει ως εξής. Ξεκινάμε παρατηρώντας ότι, για οποιοδήποτε $x \in \mathbb{R}$,

$$\begin{aligned} \Pr(\bar{S}_n \leq x) &= \Pr\left(\frac{S_n - np}{\sqrt{np(1-p)}} \leq x\right) \\ &= \Pr\left(S_n \leq \lfloor np + x\sqrt{np(1-p)} \rfloor\right) \\ &= \sum_{k=0}^{\lfloor np + x\sqrt{np(1-p)} \rfloor} \Pr(S_n = k). \end{aligned}$$

Αντικαθιστώντας τώρα την ασυμπτωτική έκφραση της πυκνότητας $\Pr(S_n = k)$ του S_n από την (13.6), και εξετάζοντας με λίγη προσοχή το πιο πάνω άθροισμα, εύκολα διαπιστώνουμε ότι, για μεγάλα n , κατά προσέγγιση ισούται με,

$$\begin{aligned} \Pr(\bar{S}_n \leq x) &\approx \frac{1}{\sqrt{2\pi}} \sum_{-\infty \leq y \leq x} \frac{1}{\sqrt{np(1-p)}} \exp(-y^2/2) \\ &\approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-y^2/2) dy, \end{aligned}$$

όπου στο παραπάνω άθροισμα τα διαδοχικά y που αθροίζονται έχουν μεταξύ τους απόσταση $1/\sqrt{np(1-p)}$, άρα κάθε όρος του αθροίσματος ισούται με το μήκος ενός διαστήματος επί την τιμή της συνάρτησης $\exp(-y^2/2)$ σε εκείνο το διάστημα. Συνεπώς η τελική προσέγγιση προκύπτει από τον ορισμό του ολοκληρώματος Riemann.

2. Για την απόδειξη του θεωρήματος θα χρειαστούμε την ακριβή μορφή του τύπου του Stirling, όπως δίνεται στη σημείωση που ακολουθεί το Λήμμα 7.1. Καθώς το $n \rightarrow \infty$:

$$n! \sim \sqrt{2\pi n} n^n \sqrt{n} e^{-n}. \quad (13.7)$$

Απόδειξη: Η βασική ιδέα της απόδειξης είναι απλή: Θα εφαρμόσουμε τον τύπο του Stirling στην πυκνότητα της διωνυμικής κατανομής. Κατ' αρχάς, λοιπόν, παρατηρούμε πως $S_n \sim \Delta_{\text{ων}}(n, p)$. Για λόγους ευκολίας, γράφουμε $q = 1 - p$, και ορίζουμε μια ακολουθία ακεραίων $k_n = np + x_n$, για κάθε $n \geq 1$, όπου τα $\{x_n\}$ είναι τέτοια ώστε η ακολουθία $\{x_n/\sqrt{n}\}$ να είναι φραγμένη.

Χρησιμοποιώντας τον τύπο της πυκνότητας της διωνυμικής κατανομής έχουμε,

$$\begin{aligned} \Pr(S_n = k_n) &= \binom{n}{k_n} p^{k_n} q^{n-k_n} \\ &= \frac{n!}{k_n!(n-k_n)!} p^{np+x_n} q^{nq-x_n} \\ &= \frac{n!}{(np+x_n)!(nq-x_n)!} p^{np+x_n} q^{nq-x_n}, \end{aligned}$$

εφαρμόζοντας την ακριβέστερη μορφή του τύπου του Stirling όπως δίνεται στη (13.7) βρίσκουμε,

$$\Pr(S_n = k_n) \sim \frac{\sqrt{2\pi}\sqrt{nn^n}e^{-n} p^{np+x_n}q^{nq-x_n}}{\sqrt{2\pi}\sqrt{np+x_n}(np+x_n)^{(np+x_n)}e^{-(np+x_n)}\sqrt{2\pi}\sqrt{nq-x_n}(nq-x_n)^{(nq-x_n)}e^{-(nq-x_n)}},$$

και απλοποιώντας,

$$\Pr(S_n = k_n) \sim \frac{1}{\sqrt{2\pi}}\sqrt{\frac{n}{(np+x_n)(nq-x_n)}}\exp\left\{(-np-x_n)\log\left(1+\frac{x_n}{np}\right)+(-nq+x_n)\log\left(1-\frac{x_n}{nq}\right)\right\}.$$

Στη συνέχεια θα απλοποιήσουμε περαιτέρω την έκφραση που βρίσκεται στον εκθέτη της τελευταίας παραπάνω έκφρασης. Χρησιμοποιώντας το απλό αποτέλεσμα που δίνεται στη σχέση (13.9) του Λήμματος 13.1 παρακάτω, ο εκθέτης αυτός ισούται με,

$$\begin{aligned} & -(x_n+np)\left(\frac{x_n}{np}-\frac{x_n^2}{2n^2p^2}+o\left(\frac{x_n^2}{n^2p^2}\right)\right)+(nq-x_n)\left(\frac{x_n}{nq}+\frac{x_n^2}{2n^2q^2}+o\left(\frac{x_n^2}{n^2q^2}\right)\right) \\ &= -\frac{x_n^2}{np}-x_n+\frac{x_n^3}{2n^2p^2}+\frac{x_n^2}{2np}+o\left(\frac{x_n^3}{n^2p^2}\right)+o\left(\frac{x_n^2}{np}\right) \\ & \quad +x_n-\frac{x_n^2}{nq}+\frac{x_n^2}{2nq}-\frac{x_n^3}{2n^2q^2}+o\left(\frac{x_n^2}{nq}\right)+o\left(\frac{x_n^3}{n^2q^2}\right), \end{aligned}$$

όπου παρατηρούμε πως, από την υπόθεσή μας ότι η ακολουθία $\{x_n/\sqrt{n}\}$ είναι φραγμένη, προκύπτει πως το άθροισμα των τεσσάρων όρων της μορφής $o(\dots)$ παραπάνω ισούται με έναν όρο που τείνει στο μηδέν. Άρα, ο εκθέτης που εξετάζουμε τελικά ισούται με,

$$\frac{x_n^2}{n}\left(-\frac{1}{p}+\frac{1}{2p}-\frac{1}{q}+\frac{1}{2q}\right)+o(1)=-\frac{x_n^2}{2npq}+o(1),$$

και, αντικαθιστώντας στην προηγούμενη έκφραση για την πυκνότητα του S_n , έχουμε,

$$\begin{aligned} \Pr(S_n = k_n) & \sim \frac{1}{\sqrt{2\pi}}\sqrt{\frac{n}{(np+x_n)(nq-x_n)}}\exp\left[-\frac{x_n^2}{2npq}+o(1)\right] \\ & \sim \frac{1}{\sqrt{2\pi npq}}\exp\left[-\frac{x_n^2}{2npq}\right]. \end{aligned} \tag{13.8}$$

Τώρα μπορούμε να εφαρμόσουμε την πιο πάνω σχέση στην περίπτωση όπου το $k(n)$ είναι της μορφής $k(n) = \lfloor np + y\sqrt{npq} \rfloor$ για κάποια σταθερά $y \in \mathbb{R}$, το οποίο είναι προφανώς της μορφής $k(n) = np + x_n$ που έχουμε υποθέσει. Σε αυτή την περίπτωση, η (13.8) μας δίνει,

$$\Pr(S_n = \lfloor np + y\sqrt{npq} \rfloor) \sim \frac{1}{\sqrt{2\pi npq}}\exp(-y^2/2),$$

αποδεικνύοντας το θεώρημα. ■

Λήμμα 13.1 Καθώς $n \rightarrow \infty$,

$$\log\left(1 + \frac{1}{n}\right) = \frac{1}{n} - \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right),$$

και, κατά συνέπεια, για οποιαδήποτε ακολουθία $\{a_n\}$ που τείνει στο άπειρο παρομοίως έχουμε,

$$\log\left(1 + \frac{1}{a_n}\right) = \frac{1}{a_n} - \frac{1}{2a_n^2} + o\left(\frac{1}{a_n^2}\right). \quad (13.9)$$

Απόδειξη: Η σχέση (13.9) είναι προφανής συνέπεια της πρώτης πρότασης του λήμματος, οπότε αρκεί να δείξουμε ότι:

$$\lim_{n \rightarrow \infty} \frac{\log(1 + \frac{1}{n}) - \frac{1}{n} + \frac{1}{2n^2}}{\frac{1}{n^2}} = 0.$$

Σε αυτό το όριο εμφανίζεται απροσδιοριστία της μορφής $0/0$, η οποία μπορεί να αφαιρεθεί αν κάνουμε χρήση του κανόνα του L'Hôpital. Πράγματι, τότε έχουμε,

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\log(1 + \frac{1}{n}) - \frac{1}{n} + \frac{1}{2n^2}}{\frac{1}{n^2}} &= \lim_{n \rightarrow \infty} \frac{\frac{1}{1+1/n}(-1/n^2) + 1/n^2 - 1/n^3}{(-2/n^3)} \\ &= \lim_{n \rightarrow \infty} \frac{n^3}{2} \left[\frac{1}{n(n+1)} + \frac{1}{n^3} - \frac{1}{n^2} \right] \\ &= \lim_{n \rightarrow \infty} \frac{1}{2(n+1)} = 0, \end{aligned}$$

άρα αποδείξαμε το ζητούμενο. ■

13.4 Το θεώρημα του Lindeberg

Στην τελευταία αυτή ενότητα του κεφαλαίου, θα δώσουμε μια πλήρη απόδειξη του Κ.Ο.Θ., και μάλιστα κάτω από γενικές συνθήκες. Ακολουθώντας τις βασικές γραμμές μιας απόδειξης του Lindeberg (1922), θα αποδείξουμε τη γενική μορφή του Κ.Ο.Θ. όπως ακριβώς είναι διατυπωμένο στο Θεώρημα 12.2 του Κεφαλαίου 12. Όπως αναφέραμε πιο πάνω, αν και η απόδειξη είναι μακροσκελής, δεν απαιτεί τη χρήση ιδιαίτερων μαθηματικών εργαλείων πέρα από κάποια απλά αποτελέσματα του διαφορικού λογισμού για συναρτήσεις μίας μεταβλητής. Επιπλέον, έχει το πλεονέκτημα ότι εύκολα γενικεύεται σε πολλές περιπτώσεις ακολουθιών τυχαίων μεταβλητών οι οποίες δεν ικανοποιούν τις παρούσες συνθήκες. Τέτοια θέματα, όμως, δεν θα μας απασχολήσουν εδώ περαιτέρω.

Απόδειξη του Lindeberg για το Κ.Ο.Θ.: Χωρίζουμε την απόδειξη σε έξι βήματα, χάριν ευκολίας στην ανάγνωση και την κατανόησή της. Όπως θα καταστεί σαφές, το σημαντικότερο και πιο πρωτότυπο βήμα είναι το πέμπτο.

Θυμίζουμε τις υποθέσεις του Θεωρήματος 12.2: Οι τυχαίες μεταβλητές $\{X_n\}$ είναι ανεξάρτητες, έχουν όλες την ίδια κατανομή και, κατά συνέπεια, έχουν κοινή μέση τιμή $\mu = E(X_i)$ και διασπορά $\sigma^2 = \text{Var}(X)$, για κάθε i . Ο στόχος μας είναι να δείξουμε ότι, για κάθε $x \in \mathbb{R}$,

$$\Pr(\bar{S}_N \leq x) = \Pr\left(\frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N (X_i - \mu) \leq x\right) \rightarrow \Phi(x), \quad \text{καθώς } N \rightarrow \infty. \quad (13.10)$$

ΒΗΜΑ 1. Ισχυριζόμαστε πως αρκεί να δείξουμε τη (13.10) για την περίπτωση $\mu = 0$ και $\sigma^2 = 1$: Έστω ότι, για οποιαδήποτε ακολουθία ανεξάρτητων Τ.Μ. $\{Y_i\}$ με μέση τιμή μηδέν και διασπορά ίση με ένα, γνωρίζουμε πως,

$$\Pr\left(\frac{1}{\sqrt{N}} \sum_{i=1}^N Y_i \leq x\right) \rightarrow \Phi(x), \quad \text{καθώς } N \rightarrow \infty. \quad (13.11)$$

Δεδομένης της ακολουθίας $\{X_n\}$ που μας ενδιαφέρει, αν ορίσουμε τις νέες Τ.Μ. $Y_n = \frac{X_n - \mu}{\sigma}$, για $n = 1, 2, \dots$, τότε (βλ. Παρατήρηση 2 στην Ενότητα 9.2 και Θεώρημα 15.1 στο Κεφάλαιο 15) οι $\{Y_n\}$ είναι ανεξάρτητες, και, από τις ιδιότητες των Θεωρημάτων 6.1 και 11.2, επίσης έχουμε,

$$\begin{aligned} E(Y_i) &= E\left(\frac{X_i - \mu}{\sigma}\right) = \frac{E(X_i) - \mu}{\sigma} = 0 \\ \text{και } \text{Var}(Y_i) &= \text{Var}\left(\frac{X_i - \mu}{\sigma}\right) = \frac{1}{\sigma^2} \text{Var}(X_i) = 1. \end{aligned}$$

Μπορούμε άρα να εφαρμόσουμε τη (13.11) στις $\{Y_n\}$, η οποία αμέσως μας δίνει τη (13.10).

ΒΗΜΑ 2. Έστω οι ανεξάρτητες Τ.Μ. $\{Y_n\}$, όλες με την ίδια κατανομή και με $\mu = E(Y_i) = 0$ και $\sigma^2 = \text{Var}(Y_i) = 1$ για κάθε i . Συμβολίζουμε με \bar{T}_N το κανονικοποιημένο άθροισμά τους,

$$\bar{T}_N = \frac{1}{\sqrt{N}} \sum_{i=1}^N Y_i,$$

και με $F_N(x)$ τη συνάρτηση κατανομής του $\bar{T}_N(x)$. Τότε η ζητούμενη σχέση (13.11) γίνεται,

$$F_N(x) = \Pr(\bar{T}_N \leq x) \rightarrow \Phi(x), \quad \text{καθώς } N \rightarrow \infty. \quad (13.12)$$

Έστω τώρα $Z \sim N(0, 1)$ με συνάρτηση κατανομής $\Phi(x)$. Όπως στην Άσκηση 2 του Κεφαλαίου 7, ορίζουμε, για κάθε $x \in \mathbb{R}$, τις δείκτριες συναρτήσεις, $h_x : \mathbb{R} \rightarrow \{0, 1\}$ ως,

$$h_x(y) = \begin{cases} 1, & \text{αν } y \leq x, \\ 0, & \text{αν } y > x, \end{cases} \quad (13.13)$$

και παρατηρούμε πως οι $F_N(x)$ και $\Phi(x)$ μπορούν να εκφραστούν, αντίστοιχα, ως,

$$F_N(x) = E[h_x(\bar{T}_N)] \quad \text{και} \quad \Phi(x) = E[h_x(Z)].$$

Άρα, το ζητούμενό μας είναι να αποδείξουμε πως,

$$E[h_x(\bar{T}_N)] \rightarrow E[h_x(Z)], \quad \text{καθώς } N \rightarrow \infty, \quad (13.14)$$

για κάθε δείκτρια συνάρτηση h_x της μορφής (13.13).

ΒΗΜΑ 3. Εδώ ισχυριζόμαστε πως, αντί να αποδείξουμε τη σύγκλιση της σχέσης (13.14) για κάθε δείκτρια συνάρτηση $h_x(y)$, αρκεί να αποδείξουμε πως,

$$E[h(\bar{T}_N)] \rightarrow E[h(Z)], \quad \text{καθώς } N \rightarrow \infty, \quad (13.15)$$

για όλες τις συναρτήσεις $h : \mathbb{R} \rightarrow \mathbb{R}$ οι οποίες ικανοποιούν: (1.) Η τρίτη παράγωγος $h'''(y)$ υπάρχει για όλα τα $y \in \mathbb{R}$ και είναι φραγμένη, και (2.) Για κάποιο $M > 0$, η $h(y) = 0$ για όλα τα $y \notin [-M, M]$. Το ότι η (13.15) πράγματι ισχύει είναι το πιο ενδιαφέρον μέρος της απόδειξης και θα το δείξουμε στο πέμπτο βήμα. Το ότι η (13.15) για όλες αυτές τις συναρτήσεις είναι αρκετή για να αποδειχθεί η ζητούμενη σύγκλιση (13.14) για κάθε $x \in \mathbb{R}$ είναι το λιγότερο ενδιαφέρον μέρος της απόδειξης, και θα δειχθεί στο τελευταίο βήμα.

Παρατηρούμε πως, αν η $h(y)$ ικανοποιεί τις δύο παραπάνω συνθήκες, τότε και η $h(y)$ και οι τρεις πρώτοι παράγωγοί της είναι φραγμένες συναρτήσεις. Επιπλέον, από το ανάπτυγμα Taylor έχουμε ότι, για κάθε x, y ,

$$h(x+y) = h(x) + yh'(x) + \frac{y^2}{2!}h''(x) + \frac{y^3}{3!}h'''(\xi),$$

για κάποιο ξ μεταξύ των τιμών x και $x+y$. Το ότι η $h'''(y)$ είναι φραγμένη σημαίνει πως υπάρχει κάποια σταθερά $C_3 < \infty$ τέτοια ώστε $|h'''(z)| \leq C_3$ για όλα τα $z \in \mathbb{R}$. Οπότε, δεδομένου αυθαίρετου $\epsilon > 0$, αν θέσουμε $\delta = 6\epsilon/C_3 > 0$, το παραπάνω ανάπτυγμα μας δίνει,

$$\begin{aligned} \left| h(x+y) - h(x) - yh'(x) - \frac{y^2}{2}h''(x) \right| &\leq y^2 \frac{C_3|y|}{6} \\ &\leq \begin{cases} \epsilon y^2, & \text{αν } |y| < \delta, \\ C y^2, & \text{για όλα τα } y \in \mathbb{R}, \end{cases} \\ &\leq \epsilon y^2 + C y^2 [1 - h_\delta(|y|)], \end{aligned} \quad (13.16)$$

όπου η σταθερά $C = MC_3/6 < \infty$, με M το φράγμα που έχουμε από το γεγονός ότι η $h(y)$ ικανοποιεί τη δεύτερη από τις παραπάνω συνθήκες, και όπου η h_δ είναι δείκτρια συνάρτηση της μορφής που ορίσαμε στη (13.13).

ΒΗΜΑ 4. Πριν προχωρήσουμε στο πέμπτο και σημαντικότερο βήμα, θυμίζουμε μια απλή ιδιότητα που είδαμε, στην περίπτωση των διακριτών Τ.Μ., στην Άσκηση 3 του Κεφαλαίου 7: Έστω μια οποιαδήποτε διακριτή Τ.Μ. V τέτοια ώστε $E(|V|) < \infty$. Δεδομένης μιας δείκτριας συνάρτησης h_x όπως στα προηγούμενα δύο βήματα, για οποιοδήποτε δεδομένο $x > 0$, ορίζουμε επίσης τη συμπληρωματική δείκτρια συνάρτηση,

$$H_x(v) = 1 - h_x(v) = \begin{cases} 1, & \text{αν } v > x, \\ 0, & \text{αν } v \leq x. \end{cases} \quad (13.17)$$

Από το αποτέλεσμα της Άσκησης 3 του Κεφαλαίου 7, με $Y = |V|$, έχουμε,

$$\lim_{x \rightarrow \infty} E[|V|H_x(|V|)] = 0.$$

Επιπλέον, στην Άσκηση 4 στο τέλος αυτού του κεφαλαίου θα δείξουμε ότι το ίδιο ισχύει αν η V είναι συνεχής. Και μια και το ίδιο ακριβώς αποτέλεσμα μπορεί να εφαρμοστεί στη V^2 , αν θεωρήσουμε ότι ικανοποιεί $E(V^2) < \infty$, έχουμε παρομοίως ότι, για οποιαδήποτε (συνεχή ή διακριτή) Τ.Μ. με $E(V^2) < \infty$,

$$\lim_{x \rightarrow \infty} E(V^2 H_x(|V|)) = \lim_{x \rightarrow \infty} E(V^2 H_{x^2}(V^2)) = 0, \quad (13.18)$$

γεγονός το οποίο θα χρησιμοποιήσουμε παρακάτω.

ΒΗΜΑ 5. Σε αυτό το βήμα θα δούμε το συλλογισμό που βρίσκεται στην καρδιά της απόδειξης του Lindeberg. Θα αποδείξουμε τη σχέση (13.15) όπως απαιτείται από το τρίτο βήμα. Έστω λοιπόν μια αυθαίρετη συνάρτηση $h : \mathbb{R} \rightarrow \mathbb{R}$ που ικανοποιεί τις δύο συνθήκες του τρίτου βήματος. Επιπλέον, ορίζουμε τις ανεξάρτητες Τ.Μ. Z_1, Z_2, \dots, Z_N με κάθε $Z_i \sim N(0, 1)$.

Από την τέταρτη ιδιότητα του Θεωρήματος 12.1, ξέρουμε ότι το κανονικοποιημένο άθροισμα $\bar{R}_N = \frac{1}{\sqrt{N}} \sum_{i=1}^N Z_i$ των Z_i έχει κατανομή $\bar{R}_N \sim N(0, 1)$. Άρα, το ζητούμενό μας είναι να δείξουμε πως, καθώς $N \rightarrow \infty$,

$$E[h(\bar{T}_N)] - E[h(\bar{R}_N)] = E\left\{h\left[\frac{1}{\sqrt{N}} \sum_{i=1}^N Y_i\right]\right\} - E\left\{h\left[\frac{1}{\sqrt{N}} \sum_{i=1}^N Z_i\right]\right\} \rightarrow 0,$$

ή, θέτοντας, για ευκολία, $Y'_i = Y_i/\sqrt{N}$ και $Z'_i = Z_i/\sqrt{N}$, ότι, καθώς $N \rightarrow \infty$,

$$\Delta_N = \left| E\left\{h\left[\sum_{i=1}^N Y'_i\right]\right\} - E\left\{h\left[\sum_{i=1}^N Z'_i\right]\right\} \right| \rightarrow 0. \quad (13.19)$$

Η κεντρική ιδέα της απόδειξης είναι η πιο κάτω έξυπνη αναπαράσταση, όπου αντικαθιστούμε

έναν-έναν τους όρους Y'_i του αθροίσματος με τις αντίστοιχες κανονικές Τ.Μ.:

$$\begin{aligned}\Delta_N &= \left| \sum_{k=1}^N \left(E \left\{ h \left[\sum_{i=1}^k Y'_i + \sum_{i=k+1}^N Z'_i \right] \right\} - E \left\{ h \left[\sum_{i=1}^{k-1} Y'_i + \sum_{i=k}^N Z'_i \right] \right\} \right) \right| \\ &\leq \sum_{k=1}^N \left| E \left\{ h \left[\sum_{i=1}^k Y'_i + \sum_{i=k+1}^N Z'_i \right] \right\} - E \left\{ h \left[\sum_{i=1}^{k-1} Y'_i + \sum_{i=k}^N Z'_i \right] \right\} \right| \\ &= \sum_{k=1}^N \left| E \left\{ h \left[\left(\sum_{i=1}^{k-1} Y'_i + \sum_{i=k+1}^N Z'_i \right) + Y'_k \right] - h \left[\left(\sum_{i=1}^{k-1} Y'_i + \sum_{i=k+1}^N Z'_i \right) + Z'_k \right] \right\} \right|.\end{aligned}$$

Δεδομένου αυθαίρετου $\epsilon > 0$, χρησιμοποιώντας δύο φορές το φράγμα που βρήκαμε στη σχέση (13.16) από το ανάπτυγμα Taylor, ο κάθε όρος στο τελευταίο πιο πάνω άθροισμα μπορεί να εκφραστεί ως,

$$\begin{aligned}E \left\{ (Y'_k - Z'_k) h' \left[\sum_{i=1}^{k-1} Y'_i + \sum_{i=k+1}^N Z'_i \right] \right\} + \frac{1}{2} E \left\{ (Y'^2_k - Z'^2_k) h'' \left[\sum_{i=1}^{k-1} Y'_i + \sum_{i=k+1}^N Z'_i \right] \right\} \\ + E(A_k) + E(B_k),\end{aligned}\tag{13.20}$$

όπου οι όροι των «υπολοίπων» A_k και B_k που προέρχονται από το ανάπτυγμα Taylor ικανοποιούν,

$$\begin{aligned}|A_k| &\leq \epsilon Y'^2_k + C Y'^2_k H_\delta(|Y'_k|) \\ \text{και } |B_k| &\leq \epsilon Z'^2_k + C Z'^2_k H_\delta(|Z'_k|),\end{aligned}$$

με H_δ τη συμπληρωματική δείκτρια συνάρτηση που ορίσαμε στην (13.17).

Η κρίσιμη παρατήρηση σε αυτό το σημείο είναι πως, λόγω της ανεξαρτησίας των Τ.Μ. Y'_i και Z'_i , και επειδή όλες αυτές οι Τ.Μ. έχουν μέση τιμή μηδέν και ίδια διασπορά $1/N$, οι δύο πρώτες μέσες τιμές στην έκφραση (13.20) είναι ίσες με μηδέν. Οπότε, αντικαθιστώντας,

$$\begin{aligned}\Delta_N &\leq \sum_{k=1}^N \left\{ \epsilon E(Y'^2_k) + C E [Y'^2_k H_\delta(|Y'_k|)] + \epsilon E(Z'^2_k) + C E [Z'^2_k H_\delta(|Z'_k|)] \right\} \\ &\stackrel{(a)}{=} \epsilon \sum_{k=1}^N \left[E \left(\frac{Y'^2_k}{N} \right) + E \left(\frac{Z'^2_k}{N} \right) \right] + C \sum_{k=1}^N \left[\frac{1}{N} E \left(Y'^2_k H_\delta \left(\frac{|Y'_k|}{N} \right) \right) + \frac{1}{N} E \left(Z'^2_k H_\delta \left(\frac{|Z'_k|}{N} \right) \right) \right] \\ &\stackrel{(b)}{=} 2\epsilon + \frac{C}{N} \sum_{k=1}^N \left[E \left(Y'^2_k H_{N\delta}(|Y_k|) \right) + E \left(Z'^2_k H_{N\delta}(|Z_k|) \right) \right] \\ &\stackrel{(c)}{=} 2\epsilon + C \left[E \left(Y_1^2 H_{N\delta}(|Y_1|) \right) + E \left(Z_1^2 H_{N\delta}(|Z_1|) \right) \right],\end{aligned}$$

όπου για την ισότητα (a) αντικαταστήσαμε τους ορισμούς των Y'_k, Z'_k , για την (b) χρησιμοποιήσαμε το γεγονός ότι όλες αυτές οι Τ.Μ. έχουν μέση τιμή μηδέν και διασπορά ένα και την παρατήρηση

πως, εξ ορισμού, $H_\delta(|a|/N) = H_{N\delta}(|a|)$, και για την (c) χρησιμοποιήσαμε το ότι όλοι οι όροι του τελευταίου αθροίσματος είναι ίδιοι για κάθε k .

Εφαρμόζοντας τώρα το αποτέλεσμα (13.18) του τέταρτου βήματος πιο πάνω, έχουμε πως,

$$\limsup_{N \rightarrow \infty} \Delta_N \leq 2\epsilon,$$

και, αφού το $\epsilon > 0$ ήταν αυθαίρετο, έχουμε τελικά αποδείξει το ζητούμενο αποτέλεσμα της σχέσης (13.19): $\lim_{N \rightarrow \infty} \Delta_N = 0$.

ΒΗΜΑ 6. Το μόνο που απομένει να κάνουμε είναι να τεκμηριώσουμε τον ισχυρισμό του βήματος 3, δηλαδή ότι αν,

$$E[h(\bar{T}_N)] \rightarrow E[h(Z)], \quad \text{καθώς } N \rightarrow \infty,$$

για όλες τις συναρτήσεις $h: \mathbb{R} \rightarrow \mathbb{R}$ για τις οποίες η τρίτη παράγωγος $h'''(y)$ υπάρχει για όλα τα $y \in \mathbb{R}$, και $h(y) = 0$ για όλα τα $y \notin [-M, M]$, για κάποιο $M > 0$, τότε έχουμε και ότι,

$$E[h_x(\bar{T}_N)] \rightarrow E[h_x(Z)], \quad \text{καθώς } N \rightarrow \infty, \quad (13.21)$$

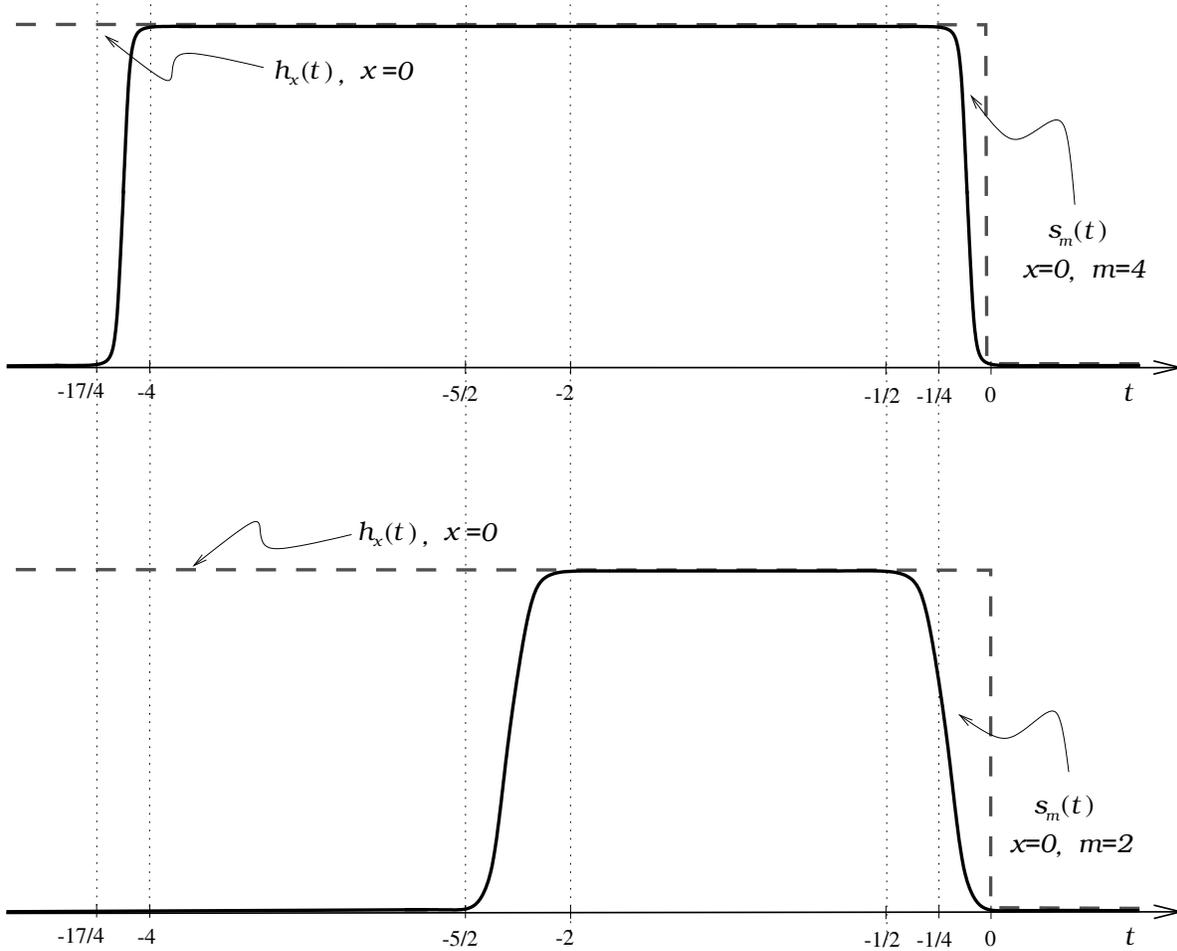
για κάθε δείκτρια συνάρτηση h_x της μορφής (13.13).

Η απόδειξη αυτού του ισχυρισμού είναι κυρίως τεχνική και ελάχιστα σχετίζεται με ιδέες και εργαλεία των πιθανοτήτων. Δεν θα παρεξηγήσουμε λοιπόν τους αναγνώστες που θα τη δεχτούν χωρίς να τη διαβάσουν, και συνεχίζουμε μόνο για τους πιο φανατικούς της μαθηματικής ανάλυσης.

Έστω μια δείκτρια συνάρτηση h_x για κάποιο δεδομένο $x \in \mathbb{R}$. Θα προσεγγίσουμε την $h_x(y)$ μέσω της ακολουθίας συναρτήσεων $\{s_m\}$, όπου, για $m > |x| + 1$,

$$s_m(t) = \begin{cases} 0, & \text{αν } t \leq -m - \frac{1}{m}, \\ A^{-1} \int_{-m(t+m)}^1 \exp\left(-\frac{1}{z(1-z)}\right) dz, & \text{αν } -m - \frac{1}{m} < t < -m, \\ 1, & \text{αν } -m \leq t \leq x - \frac{1}{m}, \\ A^{-1} \int_{m(t-x)+1}^1 \exp\left(-\frac{1}{z(1-z)}\right) dz, & \text{αν } x - \frac{1}{m} < t < x, \\ 0, & \text{αν } t \geq x, \end{cases}$$

και η σταθερά $A = \int_0^1 \exp\left(-\frac{1}{z(1-z)}\right) dz$. Από τον ορισμό της είναι εύκολο να διαπιστώσουμε ότι, για μεγάλα m , η $s_m(y)$ προσεγγίζει όλο και περισσότερο την $h_x(y)$: Καθώς το m μεγαλώνει, αφενός το διάστημα $[-m, x - 1/m]$ στο οποίο η $s_m(y) = 1$ προσεγγίζει όλο και περισσότερο το αντίστοιχο διάστημα $(-\infty, x]$ όπου η $h_x(y) = 1$, και αφετέρου, τα διαστήματα όπου $s_m(y)$ παίρνει τιμές αυστηρά μεταξύ 0 και 1 γίνονται όλο και μικρότερα ώσπου στο όριο εξαφανίζονται. Ένα παράδειγμα αυτής της συμπεριφοράς φαίνεται στο Σχήμα 13.1, όπου, για την περίπτωση $x = 0$, έχουμε σχεδιάσει τα γραφήματα των συναρτήσεων $s_m(y)$ για $m = 2$ και $m = 4$.



Σχήμα 13.1: Τα γραφήματα των συναρτήσεων $s_m(t)$ στην περίπτωση $x = 0$, για $m = 2$ (κάτω) και για $m = 4$ (επάνω). Στην απόδειξη η ακολουθία $\{s_m\}$ χρησιμοποιείται ως προσέγγιση της δείκτριας συνάρτησης $h_x(t)$, της οποίας το γράφημα αναπαρίσταται με τη διακεκομμένη γραμμή.

Από τον ορισμό της s_m προφανώς έχουμε $s_m(t) = 0$ για $t \notin [-(m+1), m+1]$ και εύκολα μπορούμε να διαπιστώσουμε ότι η τρίτη της παράγωγος $s_m'''(t)$ υπάρχει για κάθε $t \in \mathbb{R}$. Άρα για κάθε δεδομένο m , από το προηγούμενο βήμα έχουμε ότι,

$$E[s_m(\bar{T}_N)] \rightarrow E[s_m(Z)], \quad \text{καθώς } N \rightarrow \infty. \quad (13.22)$$

Επιπλέον, έχουμε $0 \leq s_m(t) \leq s_{m+1}(t) \leq h_x(t)$ για κάθε m και κάθε $t \in \mathbb{R}$, αλλά αφού $s_m(t) = 1$ για όλα τα $t \in [-m, x - 1/m]$, έχουμε και ότι $s_m(t) \geq h_{x-1/m}(t) - h_{-m}(t)$, οπότε,

$$E[s_m(Z)] \geq E[h_{x-1/m}(Z) - h_{-m}(Z)] = \Phi(x - 1/m) - \Phi(-m).$$

Άρα, δεδομένου αυθαίρετου $\epsilon > 0$, αφού η $\Phi(z)$ είναι συνεχής και $\lim_{z \rightarrow -\infty} \Phi(z) = 0$, μπορούμε να επιλέξουμε m_0 αρκετά μεγάλο ώστε να έχουμε $E[s_m(Z)] \geq \Phi(x) - \epsilon$ για όλα τα $m \geq m_0$. Συνδυάζοντας τις παραπάνω παρατηρήσεις, προκύπτει ότι,

$$\liminf_{N \rightarrow \infty} E[h_x(\bar{T}_N)] \geq \liminf_{N \rightarrow \infty} E[s_m(\bar{T}_N)] = E[s_m(Z)] \geq \Phi(x) - \epsilon = E[h_x(Z)] - \epsilon,$$

για κάθε $m \geq m_0$. Και εφόσον το $\epsilon > 0$ ήταν αυθαίρετο, έχουμε ότι,

$$\liminf_{N \rightarrow \infty} E[h_x(\bar{T}_N)] \geq E[h_x(Z)] - \epsilon. \quad (13.23)$$

Τέλος (επιτέλους!), για να δείξουμε το αντίστοιχο άνω φράγμα, παρατηρούμε ότι, για κάθε m και t , $s_m(t) \geq h_{x-1/m}(t) - h_{-m}(t)$, οπότε,

$$E[h_{x-1/m}(\bar{T}_N)] \leq E[s_m(\bar{T}_N)] + E[h_{-m}(\bar{T}_N)].$$

Για τον τελευταίο όρο, χρησιμοποιώντας το γεγονός ότι το \bar{T}_N έχει μέση τιμή μηδέν και διασπορά ίση με ένα, και εφαρμόζοντας ανισότητα Chebychev, έχουμε,

$$E[h_{-m}(\bar{T}_N)] = \Pr(\bar{T}_N \leq -m) \leq \Pr(|\bar{T}_N| \geq m) \leq \frac{E(\text{Var}(\bar{T}_N))}{m^2} = \frac{1}{m^2},$$

οπότε, αντικαθιστώντας,

$$E[h_{x-1/m}(\bar{T}_N)] \leq E[s_m(\bar{T}_N)] + \frac{1}{m^2}.$$

Παίρνοντας το όριο $N \rightarrow \infty$, από το αποτέλεσμα (13.22) του τρίτου βήματος,

$$\limsup_{N \rightarrow \infty} E[h_{x-1/m}(\bar{T}_N)] \leq E[s_m(Z)] + \frac{1}{m^2} \leq E[h_x(Z)] + \frac{1}{m^2},$$

και εφόσον αυτή η σχέση ισχύει για κάθε x και κάθε m , αντικαθιστώντας το x με το $x + 1/m$,

$$\limsup_{N \rightarrow \infty} E[h_x(\bar{T}_N)] \leq E[h_{x+1/m}(Z)] + \frac{1}{m^2} \leq \Phi(x + 1/m) + \frac{1}{m^2}.$$

Άρα, δεδομένου αυθαίρετου $\epsilon > 0$, από το γεγονός ότι η $\Phi(x)$ είναι συνεχής συνάρτηση, μπορούμε να επιλέξουμε m_1 αρκετά μεγάλο ώστε $1/m^2 \leq \epsilon/2$ και $\Phi(x + 1/m) \leq \Phi(x) + \epsilon/2$, για κάθε $m \geq m_1$. Τότε,

$$\limsup_{N \rightarrow \infty} E[h_x(\bar{T}_N)] \leq \Phi(x) + \epsilon = E[h_x(Z)] + \epsilon,$$

και επειδή το $\epsilon > 0$ ήταν αυθαίρετο, έχουμε τελικά αποδείξει ότι,

$$\limsup_{N \rightarrow \infty} E[h_x(\bar{T}_N)] \leq E[h_x(Z)]. \quad (13.24)$$

Ο συνδυασμός των αποτελεσμάτων (13.23) και (13.24) συνεπάγεται ότι,

$$\lim_{N \rightarrow \infty} E[h_x(\bar{T}_N)] = E[h_x(Z)],$$

το οποίο αποδεικνύει τον ισχυρισμό του τρίτου βήματος και ολοκληρώνει την απόδειξη. ■

13.5 Ασκήσεις

1. Σύγκλιση κατά πιθανότητα \Rightarrow σύγκλιση κατά κατανομή. Αποδείξτε πως, αν οι Τ.Μ. $\{X_n\}$ τείνουν κατά πιθανότητα σε κάποια Τ.Μ. X , τότε τείνουν στη X και κατά κατανομή.
2. Σύγκλιση κατά κατανομή $\not\Rightarrow$ σύγκλιση κατά πιθανότητα. Εδώ θα δούμε πως η αντίστροφη πρόταση εκείνης που αποδείξαμε στην προηγούμενη άσκηση δεν ισχύει πάντα. Έστω δύο ανεξάρτητες $\text{Bern}(1/2)$ Τ.Μ. Y, Z . Ορίζουμε την ακολουθία Τ.Μ. $X_n = Y + \frac{1}{n}Z$, για κάθε $n = 1, 2, \dots$. Αποδείξτε πως οι $\{X_n\}$:

- (α) Συγκλίνουν κατά κατανομή στην Y .
- (β) Συγκλίνουν κατά κατανομή στη Z .
- (γ) Συγκλίνουν κατά πιθανότητα στην Y .
- (δ) Δεν συγκλίνουν κατά πιθανότητα στη Z .

Άρα έχουμε $X_n \rightarrow Z$ κατά κατανομή αλλά όχι κατά πιθανότητα.

3. Σύγκλιση κατά κατανομή και κατά πιθανότητα σε σταθερά. Αντίθετα με τη γενική περίπτωση της προηγούμενης άσκησης, εδώ θα δείξετε πως, στην ειδική περίπτωση που μια ακολουθία τυχαίων μεταβλητών $\{X_n\}$ συγκλίνει κατά κατανομή σε μια σταθερά $c \in \mathbb{R}$, τότε οι $\{X_n\}$ συγκλίνουν και κατά πιθανότητα στην c .
4. Η ουρά της μέσης τιμής μιας συνεχούς Τ.Μ. Έστω μια συνεχής Τ.Μ. Y η οποία παίρνει πάντα τιμές μεγαλύτερες ή ίσες του μηδενός και έχει πεπερασμένη μέση τιμή $E(Y)$. Όπως είδαμε και στην Άσκηση 3 του Κεφαλαίου 7 για διακριτές Τ.Μ., δείξτε και εδώ πως, καθώς το $x \rightarrow \infty$,

$$E[YH_x(Y)] \rightarrow 0,$$

όπου $H_x(y)$ είναι οι συμπληρωματικές δείκτριες συναρτήσεις τις οποίες ορίσαμε στη σχέση (13.17).

5. Σύγκλιση ως προς την απόσταση χ^2 . Στην Άσκηση 18 του Κεφαλαίου 6, για δύο οποιεσδήποτε διακριτές πυκνότητες $P(x)$ και $Q(x)$ στο ίδιο πεπερασμένο σύνολο S , ορίσαμε τη χ^2 -απόσταση της Q από την P ως,

$$d_{\chi^2}(P, Q) = \sum_{x \in S} \frac{(P(x) - Q(x))^2}{Q(x)}.$$

Έστω μια ακολουθία διακριτών Τ.Μ. $\{X_n\}$ όπου το κάθε X_n έχει πυκνότητα P_n στο πεπερασμένο σύνολο τιμών S , και έστω μια άλλη Τ.Μ. X με πυκνότητα P στο S .

Αποδείξτε πως αν, καθώς $n \rightarrow \infty$, οι αποστάσεις $d_{\chi^2}(P_n, P) \rightarrow 0$, τότε και οι Τ.Μ. $\{X_n\}$ τείνουν στη X κατά κατανομή.

6. Σύγκλιση της διωνυμικής στην *Poisson*. Στο Πρόγραμμα 7.1 του Κεφαλαίου 7 δείξαμε πως, αν για κάθε n η $P_n(k)$ είναι η πυκνότητα μιας Τ.Μ. X_n με Διων($n, \lambda/n$) κατανομή και $P(x)$ είναι η πυκνότητα της $Z \sim \text{Poisson}(\lambda)$, τότε οι Τ.Μ. $\{X_n\}$ τείνουν στη Z υπό την έννοια ότι οι πυκνότητες P_n συγκλίνουν στην P :

$$\lim_{n \rightarrow \infty} P_n(k) = P(k), \quad \text{για κάθε } k = 0, 1, \dots \quad (13.25)$$

- (α') Δείξτε ότι οι $\{X_n\}$ τείνουν στη Z και κατά κατανομή.
- (β') Γενικότερα, έστω $\{X_n\}$ μια οποιαδήποτε ακολουθία Τ.Μ. με αντίστοιχες πυκνότητες $P_n(k)$ και σύνολο τιμών $S = \{0, 1, 2, \dots\}$. Δείξτε πως, αν οι πυκνότητες P_n συγκλίνουν όπως στη (13.25) στην πυκνότητα P κάποιας Τ.Μ. Z με το ίδιο σύνολο τιμών S , τότε οι $\{X_n\}$ τείνουν στη Z και κατά κατανομή.

Κεφάλαιο 14

Παραδείγματα εφαρμογών στη στατιστική

Η στατιστική επιστήμη είναι μία από τις σημαντικότερες περιοχές στις οποίες εφαρμόζονται οι ιδέες και τα μαθηματικά αποτελέσματα των πιθανοτήτων. Στη στατιστική, οι πιθανότητες δεν χρησιμοποιούνται απλά ως μια συλλογή τεχνικών εργαλείων για την επίλυση προβλημάτων, αλλά αποτελούν αναπόσπαστο στοιχείο της ίδιας της θεμελίωσής της: Όπως, γενικώς, λέμε πως τα μαθηματικά αποτελούν τη γλώσσα των θετικών επιστημών, έτσι και η θεωρία πιθανοτήτων είναι η γλώσσα της στατιστικής.

Σε αυτό το κεφάλαιο έχουμε δύο στόχους. Ο πρώτος είναι να αναδείξουμε τη σύνδεση των πιθανοτήτων με κάποιους από τους βασικούς πυλώνες της στατιστικής, και ο δεύτερος είναι να παρουσιάσουμε, κυρίως μέσω επιλεγμένων παραδειγμάτων, μια συνοπτική εισαγωγή στον στατιστικό τρόπο σκέψης.

Σε καθεμία από τις Ενότητες 14.1, 14.2 και 14.3, θα παρουσιάσουμε ένα θέμα της κλασικής στατιστικής και θα δείξουμε τον κεντρικό ρόλο που παίζουν στην ανάλυσή του κάποια από τα αποτελέσματα και τις μεθόδους που έχουμε ήδη συναντήσει σε προηγούμενα κεφάλαια. Ιδιαίτερα στις Ενότητες 14.1 (διαστήματα εμπιστοσύνης) και 14.2 (έλεγχοι υποθέσεων), θα εξετάσουμε προβλήματα των οποίων η περιγραφή, η ανάλυση και η λύση θα αναπτυχθούν με τέτοιο τρόπο ώστε η κύρια έμφαση να είναι στο ευρύτερο πλαίσιο και στη γενική μεθοδολογία με την οποία αντιμετωπίζονται τα περισσότερα ερωτήματα που προκύπτουν στις αντίστοιχες περιοχές.

Η στατιστική είναι, δυστυχώς, μια πολύ – και πολύ άδικα – δυσφημισμένη επιστήμη. Αν και το παρόν βιβλίο μάλλον δεν είναι το καταλληλότερο βήμα, δεν μπορούμε να αντισταθούμε στον πειρασμό να υπενθυμίσουμε τον τεράστιο βαθμό στον οποίο η στατιστική επηρεάζει την καθημερινότητα όλων μας: Από τις κλινικές μελέτες της ιατρικής που μας έχουν διπλασιάσει τον μέσο όρο ζωής, μέχρι τις μεγαλύτερες πολιτικές αποφάσεις, οι οποίες σχεδόν πάντα βασίζονται, εν μέρει τουλάχιστον, σε στατιστικά συμπεράσματα οικονομικών προβλέψεων.

Είναι σχεδόν αδύνατον να διαβάσει κανείς τους κύριους τίτλους ενός ενημερωτικού σάιτ, να δει ειδήσεις στην τηλεόραση ή να ξεφυλλίσει μια εφημερίδα, χωρίς να συναντήσει όχι μόνο μία, αλλά αρκετές αναφορές σε στατιστικά δεδομένα της Eurostat, αποτελέσματα δημοσκοπήσεων, εντυπωσιακές ανακαλύψεις κλινικών ερευνών για καινούργια φάρμακα και ούτω καθεξής. Από αυτήν τη σκοπιά θα μπορούσε να πει κανείς πως μια έστω στοιχειώδης κατανόηση των βασικών στατιστικών εννοιών είναι απαραίτητη ώστε να μπορεί οποιοσδήποτε να λειτουργήσει ως κανονικός πολίτης στη σημερινή κοινωνία.

Αυτό ήταν, σε γενικές γραμμές, το σκεπτικό μιας επιτροπής διασήμων επιστημόνων οι οποίοι, όταν κλήθηκαν πριν 15 περίπου χρόνια να συμβουλέψουν το αμερικανικό κογκρέσο για το περιεχόμενο της ύλης των μαθηματικών που θα έπρεπε να διδάσκεται στα σχολεία, πρότειναν, στην πέμπτη δημοτικού, αμέσως μετά τα κλάσματα και πριν από οτιδήποτε άλλο, να μαθαίνουν τα παιδιά στατιστική. Αυτό το σκεπτικό έχει επηρεάσει και τη δομή των αποτελεσμάτων που θα δούμε σε αυτό το κεφάλαιο. Αφενός θέλουμε να παρουσιάσουμε μια σειρά από ενδιαφέρουσες εφαρμογές ώστε να εμπλουτιστεί η κατανόησή μας για το αντικείμενο των πιθανοτήτων, και αφετέρου ελπίζουμε η επιλογή και ο τρόπος με τον οποίο θα αναπτυχθούν αυτές οι εφαρμογές να αναδειξει τη σημασία και το ευρύτερο, ενδογενές επιστημονικό ενδιαφέρον της στατιστικής.

Ως «απεριτίφ» για το τι θα ακολουθήσει, και ως προειδοποίηση για το πόσο συχνά προκύπτουν λεπτά ζητήματα στατιστικής φύσης στην πράξη, περιγράφουμε συνοπτικά ένα απλό παράδειγμα στο οποίο θα επανέλθουμε πιο διεξοδικά στην Ενότητα 14.2.

Παράδειγμα 14.1 (Καινούργιο φάρμακο) Για να ελεγχθεί η αποτελεσματικότητα ενός νέου φαρμάκου, η φαρμακευτική εταιρία που το παρασκευάζει πραγματοποιεί μια προκαταρκτική κλινική μελέτη με 40 ασθενείς. Σε 20 από αυτούς χορηγείται το πραγματικό φάρμακο, και στους υπόλοιπους 20 χορηγείται εικονικό φάρμακο, το λεγόμενο placebo. Μετά από ένα εύλογο χρονικό διάστημα παρατηρούμε ότι το 60% εκείνων που παίρνουν το φάρμακο σημειώνουν βελτίωση, ενώ το αντίστοιχο ποσοστό για εκείνους που παίρνουν το placebo είναι μόνο 40%. Πόσο αξιόπιστη ένδειξη της επιτυχίας του φαρμάκου μπορεί να θεωρηθεί το παραπάνω αποτέλεσμα; Μπορείτε να βρείτε και να τεκμηριώσετε μια ποσοτική απάντηση σε αυτό το ερώτημα;

14.1 Διαστήματα εμπιστοσύνης

Στα Κεφάλαια 9 και 12 είδαμε πολλά παραδείγματα όπου, προκειμένου να εκτιμήσουμε μια άγνωστη ποσότητα, χρησιμοποιήσαμε τον εμπειρικό μέσο όρο,

$$\bar{X}_N = \frac{1}{N} \sum_{i=1}^N X_i,$$

N ανεξάρτητων τυχαίων δειγμάτων X_1, X_2, \dots, X_N . Για παράδειγμα:

- **Δημοσκόπηση.** Αν σε μια δημοσκόπηση το κάθε X_i περιγράφει την απάντηση ενός ψηφοφόρου για κάποιο συγκεκριμένο κόμμα ($X_i = 1$ αν σκοπεύει να το ψηφίσει, $X_i = 0$ αν όχι), τότε ο εμπειρικός μέσος όρος \bar{X}_N μας δίνει μια εκτίμηση για το ποσοστό p των ψήφων που θα πάρει αυτό το κόμμα στις εκλογές.
- **Εκτίμηση πιθανότητας.** Γενικότερα, αν τα X_i είναι «δείκτριες» Τ.Μ., δηλαδή το κάθε X_i ισούται με 1 όποτε κάποιο ενδεχόμενο A συμβαίνει και με 0 όποτε όχι, τότε το \bar{X}_N μας δίνει μια εκτίμηση της πιθανότητας $p = \Pr(A)$ με την οποία συμβαίνει το A .
- **Δημογραφική δειγματοληψία.** Όταν επιλέγονται άτομα από έναν πληθυσμό τυχαία και το κάθε X_i αναπαριστά την τιμή κάποιου χαρακτηριστικού που μας ενδιαφέρει – π.χ., το εισόδημα, την ηλικία ή το πλήθος των μελών της οικογένειας του ατόμου i – τότε το \bar{X}_N δίνει μια εκτίμηση της μέσης τιμής του αντίστοιχου χαρακτηριστικού του πληθυσμού, δηλαδή του μέσου εισοδήματος, της μέσης ηλικίας, του μέσου μεγέθους της οικογένειας, κλπ.
- **Εμπειρικές μετρήσεις.** Σε περιπτώσεις που το ζητούμενο είναι να εκτιμηθεί η μέση τιμή μ κάποιας ποσότητας, όπως για παράδειγμα η μέση διάρκεια μιας τηλεφωνικής κλήσης σε κάποιο δίκτυο, η μέση ατμοσφαιρική θερμοκρασία στην ανταρκτική, η μέση τιμή πώλησης ενός χρηματιστηριακού προϊόντος ή η μέση περιεκτικότητα του νερού της ΕΥΔΑΠ σε μόλυβδο, τότε ο εμπειρικός μέσος όρος \bar{X}_N από N μετρήσεις X_1, X_2, \dots, X_N της αντίστοιχης ποσότητας μας δίνει μια εκτίμηση της άγνωστης μέσης τιμής $\mu = E(X_i)$.

Σε καθένα από τα παραπάνω παραδείγματα, αν τα X_i είναι επιλεγμένα έτσι ώστε να είναι ανεξάρτητα και να έχουν όλα την ίδια κατανομή, ο Ν.Μ.Α. μάς διαβεβαιώνει πως, για μεγάλα μεγέθη δειγμάτων N :

Ο εμπειρικός μέσος όρος $\bar{X}_N \approx \mu$, με μεγάλη πιθανότητα.

Εδώ $\mu = E(X_i)$, η κοινή μέση τιμή των X_i , είναι η ποσότητα που θέλουμε να εκτιμήσουμε. Επιπλέον, στην απόδειξη του Ν.Μ.Α. στο Κεφάλαιο 9 είδαμε πως η εκτίμηση \bar{X}_N έχει,

$$E(\bar{X}_N) = \mu \text{ και } \text{Var}(\bar{X}_N) = \frac{\sigma^2}{N}, \quad (14.1)$$

όπου $\sigma^2 = \text{Var}(X_i)$ είναι η κοινή διασπορά των X_i . Και σύμφωνα με το Κ.Ο.Θ. του Κεφαλαίου 12, το \bar{X}_N έχει κατά προσέγγιση κανονική κατανομή. Για μεγάλα N :

$$\text{Η κατανομή του } \bar{X}_N \text{ είναι } \approx N\left(\mu, \frac{\sigma^2}{N}\right). \quad (14.2)$$

Το κύριο αντικείμενο αυτής της ενότητας είναι να εξετασθεί το πώς μπορεί να διατυπωθεί με ακρίβεια το ενδεχόμενο μέγεθος και η πιθανότητα του «στατιστικού σφάλματος» σε τέτοιου είδους εκτιμήσεις.

Παράδειγμα 14.2 (Μια απλή δημοσκόπηση) Στις εθνικές εκλογές του Ιανουαρίου του 2015, το κόμμα της Νέας Δημοκρατίας (ΝΔ) έλαβε ποσοστό 27.81% επί των (έγκυρων) ψηφοδελτίων. Λίγο πριν τις εκλογές η γνωστή εταιρία δημοσκοπήσεων alco, βασισμένη σε ένα τυχαίο δείγμα $N = 1000$ ψηφοφόρων, ανακοίνωσε, μεταξύ άλλων αποτελεσμάτων, την πρόβλεψη πως το ποσοστό της ΝΔ θα ήταν 30.5% με «στατιστικό σφάλμα» $\pm 3.1\%$ και «διάστημα βεβαιότητας» 95%. Τι ακριβώς σημαίνει στατιστικό σφάλμα και διάστημα βεβαιότητας;

Ας μπούμε για λίγο στην θέση των δημοσκόπων. Έστω $p = 0.2781$ το πραγματικό ποσοστό των ψηφοφόρων της ΝΔ, και έστω X_1, X_2, \dots, X_N οι απαντήσεις $N = 1000$ ατόμων, όπου κάθε $X_i = 1$ αν το άτομο i είναι ψηφοφόρος της ΝΔ, και $X_i = 0$ αν όχι. Λογικά θεωρούμε ότι τα X_i είναι ανεξάρτητες $\text{Bern}(p)$ Τ.Μ. και η πρόβλεψή μας βάσει αυτών των στοιχείων είναι ο εμπειρικός μέσος \bar{X}_N . Για να είμαστε, τώρα, σε θέση να ανακοινώσουμε τα αποτελέσματά μας στα ΜΜΕ, αναρωτιόμαστε πώς θα μπορούσαμε να βρούμε ένα διάστημα τιμών της μορφής $[\bar{X}_N - e, \bar{X}_N + e]$, τέτοιο ώστε η πιθανότητα να περιέχει το πραγματικό αποτέλεσμα p να είναι τουλάχιστον 95%. Με άλλα λόγια, πόσο μεγάλο «στατιστικό σφάλμα e » πρέπει να δεχθούμε ώστε το «διάστημα εμπιστοσύνης» $[\bar{X}_N - e, \bar{X}_N + e]$ να έχει πιθανότητα τουλάχιστον 95%;

Σε αυτό το (ομολογουμένως, υπεραπλουστευμένο) παράδειγμα, όπου ξέρουμε το πραγματικό αποτέλεσμα p , μπορούμε εύκολα να απαντήσουμε, με μια απλή εφαρμογή του Κ.Ο.Θ. [Δείτε την Άσκηση 4 στο τέλος του κεφαλαίου για την πιο ρεαλιστική περίπτωση όπου δεν ξέρουμε εκ των προτέρων το p .] Έστω $Y = N\bar{X}_N$ το πλήθος των ψηφοφόρων της ΝΔ στην έρευνά μας, οπότε η $Y \sim \text{Διων}(N, p)$. Από το Πόρισμα 12.1, έχουμε ότι η πιθανότητα που μας ενδιαφέρει,

$$\Pr(p \in [\bar{X}_N - e, \bar{X}_N + e]) = \Pr(\bar{X}_N \in [p - e, p + e]) = \Pr((p - e)N \leq Y \leq N(p + e)),$$

κατά προσέγγιση ισούται με,

$$\begin{aligned} & \Phi\left(\frac{N(p + e) - Np + \frac{1}{2}}{\sqrt{p(1 - p)N}}\right) - \Phi\left(\frac{N(p - e) - Np - \frac{1}{2}}{\sqrt{p(1 - p)N}}\right) \\ &= \Phi\left(\frac{Ne + \frac{1}{2}}{\sqrt{Np(1 - p)}}\right) - \Phi\left(\frac{-Ne - \frac{1}{2}}{\sqrt{Np(1 - p)}}\right) \\ &= 2\Phi\left(\frac{Ne + \frac{1}{2}}{\sqrt{Np(1 - p)}}\right) - 1, \end{aligned}$$

το οποίο, για να ισούται με 95%, πρέπει να έχουμε,

$$\Phi\left(\frac{Ne + \frac{1}{2}}{\sqrt{Np(1 - p)}}\right) \approx 0.975, \text{ δηλαδή, } \frac{Ne + \frac{1}{2}}{\sqrt{Np(1 - p)}} \approx 1.96,$$

όπου χρησιμοποιήσαμε τους πίνακες τιμών της τυπικής κανονικής κατανομής από το Κεφάλαιο 12 για να βρούμε την τιμή $z = 1.96$ για την οποία η $\Phi(z)$ είναι όσο το δυνατόν πιο κοντά στο 0.975.

Λύνοντας ως προς e βρίσκουμε,

$$e = \frac{1}{N} \left[1.96\sqrt{Np(1-p)} - \frac{1}{2} \right] \approx 0.0273 = 2.73\%.$$

Οπότε καταλήγουμε στο εξής συμπέρασμα:

Η πρόβλεψή μας είναι πως η ΝΔ θα λάβει ποσοστό $\hat{p} = \bar{X}_N$, με διάστημα εμπιστοσύνης $\hat{p} \pm 2.73\%$, ή $[\hat{p} - 0.023, \hat{p} + 0.023]$, το οποίο έχει επίπεδο εμπιστοσύνης 95%.

Παρατήρηση: Όπως είδαμε στις σχέσεις (14.1) και (14.2) παραπάνω, η ακρίβεια της εκτίμησης \bar{X}_N για τη μέση τιμή μ των X_i περιγράφεται πρωτίστως από τη διασπορά $\text{Var}(\bar{X}_N) = \sigma^2/N$, η οποία εξαρτάται από τη διασπορά $\sigma^2 = \text{Var}(X_i)$ των X_i και το μέγεθος N του δείγματος. Άρα, προκειμένου να είναι σχετικά ακριβής η εκτίμηση $\bar{X}_N \approx \mu$, θα πρέπει είτε το μέγεθος N του δείγματος να είναι σχετικά μεγάλο, είτε η διασπορά σ^2 να είναι αντίστοιχα μικρή. Στην Ενότητα 14.3 θα δούμε μια απλή μέθοδο μείωσης της διασποράς, η οποία χρησιμοποιείται κυρίως σε περιπτώσεις προσομοιωμένων δειγμάτων X_i . Στο ακόλουθο παράδειγμα θα δούμε με ποιον τρόπο ο υπολογισμός του διαστήματος εμπιστοσύνης σχετίζεται με την επιλογή του μεγέθους N του δείγματος.

Παράδειγμα 14.3 Για να μετρηθεί η δυσκολία μιας άσκησης, καταγράφηκαν οι χρόνοι X_i , σε λεπτά, που χρειάστηκαν $N = 114$ πρωτοετείς φοιτητές για να τη λύσουν ($i = 1, 2, \dots, N$). Υποθέτοντας ότι η τυπική απόκλιση $\sigma = \pm 6$ λεπτά του τυχαίου χρόνου που απαιτεί η λύση της άσκησης είναι γνωστή, εκτιμήσαμε τον μέσο χρόνο λύσης μ μέσω του εμπειρικού μέσου \bar{X}_N των X_i , ο οποίος βρέθηκε να είναι $\bar{X}_N = 9.31$ λεπτά.

Εστω πως θέλουμε να δώσουμε ένα διάστημα εμπιστοσύνης της μορφής $[\bar{X}_N - e, \bar{X}_N + e]$ για την εκτίμηση \bar{X}_N , με επίπεδο εμπιστοσύνης $c = 90\%$, δηλαδή θέλουμε να ισχύει ότι,¹

$$\Pr(\mu \in [\bar{X}_N - e, \bar{X}_N + e]) = c. \quad (14.3)$$

Βάσει του Κ.Ο.Θ., αυτή η πιθανότητα μπορεί να προσεγγιστεί ως,

$$\begin{aligned} \Pr(\mu \in [\bar{X}_N - e, \bar{X}_N + e]) &= \Pr(|\bar{X}_N - \mu| \leq e) \\ &= \Pr\left(\left|\frac{\sqrt{N}}{\sigma}(\bar{X}_N - \mu)\right| \leq \frac{e\sqrt{N}}{\sigma}\right) \\ &\approx \Pr\left(|Z| \leq \frac{e\sqrt{N}}{\sigma}\right) \\ &= 1 - 2\Phi\left(-\frac{e\sqrt{N}}{\sigma}\right), \end{aligned}$$

¹Επιλέγουμε το συμβολισμό e για το στατιστικό σφάλμα από το αρχικό γράμμα της αγγλικής λέξης «error», δηλαδή «σφάλμα», και το συμβολισμό c για την πιθανότητα του διαστήματος εμπιστοσύνης από την αγγλική λέξη «confidence» που σημαίνει «εμπιστοσύνη».

όπου η Z έχει τυπική κανονική κατανομή και χρησιμοποιήσαμε και πάλι την παρατήρηση ότι το κανονικοποιημένο άθροισμα \bar{S}_N των X_i μπορεί να εκφραστεί (βλ. (12.3) στο Κεφάλαιο 12) ως,

$$\bar{S}_N = \frac{\sqrt{N}}{\sigma} (\bar{X}_N - \mu).$$

Άρα, για να ισχύει η (14.3) πρέπει να έχουμε, $1 - 2\Phi(-e\sqrt{N}/\sigma) = c$ ή, ισοδύναμα,

$$e\sqrt{N} = -\sigma\Phi^{-1}\left(\frac{1-c}{2}\right), \quad (14.4)$$

όπου Φ^{-1} είναι η αντίστροφη συνάρτηση της τυπικής κανονικής συνάρτησης κατανομής Φ (δηλαδή, $\Phi^{-1}(x) = z$ αν και μόνο αν $\Phi(z) = x$). Με τα δεδομένα του προβλήματος, $N = 114$, $\sigma = 6$ και $c = 0.9$, η (14.4) μας δίνει,

$$e \approx -\frac{6}{\sqrt{114}} \times \Phi^{-1}(0.05) \approx -0.562 \times (-1.645) \approx 0.924,$$

όπου από τον πίνακα τιμών της τυπικής κανονικής συνάρτησης κατανομής στην Ενότητα 12.4 επιλέξαμε την τιμή -1.645 ως μια καλή προσέγγιση στην $\Phi^{-1}(0.05)$, μια που $\Phi(-1.64) = 0.0505$ και $\Phi(-1.65) = 0.0495$. Άρα, συμπεραίνουμε πως ο μέσος χρόνος για τη λύση της συγκεκριμένης άσκησης είναι 9.31 λεπτά, με «στατιστικό σφάλμα» ± 0.924 λεπτά και «επίπεδο εμπιστοσύνης» 90%.

Έστω όμως πως δεν είμαστε ικανοποιημένοι με αυτό το αποτέλεσμα και θέλουμε οπωσδήποτε ένα μικρότερο διάστημα εμπιστοσύνης με απόκλιση $e =$ μισό λεπτό. Προκειμένου να είναι ακριβέστερη η εκτίμηση, πόσο μεγαλύτερο μέγεθος δείγματος είναι απαραίτητο; Αν θέσουμε $e = 0.5$, καταφεύγοντας στη γενική έκφραση (14.4), μπορούμε να λύσουμε ως προς N :

$$N = \left[\frac{\sigma}{e} \Phi^{-1}\left(\frac{1-c}{2}\right) \right]^2 \approx \left[\frac{6}{1/2} \times (-1.645) \right]^2 \approx 389.67.$$

Συνεπώς, προκειμένου να είμαστε σε θέση να δώσουμε ένα διάστημα εμπιστοσύνης της μορφής $\bar{X}_N \pm 0.5$ με επίπεδο εμπιστοσύνης 90%, απαιτούνται τουλάχιστον 390 δείγματα X_i .

Παρατήρηση: Αν και ο υπολογισμός που οδήγησε στη σχέση (14.4) έγινε με αφορμή το Παράδειγμα 14.3, υπογραμμίζουμε ότι μέχρι εκείνο το σημείο δεν είχε χρησιμοποιηθεί κανένα από τα δεδομένα του συγκεκριμένου προβλήματος. Άρα η συσχέτιση του στατιστικού σφάλματος e με το επίπεδο εμπιστοσύνης c και το μέγεθος του δείγματος N , όπως εκφράζεται από τη (14.4), ισχύει κάτω από γενικές συνθήκες. Οι γενικές συνέπειες αυτής της σχέσης παρατίθενται επιγραμματικά παρακάτω.

Διαστήματα εμπιστοσύνης: Γνωστή διασπορά. Από τα τυχαία δείγματα X_1, \dots, X_N με γνωστή διασπορά $\sigma^2 = \text{Var}(X_i)$ υπολογίζουμε τον εμπειρικό μέσο \bar{X}_N προκειμένου να εκτιμήσουμε την άγνωστη μέση τιμή $\mu = E(X_i)$.

1. Για το επιθυμητό **επίπεδο εμπιστοσύνης** c βρίσκουμε την τιμή $z^* = -\Phi^{-1}\left(\frac{1-c}{2}\right)$. (Στο Σχήμα 14.1 δίνονται οι πιο συνηθισμένες τιμές του z^* .)
2. Αν το μέγεθος του δείγματος N είναι δεδομένο, τότε από τη σχέση (14.4) υπολογίζουμε το μέγεθος του **στατιστικού σφάλματος** e ,

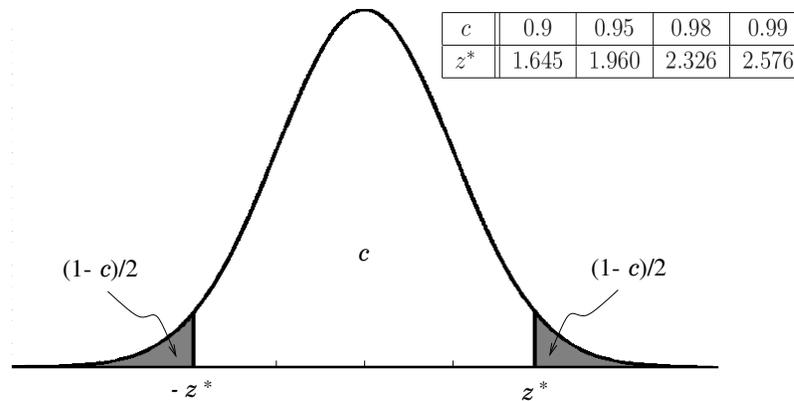
$$e = \frac{\sigma z^*}{\sqrt{N}}, \quad (14.5)$$

και δίνουμε ως **διάστημα εμπιστοσύνης με επίπεδο εμπιστοσύνης** c ή, πιο απλά, ως **c -διάστημα εμπιστοσύνης**, το $[\bar{X}_N - e, \bar{X}_N + e]$.

3. Αν το ζητούμενο είναι να επιτευχθεί κάποιο επιθυμητό στατιστικό σφάλμα e , από τη σχέση (14.4) υπολογίζουμε το **απαραίτητο μέγεθος του δείγματος** N ,

$$N = \left(\frac{\sigma z^*}{e}\right)^2, \quad (14.6)$$

ώστε το διάστημα εμπιστοσύνης $[\bar{X}_N - e, \bar{X}_N + e]$ να έχει επίπεδο εμπιστοσύνης c .



Σχήμα 14.1: Γραφική αναπαράσταση του υπολογισμού της τιμής $z^* = -\Phi^{-1}\left(\frac{1-c}{2}\right)$ που αντιστοιχεί στο επίπεδο εμπιστοσύνης c . Στον πίνακα σημειώνονται οι πιο συνηθισμένες τιμές του z^* .

Παρατήρηση: Σε ρεαλιστικά προβλήματα όπου η τυπική απόκλιση σ των δειγμάτων συχνά είναι άγνωστη, η «ορθόδοξη» στατιστική πρακτική είναι να εκτιμήσουμε πρώτα τη σ και κατόπιν να υπολογίσουμε το σχετικό διάστημα εμπιστοσύνης βάσει του κανονικοποιημένου αθροίσματος,

$$\hat{S}_N = \frac{\sqrt{N}}{\hat{\sigma}}(\bar{X}_N - \mu),$$

όπου $\hat{\sigma}$ είναι η εκτίμησή μας για την τυπική απόκλιση των X_i . Αλλά, σε αντίθεση με το \bar{S}_N , για το οποίο το Κ.Ο.Θ. μάς λέει πως έχει κατά προσέγγιση κανονική κατανομή, η κατανομή του νέου κανονικοποιημένου αθροίσματος \bar{S}_N είναι διαφορετική. Αν και δεν θα επεκταθούμε περαιτέρω σε αυτή την κατεύθυνση, θα δούμε έναν πιο απλό (αλλά λιγότερο ακριβή) τρόπο με τον οποίο μπορεί να αντιμετωπιστεί το πρόβλημα της άγνωστης διασποράς σε ορισμένες περιπτώσεις.

Παράδειγμα 14.4 Η αμερικανική περιοδική έκδοση *Student Monitor* διανέμει ερωτηματολόγια σε $N = 1200$ φοιτητές 100 διαφορετικών πανεπιστημίων των Η.Π.Α. δύο φορές το χρόνο, προκειμένου να καταγράψει τις νέες τάσεις στον τρόπο ζωής τους. Πρόσφατα ανακοίνωσε την εκτίμηση ότι κατά μέσο όρο οι φοιτητές περνούν 15.1 ώρες την εβδομάδα «σερφάροντας» στο διαδίκτυο. Εδώ, αν η τυπική απόκλιση σ του τυχαίου χρόνου X_i που περνά στο διαδίκτυο ο κάθε φοιτητής i ήταν γνωστή, θα μπορούσαμε υπολογίσουμε το στατιστικό σφάλμα e με επίπεδο εμπιστοσύνης $c = 95\%$ για την εκτίμηση $\bar{X}_N = 15.1$ από την έκφραση (14.5) παραπάνω, $e = 1.96\sigma/\sqrt{1200}$.

Ας υποθέσουμε πως, αν και άγνωστη, η διασπορά $\sigma^2 = \text{Var}(X_i)$ είναι μεταξύ 4 και 7.5, δηλαδή η αντίστοιχη τυπική απόκλιση είναι $2 \leq \sigma \leq \sqrt{7.5}$. Τότε, από τον παραπάνω υπολογισμό έχουμε ότι το στατιστικό σφάλμα $1.96\sigma/\sqrt{1200}$ είναι μεταξύ 0.113 και 0.155, και, προκειμένου να είμαστε σίγουροι για το αποτέλεσμά μας, υιοθετούμε τη χειρότερη περίπτωση, λέγοντας πως η εκτίμηση $\bar{X}_N = 15.1$ έχει στατιστικό σφάλμα το πολύ ± 0.155 ωρών, με επίπεδο εμπιστοσύνης 95%.

Η ίδια λογική οδηγεί και στην παρακάτω γενική μέθοδο.

Διαστήματα εμπιστοσύνης: Άγνωστη διασπορά. Υπολογίζουμε τον εμπειρικό μέσο \bar{X}_N των τυχαίων δειγμάτων X_1, X_2, \dots, X_N προκειμένου να εκτιμήσουμε την άγνωστη μέση τιμή $\mu = E(X_i)$.

1. Από τα δεδομένα του προβλήματος υπολογίζουμε ένα άνω φράγμα $\sigma \leq \sigma_{\max}$ για την τυπική απόκλιση $\sigma = \sqrt{\text{Var}(X_i)}$ των X_i .
2. Για το επιθυμητό επίπεδο εμπιστοσύνης c βρίσκουμε την τιμή $z^* = -\Phi^{-1}\left(\frac{1-c}{2}\right)$.
3. Αν το μέγεθος του δείγματος N είναι δεδομένο, τότε μπορούμε να υπολογίσουμε ένα άνω φράγμα για το μέγεθος του στατιστικού σφάλματος e ,

$$e \leq \frac{\sigma_{\max} z^*}{\sqrt{N}}, \quad (14.7)$$

και να πούμε πως έχουμε c -διάστημα εμπιστοσύνης (ή διάστημα εμπιστοσύνης με επίπεδο εμπιστοσύνης c) όχι μεγαλύτερο από το $[\bar{X}_N - e, \bar{X}_N + e]$.

4. Αν το ζητούμενο είναι να επιτευχθεί κάποιο επιθυμητό στατιστικό σφάλμα e , τότε μπορούμε να υπολογίσουμε ένα άνω φράγμα για το απαραίτητο μέγεθος του δείγματος N ,

$$N \leq \left(\frac{\sigma_{\max} z^*}{e}\right)^2, \quad (14.8)$$

ώστε το διάστημα εμπιστοσύνης $[\bar{X}_N - e, \bar{X}_N + e]$ να έχει επίπεδο εμπιστοσύνης c .

Οι ασκήσεις στο τέλος του κεφαλαίου περιέχουν αρκετά ακόμα ενδιαφέροντα παραδείγματα διαστημάτων εμπιστοσύνης. Κλείνουμε αυτή την ενότητα με ένα παράδειγμα διαφορετικού τύπου.

Παράδειγμα 14.5 Για μια συγκεκριμένη μετοχή στο χρηματιστήριο, θέλουμε να προβλέψουμε αν η τιμή της την επόμενη μέρα θα ανέβει ή θα πέσει. Έστω $Y_i = 1$ αν η τιμή ανέβηκε τη μέρα i και $Y_i = 0$ αν έπεσε. (Για ευκολία υποθέτουμε ότι η τιμή δεν μένει ποτέ σταθερή από μέρα σε μέρα· παρότι αυτό δεν είναι απολύτως ακριβές, δεν επηρεάζει την ουσία του προβλήματος.) Σε πολλά εμπειρικά δεδομένα έχει παρατηρηθεί ότι τα στατιστικά χαρακτηριστικά τέτοιων $\{Y_i\}$ είναι σχεδόν ίδια με αυτά μιας ακολουθίας ανεξάρτητων $\text{Bern}(1/2)$ Τ.Μ., και γι' αυτόν το λόγο θεωρείται σχεδόν αδύνατη η πρόβλεψή τους.

Έστω, τώρα, πως κάποιος χρηματιστής ισχυρίζεται ότι μπορεί να προβλέψει τα Y_i με ακρίβεια μεγαλύτερη από 50%. Κάτι τέτοιο θα συνεπαγόταν πολύ μεγάλα χρηματικά κέρδη, οπότε θέλουμε να βρούμε έναν τρόπο να ελέγξουμε τον ισχυρισμό του. Του ζητάμε λοιπόν να προβλέψει, για N συνεχόμενες μέρες, αν η τιμή της μετοχής θα ανέβει ή θα πέσει και καταγράφουμε τα αποτελέσματα $\{X_i\}$, όπου $X_i = 1$ αν τη μέρα i η πρόβλεψή του ήταν σωστή και $X_i = 0$ αν όχι. Υπάρχει κάποιος τρόπος, βάσει των $\{X_i\}$, να επιβεβαιώσουμε ή να απορρίψουμε τον ισχυρισμό του χρηματιστή; Για παράδειγμα, αν μεταξύ $N = 120$ ημερών κατάφερε να προβλέψει σωστά τις 71 (δηλαδή τα 71 από τα X_i είναι 1 και μόνο τα 49 είναι 0), μπορούμε να πούμε πως πράγματι η φαινομενική επιτυχία της πρόβλεψής του δεν δικαιολογείται από τις τυχαίες διακυμάνσεις των τιμών της μετοχής;

Για τέτοιες περιπτώσεις στατιστικών προβλημάτων, στις οποίες η ζητούμενη απάντηση δεν είναι ποσοτική (όπως τα διαστήματα εμπιστοσύνης) αλλά της μορφής ΝΑΙ/ΟΧΙ, απαιτείται η ανάπτυξη μιας διαφορετικής μεθοδολογίας, αυτή των *ελέγχων υποθέσεων*.

14.2 Έλεγχος υποθέσεων

Παράδειγμα 14.6 (Καινούργιο φάρμακο) Επανερχόμαστε στο ερώτημα του Παραδείγματος 14.1. Σε 20 ασθενείς χορηγείται ένα νέο φάρμακο, σε άλλους 20 χορηγείται εικονικό φάρμακο (placebo), και παρατηρούμε ότι 60% εκείνων που παίρνουν το φάρμακο σημειώνουν βελτίωση ενώ το αντίστοιχο ποσοστό για εκείνους που παίρνουν το placebo είναι μόνο 40%. Θέλουμε να εξετάσουμε κατά πόσο αυτή η ένδειξη για την αποτελεσματικότητα του φαρμάκου είναι στατιστικά αξιόπιστη. Με άλλα λόγια, αναρωτιόμαστε αν, ακόμα και στην περίπτωση που το φάρμακο δεν έχει καμία επίπτωση στην πορεία της συγκεκριμένης ασθένειας, το αποτέλεσμα δικαιολογείται από απλές τυχαίες διακυμάνσεις.

Σε αυτό το απλό παράδειγμα μπορεί να δοθεί μια ακριβής απάντηση σχετικά εύκολα. Ας υποθέσουμε ότι η πιθανότητα να βελτιωθεί η υγεία ενός οποιουδήποτε ασθενή είναι 50% και ότι το φάρμακο δεν επηρεάζει καθόλου την πορεία της ασθένειας. Ορίζοντας για κάθε i μια Τ.Μ. X_i η οποία ισούται με 1 αν ο ασθενής i παρουσίασε βελτίωση και $X_i = 0$ αν όχι, οι παραπάνω υποθέσεις μάς λένε πως οι X_1, X_2, \dots, X_{2M} , με $M = 20$, είναι ανεξάρτητες $\text{Bern}(1/2)$ Τ.Μ. Επιπλέον, θεωρούμε ότι οι ασθενείς $i = 1$ έως 20 πήραν το φάρμακο ενώ οι $i = 21$ έως 40 πήραν

το placebo, και θέτουμε,

$$Y = X_1 + X_2 + \dots + X_N = \text{πλήθος ασθενών που έδειξαν βελτίωση με το φάρμακο,}$$

$$Z = X_{N+1} + X_{N+2} + \dots + X_{2M} = \text{πλήθος ασθενών που έδειξαν βελτίωση με το placebo,}$$

έτσι ώστε οι T.M. Y και Z είναι ανεξάρτητες και έχουν $\Delta_{\text{ιων}}(N, 1/2)$ κατανομή. Με αυτόν το συμβολισμό το αποτέλεσμα που παρατηρήθηκε είναι ότι $Y = 12$ και $Z = 8$, και το αρχικό ερώτημα γίνεται: Δεδομένης της υπόθεσης ότι το φάρμακο δεν έχει καμία επίδραση, είναι το αποτέλεσμα « $Y = 12$ και $Z = 8$ » αρκετά «απίθανο» ώστε να απορρίψουμε αυτή την υπόθεση; Η απόκλιση των τιμών Y και Z των δύο ομάδων ασθενών είναι στατιστικά σημαντική ή όχι;

Για να απαντήσουμε, υπολογίζουμε την πιθανότητα να παρατηρηθεί εντελώς τυχαία μια τόσο μεγάλη ή ακόμα μεγαλύτερη απόκλιση, δηλαδή την,

$$\begin{aligned} \Pr(|Y - Z| \geq 4) &= 1 - \Pr(|Y - Z| \leq 3) \\ &= 1 - \sum_{k=0}^{20} \Pr(|Y - Z| \leq 3, Y = k) \\ &= 1 - \sum_{k=0}^{20} \Pr(|Z - k| \leq 3, Y = k) \\ &= 1 - \sum_{k=0}^{20} \Pr(|Z - k| \leq 3) \Pr(Y = k) \\ &= 1 - \sum_{k=0}^{20} \Pr(Y = k) \sum_{\ell=k-3}^{k+3} \Pr(Z = \ell), \end{aligned}$$

όπου χρησιμοποιήσαμε τον κανόνα συνολικής πιθανότητας και την ανεξαρτησία των Y, Z . Εφόσον όλες οι πιθανότητες $\Pr(Y = k)$ και $\Pr(Z = \ell)$ μας είναι γνωστές από τον τύπο της $\Delta_{\text{ιων}}(20, 1/2)$ πυκνότητας, κάνοντας τον σχετικό (μακροσκελή αλλά απλό) υπολογισμό βρίσκουμε ότι,

$$\Pr(|Y - Z| \geq 4) \approx 0.268.$$

Συνεπώς: Υπάρχει πιθανότητα μεγαλύτερη από μία στις τέσσερις να παρουσιαστεί από καθαρή τυχαιότητα απόκλιση στις τιμές των δύο ομάδων ασθενών τόσο μεγάλη όσο αυτή που παρατηρήσαμε. Έτσι, συμπεραίνουμε πως το αποτέλεσμα αυτής της προκαταρκτικής μελέτης μπορεί να αποδοθεί σε τυχαίο γεγονός με μη αμελητέα πιθανότητα.

Αν και η παραπάνω απάντηση είναι απολύτως ικανοποιητική, η εφαρμογή της μεθόδου με την οποία προσεγγίσαμε το ερώτημα είναι σαφώς περιορισμένη στο συγκεκριμένο πρόβλημα. Παρακάτω, στην Ενότητα 14.2.2, θα δούμε ποια είναι η συνηθισμένη «ορθόδοξη» προσέγγιση για γενικά προβλήματα αυτής της μορφής.

Παράδειγμα 14.7 (Δίκαιο ή κάλπικο νόμισμα;) Πριν παίξουμε με κάποιον Κορώνα-Γράμματα, θέλουμε να ελέγξουμε αν το νόμισμα είναι δίκαιο ή όχι. Το στρίβουμε 210 φορές και φέρνουμε 111 Κ και 99 Γ. Τι μπορούμε να συμπεράνουμε;

Πριν απαντήσουμε, παρατηρούμε ότι η μορφή αυτού του ερωτήματος είναι μαθηματικά ακριβώς ίδια με το πρόβλημα του χρηματιστή στο Παράδειγμα 14.5, όπως και με πολλά άλλα προβλήματα που εμφανίζονται συχνά στην πράξη. Για παράδειγμα, κάποιος ισχυρίζεται ότι είναι μέντιουμ και μπορεί να προβλέψει τα αποτελέσματα ενός δίκαιου νομίσματος με πιθανότητα μεγαλύτερη από 50%. Τον προκαλούμε να μας κάνει μια επίδειξη, και πανηγυρίζει μαντεύοντας σωστά τις 111 από τις 210 φορές. Είναι πράγματι μέντιουμ ή όχι;

Σε όλες αυτές τις περιπτώσεις, έχουμε N ανεξάρτητες $\text{Bern}(p)$ Τ.Μ. X_1, X_2, \dots, X_N , και αφού παρατηρήσουμε τις τιμές τους καλούμαστε να απαντήσουμε το ερώτημα, «είναι το $p = 1/2$ ή όχι;». Μια πρώτη προσέγγιση του προβλήματος είναι μέσω των διαστημάτων εμπιστοσύνης. Ο εμπειρικός μέσος των $N = 210$ ρίψεων ισούται με $\bar{X}_N = 111/210 \approx 0.5286$ και, προκειμένου να αποφασίσουμε πόσο πιστευτό είναι το ενδεχόμενο να έχουμε $p = 1/2$, βρίσκουμε ένα 95%-διάστημα εμπιστοσύνης για την εκτίμηση $\hat{p} = 0.5286$. Από τη γενική μέθοδο της προηγούμενης παραγράφου, για $c = 0.95$ βρίσκουμε την τιμή $z^* = 1.96$ και υπολογίζουμε το στατιστικό σφάλμα e από τη σχέση $e = \sigma_{\max} z^* / \sqrt{N}$, όπου για την τιμή της τυπικής απόκλισης $\sigma = \sqrt{\text{Var}(X_i)} = \sqrt{p(1-p)}$ επιλέγουμε τη μέγιστη δυνατή τιμή της, $\sigma_{\max} = 1/2$, η οποία αντιστοιχεί στην περίπτωση $p = 1/2$. Έτσι, βρίσκουμε ότι $e \approx 0.0676$ και καταλήγουμε στο 95%-διάστημα εμπιστοσύνης $[0.461, 0.596]$.

Τι μπορούμε να συμπεράνουμε από τον παραπάνω υπολογισμό; Η απάντηση αγγίζει ένα από τα πιο λεπτά ζητήματα της στατιστικής. Εφόσον το $p = 1/2$ είναι εντός του 95%-διαστήματος εμπιστοσύνης $[0.461, 0.596]$, ένα σαφές πρώτο συμπέρασμα είναι ότι δεν μπορούμε να απορρίψουμε την περίπτωση $p = 1/2$ το νόμισμα να είναι δίκαιο. Αυτό που ίσως εκ πρώτης όψεως να μην είναι προφανές, είναι πως το παραπάνω είναι το *μόνο* συμπέρασμα που μπορούμε να συνάγουμε από τις παρατηρήσεις μας. Συγκεκριμένα, δεν μπορούμε να πούμε ότι «επιβεβαιώσαμε» ή ότι «αποδεχόμαστε» την υπόθεση ότι τα δεδομένα είναι ανεξάρτητα με κατανομή $\text{Bern}(1/2)$. Το μόνο στατιστικά ασφαλώς τεκμηριωμένο συμπέρασμα εδώ είναι ότι: *Βάσει των παρατηρήσεων $\{X_i\}$, σε επίπεδο σημαντικότητας 5% δεν μπορούμε να απορρίψουμε την υπόθεση $p = 1/2$.*

Όπως και στο προηγούμενο παράδειγμα, η απάντηση που δόθηκε εδώ είναι πλήρης από τη στατιστική σκοπιά, αλλά η μέθοδος με την οποία βρέθηκε είναι περιορισμένη στο συγκεκριμένο πρόβλημα. Στις επόμενες δύο παραγράφους θα αναπτύξουμε μια γενική μεθοδολογία για την επίλυση προβλημάτων ελέγχου υποθέσεων.

14.2.1 Έλεγχος παραμέτρου Bernoulli

Εδώ θα δώσουμε μια πιο συστηματική απάντηση στη γενική μορφή του ερωτήματος που είδαμε στο παραπάνω παράδειγμα. Ξεκινάμε από τα **δεδομένα** X_1, X_2, \dots, X_N , όπου το κάθε $X_i = 0$ ή 1 , και θεωρούμε πως προήλθαν από N ανεξάρτητες Τ.Μ. $X_i \sim \text{Bern}(p)$. Η κατανομή αυτή είναι το **μοντέλο** μας για τα δεδομένα και η $p \in (0, 1)$ είναι μια άγνωστη **παράμετρος**.

Θα εξετάσουμε δύο υποθέσεις. Για κάποιο συγκεκριμένο p^* :

- Μηδενική υπόθεση $H_0 : p = p^*$.
- Εναλλακτική υπόθεση $H_1 : p \in (0, 1)$.

Είναι δυνατόν, βάσει των παρατηρήσεων X_1, X_2, \dots, X_N , να απορρίψουμε τη μηδενική υπόθεση H_0 με κάποιο επιθυμητό επίπεδο «στατιστικής σημαντικότητας»; Για να απαντήσουμε με οργανωμένο τρόπο, θα ορίσουμε την εμπειρική κατανομή των δεδομένων και θα εξετάσουμε πόσο κοντά ή μακριά είναι από την $\text{Bern}(p^*)$ κατανομή. Υπολογίζουμε τον εμπειρικό μέσο \bar{X}_N και ορίζουμε την πυκνότητα \hat{P}_N με τιμές $\hat{P}_N(1) = \bar{X}_N$ και $\hat{P}_N(0) = 1 - \bar{X}_N$. Η εμπειρική κατανομή των δεδομένων είναι η $\text{Bern}(\bar{X}_N)$, με πυκνότητα \hat{P}_N .

Υπενθύμιση: Απόσταση χ^2 και κατανομή $\chi^2(k)$.

1. Στην Άσκηση 18 του Κεφαλαίου 6 ορίσαμε την χ^2 -απόσταση μεταξύ δύο πυκνοτήτων P και Q στο ίδιο πεπερασμένο σύνολο τιμών S ως,

$$d_{\chi^2}(P, Q) = \sum_{x \in S} \frac{(P(x) - Q(x))^2}{Q(x)},$$

και παρατηρήσαμε πως πάντοτε έχουμε $d_{\chi^2}(P, Q) \geq 0$, με $d_{\chi^2}(P, Q) = 0$ αν και μόνο αν οι δύο πυκνότητες είναι ίδιες, δηλαδή $P(x) = Q(x)$, για κάθε $x \in S$.

2. Στην Άσκηση 5 του Κεφαλαίου 12 ορίσαμε την κατανομή χ^2 με k βαθμούς ελευθερίας ($\chi^2(k)$ για συντομία), και είδαμε πως στην περίπτωση $k = 1$ η $\chi^2(1)$ έχει συνάρτηση κατανομής,

$$F(z) = \Phi(\sqrt{z}) - \Phi(-\sqrt{z}) = 2\Phi(\sqrt{z}) - 1, \quad \text{για } z \geq 0,$$

και $F(z) = 0$ για $z < 0$, όπου $\Phi(z)$ είναι η $N(0, 1)$ συνάρτηση κατανομής.

Προκειμένου να συγκρίνουμε την εμπειρική κατανομή \hat{P}_N με την κατανομή $P_0 \sim \text{Bern}(p^*)$ της μηδενικής υπόθεσης, ορίζουμε τη **στατιστική συνάρτηση** Δ_N ως N φορές τη χ^2 -απόστασή τους,

$$\Delta_N = N d_{\chi^2}(\hat{P}_N, P_0) = \frac{N(\hat{P}_N(1) - P_0(1))^2}{P_0(1)} + \frac{N(\hat{P}_N(0) - P_0(0))^2}{P_0(0)},$$

η οποία μπορεί να εκφραστεί ως μια απλή συνάρτηση του εμπειρικού μέσου \bar{X}_N :

$$\Delta_N = \frac{N(\bar{X}_N - p^*)^2}{p^*} + \frac{N(1 - \bar{X}_N - (1 - p^*))^2}{1 - p^*} = \frac{N(\bar{X}_N - p^*)^2}{p^*(1 - p^*)}. \quad (14.9)$$

Το παρακάτω θεώρημα (που θα αποδειχθεί στο τέλος αυτής της ενότητας) μας λέει πως, κάτω από τη μηδενική υπόθεση, η κατανομή του Δ_N είναι κατά προσέγγιση $\chi^2(1)$.

Θεώρημα 14.1 Κάτω από τη μηδενική υπόθεση H_0 , δηλαδή για δεδομένα X_1, X_2, \dots, X_N ανεξάρτητα και με κατανομή $X_i \sim \text{Bern}(p^*)$, η στατιστική συνάρτηση,

$$\Delta_N = Nd_{\chi^2}(\hat{P}_N, P_0) = \frac{N(\bar{X}_N - p^*)^2}{p^*(1 - p^*)},$$

συγκλίνει κατά κατανομή στη $\chi^2(1)$. Για κάθε $\delta \geq 0$:

$$\Pr(\Delta_N \leq \delta | H_0) \rightarrow 2\Phi(\sqrt{\delta}) - 1, \quad \text{καθώς } N \rightarrow \infty.$$

Είμαστε τώρα σε θέση να περιγράψουμε τη γενική μέθοδο του χ^2 -ελέγχου υπόθεσης.

χ^2 -έλεγχος παραμέτρου Bernoulli

1. Από τα δεδομένα X_1, X_2, \dots, X_N , υπολογίζουμε τον εμπειρικό μέσο \bar{X}_N και την τιμή δ της στατιστικής συνάρτησης, $\delta = N(\bar{X}_N - p^*)^2 / (p^*(1 - p^*))$.
2. Επιλέγουμε το επιθυμητό **επίπεδο σημαντικότητας** $\alpha \in (0, 1)$.
3. Υπολογίζουμε την **p -τιμή**, δηλαδή την πιθανότητα, κάτω από τη μηδενική υπόθεση H_0 , η Δ_N να είναι τόση ή μεγαλύτερη από την τιμή δ που υπολογίσαμε. Από το Θεώρημα 14.1:

$$p\text{-τιμή} = \Pr(\Delta_N \geq \delta | H_0) \approx 2\Phi(-\sqrt{\delta}).$$

4. Αν η p -τιμή $\leq \alpha$, η H_0 απορρίπτεται σε επίπεδο σημαντικότητας α .
Αν η p -τιμή $> \alpha$, η H_0 δεν απορρίπτεται σε επίπεδο σημαντικότητας α .

Παρατηρήσεις

1. Υπογραμμίζουμε και πάλι πως τα **μόνα** δύο δυνατά συμπεράσματα του ελέγχου υπόθεσης είναι, με κάποιο προαποφασισμένο επίπεδο σημαντικότητας α : Είτε να απορρίψουμε τη μηδενική υπόθεση H_0 (αν η απόσταση Δ_N της εμπειρικής κατανομής των δεδομένων από την κατανομή της μηδενικής υπόθεσης είναι αρκετά μεγάλη ώστε η p -τιμή να είναι αντίστοιχα μικρή), είτε να μην απορρίψουμε τη μηδενική υπόθεση H_0 . Σε **καμία** περίπτωση δεν καταλήγουμε στο συμπέρασμα ότι «αποδεχόμαστε» την H_0 .
2. Σημειώνουμε πως η μηδενική υπόθεση H_0 είναι ειδική περίπτωση της εναλλακτικής υπόθεσης H_1 . Αν και εκ πρώτης όψεως αυτό μπορεί να μοιάζει παράξενο, είναι πολύ συνηθισμένο και δεν δημιουργεί κανένα πρόβλημα στην ανάλυσή μας. Αφενός διότι σχεδόν όλοι μας οι υπολογισμοί γίνονται κάτω από την H_0 , και αφετέρου γιατί, όπως παρατηρούμε από την παραπάνω συζήτηση, το μόνο σημαντικό σημείο είναι να βεβαιωθούμε πως όλες οι ενδεχόμενες εμπειρικές κατανομές τις οποίες συγκρίνουμε με την P_0 ανήκουν πράγματι στην H_1 , το οποίο φυσικά συμβαίνει εδώ.

Παράδειγμα 14.8 Με το νόμισμα του Παραδείγματος 14.7 φέραμε 111 Κ σε $N = 210$ ρίψεις, οπότε ο εμπειρικός μέσος $\bar{X}_N = 111/210 \approx 0.5286$. Για να ελέγξουμε αν είναι δίκαιο, δηλαδή αν $p = p^* = 1/2$, επιλέγουμε το επίπεδο σημαντικότητας $\alpha = 5\%$, υπολογίζουμε τη στατιστική συνάρτηση $\delta \approx 0.686$, και για την p -τιμή από τους πίνακες του Κεφαλαίου 12 βρίσκουμε ότι ισούται με $2\Phi(-0.83) \approx 0.406$. Εφόσον η p -τιμή είναι (πολύ!) μεγαλύτερη του $\alpha = 0.05$, καταλήγουμε και πάλι στο συμπέρασμα ότι, βάσει των δεδομένων X_i , δεν μπορούμε να απορρίψουμε την μηδενική υπόθεση $p = 1/2$ σε επίπεδο σημαντικότητας 5%.

Παρομοίως, ο χρηματιστής του Παραδείγματος 14.5 προέβλεψε σωστά την πορεία της μετοχής σε 71 από τις $N = 120$ μέρες, οπότε $\bar{X}_N = 71/120 \approx 0.5917$. Για να ελέγξουμε τον ισχυρισμό του ότι προβλέπει τα Y_i με ακρίβεια σημαντικά μεγαλύτερη από 50%, με $p^* = 1/2$ και με επίπεδο σημαντικότητας $\alpha = 5\%$ υπολογίζουμε τη στατιστική συνάρτηση $\delta \approx 4.03$. Η αντίστοιχη p -τιμή εδώ είναι $2\Phi(-2.01) \approx 0.044$ που είναι (οριακά) μικρότερο από το $\alpha = 5\%$, οπότε συμπεραίνουμε πως η H_0 απορρίπτεται σε επίπεδο σημαντικότητας 5%, δηλαδή ότι ο χρηματιστής πράγματι απέδειξε τον ισχυρισμό του.

Ένα ακόμα σχετικό παράδειγμα όπου $p^* \neq 1/2$ θα δούμε στην Άσκηση 6 στο τέλος του κεφαλαίου.

Απόδειξη του Θεωρήματος 14.1: Κατ' αρχάς παρατηρούμε ότι κάτω από την υπόθεση H_0 η διασπορά των X_i είναι $\sigma^2 = p^*(1 - p^*)$, και ξεκινώντας από τη σχέση (14.9) υπολογίζουμε,

$$\Delta_N = \frac{N(\bar{X}_N - p^*)^2}{p^*(1 - p^*)} = \left[\frac{\sqrt{N}}{\sigma} (\bar{X}_N - p^*) \right]^2 = \bar{S}_N^2,$$

όπου \bar{S}_N είναι το κανονικοποιημένο άθροισμα των X_i όπως στη σχέση (12.3). Άρα, για $\delta \geq 0$,

$$\Pr(\Delta_n \leq \delta | H_0) = \Pr(\bar{S}_N^2 \leq \delta | H_0) = \Pr(-\sqrt{\delta} \leq \bar{S}_N \leq \sqrt{\delta} | H_0),$$

και από το Κ.Ο.Θ. έχουμε,

$$\Pr(\Delta_n \leq \delta | H_0) = \Pr(\bar{S}_N \leq \sqrt{\delta}) - \Pr(\bar{S}_N < -\sqrt{\delta}) \rightarrow \Phi(\sqrt{\delta}) - \Phi(-\sqrt{\delta}) = 2\Phi(\sqrt{\delta}) - 1,$$

καθώς το $N \rightarrow \infty$. ■

14.2.2 Έλεγχος ανεξαρτησίας

Έστω πώς έχουμε μια ακολουθία από δεδομένα της μορφής,

$$(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N),$$

όπου τα X_i και τα Y_i παίρνουν τις τιμές 0 και 1. Υποθέτοντας ότι, βάσει της κατανομής τους, τα ζεύγη (X_i, Y_i) είναι ανεξάρτητα μεταξύ τους για διαφορετικά $i = 1, 2, \dots, N$, ο στόχος μας είναι να ελέγξουμε αν και τα X_i είναι ανεξάρτητα από τα Y_i .

Υιοθετούμε το εξής γενικό **μοντέλο**. Θεωρούμε ότι τα Y_i προέρχονται από ανεξάρτητες $\text{Bern}(q)$ Γ.Μ. για κάποιο ενδεχομένως άγνωστο q και πως, δεδομένης της τιμής του κάθε Y_i , το αντίστοιχο $X_i \sim \text{Bern}(p_0)$ αν $Y_i = 0$, ενώ $X_i \sim \text{Bern}(p_1)$ αν $Y_i = 1$. Με αυτόν το συμβολισμό, η υπόθεση της ανεξαρτησίας των X_i από τα Y_i αντιστοιχεί στην περίπτωση όπου $p_0 = p_1$, δηλαδή η κατανομή των X_i είναι ανεξάρτητη από τις τιμές των Y_i . Οι p_0, p_1 και q είναι οι τρεις άγνωστες **παράμετροι** του μοντέλου.

Π.χ., στην περίπτωση της κλινικής μελέτης του Παραδείγματος 14.6 μπορούμε να θέσουμε $Y_i = 1$ αν στον ασθενή i χορηγείται το κανονικό φάρμακο και $Y_i = 0$ αν του χορηγείται το placebo, όπου υποθέτουμε πως η επιλογή εκείνων που θα πάρουν το φάρμακο γίνεται τυχαία, με κάποια πιθανότητα q . Επιπλέον, θέτοντας,

$$\begin{aligned} p_0 &= \text{Pr}(\text{κάποιος ασθενής θα σημειώσει βελτίωση} \mid \text{πήρε το placebo}), \\ p_1 &= \text{Pr}(\text{κάποιος ασθενής θα σημειώσει βελτίωση} \mid \text{πήρε το φάρμακο}), \end{aligned}$$

το ζητούμενο είναι να ελέγξουμε αν η πορεία της υγείας του ασθενή είναι ανεξάρτητη ή όχι από το αν πήρε το φάρμακο, δηλαδή αν $p_0 = p_1$.

Επιγραμματικά, οι δύο υποθέσεις που εξετάζουμε είναι:

- **Μηδενική υπόθεση (ανεξαρτησία)** $H_0: p_0 = p_1$.
- **Εναλλακτική υπόθεση** $H_1: p_0, p_1$ αυθαίρετα.

Όπως και στην προηγούμενη ενότητα, το βασικό ερώτημα είναι αν είναι δυνατόν, βάσει των παρατηρήσεων $(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)$, να απορρίψουμε τη μηδενική υπόθεση H_0 σε κάποιο επιθυμητό επίπεδο στατιστικής σημαντικότητας. Προκειμένου να απαντήσουμε, θα συγκρίνουμε την εμπειρική κατανομή των δεδομένων με την κατανομή που θα είχαν αν τα X_i και τα Y_i ήταν ανεξάρτητα. Συγκεκριμένα, γράφοντας \bar{X}_N, \bar{Y}_N για τους εμπειρικούς μέσους των X_i και των Y_i , ορίζουμε τις πυκνότητες $\hat{P}_{X,N}, \hat{P}_{Y,N}$ ως αυτές των κατανομών $\text{Bern}(\bar{X}_N)$ και $\text{Bern}(\bar{Y}_N)$ αντίστοιχα. Επιπλέον, ορίζουμε την από κοινού εμπειρική κατανομή,

$$\hat{P}_{XY,N}(x, y) = \frac{1}{N} [\text{πλήθος από ζεύγη } (x, y) \text{ μεταξύ των } (X_i, Y_i)],$$

για κάθε $x = 0, 1$ και $y = 0, 1$.

Τέλος, ορίζουμε τη **στατιστική συνάρτηση** Δ_N ως N φορές την χ^2 -απόσταση μεταξύ της από κοινού εμπειρικής πυκνότητας $\hat{P}_{XY,N}(x, y)$ και της από κοινού πυκνότητας $\hat{P}_{X,N}(x)\hat{P}_{Y,N}(y)$ που θα είχαν τα (X_i, Y_i) αν ήταν ανεξάρτητα:

$$\begin{aligned} \Delta_N &= Nd_{\chi^2}(\hat{P}_{XY,N}, \hat{P}_{X,N}\hat{P}_{Y,N}) \\ &= N \sum_{x=0,1, y=0,1} \frac{\left(\hat{P}_{XY,N}(x, y) - \hat{P}_{X,N}(x)\hat{P}_{Y,N}(y)\right)^2}{\hat{P}_{X,N}(x)\hat{P}_{Y,N}(y)}. \end{aligned}$$

Το ακόλουθο θεώρημα περιγράφει την κατανομή της Δ_N κάτω από τη μηδενική υπόθεση H_0 , για μεγάλα N . Η απόδειξή του βασίζεται σε έναν πανομοιότυπο (αλλά πολύ πιο μακροσκελή) υπολογισμό με εκείνον που είδαμε στην απόδειξη του Θεωρήματος 14.1, και δεν παρουσιάζει ιδιαίτερο ενδιαφέρον: γι' αυτόν το λόγο παραλείπεται.

Θεώρημα 14.2 *Κάτω από τη μηδενική υπόθεση H_0 , δηλαδή για ανεξάρτητα ζεύγη δεδομένων $(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)$ τέτοια ώστε τα X_i να είναι ανεξάρτητα από τα Y_i , η στατιστική συνάρτηση,*

$$\Delta_N = Nd_{\chi^2}(\hat{P}_{XY,N}, \hat{P}_{X,N}\hat{P}_{Y,N}),$$

συγκλίνει κατά κατανομή στη $\chi^2(1)$. Για κάθε $\delta \geq 0$:

$$\Pr(\Delta_N \leq \delta | H_0) \rightarrow 2\Phi(\sqrt{\delta}) - 1, \quad \text{καθώς } N \rightarrow \infty.$$

Η μεθοδολογία του χ^2 -ελέγχου ανεξαρτησίας βασίζεται στο αποτέλεσμα του Θεωρήματος 14.2 με τον ίδιο τρόπο που ο χ^2 -έλεγχος παραμέτρου Bernoulli προέκυψε από το Θεώρημα 14.1. Η βασική ιδέα είναι να υπολογιστεί η απόσταση της εμπειρικής κατανομής των δεδομένων από την κατανομή που θα είχαν αν τα X_i, Y_i ήταν ανεξάρτητα, και αν βρεθεί να είναι αρκετά μεγάλη ώστε η σχετική p -τιμή να είναι αντίστοιχα μικρή, τότε να απορρίψουμε τη μηδενική υπόθεση H_0 .

χ^2 -έλεγχος ανεξαρτησίας

1. Από τα δεδομένα $(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)$, υπολογίζουμε τις εμπειρικές πυκνότητες $\hat{P}_{X,N}, \hat{P}_{Y,N}$ και $\hat{P}_{XY,N}$, και την τιμή δ της **στατιστικής συνάρτησης**,

$$\delta = Nd_{\chi^2}(\hat{P}_{XY,N}, \hat{P}_{X,N}\hat{P}_{Y,N}).$$

2. Επιλέγουμε το επιθυμητό **επίπεδο σημαντικότητας** $\alpha \in (0, 1)$.
3. Υπολογίζουμε την **p -τιμή**, δηλαδή την πιθανότητα, κάτω από τη μηδενική υπόθεση H_0 , η Δ_N να είναι τόσο ή μεγαλύτερη από την τιμή δ που υπολογίσαμε. Από το Θεώρημα 14.2:

$$p\text{-τιμή} = \Pr(\Delta_N \geq \delta | H_0) \approx 2\Phi(-\sqrt{\delta}).$$

4. Αν η p -τιμή $\leq \alpha$, η H_0 απορρίπτεται σε επίπεδο σημαντικότητας α .
Αν η p -τιμή $> \alpha$, η H_0 δεν απορρίπτεται σε επίπεδο σημαντικότητας α .

Παρατήρηση: Όπως και στην περίπτωση του ελέγχου παραμέτρου Bernoulli στο Θεώρημα 14.2, η ασυμπτωτική κατανομή της στατιστικής συνάρτησης Δ_N είναι η $\chi^2(1)$. Ένα παράδειγμα ελέγχου υπόθεσης στο οποίο η ασυμπτωτική κατανομή είναι διαφορετική παρουσιάζεται στην Άσκηση 8 στο τέλος του κεφαλαίου.

Παράδειγμα 14.9 (Κλινική μελέτη με πολλούς ασθενείς) Στο Παράδειγμα 14.6, από τα κλινικά αποτελέσματα σε $N = 40$ ασθενείς καταλήξαμε στο συμπέρασμα πως δεν μπορεί να απορριφθεί η υπόθεση ότι η λήψη του φαρμάκου είναι ανεξάρτητη από την εξέλιξη της ασθένειας. Εδώ εξετάζουμε ποιο θα ήταν το αντίστοιχο συμπέρασμα για μια κλινική μελέτη με τα ίδια ποσοστά αποτελεσμάτων αλλά με σημαντικά μεγαλύτερο μέγεθος δείγματος N .

Έστω πως, από $N = 380$ ασθενείς, επιλέχθηκαν τυχαία (με $q = 1/2$) 185 ασθενείς στους οποίους χορηγήθηκε το φάρμακο, και στους υπόλοιπους 195 χορηγήθηκε placebo. Παρατηρούμε ότι το 60% από τους πρώτους, δηλαδή 111, σημείωσαν βελτίωση, ενώ το αντίστοιχο ποσοστό για τους δεύτερους ήταν μόνο 40%, δηλαδή 78. Από αυτά τα στοιχεία εύκολα υπολογίζουμε τις σχετικές εμπειρικές πυκνότητες. Έχουμε ότι η $\hat{P}_{Y,N}$ είναι η πυκνότητα της $\text{Bern}(185/380)$ κατανομής και η $\hat{P}_{X,N}$ είναι η πυκνότητα της κατανομής Bernoulli με παράμετρο $(111 + 78)/380$. Παρομοίως, για τις τιμές της από κοινού εμπειρικής πυκνότητας βρίσκουμε:

$$\hat{P}_{XY,N}(1, 1) = \frac{111}{380}, \quad \hat{P}_{XY,N}(0, 1) = \frac{74}{380}, \quad \hat{P}_{XY,N}(1, 0) = \frac{78}{380} \quad \text{και} \quad \hat{P}_{XY,N}(0, 0) = \frac{117}{380}.$$

Άρα η τιμή της στατιστικής συνάρτησης είναι,

$$\delta = 380 \times \left[\frac{\left(\frac{111}{380} - \frac{189}{380} \cdot \frac{185}{380} \right)^2}{\frac{189}{380} \cdot \frac{185}{380}} + \frac{\left(\frac{74}{380} - \frac{191}{380} \cdot \frac{185}{380} \right)^2}{\frac{191}{380} \cdot \frac{185}{380}} + \frac{\left(\frac{78}{380} - \frac{189}{380} \cdot \frac{195}{380} \right)^2}{\frac{189}{380} \cdot \frac{195}{380}} + \frac{\left(\frac{117}{380} - \frac{191}{380} \cdot \frac{195}{380} \right)^2}{\frac{191}{380} \cdot \frac{195}{380}} \right],$$

το οποίο μας δίνει $\delta \approx 11.512$.

Εφόσον το δείγμα είναι σχετικά μεγάλο, επιλέγουμε για το επίπεδο στατιστικής σημαντικότητας $\alpha = 1\%$, και υπολογίζουμε την p -τιμή $2\Phi(-3.39) = 0.0006$, η οποία είναι προφανώς πολύ μικρότερη του $\alpha = 0.01$. Άρα, σε αντίθεση με το Παράδειγμα 14.6, εδώ συμπεραίνουμε πως η μηδενική υπόθεση – ότι η δράση του φαρμάκου είναι ανεξάρτητη από την πορεία της ασθένειας – απορρίπτεται με επίπεδο σημαντικότητας 1%.

Παρατηρήσεις

1. Στο παραπάνω παράδειγμα ίσως να μπαίναμε στον πειρασμό να πούμε ότι, αφού η p -τιμή είναι μόλις 0.06%, μπορούμε να απορρίψουμε την μηδενική υπόθεση με πολύ «υψηλότερο» επίπεδο σημαντικότητας, π.χ., $\alpha = 0.1\%$. Αυτός ο συλλογισμός αγγίζει ένα λεπτό θέμα στατιστικής μεθοδολογίας στο οποίο δεν θα επεκταθούμε εδώ, πέραν του να υπογραμμίσουμε ότι η ορθή διαδικασία είναι πάντοτε να επιλέγεται πρώτα το επιθυμητό επίπεδο στατιστικής σημαντικότητας α και μετά να συγκρίνεται με την p -τιμή.
2. Από το τελευταίο παραπάνω παράδειγμα και από την περιγραφή του μοντέλου στην αρχή αυτής της ενότητας ίσως να δίνεται η εντύπωση ότι περιοριζόμαστε σε περιπτώσεις όπου οι τιμές των Y_i είναι επιλεγμένες τυχαία βάσει κάποιου αρχικού σχεδιασμού. Αυτό δεν ισχύει: στην πραγματικότητα οι υποθέσεις μας απλά απαιτούν τα Y_i να είναι ανεξάρτητες Bernoulli T.M. όλες με την ίδια παράμετρο. Ένα τέτοιο παράδειγμα όπου οι τιμές των Y_i είναι εντελώς εκτός του ελέγχου μας είναι το ακόλουθο.

Παράδειγμα 14.10 (Μορφωτικό επίπεδο και εισόδημα) Ένα πολυσυζητημένο θέμα είναι ο βαθμός στον οποίο υπάρχει (ή δεν υπάρχει) στατιστική συσχέτιση μεταξύ του μορφωτικού επιπέδου και του εισοδήματος. Μια από τις μελέτες του αμερικανικού Bureau of Labor Statistics αναφέρει τα εξής αποτελέσματα. Από ένα τυχαίο δείγμα $N = 419$ ατόμων κάποιου πληθυσμού, καταγράφονται, για κάθε άτομο $i = 1, 2, \dots, N$ τα εξής: $X_i = 1$ αν το ετήσιο εισόδημά του είναι υψηλό (δηλαδή μεγαλύτερο από 45,000\$), $X_i = 0$ αν όχι, και $Y_i = 1$ αν το μορφωτικό του επίπεδο είναι υψηλό (δηλαδή αν είναι κάτοχος πτυχίου πανεπιστημίου), $Y_i = 0$ αν όχι.

Παρατηρούμε ότι από τους 304 που έχουν υψηλό μορφωτικό επίπεδο οι 187 έχουν υψηλό εισόδημα, ενώ από τους υπόλοιπους 115 οι 43 έχουν υψηλό εισόδημα. Μπορούμε, βάσει αυτών των παρατηρήσεων, να αποδείξουμε ότι υπάρχει «στατιστικά σημαντική» συσχέτιση μεταξύ του μορφωτικού επιπέδου και του εισοδήματος; Με άλλα λόγια, μπορούμε από αυτά τα δεδομένα να απορρίψουμε τη μηδενική υπόθεση (της ανεξαρτησίας εισοδήματος-μορφωτικού επιπέδου) με επίπεδο σημαντικότητας, ας πούμε, $\alpha = 5\%$;

Όπως στο προηγούμενο παράδειγμα, από τα δεδομένα υπολογίζουμε τις εμπειρικές κατανομές,

$$P_{X,N}(1) = 1 - P_{X,N}(0) = \frac{187 + 43}{419} = \frac{230}{419},$$

$$P_{Y,N}(1) = 1 - P_{Y,N}(0) = \frac{304}{419},$$

$$P_{XY,N}(1,1) = \frac{187}{419}, \quad P_{XY,N}(0,1) = \frac{117}{419}, \quad P_{XY,N}(1,0) = \frac{43}{419}, \quad P_{XY,N}(0,0) = \frac{72}{419},$$

από τις οποίες βρίσκουμε την τιμή της στατιστικής συνάρτησης,

$$\delta = 419 \times \left[\frac{\left(\frac{187}{419} - \frac{230}{419} \cdot \frac{304}{419}\right)^2}{\frac{230}{419} \cdot \frac{304}{419}} + \frac{\left(\frac{117}{419} - \frac{189}{419} \cdot \frac{304}{419}\right)^2}{\frac{189}{419} \cdot \frac{304}{419}} + \frac{\left(\frac{43}{419} - \frac{230}{419} \cdot \frac{115}{419}\right)^2}{\frac{230}{419} \cdot \frac{115}{419}} + \frac{\left(\frac{72}{419} - \frac{189}{419} \cdot \frac{115}{419}\right)^2}{\frac{189}{419} \cdot \frac{115}{419}} \right],$$

δηλαδή $\delta \approx 19.607$. Άρα η αντίστοιχη p -τιμή είναι $2\Phi(-4.428) < 0.0001$, και συμπεραίνουμε ότι μπορούμε να απορρίψουμε (με μάλλον «πανηγυρικό» τρόπο) τη μηδενική υπόθεση της ανεξαρτησίας μεταξύ του εισοδήματος και του μορφωτικού επιπέδου, με επίπεδο εμπιστοσύνης 5%.

14.3 Μείωση διασποράς

Παράδειγμα 14.11 Έστω πως θέλουμε να υπολογίσουμε το ολοκλήρωμα της συνάρτησης $G(x) = e^{|x+40|^{1/4}}$ στο διάστημα $[-50, 50]$,

$$\int_{-50}^{50} G(x) dx = \int_{-50}^{50} e^{|x+40|^{1/4}} dx,$$

το οποίο υποθέτουμε (σωστά) πως δεν μπορεί να εκφραστεί σε κλειστή μορφή. Παρατηρούμε όμως πως μπορεί να εκφραστεί σαν μια μέση τιμή ως προς την πυκνότητα της $U[-50, 50]$ κατανομής,

$$100 \times \int_{-50}^{50} e^{|x+40|^{1/4}} \frac{1}{100} dx = 100E \left[e^{|X+40|^{1/4}} \right],$$

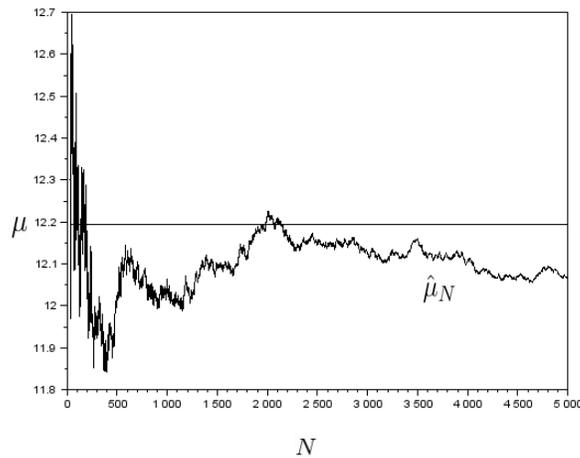
όπου $X \sim U[-50, 50]$. Άρα για να εκτιμήσουμε το αρχικό ολοκλήρωμα αρκεί να εκτιμήσουμε τη μέση τιμή,

$$\mu = E[G(X)] = E \left[e^{|X+40|^{1/4}} \right].$$

Γι' αυτόν το σκοπό παίρνουμε μέσω του υπολογιστή ένα μεγάλο πλήθος N ανεξάρτητων δειγμάτων $X_i \sim U[-50, 50]$ και υπολογίζουμε τον εμπειρικό μέσο όρο,

$$\hat{\mu}_N = \frac{1}{N} \sum_{i=1}^N G(X_i).$$

Ο Ν.Μ.Α. μάς διαβεβαιώνει ότι θα έχουμε $\hat{\mu}_N \rightarrow \mu$ κατά πιθανότητα, καθώς το $N \rightarrow \infty$, αλλά με $N = 5000$ δείγματα παρατηρούμε πως η εκτίμηση $\hat{\mu}_N$ φαίνεται να μην έχει ακόμα συγκλίνει στη ζητούμενη τιμή μ . Στο Σχήμα 14.2 βλέπουμε το αποτέλεσμα ενός τέτοιου πειράματος.



Σχήμα 14.2: Παράδειγμα 14.11: Γραφική αναπαράσταση των τιμών της εκτιμήτριας $\hat{\mu}_N$ ως συνάρτηση του πλήθους N των προσομοιωμένων δειγμάτων. Η οριζόντια γραμμή είναι η πραγματική τιμή του ζητούμενου μέσου $\mu = E[G(X)]$.

Εφόσον γνωρίζουμε πως οι Τ.Μ. X_i έχουν μέση τιμή $E(X_i) = 0$, μια ενδιαφέρουσα παρατήρηση είναι πως, για οποιαδήποτε σταθερά $\theta \in \mathbb{R}$, οι τυχαίες μεταβλητές $[G(X_i) - \theta X_i]$ έχουν κι αυτές μέση τιμή ίση με μ ,

$$E[G(X) - \theta X] = E[G(X)] - \theta E(X) = \mu,$$

άρα για κάθε θ μπορούμε να ορίσουμε μια νέα εκτιμήτρια:

$$\mu_N(\theta) = \frac{1}{N} \sum_{i=1}^N [G(X_i) - \theta X_i].$$

Και πάλι ο Ν.Μ.Α. εγγυάται ότι $\mu_N(\theta) \rightarrow \mu$ κατά πιθανότητα, καθώς το $N \rightarrow \infty$, και εύλογα διερωτόμαστε αν, με μια κατάλληλη τιμή για το θ , η $\mu_N(\theta)$ θα μπορούσε να συγκλίνει στο μ πιο «γρήγορα». Στο Σχήμα 14.3 βλέπουμε πειραματικά αποτελέσματα στα οποία συγκρίνεται η συμπεριφορά των δύο εκτιμητριών $\hat{\mu}_N$ και $\mu_N(\theta)$ για διαφορετικές τιμές του θ . Είναι εμφανές ότι για $\theta = 0.25$ η $\mu_N(\theta)$ μοιάζει να συγκλίνει ταχύτερα και να έχει μικρότερη διασπορά από τη $\hat{\mu}_N$, γεγονός το οποίο επιβεβαιώνουμε εμπειρικά και με δύο ακόμα προσομοιώσεις στο Σχήμα 14.4.

Μια και το παραπάνω σενάριο εμφανίζεται αρκετά συχνά σε προβλήματα εκτίμησης, ιδιαίτερα σε περιπτώσεις προσομοιωμένων δεδομένων, παρακάτω περιγράφουμε τη γενική του μορφή και εξετάζουμε με ποιον τρόπο μπορεί να γίνει συστηματικά η επιλογή μιας αποτελεσματικής τιμής για τη σταθερά θ .

Μέθοδος μεταβλητών ελέγχου. Έστω πως έχουμε ανεξάρτητα δείγματα X_1, \dots, X_N από μια ενδεχομένως άγνωστη κατανομή, και θέλουμε να εκτιμήσουμε τη μέση τιμή $\mu = E[G(X)]$ μιας συνάρτησης $G(x)$, όπου η X έχει την ίδια κατανομή με τα X_i . Αν υποθέσουμε πως γνωρίζουμε τη μέση τιμή $\nu = E[H(X)]$ μιας διαφορετικής συνάρτησης $H(X)$, εκτός από τον εμπειρικό μέσο,

$$\hat{\mu}_N = \frac{1}{N} \sum_{i=1}^N G(X_i),$$

μπορούμε να ορίσουμε, για οποιοδήποτε $\theta \in \mathbb{R}$, και τη νέα εκτιμήτρια,

$$\mu_N(\theta) = \frac{1}{N} \sum_{i=1}^N (G(X_i) - \theta[H(X_i) - \nu]).$$

Από τον Ν.Μ.Α. έχουμε πως και οι δύο εκτιμήτριες συγκλίνουν κατά πιθανότητα στο μ καθώς $N \rightarrow \infty$, μια και,

$$E(G(X) - \theta[H(X) - \nu]) = E[G(X)] - \theta E(H(X) - \nu) = \mu - \theta(E[H(X)] - \nu) = \mu.$$

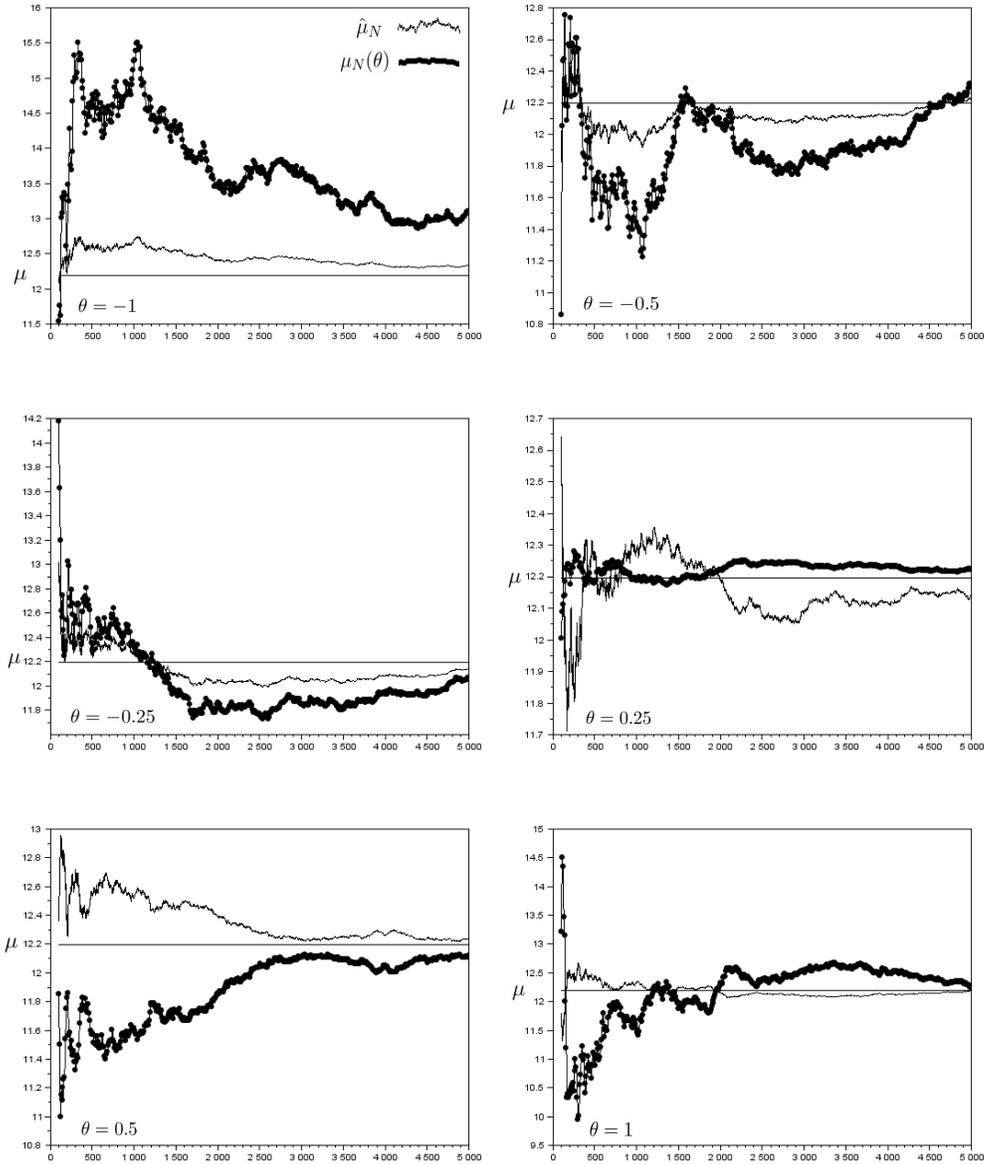
Το παρακάτω αποτέλεσμα μας λέει πως, αν οι Τ.Μ. $G(X)$ και $H(X)$ δεν έχουν μηδενική συνδιακύμανση $\text{Cov}(G(X), H(X))$, τότε πάντοτε υπάρχει κάποιο θ^* για το οποίο η $\mu_N(\theta^*)$ είναι πιο αποτελεσματική από τον κλασικό εμπειρικό μέσο $\hat{\mu}_N$.

Θεώρημα 14.3 Έστω πως οι τυχαίες μεταβλητές $G(X)$ και $H(X)$ έχουν μη μηδενική συνδιακύμανση $\text{Cov}(G(X), H(X))$, και η $H(X)$ έχει μη μηδενική διασπορά. Τότε, για κάθε N , η διασπορά της εκτιμήτριας $\mu_N(\theta)$ ελαχιστοποιείται από την τιμή,

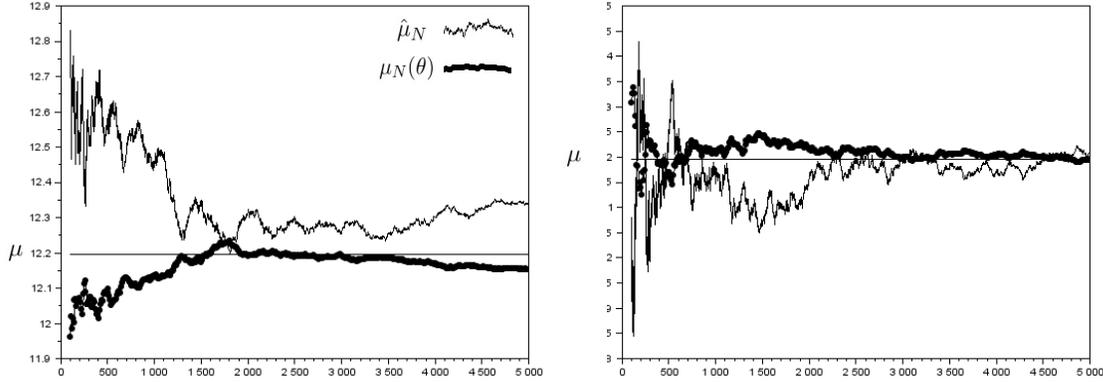
$$\theta^* = \frac{\text{Cov}(G(X), H(X))}{\text{Var}(H(X))}, \quad (14.10)$$

οπότε και είναι αυστηρά αποτελεσματικότερη από τον εμπειρικό μέσο:

$$\text{Var}(\mu_N(\theta^*)) < \text{Var}(\hat{\mu}_N). \quad (14.11)$$



Σχήμα 14.3: Παράδειγμα 14.11: Για διαφορετικές τιμές του $\theta = -1, -1/2, -1/4, 1/4, 1/2, 1$, οι τιμές των δύο εκτιμητριών $\hat{\mu}_n$ (απλή γραμμή) και $\mu_N(\theta)$ (έντονη γραμμή) αναπαρίστανται ως συναρτήσεις του πλήθους N των δειγμάτων X_i . Η οριζόντια γραμμή είναι η πραγματική τιμή του μέσου μ .



Σχήμα 14.4: Παράδειγμα 14.11: Αποτελέσματα δύο ακόμα προσομοιώσεων με $\theta = 1/4$. Η νέα εκτιμήτρια $\mu_N(\theta)$ (έντονη γραμμή) φαίνεται σαφώς να συγκλίνει ταχύτερα από τον εμπειρικό μέσο $\hat{\mu}_n$ (απλή γραμμή). Η οριζόντια γραμμή είναι η πραγματική τιμή του μέσου μ .

Παρατήρηση: Αν και ως τώρα έχουμε ορίσει τη συνδιακύμανση μόνο για διακριτές Τ.Μ., όπως θα δούμε στο επόμενο κεφάλαιο ο ορισμός παραμένει ακριβώς ίδιος και στην περίπτωση Τ.Μ. με συνεχή κατανομή: $\text{Cov}(X, Y) = E[(X - E(X))(Y - E(Y))]$, για οποιεσδήποτε X, Y .

Απόδειξη: Όπως έχουμε παρατηρήσει αρκετές φορές από την αρχή του κεφαλαίου (βλ., για παράδειγμα, (14.1)), για τη διασπορά ενός εμπειρικού μέσου όπως η $\mu_N(\theta)$ έχουμε,

$$N\text{Var}(\mu_N(\theta)) = \text{Var}(G(X) - \theta[H(X) - \nu]).$$

Άρα, για να ελαχιστοποιήσουμε τη $\text{Var}(\mu_N(\theta))$ αρκεί να βρούμε το θ που ελαχιστοποιεί την παραπάνω διασπορά, η οποία μπορεί να αναπτυχθεί ως,

$$\begin{aligned} & \text{Var}(G(X) - \theta[H(X) - \nu]) \\ &= E\left[\left(G(X) - \theta[H(X) - \nu] - E[G(X) - \theta[H(X) - \nu]]\right)^2\right] \\ &= E\left[\left(G(X) - \mu - \theta H(X) + \theta\nu + \theta\nu - \theta\nu\right)^2\right] \\ &= E\left[\left((G(X) - \mu) - \theta(H(X) - \nu)\right)^2\right] \\ &= E\left[\left(G(X) - \mu\right)^2\right] - 2\theta E\left[(G(X) - \mu)(H(X) - \nu)\right] + \theta^2 E\left[\left(H(X) - \nu\right)^2\right] \\ &= \text{Var}(H(X))\theta^2 - 2\text{Cov}(G(X), H(X))\theta + \text{Var}(G(X)), \end{aligned} \quad (14.12)$$

όπου χρησιμοποιήσαμε μόνο τους ορισμούς και τις βασικές ιδιότητες της μέσης τιμής από τα Θεωρήματα 6.1 και 11.5. Συνεπώς η $\text{Var}(\mu_N(\theta))$ ελαχιστοποιείται για το θ εκείνο που ελαχιστοποιεί το απλό τριώνυμο στη (14.12), δηλαδή ακριβώς για την τιμή θ^* της (14.10).

Τέλος, το ότι έχουμε αυστηρή ανισότητα στη σχέση (14.11) προκύπτει από το γεγονός ότι κάτω από τις παρούσες συνθήκες το $\theta^* \neq 0$, άρα,

$$\text{Var}(\mu_N(\theta^*)) < \text{Var}(\mu_N(0)) = \text{Var}(\hat{\mu}_N),$$

αφού εξ ορισμού $\mu_N(0) = \hat{\mu}_N$. ■

Το παραπάνω θεώρημα μας διαβεβαιώνει πως, κάτω από γενικές συνθήκες, η χρήση μιας «μεταβλητής ελέγχου» $H(X)$ με γνωστή μέση τιμή είναι αποτελεσματική για την ακριβέστερη εκτίμηση του $\mu = E[G(X)]$. Παρατηρούμε όμως πως το αποτέλεσμα του θεωρήματος δεν μπορεί να εφαρμοστεί απευθείας στην πράξη, διότι το θ^* όπως στη (14.10) είναι αδύνατον να υπολογιστεί αναλυτικά – ακόμα και η απλή τιμή του μέσου $E[G(X)]$ μάς είναι άγνωστη. Η συνήθης πρακτική είναι να πρώτα να υπολογίσουμε μια εκτίμηση $\hat{\theta}$ για τη βέλτιστη τιμή θ^* και μετά να χρησιμοποιήσουμε την εκτιμήτρια $\mu_N(\hat{\theta})$. Συγκεκριμένα, εφόσον,

$$\theta^* = \frac{\text{Cov}(G(X), H(X))}{\text{Var}(H(X))} = \frac{E[(G(X) - \mu)(H(X) - \nu)]}{E[(H(X) - \nu)^2]}, \quad (14.13)$$

παίρνοντας τον εμπειρικό μέσο $\hat{\mu}_N$ ως μια πρώτη εκτίμηση για το μ , εκτιμούμε την παραπάνω συνδιακύμανση ως,

$$\hat{c} = \frac{1}{N} \sum_{i=1}^N (G(X_i) - \hat{\mu})(H(X_i) - \nu), \quad (14.14)$$

και τη διασπορά της $H(X)$ ως,

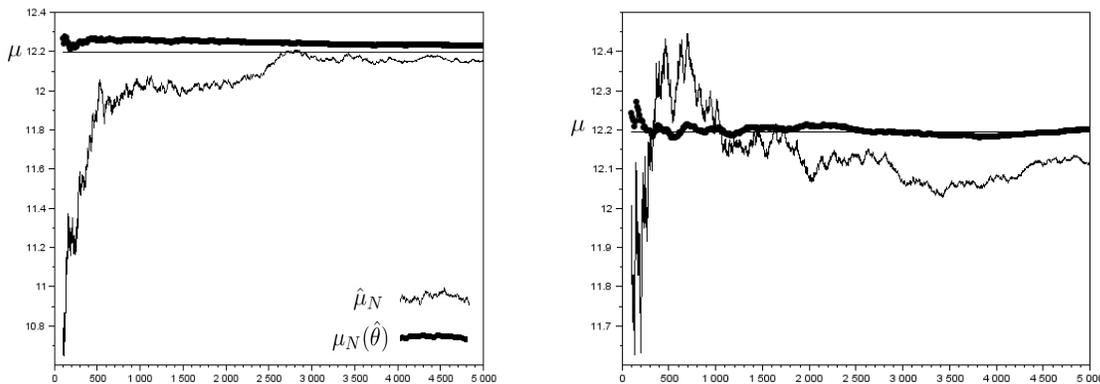
$$\hat{\tau}^2 = \frac{1}{N} \sum_{i=1}^N (H(X_i) - \nu)^2. \quad (14.15)$$

Άρα, βάσει της (14.13), μια λογική εκτίμηση για τη θ^* είναι η $\hat{\theta} = \hat{c}/\hat{\tau}^2$. Χρησιμοποιώντας αυτό το $\hat{\theta}$, ορίζουμε την **προσαρμοστική εκτιμήτρια με μεταβλητή ελέγχου** $H(X)$:

$$\mu_N(\hat{\theta}) = \frac{1}{N} \sum_{i=1}^N (G(X_i) - \hat{\theta}[H(X_i) - \nu]). \quad (14.16)$$

Παράδειγμα 14.12 Επιστρέφουμε στο Παράδειγμα 14.11, όπου το ζητούμενο είναι να υπολογιστεί η $\mu = E[G(X)]$ για τη συνάρτηση $G(x) = e^{|x+40|^{1/4}}$ μιας Τ.Μ. $X \sim U[-50, 50]$, και όπου έχουμε τη μεταβλητή ελέγχου $H(X) = X$ για την οποία $\nu = E[H(X)] = E(X) = 0$.

Στο Σχήμα 14.5 δίνονται τα αποτελέσματα δύο προσομοιώσεων όπου συγκρίνεται η προσαρμοστική εκτιμήτρια $\mu_N(\hat{\theta})$ με τον κλασικό εμπειρικό μέσο $\hat{\mu}_N$. Για μεγάλο πλήθος δειγμάτων η εκτίμηση $\hat{\theta}$ συγκλίνει στη βέλτιστη τιμή θ^* και βρίσκουμε ότι $\theta^* \approx 0.194$, το οποίο είναι αρκετά κοντά στην τιμή $\theta = 0.25$, την οποία είχαμε νωρίτερα επιλέξει πειραματικά.



Σχήμα 14.5: Παράδειγμα 14.12: Οι τιμές των εκτιμητριών $\mu_N(\hat{\theta})$ (έντονη γραμμή) και $\hat{\mu}_N$ (απλή γραμμή) ως συναρτήσεις του πλήθους N των προσομοιωμένων δειγμάτων, σε δύο διαφορετικά πειράματα προσομοίωσης. Είναι εμφανές ότι η προσαρμοστική εκτιμήτρια $\mu_N(\hat{\theta})$ είναι πολύ αποτελεσματικότερη από τον εμπειρικό μέσο $\hat{\mu}_N$. Η οριζόντια γραμμή είναι η πραγματική τιμή $\mu = E[G(X)]$.

Κλείνουμε με ένα πιο ρεαλιστικό παράδειγμα από τα χρηματοοικονομικά, μία από τις πιο συνηθισμένες περιοχές όπου χρησιμοποιείται ευρέως η μέθοδος των μεταβλητών ελέγχου.

Παράδειγμα 14.13 (Τιμολόγηση δικαιώματος αγοράς) Ένα από τα πιο κοινά χρηματιστηριακά παράγωγα είναι το λεγόμενο «δικαίωμα αγοράς». Για μια συγκεκριμένη μετοχή, μπορούμε να αγοράσουμε το δικαίωμα να αγοράσουμε τη μετοχή κάποια μελλοντική στιγμή T στην τιμή άσκησης K , όπου τα T και K είναι προσυμφωνημένα. Συμβολίζοντας με $X(t)$ την τιμή της μετοχής την κάθε χρονική στιγμή t , αν η $X(T)$ είναι μεγαλύτερη από K θα αγοράσουμε τη μετοχή στη φθηνότερη τιμή K κερδίζοντας $X(T) - K$ ευρώ, αλλιώς δεν θα εξασκήσουμε το δικαίωμά μας. Συνεπώς, η απόδοση αυτού του δικαιώματος είναι $\max\{X(T) - K, 0\}$ ευρώ, και σε μια αγορά με σταθερό επιτόκιο r η προεξοφλημένη απόδοσή του ισούται με,

$$G(X(T)) = e^{-rT} \max\{X(T) - K, 0\} \quad \text{ευρώ.} \quad (14.17)$$

Το ζητούμενο της τράπεζας που πουλάει αυτό το δικαίωμα είναι να το τιμολογήσει, δηλαδή να αποφασίσει σε ποια τιμή θα το πουλάει, και γι' αυτόν το λόγο θέλει να ξέρει ποια είναι μέση προεξοφλημένη απόδοση $\mu = E[G(X(T))]$.

Μια από τις βασικές υποθέσεις των χρηματοοικονομικών είναι η απουσία «arbitrage» ή «επιτηδειότητας», δηλαδή ότι δεν υπάρχει στην αγορά κάποια επένδυση η οποία μπορεί να επιφέρει κέρδη με 100% βεβαιότητα. Ο τρόπος με τον οποίο αυτή η υπόθεση αντανακλάται στα σχετικά μαθηματικά μοντέλα είναι μέσω της χρήσης ενός μέτρου πιθανότητας που λέγεται «μέτρο αδιάφορο κινδύνου». Στη δική μας περίπτωση, αυτή η υπόθεση συνεπάγεται ότι η μέση προεξοφλημένη τιμή της μετοχής μετά από οποιοδήποτε χρονικό διάστημα ισούται με τη σημερινή τιμή της $X(0)$ η οποία μας είναι γνωστή, δηλαδή, $E[e^{-rT}X(T)] = X(0)$. Άρα, αν ορίσουμε $H(X(T)) = X(T)$, τότε ξέρουμε τη μέση τιμή,

$$\nu = E[H(X(T))] = E[X(T)] = e^{rT}X(0).$$

Η πιο απλή υπόθεση για την τιμή της μετοχής τη χρονική στιγμή T είναι αυτή του μοντέλου «γεωμετρικής κίνησης Brown», το οποίο εδώ αντιστοιχεί στο να θεωρήσουμε ότι,

$$X(T) = X(0) \exp \left\{ \left(r - \frac{1}{2}\sigma^2 \right) T + \sigma\sqrt{T}Z \right\}, \quad (14.18)$$

όπου η $Z \sim N(0, 1)$ και η παράμετρος σ είναι η μεταβλητότητα της μετοχής. (Στην Άσκηση 11 θα δούμε ότι ο παραπάνω ορισμός (14.18) πράγματι συνεπάγεται ότι $E[e^{-rT}X(T)] = X(0)$.)

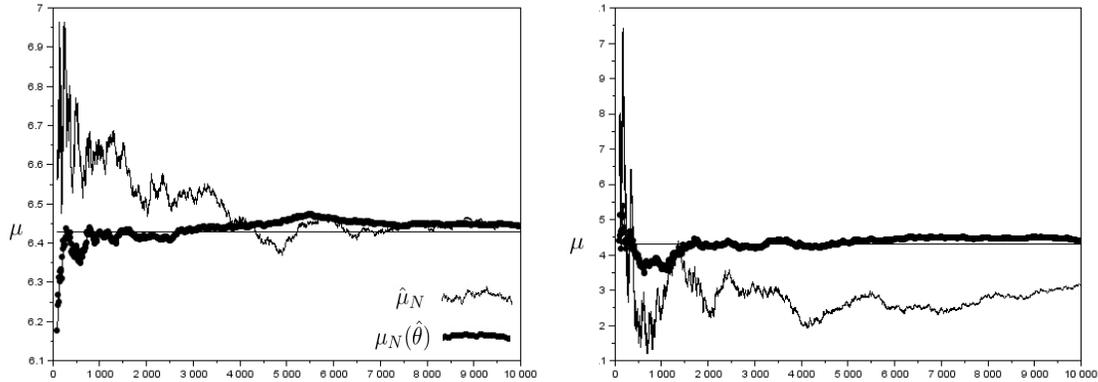
Κάτω από αυτές τις υποθέσεις, προκειμένου να εκτιμήσουμε τη μέση προεξοφλημένη απόδοση $\mu = E[G(X(T))]$, μπορούμε να πάρουμε N ανεξάρτητα δείγματα $Z_i \sim N(0, 1)$ και από αυτά να υπολογίσουμε, για κάθε i , τη μελλοντική τιμή της μετοχής $X_i(T)$ από τη σχέση (14.18) και την τιμή της αντίστοιχης προεξοφλημένης απόδοσης του δικαιώματος $G(X_i(T))$ από τη (14.17). Έτσι, έχουμε τα $G(X_i(T))$ και τις αντίστοιχες τιμές της μεταβλητής ελέγχου $H(X_i(T)) = X_i(T)$, οπότε μπορούμε να ορίσουμε την κλασική εκτιμήτρια,

$$\hat{\mu}_N = \frac{1}{N} \sum_{i=1}^N G(X_i(T)),$$

και την προσαρμοστική εκτιμήτρια,

$$\begin{aligned} \mu_N(\hat{\theta}) &= \frac{1}{N} \sum_{i=1}^N (G(X_i(T)) - \hat{\theta}[H(X_i(T)) - \nu]) \\ &= \frac{1}{N} \sum_{i=1}^N (G(X_i(T)) - \hat{\theta}[X_i(T) - e^{rT}X(0)]), \end{aligned}$$

ακολουθώντας τη γενική μέθοδο που δώσαμε στους τύπους (14.14), (14.15) και (14.16). Επιλέγοντας για τις παραμέτρους τις (ρεαλιστικές) τιμές $r = 5\%$, $\sigma = 30\%$, $T = 0.25$, $X(0) = 50$ και $K = 45$, και ξεκινώντας από $N = 10000$ προσομοιωμένα δείγματα $Z_i \sim N(0, 1)$, τα αποτελέσματα των δύο εκτιμητριών $\hat{\mu}_N$ και $\mu_N(\hat{\theta})$ σε δύο διαφορετικά πειράματα δίνονται στο Σχήμα 14.6.



Σχήμα 14.6: Σε δύο διαφορετικά πειράματα προσομοίωσης, βλέπουμε τις τιμές των εκτιμητριών $\mu_N(\hat{\theta})$ (έντονη γραμμή) και $\hat{\mu}_N$ (απλή γραμμή) για την προεξοφλημένη απόδοση $\mu = E[G(X(T))]$ μιας μετοχής στο Παράδειγμα 14.13. Είναι και εδώ εμφανές ότι η προσαρμοστική εκτιμήτρια $\mu_N(\hat{\theta})$ είναι πολύ αποτελεσματικότερη από τον εμπειρικό μέσο $\hat{\mu}_N$, όπως προβλέπει το Θεώρημα 14.3. Η βέλτιστη τιμή του θ σε αυτό το παράδειγμα είναι $\theta^* \approx 0.815$. Η οριζόντια γραμμή είναι η πραγματική τιμή $\mu = E[G(X)]$.

14.4 Ασκήσεις

- Διάρκειες τραγουδιών.** Τον Ιούλιο του 2015, η εταιρία διαχείρισης πνευματικών δικαιωμάτων *Κορδατέας Α.Ε.* πραγματοποίησε δειγματοληπτικές μετρήσεις της διάρκειας τραγουδιών που παίζονταν στο ραδιόφωνο. Με $N = 145$ δείγματα, η έρευνα κατέληξε στην εκτίμηση ότι η μέση διάρκεια των τραγουδιών ήταν 3.12 λεπτά.

 - Αν η τυπική απόκλιση της τυχαίας διάρκειας ενός τραγουδιού είναι 54 δευτερόλεπτα, υπολογίστε ένα διάστημα εμπιστοσύνης για την παραπάνω εκτίμηση, με επίπεδο εμπιστοσύνης $c = 98\%$.
 - Πόσα δείγματα θα χρειάζονταν ώστε να μπορεί η εταιρία να δώσει ένα 98%-διάστημα εμπιστοσύνης με στατιστικό σφάλμα μόλις ± 5 δευτερόλεπτα;
- Ελέφαντες.** Προκειμένου να εκτιμήσουμε το μέσο μέγεθος μ μιας αγέλης ελεφάντων, καταγράφουμε τα μεγέθη X_1, X_2, \dots, X_N από $N = 83$ αγέλες που συναντάμε, και εκτιμούμε το μ μέσω του εμπειρικού μέσου όρου, ο οποίος υπολογίζουμε πως είναι $\bar{X}_N = 7.2$.

Αν η τυπική απόκλιση των τυχαίων X_i είναι μεταξύ 5 και 7, βρείτε ένα 90%-διάστημα εμπιστοσύνης για αυτή την εκτίμηση.
- Ναύτες.** Θέλουμε να εκτιμήσουμε το μέσο ύψος μ μεταξύ των 4000 νεοσύλλεκτων σε ένα στρατόπεδο του Ναυτικού. Επιλέγουμε $N = 108$ από αυτούς τυχαία (χωρίς επαναποθέτηση) και εκτιμούμε ότι το μ είναι κατά προσέγγιση $\bar{X}_N = 1.77$ μέτρα. Επιπλέον, θεωρούμε ότι η τυπική απόκλιση του ύψους των 4000 ναυτών είναι 3 με 9 εκατοστά.

 - Υπολογίστε ένα διάστημα εμπιστοσύνης $[a, b]$ για την εκτίμηση $\mu \approx 1.77$, με επίπεδο εμπιστοσύνης 98%.
 - Μπορούμε να συμπεράνουμε από το παραπάνω αποτέλεσμα ότι το ύψος του 98% των ναυτών στο στρατόπεδο είναι μεταξύ a και b ; Τεχνηρώστε την απάντησή σας.
 - Αν θέλουμε ένα 99%-διάστημα εμπιστοσύνης γι' αυτή την εκτίμηση, πόσο μεγάλο μέγεθος δείγματος απαιτείται;
 - Ο διοικητής του στρατοπέδου Αρχιπλοίαρχος Παν. Τζίφας επιμένει ότι θέλει αποτελέσματα με στατιστικό σφάλμα το πολύ ± 1 εκατοστό, και δεν μας επιτρέπει να πάρουμε επιπλέον δείγματα από τα $N = 108$ που είχαμε στην αρχή. Ποιο είναι το καλύτερο αποτέλεσμα που μπορούμε να του δώσουμε;
- Μια πιο ρεαλιστική δημοσκόπηση.** Όπως στο Παράδειγμα 14.2, από ένα τυχαίο δείγμα $N = 1000$ ψηφοφόρων θέλουμε να εκτιμήσουμε το ποσοστό που θα πάρει η ΝΔ στις εκλογές και βρίσκουμε πως $\bar{X}_N = 30.5\%$.

 - Βρείτε ένα 95%-διάστημα εμπιστοσύνης για την εκτίμηση \bar{X}_N όταν το πραγματικό ποσοστό p των ψηφοφόρων της ΝΔ μας είναι εντελώς άγνωστο.

(β') Βρείτε ένα αντίστοιχο διάστημα εμπιστοσύνης στην περίπτωση που (εύλογα) υποθέτουμε πως το πραγματικό ποσοστό p είναι μεταξύ 20% και 35%. Σχολιάστε τη διαφορά του αποτελέσμάτος σας με αυτό του προηγούμενου σκέλους.

5. * *Εκτίμηση διασποράς.* Έστω X_1, X_2, \dots, X_N ανεξάρτητες τυχαίες μεταβλητές, όλες με την ίδια (άγνωστη) κατανομή και με (επίσης άγνωστες) μέση τιμή $\mu = E(X_i)$ και διασπορά $\sigma^2 = \text{Var}(X_i)$. Η πιο συνηθισμένη εκτιμήτρια για τη μέση τιμή τους είναι ο εμπειρικός μέσος όρος, $\hat{\mu} = \bar{X}_N$, αλλά στη στατιστική βιβλιογραφία συναντάμε δύο διαφορετικές εκτιμήτριες για τη διασπορά:

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^N (X_i - \hat{\mu})^2,$$

$$\text{και } \hat{s}^2 = \frac{1}{N-1} \sum_{i=1}^N (X_i - \hat{\mu})^2.$$

(α') Υπολογίστε τη μέση τιμή $E[(X_1 - \hat{\mu})^2]$.

(β') Υπολογίστε τη μέση τιμή $E(\hat{\sigma}^2)$ και $E(\hat{s}^2)$ των δύο παραπάνω εκτιμητριών.

(γ') Μπορείτε να βρείτε κάποιο συγκριτικό πλεονέκτημα της κάθε εκτιμήτριας ως προς την άλλη;

6. *Τυχαίο; Νομίζω.* Όλοι έχουμε έναν φίλο ο οποίος ισχυρίζεται ότι φέρνει πολύ συχνά διπλές στο τάβλι. Για να ελέγξουμε τις «δυνατότητες» ενός τέτοιου ταβλαδόρου, μετράμε ότι έφερε 62 φορές διπλές στις 275 ζαριές που έριξε παίζοντας πέντε διαδοχικές παρτίδες. Αυτός θεωρεί ότι μας απέδειξε τον ισχυρισμό του διότι, λέει, «62 στις 275 είναι πολύ περισσότερες από μία στις έξι». Εμείς τι μπορούμε να του απαντήσουμε;
7. *Στυτική δυσλειτουργία.* Εξετάζοντας τα ερωτηματολόγια της κλινικής μελέτης ενός νέου αντικαταθλιπτικού φαρμάκου παρατηρούμε ότι, από τους 66 (τυχαία επιλεγμένους, άντρες) ασθενείς που παίρνουν το φάρμακο, οι 36 παρουσίασαν συμπτώματα στυτικής δυσλειτουργίας, ενώ από τους 59 που παίρνουν το placebo μόνο οι 28 αντιμετώπισαν αντίστοιχα προβλήματα. Μπορούμε να συμπεράνουμε με επίπεδο σημαντικότητας 5% ότι η λήψη του φαρμάκου σχετίζεται με το πρόβλημα, ή όχι;
8. *Γιωργάκη, θα πουντιάσεις!* Έχουν γίνει πολλές μελέτες με σκοπό να ελεγχθεί ο στερεότυπος ισχυρισμός του γονιού που λέει στο παιδί του ότι θα αρρωστήσει αν βγει στο κρύο με βρεγμένα μαλλιά.

Έστω πως πραγματοποιείται μια τέτοια μελέτη με $N = 173$ άτομα ηλικίας 18-25 ετών. Αφού βρέξουν τα μαλλιά τους με κρύο νερό, για τον καθένα επιλέγεται τυχαία, με κάποια πιθανότητα q , αν (1) θα κάτσει για μία ώρα δίπλα στο τζάκι σε ένα ζεστό δωμάτιο, ή (2) θα βγει για μία ώρα έξω σε θερμοκρασία 2°C , με πολύ ελαφρύ ντύσιμο. Επιπλέον, για

κάθε άτομο $i = 1, 2, \dots, N$ καταγράφουμε δύο μεταβλητές: Την Y_i η οποία ισούται με 1 ή 0 αν κάθισε στο ζεστό δωμάτιο ή βγήκε έξω αντίστοιχα, και την X_i η οποία παίρνει τις τιμές $-1, 0$ και $+1$ ανάλογα με το αν η κατάσταση της υγείας του μετά από 48 ώρες έχει χειροτερέψει, έχει μείνει σταθερή ή έχει καλύτερη.

- (α') Ορίστε ένα μοντέλο για την κατανομή των δεδομένων $(X_1, Y_1), \dots, (X_N, Y_N)$.
- (β') Θεωρώντας πως το ζητούμενο είναι να ελεγχθεί αν η θερμοκρασία του περιβάλλοντος επηρεάζει το αν κάποιος πρόκειται να αρρωστήσει ή όχι, περιγράψτε λεπτομερώς τις παραμέτρους του μοντέλου, τη μηδενική υπόθεση και την εναλλακτική υπόθεση.
- (γ') Ακολουθώντας το ίδιο σκεπτικό όπως στις Ενότητες 14.2.1 και 14.2.2, ορίστε μια στατιστική συνάρτηση Δ_N γι' αυτό το πρόβλημα και περιγράψτε τις αντίστοιχες εμπειρικές κατανομές που θα χρειαστείτε.
- (δ') Μπορεί να αποδειχθεί ότι, κάτω από τη μηδενική υπόθεση, η Δ_N συγκλίνει κατά κατανομή στην $\chi^2(2)$, καθώς το $N \rightarrow \infty$. Περιγράψτε μια μέθοδο ελέγχου ανεξαρτησίας για τις παραπάνω υποθέσεις, βασισμένη στη Δ_N του προηγούμενου ερωτήματος. Πώς υπολογίζεται η p -τιμή εδώ; [Υπόδειξη. Η Άσκηση 2 του Κεφαλαίου 15 δίνει κάποιες πληροφορίες για την κατανομή $\chi^2(2)$.]
- (ε') Τα δεδομένα από την παραπάνω μελέτη συνοψίζονται στον πίνακα:

ΠΛΗΘΟΣ ΑΤΟΜΩΝ ΑΝΑ ΠΕΡΙΠΤΩΣΗ	χειρότερα ($X = -1$)	το ίδιο ($X = 0$)	καλύτερα ($X = +1$)
τζάκι ($Y = 1$)	6	65	8
κρύο ($Y = 0$)	13	71	10

Εφαρμόστε τη μέθοδο ελέγχου του προηγούμενου ερωτήματος σε αυτές τις παρατηρήσεις. Τι μπορείτε να συμπεράνετε; Έχει δίκιο η περιβόητη Ελληνίδα μάνα;

9. *Μεταβλητές ελέγχου.* Χρησιμοποιώντας το scilab, το matlab ή όποιο άλλο περιβάλλον προγραμματισμού σάς βολεύει, υλοποιήστε το προγραμματιστικό μέρος των Παραδειγμάτων 14.12 και 14.13 και συγκρίνετε τα αποτελέσματά σας με αυτά που είδαμε στην Ενότητα 14.3.
10. *Διαφορετική μεταβλητή ελέγχου.* Αν στο πρόβλημα του Παραδείγματος 14.12 αντί για τη μεταβλητή ελέγχου $H(x) = x$ θέλουμε να χρησιμοποιήσουμε την $H(x) = x^3$, πώς εφαρμόζεται η γενική μέθοδος που περιγράφεται στην Ενότητα 14.3; Χρησιμοποιώντας και πάλι όποιο προγραμματιστικό περιβάλλον σάς βολεύει, υλοποιήστε το αντίστοιχο πείραμα και συγκρίνετε τα αποτελέσματα που δίνει η προσαρμοστική εκτιμήτρια για την κάθε μια από τις δύο μεταβλητές ελέγχου. Ποια σας φαίνεται πιο αποτελεσματική;
11. *Προεξοφλημένη τιμή μετοχής.* Αποδείξτε απευθείας από τη σχέση (14.18) ότι η προεξοφλημένη τιμή $E[e^{-rT} X(T)]$ της μετοχής πράγματι ισούται με τη σημερινή τιμή της $X(0)$.

Κεφάλαιο 15

Συνεχής από κοινού κατανομή

Στα Κεφάλαια 9 έως 12 συναντήσαμε μια σειρά ιδιοτήτων της από κοινού κατανομής δύο ή περισσότερων διακριτών Τ.Μ. Εδώ θα αναπτύξουμε τις αντίστοιχες ιδιότητες για συνεχείς Τ.Μ. Θα ορίσουμε την από κοινού πυκνότητα ενός πεπερασμένου πλήθους Τ.Μ. με συνεχή κατανομή, και θα αποδείξουμε τις βασικές ιδιότητες της από κοινού πυκνότητας, της μέσης τιμής και της διασποράς τους. Μια τεχνική διαφορά που προκύπτει είναι ότι συχνά σε υπολογισμούς πιθανοτήτων εμφανίζονται διπλά ή πολλαπλά ολοκληρώματα. Κατά συνέπεια, ο μαθηματικός χειρισμός που απαιτείται βασίζεται στα εργαλεία της αρκετά πιο σύνθετης περιοχής του απειροστικού λογισμού συναρτήσεων πολλών μεταβλητών, και αυτό το κεφάλαιο είναι μάλλον πιο τεχνικό από τα περισσότερα προηγούμενα. Παρ' όλα αυτά θα προσπαθήσουμε να διατηρήσουμε τις μαθηματικές μας απαιτήσεις στο ελάχιστο. Γι' αυτόν το λόγο επικεντρωνόμαστε κυρίως στη μελέτη της από κοινού κατανομής δύο μόνο Τ.Μ., και στο Παράρτημα Α.1 συμπεριλαμβάνουμε μια σύντομη επισκόπηση κάποιων από τις απλούστερες τεχνικές υπολογισμού διπλών ολοκληρωμάτων.

15.1 Από κοινού πυκνότητα

Ορισμός 15.1 1. Δύο τυχαίες μεταβλητές X και Y είναι από κοινού συνεχείς όταν υπάρχει μια συνάρτηση $f_{XY} : \mathbb{R} \times \mathbb{R} \rightarrow [0, \infty)$, η από κοινού πυκνότητα των X, Y , τέτοια ώστε, για οποιαδήποτε $a < b$ και $c < d$, να ισχύει ότι,

$$\Pr(a \leq X \leq b, c \leq Y \leq d) = \int_c^d \int_a^b f_{XY}(x, y) dx dy, \quad (15.1)$$

και γενικότερα, για οποιοδήποτε σύνολο R στο επίπεδο, $R \subseteq \mathbb{R} \times \mathbb{R}$:

$$\Pr((X, Y) \in R) = \iint_R f_{XY}(x, y) dx dy. \quad (15.2)$$

2. Οι τυχαίες μεταβλητές X_1, X_2, \dots, X_N είναι **από κοινού συνεχείς** όταν υπάρχει μια συνάρτηση $f_{X_1 X_2 \dots X_N} : \mathbb{R}^N \rightarrow [0, \infty)$, η **από κοινού πυκνότητα** των X_1, X_2, \dots, X_N , τέτοια ώστε, για οποιαδήποτε $a_i < b_i, i = 1, 2, \dots, N$, να ισχύει ότι,

$$\begin{aligned} \Pr(a_1 \leq X_1 \leq b_1, a_2 \leq X_2 \leq b_2, \dots, a_N \leq X_N \leq b_N) \\ = \int_{a_N}^{b_N} \cdots \int_{a_2}^{b_2} \int_{a_1}^{b_1} f_{X_1 X_2 \dots X_N}(x_1, x_2, \dots, x_N) dx_1 dx_2 \cdots dx_N, \end{aligned}$$

και γενικότερα, για οποιαδήποτε $R \subseteq \mathbb{R}^N$:

$$\Pr((X_1, X_2, \dots, X_N) \in R) = \int_R \cdots \int f_{X_1 X_2 \dots X_N}(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n. \quad (15.3)$$

Παράδειγμα 15.1 Ως μια από τις απλούστερες δυνατές περιπτώσεις, εξετάζουμε δύο Τ.Μ. X, Y «ομοιόμορφα» κατανεμημένες στο $S = [0, 1]^2$, δηλαδή με από κοινού πυκνότητα $f_{XY}(x, y)$ που ισούται με 1 αν $0 \leq x \leq 1$ και $0 \leq y \leq 1$, και $f_{XY}(x, y) = 0$ για όλα τα υπόλοιπα (x, y) . Εδώ πολύ εύκολα μπορούμε, π.χ., να υπολογίσουμε την πιθανότητα,

$$\begin{aligned} \Pr(X \in [0, 1/2] \text{ και } Y \in [0, 1/2]) &= \Pr((X, Y) \in [0, 1/2] \times [0, 1/2]) \\ &= \int_0^{1/2} \int_0^{1/2} 1 dy dx \\ &= \int_0^{1/2} [y]_0^{1/2} dx \\ &= \int_0^{1/2} \frac{1}{2} dx \\ &= \frac{1}{4}, \end{aligned}$$

και παρομοίως βρίσκουμε,

$$\begin{aligned} \Pr(0 \leq Y \leq X \leq 3/4) &= \int_0^{3/4} \int_0^x 1 dy dx \\ &= \int_0^{3/4} [y]_0^x dx \\ &= \int_0^{3/4} x dx \\ &= \left[\frac{x^2}{2} \right]_0^{3/4} \\ &= \frac{9}{32}. \end{aligned}$$

Παρατηρήσεις

1. Όπως στον υπολογισμό της πρώτης πιθανότητας στο παραπάνω παράδειγμα, η πιο απλή (και αρκετά συνηθισμένη) περίπτωση υπολογισμού μιας από κοινού πιθανότητας για δύο συνεχείς Τ.Μ. X, Y είναι αυτή που είδαμε στην πρώτη περίπτωση του ορισμού, στη σχέση (15.1): Για $a < b$ και $c < d$, η πιθανότητα,

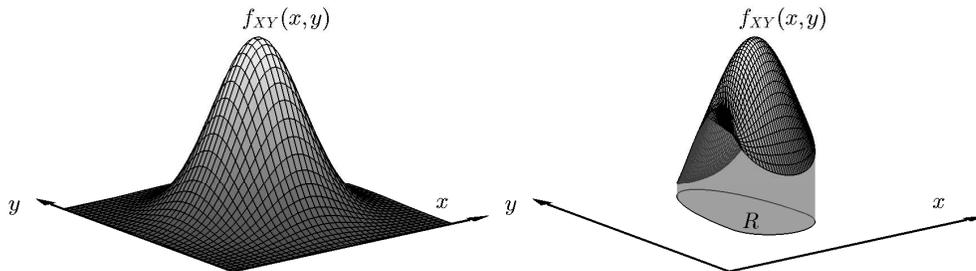
$$\Pr(a \leq X \leq b, c \leq Y \leq d) = \Pr((X, Y) \in [a, b] \times [c, d]), \tag{15.4}$$

μπορεί ισοδύναμα να εκφραστεί ως,

$$\Pr(a \leq X \leq b, c \leq Y \leq d) = \int_a^b \left(\int_c^d f_{XY}(x, y) dy \right) dx = \int_c^d \left(\int_a^b f_{XY}(x, y) dx \right) dy,$$

όπου το γεγονός πως οι δύο τελευταίες εκφράσεις είναι ίδιες και ίσες με τον ορισμό, προκύπτει από το θεώρημα του Fubini. (Βλ. Παράρτημα Α'1.)

2. Όταν κάποια πιθανότητα που μας ενδιαφέρει δεν είναι της παραπάνω απλής μορφής, υπάρχει μια γεωμετρική περιγραφή που ορισμένες φορές διευκολύνει τον υπολογισμό της. Όπως είδαμε στο Κεφάλαιο 10, οι πιθανότητες που αφορούν μια συνεχή Τ.Μ. υπολογίζονται μέσω απλών ολοκληρωμάτων της πυκνότητάς της και μπορούν να ερμηνευτούν γραφικά ως εμβαδά. Παρομοίως, όταν έχουμε $N = 2$ ή περισσότερες συνεχείς Τ.Μ., οι πιθανότητες υπολογίζονται μέσω πολλαπλών ολοκληρωμάτων και επομένως γεωμετρικά αντιστοιχούν σε $(N + 1)$ -διάστατους όγκους. Έστω, για παράδειγμα, δύο Τ.Μ. X, Y με από κοινού πυκνότητα $f_{XY}(x, y)$ όπως στο Σχήμα 15.1. Η πιθανότητα το (X, Y) να παίρνει τιμές στο σύνολο $R \subseteq \mathbb{R}^2$ ισούται με τον όγκο του στερεού που οριοθετείται από το R και το γράφημα της συνάρτησης· βλ. Σχήμα 15.1. Ένας απλός τέτοιος υπολογισμός δίνεται στο αμέσως επόμενο παράδειγμα.

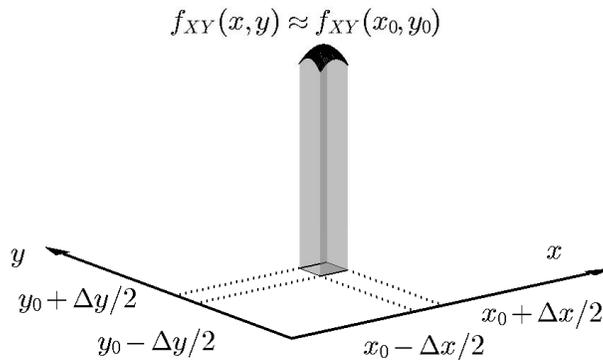


Σχήμα 15.1: Παράδειγμα μιας από κοινού πυκνότητας $f_{XY}(x, y)$ (αριστερά). Η πιθανότητα $\Pr((X, Y) \in R)$ τα (X, Y) να πάρουν τιμές στο R δίνεται από τον όγκο του σκιασμένου στερεού που βρίσκεται ανάμεσα στο γράφημα της $f_{XY}(x, y)$ και το R (δεξιά).

3. Διαισθητικά, η τιμή $f_{XY}(x_0, y_0)$ της από κοινού πυκνότητας δύο Τ.Μ. X, Y στο σημείο (x_0, y_0) , εκφράζει τη σχετική πιθανότητα οι Τ.Μ. (X, Y) να πάρουν τιμές «κοντά» στο (x_0, y_0) . Πράγματι, έστω πως η $f_{XY}(x, y)$ είναι συνεχής στο (x_0, y_0) . Τότε, από το θεμελιώδες θεώρημα του ολοκληρωτικού λογισμού,

$$\begin{aligned} \Pr \left(x_0 - \frac{\Delta x}{2} \leq X \leq x_0 + \frac{\Delta x}{2}, y_0 - \frac{\Delta y}{2} \leq Y \leq y_0 + \frac{\Delta y}{2} \right) \\ = \int_{y_0 - \Delta y/2}^{y_0 + \Delta y/2} \int_{x_0 - \Delta x/2}^{x_0 + \Delta x/2} f_{XY}(x, y) dx dy \\ \approx f_{XY}(x_0, y_0) \Delta x \Delta y, \end{aligned}$$

όπως αναπαρίσταται και στο Σχήμα 15.2. Συνεπώς, όσο πιο μεγάλη είναι η τιμή της πυκνότητας, τόσο πιο πιθανό είναι τα (X, Y) να βρεθούν κοντά στο (x_0, y_0) .



Σχήμα 15.2: Καθώς τα $\Delta x, \Delta y \rightarrow 0$, το στερεό που δημιουργείται ανάμεσα στο γράφημα της $f_{XY}(x, y)$ και το ορθογώνιο $\left[x_0 - \frac{\Delta x}{2}, x_0 + \frac{\Delta x}{2} \right] \times \left[y_0 - \frac{\Delta y}{2}, y_0 + \frac{\Delta y}{2} \right]$ προσεγγίζει ένα ορθογώνιο παραλληλεπίπεδο, του οποίου ο όγκος του ισούται με $f_{XY}(x_0, y_0) \Delta x \Delta y$.

4. Στο Κεφάλαιο 10, ο ορισμός της πυκνότητας μιας συνεχούς Τ.Μ. X διατυπώθηκε μέσω του ορισμού της πιθανότητας για ενδεχόμενα μόνο της απλής μορφής $\{a \leq X \leq b\}$. Αντίθετα, στον Ορισμό 15.1 της από κοινού πυκνότητας, εκτός από τα απλά ενδεχόμενα $\{a_1 \leq X_1 \leq b_1, a_2 \leq X_2 \leq b_2, \dots, a_N \leq X_N \leq b_N\}$ συμπεριλάβαμε στις σχέσεις (15.3) και (15.3) την περίπτωση υπολογισμού της πιθανότητας των πιο γενικών ενδεχομένων $\{(X_1, X_2, \dots, X_N) \in R\}$, για αυθαίρετα σύνολα $R \subseteq \mathbb{R}^N$. Αν και εκ πρώτης όψευς ο Ορισμός 15.1 φαίνεται πιο γενικός, κάτω από τους περιορισμούς που θέτουμε στις Ενότητες 10.3 και 15.4 μπορεί ναδειχθεί πως οι δύο ορισμοί είναι ισοδύναμοι. Κάποιες περαιτέρω λεπτομέρειες δίνονται στην Ενότητα 15.4.

Παράδειγμα 15.2 Έστω πως οι Τ.Μ. (X, Y) είναι συνεχείς, παίρνουν τιμές στο δίσκο με ακτίνα 1, $S = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ και έχουν από κοινού πυκνότητα,

$$f_{XY}(x, y) = \begin{cases} \frac{3}{2\pi} \sqrt{1 - x^2 - y^2}, & \text{αν } (x, y) \in S, \\ 0, & \text{αν } (x, y) \notin S. \end{cases}$$

Το γράφημα της f_{XY} είναι απλά η επιφάνεια του «βόρειου ημισφαιρίου» της σφαίρας με ακτίνα 1 και κέντρο το κέντρο των αξόνων.

Έστω πως θέλουμε να υπολογίσουμε την πιθανότητα $\Pr(X > 0, Y > 0)$. Από τη δεύτερη παρατήρηση παραπάνω, αυτή ισούται με τον όγκο που αντιστοιχεί στο μέρος της σφαίρας το οποίο αποτελείται από σημεία (x, y, z) με θετικές συντεταγμένες, δηλαδή, στο 1/8 της πλήρους σφαίρας. Και αφού, ως γνωστόν, η σφαίρα με ακτίνα 1 έχει συνολικό όγκο $4\pi/3$, η πιθανότητα $\Pr(X > 0, Y > 0)$ ισούται με $\pi/6$.

Συγκεντρώνουμε πιο κάτω κάποιες από τις απλούστερες ιδιότητες της από κοινού πυκνότητας. Οι αποδείξεις, εκτός από εκείνη της τέταρτης ιδιότητας, παραλείπονται, μια και είναι ακριβώς ίδιες με εκείνες των αντίστοιχων αποτελεσμάτων που συναντήσαμε στα Κεφάλαια 9 έως 12

Ορισμός 15.2 Η περιθώρια πυκνότητα της X_i είναι η πυκνότητα $f_{X_i}(x_i)$ μίας από N συνεχείς τυχαίες μεταβλητές X_1, X_2, \dots, X_N με από κοινού πυκνότητα $f_{X_1 X_2 \dots X_N}(x_1, x_2, \dots, x_N)$.

Συνεχής από κοινού πυκνότητα: Βασικές ιδιότητες. Για οποιοδήποτε ζεύγος από συνεχείς Τ.Μ. X, Y με από κοινού πυκνότητα $f_{XY}(x, y)$, και για οποιοδήποτε πεπερασμένο πλήθος από συνεχείς Τ.Μ. X_1, X_2, \dots, X_N με από κοινού πυκνότητα $f_{X_1 X_2 \dots X_N}(x_1, x_2, \dots, x_N)$, έχουμε:

$$1. \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y) dx dy = 1 \text{ και } \int_{\mathbb{R}^N} \dots \int f_{X_1 X_2 \dots X_N}(x_1, x_2, \dots, x_N) dx_1 dx_2 \dots dx_N = 1.$$

2. Για οποιοδήποτε μεμονωμένο σημείο $(x_0, y_0) \in \mathbb{R}^2$, η πιθανότητα,

$$\Pr((X, Y) = (x_0, y_0)) = \iint_{\{(x_0, y_0)\}} f_{XY}(x, y) dx dy = 0,$$

και αντίστοιχα η πιθανότητα η N -άδα Τ.Μ. X_1, X_2, \dots, X_N να πάρει μια οποιαδήποτε συγκεκριμένη τιμή (x_1, x_2, \dots, x_N) ισούται με μηδέν.

Κατά συνέπεια, αν αλλάξουμε τις τιμές της από κοινού πυκνότητας σε ένα, δύο ή οποιοδήποτε πεπερασμένο πλήθος σημείων, δεν αλλάζει καμία από τις πιθανότητες που προκύπτουν για τις αντίστοιχες Τ.Μ.

3. Για οποιαδήποτε $a < b$ και $c < d$, η πιθανότητα $\Pr(a \leq X \leq b, c \leq Y \leq d)$ παραμένει η ίδια αν ένα ή περισσότερα από τα « \leq » αντικατασταθούν με « $<$ », και ισούται με καθεμία από τις εναλλακτικές εκφράσεις που δώσαμε στην πρώτη παρατήρηση παραπάνω, ανεξάρτητα από το αν και πόσα από τα οριακά σημεία του ορθογωνίου $[a, b] \times [c, d]$ περιέχονται ή όχι στον υπολογισμό της πιθανότητας. [Άσκηση. Διατυπώστε και τεκμηριώστε την προφανή γενίκευση αυτής της ιδιότητας για N συνεχείς Τ.Μ.]

4. Η X και η Y είναι συνεχείς Τ.Μ. με περιθώριες πυκνότητες, αντίστοιχα,

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy \quad \text{και} \quad f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y) dx,$$

και στη γενική περίπτωση, η κάθε X_i είναι συνεχής, με περιθώρια πυκνότητα,

$$\int_{\mathbb{R}^{N-1}} \cdots \int f_{X_1 X_2 \dots X_N}(x_1, x_2, \dots, x_N) dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_N.$$

Θα αποδείξουμε μόνο την πρώτη από τις παραπάνω σχέσεις: οι αποδείξεις των υπολοίπων είναι πανομοιότυπες. Παρατηρούμε πως, για οποιαδήποτε $a < b$, η $\Pr(a \leq X \leq b)$ μπορεί να εκφραστεί,

$$\Pr(a \leq X \leq b) = \Pr(a \leq X \leq b, -\infty < Y < \infty) = \int_a^b \left(\int_{-\infty}^{\infty} f_{XY}(x, y) dy \right) dx,$$

άρα, η συνάρτηση $f_X(x)$ όπως ορίζεται πιο πάνω, ικανοποιεί τον ορισμό της πυκνότητας μιας συνεχούς Τ.Μ. X , όπως δίνεται στο Κεφάλαιο 10.

Παράδειγμα 15.3 Έστω δύο συνεχείς Τ.Μ. (X, Y) με από κοινού πυκνότητα,

$$f_{XY}(x, y) = \begin{cases} \frac{1}{3}(2x + 4y), & \text{για } (x, y) \in [0, 1]^2, \\ 0, & \text{για } (x, y) \notin [0, 1]^2, \end{cases}$$

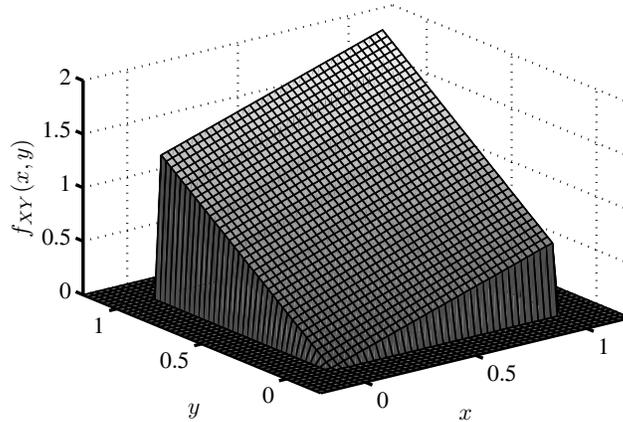
η οποία απεικονίζεται στο Σχήμα 15.3. Το ολοκλήρωμά της σε ολόκληρο το \mathbb{R}^2 είναι,

$$\iint_{\mathbb{R} \times \mathbb{R}} f_{XY}(x, y) dx dy = \int_0^1 \int_0^1 f_{XY}(x, y) dx dy = \int_0^1 \int_0^1 \frac{2x}{3} dx dy + \int_0^1 \int_0^1 \frac{4y}{3} dx dy,$$

το οποίο εύκολα υπολογίζεται,

$$\begin{aligned} \int_0^1 \frac{2x}{3} \left(\int_0^1 dy \right) dx + \int_0^1 \frac{4y}{3} \left(\int_0^1 dx \right) dy &= \int_0^1 \frac{2x}{3} dx + \int_0^1 \frac{4y}{3} dy \\ &= \left[\frac{x^2}{3} \right]_0^1 + \left[\frac{2y^2}{3} \right]_0^1 = \frac{1}{3} + \frac{2}{3} = 1, \end{aligned}$$

και φυσικά ισούται με 1 όπως προβλέπει η πρώτη ιδιότητα της από κοινού πυκνότητας.



Σχήμα 15.3: Η από κοινού πυκνότητα του Παραδείγματος 15.3. Εκτός του $S = [0, 1] \times [0, 1]$ η $f_{XY}(x, y)$ είναι ίση με μηδέν, και στο S είναι γραμμική ως προς τα x, y .

Θα υπολογίσουμε δύο απλές πιθανότητες. Από τον ορισμό της από κοινού πυκνότητας εύκολα βρίσκουμε πως η $\Pr(0 \leq X \leq 1/2, 0 \leq Y \leq 1/2)$ ισούται με,

$$\int_0^{1/2} \int_0^{1/2} \left(\frac{2x}{3} + \frac{4y}{3} \right) dx dy = \int_0^{1/2} \left(\int_0^{1/2} \left(\frac{2x}{3} + \frac{4y}{3} \right) dy \right) dx,$$

όπου τα δύο παραπάνω ολοκληρώματα υπολογίζονται ως,

$$\int_0^{1/2} \left[\frac{2xy}{3} + \frac{2y^2}{3} \right]_0^{1/2} dx = \int_0^{1/2} \left(\frac{x}{3} + \frac{1}{6} \right) dx = \left[\frac{x^2}{6} + \frac{x}{6} \right]_0^{1/2} = \frac{1}{8}.$$

Το στερεό του οποίου τον όγκο υπολογίσαμε έχει σχεδιαστεί στο Σχήμα 15.4.

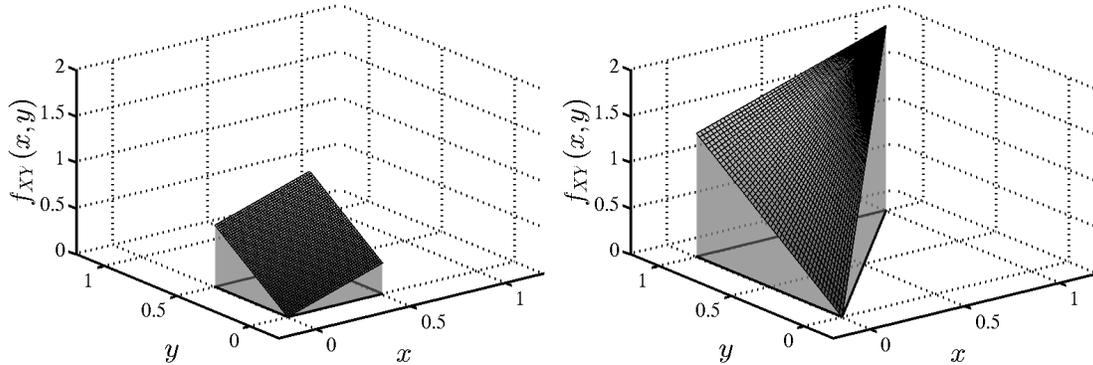
Παρομοίως υπολογίζουμε και την πιθανότητα $\Pr(X < Y)$ ως:

$$\int_0^1 \int_0^y \left(\frac{2x}{3} + \frac{4y}{3} \right) dx dy = \int_0^1 \left[\frac{x^2}{3} + \frac{4yx}{3} \right]_0^y dy = \int_0^1 \left(\frac{y^2}{3} + \frac{4y^2}{3} \right) dy = \left[\frac{5y^3}{9} \right]_0^1 = \frac{5}{9},$$

όπου, και εδώ, το στερεό του οποίου τον όγκο υπολογίσαμε φαίνεται σκιασμένο στο Σχήμα 15.4.

Τέλος, θα βρούμε τις περιθώριες πυκνότητες των X, Y βάσει της τέταρτης από τις παραπάνω ιδιότητες. Για να υπολογίσουμε την $f_X(x)$ παρατηρούμε πως, για $x < 0$ ή $x > 1$, η $f_X(x)$ ισούται με το ολοκλήρωμα μιας συνάρτησης που είναι παντού 0, δηλαδή έχουμε $f_X(x) = 0$. Ενώ για $0 \leq x \leq 1$,

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy = \int_0^1 \left(\frac{2x}{3} + \frac{4y}{3} \right) dy = \left[\frac{2xy}{3} + \frac{2y^2}{3} \right]_{y=0}^1 = \frac{2}{3}(x + 1).$$



Σχήμα 15.4: Παράδειγμα 15.3: Ο όγκος του στερεού αριστερά είναι ίσος με την πιθανότητα $\Pr(0 \leq X \leq \frac{1}{2}, 0 \leq Y \leq \frac{1}{2})$, και του στερεού δεξιά με την $\Pr(X < Y)$.

Εξίσου εύκολα βρίσκουμε και πως,

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y) dx = \begin{cases} \frac{1}{3}(4y + 1), & y \in [0, 1], \\ 0, & y \notin [0, 1]. \end{cases}$$

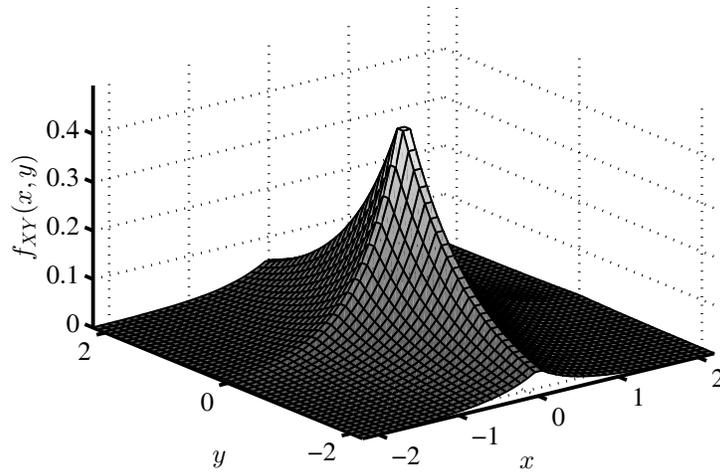
Παράδειγμα 15.4 Έστω δύο συνεχείς Τ.Μ. X, Y με από κοινού πυκνότητα την,

$$f_{XY}(x, y) = \frac{1}{2} e^{-2|x|-|y|}, \quad x, y \in \mathbb{R},$$

η οποία έχει σχεδιαστεί στο Σχήμα 15.5. Όπως και στο προηγούμενο παράδειγμα, θα υπολογίσουμε τις πιθανότητες $\Pr(0 \leq X \leq \frac{1}{2}, 0 \leq Y \leq \frac{1}{2})$ και $\Pr(X < Y)$, και θα βρούμε τις περιθώριες πυκνότητες των X και Y .

Για την πρώτη πιθανότητα έχουμε:

$$\begin{aligned} \Pr\left(0 \leq X \leq \frac{1}{2}, 0 \leq Y \leq \frac{1}{2}\right) &= \int_0^{\frac{1}{2}} \int_0^{\frac{1}{2}} \frac{e^{-2x-y}}{2} dy dx \\ &= \left(\int_0^{\frac{1}{2}} \frac{e^{-2x}}{2} dx\right) \left(\int_0^{\frac{1}{2}} e^{-y} dy\right) \\ &= \left[-\frac{1}{4}e^{-2x}\right]_0^{\frac{1}{2}} \cdot [-e^{-y}]_0^{\frac{1}{2}} \\ &= \frac{(1 - e^{-1})(1 - e^{-1/2})}{4}. \end{aligned}$$



Σχήμα 15.5: Η από κοινού πυκνότητα f_{XY} του Παραδείγματος 15.4. Η $f_{XY}(x, y)$ είναι θετική παντού στο $\mathbb{R} \times \mathbb{R}$.

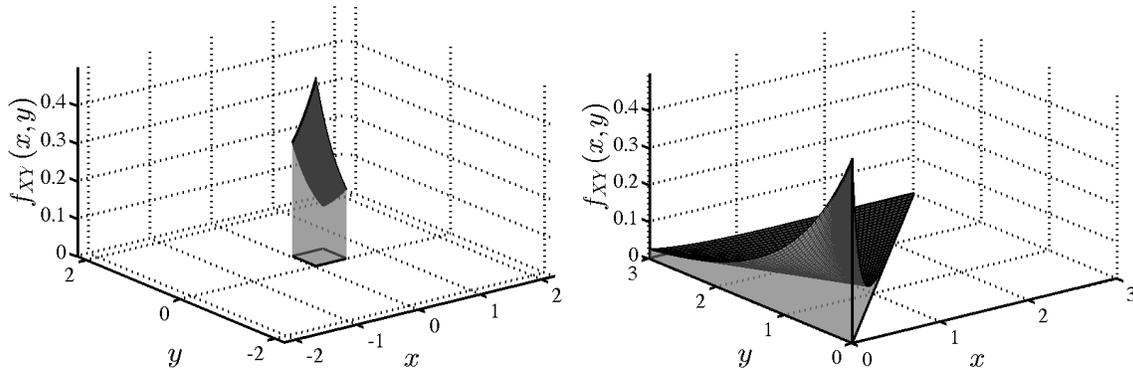
Το στερεό του οποίου τον όγκο υπολογίσαμε έχει σχεδιαστεί στο Σχήμα 15.6. Παρομοίως, για τη δεύτερη πιθανότητα βρίσκουμε:

$$\begin{aligned}
 \Pr(X < Y) &= \int_0^\infty \left(\int_0^y \frac{1}{2} e^{-2x-y} dx \right) dy = \int_0^\infty \frac{1}{2} e^{-y} \left(\int_0^y e^{-2x} dx \right) dy \\
 &= \int_0^\infty \frac{1}{2} e^{-y} \left[-\frac{e^{-2x}}{2} \right]_0^y dy = \int_0^\infty \frac{1}{2} e^{-y} \left(\frac{1}{2} - e^{-2y} \right) dy \\
 &= \int_0^\infty \left(\frac{1}{4} e^{-y} - \frac{1}{2} e^{-3y} \right) dy = \left[-\frac{1}{4} e^{-y} + \frac{1}{6} e^{-3y} \right]_0^\infty = \frac{1}{4} - \frac{1}{6} = \frac{1}{12}.
 \end{aligned}$$

Και πάλι το στερεό του οποίου τον όγκο υπολογίσαμε φαίνεται σκιασμένο στο Σχήμα 15.6. Τέλος, υπολογίζουμε τις περιθώριες πυκνότητες, χρησιμοποιώντας την τέταρτη ιδιότητα της από κοινού πυκνότητας,

$$\begin{aligned}
 f_X(x) &= \int_{-\infty}^\infty \frac{e^{-2|x|-|y|}}{2} dy = \frac{e^{-2|x|}}{2} \int_{-\infty}^\infty e^{-|y|} dy = e^{-2|x|}, \\
 f_Y(y) &= \int_{-\infty}^\infty \frac{e^{-2|x|-|y|}}{2} dx = \frac{e^{-|y|}}{2} \int_{-\infty}^\infty e^{-2|x|} dx = \frac{e^{-|y|}}{2},
 \end{aligned}$$

για κάθε $x, y \in \mathbb{R}$.



Σχήμα 15.6: Παράδειγμα 15.3: Ο όγκος του στερεού αριστερά είναι ίσος με την πιθανότητα $\Pr(0 \leq X \leq \frac{1}{2}, 0 \leq Y \leq \frac{1}{2})$, και του στερεού δεξιά με την $\Pr(X < Y)$.

15.2 Μέση τιμή, διασπορά και συνδιακύμανση

Πριν ορίσουμε τη συνδιακύμανση μεταξύ δύο Τ.Μ. X και Y με συνεχή από κοινού κατανομή, παραθέτουμε την πιο κάτω βασική ιδιότητα:

Ιδιότητα 15.1 Η μέση τιμή μιας συνάρτησης $g(x, y)$ δύο Τ.Μ. X, Y με από κοινού πυκνότητα $f_{XY}(x, y)$ υπολογίζεται ως,

$$E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{XY}(x, y) dx dy, \quad (15.5)$$

και παρομοίως η μέση τιμή μιας συνάρτησης $g(x_1, x_2, \dots, x_N)$ των N Τ.Μ. X_1, X_2, \dots, X_N με από κοινού πυκνότητα $f_{X_1 X_2 \dots X_N}(x_1, x_2, \dots, x_N)$ υπολογίζεται ως,

$$E[g(X_1, X_2, \dots, X_N)] = \int_{\mathbb{R}^N} \dots \int g(x_1, x_2, \dots, x_N) f_{X_1 X_2 \dots X_N}(x_1, x_2, \dots, x_N) dx_1 dx_2 \dots dx_N.$$

Όπως παρατηρήσαμε στην Ενότητα 10.2, η αντίστοιχη σχέση (10.10) για τον υπολογισμό της μέσης τιμής $E[g(X)]$ της συνάρτησης $g(X)$ μιας συνεχούς Τ.Μ. X , είναι συνέπεια του βασικού ορισμού της μέσης τιμής. Αλλά λόγω του ότι η απόδειξη της (10.10) από τον ορισμό είναι τεχνικής φύσεως και ξεφεύγει από τους παρόντες στόχους μας, χάριν ευκολίας επιλέξαμε να δεχθούμε τον τύπο (10.10) ως δεδομένο, παραλείποντας την απόδειξή του. Για τους ίδιους λόγους, και στη γενικότερη περίπτωση της Ιδιότητας 15.1 θα δεχτούμε τη σχέση (15.5) ως δεδομένη, παραλείποντας και πάλι την απόδειξή της.

Τώρα είμαστε σε θέση να αποδείξουμε το πρώτο σκέλος του Θεωρήματος 11.5, το οποίο διατυπώθηκε στο Κεφάλαιο 11 χωρίς απόδειξη. Η Πρόταση 15.1 που ακολουθεί δίνει μια γενικότερη

μορφή του αποτελέσματος, για οποιοδήποτε πεπερασμένο πλήθος Τ.Μ.

Πρόταση 15.1 Για οποιοδήποτε συνεχείς Τ.Μ. X_1, X_2, \dots, X_N και σταθερές a_1, a_2, \dots, a_N , έχουμε:

$$E \left[\sum_{i=1}^N a_i X_i \right] = \sum_{i=1}^N a_i E(X_i).$$

Απόδειξη: Έστω $f_{X_1 X_2 \dots X_N}(x_1, x_2, \dots, x_N)$ η από κοινού πυκνότητα των X_1, X_2, \dots, X_N . Εφαρμόζοντας την Ιδιότητα 15.1,

$$\begin{aligned} E \left[\sum_{i=1}^N a_i X_i \right] &= \int \cdots \int_{\mathbb{R}^N} \left(\sum_{i=1}^N a_i x_i \right) f_{X_1 X_2 \dots X_N}(x_1, x_2, \dots, x_N) dx_1 dx_2 \cdots dx_N \\ &= \sum_{i=1}^N \left(\int \cdots \int_{\mathbb{R}^N} a_i x_i f_{X_1 X_2 \dots X_N}(x_1, x_2, \dots, x_N) dx_1 dx_2 \cdots dx_N \right) \\ &= \sum_{i=1}^N a_i \int_{-\infty}^{\infty} x_i \left(\int \cdots \int_{\mathbb{R}^{N-1}} f_{X_1 X_2 \dots X_N}(x_1, x_2, \dots, x_N) dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_N \right) dx_i, \end{aligned}$$

όπου, από την τέταρτη βασική ιδιότητα της από κοινού πυκνότητας, η τελευταία έκφραση εντός της παραπάνω παρένθεσης ισούται με την περιθώρια πυκνότητα $f_{X_i}(x_i)$ της X_i . Συνεπώς,

$$E \left[\sum_{i=1}^N a_i X_i \right] = \sum_{i=1}^N a_i \int_{-\infty}^{\infty} f_{X_i}(x_i) dx_i = \sum_{i=1}^N a_i E(X_i),$$

που είναι ακριβώς το ζητούμενο αποτέλεσμα. ■

Παράδειγμα 15.5 Για τις Τ.Μ. X, Y του Παραδείγματος 15.3, μπορούμε να υπολογίσουμε τη μέση τιμή του γινομένου τους ως,

$$\begin{aligned} E(XY) &= \frac{1}{3} \int_0^1 \int_0^1 xy(2x+4y) dy dx \\ &= \frac{1}{3} \int_0^1 \left[x^2 y^2 + \frac{4}{3} xy^3 \right]_0^1 dx \\ &= \frac{1}{3} \int_0^1 \left(x^2 + \frac{4x}{3} \right) dx \\ &= \frac{1}{3} \left[\frac{x^3}{3} + \frac{2x^2}{3} \right]_0^1 = \frac{1}{3}. \end{aligned}$$

Ορισμός 15.3 Για δύο συνεχείς Τ.Μ. X, Y , η συνδιακύμανση $\text{Cov}(X, Y)$ μεταξύ τους ορίζεται, ακριβώς όπως και στην διακριτή περίπτωση, ως:

$$\text{Cov}(X, Y) = E[(X - E(X))(Y - E(Y))].$$

Παρατήρηση: Όπως και για την περίπτωση διακριτών Τ.Μ., η διαισθητική ερμηνεία της συνδιακύμανσης είναι πως, όταν $\text{Cov}(X, Y) > 0$, τότε οι δύο Τ.Μ. X, Y τείνουν να παίρνουν και οι δύο ταυτόχρονα μεγάλες τιμές, ή και οι δύο ταυτόχρονα μικρές τιμές. Αντίστοιχα, όταν $\text{Cov}(X, Y) < 0$, τότε, όταν η μία Τ.Μ. παίρνει μεγάλες τιμές, η άλλη τείνει να παίρνει μικρές τιμές. Άρα η συνδιακύμανση $\text{Cov}(X, Y)$ μάς παρέχει μια πρώτη ένδειξη για το είδος της συσχέτισης μεταξύ της X και της Y .

Οι αποδείξεις των παρακάτω απλών ιδιοτήτων της συνδιακύμανσης παραλείπονται, μια και είναι ακριβώς ίδιες με εκείνες που είδαμε στη διακριτή περίπτωση στο Κεφάλαιο 9, και συγκεκριμένα στο Παράδειγμα 9.9 και στις Ιδιότητες 1 και 4 της Ενότητας 9.2.

Πρόταση 15.2 Για οποιεσδήποτε δύο συνεχείς Τ.Μ. X και Y έχουμε:

1. $\text{Cov}(X, X) = \text{Var}(X)$,
2. $\text{Cov}(X, -X) = -\text{Var}(X)$,
3. $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$,
4. $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$.

Παράδειγμα 15.6 Για τις Τ.Μ. X, Y του Παραδείγματος 15.3, χρησιμοποιώντας τις περιθώριες πυκνότητες που υπολογίσαμε εκεί, έχουμε,

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x \frac{2}{3}(x+1) dx = \left[\frac{2x^3}{9} + \frac{x^2}{3} \right]_0^1 = \frac{5}{9},$$

$$\text{και } E(Y) = \int_{-\infty}^{\infty} y f_Y(y) dy = \int_0^1 y \frac{1}{3}(4y+1) dy = \frac{1}{3} \left[\frac{4y^3}{3} + \frac{y^2}{2} \right]_0^1 = \frac{11}{18}.$$

Και εφόσον στο Παράδειγμα 15.5 βρήκαμε πως $E(XY) = \frac{1}{3}$, μπορούμε εύκολα να υπολογίσουμε τη συνδιακύμανση,

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = \frac{1}{3} - \frac{5}{9} \times \frac{11}{18} = -\frac{1}{162}.$$

15.3 Ανεξαρτησία

Όπως είδαμε στον Ορισμό 11.3 του Κεφαλαίου 11, δύο συνεχείς τυχαίες μεταβλητές X, Y είναι ανεξάρτητες αν και μόνο αν, για κάθε $a \leq b, a' \leq b'$,

$$\Pr(a \leq X \leq b, a' \leq Y \leq b') = \Pr(a \leq X \leq b) \Pr(a' \leq Y \leq b'). \quad (15.6)$$

Στο επόμενο αποτέλεσμα δείχνουμε πως η ανεξαρτησία μπορεί να εκφραστεί και ως ιδιότητα της από κοινού πυκνότητας των X, Y .

Κριτήριο ανεξαρτησίας για συνεχείς Τ.Μ. Δύο συνεχείς Τ.Μ. X, Y , με περιθώριες πυκνότητες $f_X(x), f_Y(y)$ αντίστοιχα, είναι ανεξάρτητες αν και μόνο αν η από κοινού πυκνότητά τους $f_{XY}(x, y)$ μπορεί να εκφραστεί ως,

$$f_{XY}(x, y) = f_X(x)f_Y(y), \quad \text{για κάθε } x, y \in \mathbb{R}. \quad (15.7)$$

Απόδειξη: Έστω πως η από κοινού πυκνότητα των X, Y ικανοποιεί τη σχέση (15.7). Τότε, για οποιαδήποτε $a \leq b, a' \leq b'$, έχουμε,

$$\begin{aligned} \Pr(a \leq X \leq b, a' \leq Y \leq b') &= \int_a^b \int_{a'}^{b'} f_X(x)f_Y(y) dx dy \\ &= \left(\int_a^b f_X(x) dx \right) \left(\int_{a'}^{b'} f_Y(y) dy \right) \\ &= \Pr(a \leq X \leq b) \Pr(a' \leq Y \leq b'), \end{aligned}$$

και άρα οι X και Y είναι ανεξάρτητες.

Αντίστροφα, έστω πως οι X, Y είναι ανεξάρτητες με περιθώριες πυκνότητες $f_X(x), f_Y(y)$ αντίστοιχα. Από τον ορισμό της ανεξαρτησίας έχουμε, για οποιαδήποτε $a \leq b, a' \leq b'$,

$$\begin{aligned} \Pr(a \leq X \leq b, a' \leq Y \leq b') &= \Pr(a \leq X \leq b) \Pr(a' \leq Y \leq b') \\ &= \left(\int_a^b f_X(x) dx \right) \left(\int_{a'}^{b'} f_Y(y) dy \right) \\ &= \int_a^b \int_{a'}^{b'} f_X(x)f_Y(y) dx dy, \end{aligned}$$

άρα η συνθήκη (15.1) στον ορισμό της από κοινού πυκνότητας ικανοποιείται από τη συνάρτηση $f_X(x)f_Y(y)$, και συνεπώς για την από κοινού πυκνότητα των X, Y μπορούμε να θέσουμε $f_{XY}(x, y) = f_X(x)f_Y(y)$. ■

Στο ακόλουθο θεώρημα αποδεικνύονται κάποιες βασικές συνέπειες της ανεξαρτησίας, τις οποίες έχουμε ήδη συναντήσει σε προηγούμενα κεφάλαια.

Θεώρημα 15.1 (Ιδιότητες ανεξαρτησίας) Αν οι X, Y είναι ανεξάρτητες συνεχείς Τ.Μ., τότε:

1. $E(XY) = E(X)E(Y)$.
2. $\text{Cov}(X, Y) = 0$.
3. $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$.
4. Για οποιοσδήποτε συναρτήσεις $g, h : \mathbb{R} \rightarrow \mathbb{R}$, οι τυχαίες μεταβλητές $g(X)$ και $h(Y)$ είναι ανεξάρτητες.

Πριν δώσουμε την απόδειξη παρατηρούμε πως, για την περίπτωση διακριτών Τ.Μ. X, Y : την Ιδιότητα 1 πιο πάνω την έχουμε συναντήσει ήδη ως μέρος της πρώτης ιδιότητας της συνδιακύμανσης στο Κεφάλαιο 9· η Ιδιότητα 2 αντιστοιχεί στην Ιδιότητα 2 της Ενότητας 9.2· η Ιδιότητα 3 είναι ακριβώς η πέμπτη ιδιότητα της συνδιακύμανσης στο Κεφάλαιο 9· και η Ιδιότητα 4 δίνεται στην Παρατήρηση 2 της Ενότητας 9.2. Επίσης σημειώνουμε ότι η τρίτη ιδιότητα έχει ήδη διατυπωθεί στη συνεχή περίπτωση, χωρίς απόδειξη, στο δεύτερο σκέλος του Θεωρήματος 11.5 στο Κεφάλαιο 11.

Απόδειξη: Ξεκινάμε αποδεικνύοντας την τέταρτη (και σημαντικότερη) ιδιότητα. Έστω δύο οποιοσδήποτε συναρτήσεις $g, h : \mathbb{R} \rightarrow \mathbb{R}$, και δύο αυθαίρετα υποσύνολα R_1, T_1 του \mathbb{R} . Αν ορίσουμε τα «αντίστροφα» σύνολα,

$$R_2 = g^{-1}(R_1) = \{x \in \mathbb{R} : g(x) \in R_1\} \quad \text{και} \quad T_2 = g^{-1}(T_1) = \{x \in \mathbb{R} : g(x) \in T_1\},$$

τότε απλά από τους ορισμούς έχουμε,

$$\Pr(g(X) \in R_1, h(Y) \in T_1) = \Pr(X \in g^{-1}(R_1), Y \in h^{-1}(T_1)) = \Pr(X \in R_2, Y \in T_2),$$

και από την ανεξαρτησία των X, Y ,

$$\Pr(g(X) \in R_1, h(Y) \in T_1) = \Pr(X \in R_2) \Pr(Y \in T_2) = \Pr(g(X) \in R_1) \Pr(h(Y) \in T_1).$$

Εφόσον η παραπάνω σχέση ισχύει για αυθαίρετα R_1, R_2 , οι $g(X)$ και $h(Y)$ είναι πράγματι ανεξάρτητες Τ.Μ.

Για την πρώτη ιδιότητα, αφού οι X, Y είναι ανεξάρτητες, το κριτήριο (15.7) μας λέει πως μπορούμε να θεωρήσουμε ότι η από κοινού πυκνότητά τους είναι της μορφής $f_{XY}(x, y) = f_X(x)f_Y(y)$. Οπότε βρίσκουμε ότι,

$$\begin{aligned} E(XY) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{XY}(x, y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_X(x) f_Y(y) dx dy \\ &= \left(\int_{-\infty}^{\infty} x f_X(x) dx \right) \left(\int_{-\infty}^{\infty} y f_Y(y) dy \right) = E(X)E(Y). \end{aligned}$$

Η δεύτερη ιδιότητα είναι άμεση συνέπεια της πρώτης σε συνδυασμό με το αποτέλεσμα του τρίτου σκέλους της Πρότασης 15.2, και η τρίτη ιδιότητα παρομοίως προκύπτει από το συνδυασμό της δεύτερης με το αποτέλεσμα του τέταρτου σκέλους της Πρότασης 15.2. ■

Παράδειγμα 15.7 Για τις Τ.Μ. X, Y του Παραδείγματος 15.3 υπολογίσαμε στο Παράδειγμα 15.6 πως έχουν μη μηδενική συνδιακύμανση, άρα, βάσει του δεύτερου σκέλους του Θεωρήματος 15.1, δεν είναι ανεξάρτητες. Εναλλακτικά, θα μπορούσαμε να καταφύγουμε απευθείας στον ορισμό της ανεξαρτησίας. Για παράδειγμα, χρησιμοποιώντας τις περιθώριες πυκνότητες των X, Y , που βρήκαμε στο Παράδειγμα 15.3, υπολογίζουμε τις πιθανότητες,

$$\begin{aligned}\Pr(0 \leq X \leq 1/2) &= \int_0^{1/2} \frac{2}{3}(x+1) dx = \frac{2}{3} \left[\frac{x^2}{2} + x \right]_0^{1/2} = \frac{5}{12}, \\ \Pr(0 \leq Y \leq 1/2) &= \int_0^{1/2} \frac{1}{3}(4y+1) dy = \frac{1}{3} [2y^2 + y]_0^{1/2} = \frac{1}{3},\end{aligned}$$

και παρατηρούμε πως,

$$\Pr(0 \leq X \leq 1/2) \Pr(0 \leq Y \leq 1/2) = \frac{5}{36} \neq \frac{1}{8} = \Pr(0 \leq X \leq 1/2, 0 \leq Y \leq 1/2),$$

όπου η τιμή της πιθανότητας $\Pr(0 \leq X \leq 1/2, 0 \leq Y \leq 1/2)$ έχει επίσης ήδη υπολογιστεί στο Παράδειγμα 15.3. Άρα, και από τον ορισμό της ανεξαρτησίας προκύπτει ότι οι X, Y δεν είναι ανεξάρτητες.

Για την περίπτωση των Τ.Μ. X, Y του Παραδείγματος 15.4, εύκολα διαπιστώνουμε πως το γινόμενο των περιθωρίων πυκνοτήτων τους είναι ίσο με την από κοινού πυκνότητά τους σε όλο το \mathbb{R}^2 , άρα, από το κριτήριο (15.7) έπεται πως είναι ανεξάρτητες.

Τα τελευταία δύο αποτελέσματα αυτού του κεφαλαίου μάς λένε με ποιον τρόπο μπορεί να υπολογιστεί η πυκνότητα του αθροίσματος δύο ανεξάρτητων Τ.Μ., και ότι το άθροισμα ανεξάρτητων, κανονικών Τ.Μ. έχει κι αυτό πάντα κανονική κατανομή.

Θεώρημα 15.2 (Άθροισμα ανεξάρτητων Τ.Μ.) Έστω δύο συνεχείς Τ.Μ. X, Y με πυκνότητες $f_X(x), f_Y(y)$ αντίστοιχα. Αν οι X, Y είναι ανεξάρτητες, τότε η πυκνότητα του αθροίσματός τους $Z = X + Y$ ισούται με:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx.$$

Παρατηρήσεις

1. Γενικά, για οποιεσδήποτε συναρτήσεις $f, g : \mathbb{R} \rightarrow \mathbb{R}$, η συνεχής συνέλιξη των $f(x), g(x)$ είναι η νέα συνάρτηση $h(x) = \int_{-\infty}^{\infty} f(t)g(x-t) dt$. Είναι εμφανής η συνάφεια της πιο πάνω έκφρασης με τη διακριτή μορφή της συνέλιξης, την οποία έχουμε ήδη συναντήσει στην Άσκηση 9 του Κεφαλαίου 6.

2. Πριν δούμε την απόδειξη του θεωρήματος, θα διατυπώσουμε μια σημαντική και χρήσιμη συνέπειά του, την οποία ήδη χρησιμοποιήσαμε στην απόδειξη του θεωρήματος του Lindeberg στο Κεφάλαιο 13.

Πόρισμα 15.1 (Άθροισμα ανεξάρτητων κανονικών Τ.Μ.) Αν οι Y_1, Y_2, \dots, Y_N είναι ανεξάρτητες Τ.Μ., όπου η κάθε $Z_i \sim N(a_i, \sigma_i^2)$, για $i = 1, 2, \dots, N$, τότε το άθροισμά τους έχει κι αυτό κανονική κατανομή, και συγκεκριμένα,

$$Z = \sum_{i=1}^N Y_i \sim N(a, \sigma^2), \quad \text{όπου } a = \sum_{i=1}^N a_i, \quad \sigma^2 = \sum_{i=1}^N \sigma_i^2.$$

Απόδειξη του Θεωρήματος 15.2: Για να βρούμε την πυκνότητα του Z , κατ' αρχάς παρατηρούμε πως η από κοινού πυκνότητα των X, Y μπορεί να εκφραστεί ως $f_{XY}(x, y) = f_X(x)f_Y(y)$ και εξετάζουμε τη συνάρτηση κατανομής $F_Z(z)$ του Z . Για δεδομένο $z \in \mathbb{R}$, αν ορίσουμε το τρίγωνο $R_z = \{(x, y) : x + y \leq z\} \subset \mathbb{R}^2$, τότε έχουμε,

$$F_Z(z) = \Pr(Z \leq z) = \Pr(X + Y \leq z) = \iint_{R_z} f_{XY}(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} f_X(x)f_Y(y) dy dx,$$

και, συμβολίζοντας $F_Y(y)$ τη συνάρτηση κατανομής του Y ,

$$F_Z(z) = \int_{-\infty}^{\infty} f_X(x) \left(\int_{-\infty}^{z-x} f_Y(y) dy \right) dx = \int_{-\infty}^{\infty} f_X(x)F_Y(z-x) dx.$$

Τέλος, παίρνοντας την παράγωγο ως προς z και στα δύο παραπάνω μέρη,

$$f_Z(z) = F'_Z(z) = \frac{d}{dz} \left(\int_{-\infty}^{\infty} f_X(x)F_Y(z-x) dx \right) = \int_{-\infty}^{\infty} f_X(x)F'_Y(z-x) dx,$$

το οποίο φυσικά ισούται με $\int_{-\infty}^{\infty} f_X(x)f_Y(z-x) dx$, ολοκληρώνοντας την απόδειξη. ■

Απόδειξη του Πορίσματος 15.1: Κατ' αρχάς παρατηρούμε πως η γενική περίπτωση μπορεί να αποδειχθεί με επαγωγική εφαρμογή του αποτελέσματος για $N = 2$ Τ.Μ., οπότε αρκεί να δείξουμε ότι, αν οι $X \sim N(a, \sigma^2)$ και $Y \sim N(b, \tau^2)$ είναι ανεξάρτητες, τότε η $Z = X + Y$ έχει κατανομή $N(a + b, \sigma^2 + \tau^2)$.

Ξεκινάμε με την ειδική περίπτωση της $W = X' + Y'$, όπου οι $X' \sim N(0, \phi^2)$, $Y' \sim N(0, 1)$ είναι ανεξάρτητες. Εφαρμόζοντας τον τύπο της συνέλιξης από το Θεώρημα 15.2 και αντικαθιστώντας τις αντίστοιχες πυκνότητες της κανονικής κατανομής, η πυκνότητα της W είναι,

$$f_W(w) = \int_{-\infty}^{\infty} f_{X'}(x)f_{Y'}(w-x)dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\phi^2}} \exp\left\{-\frac{1}{2\phi^2}x^2\right\} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(w-x)^2\right\} dx,$$

και θέτοντας $\rho^2 = \phi^2/(\phi^2 + 1)$ και απλοποιώντας βρίσκουμε,

$$\begin{aligned} f_W(w) &= \frac{1}{\sqrt{2\pi(\phi^2 + 1)}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\rho^2}} \exp \left\{ -\frac{1}{2\phi^2}x^2 - \frac{1}{2}(x-w)^2 \right\} dx \\ &= \frac{1}{\sqrt{2\pi(\phi^2 + 1)}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\rho^2}} \exp \left\{ -\frac{1}{2\rho^2} [x^2 - 2\rho^2xw + \rho^2w^2] \right\} dx \\ &= \frac{1}{\sqrt{2\pi(\phi^2 + 1)}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\rho^2}} \exp \left\{ -\frac{1}{2\rho^2} [x - \rho^2w]^2 \right\} \exp \left\{ -\frac{1}{2} [(1 - \rho^2)w^2] \right\} dx \\ &= \frac{1}{\sqrt{2\pi(\phi^2 + 1)}} \exp \left\{ -\frac{1}{2(\phi^2 + 1)}w^2 \right\} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\rho^2}} \exp \left\{ -\frac{1}{2\rho^2} [x - \rho^2w]^2 \right\} dx. \end{aligned}$$

Παρατηρούμε τώρα πως το τελευταίο από τα παραπάνω ολοκληρώματα είναι το ολοκλήρωμα της πυκνότητας της κατανομής $N(\rho^2w, \rho^2)$ σε όλο το \mathbb{R} , άρα ισούται με 1, και πως η έκφραση πριν το ολοκλήρωμα είναι ακριβώς η πυκνότητα της κατανομής $N(0, \phi^2 + 1)$. Άρα, έχουμε ότι $W \sim N(0, \phi^2 + 1)$.

Επιστρέφοντας τώρα στη γενική περίπτωση, έστω $X' \sim N(0, \phi^2)$ και $Y' \sim N(0, 1)$ ανεξάρτητες, με $\phi^2 = \sigma^2/\tau^2$. Αν ορίσουμε $X = \tau X' + a$ και $Y = \tau Y' + b$, τότε, από την απλή ιδιότητα των γραμμικών μετασχηματισμών για κανονικές Τ.Μ. (βλ. Άσκηση 7 του Κεφαλαίου 12), έχουμε $X \sim N(a, \sigma^2)$, $Y \sim N(b, \tau^2)$ και, από το Θεώρημα 15.1, οι X, Y είναι ανεξάρτητες. Επιπλέον, από την ειδική περίπτωση πιο πάνω, έχουμε ότι η $X' + Y' \sim N(0, \phi^2 + 1)$, αλλά από τους ορισμούς,

$$X' + Y' = \frac{X - a}{\tau} + \frac{Y - b}{\tau} = \frac{1}{\tau}(X + Y) - \frac{a + b}{\tau},$$

οπότε η Τ.Μ.,

$$\frac{1}{\tau}Z - \frac{a + b}{\tau} \sim N\left(0, \frac{\sigma^2}{\tau^2} + 1\right),$$

και εφαρμόζοντας και πάλι την ιδιότητα του γραμμικού μετασχηματισμού προκύπτει πως,

$$Z = \tau\left(\frac{Z}{\tau} - \frac{a + b}{\tau}\right) + (a + b) \sim N(a + b, \sigma^2 + \tau^2),$$

το οποίο είναι και το ζητούμενο αποτέλεσμα. ■

15.4 * Μετρησιμότητα και άπειρες τιμές

Στις Ενότητες 6.3 και 10.3 ήδη συζητήσαμε κάποιες μαθηματικές πτυχές των τεχνικής φύσεως προβλημάτων που προκύπτουν όταν ο χώρος πιθανότητας Ω στον οποίο ορίζονται οι Τ.Μ. που εξετάζουμε δεν είναι πεπερασμένος. Δεδομένου ότι οποτεδήποτε οι Τ.Μ. είναι συνεχείς, ο χώρος πιθανότητας στον οποίο θα οριστούν είναι απαραίτητα άπειρος και μη αριθμήσιμος, προκειμένου

να αποφευχθούν ενδεχόμενες παθολογίες, εισάγουμε τρεις απλούς περιορισμούς παρόμοιους με εκείνους που είδαμε στο Κεφάλαιο 10:

(1.) Τα μόνα σύνολα τιμών που επιτρέπουμε για συνεχείς Τ.Μ. είναι εκείνα που μπορούν να εκφραστούν ως ενώσεις ενός πεπερασμένου πλήθους διαστημάτων στο \mathbb{R} .

(2.) Κάθε N -άδα από συνεχείς Τ.Μ. X_1, X_2, \dots, X_N ορίζεται πάντοτε σε συνδυασμό με την από κοινού πυκνότητά τους, $f_{X_1 X_2 \dots X_N}(x_1, x_2, \dots, x_N)$.

(3.) Τα μόνα ενδεχόμενα των οποίων υπολογίζουμε τις πιθανότητες (ή τις δεσμευμένες πιθανότητες) είναι της μορφής $\{(X_1, X_2, \dots, X_N) \in R\}$ για σύνολα $R \subset \mathbb{R}^N$ τα οποία είτε (α') μπορούν να εκφραστούν ως ενώσεις ενός πεπερασμένου πλήθους ορθογωνίων της μορφής,

$$\{(x_1, x_2, \dots, x_N) \in \mathbb{R}^N : a_1 \leq x_1 \leq b_1, a_2 \leq x_2 \leq b_2, \dots, a_N \leq x_N \leq b_N\},$$

είτε (β') είναι χωρία $R \subseteq \mathbb{R}^N$ όπως αυτά που περιγράφονται στο Θεώρημα του Fubini. (Βλ. Θεώρημα Α'.2 του Παραρτήματος Α'.1.)

Παρότι οι πιο πάνω υποθέσεις είναι πιο περιοριστικές από όσο είναι απαραίτητο για να αναπτυχθεί η σχετική θεωρία, είναι εύκολο να διατυπωθούν και να ελεγχθούν στην πράξη και είναι αρκετά γενικές ώστε να συμπεριλαμβάνουν όλες τις σημαντικές εφαρμογές που μας ενδιαφέρουν. Συγκεκριμένα, κάτω από αυτές τις συνθήκες, για οποιαδήποτε δεδομένη από κοινού πυκνότητα $f_{X_1 X_2 \dots X_N}(x_1, x_2, \dots, x_N)$, είναι εύκολο να κατασκευαστεί ένας χώρος πιθανότητας Ω και να οριστούν συναρτήσεις $X_i : \Omega \rightarrow \mathbb{R}$, για $i = 1, 2, \dots, N$ και ένα μέτρο πιθανότητας \mathbb{P} στο Ω , έτσι ώστε οι X_1, X_2, \dots, X_N να έχουν από κοινού πυκνότητα την $f_{X_1 X_2 \dots X_N}$. Η βασική ιδέα είναι αντίστοιχη της κατασκευής που είδαμε στην Ενότητα 10.3, και δεν θα δώσουμε εδώ περαιτέρω λεπτομέρειες. Αναφέρουμε μόνο πως το δύσκολο και μαθηματικά ενδιαφέρον βήμα είναι η περιγραφή μιας αρκετά πλούσιας ομάδας υποσυνόλων \mathcal{F} του Ω στην οποία μπορεί να οριστεί το μέτρο πιθανότητας \mathbb{P} , και η απόδειξη του ότι το \mathbb{P} πράγματι έχει τις απαιτούμενες ιδιότητες.

Τέλος, για να αποφύγουμε τεχνικές λεπτομέρειες που ξεφεύγουν από τα ζητούμενα του παρόντος βιβλίου, υιοθετούμε τις ίδιες συμβατικές υποθέσεις όπως κάναμε στο Κεφάλαιο 10: Συνοπτικά, υποθέτουμε εμμέσως πως όλα τα ολοκληρώματα που εμφανίζονται σε υπολογισμούς πιθανοτήτων, μέσω τιμών και πυκνοτήτων υπάρχουν και η τιμή τους είναι πεπερασμένη.

15.5 Ασκήσεις

1. Παραδείγματα ανεξαρτησίας.

(α') Είναι ή όχι οι Τ.Μ. X, Y του Παραδείγματος 15.1 ανεξάρτητες;

(β') Είναι ή όχι οι Τ.Μ. X, Y του Παραδείγματος 15.2 ανεξάρτητες;

Αποδείξτε τις απαντήσεις σας.

2. Περισσότερα για την κατανομή $\chi^2(k)$. Στην Άσκηση 5 του Κεφαλαίου 12 ορίσαμε την κατανομή χ^2 με k βαθμούς ελευθερίας, $\chi^2(k)$.

(α') Υπολογίστε τη διασπορά της κατανομής $\chi^2(k)$.

(β') Βρείτε την πυκνότητα της $\chi^2(k)$ για $k = 1, 2$ και 3. Σχολιάστε.

3. Ασυσχετίστες αλλά όχι ανεξάρτητες συνεχείς Τ.Μ. Έστω πως οι Τ.Μ. (X, Y) είναι «ομοιόμορφα» κατανεμημένες στο δίσκο $S = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$, δηλαδή έχουν από κοινού πυκνότητα,

$$f_{XY}(x, y) = \begin{cases} \frac{1}{\pi}, & \text{αν } (x, y) \in S, \\ 0, & \text{αν } (x, y) \notin S. \end{cases}$$

(α') Βρείτε τις περιθώριες πυκνότητες των X και Y .

(β') Δείξτε πως η συνδιακύμανση $\text{Cov}(X, Y) = 0$, αλλά οι X, Y δεν είναι ανεξάρτητες.

4. Χρόνοι διεργασιών. Δύο διεργασίες που πραγματοποιούνται σε ένα δίκτυο επεξεργασιών είναι προγραμματισμένες έτσι ώστε να τελειώσουν ταυτόχρονα και σε μια συγκεκριμένη ώρα. Έστω X, Y οι χρόνοι καθυστέρησης του τερματισμού τους, σε δευτερόλεπτα, όπου οι Τ.Μ. X, Y είναι ανεξάρτητες και ομοιόμορφα κατανεμημένες από το 0 ως το 1.

(α') Ποια είναι η πιθανότητα η διαφορά στους χρόνους τερματισμού, δηλαδή το $|X - Y|$, να είναι λιγότερο από μισό δευτερόλεπτο;

(β') Ποια είναι η μέση τιμή της διαφοράς $|X - Y|$ των χρόνων τερματισμού;

(γ') Η καθυστέρηση στον τερματισμό των διεργασιών έχει κόστος $Z = 20X + 10Y$ ευρώ. Ποιο είναι το μέσο κόστος που προκύπτει κάθε φορά που εκτελούνται;

(δ') Ποια είναι η πιθανότητα το κόστος να ξεπερνά τα 20 ευρώ;

(ε') Υπολογίστε τη συνδιακύμανση $\text{Cov}(X, Y)$.

5. Θεραπείες. Σε κάποιον ασθενή χορηγούνται δύο θεραπείες, με ποσοστά επιτυχίας X και Y αντίστοιχα, για τα οποία γνωρίζουμε ότι η από κοινού πυκνότητα τους είναι:

$$f_{XY}(x, y) = \begin{cases} cx(1 - y), & (x, y) \in [0, 1] \times [0, 1], \\ 0, & (x, y) \notin [0, 1] \times [0, 1]. \end{cases}$$

- (α') Βρείτε την τιμή της σταθεράς c .
- (β') Ποιες είναι οι περιθώριες πυκνότητες των X, Y ;
- (γ') Υπολογίστε τη συνδιακύμανση $\text{Cov}(X, Y)$ και δείξτε αν οι Τ.Μ. X, Y είναι ανεξάρτητες ή όχι.
- (δ') Ποια είναι η πιθανότητα $\text{Pr}(X \leq Y)$ η δεύτερη θεραπεία να είναι αποτελεσματικότερη από την πρώτη;

6. *Εναλλακτικές θεραπείες.* Όπως στην Άσκηση 5, θεωρούμε πως σε κάποιον ασθενή χορηγούνται δύο θεραπείες με ποσοστά επιτυχίας X και Y αντίστοιχα, οι οποίες έχουν ελαφρώς διαφορετική πυκνότητα από πριν:

$$f_{XY}(x, y) = \begin{cases} cx(1-y), & 0 \leq x, y \leq 1, x+y \leq 1, \\ 0, & \text{αλλού.} \end{cases}$$

- (α') Ποια είναι η τιμή της σταθεράς c ;
- (β') Ποιες είναι οι περιθώριες πυκνότητες των X, Y ;
- (γ') Υπολογίστε τη συνδιακύμανση $\text{Cov}(X, Y)$ και δείξτε αν οι Τ.Μ. είναι X, Y ανεξάρτητες ή όχι.

7. *Τριπλά ολοκληρώματα.* Έστω πως τρεις συνεχείς Τ.Μ. X, Y, Z έχουν από κοινού πυκνότητα:

$$f_{XYZ}(x, y, z) = \begin{cases} \frac{1}{k}(2x + 3y^2 + 4z^3), & 0 \leq x \leq 3, 0 \leq y \leq 2, 0 \leq z \leq 1, \\ 0, & \text{αλλού.} \end{cases}$$

- (α') Βρείτε την τιμή της σταθεράς k .
- (β') Υπολογίστε τη μέση τιμή $E(XYZ)$.
- (γ') Υπολογίστε την πιθανότητα $\text{Pr}(X + Y + Z \leq 1)$.

8. *Περιθώρια και από κοινού πυκνότητα.* Έστω τρεις συνεχείς Τ.Μ. X, Y και Z με από κοινού πυκνότητα $f_{XYZ}(x, y, z)$. Βρείτε έναν τρόπο να εκφράσετε την από κοινού πυκνότητα $f_{XY}(x, y)$ μόνο δύο εκ των τριών Τ.Μ., συναρτήσει της $f_{XYZ}(x, y, z)$. Αποδείξτε ότι το αποτέλεσμά σας ισχύει.

9. *Μίξη πυκνοτήτων.* Έστω τρεις ανεξάρτητες Τ.Μ. X, Y, Z με κατανομές, $X \sim \text{Exp}(1)$, $Y \sim U[1, 2]$ και $Z \sim \text{Bern}(1/2)$ αντίστοιχα. Βρείτε την πυκνότητα της Τ.Μ.,

$$W = ZX + (1 - Z)Y,$$

και εξηγήστε διαισθητικά τι περιγράφει η W .

10. ★ Φτου φτου φτου. Η madame Depy Sisini – η διασημότερη επαγγελματίας ξεματιάστρα Καλλιθέας και περιχώρων – έχει διαπιστώσει ότι το επίπεδο ικανοποίησης του κάθε πελάτη της, έστω Z , εξαρτάται από τις τιμές δύο άλλων Τ.Μ.: Του χρόνου X (σε λεπτά) που διαρκούν τα χασμουρητά της ενώ τον ξεματιάζει, και του συνολικού χρόνου Y (επίσης σε λεπτά) που διαρκεί η διαδικασία του ξεματιάσματος.

Έστω πως οι X, Y, Z έχουν από κοινού πυκνότητα:

$$f_{XYZ}(x, y, z) = \begin{cases} \frac{x}{50\sqrt{2\pi}} \exp \left\{ -\frac{x}{2} - \frac{(y-z)^2}{50} \right\}, & \text{αν } x > 0 \text{ και } 5/x \leq y \leq 10/x, \\ 0, & \text{αν όχι.} \end{cases}$$

- (α') Βρείτε την περιθώρια πυκνότητα και τη μέση τιμή του χρόνου X των χασμουρητών.
 (β') Δεδομένου ότι σε κάποιο ξεμάτιασμα η madame Depy χασμουρήθηκε για λιγότερο από ένα λεπτό, ποια η πιθανότητα το όλο ξεμάτιασμα να διήρκεσε το πολύ 10 λεπτά;

11. Συσχέτιση, συνδιακύμανση, διασπορά. [Εδώ θα δούμε τις προφανείς επεκτάσεις κάποιων αποτελεσμάτων που αποδείχθηκαν για διακριτές Τ.Μ. στην Άσκηση 10 του Κεφαλαίου 9 και στην Άσκηση 17 του Κεφαλαίου 11.] Για δύο συνεχείς Τ.Μ. X, Y , ο συντελεστής συσχέτισης μεταξύ τους ορίζεται ακριβώς όπως και στη διακριτή περίπτωση:

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{[\text{Var}(X)\text{Var}(Y)]^{\frac{1}{2}}}.$$

- (α') Αποδείξτε ότι, αν $Y = aX + b$, όπου η σταθερά $a \neq 0$, τότε το $\rho_{X,Y}$ ισούται με $+1$ ή με -1 . Πότε ισούται με $+1$ και πότε με -1 ;
 (β') Δείξτε ότι για τη συνδιακύμανση μεταξύ των X και Y έχουμε:

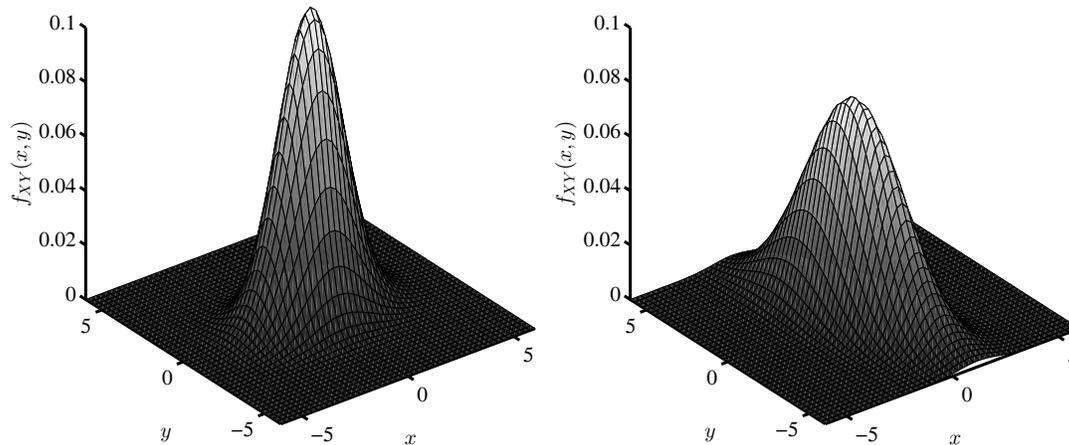
$$\text{Cov}(X, Y) \leq \sqrt{\text{Var}(X)\text{Var}(Y)}.$$

- (γ') Αποδείξτε πως ο συντελεστής συσχέτισης $\rho_{X,Y}$ πάντα ικανοποιεί $|\rho_{X,Y}| \leq 1$.

12. ★ Η διμεταβλητή κανονική κατανομή. Δύο Τ.Μ. X, Y λέμε πως έχουν διμεταβλητή κανονική κατανομή με αντίστοιχους μέσους μ, ν , διασπορές $\sigma^2 > 0, \tau^2 > 0$, και συνδιακύμανση c , αν έχουν από κοινού πυκνότητα,

$$\begin{aligned} f_{XY}(x, y) &= \frac{1}{2\pi\sigma\tau\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[\frac{(x-\mu)^2}{\sigma^2} + \frac{(y-\nu)^2}{\tau^2} - \frac{2\rho(x-\mu)(y-\nu)}{\sigma\tau} \right] \right\}, \end{aligned} \quad (15.8)$$

για κάθε $x, y \in \mathbb{R}$, όπου $\rho = c/(\sigma\tau)$, και η συνδιακύμανσή τους ικανοποιεί $|c| < \sigma\tau$. Δύο παραδείγματα αυτής της πυκνότητας για διαφορετικές τιμές των παραμέτρων απεικονίζονται στο Σχήμα 15.7.



Σχήμα 15.7: Παραδείγματα της διμεταβλητής κανονικής πυκνότητας (15.8) με μέσους $\mu = \nu = 0$. Στην πρώτη περίπτωση οι διασπορές είναι $\sigma^2 = 1$, $\tau^2 = 2$ και η συνδιακύμανση $c = 0.1$ (αριστερά) και στην δεύτερη περίπτωση έχουμε $\sigma^2 = 1$, $\tau^2 = 5$, $c = -1$ (δεξιά).

Εδώ θα δούμε έναν απλό τρόπο για να κατασκευάσουμε δύο Τ.Μ. X, Y με τις πιο πάνω ιδιότητες. Έστω Z, W δύο ανεξάρτητες Τ.Μ. με τυπική κανονική κατανομή $N(0, 1)$. Ορίζουμε τις δύο νέες Τ.Μ.,

$$X = \mu + \sigma Z, \quad Y = \nu + \frac{c}{\sigma} Z + \sqrt{\tau^2 - \frac{c^2}{\sigma^2}} W.$$

(α') Δείξτε πως οι X, Y έχουν τις επιθυμητές περιθώριες κατανομές, $X \sim N(\mu, \sigma^2)$ και $Y \sim N(\nu, \tau^2)$.

(β') Υπολογίστε πως η συνδιακύμανσή τους $\text{Cov}(X, Y)$ πράγματι ισούται με c .

(γ') Εξηγήστε γιατί πάντοτε έχουμε $c^2 \leq \sigma^2 \tau^2$ και σχολιάστε τον περιορισμό $|c| < \sigma \tau$.

(δ') Αποδείξτε ότι οι X, Y έχουν από κοινού πυκνότητα την παραπάνω $f_{XY}(x, y)$.

13. Το Κ.Ο.Θ. χωρίς το «Ο». Έστω πως οι Τ.Μ. $\{X_n\}$ στο Κ.Ο.Θ. (Θεώρημα 12.2) έχουν όλες κατανομή $X_n \sim N(\mu, \sigma^2)$. Εξηγήστε γιατί σε αυτή την ειδική περίπτωση δεν χρειάζεται καν να πάρουμε το όριο $N \rightarrow \infty$ για να δείξουμε ότι τα κανονικοποιημένα αθροίσματα \bar{S}_N συγκλίνουν κατά κατανομή στην τυπική κανονική κατανομή.

Παράρτημα Α΄

Α΄.1 Διπλά ολοκληρώματα: Συνοπτική επισκόπηση

Για μια οποιαδήποτε συνάρτηση $f(x, y)$ δύο μεταβλητών, το ολοκλήρωμα της $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ στο υποσύνολο $R \subseteq \mathbb{R} \times \mathbb{R}$, συμβολίζεται ως,

$$\iint_R f(x, y) dA, \quad \text{ή πιο απλά,} \quad \iint_R f(x, y) dx dy.$$

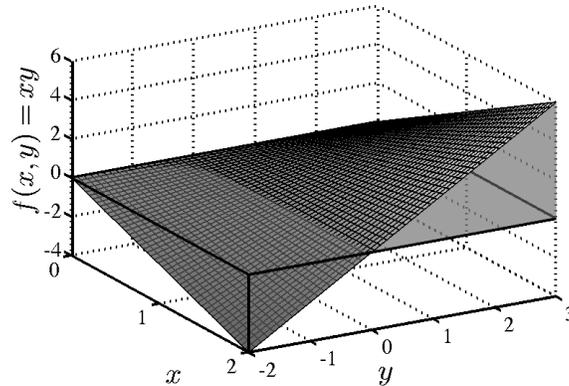
Αν η $f(x, y)$ παίρνει θετικές τιμές, η γεωμετρική ερμηνεία του διπλού ολοκληρώματος είναι πως η τιμή του ισούται με τον όγκο του στερεού που περικλείεται μεταξύ του συνόλου R στο επίπεδο x - y και του γραφήματος $z = f(x, y)$. Γενικά, η τιμή του διπλού ολοκληρώματος ισούται με τον όγκο του στερεού μεταξύ του R και της $f(x, y)$ όπου αυτή είναι θετική, μείον τον όγκο μεταξύ του R και της $f(x, y)$ όπου αυτή είναι αρνητική. Για παράδειγμα, το διπλό ολοκλήρωμα της συνάρτησης $f(x, y) = xy$ στο $R = [0, 2] \times [-2, 3]$ είναι ο όγκος του στερεού που βρίσκεται μεταξύ του $R_1 = [0, 2] \times [0, 3]$, όπου η συνάρτηση είναι θετική, και του γραφήματος της $f(x, y)$, μείον τον όγκο του στερεού που βρίσκεται μεταξύ του $R_2 = [0, 2] \times [-2, 0]$, όπου η συνάρτηση είναι αρνητική, και του γραφήματος της $f(x, y)$, όπως απεικονίζεται στο Σχήμα Α΄.1.

Στο ακόλουθο θεώρημα συγκεντρώνουμε κάποιες από τις απλούστερες ιδιότητες των διπλών ολοκληρωμάτων. Επίσης σημειώνουμε πως, για να αποφύγουμε τεχνικές δυσκολίες που ξεπερνούν τους στόχους αυτού του βιβλίου, εμμέσως υποθέτουμε πως υπάρχουν όλα τα ολοκληρώματα που εμφανίζονται.

Θεώρημα Α΄.1 Για κάθε ζεύγος συναρτήσεων f και g στο $\mathbb{R} \times \mathbb{R}$, για οποιεσδήποτε σταθερές $a, b \in \mathbb{R}$, και για οποιαδήποτε υποσύνολα $R, S \subseteq \mathbb{R} \times \mathbb{R}$, έχουμε:

1. Το ολοκλήρωμα είναι γραμμικό:

$$\iint_R (af(x, y) + bg(x, y)) dx dy = a \iint_R f(x, y) dx dy + b \iint_R g(x, y) dx dy.$$



Σχήμα Α΄.1: Το διπλό ολοκλήρωμα της $f(x, y) = xy$ ισούται με τον όγκο του στερεού πάνω από το ορθογώνιο $[0, 2] \times [0, 3]$ μείον τον όγκο του στερεού κάτω από το ορθογώνιο $[0, 2] \times [-2, 0]$.

2. Αν $f(x, y) \leq g(x, y)$ για κάθε $(x, y) \in R$, τότε:

$$\iint_R f(x, y) dx dy \leq \iint_R g(x, y) dx dy.$$

3. Αν $f(x, y) \geq 0$ για κάθε $(x, y) \in S$ και το R είναι υποσύνολο του S , τότε:

$$\iint_R f(x, y) dx dy \leq \iint_S f(x, y) dx dy.$$

4. Αν τα R και S είναι ξένα, τότε για οποιαδήποτε συνάρτηση $f(x, y)$:

$$\iint_{R \cup S} f(x, y) dx dy = \iint_R f(x, y) dx dy + \iint_S f(x, y) dx dy.$$

5. Για οποιαδήποτε $R \subseteq \mathbb{R} \times \mathbb{R}$: $\iint_R 0 dx dy = 0$.

6. Αν το $R \subseteq \mathbb{R} \times \mathbb{R}$ έχει εμβαδόν ίσο με μηδέν, τότε για οποιαδήποτε συνάρτηση $f(x, y)$:

$$\iint_R f(x, y) dx dy = 0.$$

Το επόμενο αποτέλεσμα αποτελεί ένα πολύ χρήσιμο και βασικό εργαλείο για τον υπολογισμό διπλών ολοκληρωμάτων. Αμέσως μετά θα δούμε δύο παραδείγματα όπου εφαρμόζονται οι πιο πάνω ιδιότητες και το θεώρημα του Fubini.

Θεώρημα Α'.2 (Θεώρημα του Fubini) 1. Έστω ένα υποσύνολο $R \subseteq \mathbb{R} \times \mathbb{R}$ στο επίπεδο, το οποίο μπορεί να εκφραστεί ως,

$$R = \{(x, y) : a \leq x \leq b, \phi_1(x) \leq y \leq \phi_2(x)\}, \quad (\text{A'.1})$$

για κάποιες σταθερές $a < b$ και κάποιες συνεχείς συναρτήσεις $\phi_1, \phi_2 : [a, b] \rightarrow \mathbb{R}$ τέτοιες ώστε $\phi_1(x) \leq \phi_2(x)$ για κάθε $x \in [a, b]$. (Ένα τέτοιο R απεικονίζεται στο Σχήμα Α'.2.) Αν η $f(x, y)$ είναι συνεχής στο R , τότε το διπλό ολοκλήρωμά της στο R υπάρχει και μπορεί να υπολογιστεί ως,

$$\iint_R f(x, y) dx dy = \int_a^b \left(\int_{\phi_1(x)}^{\phi_2(x)} f(x, y) dy \right) dx.$$

2. Παρομοίως, έστω ένα υποσύνολο $R \subseteq \mathbb{R} \times \mathbb{R}$ στο επίπεδο, το οποίο μπορεί να εκφραστεί ως,

$$R = \{(x, y) : a \leq y \leq b, \phi_1(y) \leq x \leq \phi_2(y)\}, \quad (\text{A'.2})$$

για κάποιες σταθερές $a < b$ και κάποιες συνεχείς συναρτήσεις $\phi_1, \phi_2 : [a, b] \rightarrow \mathbb{R}$ τέτοιες ώστε $\phi_1(y) \leq \phi_2(y)$ για κάθε $y \in [a, b]$. (Ένα τέτοιο R απεικονίζεται στο Σχήμα Α'.2.) Αν η $f(x, y)$ είναι συνεχής στο R , τότε το διπλό ολοκλήρωμά της στο R υπάρχει και μπορεί να υπολογιστεί ως,

$$\iint_R f(x, y) dx dy = \int_a^b \left(\int_{\phi_1(y)}^{\phi_2(y)} f(x, y) dx \right) dy.$$

3. Στην ειδική περίπτωση που το R είναι ένα ορθογώνιο, $R = [a, b] \times [c, d]$, έχουμε:

$$\iint_{[a,b] \times [c,d]} f(x, y) dx dy = \int_a^b \left(\int_c^d f(x, y) dy \right) dx = \int_c^d \left(\int_a^b f(x, y) dx \right) dy.$$

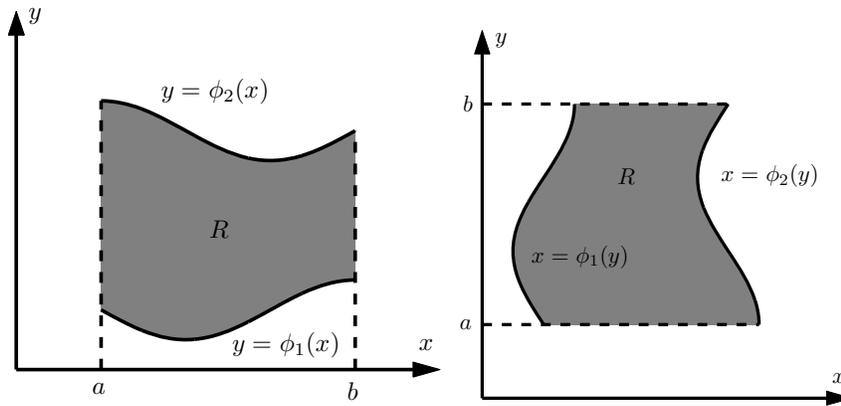
Παράδειγμα Α'.1 Θα υπολογίσουμε το διπλό ολοκλήρωμα της συνάρτησης $f(x, y) = y^2 \sqrt{x}$, $x \geq 0, y \in \mathbb{R}$, σε τρία διαφορετικά χωρία R .

1. Το ορθογώνιο R_1 με γωνίες τα σημεία $(0, 0)$, $(2, 0)$, $(0, 1)$, και $(2, 1)$.

2. Το τρίγωνο R_2 με γωνίες τα σημεία $(0, 0)$, $(0, 1)$, και $(2, 1)$.

3. Το χωρίο R_3 που περικλείεται από τις ευθείες $x = 0$, $y = 1$, και την καμπύλη $y = \sqrt{x/2}$.

Σε κάθε περίπτωση, το πιο σημαντικό βήμα είναι η έκφραση του αντίστοιχου R_i σε μία από τις μορφές (Α'.1) ή (Α'.2), όπως απαιτεί η εφαρμογή του Θεωρήματος Α'.2. Έχουμε:



Σχήμα Α'.2: Η πρώτη (αριστερά) και η δεύτερη (δεξιά) περίπτωση του θεωρήματος του Fubini.

1. Το R_1 μπορεί να γραφεί ως $R_1 = \{(x, y) : 0 \leq x \leq 2, 0 \leq y \leq 1\}$, άρα,

$$\begin{aligned}
 \iint_{R_1} f(x, y) dx dy &= \int_0^2 \left(\int_0^1 y^2 \sqrt{x} dy \right) dx \\
 &= \int_0^2 \sqrt{x} \left[\frac{y^3}{3} \right]_0^1 dy dx \\
 &= \int_0^2 \frac{\sqrt{x}}{3} dx \\
 &= \frac{2}{9} \left[x^{3/2} \right]_0^2 = \frac{4\sqrt{2}}{9}.
 \end{aligned}$$

Το R_1 και ο όγκος του στερεού που μόλις υπολογίσαμε απεικονίζονται στο Σχήμα Α'.3.

2. Για το R_2 , παρατηρούμε πως μπορεί να εκφραστεί ως $\{(x, y) : 0 \leq x \leq 2, x/2 \leq y \leq 1\}$, οπότε,

$$\begin{aligned}
 \iint_{R_2} f(x, y) dx dy &= \int_0^2 \left(\int_{x/2}^1 y^2 \sqrt{x} dy \right) dx \\
 &= \int_0^2 \sqrt{x} \left[\frac{y^3}{3} \right]_{x/2}^1 dy dx \\
 &= \int_0^2 \sqrt{x} \left(\frac{1}{3} - \frac{x^3}{24} \right) dx \\
 &= \left[\frac{2}{9} x^{3/2} - \frac{1}{108} x^{9/2} \right]_0^2 = \frac{8\sqrt{2}}{27}.
 \end{aligned}$$

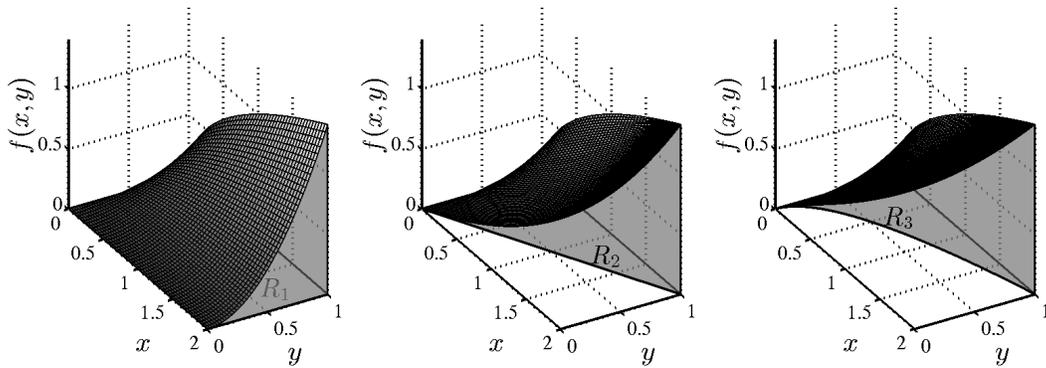
Όπως πριν, το R_2 και ο όγκος του στερεού που μόλις υπολογίσαμε απεικονίζονται στο Σχήμα A.3.

3. Στην τελευταία περίπτωση, όπως φαίνεται και στο Σχήμα A.3, έχουμε,

$$R_3 = \left\{ (x, y) : 0 \leq x \leq 2, \sqrt{\frac{x}{2}} \leq y \leq 1 \right\},$$

και εφαρμόζοντας και πάλι το θεώρημα του Fubini,

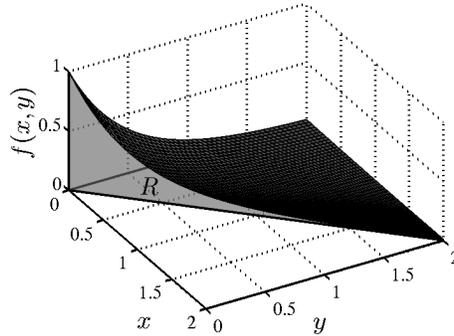
$$\begin{aligned} \iint_{R_3} f(x, y) \, dx \, dy &= \int_0^2 \left(\int_{\sqrt{x/2}}^1 y^2 \sqrt{x} \, dy \right) dx \\ &= \int_0^2 \sqrt{x} \left[\frac{y^3}{3} \right]_{\sqrt{x/2}}^1 dx \\ &= \int_0^2 \left(\frac{\sqrt{x}}{3} - \frac{1}{6\sqrt{2}} x^2 \right) dx \\ &= \left[\frac{2}{9} x^{3/2} - \frac{1}{18\sqrt{2}} x^3 \right]_0^2 = \frac{2\sqrt{2}}{9}. \end{aligned}$$



Σχήμα A.3: Τα στερεά των ολοκληρωμάτων στο Παράδειγμα A.1.

Το Θεώρημα A.2 μπορεί να εφαρμοστεί ακόμη και αν το διπλό ολοκλήρωμα είναι καταχρηστικό, δηλαδή σε περιπτώσεις που είτε το R δεν είναι πεπερασμένο, είτε η συνάρτηση $f(x, y)$ δεν είναι φραγμένη. Ακολουθεί ένα τέτοιο παράδειγμα.

Παράδειγμα A.2 Θα υπολογίσουμε το ολοκλήρωμα της $f(x, y) = e^{-x-2y}$, $x, y \in \mathbb{R}$ στο τρίγωνο R που περικλείεται μεταξύ των ευθειών $x = 0$ και $y = x$. Το R και το αντίστοιχο στερεό που δημιουργείται έχουν σχεδιαστεί στο Σχήμα A.4.



Σχήμα Α΄.4: Το τρίγωνο R και το αντίστοιχο στερεό που προκύπτουν στο Παράδειγμα Α΄.2.

Αρχικά παρατηρούμε πως,

$$R = \{(x, y) : 0 \leq y \leq \infty, 0 \leq x \leq y\},$$

και, χρησιμοποιώντας τη δεύτερη περίπτωση του θεωρήματος του Fubini, έχουμε,

$$\begin{aligned} \iint_R f(x, y) dx dy &= \int_0^\infty \left(\int_0^y e^{-x-2y} dx \right) dy \\ &= \int_0^\infty e^{-2y} \left(\int_0^y e^{-x} dx \right) dy \\ &= \int_0^\infty e^{-2y} [-e^{-x}]_0^y dy \\ &= \int_0^\infty e^{-2y} (1 - e^{-y}) dy \\ &= \left[\frac{1}{3} e^{-3y} - \frac{1}{2} e^{-2y} \right]_0^\infty = \frac{1}{6}. \end{aligned}$$

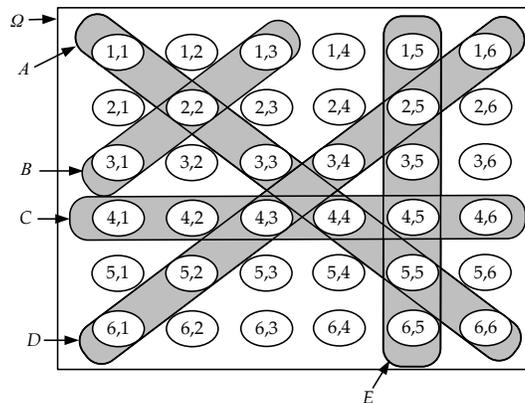
A'.2 Ασκήσεις Κεφαλαίου 2

1. *Τυχαία παιδιά.* Ο χώρος πιθανότητας είναι το σύνολο Ω που αποτελείται από όλες τις δυνατές n -άδες της μορφής (X, X, \dots, X) , όπου $X = A$ ή $X = K$. Άρα περιέχει 2^n τέτοια στοιχεία. Π.χ., το στοιχείο $\omega = (K, A, A, \dots, A)$ περιγράφει την περίπτωση όπου το ζευγάρι έκανε πρώτα ένα κορίτσι και μετά $(n - 1)$ αγόρια.
2. *Κι άλλα τυχαία παιδιά.* Ο χώρος πιθανότητας Ω εδώ περιέχει δύο ειδών στοιχεία: Εκείνα που αντιστοιχούν σε αποτελέσματα πεπερασμένου μήκους $(K), (A, K), (A, A, K), (A, A, A, K)$ κ.ο.κ., και ένα που αντιστοιχεί στην περίπτωση όπου το ζευγάρι κάνει άπειρα αγόρια, το (A, A, A, \dots) . Δηλαδή,

$$\Omega = \{(K), (A, K), (A, A, K), \dots, (A, \dots, A, K), \dots\} \cup \{(A, A, A, \dots)\}.$$

Μια εναλλακτική, ισοδύναμη περιγραφή του αποτελέσματος αυτού του πειράματος μπορεί να δοθεί με το να καταγράψουμε απλώς το πόσα παιδιά συνολικά έκανε το ζευγάρι. Τότε, ο αντίστοιχος χώρος πιθανότητας θα μπορούσε να οριστεί ως $\Omega' = \{1, 2, 3, \dots\} \cup \{\infty\}$. Η αντιστοιχία των δύο χώρων πιθανότητας Ω και Ω' είναι προφανής.

3. *Δύο διαδοχικές ζαριές.* Ο χώρος πιθανότητας Ω και όλα τα ζητούμενα ενδεχόμενα έχουν σχεδιαστεί στο Σχήμα A'.5.

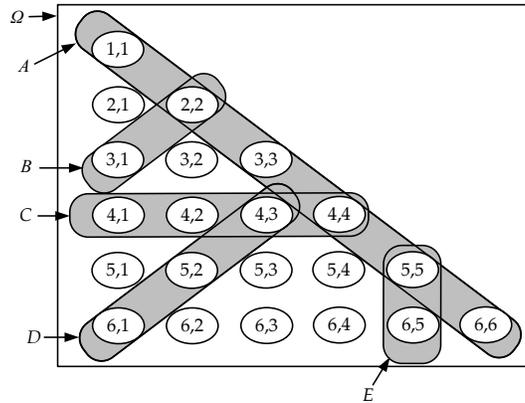


Σχήμα A'.5: Λύση Άσκησης 3.

4. *Υπάρχουν και περίεργοι χώροι πιθανότητας.* Εδώ ο χώρος πιθανότητας Ω μπορεί να περιγραφεί ως το ακόλουθο σύνολο:

$$\Omega = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 20^2\} \cup \{\text{ΣΚΥΛΟΣ}\}.$$

5. Δύο ταυτόχρονες ζαριές. Ο χώρος πιθανότητας Ω και όλα τα ζητούμενα ενδεχόμενα εμφανίζονται στο Σχήμα Α΄.6. Σε αυτή την περίπτωση, οι αριθμοί που προκύπτουν δεν είναι διατεταγμένες δυάδες, και συνεπώς δεν έχει νόημα να υπάρχουν στο χώρο πιθανότητας ταυτόχρονα οι δυάδες (1, 2) και (2, 1), (1, 3) και (3, 1) κ.ο.κ.



Σχήμα Α΄.6: Λύση Άσκησης 5.

6. Τρία νομίσματα. Έστω ΚΚΚ το αποτέλεσμα να έρθουν τρεις Κορώνες, ΚΚΓ να έρθουν δύο Κορώνες και κατόπιν Γράμματα κλπ. Ο χώρος πιθανότητας είναι το,

$$\Omega = \{ΚΚΚ, ΚΚΓ, ΚΓΚ, ΚΓΓ, ΓΚΚ, ΓΚΓ, ΓΓΚ, ΓΓΓ\},$$

και τα ζητούμενα ενδεχόμενα είναι τα,

$$A = \{ΚΚΚ, ΓΓΓ\}, B = \{ΓΓΓ, ΓΓΚ\}, C = \{ΚΚΓ, ΚΓΚ, ΓΚΚ, ΚΚΚ\}.$$

7. Άσπρες και μαύρες μπάλες.

(α΄) Ο χώρος πιθανότητας του πειράματος είναι ο $\Omega_1 = \{AM, MA, MM\}$. Εφόσον υποθέσαμε ότι οι μαύρες είναι πανομοιότυπες, δεν μπορούμε να διακρίνουμε μεταξύ τους. Αν μπορούσαμε να τις διακρίνουμε, τότε ο χώρος πιθανότητας θα ήταν το σύνολο:

$$\{AM_1, AM_2, AM_3, M_1A, M_2A, M_3A, M_1M_2, M_1M_3, M_2M_1, M_2M_3, M_3M_1, M_3M_2\}.$$

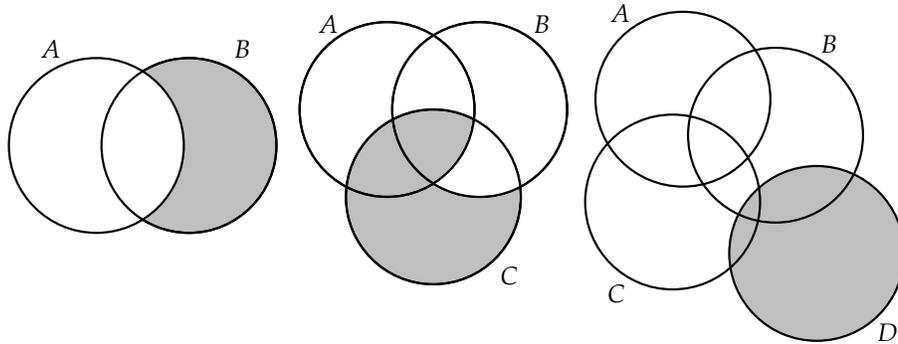
(β΄) Το ενδεχόμενο να κερδίσουμε συνολικά 10 ευρώ είναι προφανώς το $A = \{MM\}$.

(γ΄) Ο νέος χώρος πιθανότητας είναι ο,

$$\Omega_2 = \{AM, MA, AA, MM\},$$

και το ζητούμενο ενδεχόμενο είναι το $B = \{AM, MA\}$.

8. *Λειτουργία δικτύου.* Θυμηθείτε την Παρατήρηση 2 του Κεφαλαίου 2. Κατά συνέπεια, έχουμε: $D = A \cap B$, $E = A \cap B'$, $F = A \cap B \cap C'$, $G = (A \cap B') \cup (A \cap C) \cup A'$.
9. *Απλά διαγράμματα ενδεχομένων.* Τα ζητούμενα ενδεχόμενα εμφανίζονται σκιασμένα στο Σχήμα A'.7.



Σχήμα A'.7: Λύση Άσκησης 9.

10. *Τρεις ζαριές.* Ο χώρος πιθανότητας είναι το σύνολο που περιέχει όλες τις τριάδες που αποτελούνται από τα στοιχεία 1,2,3,4,5,6:

$$\Omega = \{111, 112, 113, 114, \dots, 666\}.$$

Τα ζητούμενα ενδεχόμενα είναι τα:

$$A = \{616, 626, 636, 646, 656, 666\},$$

$$B = \{111, 122, 133, 144, 155, 166\},$$

$$C = \{222, 444, 666\}.$$

11. *Σταθερά και κινητά τηλέφωνα.* Για το χώρο πιθανότητας έχουμε,

$$\Omega = \{\text{όλα τα ζεύγη διαφορετικών αριθμών από τα } 1, 2, 3, \dots, 450\}.$$

Για το πρώτο ζητούμενο ενδεχόμενο,

$$A = \text{«επιλέξαμε ένα σταθερό και ένα κινητό»}$$

$$= \{(1, 401), (1, 402), \dots, (1, 450),$$

$$(2, 401), (2, 402), \dots, (2, 450),$$

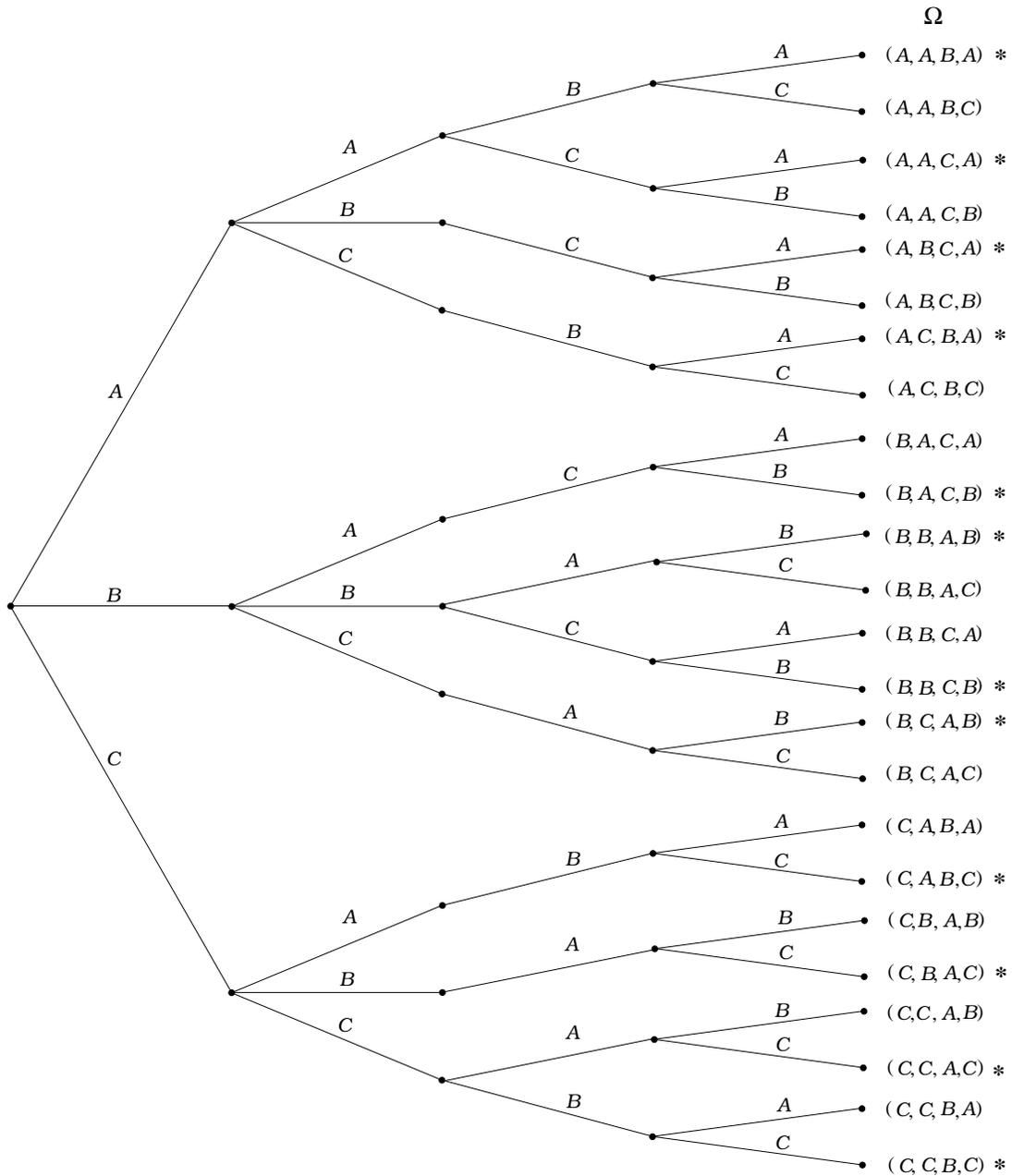
$$\vdots \quad \vdots \quad \vdots$$

$$(400, 401), \dots, (400, 450)\},$$

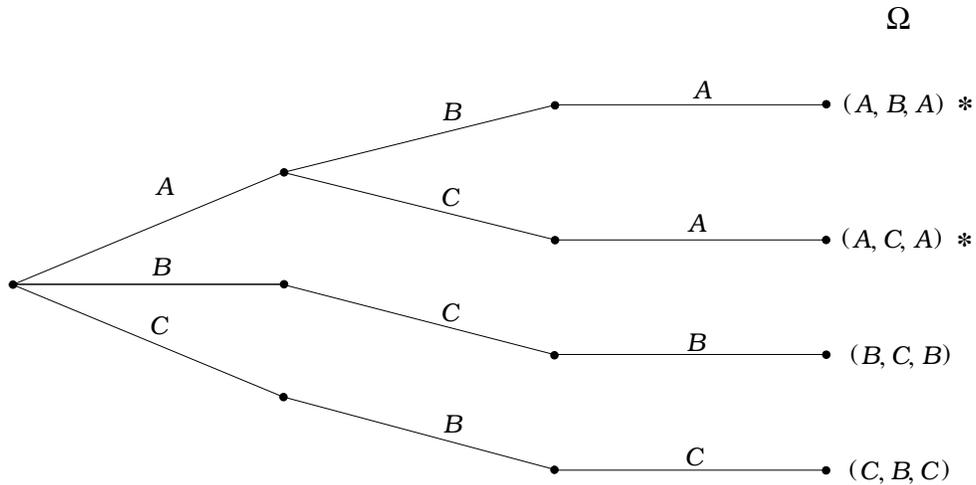
δηλαδή το A αποτελείται από όλα τα μη διατεταγμένα ζεύγη ακεραίων (m, n) που έχουν $1 \leq m \leq 400$ και $401 \leq n \leq 450$.

Τέλος, $B = \emptyset$, γιατί δεν υπάρχει συνδυασμός τηλεφώνων που να δίνει αυτό το κόστος.

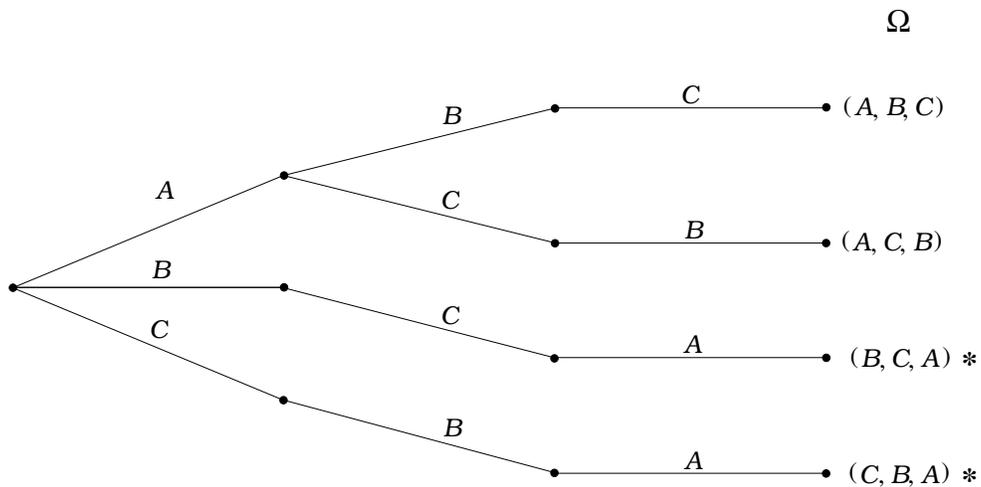
12. *Το πρόβλημα των τριών φυλακισμένων.* Το πρόβλημα αυτό είναι εντελώς ανάλογο του προβλήματος Monty Hall στο Παράδειγμα 2.4, όπου στη θέση του δώρου έχουμε την απονομή χάριτος, στη θέση του διαγωνιζόμενου τον φυλακισμένο, και στη θέση του παρουσιαστή τον δεσμοφύλακα. Μπορούμε και πάλι να περιγράψουμε το χώρο πιθανότητας με χρήση του Σχήματος 2.4, εφόσον ορίσουμε ως A τον φυλακισμένο που θα ελευθερωθεί και ως B, C τους άλλους δύο. Εδώ, το πρώτο στοιχείο της τριάδας είναι το όνομα του φυλακισμένου που ρωτά, το δεύτερο στοιχείο της τριάδας είναι το όνομα του φυλακισμένου που ο δεσμοφύλακας αποκαλύπτει ότι δεν θα πάρει χάρη, και το τρίτο στοιχείο είναι το όνομα του φυλακισμένου του οποίου θα πάρει τη θέση ο φυλακισμένος που ρωτά.
13. *Monty Hall 2.* Ο νέος χώρος πιθανότητας και τα αποτελέσματα που οδηγούν σε νίκη εμφανίζονται στο Σχήμα Α΄.8. Παρατηρήστε ότι ο νέος χώρος πιθανότητας έχει τη μορφή δέντρου με τρεις αρχικούς κλάδους, ο καθένας εν των οποίων είναι ακριβώς ίδιος με το χώρο πιθανότητας του Παραδείγματος 2.4.
14. *Monty Hall 3.* Ο νέος χώρος πιθανότητας και τα αποτελέσματα που οδηγούν σε νίκη εμφανίζονται στο Σχήμα Α΄.9. Παρατηρήστε ότι ο νέος χώρος πιθανότητας έχει μόλις 4 στοιχεία.
15. *Monty Hall 4.* Ο νέος χώρος πιθανότητας και τα αποτελέσματα που οδηγούν σε νίκη εμφανίζονται στο Σχήμα Α΄.10. Παρατηρήστε ότι και αυτός ο χώρος πιθανότητας έχει μόνο 4 στοιχεία.



Σχήμα Α'.8: Ο χώρος πιθανότητας της Άσκησης 13. Στο πρώτο βήμα επιλέγεται η κουρτίνα στην οποία θα τοποθετηθεί το δώρο, στο δεύτερο ο διαγωνιζόμενος επιλέγει μια κουρτίνα, στο τρίτο ο παρουσιαστής ανοίγει μία από τις άλλες δύο, και στο τέταρτο ο διαγωνιζόμενος αλλάζει αν θέλει την επιλογή του. Τα αποτελέσματα που αντιστοιχούν σε «νίκη» είναι σημειωμένα με *.



Σχήμα Α΄.9: Ο χώρος πιθανότητας της Άσκησης 14. Στο πρώτο βήμα ο διαγωνιζόμενος επιλέγει μια κουρτίνα, στο δεύτερο ο παρουσιαστής ανοίγει μία από τις άλλες δύο, και στο τρίτο ο διαγωνιζόμενος διατηρεί την αρχική επιλογή του. Τα αποτελέσματα που αντιστοιχούν σε «νίκη» είναι σημειωμένα με *.



Σχήμα Α΄.10: Ο χώρος πιθανότητας της Άσκησης 15. Στο πρώτο βήμα ο διαγωνιζόμενος επιλέγει μια κουρτίνα, στο δεύτερο ο παρουσιαστής ανοίγει μία από τις άλλες δύο, και στο τρίτο ο διαγωνιζόμενος αλλάζει την αρχική επιλογή του. Τα αποτελέσματα που αντιστοιχούν σε «νίκη» είναι σημειωμένα με *.

Α'.3 Ασκήσεις Κεφαλαίου 3

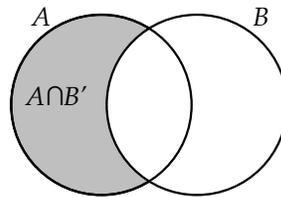
1. Η πιθανότητα της διαφοράς. Έστω Ω ο χώρος πιθανότητας όπου ορίζονται τα ενδεχόμενα E και F . Έχουμε κατ' αρχάς ότι,

$$E = E \cap \Omega = E \cap (F \cup F') = (E \cap F) \cup (E \cap F').$$

Όμως τα ενδεχόμενα $E \cap F$ και $E \cap F'$ είναι ξένα, συνεπώς, από τον κανόνα πιθανότητας #3,

$$\Pr(E) = \Pr(E \cap F) + \Pr(E \cap F'),$$

που μας δίνει ακριβώς τη ζητούμενη σχέση. Δείτε το διάγραμμα του Σχήματος Α'.11.



Σχήμα Α'.11: Διάγραμμα για την Άσκηση 1.

2. Περίεργα ζάρια. Έστω ο χώρος πιθανότητας $\Omega = \{1, 2, 3, 4, 5, 6\}$. Για $1 \leq i \leq 6$ ορίζουμε τα στοιχειώδη ενδεχόμενα $E_i = \{i\} = \text{«το ζάρι ήρθε } i\text{»}$. Έστω \mathbb{P} το μέτρο πιθανότητας που περιγράφει τις πιθανότητες αυτού του ζαριού. Μας έχει δοθεί ότι:

$$\begin{aligned} \mathbb{P}(E_1 \cup E_2) &= \mathbb{P}(E_1) + \mathbb{P}(E_2) = 1/3, \\ \text{και } \mathbb{P}(E_2 \cup E_3) &= \mathbb{P}(E_2) + \mathbb{P}(E_3) = 1/3. \end{aligned}$$

Για το κάτω φράγμα, προφανώς έχουμε $\mathbb{P}(E_2) \geq 0$, και η τιμή $\mathbb{P}(E_2) = 0$ είναι εφικτή αν επιλέξουμε, για παράδειγμα, ως μέτρο πιθανότητας το ακόλουθο,

$$\begin{aligned} \mathbb{P}(E_1) &= \mathbb{P}(E_3) = \mathbb{P}(E_5) = 1/3, \\ \text{και } \mathbb{P}(E_2) &= \mathbb{P}(E_4) = \mathbb{P}(E_6) = 0, \end{aligned}$$

το οποίο είναι συμβατό με τα δεδομένα του προβλήματος.

Για το άνω φράγμα, παρατηρούμε ότι θα πρέπει η πιθανότητα του E_2 , δηλαδή του να έρθει 2, να είναι το πολύ ίση με την πιθανότητα του ενδεχόμενου να έρθει 1 ή 2, η οποία είναι ίση με $1/3$. (Αυτό προκύπτει από τον κανόνα πιθανότητας #2). Αυτό το άνω φράγμα είναι εφικτό και επιτυγχάνεται, για παράδειγμα, για το ακόλουθο μέτρο πιθανότητας:

$$\mathbb{P}(E_1) = \mathbb{P}(E_3) = 0, \quad \mathbb{P}(E_2) = \mathbb{P}(E_4) = 1/3, \quad \mathbb{P}(E_5) = \mathbb{P}(E_6) = 1/6.$$

Παρατηρήστε ότι τα παραπάνω μέτρα πιθανότητας ικανοποιούν όλα τα αξιώματα πιθανοτήτων, και επομένως είναι αποδεκτά. Το κατά πόσο θα είναι χρήσιμα, εξαρτάται από το πόσο ανταποκρίνονται στην πραγματικότητα, και μπορούμε συνεπώς να τα χρησιμοποιήσουμε ως μοντέλα. [Από αυτή την άποψη, στις περισσότερες περιπτώσεις δεν είναι αποδεκτά, καθώς τα περισσότερα ζάρια ανταποκρίνονται στις συνθήκες $\mathbb{P}(E_i) \approx 1/6$ για κάθε $1 \leq i \leq 6$.]

3. Τυχαία συνάντηση.

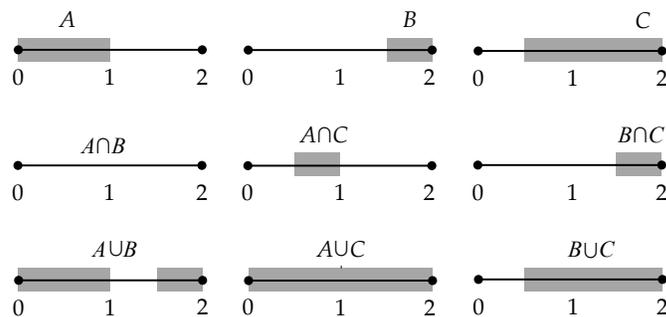
(α') Μπορούμε να χρησιμοποιήσουμε ως χώρο πιθανότητας το διάστημα $\Omega = [0, 2]$. Επειδή ο Γιάννης δεν έχει προτίμηση σε κάποιο διάστημα, είναι λογικό να υποθέσουμε ότι η πιθανότητα ενός οποιουδήποτε διαστήματος E είναι ανάλογη του μήκους του $\ell(E)$. Προκειμένου να είναι η πιθανότητα όλου του χώρου πιθανότητας Ω ίση με τη μονάδα, προκύπτει ότι πρέπει $\mathbb{P}(E) = \ell(E)/2$. Άρα, έχουμε ότι:

$$\begin{aligned} A &= [0, 1], & \mathbb{P}(A) &= 1/2, \\ B &= [3/2, 2], & \mathbb{P}(B) &= 1/4, \\ C &= [1/2, 2], & \mathbb{P}(C) &= 3/4, \\ A \cap B &= \emptyset, & \mathbb{P}(A \cap B) &= 0, \\ A \cap C &= [1/2, 1], & \mathbb{P}(A \cap C) &= 1/4, \\ B \cap C &= [3/2, 2], & \mathbb{P}(B \cap C) &= 1/4. \end{aligned}$$

(β') Χρησιμοποιώντας τον κανόνα πιθανότητας #4, βρίσκουμε,

$$\begin{aligned} \mathbb{P}(A \cup B) &= \mathbb{P}([0, 1] \cup [3/2, 2]) = \mathbb{P}([0, 1]) + \mathbb{P}([3/2, 2]) = 1/2 + 1/4 = 3/4, \\ \mathbb{P}(A \cup C) &= \mathbb{P}([0, 1] \cup [1/2, 2]) = \mathbb{P}([0, 2]) = 1, \\ \mathbb{P}(B \cup C) &= \mathbb{P}([3/2, 2] \cup [1/2, 2]) = \mathbb{P}([1/2, 2]) = 3/4. \end{aligned}$$

Όλα τα πιο πάνω ενδεχόμενα εμφανίζονται στο Σχήμα Α΄.12.



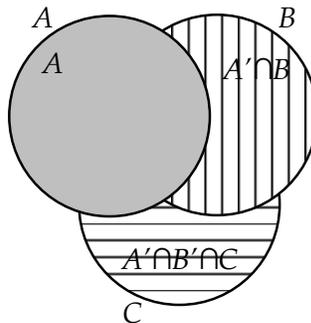
Σχήμα Α΄.12: Ο χώρος πιθανότητας και τα ενδεχόμενα της Άσκησης 3.

4. Ένωση τριών ενδεχόμενων. Παρατηρήστε κατ' αρχάς πως το $A \cup B \cup C$ είναι ίσο με την ένωση των τριών συνόλων A , $A' \cap B$, και $A' \cap B' \cap C$, δηλαδή

$$A \cup B \cup C = A \cup (A' \cap B) \cup (A' \cap B' \cap C).$$

[Αν θέλουμε να είμαστε σχολαστικοί, αυτό μπορεί να αποδειχθεί με τη συνήθη μέθοδο, δηλαδή δείχνοντας ότι κάθε στοιχείο που ανήκει στο $A \cup B \cup C$ θα ανήκει και στην ένωση, και αντιστρόφως. Πράγματι, έστω ένα στοιχείο που ανήκει στο $A \cup B \cup C$. Τότε, αν ανήκει στο A , ανήκει και στην ένωση. Αν όχι, τότε, αν ανήκει στο B , θα ανήκει στο $A' \cap B$ άρα και στην ένωση, ενώ, αν δεν ανήκει ούτε στο A ούτε στο B , τότε σίγουρα θα ανήκει στο C , οπότε και στο $A' \cap B' \cap C$, άρα και στην ένωση. Αντιστρόφως, αν ένα στοιχείο ανήκει στην ένωση, τότε είτε θα ανήκει στο A , είτε στο $A' \cap B$, άρα και στο B , είτε θα ανήκει στο $A' \cap B' \cap C$, άρα και στο C . Άρα, θα ανήκει στο $A \cup B \cup C$.]

Επιπλέον, τα τρία σύνολα της ένωσης στη σχέση (3.2) είναι ξένα μεταξύ τους, όπως είναι εύκολο να δείχτεί, παίρνοντας τις τομές ανά δύο. Άρα, η πιθανότητα της ένωσής τους $A \cup B \cup C$, θα ισούται με το άθροισμα των πιθανοτήτων τους από τον κανόνα πιθανότητας #3, και έτσι προκύπτει το ζητούμενο. Τα τρία ξένα σύνολα έχουν σχεδιαστεί στο Σχήμα A'.13.



Σχήμα A'.13: Διάγραμμα για την Άσκηση 4.

5. Ένας χώρος πιθανότητας με 3 στοιχεία. Τα στοιχειώδη ενδεχόμενα είναι ξένα μεταξύ τους, συνεπώς η πιθανότητα της ένωσής τους ισούται με το άθροισμα των πιθανοτήτων τους:

$$\frac{9}{16} = \mathbb{P}(\{a, c\}) = \mathbb{P}(\{a\} \cup \{c\}) = \mathbb{P}(\{a\}) + \mathbb{P}(\{c\}). \quad (A'.3)$$

Ομοίως προκύπτει και ότι:

$$\mathbb{P}(\{a\}) + \mathbb{P}(\{b\}) = \frac{3}{4}. \quad (A'.4)$$

Τέλος, ο χώρος πιθανότητας Ω έχει $\mathbb{P}(\Omega) = 1$, συνεπώς έχουμε:

$$\mathbb{P}(\{a\}) + \mathbb{P}(\{b\}) + \mathbb{P}(\{c\}) = \mathbb{P}(\{a\} \cup \{b\} \cup \{c\}) = \mathbb{P}(\Omega) = 1. \quad (A'.5)$$

Οι εξισώσεις (Α΄.3), (Α΄.4) και (Α΄.5) είναι ένα απλό σύστημα τριών εξισώσεων με τρεις αγνώστους, το οποίο εύκολα υπολογίζουμε ότι έχει μόνη λύση τις τιμές:

$$\mathbb{P}(\{a\}) = \frac{5}{16}, \quad \mathbb{P}(\{b\}) = \frac{7}{16}, \quad \mathbb{P}(\{c\}) = \frac{1}{4}.$$

6. Τι λένε οι πιθανοθεωρίστες στα παιδιά τους. Έστω ότι συμβαίνει ένα τυχαίο έγκλημα κάπου στην Αττική. Ορίζουμε ως A το ενδεχόμενο το έγκλημα να έγινε στην Αθήνα, και ως N το ενδεχόμενο το έγκλημα να έγινε νύχτα. Μας δίνεται ότι $\Pr(N) = 0.95$, $\Pr(A) = 0.54$ και $\Pr(N' \cap A) = 0.02$, και ζητούνται οι πιθανότητες $\Pr(N \cap A)$ και $\Pr(N \cap A')$.

Παρατηρούμε ότι,

$$\Pr(A) = \Pr((N \cap A) \cup (N' \cap A)) = \Pr(N \cap A) + \Pr(N' \cap A),$$

καθώς τα ενδεχόμενα $N \cap A$ και $N' \cap A$ είναι ζένα. Συνεπώς,

$$\Pr(N \cap A) = \Pr(A) - \Pr(N' \cap A) = 0.52.$$

Παρομοίως, έχουμε,

$$\Pr(N) = \Pr((N \cap A) \cup (N \cap A')) = \Pr(N \cap A) + \Pr(N \cap A'),$$

και άρα,

$$\Pr(N \cap A') = \Pr(N) - \Pr(N \cap A) = 0.95 - 0.52 = 0.43.$$

7. Διαιρέτες. Εφόσον η επιλογή είναι τυχαία, όλα τα στοιχειώδη ενδεχόμενα είναι ισοπίθανα και συνεπώς οι υπολογισμοί των πιθανοτήτων θα βασιστούν στον κανόνα πιθανότητας #5.

(α') Έστω A το σύνολο των αριθμών στο Ω που διαιρούνται με το 2. Έχουμε $\#A = 300$, συνεπώς η ζητούμενη πιθανότητα είναι $\Pr(A) = \frac{\#A}{\#\Omega} = 1/2$.

(β') Έστω B το σύνολο των αριθμών στο Ω που διαιρούνται με το 3. Έχουμε $\#B = 200$, συνεπώς η ζητούμενη πιθανότητα είναι $\Pr(B) = \frac{\#B}{\#\Omega} = 1/3$.

(γ') Οι αριθμοί που διαιρούνται και με το 2 και με το 3 είναι ακριβώς αυτοί που διαιρούνται με το 6. Έστω $C = A \cap B$ το σύνολο των αριθμών στο Ω που διαιρούνται με το 6. Έχουμε $\#C = 100$, συνεπώς η ζητούμενη πιθανότητα είναι $\Pr(C) = \frac{\#C}{\#\Omega} = 1/6$.

(δ') Έστω $D = A \cup B$ το σύνολο των στοιχείων του Ω που διαιρούνται με τουλάχιστον έναν από τους αριθμούς. Έχουμε,

$$\begin{aligned} \#D &= \#(A \cup B) \\ &= \#A + \#B - \#(A \cap B) \\ &= \#A + \#B - \#C \\ &= 300 + 200 - 100 = 400. \end{aligned}$$

Συνεπώς η ζητούμενη πιθανότητα είναι $\Pr(D) = \frac{\#D}{\#\Omega} = \frac{400}{600} = 2/3$.

(ε') Έστω E το σύνολο των στοιχείων που διαιρούνται με το 2 αλλά όχι με το 3. Παρατηρήστε ότι $D = E \cup B$, και ότι $E \cap B = \emptyset$. Συνεπώς $\Pr(D) = \Pr(E) + \Pr(B)$, άρα,

$$\Pr(E) = \Pr(D) - \Pr(B) = 2/3 - 1/3 = 1/3.$$

8. * Άλλη μία τυχαία συνάντηση. Ο χώρος πιθανότητας Ω αποτελείται από ζεύγη της μορφής (t_1, t_2) , όπου $t_1, t_2 \in [0, 1]$ είναι οι χρονικές στιγμές κατά τις οποίες καταφθάνουν στο μπαρ ο Σταύρος και ο Γιάννης αντίστοιχα. Δηλαδή το Ω είναι το τετράγωνο του επιπέδου μεταξύ 0 και 1: $\Omega = [0, 1] \times [0, 1] = \{(t_1, t_2) : 0 \leq t_1 \leq 1, 0 \leq t_2 \leq 1\} \subset \mathbb{R}^2$.

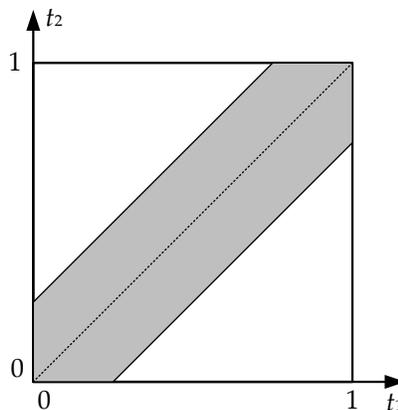
Η υπόθεσή μας ότι οι φίλοι δεν έχουν προτίμηση στη στιγμή που θα έρθουν, εντός της ώρας, μεταφράζεται στο ότι όλα τα ζεύγη τιμών έχουν την ίδια πιθανότητα. [Παρατηρήστε ότι αναγκαστικά αυτό σημαίνει ότι όλα τα ζεύγη τιμών έχουν πιθανότητα ακριβώς 0, αλλιώς έχουμε άτοπο, γιατί το άθροισμα των πιθανοτήτων όλων των αποτελεσμάτων θα είναι άπειρο και όχι 1.]

Για να μπορέσουμε να λύσουμε την άσκηση, κάνουμε και την επιπλέον υπόθεση ότι ένα σύνολο σημείων έχει πιθανότητα να προκύψει ίση με το εμβαδόν του. Πράγματι, αυτός ο κανόνας δίνει στο τετράγωνο πιθανότητα 1, και ικανοποιεί και τη διαίσθησή μας σχετικά με το ότι όλα τα ζεύγη τιμών προτιμώνται εξίσου από τους δύο φίλους. Με αυτή την υπόθεση, εύκολα προκύπτει ότι η πιθανότητα του ενδεχομένου,

$$A = \{\text{δεν θα περιμένει ο πρώτος τον δεύτερο για πάνω από ένα τέταρτο της ώρας}\},$$

ισούται με το σκιασμένο εμβαδόν του Σχήματος Α.14, το οποίο είναι ίσο (γιατί;) με,

$$\Pr(A) = 1 - 2(1/2)(3/4)(3/4) = 5/16.$$



Σχήμα Α.14: Άσκηση 8: Ο χώρος πιθανότητας Ω είναι το τετράγωνο $[0, 1] \times [0, 1]$, και το ενδεχόμενο A είναι το σκιασμένο υποσύνολο του Ω .

9. ★ Το μέτρο πιθανότητας είναι συνεχής συνάρτηση.

(α') Έχουμε,

$$\begin{aligned} \mathbb{P}\left(\lim_{i \rightarrow \infty} A_i\right) &= \mathbb{P}\left(A_1 \cup \left(\bigcup_{i=2}^{\infty} (A_i \cap A'_{i-1})\right)\right) \\ &= \mathbb{P}(A_1) + \sum_{i=2}^{\infty} \mathbb{P}(A_i \cap A'_{i-1}) \\ &= \mathbb{P}(A_1) + \sum_{i=2}^{\infty} (\mathbb{P}(A_i) - \mathbb{P}(A_{i-1})) \\ &= \lim_{n \rightarrow \infty} \mathbb{P}(A_n), \end{aligned}$$

όπου η πρώτη ισότητα προκύπτει από την υπόδειξη της άσκησης, η δεύτερη από τον κανόνα πιθανότητας #3, η τρίτη από το αποτέλεσμα της Άσκησης 1, και η τελευταία από το γεγονός ότι η εκεί σειρά είναι «τηλεσκοπική» δηλαδή ο κάθε όρος εμφανίζεται αρχικά με θετικό πρόσημο και αμέσως μετά με αρνητικό.

(β') Παρατηρήστε ότι αν τα B_i μικραίνουν, τότε τα B'_i μεγαλώνουν, και μπορούμε να εφαρμόσουμε το πρώτο σκέλος γι' αυτά. Επιπλέον, παρατηρήστε πως, από τις στοιχειώδεις ιδιότητες των πράξεων συνόλων και τον ορισμό του B , έχουμε,

$$B' = \bigcup_{i=1}^{\infty} B'_n.$$

Άρα,

$$\begin{aligned} \lim_{i \rightarrow \infty} \mathbb{P}(B_i) &= \lim_{i \rightarrow \infty} [1 - \mathbb{P}(B'_i)] \\ &= 1 - \lim_{i \rightarrow \infty} \mathbb{P}(B'_i) \\ &= 1 - \mathbb{P}\left(\lim_{i \rightarrow \infty} B'_i\right) \\ &= 1 - \mathbb{P}(B') = \mathbb{P}(B). \end{aligned}$$

Α'.4 Ασκήσεις Κεφαλαίου 4

1. Μέτρηση.

- (α') Έχουμε 12 δυνατά αποτελέσματα για την πρώτη επιλογή, 11 για τη δεύτερη, και 10 για την τρίτη. Άρα υπάρχουν $12 \times 11 \times 10 = 1320$ δυνατές εκδοχές. Πιο συστηματικά, από τον κανόνα αρίθμησης #3, το πλήθος των διατεταγμένων 3άδων είναι:

$$\frac{12!}{(12-3)!} = \frac{12 \times 11 \times 10 \times 9!}{9!} = 1320.$$

- (β') Καθεμία από τις 6 ρίψεις μπορεί να καταλήξει σε 2 αποτελέσματα, άρα, από τον κανόνα αρίθμησης #1, το πλήθος των δυνατών αποτελεσμάτων είναι:

$$2 \times 2 \times 2 \times 2 \times 2 \times 2 = 2^6 = 32.$$

- (γ') Εφόσον δεν μας ενδιαφέρει η σειρά επιλογής, το πλήθος των μη διατεταγμένων 20άδων, από τον κανόνα αρίθμησης #4, είναι:

$$\binom{100}{20} = \frac{100!}{(100-20)!20!} \approx 5.359 \times 10^{20}.$$

- (δ') Κάθε ζαριά έχει 6 δυνατά αποτελέσματα, συνεπώς από τον κανόνα αρίθμησης #1 έχουμε $6 \times 6 \times 6 \times 6 \times 6 \times 6 \times 6 = 6^7$ δυνατές εκδοχές.

- (ε') Έχουμε 13 επιλογές για τον πρώτο που θα επιλέξουμε, 12 επιλογές για τον δεύτερο κ.ο.κ. Συνεπώς έχουμε συνολικά $13 \times 12 \times 11 \times \dots \times 1 = 13! = 6,227,020,800$ επιλογές. Το ίδιο αποτέλεσμα προκύπτει και απευθείας από τον κανόνα αρίθμησης #2.

- (ς') Παρατηρήστε ότι τα φύλλα μοιράζονται με τη σειρά, συνεπώς ψάχνουμε για διατεταγμένες οκτάδες. Έχουμε 52 επιλογές για το πρώτο χαρτί, 51 επιλογές για το δεύτερο χαρτί κ.ο.κ., συνεπώς, από τον κανόνα αρίθμησης #3, το πλήθος από διαφορετικές οκτάδες είναι $\frac{52!}{(52-8)!} = 30,342,338,208,000$.

- (ζ') Κάθε νόμισμα έχει 2 δυνατά αποτελέσματα και κάθε ζάρι 6 δυνατά αποτελέσματα, άρα από τον κανόνα αρίθμησης #1 έχουμε $2 \times 2 \times 2 \times 2 \times 6 \times 6 = 576$ δυνατές εκδοχές.

2. * *Επιλογές με επανατοποθέτηση.* Ένας βολικός τρόπος να αναπαραστήσουμε μια οποιαδήποτε επιλογή, είναι με μια ακολουθία από ακριβώς $n-1$ κάθετους (/) και ακριβώς k αστεράκια (*), σε αυθαίρετη σειρά. Ο αριθμός των φορών που επιλέξαμε το αντικείμενο i ισούται με τον αριθμό από αστεράκια ανάμεσα στην κάθετο $i-1$ και την κάθετο i . Ειδικά για το πρώτο αντικείμενο, μετράμε τα αστεράκια πριν την πρώτη κάθετο, και για το τελευταίο αντικείμενο μετράμε τα αστεράκια μετά την τελευταία κάθετο. Για παράδειγμα, στην περίπτωση του πρώτου παραδείγματος, η ακολουθία ***////**/** σημαίνει ότι επιλέξαμε δύο

φορές το πρώτο υλικό, μία φορά το πέμπτο, μία το έκτο, και μία το έβδομο. Όμως, ο αριθμός των διαφορετικών ακολουθιών αυτού του τύπου είναι ίσος με τον αριθμό των τρόπων που μπορούμε να τοποθετήσουμε k αστερίσκους σε $n + k - 1$ θέσεις, οπότε ο ζητούμενος αριθμός από συνδυασμούς αυτού του τύπου είναι ίσος με:

$$\binom{n+k-1}{k}.$$

3. *1-2-X*. Εδώ απλά θέλουμε να μοιράσουμε 9 «αντικείμενα» (τα 9 ματς) σε τρεις κατηγορίες των τεσσάρων, τριών και δύο. Συνεπώς, υπάρχουν,

$$\binom{9}{4\ 3\ 2} = \frac{9!}{4!3!2!} = 1260 \quad \text{τρόποι.}$$

4. *Τράπουλα*. Κατ' αρχάς παρατηρούμε ότι ο χώρος πιθανότητας Ω του προβλήματος περιλαμβάνει όλες τις δυνατές (μη διατεταγμένες) δεκάδες φύλλων που μπορούν να επιλεγούν από 52 φύλλα, οπότε, $\#\Omega = \binom{52}{10} = \frac{52!}{(52-10)!10!} = 15,820,024,220$.

(α') Έστω A το ενδεχόμενο να μην επιλεγεί κανένας άσος. Υπάρχουν $\binom{48}{10}$ τρόποι με τους οποίους μπορεί να σχηματισθεί το A , ένας για καθεμία επιλογή των 10 φύλλων από τα υπόλοιπα 48. Συνεπώς, από τον κανόνα πιθανότητας #5, $\Pr(A) = \binom{48}{10} / \binom{52}{10} \approx 0.41$.

(β') Έστω B το ενδεχόμενο να πάρουμε το πολύ τρεις άσους. Το ενδεχόμενο B' περιέχει όλες τις μοιρασιές στις οποίες πήραμε και τους 4 άσους, και ο υπολογισμός της πιθανότητάς του είναι απλούστερος. Πράγματι, υπάρχουν $\binom{4}{4}$ τρόποι να επιλέξουμε τους 4 άσους και $\binom{48}{6}$ τρόποι να επιλέξουμε τα υπόλοιπα 6 φύλλα. Συνεπώς, από τον κανόνα πιθανότητας #5,

$$\Pr(B) = 1 - \Pr(B') = 1 - \frac{\binom{4}{4}\binom{48}{6}}{\binom{52}{10}} \approx 0.992.$$

5. *Poker*. Σε αυτό το παράδειγμα, ο χώρος πιθανότητας Ω αποτελείται από όλες τις $\binom{52}{5}$ δυνατές πεντάδες φύλλων που μπορεί να έχει ο παίκτης, όπως προκύπτει από τον κανόνα αρίθμησης #4.

(α') Έστω A το σύνολο των 5άδων που περιέχουν καρέ. Κατ' αρχάς παρατηρούμε πως υπάρχουν 48 διαφορετικές πεντάδες με καρέ του άσου, μία για κάθε ένα από τα 48 φύλλα με τα οποία συμπληρώνουμε την πεντάδα. Υπάρχουν άλλες 48 πεντάδες με καρέ του ρήγα κ.ο.κ. Και εφόσον υπάρχουν 13 διαφορετικά καρέ, το συνολικό πλήθος των στοιχείων του A είναι 13×48 . Άρα, από τον κανόνα πιθανότητας #5, η ζητούμενη πιθανότητα είναι:

$$\Pr(A) = \frac{\#A}{\#\Omega} = \frac{13 \times 48}{\binom{52}{5}} = \frac{13 \times 48 \times 5! \times 47!}{52!} \approx 2.4 \times 10^{-4}.$$

(β') Έστω B το σύνολο των 5άδων που αντιστοιχούν σε φουλ. Κατ' αρχάς, πόσες πεντάδες αντιστοιχούν σε φουλ του άσου με ρηγάδες; Υπάρχουν $\binom{4}{3}$ τρόποι να επιλέξουμε 3 άσους και $\binom{4}{2}$ τρόποι να επιλέξουμε 2 ρηγάδες. Άρα υπάρχουν $\binom{4}{3} \times \binom{4}{2}$ πεντάδες που περιέχουν φουλ του άσου με ρηγάδες. Επιπλέον, υπάρχουν 13×12 διαφορετικά είδη φουλ: Φουλ του άσου με ρηγάδες, φουλ του άσου με ντάμες κ.ο.κ. Άρα, το πλήθος όλων των δυνατών 5άδων που αντιστοιχούν σε φουλ είναι $\#B = 13 \times 12 \times \binom{4}{3} \times \binom{4}{2}$ και, από τον κανόνα πιθανότητας #5, η πιθανότητα για φουλ είναι:

$$\begin{aligned} \Pr(B) &= \frac{\#B}{\#\Omega} = \frac{13 \times 12 \times \binom{4}{3} \times \binom{4}{2}}{\binom{52}{5}} \\ &= \frac{13 \times 12 \times 4! \times 4! \times 5! \times 47!}{3! \times 1! \times 2! \times 2! \times 52!} \approx 0.0014. \end{aligned}$$

(γ') Τέλος, έστω C το σύνολο των 5άδων που αντιστοιχούν σε χρώμα. Υπάρχουν $\binom{13}{5}$ πεντάδες από κούπες, και το ίδιο πλήθος για τα μπαστούνια, σπαθιά και καρό, άρα, $\#C = 4 \times \binom{13}{5}$. Οπότε, από τον κανόνα πιθανότητας #5, η πιθανότητα του χρώματος είναι:

$$\Pr(C) = \frac{\#C}{\#\Omega} = \frac{4 \times \binom{13}{5}}{\binom{52}{5}} = \frac{4 \times 13! \times 5! \times 47!}{5! \times 8! \times 52!} \approx 0.002.$$

6. ★ Ξενοδοχείο Ακρόπολις.

(α') Έστω πως δίνουμε έναν αριθμό από το 1 ως το 4 στα διαφορετικά ξενοδοχεία. Το τυχαίο πείραμα εδώ συνίσταται στην τυχαία επιλογή του ξενοδοχείου μετάβασης για καθέναν από τους 6 φίλους. Το σύνολο των δυνατών αποτελεσμάτων θα αποτελείται από όλες τις διατεταγμένες εξάδες $(a_1, a_2, a_3, a_4, a_5, a_6)$ ξενοδοχείων (όπου το κάθε $a_i = 1, 2, 3$ ή 4). Π.χ. το στοιχείο $(2, 3, 2, 1, 4, 4)$ αντιστοιχεί στο ενδεχόμενο ο πρώτος και ο τρίτος να πήγαν στο 2ο ξενοδοχείο, ο δεύτερος στο 3ο, ο τέταρτος στο 1ο, και οι πέμπτος και έκτος στο 4ο. Ο χώρος πιθανότητας Ω επομένως θα αποτελείται από όλες τις $4^6 = 4096$ τέτοιες εξάδες.

(β') Έστω A το ενδεχόμενο να καταλήξουν οι φίλοι σε τρία ζεύγη. Παρατηρήστε ότι, πριν σκεφτούμε σε ποια ξενοδοχεία θα πάνε, υπάρχουν $\binom{6}{2} \binom{4}{2}$ τρόποι να μοιραστούν 6 άνθρωποι σε 3 ζεύγη, αν μας ενδιαφέρει η σειρά σχηματισμού. [Εναλλακτικά, από τον κανόνα πιθανότητας #5, υπάρχουν $\binom{6}{2 \ 2 \ 2}$ τρόποι, το οποίο μας δίνει το ίδιο αποτέλεσμα.] Συνεπώς υπάρχουν $\binom{6}{2} \binom{4}{2} / 3!$ αν, όπως στην προκειμένη περίπτωση, δεν μας ενδιαφέρει η σειρά σχηματισμού. [Θυμηθείτε το συλλογισμό που μας οδήγησε από τον κανόνα πιθανότητας #3 στον #4.]

Αφού έχουμε τα τρία ζευγάρια, υπάρχουν $4 \times 3 \times 2 = 24$ τρόποι να μοιραστούν σε ξενοδοχεία (4 επιλογές για το πρώτο, 3 για το δεύτερο, 2 για το τρίτο). Άρα, το πλήθος

των στοιχείων του A είναι $\#A = 24 \times \binom{4}{2} \binom{6}{2} / 3!$, και, από τον κανόνα πιθανότητας #5, η ζητούμενη πιθανότητα ισούται με,

$$\Pr(A) = \frac{\#A}{\#\Omega} = \frac{24 \times \binom{4}{2} \binom{6}{2} / 3!}{4^6} = \frac{45}{512} \approx 0.0879.$$

(γ) Έστω B το ενδεχόμενο να βρεθούν δύο μόνοι τους και άλλοι τέσσερις σε δύο ζεύγη. Παρατηρήστε ότι υπάρχουν $\binom{6}{2} \binom{4}{2}$ τρόποι να σχηματιστούν τα δύο ζεύγη αν μας ενδιαφέρει η σειρά σχηματισμού, και $\binom{6}{2} \binom{4}{2} / 2$ αν δεν μας ενδιαφέρει. Οπότε υπάρχουν $4 \times 3 \times 2 = 24$ τρόποι να μοιραστούν οι τέσσερις ομάδες (δύο ζεύγη και δύο άτομα μόνα) σε ξενοδοχεία (4 επιλογές για το πρώτο, 3 για το δεύτερο, 2 για το τρίτο). Άρα, από τον κανόνα πιθανότητας #5, η ζητούμενη πιθανότητα ισούται με:

$$\Pr(B) = \frac{12 \times \binom{4}{2} \binom{6}{2}}{4^6} = \frac{135}{512} \approx 0.2637.$$

7. *Superleague.* Υπάρχουν δύο τρόποι να λύσουμε την άσκηση. Πρώτον, παρατηρήστε πως υπάρχουν συνολικά $\binom{16}{2}$ ζεύγη ομάδων, και άρα πρέπει να γίνουν $2 \times \binom{16}{2} = 240$ αγώνες. Εναλλακτικά, παρατηρήστε πως καθεμία από τις 16 ομάδες θα υποδεχτεί καθεμία από τις άλλες 15 στο γήπεδό της, άρα θα γίνουν $16 \times 15 = 240$ αγώνες.

8. *Λόττο.* Υπάρχουν $\binom{49}{6} = 13,983,816$ δυνατοί συνδυασμοί, όλοι τους ισοπίθανοι. Συνεπώς, από τον κανόνα πιθανότητας #5, η πιθανότητα να κερδίσουμε με έναν μόνο λαχνό είναι,

$$\frac{1}{\binom{49}{6}} \approx 7.15112384 \times 10^{-8},$$

δηλαδή μικρότερη από μία στα 10 εκατομμύρια.

9. *Εύκολες και δύσκολες ασκήσεις.* Εδώ ο χώρος πιθανότητας Ω αποτελείται από όλα τα $\#\Omega = \binom{18}{6\ 6\ 6}$ δυνατά αποτελέσματα της μοιρασιάς. Εφόσον οι ασκήσεις μοιράζονται τυχαία, θα χρησιμοποιήσουμε τον κανόνα πιθανότητας #5.

Έστω A το ενδεχόμενο να έχει ακριβώς μία εύκολη άσκηση κάθε ομάδα. Το πλήθος όλων των δυνατών συνδυασμών με τους οποίους οι ομάδες μπορούν να πάρουν από μία εύκολη άσκηση είναι $\binom{3}{1\ 1\ 1}$, και οι συνδυασμοί με τους οποίους μπορούν να πάρουν τις δύσκολες ασκήσεις είναι $\binom{15}{5\ 5\ 5}$. Άρα, τελικά έχουμε:

$$\Pr(A) = \frac{\#A}{\#\Omega} = \frac{\binom{3}{1\ 1\ 1} \binom{15}{5\ 5\ 5}}{\binom{18}{6\ 6\ 6}} = \frac{3!15!6!6!}{5!5!5!18!} \approx 0.2647.$$

10. Μέτρηση αποτελεσμάτων.

- (α') Στη λέξη ΚΥΠΡΟΣ υπάρχουν 6 γράμματα, όλα διαφορετικά μεταξύ τους. Έχουμε 6 επιλογές για το ποιο γράμμα θα μπει πρώτο, 5 επιλογές για το ποιο γράμμα θα μπει δεύτερο κ.ο.κ. Συνεπώς, συνολικά έχουμε $6! = 720$ αναγραμματισμούς. Στη λέξη ΣΤΑΥΡΟΣ υπάρχουν 7 γράμματα, που ομοίως μπορούν να παρατεθούν με $7!$ τρόπους. Δύο από αυτά όμως (τα Σ), είναι όμοια, συνεπώς μπορούμε να αλλάξουμε τη σειρά τους, χωρίς να αλλάξει ο αναγραμματισμός. Επειδή τα δυο Σ μπορούν να τοποθετηθούν με δύο διαφορετικούς τρόπους, έχουμε συνολικά $\frac{7!}{2} = 2520$ διαφορετικούς αναγραμματισμούς. Για τη λέξη ΣΙΣΙΝΙ, με παρόμοια συλλογιστική, βρίσκουμε ότι υπάρχουν $\frac{6!}{3!2!} = 60$ διαφορετικοί αναγραμματισμοί.
- (β') Μπορούμε να επιλέξουμε τις γυναίκες με $\binom{12}{3}$ διαφορετικούς τρόπους και τους άντρες με $\binom{7}{3}$ διαφορετικούς τρόπους. Από τις επιλεγμένες γυναίκες, η πρώτη μπορεί να επιλέξει τον καβαλιέρο της με 3 τρόπους, η δεύτερη με 2 τρόπους, και η τελευταία με έναν τρόπο. Συνοψίζοντας, τα ζεύγη μπορούν να σχηματιστούν με $\binom{12}{3} \binom{7}{3} 3! = 46,200$ τρόπους.
- (γ') Στην περίπτωση που τα υλικά επαναλαμβάνονται, είμαστε στην περίπτωση της Άσκησης 2, άρα υπάρχουν $\binom{20+6-1}{6} = \binom{25}{6} = 177,100$ συνδυασμοί. Στην περίπτωση που τα υλικά δεν επαναλαμβάνονται, από τον κανόνα αρίθμησης #4 έχουμε $\binom{20}{6} = 38,760$ συνδυασμούς.

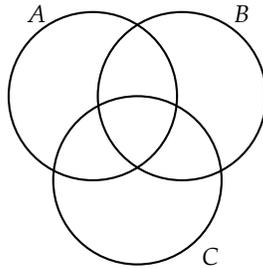
Α'.5 Ασκήσεις Κεφαλαίου 5

1. Ένωση τριών ενδεχομένων. Έχουμε:

$$\begin{aligned}
 \Pr(A \cup B \cup C) &= \Pr(A \cup (B \cup C)) \\
 &= \Pr(A) + \Pr(B \cup C) - \Pr(A \cap (B \cup C)) \\
 &= \Pr(A) + \Pr(B) + \Pr(C) - \Pr(B \cap C) - \Pr((A \cap B) \cup (A \cap C)) \\
 &= \Pr(A) + \Pr(B) + \Pr(C) - \Pr(B \cap C) \\
 &\quad - \Pr(A \cap B) - \Pr(A \cap C) + \Pr((A \cap B) \cap (A \cap C)) \\
 &= \Pr(A) + \Pr(B) + \Pr(C) - \Pr(A \cap B) \\
 &\quad - \Pr(A \cap C) - \Pr(B \cap C) + \Pr(A \cap B \cap C).
 \end{aligned}$$

Η δεύτερη και η τέταρτη ισότητα προκύπτουν από τον κανόνα πιθανότητας #8, και η τρίτη ισότητα προκύπτει από τον ίδιο κανόνα πιθανότητας και την ιδιότητα 4 για τις σχέσεις συνόλων στην τελευταία ενότητα του Κεφαλαίου 2.

Μια διαισθητική ερμηνεία του αποτελέσματος μπορεί να προκύψει κοιτάζοντας Σχήμα Α'.15: Η πιθανότητα της ένωσης ισούται με το άθροισμα των πιθανοτήτων $\Pr(A) + \Pr(B) + \Pr(C)$ των επιμέρους ενδεχομένων, εφόσον βέβαια αφαιρέσουμε τις πιθανότητες των κομματιών αυτών των ενδεχομένων που προστέθηκαν πάνω από μία φορά. Αν αφαιρέσουμε την ποσότητα $\Pr(A \cap B) + \Pr(A \cap C) + \Pr(B \cap C)$, τότε οι πιθανότητες αυτών των κομματιών θα έχουν αφαιρεθεί σωστά με εξαίρεση την πιθανότητα του $A \cap B \cap C$, η οποία αφαιρέθηκε μία φορά παραπάνω από όσες έπρεπε, και συνεπώς πρέπει να προστεθεί στην τελική έκφραση.



Σχήμα Α'.15: Διάγραμμα για την Άσκηση 1.

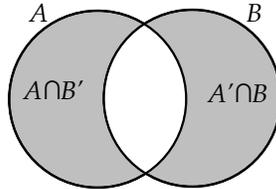
2. Πιθανότητα ακριβώς ενός ενδεχομένου. Το ενδεχόμενο C να πραγματοποιηθεί ακριβώς ένα από τα ενδεχόμενα A, B , είναι το να πραγματοποιηθεί το A και όχι το B , ή να πραγματοποιηθεί το B και όχι το A , δηλαδή, $C = (A \cap B') \cup (A' \cap B)$. Κατ' αρχάς παρατηρούμε ότι $C \cup (A \cap B) = A \cup B$ και $C \cap (A \cap B) = \emptyset$ (γιατί;). Συνεπώς, από τον κανόνα πιθανότητας #3,

$$\Pr(A \cup B) = \Pr(C \cup (A \cap B)) = \Pr(C) + \Pr(A \cap B). \quad (\text{Α'.6})$$

Επιπλέον, από τον κανόνα πιθανότητας #8,

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B). \quad (\text{A'.7})$$

Συνδυάζοντας τις (A'.6), (A'.7), προκύπτει το ζητούμενο. Το αποτέλεσμα εύκολα μπορεί να αιτιολογηθεί και διαισθητικά, με χρήση του Σχήματος A'.16.



Σχήμα A'.16: Διάγραμμα για την Άσκηση 2.

3. *Τράπουλα.* Έστω C το ενδεχόμενο να πάρουμε τουλάχιστον έναν άσο, και D το ενδεχόμενο να πάρουμε τουλάχιστον μία φιγούρα. Ζητείται η πιθανότητα του ενδεχόμενου $C \cap D$. Εδώ ο χώρος πιθανότητας Ω αποτελείται από όλες τις δυνατές (μη διατεταγμένες) δεκάδες φύλλων που έχουν να επιλεγούν από 52 φύλλα, και εφόσον η επιλογή είναι τυχαία, όλα τα στοιχειώδη ενδεχόμενα είναι ισοπίθανα.

Χρησιμοποιώντας τον κανόνα πιθανότητας #4 και μετά τον κανόνα πιθανότητας #8, έχουμε,

$$\Pr(C \cap D) = 1 - \Pr(C' \cup D') = 1 - \Pr(C') - \Pr(D') + \Pr(C' \cap D'),$$

όπου η καθεμία από τις πιο πάνω πιθανότητες μπορεί να υπολογιστεί από τον κανόνα πιθανότητας #5 σε συνδυασμό και με κάποιους από τους κανόνες αρίθμησης του Κεφαλαίου 4. Συγκεκριμένα έχουμε $\#\Omega = \binom{52}{10}$, $\#C' = \binom{48}{10}$, $\#D' = \binom{40}{10}$ και $\#(C' \cap D') = \binom{36}{10}$, οπότε:

$$\Pr(C \cap D) = 1 - \frac{\binom{48}{10}}{\binom{52}{10}} - \frac{\binom{40}{10}}{\binom{52}{10}} + \frac{\binom{36}{10}}{\binom{52}{10}} \approx 0.55.$$

4. *Ιδιότητες δεσμευμένης πιθανότητας.*

(α') Χρησιμοποιώντας απλώς τον ορισμό της δεσμευμένης πιθανότητας, έχουμε:

$$\begin{aligned} \Pr(A) \Pr(B|A) \Pr(C|A \cap B) &= \Pr(A) \cdot \frac{\Pr(A \cap B)}{\Pr(A)} \cdot \frac{\Pr(A \cap B \cap C)}{\Pr(A \cap B)} \\ &= \Pr(A \cap B \cap C). \end{aligned}$$

[Μπορείτε να γενικεύσετε την παραπάνω σχέση σε n αντί για μόνο 3 ενδεχόμενα;]

(β') Και πάλι θα χρησιμοποιήσουμε μόνο τον ορισμό της δεσμευμένης πιθανότητας. Αν $A \cap B = \emptyset$, τότε τα A και B είναι αδύνατον να συμβούν μαζί, οπότε, διαισθητικά, δεδομένου ότι έχει συμβεί το B είναι αδύνατον να συμβεί και το A . Πράγματι,

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(\emptyset)}{\Pr(B)} = 0.$$

Αν $A \subset B$, τότε:

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A)}{\Pr(B)}.$$

Αν $B \subset A$, τότε, δεδομένου ότι έχει συμβεί το B , σίγουρα θα έχει συμβεί κάποιο από τα αποτελέσματα που ανήκουν και στο A . Πράγματι,

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(B)}{\Pr(B)} = 1.$$

(γ') Χρησιμοποιώντας τον ορισμό της δεσμευμένης πιθανότητας, έχουμε,

$$\frac{\Pr(A|B)}{\Pr(A)} = \frac{\Pr(A \cap B)}{\Pr(A) \Pr(B)} = \frac{\Pr(B|A)}{\Pr(B)}.$$

Άρα, το αριστερό μέλος είναι > 1 , δηλαδή έχουμε $\Pr(A|B) > \Pr(A)$, αν και μόνο αν και το δεξί μέλος είναι > 1 , δηλαδή $\Pr(B|A) > \Pr(B)$.

5. Ρίψεις ζαριού. Ο χώρος πιθανότητας Ω αποτελείται από 36 αποτελέσματα, και συγκεκριμένα από τις διατεταγμένες δυάδες της μορφής (i, j) , όπου $i, j = 1, 2, \dots, 6$. Χρησιμοποιώντας το γεγονός ότι το ζάρι είναι δίκαιο και οι ρίψεις ανεξάρτητες, προκύπτει ότι όλα τα στοιχειώδη ενδεχόμενα έχουν την ίδια πιθανότητα, προφανώς $1/36$. Πράγματι, για παράδειγμα,

$$\begin{aligned} \Pr(\{(1, 2)\}) &= \Pr(\{\langle \text{Πρώτο ζάρι} = 1 \rangle\} \cap \{\langle \text{Δεύτερο ζάρι} = 2 \rangle\}) \\ &= \Pr(\{\langle \text{Πρώτο ζάρι} = 1 \rangle\}) \Pr(\{\langle \text{Δεύτερο ζάρι} = 2 \rangle\}) \\ &= \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}, \end{aligned}$$

όπου η δεύτερη ισότητα προκύπτει από την ανεξαρτησία και η τρίτη λόγω του ότι το ζάρι είναι δίκαιο. Βάσει των παραπάνω, μπορούμε να υπολογίσουμε την πιθανότητα οποιουδήποτε ενδεχόμενου απλά μετρώντας το πλήθος των στοιχείων του. Άρα, αφού υπάρχουν 18 μονά και 18 ζυγά αποτελέσματα, $\Pr(A) = \Pr(B) = \frac{18}{36} = \frac{1}{2}$. Επιπλέον,

$$\Pr(C) = \Pr(\{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}) = \frac{6}{36} = \frac{1}{6},$$

$$\Pr(D) = \Pr(\{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)\}) = \frac{6}{36} = \frac{1}{6},$$

$$\Pr(G) = \Pr(\{(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)\}) = \frac{5}{36},$$

ενώ προφανώς, $\Pr(E) = \Pr(F) = 1/6$.

Για να αποφανθούμε για την ανεξαρτησία των ζητούμενων ζευγαριών, χρειαζόμαστε την πιθανότητα της τομής τους, εύκολα βρίσκουμε, αν μετρήσουμε τα αποτελέσματα που ανήκουν σε κάθε τομή, ότι:

$$\begin{aligned}\Pr(A \cap B) &= \Pr(\emptyset) = 0, \\ \Pr(C \cap B) &= \Pr(C) = \frac{1}{6}, \\ \Pr(D \cap E) &= \Pr(\{(6, 6)\}) = \frac{1}{36}, \\ \Pr(C \cap F) &= \Pr(\{(1, 6)\}) = \frac{1}{36}, \\ \Pr(G, F) &= \Pr(\{(1, 5)\}) = \frac{1}{36}.\end{aligned}$$

Έτσι, συγκρίνοντας την πιθανότητα της τομής με το γινόμενο των επιμέρους πιθανοτήτων σε κάθε περίπτωση, παρατηρούμε ότι τα δύο ζεύγη D και E , C και F είναι ανεξάρτητα, ενώ όλα τα άλλα όχι.

6. Δύο δοχεία. Εφόσον υποθέτουμε πως $\Pr(F) = \Pr(F') = \frac{1}{2}$, από τον κανόνα συνολικής πιθανότητας,

$$\Pr(M) = \Pr(M|F) \Pr(F) + \Pr(M|F') \Pr(F') = \frac{1}{2} \frac{m_1}{m_1 + w_1} + \frac{1}{2} \frac{m_2}{m_2 + w_2}.$$

Επιπλέον, από τον ορισμό της δεσμευμένης πιθανότητας,

$$\Pr(M \cap F) = \Pr(M|F) \Pr(F) = \frac{1}{2} \frac{m_1}{w_1 + m_1}.$$

Άρα, για να έχουμε ανεξαρτησία των ενδεχόμενων M και F , εξ ορισμού απαιτείται να ισχύει ότι, $\Pr(M \cap F) = \Pr(M) \Pr(F)$, δηλαδή,

$$\frac{1}{2} \frac{m_1}{w_1 + m_1} = \frac{1}{2} \left(\frac{1}{2} \frac{m_1}{m_1 + w_1} + \frac{1}{2} \frac{m_2}{m_2 + w_2} \right),$$

ή, ισοδύναμα, απλοποιώντας την πιο πάνω σχέση,

$$\frac{m_1}{w_1} = \frac{m_2}{w_2}.$$

Η απάντηση αυτή είναι διαισθητικά αναμενόμενη: Για να είναι η επιλογή του τμήματος ανεξάρτητη από την επιλογή του φύλου, πρέπει η αναλογία φοιτητριών/φοιτητών στα δύο τμήματα να είναι η ίδια.

7. Poker με δεσμευμένες πιθανότητες.

- (α') Φανταστείτε πως μας αποκαλύπτονται τα φύλλα όχι ταυτόχρονα, όπως σε λύσεις που βασίζονται σε συνδυαστική, αλλά διαδοχικά. Παρατηρήστε ότι το ενδεχόμενο του καρέ μπορεί να γραφεί ως ένωση πέντε ξένων ενδεχόμενων E_i , όπου το E_i είναι το «παράταιρο» φύλλο να είναι το i -οστό που θα μας αποκαλυφθεί. Λόγω συμμετρίας, η ζητούμενη πιθανότητα είναι ίση με $5 \Pr(E_1)$ (γιατί;).

Για να υπολογίσουμε το $\Pr(E_1)$, σκεφτόμαστε ως εξής: Έστω πως έχουμε τραβήξει το πρώτο χαρτί. Η πιθανότητα να κάνουμε τελικά καρέ είναι η πιθανότητα η αξία (δηλαδή το «νούμερο», άσος, 2, 3 κλπ.) του δεύτερου φύλλου να είναι διαφορετική από του πρώτου ($48/51$), επί την πιθανότητα το τρίτο να έχει την ίδια αξία με το δεύτερο ($3/50$), επί την πιθανότητα το τέταρτο να είναι ίδιο με το δεύτερο ($2/49$), επί την πιθανότητα το τελευταίο να είναι ίδιο με το δεύτερο ($1/48$).

Αν θέλουμε να είμαστε πιο σχολαστικοί, ορίζουμε τα ακόλουθα ενδεχόμενα:

- A_2 = η αξία του δεύτερου φύλλου είναι διαφορετική από του πρώτου,
 A_3 = το τρίτο φύλλο έχει ίδια αξία με το δεύτερο,
 A_4 = το τέταρτο φύλλο έχει ίδια αξία με το δεύτερο,
 A_5 = το πέμπτο φύλλο έχει ίδια αξία με το δεύτερο,

Τότε, έχουμε, από τον ορισμό της δεσμευμένη πιθανότητας (δείτε και το πρώτο σκέλος της Άσκησης 4):

$$\begin{aligned} \Pr(E_1) &= \Pr(A_2 \cap A_3 \cap A_4 \cap A_5) \\ &= \Pr(A_2) \Pr(A_3|A_2) \Pr(A_4|A_2 \cap A_3) \Pr(A_5|A_2 \cap A_3 \cap A_4) \\ &= \frac{48}{51} \times \frac{3}{50} \times \frac{2}{49} \times \frac{1}{48}. \end{aligned}$$

Άρα, η πιθανότητα να έρθει καρέ είναι ίση με:

$$5 \Pr(E_1) = 5 \times \frac{48}{51} \times \frac{3}{50} \times \frac{2}{49} \times \frac{1}{48} \approx 2.4 \times 10^{-4}.$$

- (β') Για να υπολογίσουμε την πιθανότητα φουλ, θα υπολογίσουμε πρώτα την πιθανότητα του ενδεχόμενου B_2 να σηκώσουμε φύλλα από ακριβώς δύο αξίες. Παρατηρήστε ότι θα έχουμε φύλλα με ακριβώς δύο διαφορετικές αξίες αν και μόνο αν έχουμε καρέ ή φουλ, οπότε το ενδεχόμενο B_2 μπορεί να εκφραστεί ως ένωση δύο ξένων ενδεχομένων, $B_2 = A \cup B$, όπου B =«έχουμε φουλ» και A =«έχουμε καρέ». Την πιθανότητα του A την υπολογίσαμε ήδη, άρα για να βρούμε τη ζητούμενη πιθανότητα,

$$\Pr(B) = \Pr(B_2) - \Pr(A),$$

αρκεί να υπολογίσουμε την πιθανότητα του B_2 .

Γι' αυτόν το σκοπό θα εκφράσουμε το B_2 ως την ένωση των ξένων ενδεχομένων,

$$B_2 = B_{22} \cup B_{23} \cup B_{24} \cup B_{25},$$

όπου το B_{2i} είναι το υποσύνολο του B_2 όπου η δεύτερη από τις δύο αξίες της πεντάδας εμφανίζεται για πρώτη φορά στο φύλλο i . Παρατηρήστε πως:

$$\begin{aligned} \Pr(B_{22}) &= \frac{48}{51} \times \frac{6}{50} \times \frac{5}{49} \times \frac{4}{48}, \\ \Pr(B_{23}) &= \frac{3}{51} \times \frac{48}{50} \times \frac{5}{49} \times \frac{4}{48}, \\ \Pr(B_{24}) &= \frac{3}{51} \times \frac{2}{50} \times \frac{48}{49} \times \frac{4}{48}, \\ \Pr(B_{25}) &= \frac{3}{51} \times \frac{2}{50} \times \frac{1}{49} \times \frac{48}{48}. \end{aligned}$$

Άρα, τελικά, η ζητούμενη πιθανότητα $\Pr(B)$ ισούται με:

$$\begin{aligned} \Pr(B) &= \Pr(B_{22}) + \Pr(B_{23}) + \Pr(B_{24}) + \Pr(B_{25}) - \Pr(A) \\ &= \frac{180}{51 \times 50 \times 49} \approx 0.0014. \end{aligned}$$

(γ') Έστω ότι έχουμε τραβήξει το πρώτο χαρτί. Η πιθανότητα το δεύτερο να είναι του ίδιου χρώματος είναι $\frac{12}{51}$, γιατί έχουν μείνει 51 φύλλα προς επιλογή, εκ των οποίων 12 είναι του ίδιου χρώματος. Με δεδομένο ότι το δεύτερο φύλλο έχει το ίδιο χρώμα με το πρώτο, το τρίτο φύλλο θα είναι επίσης του ίδιου χρώματος με πιθανότητα $\frac{11}{50}$ κ.ο.κ. Πιο συστηματικά, ορίζοντας τα ενδεχόμενα,

$$Z_i = \text{«Το } i \text{ φύλλο έχει ίδιο χρώμα με το πρώτο»}, \quad i = 2, 3, 4, 5,$$

έχουμε, παρομοίως με το πρώτο σκέλος της άσκησης,

$$\begin{aligned} \Pr(\text{«χρώμα»}) &= \Pr(Z_2 \cap Z_3 \cap Z_4 \cap Z_5) \\ &= \Pr(Z_2) \Pr(Z_3|Z_2) \Pr(Z_4|Z_2 \cap Z_3) \Pr(Z_5|Z_2 \cap Z_3 \cap Z_4) \\ &= \frac{12}{51} \times \frac{11}{50} \times \frac{10}{49} \times \frac{9}{48} \\ &\approx 0.002. \end{aligned}$$

Ένα ηθικό δίδαγμα αυτής της άσκησης είναι ότι πολλά προβλήματα μπορούν να αντιμετωπιστούν είτε με την εφαρμογή κανόνων της συνδυαστικής, είτε με χρήση της δεσμευμένης πιθανότητας, αλλά πολύ συχνά η μία μέθοδος είναι σαφώς ευκολότερη.

8. *Τεστ πολλαπλών απαντήσεων.* Ορίζουμε τα ενδεχόμενα $A = \text{«Ο μαθητής απαντά σωστά»}$ και $C = \text{«Ο μαθητής γνωρίζει την απάντηση»}$, οπότε η ζητούμενη πιθανότητα είναι η $\Pr(C|A)$. Από τον κανόνα του Bayes:

$$\begin{aligned} \Pr(C|A) &= \frac{\Pr(A|C) \Pr(C)}{\Pr(A|C) \Pr(C) + \Pr(A|C') \Pr(C')} \\ &= \frac{1 \times p}{1 \times p + \frac{1}{N} \times (1 - p)} \\ &= \frac{Np}{(N - 1)p + 1}. \end{aligned}$$

9. *Δεσμευμένο μέτρο πιθανότητας.*

- (α') Από τον πρώτο κανόνα πιθανότητας έχουμε ότι $\Pr(A \cap B) \geq 0$, και υποθέτουμε ότι $\Pr(B) > 0$. Συνεπώς, από τον ορισμό της δεσμευμένης πιθανότητας:

$$\mathbb{P}(A) = \Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} \geq 0.$$

- (β') Απλά εφαρμόζοντας τον ορισμό του μέτρου πιθανότητας και τον ορισμό της δεσμευμένης πιθανότητας:

$$\mathbb{P}(\Omega) = \Pr(\Omega|B) = \frac{\Pr(\Omega \cap B)}{\Pr(B)} = \frac{\Pr(B)}{\Pr(B)} = 1.$$

- (γ') Για την τελευταία ιδιότητα έχουμε,

$$\begin{aligned} \mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) &= \Pr\left(\bigcup_{i=1}^{\infty} A_i \mid B\right) \\ &= \frac{\Pr\left(\left(\bigcup_{i=1}^{\infty} A_i\right) \cap B\right)}{\Pr(B)} \\ &= \frac{\Pr\left(\bigcup_{i=1}^{\infty} (A_i \cap B)\right)}{\Pr(B)} \\ &= \sum_{i=1}^{\infty} \frac{\Pr(A_i \cap B)}{\Pr(B)} \\ &= \sum_{i=1}^{\infty} \Pr(A_i|B) = \sum_{i=1}^{\infty} \mathbb{P}(A_i), \end{aligned}$$

όπου η πρώτη και η τελευταία ισότητα προκύπτουν από τον ορισμό του \mathbb{P} , η δεύτερη και η πέμπτη από τον ορισμό της δεσμευμένης πιθανότητας, η τρίτη από την αντίστοιχη σχέση πράξεων συνόλων, και η τέταρτη από τον κανόνα πιθανότητας #3 (διότι τα $A_i \cap B$ είναι προφανώς ξένα ενδεχόμενα).

10. Συνέπειες ανεξαρτησίας.

(α') Το ότι τα A και B είναι ανεξάρτητα σημαίνει πως $\Pr(A \cap B) = \Pr(A) \Pr(B)$. Από την ανεξαρτησία και τον κανόνα πιθανότητας #4, έχουμε,

$$\begin{aligned} \Pr(A) \Pr(B') &= \Pr(A)(1 - \Pr(B)) \\ &= \Pr(A) - \Pr(A) \Pr(B) = \Pr(A) - \Pr(A \cap B). \end{aligned}$$

Επίσης, από την υπόδειξη της άσκησης έχουμε,

$$\Pr(A) = \Pr(A \cap B) + \Pr(A \cap B'),$$

και συνδυάζοντας τις πιο πάνω σχέσεις έχουμε,

$$\Pr(A \cap B') = \Pr(A \cap B) + \Pr(A \cap B') - \Pr(A \cap B) = \Pr(A) \Pr(B'),$$

συνεπώς τα A και B' είναι ανεξάρτητα.

(β') Η ανεξαρτησία των A' και B προκύπτει όμοια με το προηγούμενο σκέλος.

(γ') Αφού τα A και B είναι ανεξάρτητα, θα είναι και τα A' και B , από το δεύτερο σκέλος. Εφαρμόζοντας τώρα το πρώτο σκέλος για τα A' και B , προκύπτει ότι θα είναι ανεξάρτητα και τα A' και B' .

Όλα τα αποτελέσματα έχουν διαισθητική εξήγηση. Για παράδειγμα, σχετικά με το πρώτο αποτέλεσμα, αν τα A, B είναι ανεξάρτητα, τότε, αν μάθουμε ότι έγινε το B , δεν αλλάζει η πιθανότητα να έχει συμβεί το A . Άρα, αν μάθουμε ότι έγινε το B' , δηλαδή ότι δεν έγινε το B , πάλι δεν θα αλλάξει η πιθανότητα του A .

11. Ανεξαρτησία και δέσμευση.

(α') Από τον κανόνα πιθανότητας #4,

$$\Pr(A \cup B \cup C) = 1 - \Pr((A \cup B \cup C)') = 1 - \Pr(A' \cap B' \cap C'),$$

και χρησιμοποιώντας το αποτέλεσμα της Άσκησης 4 (α'),

$$\Pr(A \cup B \cup C) = 1 - \Pr(A'|B' \cap C') \Pr(B'|C') \Pr(C').$$

(β') Αν το A είναι ανεξάρτητο από τον εαυτό του, τότε η πιθανότητά του ικανοποιεί,

$$\Pr(A) = \Pr(A \cap A) = \Pr(A) \Pr(A) = (\Pr(A))^2,$$

και οι μόνοι πραγματικοί αριθμοί x τέτοιοι ώστε $x^2 = x$ είναι το 0 και το 1.

- (γ') Έστω ένα οποιοδήποτε ενδεχόμενο B . Αν $\Pr(A) = 0$, τότε, εφόσον $A \cap B \subset A$, από τον κανόνα πιθανότητας #2 έχουμε $\Pr(A \cap B) \leq \Pr(A) = 0$, συνεπώς $\Pr(A \cap B) = 0$. Άρα τα δύο ενδεχόμενα είναι ανεξάρτητα:

$$\Pr(A \cap B) = 0 = 0 \cdot \Pr(B) = \Pr(A) \Pr(B).$$

Παρομοίως, αν $\Pr(A) = 1$, τότε $\Pr(A') = 1 - \Pr(A) = 1 - 1 = 0$, και άρα τα A' και B είναι ανεξάρτητα, οπότε, από το αποτέλεσμα της Άσκησης 10, τα A και B είναι επίσης ανεξάρτητα.

12. Κι άλλη τράπουλα.

- (α') Όπως στους ανάλογους υπολογισμούς σε προβλήματα του Κεφαλαίου 4, εφαρμόζοντας τον κανόνα πιθανότητας #5 σε συνδυασμό με τους κανόνες αρίθμησης #1 και #4,

$$\Pr(\text{«4 άσοι και μία φιγούρα»}) = \frac{\binom{4}{4} \binom{12}{1}}{\binom{52}{5}} = \frac{12! 5! 47!}{1! 11! 52!} \approx 0.0000046.$$

- (β') Έστω τα ενδεχόμενα $A = \text{«4 άσοι και μία φιγούρα»}$ και $B = \text{«ακριβώς μία φιγούρα»}$. Από τον ορισμό της δεσμευμένης πιθανότητας παρομοίως έχουμε:

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A)}{\Pr(B)} = \frac{\frac{\binom{4}{4} \binom{12}{1}}{\binom{52}{5}}}{\frac{\binom{12}{1} \binom{40}{4}}{\binom{52}{5}}} \approx 0.000011.$$

- (γ') Αν ορίσουμε τον ενδεχόμενο $C = \text{«4 φιγούρες και ένας άσος»}$, όπως και στο σχέλος (α') βρίσκουμε,

$$\Pr(C) = \frac{\binom{12}{4} \binom{4}{1}}{\binom{52}{5}} = \frac{12! 4!}{4! 8! 3! 1!} \approx 0.0007618.$$

- (δ') Τα ενδεχόμενα δεν είναι ανεξάρτητα γιατί είναι αδύνατον να συμβούν μαζί, ενώ το καθένα έχει θετική πιθανότητα. Συγκεκριμένα,

$$\Pr(A \cap C) = \Pr(\emptyset) = 0 \neq \Pr(A) \times \Pr(C) = 0.0000046 \times 0.0007618 > 0.$$

13. *Κανόνας δεσμευμένης συνολικής πιθανότητας.* Με συνεχόμενες χρήσεις του ορισμού της δεσμευμένης πιθανότητας, έχουμε,

$$\begin{aligned}
 & \Pr(E|G \cap F) \Pr(G|F) + \Pr(E|G' \cap F) \Pr(G'|F) \\
 &= \frac{\Pr(E \cap G \cap F) \Pr(G \cap F)}{\Pr(G \cap F) \Pr(F)} + \frac{\Pr(E \cap G' \cap F) \Pr(G' \cap F)}{\Pr(G' \cap F) \Pr(F)} \\
 &= \frac{\Pr(E \cap G \cap F) + \Pr(E \cap G' \cap F)}{\Pr(F)} \\
 &= \frac{\Pr((E \cap G \cap F) \cup (E \cap G' \cap F))}{\Pr(F)} \\
 &= \frac{\Pr((E \cap F) \cap (G \cup G'))}{\Pr(F)} \\
 &= \frac{\Pr(E \cap F)}{\Pr(F)} = \Pr(E|F),
 \end{aligned}$$

όπου η πρώτη και η τελευταία ισότητα προκύπτουν από τον ορισμό της δεσμευμένης πιθανότητας και η τρίτη από τον κανόνα πιθανότητας #3, αφού τα ενδεχόμενα $(E \cap G \cap F)$ και $(E \cap G' \cap F)$ είναι ζένα.

14. *Μεσογειακή αναιμία.*

(α') Έστω A το ενδεχόμενο να βγει το αποτέλεσμα θετικό, και S να έχει το άτομο το στίγμα. Δίνεται ότι $\Pr(S) = 0.08$, $\Pr(A|S) = 0.1$ και $\Pr(A'|S) = 0.01$. Από τον κανόνα πιθανότητας #4, έχουμε επίσης $\Pr(S') = 0.92$, και, από το Λήμμα 5.1, ότι $\Pr(A|S) = 0.99$.

Από τον κανόνα του Bayes, η ζητούμενη πιθανότητα $\Pr(S|A)$ είναι:

$$\begin{aligned}
 \Pr(S|A) &= \frac{\Pr(A|S) \Pr(S)}{\Pr(A|S) \Pr(S) + \Pr(A|S') \Pr(S')} \\
 &= \frac{0.99 \times 0.08}{0.99 \times 0.08 + 0.1 \times 0.92} \\
 &\approx 0.4626.
 \end{aligned}$$

(β') Έστω A_1 το ενδεχόμενο να βγει το αποτέλεσμα θετικό στην πρώτη εξέταση, και A_2 το ενδεχόμενο να βγει το αποτέλεσμα θετικό στη δεύτερη εξέταση. Με εφαρμογή του κανόνα του Bayes και της υπόθεσης της ανεξαρτησίας, έχουμε:

$$\begin{aligned}
 \Pr(S|A_1 \cap A_2) &= \frac{\Pr(A_1 \cap A_2|S) \Pr(S)}{\Pr(A_1 \cap A_2|S) \Pr(S) + \Pr(A_1 \cap A_2|S') \Pr(S')} \\
 &= \frac{\Pr(A_1|S) \Pr(A_2|S) \Pr(S)}{\Pr(A_1|S) \Pr(A_2|S) \Pr(S) + \Pr(A_1|S') \Pr(A_2|S') \Pr(S')} \\
 &= \frac{0.99 \times 0.01 \times 0.08}{0.99 \times 0.01 \times 0.08 + 0.1 \times 0.9 \times 0.92} \approx 0.0095.
 \end{aligned}$$

15. *Ανεπιθύμητες εγκυμοσύνες.* Ορίζουμε τα ενδεχόμενα: E = «Υπάρχει εγκυμοσύνη» και Θ = «Το τεστ είναι θετικό». Από τα δεδομένα του προβλήματος έχουμε $\Pr(E) = 0.12$, $\Pr(\Theta|E') = 0.01$ και $\Pr(\Theta|E) = 0.03$. Από τον κανόνα πιθανότητας #4 έχουμε ότι, $\Pr(E') = 0.88$, και από το Λήμμα 5.1, $\Pr(\Theta|E) = 0.97$

(α) Από τον κανόνα του Bayes και τα πιο πάνω δεδομένα:

$$\begin{aligned}\Pr(E|\Theta) &= \frac{\Pr(\Theta|E) \Pr(E)}{\Pr(\Theta|E) \Pr(E) + \Pr(\Theta|E') \Pr(E')} \\ &= \frac{0.97 \times 0.12}{0.97 \times 0.12 + 0.01 \times 0.88} \\ &\approx 92.97\%.\end{aligned}$$

(β) Έστω Θ_1, Θ_2 το ενδεχόμενο να βγει το αποτέλεσμα θετικό στην πρώτη ή στη δεύτερη εξέταση, αντίστοιχα. Από τον κανόνα του Bayes και την υπόθεση της ανεξαρτησίας, έχουμε:

$$\begin{aligned}\Pr(E|\Theta_1 \cap \Theta_2) &= \frac{\Pr(\Theta_1 \cap \Theta_2|E) \Pr(E)}{\Pr(\Theta_1 \cap \Theta_2|E) \Pr(E) + \Pr(\Theta_1 \cap \Theta_2|E') \Pr(E')} \\ &= \frac{97\% \times 3\% \times 12\%}{97\% \times 3\% \times 12\% + 1\% \times 99\% \times 88\%} \\ &\approx 28.613\%.\end{aligned}$$

16. *Απάτη.* Έστω τα ενδεχόμενα: Δ = «το ζάρι είναι δίκαιο» και A = «φέραμε 1,5,2,1,1,2». Γνωρίζουμε ότι $\Pr(\Delta) = \Pr(\Delta') = 1/2$, και επίσης ότι για το δίκαιο ζάρι έχουμε,

$$\Pr(\text{αποτέλεσμα } i|\Delta) = 1/6, \quad \text{για κάθε } i = 1, 2, \dots, 6,$$

ενώ για το κάλπικο ζάρι, $\Pr(\text{αποτέλεσμα } 1|\Delta') = 1/3$, $\Pr(\text{αποτέλεσμα } 6|\Delta') = 0$, και $\Pr(\text{αποτέλεσμα } i|\Delta') = 1/6$ για τα $i = 2, 3, 4, 5$.

Χρησιμοποιώντας αυτά τα δεδομένα μαζί με τον κανόνα του Bayes:

$$\begin{aligned}\Pr(\Delta|A) &= \frac{\Pr(A|\Delta) \Pr(\Delta)}{\Pr(A|\Delta) \Pr(\Delta) + \Pr(A|\Delta') \Pr(\Delta')} \\ &= \frac{\frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} \times \frac{1}{2}}{\frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} \times \frac{1}{2} + \frac{1}{3} \times \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} \times \frac{1}{2}} \\ &= \frac{1}{9}.\end{aligned}$$

Α'.6 Ασκήσεις Κεφαλαίου 6

1. ★ *Μέση τιμή συναρτήσεων T.M.* Έστω μια T.M. X με πυκνότητα $P(x)$ στο σύνολο τιμών S , και έστω μια συνάρτηση $f : S \rightarrow S'$. Τότε η νέα T.M. $Y = f(X)$ έχει σύνολο τιμών το S' και η πυκνότητά της, έστω $Q(y)$, έχει την εξής ιδιότητα: Για κάθε $y \in S'$ υπάρχει ένα ή περισσότερα x τέτοια ώστε $f(x) = y$. Έστω A_y το σύνολο αυτών των x :

$$A_y = \{x \in S : f(x) = y\}.$$

Τα σύνολα A_y είναι ξένα μεταξύ τους, και ικανοποιούν,

$$S = \bigcup_{y \in S'} A_y.$$

Συνεπώς, η πυκνότητα $Q(y)$ μπορεί να εκφραστεί ως,

$$Q(y) = \Pr(Y = y) = \Pr(f(X) = y) = \Pr(X \in A_y) = \sum_{x \in A_y} P(x).$$

Αν εφαρμόσουμε τον τύπο του ορισμού (6.4) για τη μέση τιμή της Y έχουμε,

$$\begin{aligned} E(Y) &= \sum_{y \in S'} y Q(y) = \sum_{y \in S'} \left[y \sum_{x \in A_y} P(x) \right] \\ &= \sum_{y \in S'} \sum_{x \in A_y} f(x) P(x) = \sum_{x \in S} f(x) P(x), \end{aligned}$$

που ισούται με τον γενικό τύπο (6.5) για τη μέση τιμή $E[f(X)]$ της $f(X)$.

2. ★ *H συνάρτηση κατανομής στο $\pm\infty$.* Έστω πως η T.M. X έχει συνάρτηση κατανομής $F(x)$ και είναι ορισμένη στο χώρο πιθανότητας Ω , στον οποίο έχουμε επίσης ορισμένο το μέτρο πιθανότητας \mathbb{P} .

Ορίζουμε τα ενδεχόμενα $A_n = \{X \leq n\}$ για $n = 1, 2, \dots$ και παρατηρούμε ότι $A_n \subset A_{n+1}$ για κάθε n και ότι, φυσικά, $F(n) = \mathbb{P}(A_n)$. Επιπλέον, χρησιμοποιώντας την ορολογία του ορίου μιας ακολουθίας συνόλων όπως στην Άσκηση 9 του Κεφαλαίου 3, η ένωση των A_n ισούται με,

$$A = \bigcup_{n=1}^{\infty} A_n = \lim_{n \rightarrow \infty} A_n = \{X < \infty\} = \Omega.$$

Η τελευταία ισότητα προκύπτει αμέσως από τον ορισμό μιας T.M. και η προτελευταία είναι συνέπεια του ορισμού των A_n . Οπότε, από το αποτέλεσμα του πρώτου σκέλους της Άσκησης 9 του Κεφαλαίου 3, έχουμε ότι,

$$\lim_{n \rightarrow \infty} \mathbb{P}(A_n) = \mathbb{P}\left(\lim_{n \rightarrow \infty} A_n\right),$$

δηλαδή,

$$\lim_{n \rightarrow \infty} F(n) = \mathbb{P}(\Omega) = 1. \quad (\text{A}'8)$$

Τέλος παρατηρούμε ότι, εφόσον η $F(x)$ είναι αύξουσα και άνω φραγμένη, $F(x) \leq 1$, το όριό της καθώς το $x \rightarrow \infty$ υπάρχει και ισούται με το όριό της κατά μήκος οποιασδήποτε ακολουθίας $\{x_n\}$ πραγματικών αριθμών που τείνουν στο $+\infty$. Αυτό, σε συνδυασμό με το αποτέλεσμα (A'8), συνεπάγεται το πρώτο αποτέλεσμα της άσκησης: $\lim_{x \rightarrow \infty} F(x) = 1$.

Η απόδειξη του δεύτερου ορίου, καθώς το $x \rightarrow -\infty$, είναι ακριβώς ανάλογη. Ορίζουμε τώρα τα ενδεχόμενα $B_n = \{X \leq -n\}$ για $n = 1, 2, \dots$, παρατηρούμε ότι $B_n \supset B_{n+1}$ για κάθε n και ότι $F(-n) = \mathbb{P}(B_n)$. Οπότε,

$$B = \bigcap_{n=1}^{\infty} B_n = \lim_{n \rightarrow \infty} A_n = \emptyset,$$

όπου η τελευταία ισότητα προκύπτει από το συλλογισμό πως, αν κάποιο $\omega \in B$, τότε θα πρέπει να έχουμε $X(\omega) \leq n$ για κάθε n , πράγμα αδύνατο. Συνεπώς, από το αποτέλεσμα του δεύτερου σκέλους της Άσκησης 9 του Κεφαλαίου 3, έχουμε ότι,

$$\lim_{n \rightarrow \infty} F(n) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n) = \mathbb{P}(\emptyset) = 0. \quad (\text{A}'9)$$

Τέλος, παρατηρούμε ότι, εφόσον η $F(x)$ είναι αύξουσα και κάτω φραγμένη, $F(x) \geq 0$, το όριό της καθώς το $x \rightarrow -\infty$ υπάρχει και ισούται με το όριό της κατά μήκος οποιασδήποτε πραγματικής ακολουθίας $\{x_n\}$ που τείνει στο $-\infty$. Αυτό, σε συνδυασμό με το αποτέλεσμα (A'9), συνεπάγεται το δεύτερο αποτέλεσμα της άσκησης: $\lim_{x \rightarrow -\infty} F(x) = 0$.

3. *Ντετερμινιστικές T.M.* Έστω μια ντετερμινιστική T.M. X με σύνολο τιμών $S_X = \{a\}$, οπότε $\Pr(X = a) = 1$, και μια οποιαδήποτε διακριτή T.M. Y με σύνολο τιμών S_Y . Η επιβεβαίωση του ότι ικανοποιείται ο ορισμός της ανεξαρτησίας για τις X, Y είναι τετριμμένη διότι, προφανώς, το ενδεχόμενο $\{X = a, Y = b\}$ ισούται με το $\{Y = b\}$ για οποιοδήποτε $b \in S_Y$. Άρα, πράγματι:

$$\Pr(X = a, Y = b) = \Pr(Y = b) = \Pr(X = a) \Pr(Y = b).$$

4. *Άπειρη μέση τιμή.*

(α') Από τη δεύτερη βασική ιδιότητα της πυκνότητας ξέρουμε πως το άθροισμα όλων των τιμών της πρέπει να ισούται με 1, συνεπώς, πρέπει να έχουμε,

$$1 = \sum_{k=1}^{\infty} P(k) = \sum_{k=1}^{\infty} \frac{C}{k^2} = C \sum_{k=1}^{\infty} \frac{1}{k^2} = C \cdot \frac{\pi^2}{6},$$

και άρα $C = 6/\pi^2 \approx 0.6079271 \dots$.

(β') Η μέση τιμή της X είναι,

$$\mu = E(X) = \sum_{k=1}^{\infty} k P(k) = \sum_{k=1}^{\infty} k \frac{6}{\pi^2} \frac{1}{k^2} = \frac{6}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{k},$$

το οποίο ισούται με $+\infty$ αφού, ως γνωστόν, η αρμονική σειρά $\sum_{k=1}^{\infty} (1/k)$ αποκλίνει.

5. ★ *Ανεξαρτησία και μέση τιμή.* Έστω πως οι τυχαίες μεταβλητές (X, Y) παίρνουν καθένα από τα 4 ζευγάρια τιμών $(0, 1), (0, -1), (1, 0), (-1, 0)$ με πιθανότητα $1/4$. Τότε η X έχει σύνολο τιμών το $\{-1, 0, 1\}$ και η πυκνότητά της παίρνει τις εξής τιμές:

$$\begin{aligned} P(-1) &= \Pr(X = -1) = \Pr((X, Y) = (-1, 0)) = 1/4, \\ P(0) &= \Pr(X = 0) = \Pr((X, Y) = (0, 1) \text{ ή } (X, Y) = (0, -1)) \\ &= \Pr((X, Y) = (0, 1)) + \Pr((X, Y) = (0, -1)) = 1/4 + 1/4 = 1/2, \\ P(1) &= \Pr(X = 1) = \Pr((X, Y) = (1, 0)) = 1/4. \end{aligned}$$

Λόγω της συμμετρίας του προβλήματος, η Y έχει το ίδιο σύνολο τιμών και την ίδια πυκνότητα με τη X . Αλλά οι X, Y δεν είναι ανεξάρτητες, αφού,

$$\Pr(X = 0, Y = 0) = 0 \neq (1/2)(1/2) = \Pr(X = 0) \Pr(Y = 0).$$

Για τις μέσες τιμές έχουμε,

$$E(X) = E(Y) = (-1) \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} = 0,$$

και επίσης,

$$E(XY) = (0 \cdot 1) \frac{1}{4} + (0 \cdot (-1)) \frac{1}{4} + (1 \cdot 0) \frac{1}{4} + ((-1) \cdot 0) \frac{1}{4} = 0,$$

άρα, παρότι οι X, Y δεν είναι ανεξάρτητες, έχουμε $E(XY) = E(X)E(Y) = 0$.

6. *Ανεξαρτησία και διασπορά.* Έστω μια Τ.Μ. X η οποία παίρνει τις τιμές 0 και 1 με πιθανότητα $1/2$ και έστω $Y = -X$, οπότε και η Y έχει το ίδιο σύνολο τιμών και την ίδια πυκνότητα με τη X . Όπως είδαμε στο Παράδειγμα 6.7, η διασπορά τους ικανοποιεί,

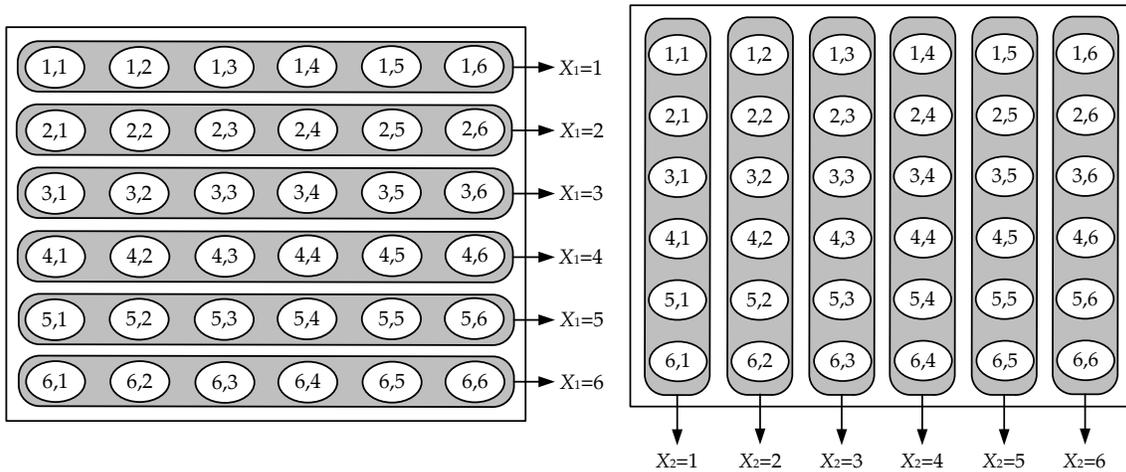
$$\text{Var}(X) = \text{Var}(Y) = (1/2)(1/2) = 1/4 > 0.$$

Αλλά, εφόσον εξ ορισμού η $X + Y = 0$ είναι απλώς μια σταθερά, έχουμε $\text{Var}(X + Y) = 0$, το οποίο φυσικά δεν ισούται με το άθροισμα των διασπορών $\text{Var}(X) + \text{Var}(Y) = 1/2!$

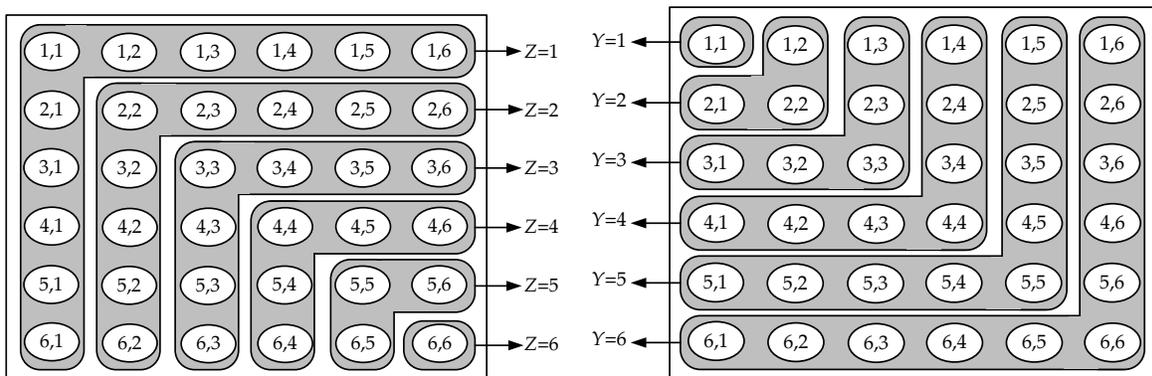
7. *Δύο ζάρια.* Κατ' αρχάς παρατηρούμε το εξής: Το γεγονός ότι οι δύο ζαριές είναι ανεξάρτητες και τα ζάρια δίκαια, είναι προφανώς ισοδύναμο με το να υποθέσουμε ότι όλα τα στοιχειώδη ενδεχόμενα είναι ισοπίθανα.

(α') Και οι τέσσερις Τ.Μ. έχουν, εξ ορισμού, το ίδιο σύνολο τιμών $S = \{1, 2, 3, 4, 5, 6\}$.

(β') Στα Σχήματα Α'.17 και Α'.18, έχουμε σχεδιάσει τα ενδεχόμενα που αντιστοιχούν σε καθεμία από τις δυνατές τιμές της X_1 , της X_2 , της Y και της Z , αντίστοιχα.



Σχήμα Α΄.17: Ο χώρος πιθανότητας Ω , τα ενδεχόμενα $\{X_1 = x\}$ που αντιστοιχούν στην καθεμία από τις δυνατές τιμές x της Τ.Μ. X_1 (αριστερά), και τα ενδεχόμενα $\{X_2 = x\}$ που αντιστοιχούν στην καθεμία από τις δυνατές τιμές x της Τ.Μ. X_2 (δεξιά) στην Άσκηση 7.



Σχήμα Α΄.18: Ο χώρος πιθανότητας Ω , τα ενδεχόμενα $\{Z = z\}$ που αντιστοιχούν στην καθεμία από τις δυνατές τιμές z της Τ.Μ. Z (αριστερά), και τα ενδεχόμενα $\{Y = y\}$ που αντιστοιχούν στην καθεμία από τις δυνατές τιμές y της Τ.Μ. Y (δεξιά) στην Άσκηση 7.

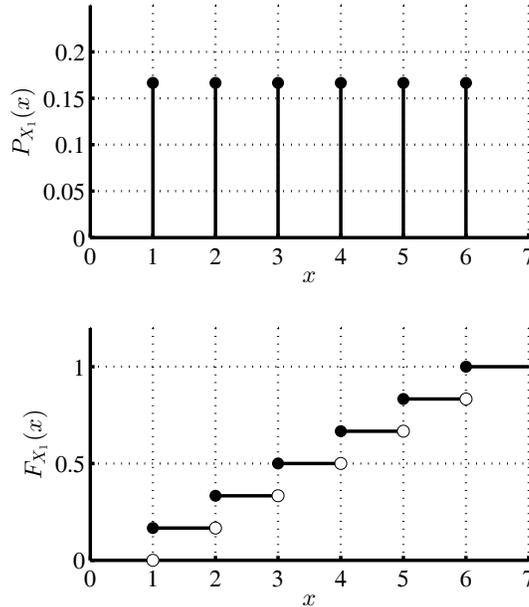
(γ) Για κάθε δυνατή τιμή $x = 1, 2, 3, 4, 5, 6$ της X_1 , υπάρχουν 6 στοιχεία του Ω που αντιστοιχούν σε αυτή την τιμή, και άρα η πιθανότητα του ενδεχομένου $\{X_1 = x\}$, από τον κανόνα πιθανότητας #5, είναι $6/36 = 1/6$. Συνεπώς, η πυκνότητα της X_1 είναι,

$$P_{X_1}(x) = \frac{1}{6}, \quad \text{για κάθε } x = 1, 2, \dots, 6.$$

Η συνάρτηση κατανομής $F_{X_1}(x)$ της X_1 εύκολα μπορεί να υπολογιστεί από τη σχέση (6.1):

$$F_{X_1}(x) = \sum_{i \in S : i \leq x} P_{X_1}(x) = \frac{\#\{i \in S : i \leq x\}}{6}.$$

Η πυκνότητα και η συνάρτηση κατανομής της X_1 έχουν σχεδιαστεί στο Σχήμα Α.19.



Σχήμα Α.19: Η πυκνότητα και η συνάρτηση κατανομής της Τ.Μ. X_1 στην Άσκηση 7.

Με ακριβώς τον ίδιο τρόπο βρίσκουμε πως η X_2 έχει την ίδια πυκνότητα και την ίδια συνάρτηση κατανομής με τη X_1 .

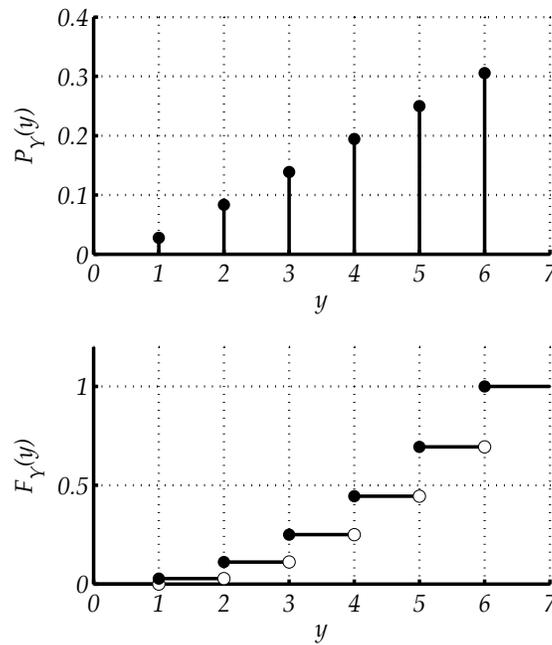
Η πυκνότητα της $Y = \max\{X_1, X_2\}$ υπολογίζεται εύκολα από τον κανόνα πιθανότητας #5, μέσω της σχέσης,

$$P_Y(y) = \frac{1}{36} \cdot \#\{Y = y\}.$$

Με τη βοήθεια του Σχήματος Α.18 βρίσκουμε:

$$\begin{aligned} P_Y(1) &= \frac{1}{36}, & P_Y(2) &= \frac{3}{36}, & P_Y(3) &= \frac{5}{36}, \\ P_Y(4) &= \frac{7}{36}, & P_Y(5) &= \frac{9}{36}, & P_Y(6) &= \frac{11}{36}, \end{aligned}$$

ενώ η συνάρτηση κατανομής προκύπτει από τη σχέση (6.1) όπως και για τη X_1 . Στο Σχήμα Α.20 έχουμε σχεδιάσει και τις δύο.



Σχήμα Α΄.20: Η πυκνότητα και η συνάρτηση κατανομής της Τ.Μ. Y στην Άσκηση 7.

Παρομοίως υπολογίζεται και η πυκνότητα της $Z = \min\{X_1, X_2\}$, από τον κανόνα πιθανότητας #5 και τη σχέση,

$$P_Z(z) = \frac{1}{36} \cdot \#\{Z = z\}.$$

Με τη βοήθεια του Σχήματος Α΄.18 βρίσκουμε,

$$\begin{aligned} P_Z(1) &= \frac{11}{36}, & P_Z(2) &= \frac{9}{36}, & P_Z(3) &= \frac{7}{36}, \\ P_Z(4) &= \frac{5}{36}, & P_Z(5) &= \frac{3}{36}, & P_Z(6) &= \frac{1}{36}. \end{aligned}$$

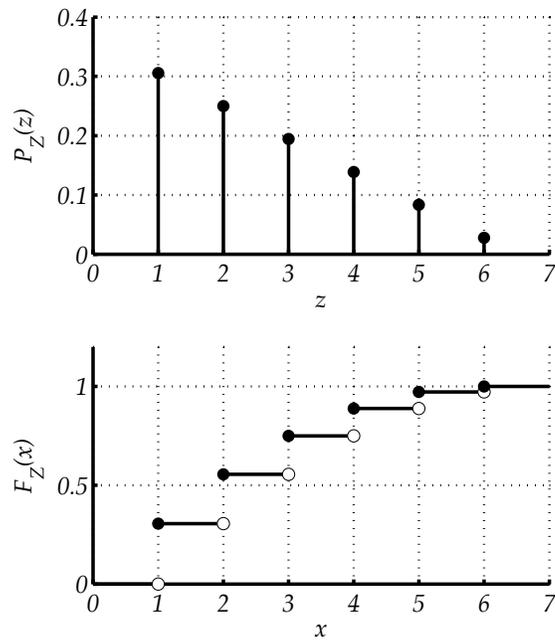
Στο Σχήμα Α΄.21 έχουν σχεδιαστεί η $P_Z(z)$ και η $F_Z(x)$.

(δ΄) Από τον ορισμό της μέσης τιμής, για τη X_1 έχουμε,

$$E(X_1) = \frac{1}{6} \times 1 + \frac{1}{6} \times 2 + \frac{1}{6} \times 3 + \frac{1}{6} \times 4 + \frac{1}{6} \times 5 + \frac{1}{6} \times 6 = \frac{21}{6}.$$

Για τη διασπορά θα χρησιμοποιήσουμε την εναλλακτική έκφραση (6.8), οπότε υπολογίζουμε πρώτα τη μέση τιμή της X_1^2 ,

$$E(X_1^2) = \frac{1}{6} \times 1^2 + \frac{1}{6} \times 2^2 + \frac{1}{6} \times 3^2 + \frac{1}{6} \times 4^2 + \frac{1}{6} \times 5^2 + \frac{1}{6} \times 6^2 = \frac{91}{6},$$



Σχήμα Α.21: Άσκηση 7: Η πυκνότητά και η συνάρτηση κατανομής της Τ.Μ. Z .

βρίσκουμε πως,

$$\text{Var}(X_1) = E(X_1^2) - (E(X_1))^2 = \frac{91}{6} - \left(\frac{21}{6}\right)^2 = \frac{35}{12}.$$

Εφόσον η X_2 έχει την ίδια πυκνότητα με την X_1 , προφανώς έχει την ίδια μέση τιμή και την ίδια διασπορά.

Με τον ίδιο τρόπο, για τη μέση τιμή της Y και της Y^2 έχουμε,

$$\begin{aligned} E(Y) &= \frac{1}{36} \times 1 + \frac{3}{36} \times 2 + \frac{5}{36} \times 3 + \frac{7}{36} \times 4 + \frac{9}{36} \times 5 + \frac{11}{36} \times 6 \\ &= \frac{161}{36}, \\ E(Y^2) &= \frac{1}{36} \times 1^2 + \frac{3}{36} \times 2^2 + \frac{5}{36} \times 3^2 + \frac{7}{36} \times 4^2 + \frac{9}{36} \times 5^2 + \frac{11}{36} \times 6^2 \\ &= \frac{791}{36}, \end{aligned}$$

οπότε η διασπορά της Y είναι,

$$\text{Var}(Y) = E(Y^2) - (E(Y))^2 = \frac{791}{36} - \left(\frac{161}{36}\right)^2 = \frac{2555}{1296} \approx 1.97.$$

Τέλος, παρομοίως υπολογίζονται η μέση τιμή και η διασπορά της Z :

$$\begin{aligned} E(Z) &= \frac{11}{36} \times 1 + \frac{9}{36} \times 2 + \frac{7}{36} \times 3 + \frac{5}{36} \times 4 + \frac{3}{36} \times 5 + \frac{1}{36} \times 6 \\ &= \frac{91}{36}, \\ E(Z^2) &= \frac{11}{36} \times 1^2 + \frac{9}{36} \times 2^2 + \frac{7}{36} \times 3^2 + \frac{5}{36} \times 4^2 + \frac{3}{36} \times 5^2 + \frac{1}{36} \times 6^2 \\ &= \frac{301}{36}, \\ \text{Var}(Z) &= E(Z^2) - (E(Z))^2 = \frac{301}{36} - \left(\frac{91}{36}\right)^2 = \frac{2555}{1296} \approx 1.97. \end{aligned}$$

Παρατηρήστε πως, παρότι η Y και η Z έχουν διαφορετική μέση τιμή, και επιπλέον $E(Y^2) \neq E(Z^2)$, έχουν την ίδια διασπορά. Μπορείτε να το εξηγήσετε διαισθητικά;

8. ★ Πού να βρω γυναίκα να σου μοιάζει. Χρησιμοποιώντας τις τυχαίες μεταβλητές της υπόδειξης έχουμε ότι $X = X_1 + X_2 + \dots + X_N$. Επιπλέον, λόγω της συμμετρίας του προβλήματος, ισχυριζόμαστε πως η πιθανότητα οποιαδήποτε γυναίκα i να επιλέξει τον άντρα της είναι $\frac{1}{N}$. Για την πρώτη γυναίκα αυτό είναι προφανές. Γιατί ισχύει και για τις υπόλοιπες; Ένας τρόπος να πεισθούμε γι' αυτό είναι να σκεφτούμε πως, αντί με τη σειρά οι γυναίκες να επιλέγουν, η μία μετά την άλλη, έναν άντρα, ισοδύναμα βάζουμε τους άντρες σε μια τυχαία διάταξη απέναντι από τις γυναίκες και η καθεμία παίρνει αυτόν που είναι μπροστά της. Προφανώς, για την κάθε γυναίκα i , η πιθανότητα απέναντί της να είναι ο άντρας της είναι ίση με $1/N$.

Συνεπώς, όπως στο Παράδειγμα 6.7, η μέση τιμή της κάθε X_i είναι,

$$E(X_i) = 1 \times (1/N) + 0 \times (1 - 1/N) = 1/N,$$

και από την πρώτη ιδιότητα του Θεωρήματος 6.1,

$$E(X) = E(X_1 + X_2 + \dots + X_N) = E(X_1) + E(X_2) + \dots + E(X_N) = N \times \frac{1}{N} = 1.$$

Άρα, κατά μέσο όρο, μόνο ένα σωστό ζευγάρι θα φύγει από το πάρτυ!

Για να υπολογίσουμε τη διασπορά παρατηρούμε κατ' αρχάς πως, εφόσον οι X_i είναι δυαδικές Τ.Μ., πάντοτε έχουμε $X_i^2 = X_i$ και συνεπώς, $E(X_i^2) = E(X_i) = 1/N$. Επιπλέον, από τον ορισμό της δεσμευμένης πιθανότητας,

$$\begin{aligned} E(X_1 X_2) &= \Pr(X_1 = X_2 = 1) \cdot 1 + (1 - \Pr(X_1 = X_2 = 1)) \cdot 0 \\ &= \Pr(X_2 = 1 \text{ και } X_1 = 1) \\ &= \Pr(X_1 = 1) \Pr(X_2 = 1 | X_1 = 1) \\ &= \frac{1}{N} \frac{1}{N-1}. \end{aligned}$$

Ισχυριζόμαστε πως το ίδιο ισχύει για οποιαδήποτε $i \neq j$:

$$E(X_i X_j) = \frac{1}{N} \frac{1}{N-1}.$$

Αν σκεφτούμε, όπως πριν, ότι οι άντρες τοποθετούνται σε μια τυχαία διάταξη απέναντι από τις γυναίκες και η καθεμία παίρνει αυτόν που είναι μπροστά της, τότε η πιθανότητα οι γυναίκες i και j να έχουν απέναντί τους τους άντρες τους είναι $(\frac{1}{N})(\frac{1}{N-1})$. Οπότε,

$$E(X_i X_j) = \Pr(X_i = 1, X_j = 1) = \frac{1}{N} \frac{1}{N-1}.$$

Τώρα είμαστε έτοιμοι να υπολογίσουμε τη διασπορά της X . Για τη μέση τιμή της X^2 παρατηρούμε πως,

$$\begin{aligned} E(X^2) &= E\left(\left(\sum_{i=1}^N X_i\right)^2\right) \\ &= E\left(\left(\sum_{i=1}^N X_i\right)\left(\sum_{j=1}^N X_j\right)\right) \\ &= E\left(\sum_{i=1}^N \sum_{j=1}^N X_i X_j\right) \\ &= \sum_{i=1}^N \sum_{j=1}^N E(X_i X_j), \end{aligned}$$

όπου και πάλι χρησιμοποιήσαμε την πρώτη ιδιότητα του Θεωρήματος 6.1. Στο τελευταίο παραπάνω διπλό άθροισμα υπάρχουν N όροι της μορφής $E(X_i^2)$ και $N(N-1)$ όροι της μορφής $E(X_i X_j)$ για $i \neq j$. Άρα, χρησιμοποιώντας τα αποτελέσματα που βρήκαμε πιο πάνω,

$$E(X^2) = E\left(\left(\sum_{i=1}^N X_i\right)^2\right) = N(N-1)\frac{1}{N(N-1)} + N\frac{1}{N} = 2,$$

και από την έκφραση (6.8) για τη διασπορά έχουμε,

$$\text{Var}(X) = E(X^2) - (E(X))^2 = 1.$$

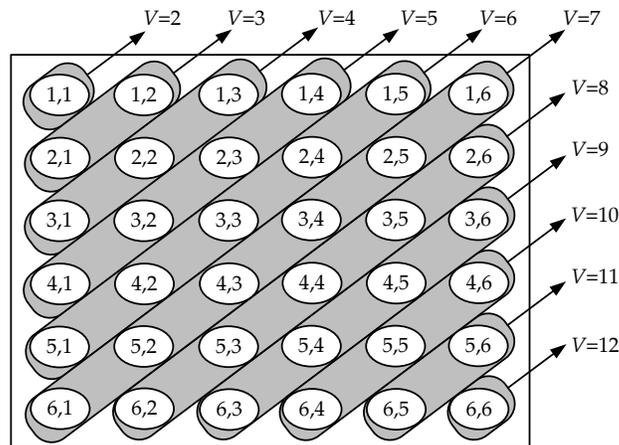
Σημείωση. Οι πιο πάνω υπολογισμοί μπορούν να θεωρηθούν ειδικές περιπτώσεις των αντίστοιχων υπολογισμών που θα κάνουμε στο επόμενο κεφάλαιο, και συγκεκριμένα στην Ενότητα 7.2, για τη μέση τιμή και τη διασπορά της λεγόμενης υπεργεωμετρικής κατανομής.

9. *Συνέλιξη.* Παρατηρούμε πως, για οποιοδήποτε τιμή $m \in \mathbb{Z}$, η $X + Y$ ισούται με m αν και μόνο αν η $X = k$ και η $Y = m - k$ για κάποιο $k \in \mathbb{Z}$. Συνεπώς:

$$\begin{aligned}
 P_{X+Y}(m) &= \Pr(X + Y = m) \\
 &= \Pr\left(\bigcup_{k=-\infty}^{\infty} (\{X = k\} \cap \{Y = m - k\})\right) \\
 &= \sum_{k=-\infty}^{\infty} \Pr(\{X = k\} \cap \{Y = m - k\}) \\
 &= \sum_{k=-\infty}^{\infty} \Pr(\{X = k\}) \Pr(\{Y = m - k\}) \\
 &= \sum_{k=-\infty}^{\infty} P_X(k) P_Y(m - k).
 \end{aligned}$$

Η πρώτη και η τελευταία ισότητα προκύπτουν από τον ορισμό της πυκνότητας, η δεύτερη προκύπτει γράφοντας το ενδεχόμενο $\{X + Y = m\}$ σαν ένωση ξένων ενδεχόμενων, η τρίτη προκύπτει ακριβώς επειδή τα ενδεχόμενα είναι ξένα, ενώ η τέταρτη λόγω της ανεξαρτησίας των X, Y .

10. *Άλλα δύο ζάρια.* Για τη V , αρχικά παρατηρούμε πως το σύνολο τιμών της είναι το $S_V = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$, ενώ τα ενδεχόμενα $\{V = v\}$ στα οποία αντιστοιχεί η κάθε τιμή $v \in S_V$ φαίνονται στο Σχήμα Α΄.22.



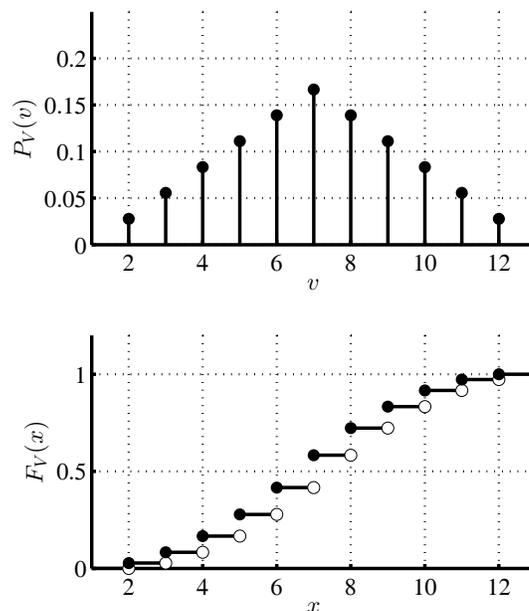
Σχήμα Α΄.22: Ο χώρος πιθανότητας Ω και τα ενδεχόμενα $\{V = v\}$ που αντιστοιχούν στην καθένα από τις δυνατές τιμές v της Τ.Μ. V στην Άσκηση 10.

Όπως και στην Άσκηση 7, λαμβάνοντας υπόψη ότι τα αποτελέσματα είναι ισοπίθανα, εύκολα προκύπτει πως:

$$P_V(2) = \frac{1}{36}, \quad P_V(3) = \frac{2}{36}, \quad P_V(4) = \frac{3}{36}, \quad P_V(5) = \frac{4}{36}, \quad P_V(6) = \frac{5}{36}, \quad P_V(7) = \frac{6}{36},$$

$$P_V(8) = \frac{5}{36}, \quad P_V(9) = \frac{4}{36}, \quad P_V(10) = \frac{3}{36}, \quad P_V(11) = \frac{2}{36}, \quad P_V(12) = \frac{1}{36}.$$

Επιπλέον, η συνάρτηση κατανομής υπολογίζεται απλά από την πυκνότητα, μέσω της σχέσης (6.1). Και οι δύο έχουν σχεδιαστεί στο Σχήμα A'.23.



Σχήμα A'.23: Η πυκνότητα και η συνάρτηση κατανομής της T.M. V στην Άσκηση 10.

Για τη μέση τιμή και τη διασπορά της V , από τους ορισμούς και την έκφραση (6.8),

$$E(V) = \frac{1}{36} \times 2 + \frac{2}{36} \times 3 + \frac{3}{36} \times 4 + \frac{4}{36} \times 5 + \frac{5}{36} \times 6 + \frac{6}{36} \times 7$$

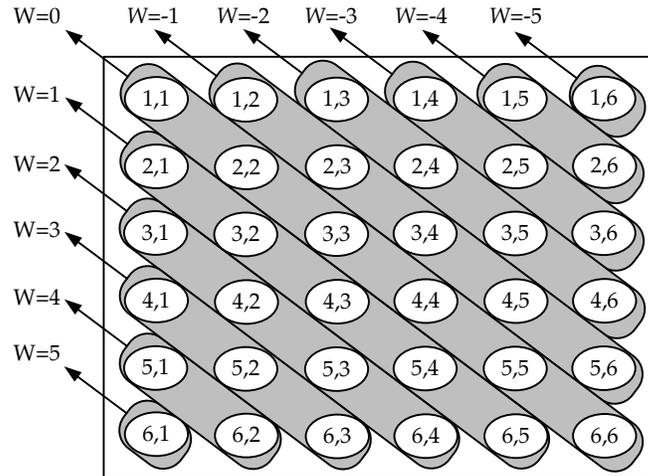
$$+ \frac{5}{36} \times 8 + \frac{4}{36} \times 9 + \frac{3}{36} \times 10 + \frac{2}{36} \times 11 + \frac{1}{36} \times 12 = 7,$$

$$\text{και, } E(V^2) = \frac{1}{36} \times 2^2 + \frac{2}{36} \times 3^2 + \frac{3}{36} \times 4^2 + \frac{4}{36} \times 5^2 + \frac{5}{36} \times 6^2 + \frac{6}{36} \times 7^2$$

$$+ \frac{5}{36} \times 8^2 + \frac{4}{36} \times 9^2 + \frac{3}{36} \times 10^2 + \frac{2}{36} \times 11^2 + \frac{1}{36} \times 12^2 = \frac{329}{6},$$

$$\text{οπότε, } \text{Var}(V) = E(V^2) - (E(V))^2 = \frac{329}{6} - 7^2 = \frac{35}{6}.$$

Παρομοίως εξετάζουμε και την Τ.Μ. W : Παρατηρούμε κατ' αρχάς πως το σύνολο τιμών της είναι το $S_W = \{-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5\}$ και τα ενδεχόμενα $\{W = w\}$ που αντιστοιχούν στην κάθε τιμή $w \in S_W$ φαίνονται στο Σχήμα Α΄.24.



Σχήμα Α΄.24: Ο χώρος πιθανότητας Ω και τα ενδεχόμενα $\{W = w\}$ που αντιστοιχούν στην καθεμία από τις δυνατές τιμές w της Τ.Μ. W στην Άσκηση 10.

Σημειώνοντας και πάλι πως όλα αποτελέσματα είναι ισοπίθανα, εύκολα βρίσκουμε:

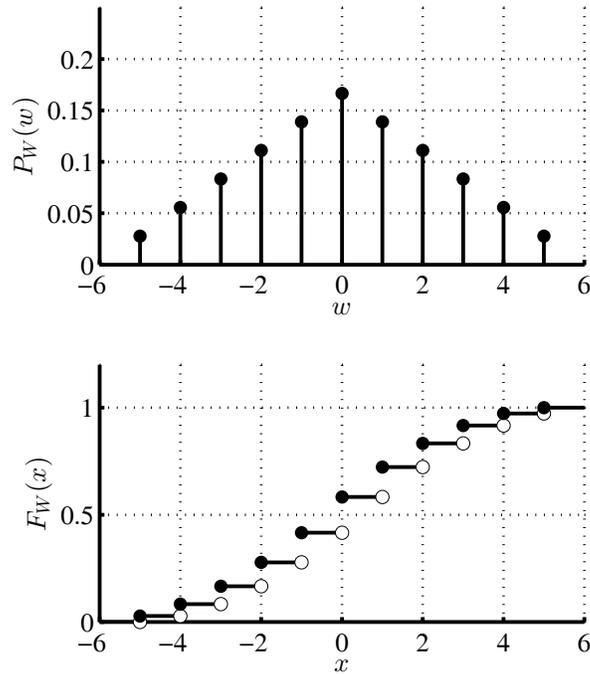
$$P_W(-5) = \frac{1}{36}, P_W(-4) = \frac{2}{36}, P_W(-3) = \frac{3}{36}, P_W(-2) = \frac{4}{36}, P_W(-1) = \frac{5}{36},$$

$$P_W(0) = \frac{6}{36}, P_W(1) = \frac{5}{36}, P_W(2) = \frac{4}{36}, P_W(3) = \frac{3}{36}, P_W(4) = \frac{2}{36}, P_W(5) = \frac{1}{36}.$$

Όπως και για τη V , η συνάρτηση κατανομής της W προκύπτει απλά από την πυκνότητα, μέσω της σχέσης (6.1). Και οι δύο έχουν σχεδιαστεί στο Σχήμα Α΄.25.

Τέλος, για τη μέση τιμή και τη διασπορά της W , από τον ορισμό της μέσης τιμής και την εναλλακτική έκφραση (6.8) για τη διασπορά, έχουμε,

$$\begin{aligned} E(W) &= \frac{1}{36} \times (-5) + \frac{2}{36} \times (-4) + \frac{3}{36} \times (-3) + \frac{4}{36} \times (-2) + \frac{5}{36} \times (-1) \\ &\quad + \frac{6}{36} \times 0 + \frac{5}{36} \times 1 + \frac{4}{36} \times 2 + \frac{3}{36} \times 3 + \frac{2}{36} \times 4 + \frac{1}{36} \times 5 \\ &= 0, \end{aligned}$$



Σχήμα Α'.25: Άσκηση 10: Η πυκνότητα και η συνάρτηση κατανομής της Τ.Μ. W .

και παρομοίως,

$$\begin{aligned}
 E(W^2) &= \frac{1}{36} \times (-5)^2 + \frac{2}{36} \times (-4)^2 + \frac{3}{36} \times (-3)^2 + \frac{4}{36} \times (-2)^2 + \frac{5}{36} \times (-1)^2 \\
 &\quad + \frac{6}{36} \times 0^2 + \frac{5}{36} \times 1^2 + \frac{4}{36} \times 2^2 + \frac{3}{36} \times 3^2 + \frac{2}{36} \times 4^2 + \frac{1}{36} \times 5^2 \\
 &= \frac{35}{6},
 \end{aligned}$$

$$\text{Var}(W) = E(W^2) - (E(W))^2 = \frac{35}{6} - 0^2 = \frac{35}{6}.$$

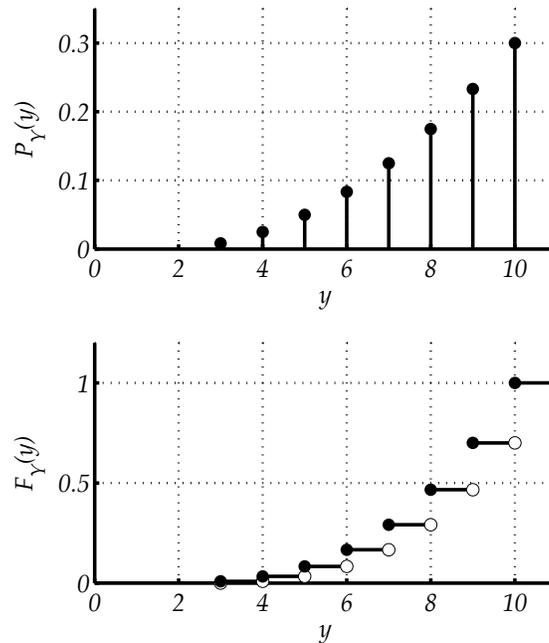
11. Δέκα μπάλες. Συνολικά υπάρχουν $\binom{10}{3}$ δυνατές επιλογές, άρα ο χώρος πιθανότητας αποτελείται από $\binom{10}{3}$ στοιχεία και τα αντίστοιχα στοιχειώδη ενδεχόμενα είναι ισοπίθανα.

Για το Y , παρατηρούμε αρχικά πως η μικρότερη δυνατή τιμή του είναι $Y = 3$, και συνεπώς έχει σύνολο τιμών το $S_Y = \{3, 4, 5, 6, 7, 8, 9, 10\}$. Τώρα, για κάθε $y \in S_Y$, οι επιλογές που αντιστοιχούν σε $Y = y$ είναι εκείνες κατά τις οποίες επιλέγουμε την μπάλα y και δύο ακόμα από τις $1, 2, \dots, y - 1$. Αυτό μπορεί να γίνει με $\binom{y-1}{2}$ τρόπους, και από τον κανόνα

πιθανότητας #5 προκύπτει πως η πυκνότητα της Y δίνεται από την έκφραση,

$$P_Y(y) = \frac{\binom{y-1}{2}}{\binom{10}{3}}, \quad y = 3, 4, \dots, 10.$$

Η συνάρτηση κατανομής μπορεί εύκολα να προσδιοριστεί από την πυκνότητα, μέσω της σχέσης (6.1). Η πυκνότητα και η συνάρτηση κατανομής έχουν σχεδιαστεί στο Σχήμα Α΄.26.

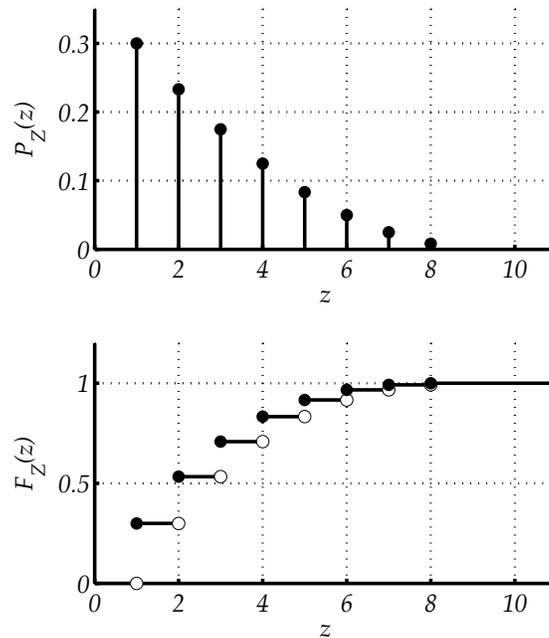


Σχήμα Α΄.26: Άσκηση 11: Η πυκνότητα και η συνάρτηση κατανομής της Τ.Μ. Y .

Σχετικά με τη Z , παρομοίως παρατηρούμε πως η μεγαλύτερη δυνατή τιμή της είναι $Z = 8$, και συνεπώς έχει σύνολο τιμών το $S_Z = \{2, 3, 4, 5, 6, 7, 8\}$. Για κάθε $z \in S_Z$, οι επιλογές που αντιστοιχούν σε $Z = z$ είναι εκείνες κατά τις οποίες επιλέγουμε την μπάλα z και δύο ακόμα από τις $z + 1, z + 2, \dots, 10$. Αυτό μπορεί να γίνει με $\binom{10-z}{2}$ τρόπους, και από τον κανόνα πιθανότητας #5 προκύπτει πως η πυκνότητα της Z δίνεται από την έκφραση,

$$P_Z(z) = \frac{\binom{10-z}{2}}{\binom{10}{3}}, \quad z = 1, 2, \dots, 8.$$

Η συνάρτηση κατανομής της Z μπορεί εύκολα να προσδιοριστεί από την πυκνότητα, μέσω της σχέσης (6.1), και οι δύο συναρτήσεις έχουν σχεδιαστεί στο Σχήμα Α΄.27.



Σχήμα Α'.27: Άσκηση 11: Η πυκνότητα και η συνάρτηση κατανομής της Τ.Μ. Z.

12. Περιορισμοί στις παραμέτρους.

(α) Από τη βασική ιδιότητα της πυκνότητας πρέπει να έχουμε,

$$\sum_{x \in S_X} P(x) = 1,$$

και, επομένως, $P(-2) + P(-1) + P(0) + P(1) + P(2) = 1$, ή, ισοδύναμα,

$$\frac{1}{10} + a + b + a + \frac{1}{10} = 1, \quad \text{δηλαδή} \quad 2a + b = \frac{4}{5}.$$

Άρα, θα πρέπει να έχουμε $a \in [0, 2/5]$, $b \in [0, 4/5]$ και $2a + b = 4/5$.

(β) Η μέση τιμή της X είναι,

$$\mu = E(X) = \sum_{x \in S_X} xP(x) = \frac{1}{10} \cdot (-2) + a \cdot (-1) + b \cdot 0 + a \cdot 1 + \frac{1}{10} \cdot 2 = 0,$$

η οποία παρατηρούμε ότι είναι ανεξάρτητη των a και b (γιατί;).

(γ) Όταν $a = 0$, απαραίτητα έχουμε ότι $b = 4/5$, από το πρώτο σκέλος, και επομένως η διασπορά της X είναι:

$$\text{Var}(X) = \sum_{x \in S_X} (x - \mu)^2 \cdot P(x) = \frac{1}{10} \cdot (-2)^2 + a \cdot (-1)^2 + b \cdot 0^2 + a \cdot 1^2 + \frac{1}{10} \cdot 2^2 = \frac{4}{5}.$$

13. ★ *Επιζώντα ζευγάρια.* Έστω οι τυχαίες μεταβλητές $M_i, F_i, i = 1, 2, \dots, n$, όπου, για κάθε i , έχουμε $M_i = 1$ (ή 0) αν ο i -οστός άνδρας επιβιώσει (ή πεθάνει), και αντιστοίχως $F_i = 1$ (ή 0) αν η i -οστή γυναίκα επιβιώσει (ή πεθάνει). Επιπλέον, ορίζουμε τις Τ.Μ. Z_i , οι οποίες παίρνουν την τιμή 1 αν το ζευγάρι i παραμείνει, αλλιώς $Z_i = 0$, και παρατηρούμε πως $Z_i = M_i F_i$, για κάθε i (γιατί;).

Το συνολικό πλήθος, έστω Z , των ζευγαριών που παραμένουν, μπορεί να εκφραστεί ως,

$$Z = \sum_{i=1}^n Z_i = \sum_{i=1}^n M_i F_i,$$

και για τον υπολογισμό της ζητούμενης μέσης τιμής $E(Z)$ αρκεί να υπολογίσουμε την $E(Z_i)$ της κάθε Z_i . Παρατηρούμε ότι, από τον ορισμό της μέσης τιμής και τον ορισμό της δεσμευμένης πιθανότητας, έχουμε,

$$\begin{aligned} E(Z_i) &= 1 \times \Pr(Z_i = 1) + 0 \times \Pr(Z_i = 0) \\ &= \Pr(\{M_i = 1\} \cap \{F_i = 1\}) \\ &= \Pr(M_i = 1) \Pr(F_i = 1 | M_i = 1). \end{aligned}$$

Ακολουθώντας ένα συλλογισμό αντίστοιχο με εκείνον της Άσκησης 8, βλέπουμε πως $\Pr(M_i = 1) = \frac{2n-m}{2n}$ και $\Pr(F_i = 1 | M_i = 1) = \frac{2n-m-1}{2n-1}$, για κάθε i . Συνοπτικά, η πιθανότητα του να επιβιώσει ο άντρας i είναι, όπως και για οποιοδήποτε άλλο μέλος του πληθυσμού, ίση με ένα μείον την πιθανότητα να πεθάνει, δηλαδή,

$$1 - \frac{m}{2n} = \frac{2n-m}{2n}.$$

Παρομοίως, αν γνωρίζουμε ότι επιβίωσε ο άνδρας i , μένουν $2n-1$ άτομα εκ των οποίων θα επιβιώσουν οι $2n-m-1$.

Τέλος, συνδυάζοντας τα πιο πάνω με την πρώτη ιδιότητα του Θεωρήματος 6.1, βρίσκουμε τη ζητούμενη μέση τιμή:

$$E(Z) = E\left(\sum_{i=1}^n Z_i\right) = \sum_{i=1}^n E(Z_i) = \frac{(2n-m)(2n-m-1)}{2(2n-1)}. \quad (\text{A'.10})$$

14. *Τρία ζάρια.* Έστω πως X, Y, Z τα αποτελέσματα των τριών ρίψεων. Παρατηρούμε ότι η $X+Y$ έχει σύνολο τιμών το $S_2 = \{2, 3, \dots, 8\}$, και από το αποτέλεσμα της Άσκησης 9 για τα X, Y , έχουμε ότι η πυκνότητα P_{X+Y} της $X+Y$, ικανοποιεί,

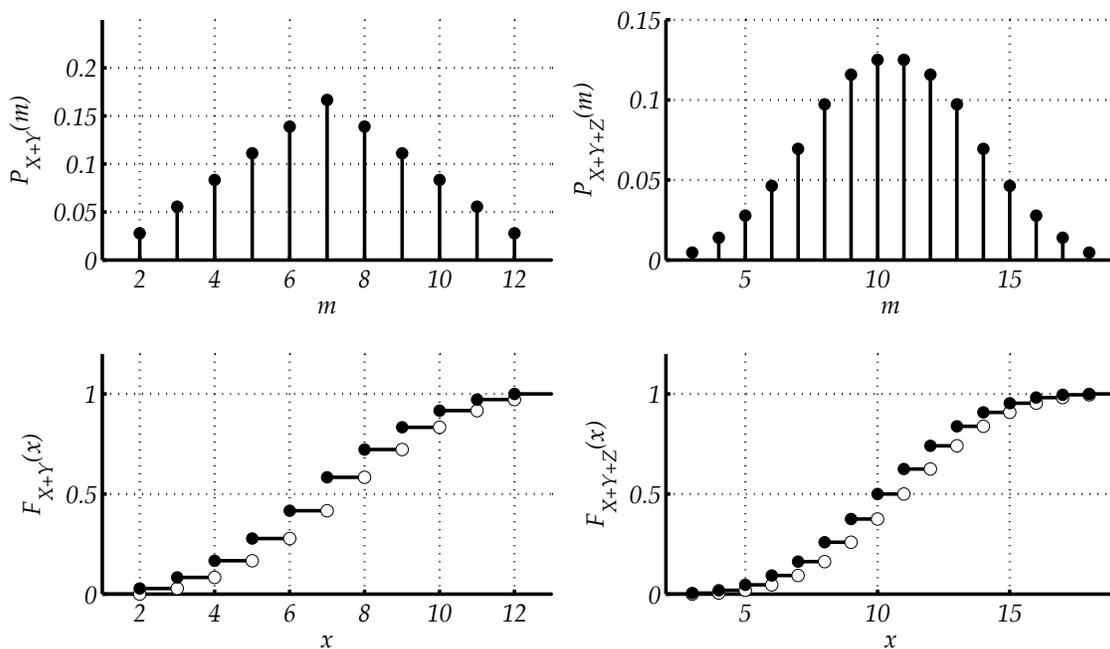
$$\begin{aligned} P_{X+Y}(2) &= P_X(1)P_Y(1) = \frac{1}{6} \times \frac{1}{6}, \\ P_{X+Y}(3) &= P_X(2)P_Y(1) + P_X(1)P_Y(2) = 2 \times \frac{1}{6} \times \frac{1}{6}, \\ P_{X+Y}(4) &= P_X(3)P_Y(1) + P_X(2)P_Y(2) + P_X(1)P_Y(3) = 3 \times \frac{1}{6} \times \frac{1}{6}, \end{aligned}$$

όπου συμβολίζουμε με P_X , P_Y και P_Z τις πυκνότητες των X , Y και Z αντίστοιχα. Συνεχίζοντας με αυτό τον τρόπο προκύπτουν και οι υπόλοιπες τιμές της πυκνότητας του $X + Y$:

$$P_{X+Y}(5) = \frac{4}{36}, \quad P_{X+Y}(6) = \frac{5}{36}, \quad P_{X+Y}(7) = \frac{6}{36}, \quad P_{X+Y}(8) = \frac{5}{36},$$

$$P_{X+Y}(9) = \frac{4}{36}, \quad P_{X+Y}(10) = \frac{3}{36}, \quad P_{X+Y}(11) = \frac{2}{36}, \quad P_{X+Y}(12) = \frac{1}{36}.$$

Η συνάρτηση κατανομής της $X + Y$ εύκολα υπολογίζεται από τη σχέση (6.1). Η πυκνότητα και η συνάρτηση κατανομής της $X + Y$ έχουν σχεδιαστεί στο Σχήμα A'28.



Σχήμα A'28: Οι πυκνότητες και οι συναρτήσεις κατανομής της Άσκησης 14.

Τώρα, για να υπολογίσουμε την πυκνότητα της $X + Y + Z$, εφαρμόζουμε το αποτέλεσμα της Άσκησης 9 για τις τυχαίες μεταβλητές $X + Y$ και Z . Αρχικά παρατηρούμε ότι η $X + Y + Z$ έχει σύνολο τιμών το $S_3 = \{3, 4, \dots, 18\}$, και η πυκνότητά της ικανοποιεί,

$$P_{X+Y+Z}(3) = P_{X+Y}(2)P_Z(1) = \frac{1}{36} \times \frac{1}{6} = \frac{1}{216},$$

$$P_{X+Y+Z}(4) = P_{X+Y}(3)P_Z(1) + P_{X+Y}(2)P_Z(2) = \frac{2}{36} \times \frac{1}{6} + \frac{1}{36} \times \frac{1}{6} = \frac{3}{216}.$$

Συνεχίζοντας με τον ίδιο τρόπο προκύπτουν και οι υπόλοιπες τιμές:

$$\begin{aligned} P_{X+Y+Z}(5) &= \frac{6}{216}, & P_{X+Y+Z}(6) &= \frac{10}{216}, & P_{X+Y+Z}(7) &= \frac{15}{216}, \\ P_{X+Y+Z}(8) &= \frac{21}{216}, & P_{X+Y+Z}(9) &= \frac{25}{216}, & P_{X+Y+Z}(10) &= \frac{27}{216}, \\ P_{X+Y+Z}(11) &= \frac{27}{216}, & P_{X+Y+Z}(12) &= \frac{25}{216}, & P_{X+Y+Z}(13) &= \frac{21}{216}, \\ P_{X+Y+Z}(14) &= \frac{15}{216}, & P_{X+Y+Z}(15) &= \frac{10}{216}, & P_{X+Y+Z}(16) &= \frac{6}{216}, \\ & & P_{X+Y+Z}(17) &= \frac{3}{216}, & P_{X+Y+Z}(18) &= \frac{1}{216}. \end{aligned}$$

Και πάλι, η συνάρτηση κατανομής της $X + Y + Z$ εύκολα υπολογίζεται από τη σχέση (6.1), βλ. Σχήμα Α΄.28.

15. Παιχνίδι.

(α) Η μέση τιμή της X σε μία παρτίδα είναι:

$$\mu = E(X) = \sum_{x \in S_X} xP(x) = (-5) \cdot \frac{1}{6} + 0 \cdot \frac{1}{2} + 2 \cdot \frac{1}{6} + 10 \cdot \frac{1}{6} = \frac{7}{6}.$$

Η διασπορά της X είναι:

$$\begin{aligned} \text{Var}(X) &= \sum_{x \in S_X} (x - \mu)^2 \cdot P(x) \\ &= \frac{1}{6} \cdot \left(-5 - \frac{7}{6}\right)^2 + \frac{1}{2} \cdot \left(0 - \frac{7}{6}\right)^2 + \frac{1}{6} \cdot \left(2 - \frac{7}{6}\right)^2 + \frac{1}{6} \cdot \left(10 - \frac{7}{6}\right)^2 \\ &= \frac{725}{36}. \end{aligned}$$

(β) Έστω Y το συνολικό κέρδος του παίκτη σε τρεις ανεξάρτητες παρτίδες και X_1, X_2, X_3 το κέρδος σε κάθε παρτίδα. Από την πρώτη ιδιότητα του Θεωρήματος 6.1, η μέση τιμή της Y είναι:

$$E(Y) = E\left(\sum_{i=1}^3 X_i\right) = \sum_{i=1}^3 E(X_i) = 3 \cdot \frac{7}{6} = \frac{7}{2}.$$

Ομοίως, αφού οι παρτίδες είναι ανεξάρτητες, από την τέταρτη ιδιότητα του Θεωρήματος 6.1, η διασπορά της Y είναι:

$$\text{Var}(Y) = \text{Var}\left(\sum_{i=1}^3 X_i\right) = \sum_{i=1}^3 \text{Var}(X_i) = 3 \cdot \frac{725}{36} = \frac{725}{12}.$$

16. Η μέθοδος της δεύτερης ροπής. Έστω $P(k)$, για $k \in S = \{0, 1, 2, \dots\}$, η πυκνότητα της X . Από τον ορισμό της διασποράς έχουμε,

$$\sigma^2 = \text{Var}(X) = \sum_{k=0}^{\infty} (k - \mu)^2 P(k),$$

και, εφόσον όλοι οι όροι του πιο πάνω αθροίσματος είναι μεγαλύτεροι ή ίσοι του μηδενός, ολόκληρο το άθροισμα θα είναι μεγαλύτερο ή ίσο του όρου $k = 0$,

$$\sigma^2 = \sum_{k=0}^{\infty} (k - \mu)^2 P(k) \geq (0 - \mu)^2 P(0) = \mu^2 \Pr(X = 0),$$

που είναι ακριβώς η ζητούμενη ανισότητα.

17. Εναλλακτική έκφραση της $E(X)$. Ξεκινάμε από τη ζητούμενη έκφραση και χρησιμοποιούμε την τρίτη βασική ιδιότητα της πυκνότητας:

$$\sum_{k=1}^{\infty} \Pr(X \geq k) = \sum_{k=1}^{\infty} \sum_{j=k}^{\infty} \Pr(X = j).$$

Παρατηρούμε πως οι όροι του αθροίσματος δεν εξαρτώνται από το k , και πως ο κάθε όρος $\Pr(X = j)$ αθροίζεται ακριβώς j φορές. Συνεπώς,

$$\sum_{k=1}^{\infty} \Pr(X \geq k) = \sum_{j=1}^{\infty} j \Pr(X = j) = E(X),$$

που είναι ακριβώς η ζητούμενη έκφραση.

18. ★ Αποστάσεις μεταξύ πυκνοτήτων. Οι παρατηρήσεις του πρώτου σκέλους είναι άμεσες συνέπειες των ορισμών. Για το δεύτερο σκέλος παρατηρούμε ότι, προφανώς, από τον ορισμό της, οποιαδήποτε διασπορά είναι μεγαλύτερη ή ίση του μηδενός. Άρα, έχουμε,

$$\begin{aligned} 0 &\leq \text{Var}(f(X)) \\ &\stackrel{(a)}{=} E[f(X)^2] - (E[f(X)])^2 \\ &\stackrel{(b)}{=} \sum_x Q(x) f(x)^2 - \left(\sum_x Q(x) f(x) \right)^2 \\ &= \sum_x Q(x) \left(\frac{|P(x) - Q(x)|}{Q(x)} \right)^2 - \left(\sum_x Q(x) \frac{|P(x) - Q(x)|}{Q(x)} \right)^2 \\ &= \sum_x \frac{(P(x) - Q(x))^2}{Q(x)} - \left(\sum_x |P(x) - Q(x)| \right)^2 \\ &= d_{\chi^2}(P, Q) - (d_{\ell_1}(P, Q))^2, \end{aligned}$$

το οποίο μας δίνει ακριβώς τη ζητούμενη ανισότητα. Στο βήμα (a) πιο πάνω χρησιμοποιήσαμε την εναλλακτική έκφραση της διασποράς (6.8), και στο βήμα (b) τον γενικό τύπο (6.5) για τη μέση τιμή κάποιας συνάρτησης μιας Τ.Μ.

Α.7 Ασκήσεις Κεφαλαίου 7

1. *Overbooking.* Το πλήθος, έστω X , των επιβατών που θα προσέλθουν ακολουθεί διωνυμική κατανομή με παραμέτρους το πλήθος των ανεξάρτητων πειραμάτων $N = 55$, και πιθανότητα επιτυχίας $p = 0.9$. Θα έχουμε πρόβλημα υπεράριθμων επιβατών αν προσέλθουν 51 με 55 επιβάτες, δηλαδή αν $X \geq 51$. Η πιθανότητα αυτή μπορεί εύκολα να υπολογιστεί από τον τύπο της πυκνότητας της διωνυμικής κατανομής:

$$\begin{aligned} \Pr(X \geq 51) &= \Pr(X = 51) + \Pr(X = 52) + \Pr(X = 53) \\ &\quad + \Pr(X = 54) + \Pr(X = 55) \\ &= \binom{55}{55} 0.9^{55} 0.1^0 + \binom{55}{54} 0.9^{54} 0.1^1 + \binom{55}{53} 0.9^{53} 0.1^2 \\ &\quad + \binom{55}{52} 0.9^{52} 0.1^3 + \binom{55}{51} 0.9^{51} 0.1^4 \\ &\approx 0.3451. \end{aligned}$$

2. Δείκτριες τυχαίες μεταβλητές.

(α) Εφόσον η X παίρνει μόνο τις τιμές 0 και 1, προφανώς έχει $\text{Bern}(p)$ κατανομή, όπου η παράμετρος της $p = \Pr(X = 1) = \mathbb{P}(A)$.

(β') Και πάλι, εφόσον η συνάρτηση $h_x(y)$ παίρνει μόνο τις τιμές 0 και 1, η Τ.Μ. $h_x(Y)$ έχει κατανομή Bernoulli με παράμετρο $p = \Pr(h_x(Y) = 1) = \Pr(Y \leq x) = F(x)$.

3. ★ *Η ουρά της μέσης τιμής.* Έστω πως η Τ.Μ. Y έχει πυκνότητα $P(y)$ και σύνολο τιμών το $S = \{a_1, a_2, \dots\}$, όπου όλα τα $a_i \geq 0$ και για ευκολία θεωρούμε ότι $0 \leq a_1 < a_2 < \dots$. Κατ' αρχάς, από τον ορισμό της $h_x(y)$, παρατηρούμε πως η νέα Τ.Μ. $Z = Y[1 - h_x(Y)]$ ισούται με την Y όταν $Y > x$, αλλιώς ισούται με μηδέν. Άρα, για μεγάλες τιμές του x , διαισθητικά περιμένουμε ότι η Z θα είναι μη μηδενική με πολύ μικρή πιθανότητα. Με αυτό το σκεπτικό, το ζητούμενο αποτέλεσμα δεν μας εκπλήσσει.

Πιο αναλυτικά τώρα, η μέση τιμή που μας ενδιαφέρει, από τον ορισμό της μέσης τιμής μιας συνάρτησης κάποιας Τ.Μ., είναι,

$$E(Y[1 - h_x(Y)]) = \sum_{y \in S} yP(y)[1 - h_x(y)].$$

Αν το S είναι πεπερασμένο με μέγιστο στοιχείο ίσο με M , ή αν είναι άπειρο και όλα του τα στοιχεία είναι $a_i \leq M$, για κάποια σταθερά M , το αποτέλεσμα είναι τετριμμένο: Για κάθε $x > M$, όλες οι τιμές $1 - h_x(y)$ είναι ίσες με μηδέν, οπότε το άθροισμα προφανώς τείνει στο μηδέν καθώς το $x \rightarrow \infty$.

Τέλος, αν το S είναι άπειρο και όχι φραγμένο ως σύνολο, τότε για κάθε x υπάρχει κάποιο $i(x)$ τέτοιο ώστε να έχουμε $a_i > x$ αν και μόνο αν $i \geq i(x)$. Και επιπλέον θα έχουμε ότι $i(x) \rightarrow \infty$ καθώς το $x \rightarrow \infty$. Με αυτόν το συμβολισμό έχουμε,

$$E(Y[1 - h_x(Y)]) = \sum_{i=1}^{\infty} a_i P(a_i)[1 - h_x(a_i)] = \sum_{i=i(x)}^{\infty} a_i P(a_i),$$

αλλά επίσης ξέρουμε πως ολόκληρη η σειρά,

$$E(Y) = \sum_{i=1}^{\infty} a_i P(a_i),$$

συγκλίνει σε ένα πεπερασμένο άθροισμα. Άρα απαραίτητα η «ουρά» της σειράς θα τείνει στο μηδέν,

$$E(Y[1 - h_x(Y)]) = \sum_{i=i(x)}^{\infty} a_i P(a_i) \rightarrow 0,$$

αφού το $i(x) \rightarrow \infty$ καθώς το $x \rightarrow \infty$.

4. *Ελάχιστο δύο γεωμετρικών τυχαίων μεταβλητών.* Ακολουθώντας την υπόδειξη, παρατηρούμε ότι το ελάχιστο Z είναι τουλάχιστον k , αν και μόνο αν και οι δύο Τ.Μ. X, Y παίρνουν τιμές τουλάχιστον k . Οπότε, για κάθε $k \geq 1$,

$$\begin{aligned} \Pr(Z \geq k) &= \Pr(\min(X, Y) \geq k) \\ &= \Pr(X \geq k, Y \geq k) \\ &= \Pr(X \geq k) \Pr(Y \geq k) \\ &= (1 - p_1)^{k-1} (1 - p_2)^{k-1} \\ &= [(1 - p_1)(1 - p_2)]^{k-1}, \end{aligned}$$

όπου η τρίτη ισότητα ισχύει λόγω της πέμπτης ιδιότητας της ανεξαρτησίας στο Θεώρημα 6.1, και η τέταρτη από την πρώτη ιδιότητα της γεωμετρικής κατανομής στο Θεώρημα 7.1.

Αν τώρα ορίσουμε $q = 1 - (1 - p_1)(1 - p_2)$, εφόσον,

$$\Pr(Z \geq k) = \Pr(Z = k) + \Pr(Z \geq k + 1),$$

έχουμε,

$$\Pr(Z = k) = \Pr(Z \geq k) - \Pr(Z \geq k + 1) = (1 - q)^{k-1} - (1 - q)^k = q(1 - q)^{k-1}.$$

Συνεπώς, και η Z έχει γεωμετρική κατανομή, αλλά με παράμετρο $q = 1 - (1 - p_1)(1 - p_2)$.

Αυτό εξηγείται εύκολα: Αν διεξάγουμε ταυτόχρονα δύο ακολουθίες ανεξάρτητων πειραμάτων, μπορούμε να ορίσουμε ένα ολικό πείραμα στο οποίο θα έχουμε επιτυχία την πρώτη φορά που θα έχει επιτυχία ένα από τα δύο πειράματα. Σε αυτή την περίπτωση, το πλήθος των επαναλήψεων που απαιτούνται για την πρώτη επιτυχία του ολικού πειράματος είναι $Z = \min(X, Y)$ όπου X, Y είναι το πλήθος επαναλήψεων μέχρι την πρώτη επιτυχία στο πρώτο και το δεύτερο πείραμα, αντίστοιχα. Το ολικό πείραμα αποτυγχάνει όταν αποτύχουν και τα δύο επί μέρους, δηλαδή με πιθανότητα $(1 - p_1)(1 - p_2)$, άρα η πιθανότητα επιτυχίας είναι $1 - (1 - p_1)(1 - p_2) = q$. Συνεπώς, το πλήθος των προσπαθειών του ολικού πειράματος μέχρι την πρώτη επιτυχία ακολουθεί γεωμετρική κατανομή με παράμετρο q .

5. *Άθροισμα Poisson.* Έστω $P_X(x)$ και $P_Y(y)$ οι πυκνότητες των X και Y αντίστοιχα. Παρατηρούμε (όπως και στην Άσκηση 9 του Κεφαλαίου 6) πως η $X + Y$ ισούται με z αν και μόνο αν για κάποιο k έχουμε $X = k$ και $Y = z - k$. Αν $P_{X+Y}(z)$ είναι η πυκνότητα της $X + Y$, τότε έχουμε,

$$\begin{aligned} P_{X+Y}(z) &= \sum_{x=0}^z P_X(x) P_Y(z-x) \\ &= \sum_{x=0}^z \left(\frac{e^{-\mu} \mu^x}{x!} \right) \left(\frac{e^{-\lambda} \lambda^{z-x}}{(z-x)!} \right) \\ &= \frac{e^{-(\mu+\lambda)}}{z!} \sum_{x=0}^z \frac{z!}{x!(z-x)!} \mu^x \lambda^{z-x} \\ &= \frac{e^{-(\mu+\lambda)}}{z!} \sum_{x=0}^z \binom{z}{x} \mu^x \lambda^{z-x} \\ &= \frac{e^{-(\mu+\lambda)}}{z!} (\mu + \lambda)^z, \end{aligned}$$

όπου η τελευταία ισότητα προκύπτει από το διωνυμικό θεώρημα. Συνεπώς, η τυχαία μεταβλητή $X + Y \sim \text{Poisson}(\lambda + \mu)$.

6. *Ταυτότητα Vandermonde.* Πρώτα εξετάζουμε την περίπτωση $m \leq n_1$ και $m \leq n_2$. Θα δείξουμε ότι τα δυο μέρη της σχέσης (7.14) εκφράζουν το ίδιο πλήθος συνδυασμών. Έστω πως έχουμε n_1 άντρες και n_2 γυναίκες. Το πλήθος των τρόπων με τους οποίους μπορούμε να δημιουργήσουμε μια επιτροπή με m από αυτά τα άτομα είναι προφανώς $\binom{n_1+n_2}{m}$. Συγχρόνως όμως, για κάθε k , υπάρχουν $\binom{n_1}{k} \binom{n_2}{m-k}$ τρόποι να δημιουργήσουμε μια επιτροπή που να αποτελείται από k άντρες και $m - k$ γυναίκες. Οπότε το αριστερό σκέλος είναι το άθροισμα του πλήθους των τρόπων με τους οποίους μπορούμε να φτιάξουμε μια επιτροπή m ατόμων, απλώς ομαδοποιώντας τους συνδυασμούς ανάλογα με το πόσους άνδρες θα έχει η επιτροπή.

Αν το m είναι μεγαλύτερο του $n_1 + n_2$, τότε και τα δύο σκέλη είναι ίσα με μηδέν.

Τέλος, αν το m είναι μικρότερο ή ίσο του $n_1 + n_2$, αλλά μεγαλύτερο του n_1 ή του n_2 , τότε κάποιοι από τους όρους του αριστερού σκέλους είναι ίσοι με μηδέν, αλλά η πιο πάνω απόδειξη εξακολουθεί να ισχύει για τους όρους που απομένουν. Για παράδειγμα, αν $n_1 = 4$, $n_2 = 4$, $m = 6$, τότε θα πρέπει να ισχύει:

$$\binom{4}{2} \binom{4}{4} + \binom{4}{3} \binom{4}{3} + \binom{4}{4} \binom{4}{2} = \binom{8}{6}.$$

7. Άθροισμα διωνυμικών Τ.Μ. Έστω $P_X(x)$ και $P_Y(y)$ οι πυκνότητες των X και Y , αντίστοιχα. Παρατηρούμε ότι η $X+Y$ έχει σύνολο τιμών το $S = \{0, 1, \dots, n_1 + n_2\}$. Επιπλέον, από το αποτέλεσμα της Άσκησης 9 του Κεφαλαίου 6 (βλ. και την παρόμοια λύση της παραπάνω Άσκησης 5), για κάθε $m \in S$ έχουμε,

$$\begin{aligned} P_{X+Y}(m) &= \sum_{k=-\infty}^{\infty} P_X(k)P_Y(m-k) \\ &= \sum_{k=0}^m \binom{n_1}{k} p^k (1-p)^{n_1-k} \binom{n_2}{m-k} p^{m-k} (1-p)^{n_2-m+k} \\ &= (p^m (1-p)^{n_1+n_2-m}) \sum_{k=0}^m \binom{n_1}{k} \binom{n_2}{m-k} \\ &= \binom{n_1+n_2}{m} p^m (1-p)^{n_1+n_2-m}, \end{aligned}$$

όπου στην τελευταία ισότητα χρησιμοποιήσαμε την ταυτότητα του Vandermonde. Συνεπώς η πυκνότητα της $X+Y$ είναι αυτή της Διων(n_1+n_2, p) κατανομής.

Εναλλακτικά, θα μπορούσαμε να εκφράσουμε τις Τ.Μ. X, Y ως,

$$X = \sum_{i=1}^{n_1} Z_i, \quad \text{και} \quad Y = \sum_{i=n_1+1}^{n_1+n_2} Z_i,$$

όπου οι Z_i είναι ανεξάρτητες Bern(p) Τ.Μ. Άρα, το άθροισμά τους $X+Y$ μπορεί και αυτό να εκφραστεί ως το άθροισμα (n_1+n_2) ανεξάρτητων Bern(p) Τ.Μ., και συνεπώς έχει Διων(n_1+n_2, p) κατανομή.

Τέλος, παρατηρούμε πως το αποτέλεσμα αυτό ισχύει μόνο αν η παράμετρος p και των αρχικών τυχαίων μεταβλητών X, Y είναι η ίδια.

8. Διάφορες κατανομές.

(α) Η Y ακολουθεί γεωμετρική κατανομή με παράμετρο $p = 1/6$ αφού η πιθανότητα να φέρουμε διπλή είναι $1/6$ (γιατί;), και τα πειράματα (δηλαδή οι διαδοχικές ζαριές) είναι ανεξάρτητα μεταξύ τους.

- (β') Η X ακολουθεί διωνυμική κατανομή με παραμέτρους $N = 6$ και $p = 1/4$, αφού έχουμε το πλήθος επιτυχιών σε $N = 6$ όμοια, ανεξάρτητα πειράματα με πιθανότητα επιτυχίας $p = 13/52 = 1/4$.
- (γ') Η X ακολουθεί υπεργεωμετρική κατανομή με παραμέτρους $(52, 13, 6)$ αφού επιλέγουμε χωρίς επαναποθέτηση 6 φύλλα από τα 52 που έχει συνολικά μια τράπουλα, εκ των οποίων τα 13 είναι κούπες.
- (δ') Η X ακολουθεί διωνυμική κατανομή με παραμέτρους $N = 20$ και $p = 0.7$, αφού έχουμε το πλήθος επιτυχιών σε $N = 20$ όμοια, ανεξάρτητα πειράματα με πιθανότητα επιτυχίας $p = \Pr(\text{Γράμματα}) = 1 - 0.3 = 0.7$.

9. Ουρά στην τράπεζα.

- (α') Οι 2 ώρες έχουν 24 πεντάλεπτα, κι εφόσον η πιθανότητα επιτυχίας, δηλαδή του να έρθει ένας νέος πελάτης, είναι 0.05, η κατανομή της X είναι $X \sim \text{Διων}(24, 0.05)$.
- (β') Εφόσον $X =$ πλήθος πελατών που έφτασαν τις πρώτες 2 ώρες, η πιθανότητα να έρθουν ακριβώς 3 πελάτες τις πρώτες δύο ώρες, από τον τύπο της πυκνότητας της διωνυμικής κατανομής είναι,

$$\Pr(X = 3) = P(3) = \binom{24}{3} (0,05)^3 (1 - 0,05)^{24-3} \approx 0,0862.$$

- (γ') Η Y περιγράφει τη χρονική στιγμή της πρώτης επιτυχίας, άρα έχει γεωμετρική κατανομή με παράμετρο $p = 0.05$: $Y \sim \text{Γεωμ}(0.05)$.
- (δ') Τα λεπτά που θα περιμένουν οι υπάλληλοι μέχρι την άφιξη του πρώτου πελάτη κατά μέσο όρο είναι πέντε φορές η μέση τιμή της Y . Έχουμε $E(Y) = 1/0.05 = 100/5 = 20$ πεντάλεπτα, άρα ο αναμενόμενος χρόνος μέχρι τον πρώτο πελάτη είναι $5 \times 20 = 100$ λεπτά.
- (ε') Παρατηρήστε ότι οι 2 ώρες έχουν 24 πεντάλεπτα και οι 2.5 ώρες έχουν 30 πεντάλεπτα. Λόγω της ιδιότητας έλλειψης μνήμης έχουμε:

$$\begin{aligned} \Pr(X > 30 | X > 24) &= \Pr(X \geq 31 | X > 24) \\ &= \Pr(X \geq 24 + 7 | X > 24) \\ &= \Pr(X \geq 7) \\ &= \Pr(X > 6) = (1 - 0,05)^6 = (0,95)^6 \approx 0.7351. \end{aligned}$$

10. Επιθέσεις μηνυμάτων spam. Έστω X το πλήθος των χρηστών που ανταποκρίνονται στο email. Το X εκφράζει το πλήθος των επιτυχιών σε $N = 10$ χιλιάδες όμοια, ανεξάρτητα πειράματα, που το καθένα έχει πιθανότητα επιτυχίας $p = 0.00018$. Συνεπώς $X \sim \text{Διων}(10000, 0.00018)$.

Εφόσον $Np = 10000 \times 0.00018 = 1.8$ είναι «της τάξεως του 1», προφανώς ικανοποιούνται όλες οι συνθήκες του Πορίσματος 7.1, και μπορούμε να προσεγγίσουμε την κατανομή της X μέσω της Poisson(λ), με $\lambda = Np = 10000 \times 0.00018 = 1.8$. Από τον τύπο της πυκνότητας της κατανομής Poisson έχουμε,

$$\begin{aligned} \Pr(X \geq 3) &= 1 - \Pr(X \leq 2) \\ &= 1 - \Pr(X = 0) - \Pr(X = 1) - \Pr(X = 2) \\ &\approx 1 - e^{-1.8} \cdot \frac{1.8^0}{0!} - e^{-1.8} \cdot \frac{1.8^1}{1!} - e^{-1.8} \cdot \frac{1.8^2}{2!} \\ &\approx 0.2694 \approx 26.9\%. \end{aligned}$$

[Δεδομένου του μηδαμινού κόστους της αποστολής των μηνυμάτων, η πρόβλεψη για τον διαφημιστή δεν είναι κι άσχημη!]

11. Τα «ν» του Βαρουφάκη.

- (α') Έστω οι ανεξάρτητες Τ.Μ. X_1, X_2, \dots, X_N με κατανομή $X_i \sim \text{Bern}(p)$ για κάθε i . Αν $N = 800$ και $p = 0.003$, μπορούμε να θεωρήσουμε πως $X_i = 1$ αν και μόνο αν τη φορά i ο διορθωτής έκανε λάθος και έγραψε το όνομα «Γιάννης» αντί για «Γιάνης». Άρα, αν ορίσουμε ως Y την Τ.Μ. $Y = X_1 + X_2 + \dots + X_N$, τότε η $Y \sim \text{Διων}(N, p)$ και το συνολικό πλήθος από «ν» που χρησιμοποιήθηκαν θα είναι $Z = N + Y$.
- (β') Εφόσον $N = 800$, $p = 0.003$ και $\lambda = Np = 2.4$, προφανώς ικανοποιούνται οι συνθήκες του Πορίσματος 7.1, άρα μπορούμε να πούμε πως η Y έχει, κατά προσέγγιση, ίδια κατανομή με μια Τ.Μ. $W \sim \text{Poisson}(\lambda)$, οπότε έχουμε,

$$\Pr(Z = 810) = \Pr(800 + Y = 810) = \Pr(Y = 10) \approx \Pr(W = 10),$$

το οποίο ισούται με,

$$e^{-\lambda} \frac{\lambda^{10}}{10!} \approx 0.0001585 \approx 0.016\%.$$

12. Άσχετος φοιτητής. Ορίζουμε τις ανεξάρτητες Τ.Μ. X_1, X_2, \dots, X_N , όλες με την ίδια κατανομή $X_i \sim \text{Bern}(p)$ για κάθε i , όπου $N = 250$, $p = 0.005$ και θεωρούμε πως $X_i = 1$ αν και μόνο αν ο εξεταζόμενος απάντησε σωστά στην ερώτηση i .

Τότε το πλήθος $Y = X_1 + X_2 + \dots + X_N$ των σωστών απαντήσεων του έχει κατανομή $Y \sim \text{Διων}(N, p)$ και, εφόσον $N = 250$, $p = 0.005$ και $\lambda = Np = 1.25$, βλέπουμε πως ικανοποιούνται οι συνθήκες του Πορίσματος 7.1. Συνεπώς, η κατανομή της Y είναι, κατά προσέγγιση, Poisson(λ), και άρα:

$$\Pr(Y = 3) \approx e^{-\lambda} \frac{\lambda^3}{3!} \approx 0.0933.$$

13. Δύο διαφορετικές δημοσκοπήσεις.

(α') Όπως περιγράψαμε στο Παράδειγμα 7.2, το συνολικό πλήθος επιτυχιών σε προβλήματα επιλογής με επανατοποθέτηση κατά κανόνα έχει διωνυμική κατανομή. Εδώ, αν ορίσουμε τις ανεξάρτητες Τ.Μ. X_1, X_2, \dots, X_N , όπου το κάθε $X_i = 1$ αν το άτομο i στο δείγμα είναι ψηφοφόρος του κόμματος και $X_i = 0$ αν όχι, τότε μπορούμε να εκφράσουμε την Y ως το άθροισμα των X_i , οπότε $Y \sim \text{Διων}(N, p)$, όπου η παράμετρος p ισούται με την πιθανότητα ένα τυχαία επιλεγμένο άτομο να είναι ψηφοφόρος του κόμματος, δηλαδή $p = 0.3$. Συνεπώς, για τη μέση τιμή και τη διασπορά της πρόβλεψης, έχουμε,

$$\begin{aligned} E(Y/N) &= \frac{1}{N} E(Y) = \frac{1}{N} \cdot Np = 0.3, \\ \text{Var}(Y/N) &= \frac{1}{N^2} \text{Var}(Y) = \frac{1}{N^2} \cdot Np(1-p) = \frac{0.21}{N}, \end{aligned}$$

όπου χρησιμοποιήσαμε τον τύπο για τη μέση τιμή και τη διασπορά της διωνυμικής κατανομής, και τις πρώτες δύο ιδιότητες του Θεωρήματος 6.1.

Πριν προχωρήσουμε αξίζει να κάνουμε ορισμένες παρατηρήσεις. Πρώτον, η μέση τιμή $E(Y/N)$ της εκτίμησης του εκλογικού αποτελέσματος είναι ίση με το πραγματικό αποτέλεσμα. Δεύτερον, η διασπορά $\text{Var}(Y/N)$ της εκτίμησης μικραίνει, και συνεπώς η εκτίμηση γίνεται ακριβέστερη, καθώς το μέγεθος N του δείγματός μας μεγαλώνει. Και τρίτον, η μέση τιμή και η διασπορά της εκτίμησης δεν εξαρτώνται από το μέγεθος M του πληθυσμού.

(β') Από το Παράδειγμα 7.7 και τον ορισμό της υπεργεωμετρικής κατανομής, άμεσα έπεται πως εδώ η $Y \sim \text{Υπερ}(M, 3M/10, N)$, οπότε από τις σχέσεις (7.2) και (7.3), αντικαθιστώντας τις γνωστές τιμές του προβλήματος, βρίσκουμε,

$$\begin{aligned} E(Y/N) &= \frac{1}{N} E(Y) = \frac{1}{N} \cdot N \cdot \frac{3M}{10} \cdot \frac{1}{M} = 0.3, \\ \text{Var}(Y/N) &= \frac{1}{N^2} \text{Var}(Y) \\ &= \frac{1}{N^2} \cdot \frac{N(3M/10)(M - 3M/10)(M - N)}{M^2(M - 1)} \\ &= \frac{0.21}{N} \cdot \frac{(M - N)}{(M - 1)}, \end{aligned}$$

όπου και πάλι χρησιμοποιήσαμε τις πρώτες δύο ιδιότητες του Θεωρήματος 6.1.

Από τα πιο πάνω αποτελέσματα παρατηρούμε τα εξής. Πρώτον, και πάλι η μέση τιμή $E(Y/N)$ της εκτίμησης του εκλογικού αποτελέσματος είναι ίση με το πραγματικό αποτέλεσμα. Δεύτερον, και πάλι η διασπορά $\text{Var}(Y/N)$ μικραίνει καθώς το μέγεθος

N του δείγματος μεγαλώνει, αλλά αυτήν τη φορά εξαρτάται και από το μέγεθος M του πληθυσμού.

Συγκεκριμένα, η διασπορά εδώ είναι μικρότερη από την περίπτωση δειγματοληψίας με επανατοποθέτηση, άρα, υπ' αυτήν την έννοια, είναι προτιμότερη η επιλογή δειγμάτων χωρίς επανατοποθέτηση, πράγμα και διαισθητικά αναμενόμενο, μια που σ' αυτή την περίπτωση είμαστε βέβαιοι ότι έχουμε εξετάσει N διαφορετικά άτομα χωρίς επαναλήψεις. Τέλος, παρατηρούμε ότι η διαφορά της εδώ διασποράς σε σχέση με το προηγούμενο ερώτημα είναι μεγάλη όταν το μέγεθος του δείγματος είναι σχετικά μεγάλο σε σχέση με τον συνολικό πληθυσμό. Αντίθετα, όταν το N είναι σημαντικά μικρότερο του M , καθώς το μέγεθος του πληθυσμού M αυξάνει, η διασπορά επίσης αυξάνεται και τείνει σ' εκείνη που είχαμε στην περίπτωση επιλογής με επανατοποθέτηση.

14. *Βελάκια.* Έστω X το πλήθος από βελάκια που πέτυχαν το στόχο.

(α') Εδώ το X εκφράζει το πλήθος των επιτυχιών σε $N = 10$ όμοια, ανεξάρτητα πειράματα, που το καθένα έχει πιθανότητα επιτυχίας $p = 0.01$. Συνεπώς $X \sim \text{Διων}(10, 0.01)$, και από τον τύπο της πυκνότητάς της,

$$\Pr(X = 2) = \binom{10}{2} \cdot 0.01^2 \cdot (1 - 0.01)^{10-2} \approx 0.0041524 \approx 0.42\%.$$

(β') Παρομοίως, εδώ $X \sim \text{Διων}(140, 0.01)$. Εφόσον $140 \times 0.01 = 1.4$, προφανώς ικανοποιούνται όλες οι συνθήκες του Πορίσματος 7.1, και μπορούμε να προσεγγίσουμε την κατανομή της X μέσω της Poisson(λ), με $\lambda = 140 \times 0.01 = 1.4$. Από τον τύπο της πυκνότητάς της κατανομής Poisson έχουμε,

$$\Pr(X = 2) \approx e^{-1.4} \cdot \frac{1.4^2}{2!} \approx 0.2417 \approx 24.2\%.$$

15. *Γινόμενο Bernoulli.* Κατ' αρχάς παρατηρούμε πως το γινόμενο μπορεί να πάρει μόνο δύο τιμές, την 0 και την 1, επομένως έχει κατανομή Bernoulli. Έστω q η ζητούμενη παράμετρος. Θα δώσουμε δύο λύσεις:

Πρώτα παρατηρούμε πως η παράμετρος μιας T.M. Bernoulli είναι ίση με τη μέση τιμή της. Χρησιμοποιώντας την τρίτη ιδιότητα του Θεωρήματος 6.1 (λόγω της ανεξαρτησίας των X_i), βρίσκουμε,

$$q = E(Z) = E(X_1 X_2 \cdots X_n) = E(X_1) E(X_2) \cdots E(X_n) = p_1 p_2 \cdots p_n.$$

Εναλλακτικά, παρατηρούμε πως,

$$\begin{aligned} q = \Pr(Z = 1) &= \Pr(X_1 X_2 \cdots X_n = 1) = \Pr(X_1 = 1, X_2 = 1, \dots, X_n = 1) \\ &= \Pr(X_1 = 1) \Pr(X_2 = 1) \cdots \Pr(X_n = 1) = p_1 p_2 \cdots p_n, \end{aligned}$$

όπου η τέταρτη ισότητα προκύπτει λόγω της ανεξαρτησίας των X_i .

16. Χρηματιστήριο.

(α') Η κάθε Y_i ορίζεται ως,

$$Y_i = \begin{cases} 1, & \text{με πιθανότητα } 14\%, \\ 0, & \text{με πιθανότητα } 86\%, \end{cases}$$

άρα, η κάθε Y_i έχει κατανομή Bernoulli με παράμετρο $p = 0.14$. Επιπλέον, από την εκφώνηση του προβλήματος έχουμε υποθέσει ότι οι Y_i είναι ανεξάρτητες μεταξύ τους.

(β') Η X ισούται με το άθροισμα 30 ανεξάρτητων τυχαίων μεταβλητών Bernoulli,

$$X = \sum_{i=1}^{30} Y_i,$$

κι άρα η κατανομή της είναι διωνυμική, με παραμέτρους $N = 30$ και $p = 0.14$.

(γ') Από τον τύπο της πυκνότητας της διωνυμικής κατανομής,

$$\begin{aligned} \Pr(E) &= \Pr(10 \leq X \leq 12) = P(10) + P(11) + P(12) \\ &= \binom{30}{10} (0.14)^{10} (0.86)^{20} + \binom{30}{11} (0.14)^{11} (0.86)^{19} + \binom{30}{12} (0.14)^{12} (0.86)^{18} \\ &\approx 0.0058. \end{aligned}$$

(δ') Η τιμή Z της μετοχής μετά από 60 μέρες μπορεί να εκφραστεί ως,

$$Z = 100 + \sum_{i=1}^{60} Y_i,$$

όπου το άθροισμα των Y_i έχει Διων(N, p) κατανομή με παραμέτρους $N = 60$ και $p = 0.14$. Επομένως, από την πρώτη ιδιότητα του Θεωρήματος 6.1 και τον τύπο της μέσης τιμής της διωνυμικής κατανομής,

$$E(Z) = E\left(100 + \sum_{i=1}^{60} Y_i\right) = 100 + E\left(\sum_{i=1}^{60} Y_i\right) = 100 + 60 \times 0.14 = 108.4.$$

Παρομοίως, από τη δεύτερη ιδιότητα του Θεωρήματος 6.1 και τον τύπο της διασποράς της διωνυμικής κατανομής,

$$\text{Var}(Z) = \text{Var}\left(100 + \sum_{i=1}^{60} Y_i\right) = \text{Var}\left(\sum_{i=1}^{60} Y_i\right) = 60 \times 0.14 \times (1 - 0.14) = 7.224.$$

(ε΄) Σε αυτή την περίπτωση έχουμε,

$$Z = 100 + \sum_{i=1}^{60} 2Y_i = 100 + 2 \sum_{i=1}^{60} Y_i,$$

όπου το άθροισμα των Y_i έχει όπως πριν $\Delta\text{ων}(60, 0.14)$ κατανομή. Επομένως, εφαρμόζοντας τις ίδιες ιδιότητες, βρίσκουμε,

$$E(Z) = E\left(100 + 2 \sum_{i=1}^{60} Y_i\right) = 100 + 2E\left(\sum_{i=1}^{60} Y_i\right) = 100 + 2 \times 60 \times 0.14 = 116.8,$$

και για τη διασπορά,

$$\begin{aligned} \text{Var}(Z) &= \text{Var}\left(100 + 2 \sum_{i=1}^{60} Y_i\right) \\ &= 2^2 \text{Var}\left(\sum_{i=1}^{60} Y_i\right) \\ &= 4 \times 60 \times 0.14 \times (1 - 0.14) = 28.896. \end{aligned}$$

17. XOR Bernoulli.

(α΄) Για να υπολογίσουμε την παράμετρο, έστω p , της Z χρησιμοποιούμε την ανεξαρτησία των X, Y :

$$p = \Pr(Z = 1) = \Pr(X = 1, Y = 0) + \Pr(X = 0, Y = 1) = \frac{1}{4} \times \frac{1}{2} + \frac{3}{5} \times \frac{1}{2} = \frac{1}{8}.$$

(β΄) Θα δείξουμε πως, παρότι εκ πρώτης όψεως μοιάζουν να σχετίζονται, οι X και Z είναι ανεξάρτητες. Αυτό μπορεί να αποδειχθεί μέσω του ορισμού της ανεξαρτησίας, δείχνοντας πως,

$$\Pr(X = x, Z = z) = \Pr(X = x) \Pr(Z = z),$$

και για τα 4 δυνατά ζευγάρια τιμών (x, z) . Πράγματι, έχουμε,

$$\Pr(X = 1, Z = 1) = \Pr(X = 1, Y = 0) = \Pr(X = 1) \Pr(Y = 0) = \frac{1}{4} \times \frac{1}{2} = \frac{1}{8},$$

ενώ και,

$$\Pr(X = 1) \Pr(Z = 1) = \frac{1}{4} \times \frac{1}{2} = \frac{1}{8}.$$

Παρομοίως,

$$\Pr(X = 0, Z = 1) = \Pr(X = 0, Y = 1) = \Pr(X = 0) \Pr(Y = 1) = \frac{3}{4} \times \frac{1}{2} = \frac{3}{8},$$

ενώ και,

$$\Pr(X = 0) \Pr(Z = 1) = \frac{3}{4} \times \frac{1}{2} = \frac{3}{8}.$$

Επίσης,

$$\Pr(X = 1, Z = 0) = \Pr(X = 1, Y = 1) = \Pr(X = 1) \Pr(Y = 1) = \frac{1}{4} \times \frac{1}{2} = \frac{1}{8},$$

ενώ και,

$$\Pr(X = 1) \Pr(Z = 0) = \frac{1}{4} \times \frac{1}{2} = \frac{1}{8}.$$

Και, τέλος,

$$\Pr(X = 0, Z = 0) = \Pr(X = 0, Y = 0) = \Pr(X = 0) \Pr(Y = 0) = \frac{3}{4} \times \frac{1}{2} = \frac{3}{8},$$

ενώ και,

$$\Pr(X = 0) \Pr(Z = 0) = \frac{3}{4} \times \frac{1}{2} = \frac{3}{8}.$$

Άρα οι X, Z είναι ανεξάρτητες.

18. ★ Διασπορά υπεργεωμετρικής κατανομής.

(α') Αυτό το βήμα προκύπτει από έναν απλό υπολογισμό, σε συνδυασμό με την πρώτη ιδιότητα του Θεωρήματος 6.1:

$$\begin{aligned} E(Y^2) &= E \left[\left(\sum_{i=1}^n X_i \right)^2 \right] = E \left[\left(\sum_{i=1}^n X_i \right) \left(\sum_{j=1}^n X_j \right) \right] \\ &= E \left[\sum_{i=1}^n \sum_{j=1}^n X_i X_j \right] = \sum_{i=1}^n \sum_{j=1}^n E(X_i X_j). \end{aligned}$$

(β') Με βάση το συλλογισμό που χρησιμοποιήσαμε για τον υπολογισμό της μέσης τιμής (στην παρατήρηση που ακολουθεί το Παράδειγμα 7.8) όλα τα X_i έχουν την ίδια κατανομή, συνεπώς $E(X_i^2) = E(X_1^2)$. Επιπλέον, αφού τα X_i παίρνουν μόνο τις τιμές 0 και 1, προφανώς πάντοτε έχουμε $X_1^2 = X_1$ και άρα $E(X_1^2) = E(X_1)$.

Με την ίδια λογική της εναλλακτικής περιγραφής του προβλήματος, οι πιθανότητες όλων των δυνατών αποτελεσμάτων για δύο (X_i, X_j) από όλες τις Τ.Μ. X_i , είναι ίδιες ανεξαρτήτως των (i, j) , και συνεπώς $E(X_i X_j) = E(X_1 X_2)$ για κάθε $i \neq j$.

(γ') Η παράμετρος της Bernoulli τυχαίας μεταβλητής X_1X_2 είναι η ,

$$\Pr(X_1X_2 = 1) = \Pr(X_1 = 1, X_2 = 1) = \Pr(X_2 = 1|X_1 = 1)\Pr(X_1 = 1),$$

όπου χρησιμοποιήσαμε τον ορισμό της δεσμευμένης πιθανότητας. Και χρησιμοποιώντας τώρα τις πιθανότητες που υπολογίστηκαν στις σχέσεις (7.1) και (7.2), έχουμε,

$$\Pr(X_1X_2 = 1) = \frac{k-1}{N-1} \frac{k}{N}.$$

(δ') Χρησιμοποιώντας τα τρία πιο πάνω βήματα βρίσκουμε,

$$\begin{aligned} E(Y^2) &= \sum_{i=1}^n \sum_{j=1}^n E(X_iX_j) \\ &= nE(X_1) + n(n-1)E(X_1X_2) \\ &= n \frac{k}{N} + n(n-1) \frac{k}{N} \frac{k-1}{N-1}. \end{aligned}$$

(ε') Από το προηγούμενο βήμα, σε συνδυασμό με τη γνωστή μέση τιμή $E(Y) = kn/N$, έχουμε,

$$\begin{aligned} \text{Var}(Y) &= E(Y^2) - [E(Y)]^2 \\ &= n \frac{k}{N} + n(n-1) \frac{k}{N} \frac{k-1}{N-1} - \frac{k^2n^2}{N^2} \\ &= \frac{nkN(N-1) + n(n-1)k(k-1)N - k^2n^2(N-1)}{N^2(N-1)} \\ &= \frac{nk(N-k)(N-n)}{N^2(N-1)}, \end{aligned}$$

όπως ακριβώς και στη ζητούμενη σχέση (7.3).

19. ★ Αναπαράσταση της e^x .

(α') Για $y \in (-1, \infty)$ ορίζουμε τη συνάρτηση $f(y) = \log(1+y)$, η οποία έχει παραγώγους,

$$f'(y) = \frac{1}{1+y} \quad \text{και} \quad f''(y) = -\frac{1}{(1+y)^2}.$$

Συνεπώς, από ανάπτυγμα Taylor έχουμε πως, για κάποιο ζ με $|\zeta| \leq |y|$,

$$\log(1+y) = f(1+y) = f(0) + yf'(0) + \frac{y^2}{2}f''(\zeta) = y - \frac{y^2}{2(1+\zeta)^2}.$$

(β') Εφόσον η ακολουθία $\{x_n\}$ συγκλίνει, θα είναι υποχρεωτικά φραγμένη, άρα θα έχουμε $|x_n/n| \in [0, 1)$ για κάθε n μεγαλύτερο ή ίσο με κάποιο n_0 . Εφαρμόζοντας το προηγούμενο αποτέλεσμα με $y = x_n/n$, για $n \geq n_0$,

$$\log \left(1 + \frac{x_n}{n} \right) = \frac{x_n}{n} - \frac{x_n^2}{2n^2(1 + \zeta_n)^2},$$

όπου τα ζ_n ικανοποιούν $|\zeta_n| \leq |x_n/n| < 1$. Θέτοντας $\xi_n = x_n^2/[(1 + \zeta_n)^2]$ και πολλαπλασιάζοντας την πιο πάνω σχέση με n βρίσκουμε, για $n \geq n_0$,

$$n \log \left(1 + \frac{x_n}{n} \right) = x_n - \frac{1}{2n} \xi_n,$$

όπου η ακολουθία $\{\xi_n\}$ είναι εξ ορισμού φραγμένη.

(γ') Τέλος, παίρνοντας το όριο καθώς $n \rightarrow \infty$ στο αποτέλεσμα του προηγούμενου σκέλους, έχουμε,

$$\lim_{n \rightarrow \infty} \log \left[\left(1 + \frac{x_n}{n} \right)^n \right] = \lim_{n \rightarrow \infty} n \log \left(1 + \frac{x_n}{n} \right) = x = \log e^x,$$

το οποίο είναι προφανώς ισοδύναμο με το ζητούμενο αποτέλεσμα της σχέσης (7.5).

Α'.8 Ασκήσεις Κεφαλαίου 8

1. Ένας άλλος *randomized* αλγόριθμος. Έστω ότι μας ζητείται να υπολογίσουμε την τιμή $F(x)$ για κάποιο συγκεκριμένο x . Επιλέγουμε τυχαία έναν ακέραιο Y στο σύνολο $S = \{0, 1, \dots, n-1\}$ και παρατηρούμε ότι το $(x+Y \bmod n)$ είναι επίσης ομοιόμορφα κατανοημένο στο S . Από τον πίνακα χρησιμοποιούμε τις τιμές των $F(x+Y \bmod n)$ και $F(Y)$ και, εφόσον η συνάρτηση F πάντα ικανοποιεί τη σχέση $F(x+Y \bmod n) = F(x)F(Y)$, ως απάντηση για το $F(x)$ δίνουμε το $F(x+Y \bmod n)/F(Y)$.

Ποια είναι η πιθανότητα να έχουμε κάνει λάθος;

Η μόνη περίπτωση να δώσουμε λάθος απάντηση είναι αν η τιμή του $F(x+Y \bmod n)$ είναι λάθος στον πίνακα ή αν η τιμή του $F(Y)$ είναι λάθος στον πίνακα. Άρα,

$$\begin{aligned} \Pr(\text{λάθος απάντηση}) &\leq \Pr(\{F(Y) \text{ λάθος}\} \cup \{F(x+Y \bmod n) \text{ λάθος}\}) \\ &\leq \Pr(F(Y) \text{ λάθος}) + \Pr(F(x+Y \bmod n) \text{ λάθος}) \\ &= \frac{1}{5} + \frac{1}{5} = 40\%, \end{aligned}$$

όπου η δεύτερη ανισότητα προκύπτει από το φράγμα ένωσης της Ενότητας 5.2, και οι δύο πιθανότητες ίσες με $1/5$ γιατί το Y και το $(x+Y \bmod n)$ είναι ομοιόμορφα κατανοημένα στο $\{0, 1, \dots, n-1\}$.

Παρατηρούμε πως, αν απλώς δώναμε την $F(x)$ για απάντηση, θα κάναμε λάθος μόνο 20% του χρόνου, οπότε γιατί να μπούμε στον κόπο του πιο πολύπλοκου *randomized* αλγόριθμου; Η διαφορά είναι ότι, αν δεν κάναμε τίποτα, τότε για τα x τα οποία έχουν λάθος τιμές στον πίνακα θα κάναμε πάντοτε λάθος. Το πλεονέκτημα του *randomized* αλγόριθμου είναι πως, ακόμα και για τις τιμές που είναι λάθος στον πίνακα, τουλάχιστον 60% του χρόνου μάς δίνει τη σωστή απάντηση!

2. Αθροίσματα.

- (α') Έστω πως $\sum_{k=1}^n k = (n/2)(n+1)$ για κάθε άρτιο n , και έστω $n = m+1$ όπου το m είναι άρτιο. Τότε προφανώς έχουμε,

$$\sum_{k=1}^n k = \sum_{k=1}^m k + n = \frac{m}{2}(m+1) + n = \left(\frac{n-1}{2}\right)n + n = \frac{n}{2}(n+1).$$

Εναλλακτικά, θα μπορούσαμε να τροποποιήσουμε την αρχική μέθοδο, προσθέτοντας, ανά ζεύγη, τους αριθμούς, $1+n$, $2+(n-1)$, κ.ο.κ., έως το $\frac{n-1}{2} + \frac{n+3}{2}$, παίρνοντας το αποτέλεσμα $[(n-1)/2]$ -φορές-το- $(n+1)$, δηλαδή, $\frac{n-1}{2}(n+1)$, και τέλος προσθέτοντας σε αυτό τον μεσαίο όρο $\frac{n+1}{2}$, καταλήγοντας στο ίδιο αποτέλεσμα.

(β') Για οποιαδήποτε $n \geq 1$ και $1 \leq m \leq n-1$, χρησιμοποιώντας το πιο πάνω αποτέλεσμα βρίσκουμε,

$$\begin{aligned} \sum_{k=m+1}^n (2k) &= \sum_{k=1}^n (2k) - \sum_{k=1}^m (2k) \\ &= 2 \sum_{k=1}^n k - 2 \sum_{k=1}^m k \\ &= 2 \frac{n}{2}(n+1) - 2 \frac{m}{2}(m+1) \\ &= n^2 + n - m^2 - m \\ &= (n+m+1)(n-m). \end{aligned}$$

3. *Μέτρηση πράξεων.* Η πρώτη μέθοδος δεν απαιτεί καμία ύψωση σε δύναμη, οπότε η πολυπλοκότητά της για μεγάλα n παραμένει,

$$\frac{n^2 - n}{2} \approx \frac{1}{2}n^2.$$

Η δεύτερη μέθοδος απαιτεί τον υπολογισμό των x^2, x^3, \dots, x^n για καθεμία από τις n ρίζες $x = b_i$, οπότε το επιπλέον κόστος ισούται με

$$n \sum_{k=2}^n \log k = n \log \left(\prod_{k=2}^n k \right) = n \log n! \approx n[(n+1/2) \log n - n] \approx n^2 \log n,$$

όπου χρησιμοποιήσαμε την προσέγγιση $\log n! \approx (n+1/2) \log n - n$, που προκύπτει από τον τύπο του Stirling. Άρα το συνολικό κόστος αυτής της μεθόδου, για μεγάλα n είναι της τάξης του $n^2 + n^2 \log n$, δηλαδή της τάξης του $n^2 \log n$.

Παρομοίως, το συνολικό κόστος του randomized αλγορίθμου είναι,

$$2n + 2 \sum_{k=2}^n \log k = 2n + 2 \log n! \approx 2n + 2[(n+1/2) \log n - n] \approx 2n \log n.$$

Συμπεραίνουμε λοιπόν πως, και με αυτή την ακριβέστερη έννοια, η πολυπλοκότητα του randomized αλγορίθμου είναι σημαντικά μικρότερη από εκείνη των δύο κλασικών μεθόδων.

4. *Μικρότερη πιθανότητα σφάλματος.* Οι δύο τρόποι είναι μάλλον προφανείς: Μπορούμε να αυξήσουμε το πλήθος M των δυνατών τιμών της τυχαίας μεταβλητής που χρησιμοποιούμε, ή μπορούμε να χρησιμοποιήσουμε την ίδια μέθοδο πολλές φορές.

Στην πρώτη περίπτωση, αν αντί για $100n$ δυνατές τιμές επιλέξουμε μια Τ.Μ. X με M δυνατές τιμές, τότε αν το M ικανοποιεί $M \geq n/p$, ο ίδιος υπολογισμός όπως πριν δείχνει πως η πιθανότητα σφάλματος είναι το πολύ p .

Στην δεύτερη περίπτωση, επαναλαμβάνουμε το ίδιο πείραμα k φορές, χρησιμοποιώντας k ανεξάρτητες Τ.Μ. X_1, X_2, \dots, X_k με την ίδια κατανομή όπως πριν. Υπολογίζουμε τις τιμές $f(X_i)$ και $g(X_i)$, και αν $f(X_i) = g(X_i)$ για κάθε $i = 1, 2, \dots, k$, δηλώνουμε πως τα δύο πολυώνυμα είναι τα ίδια· αν $f(X_i) \neq g(X_i)$ για τουλάχιστον ένα i , δηλώνουμε πως είναι διαφορετικά. Και πάλι, η μόνη περίπτωση σφάλματος είναι αν τα $f(x)$ και $g(x)$ είναι διαφορετικά, οπότε έχουμε,

$$\begin{aligned} P_e &= \Pr(\{f(X_1) = g(X_1)\} \cap \{f(X_2) = g(X_2)\} \cap \dots \cap \{f(X_k) = g(X_k)\}) \\ &\stackrel{(a)}{=} \prod_{i=1}^k \Pr(f(X_i) - g(X_i) = 0) \\ &\stackrel{(b)}{=} (\Pr(f(X_1) - g(X_1) = 0))^k \\ &\stackrel{(c)}{\leq} \left(\frac{1}{100n}\right)^k, \end{aligned}$$

όπου στο βήμα (a) χρησιμοποιήσαμε την ανεξαρτησία των X_i , στο (b) το γεγονός ότι όλα τα X_i έχουν την ίδια κατανομή, και στο (c) αντικαταστήσαμε το ήδη γνωστό φράγμα για την πιθανότητα σφάλματος από την απλή περίπτωση μίας μόνο τυχαίας μεταβλητής. Από αυτό το αποτέλεσμα είναι προφανές ότι, επιλέγοντας ένα επαρκώς μεγάλο k , μπορούμε να καταστήσουμε το φράγμα $[1/(100n)]^k$ όσο κοντά στο μηδέν επιθυμούμε.

5. Πόσα *min-cuts* υπάρχουν; Έστω C το τυχαίο cut set το οποίο μας δίνει ο αλγόριθμος του Karger. Εφόσον κάθε πιθανότητα είναι μικρότερη ή ίση του 1, έχουμε ότι,

$$\begin{aligned} 1 &\geq \Pr(\text{το } C \text{ είναι min-cut}) \\ &= \Pr(\{C = C_1\} \cup \{C = C_2\} \cup \dots \cup \{C = C_M\}) \\ &= \Pr(S(1) \cup S(2) \cup \dots \cup S(M)) \\ &= \Pr(S(1)) + \Pr(S(2)) + \dots + \Pr(S(M)), \end{aligned}$$

διότι τα ενδεχόμενα $S(j) = \{C = C_j\}$ είναι ξένα μεταξύ τους. Χρησιμοποιώντας τώρα την ανισότητα που μας υπενθύμισε η εκφώνηση, συνεχίζουμε τον πιο πάνω υπολογισμό και βρίσκουμε,

$$1 \geq M \frac{2}{n(n-1)},$$

δηλαδή,

$$M \leq \frac{n(n-1)}{2} = \frac{n(n-1)(n-2)!}{(n-2)!2!} = \binom{n}{2}.$$

6. ★ Τυχαία cut sets.

- (α') Θυμίζουμε το συλλογισμό που κάναμε πριν την περιγραφή του αλγορίθμου του Karger, σύμφωνα με τον οποίο κάθε cut set αντιστοιχεί σε ένα διαχωρισμό των κόμβων V του γράφου σε δύο (όχι κενά) υποσύνολα K_1 και K_2 . Όπως αναφέρουμε εκεί, για οποιονδήποτε τέτοιο διαχωρισμό, προφανώς το σύνολο των ακμών που συνδέουν κόμβους του K_1 με κόμβους στο K_2 είναι cut set.
- (β') Πιο συγκεκριμένα, ο παραπάνω συλλογισμός μάς λέει πως τα ζευγάρια (όχι κενών) υποσυνόλων $\{K_1, K_2\}$ του V τέτοια ώστε $K_1 \cup K_2 = V$, είναι σε 1-1 αντιστοιχία με τα cut sets του γράφου, όπου οι ρόλοι των K_1, K_2 είναι αντιστρέψιμοι. Για να αποδείξουμε ότι το τυχαίο C είναι ομοιόμορφα κατανομημένο, αρκεί να αποδείξουμε πως το αντίστοιχο τυχαίο σύνολο κόμβων K_1 είναι ομοιόμορφα κατανομημένο, δηλαδή πως κάθε K_1 επιλέγεται με την ίδια πιθανότητα. Αλλά αυτό είναι και πάλι προφανές από την κατασκευή του: Η επιλογή οποιουδήποτε συνόλου όπως περιγράφει η άσκηση έχει την ίδια πιθανότητα, δηλαδή $\frac{1}{2} \times \frac{1}{2} \times \dots \times \frac{1}{2} = 1/2^n$, και, αφού αποκλείουμε το κενό σύνολο και το V , η επιλογή οποιουδήποτε K_1 στο οποίο καταλήγουμε έχει πιθανότητα $1/[2^n - 2]$.
- (γ') Έστω A το σύνολο όλων των cut sets και B το σύνολο όλων των min-cuts. Από τον συλλογισμό που αναφέραμε πιο πάνω υπολογίσαμε πως το πλήθος των cut sets είναι $\#A = \Gamma_n = 2^{n-1} - 1$, και από την προηγούμενη άσκηση ξέρουμε πως το πλήθος των min-cuts ικανοποιεί $\#B \leq \binom{n}{2} = n(n+1)/2$. Άρα, αφού στο σκέλος (β') δείξαμε πως το C είναι τυχαία και ομοιόμορφα επιλεγμένο ανάμεσα σε όλα τα cut sets, έχουμε ότι:

$$P_{S'} = \Pr(\text{το } C \text{ είναι min-cut}) = \frac{\#B}{\#A} \leq \frac{\binom{n}{2}}{2^{n-1} - 1} = \frac{n(n-1)}{2(2^{n-1} - 1)}.$$

7. Φράγμα τομής. Χρησιμοποιώντας το φράγμα ένωσης, έχουμε,

$$\begin{aligned} \Pr\left(\bigcap_{i=1}^n A_i\right) &\stackrel{(a)}{=} 1 - \Pr\left[\left(\bigcap_{i=1}^n A_i\right)'\right] \\ &= 1 - \Pr\left(\bigcup_{i=1}^n A_i'\right) \\ &\stackrel{(b)}{\geq} 1 - \sum_{i=1}^n \Pr(A_i') \\ &\stackrel{(c)}{=} 1 - \sum_{i=1}^n [1 - \Pr(A_i)] = 1 + \sum_{i=1}^n \Pr(A_i) - n, \end{aligned}$$

όπου τα βήματα (a) και (c) προκύπτουν από τον κανόνα πιθανότητας #4 και το βήμα (b) από το φράγμα ένωσης (Λήμμα 5.2).

8. ★ *Μια γενικότερη πιθανοκρατική ανάλυση.* Για οποιοδήποτε (αυθαίρετο) $\epsilon > 0$, θέτουμε $k = \left(\frac{1+\epsilon}{\log_2(1/p)}\right) \log_2 n$. Θα αποδείξουμε κάτω από τις παρούσες συνθήκες πως,

$$\Pr\left(L_n > \left(\frac{1+\epsilon}{\log_2(1/p)}\right) \log_2 n\right) = \Pr(L_n > k) \leq \frac{1}{n^\epsilon}, \quad (\text{A'.11})$$

το οποίο πράγματι τείνει στο μηδέν καθώς το $n \rightarrow \infty$. Ορίζουμε τα ενδεχόμενα E_i ακριβώς όπως πριν. Με τον ίδιο ακριβώς συλλογισμό και χρησιμοποιώντας το φράγμα ένωσης του Λήμματος 5.2, έχουμε,

$$\Pr(L_n > k) = \Pr\left(\bigcup_{i=1}^{n-k+1} E_i\right) \leq \sum_{i=1}^{n-k+1} \Pr(E_i),$$

όπου εδώ η πιθανότητα του κάθε ενδεχομένου E_i είναι,

$$\Pr(E_i) = \Pr(X_i = 1, X_{i+1} = 1, \dots, X_{i+k-1} = 1) = p^k,$$

οπότε,

$$\Pr\left(L_n > \left(\frac{1+\epsilon}{\log_2(1/p)}\right) \log_2 n\right) \leq (n-k+1)p^k.$$

Τέλος, χρησιμοποιώντας το προφανές φράγμα $(n-k+1) \leq n$, γράφοντας $p^k = 2^{-k \log_2(1/p)}$, και αντικαθιστώντας την τιμή του k ,

$$\Pr\left(L_n > \left(\frac{1+\epsilon}{\log_2(1/p)}\right) \log_2 n\right) \leq n 2^{-k \log_2(1/p)} = n 2^{-(1+\epsilon) \log_2(n)} = n \cdot \frac{1}{n^{1+\epsilon}},$$

και έχουμε ακριβώς το ζητούμενο φράγμα (A'.11).

Το αποτέλεσμα αυτό μας λέει πως, για μεγάλα n , η τιμή του L_n είναι μικρότερη από,

$$\frac{1}{\log_2(1/p)} \log_2 n, \quad \text{με μεγάλη πιθανότητα.}$$

Παρατηρούμε πως ο συντελεστής του $(\log_2 n)$ είναι αύξουσα συνάρτηση του p , άρα όσο μεγαλώνει το p , δηλαδή όσο μεγαλώνει η πιθανότητα του «1» στην ακολουθία των X_i , τόσο μεγαλώνει και το φράγμα μας για το μήκος L_n της μεγαλύτερης ακολουθίας συνεχόμενων «1» ανάμεσα στα X_i , το οποίο είναι διαισθητικά απολύτως λογικό.

9. *Προσομοίωση.* Και τα τρία ερωτήματα είναι καθαρά προγραμματιστικά. Για το τελευταίο μέρος του (γ'), σημειώνουμε πως μπορεί να αποδειχθεί ότι, πράγματι,

$$\frac{L_n}{\log_2 n} \rightarrow \frac{1}{\log_2(1/p)}, \quad \text{καθώς το } n \rightarrow \infty,$$

αλλά αυτή η σύγκλιση συμβαίνει με πολύ αργό ρυθμό, δηλαδή για πολύ μεγάλες τιμές του n . Γι' αυτόν το λόγο, ίσως να μην είναι προφανής από τα αποτελέσματα βάσει των τιμών που δίνονται στην άσκηση.

10. ★ *Μεγάλοι χρόνοι αναμονής.* Έχουμε X_1, X_2, \dots ανεξάρτητες $\text{Bern}(1/2)$ Τ.Μ., και η Τ.Μ. W_N περιγράφει την πρώτη χρονική στιγμή κατά την οποία εμφανίζεται το μοτίβο $0 \dots 01 \dots 1$ μήκους $2N$ στα X_i . Θα αποδείξουμε την (ισοδύναμη με το ζητούμενο του προβλήματος) σχέση:

$$\Pr(W_N \leq 2^N) \leq 2^{-N}. \quad (\text{A'.12})$$

Για να φράξουμε την πιθανότητα $\Pr(W_N \leq 2^N)$, ορίζουμε για κάθε $i = 1, 2, \dots$ τα ενδεχόμενα,

$$A_i = \{X_i = X_{i+1} = \dots = X_{i+N-1} = 0 \text{ και } X_{i+N} = X_{i+N+1} = \dots = X_{i+2N-1} = 1\}.$$

Έχουμε,

$$\Pr(W_N \leq 2^N) = \Pr\left(\bigcup_{i=1}^{2^N} A_i\right),$$

και χρησιμοποιώντας το φράγμα ένωσης του Κεφαλαίου 5,

$$\Pr(W_N \leq 2^N) \leq \sum_{i=1}^{2^N} \Pr(A_i).$$

Εφόσον οι Τ.Μ. X_i είναι ανεξάρτητες, για κάθε ενδεχόμενο A_i βρίσκουμε,

$$\begin{aligned} \Pr(A_i) &= \Pr(X_i = 0, X_{i+1} = 0, \dots, X_{i+N-1} = 0 \\ &\quad \text{και } X_{i+N} = 1, X_{i+N+1} = 1, \dots, X_{i+2N-1} = 1) \\ &= \left(\frac{1}{2}\right)^{2N} = 2^{-2N}, \end{aligned}$$

οπότε,

$$\Pr(W_N \leq 2^N) \leq 2^N \times 2^{-2N} = 2^{-N},$$

το οποίο αποδεικνύει την (A'.12).

11. *Τυχαίες διατάξεις.*

- (α') Το στοιχείο Z_i αρχικά βρίσκεται στη θέση i . Αλλά το Z_i είναι απλά ένας από τους αριθμούς 1 έως n , οπότε μετά την ταξινόμηση θα βρίσκεται στη θέση Z_i . Άρα θα έχει μετακινηθεί κατά $|Z_i - i|$ θέσεις.
- (β') Το ότι κάθε Z_i έχει ομοιόμορφη κατανομή στο $S = \{1, 2, \dots, n\}$ είναι προφανές λόγω της συμμετρίας του προβλήματος. Για να το αποδείξουμε και αυστηρά μαθηματικά παρατηρούμε ότι προφανώς έχει σύνολο τιμών το S και ότι η $\Pr(Z_i = j)$ είναι η πιθανότητα του να επιλέξουμε μια διάταξη η οποία στη θέση i να έχει στο στοιχείο j .

Το πλήθος αυτών των διατάξεων ισούται με το πλήθος των τρόπων με τους οποίους μπορούμε να διατάξουμε τα υπόλοιπα $(n-1)$ στοιχεία στις υπόλοιπες $(n-1)$ θέσεις, δηλαδή $(n-1)!$. Και εφόσον υπάρχουν συνολικά $n!$ δυνατές διατάξεις η ζητούμενη πιθανότητα είναι, όπως περιμένουμε, ίση με $\Pr(Z_i = j) = \frac{(n-1)!}{n!} = 1/n$.

(γ') Για να υπολογίσουμε τη ζητούμενη μέση τιμή, κατ' αρχάς χρησιμοποιώντας τον ορισμό της μέσης τιμής μιας συνάρτησης μιας Τ.Μ., θα υπολογίσουμε τη μέση τιμή,

$$\begin{aligned} E(|Z_i - i|) &= \sum_{j=1}^n \Pr(Z_i = j) |j - i| \\ &= \frac{1}{n} \sum_{j=1}^n |j - i| \\ &= \frac{1}{n} \sum_{j=1}^i (i - j) + \frac{1}{n} \sum_{j=i+1}^n (j - i) \\ &= \frac{1}{n} \sum_{k=0}^{i-1} k + \frac{1}{n} \sum_{k=1}^{n-i} k \\ &= \frac{1}{n} \frac{(i-1)i}{2} + \frac{1}{n} \frac{(n-i)(n-i+1)}{2} \\ &= \frac{1}{n} i^2 - \left(\frac{n+1}{n}\right) i + \left(\frac{n+1}{2}\right), \end{aligned}$$

όπου χρησιμοποιήσαμε τον τύπο για το άθροισμα μιας αριθμητικής προόδου όπως στην Άσκηση 2 πιο πάνω. Οπότε, εφαρμόζοντας την πρώτη ιδιότητα από το Θεώρημα 6.1, βρίσκουμε,

$$\begin{aligned} E\left(\sum_{i=1}^n |Z_i - i|\right) &= \sum_{i=1}^n E(|Z_i - i|) \\ &= \frac{1}{n} \sum_{i=1}^n i^2 - \left(\frac{n+1}{n}\right) \sum_{i=1}^n i + \frac{n(n+1)}{2} \\ &= \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{2} + \frac{n(n+1)}{2} \\ &= \frac{n^2 - 1}{3}, \end{aligned}$$

όπου χρησιμοποιήσαμε και πάλι τον τύπο για το άθροισμα μιας αριθμητικής προόδου, και τον τύπο (8.7) για το άθροισμα τετραγώνων που δίνεται στην εκφώνηση. [Άσκηση. Αποδείξτε τη σχέση (8.7) με επαγωγή.]

12. *O randomized quicksort.* Η απόδειξη του Θεωρήματος 8.3 είναι ακριβώς η ίδια με εκείνη του Θεωρήματος 8.2, με μόνη απαραίτητη τροποποίηση την τεκμηρίωσή του γιατί και εδώ η πιθανότητα p_{ij} είναι ίση με $2/(j-i+1)$. Όπως και στο Θεώρημα 8.2, λόγω της συμμετρίας του προβλήματος, αφού η επιλογή των οδηγών είναι τυχαία και ομοιόμορφη, και πάλι συμπεραίνουμε πως η πιθανότητα να επιλεγεί οποιοδήποτε από τα στοιχεία $\{y_i, y_{i+1}, \dots, y_j\}$ πρώτο ως οδηγός, είναι η ίδια. Συνεπώς, όπως και πριν, $p_{ij} = 2/(j-i+1)$.

Α'.9 Ασκήσεις Κεφαλαίου 9

1. *Αφρικανικό χελιδόσι.* Διαισθητικά, ο ισχυρισμός δεν μπορεί να ευσταθεί, γιατί ακόμα και αν το 60% των χελιδονιών είχε βάρος ακριβώς 200 γραμμάρια, και το υπόλοιπο 40% ακριβώς μηδέν (!) τότε το μέσο βάρος θα ήταν $0.6 \cdot 200 + 0.4 \cdot 0 = 120$ γραμμάρια, ενώ σε όλες τις άλλες περιπτώσεις το μέσο βάρος θα ήταν ακόμα μεγαλύτερο.

Στην γλώσσα των πιθανοτήτων, παραβιάζεται η ανισότητα του Markov: Αν X είναι το βάρος ενός τυχαίου χελιδονιού, θα πρέπει να έχουμε,

$$\Pr(X \geq 200) \leq \frac{E(X)}{200} = \frac{100}{200} = \frac{1}{2},$$

που δεν συμφωνεί με τον ισχυρισμό του βιολόγου ότι αυτή η πιθανότητα είναι ίση με 60%.

2. *Απόσταση T.M. από τη μέση τιμή της.* Αφού έχουμε $\sigma^2 \ll \mu^2$ (δηλαδή η διασπορά της X είναι «πολύ μικρότερη» από τη μέση τιμή στο τετράγωνο), από την ανισότητα του Chebychev προκύπτει πως:

$$\Pr(|X - \mu| \geq 3\mu) \leq \frac{\sigma^2}{(3\mu)^2} \ll \frac{\mu^2}{9\mu^2} = \frac{1}{9}.$$

Επομένως, ο στατιστικολόγος έχει δίκιο.

3. *Ιδιότητες από κοινού πυκνότητας.* Έστω N διακριτές τυχαίες μεταβλητές X_1, X_2, \dots, X_N με σύνολα τιμών $S_{X_1}, S_{X_2}, \dots, S_{X_N}$ και πυκνότητες $P_{X_1}(x_1), P_{X_2}(x_2), \dots, P_{X_N}(x_N)$ αντίστοιχα.

Η γενική μορφή των ιδιοτήτων της από κοινού πυκνότητάς τους $P(x_1, x_2, \dots, x_N)$ είναι ως εξής:

$$(\alpha) \quad \sum_{x_1 \in S_{X_1}, x_2 \in S_{X_2}, \dots, x_N \in S_{X_N}} P(x_1, x_2, \dots, x_N) = 1.$$

(β') Για κάθε $i = 1, 2, \dots, N$:

$$P_{X_i}(x_i) = \sum_{x_1 \in S_{X_1}, \dots, x_{i-1} \in S_{X_{i-1}}, x_{i+1} \in S_{X_{i+1}}, \dots, x_N \in S_{X_N}} P(x_1, x_2, \dots, x_N).$$

(γ') Οι X_1, X_2, \dots, X_N είναι ανεξάρτητες αν και μόνο αν,

$$P(x_1, x_2, \dots, x_N) = P_{X_1}(x_1)P_{X_2}(x_2) \cdots P_{X_N}(x_N),$$

για κάθε N -άδα τιμών $x_1 \in S_{X_1}, x_2 \in S_{X_2}, \dots, x_N \in S_{X_N}$.

Οι αποδείξεις είναι παρόμοιες με εκείνες που είδαμε για την περίπτωση των $N = 2$ Τ.Μ. Εφόσον για διαφορετικές N -άδες τιμών (x_1, x_2, \dots, x_N) τα ενδεχόμενα,

$$\{X_1 = x_1, X_2 = x_2, \dots, X_N = x_N\},$$

είναι ξένα μεταξύ τους, και προφανώς η ένωσή τους ισούται με όλο το Ω , από τον κανόνα συνολικής πιθανότητας έχουμε,

$$\begin{aligned} 1 &= \Pr(\Omega) \\ &= \Pr\left(\bigcup_{x_1 \in S_{X_1}, x_2 \in S_{X_2}, \dots, x_N \in S_{X_N}} \{X_1 = x_1, X_2 = x_2, \dots, X_N = x_N\}\right) \\ &= \sum_{x_1 \in S_{X_1}, x_2 \in S_{X_2}, \dots, x_N \in S_{X_N}} \Pr(X_1 = x_1, X_2 = x_2, \dots, X_N = x_N) \\ &= \sum_{x_1 \in S_{X_1}, x_2 \in S_{X_2}, \dots, x_N \in S_{X_N}} P(x_1, x_2, \dots, x_N), \end{aligned}$$

που αποδεικνύει την πρώτη ιδιότητα.

Παρομοίως, για κάθε δεδομένο x_i , τα ενδεχόμενα $\{X_1 = x_1, X_2 = x_2, \dots, X_N = x_N\}$ είναι ξένα μεταξύ τους, και η ένωσή τους για διαφορετικά $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_N$ ισούται με το ενδεχόμενο $\{X_i = x_i\}$. Άρα, πάλι από τον κανόνα συνολικής πιθανότητας, έχουμε τη δεύτερη ιδιότητα: Η $P_{X_i}(x_i) = \Pr(X_i = x_i)$ ισούται με:

$$\begin{aligned} &\Pr\left(\bigcup_{x_1 \in S_{X_1}, \dots, x_{i-1} \in S_{X_{i-1}}, x_{i+1} \in S_{X_{i+1}}, \dots, x_N \in S_{X_N}} \{X_1 = x_1, X_2 = x_2, \dots, X_N = x_N\}\right) \\ &= \sum_{x_1 \in S_{X_1}, \dots, x_{i-1} \in S_{X_{i-1}}, x_{i+1} \in S_{X_{i+1}}, \dots, x_N \in S_{X_N}} \Pr(X_1 = x_1, X_2 = x_2, \dots, X_N = x_N) \\ &= \sum_{x_1 \in S_{X_1}, \dots, x_{i-1} \in S_{X_{i-1}}, x_{i+1} \in S_{X_{i+1}}, \dots, x_N \in S_{X_N}} P(x_1, x_2, \dots, x_N). \end{aligned}$$

Τέλος, η τρίτη ιδιότητα είναι απλά αναδιατυπωμένος ο ορισμός της ανεξαρτησίας N διακριτών τυχαίων μεταβλητών.

4. Χρόνος εκτέλεσης αλγορίθμου.

(α') Κατ' αρχάς, αν ορίσουμε, $Z = \sum_{i=1}^n X_i$, τότε η Z έχει κατανομή $\text{Διων}(n, 1/4)$, και συνεπώς,

$$E(Z) = \frac{n}{4} \quad \text{και} \quad \text{Var}(Z) = n \cdot \frac{1}{4} \cdot \left(1 - \frac{1}{4}\right) = \frac{3n}{16}.$$

Επιπλέον, από την εναλλακτική έκφραση για τη διασπορά στη σχέση (6.8), έχουμε,

$$E(Z^2) = \text{Var}(Z) + [E(Z)]^2 = \frac{3n}{16} + \frac{n^2}{16}.$$

Παρατηρώντας πως ο χρόνος εκτέλεσης T μπορεί να εκφραστεί ως $T = Y + 2Z + Z^2$, από την Ιδιότητα 1 του Θεωρήματος 6.1, βρίσκουμε,

$$\begin{aligned} E(T) &= E(Y + 2Z + Z^2) \\ &= E(Y) + 2E(Z) + E(Z^2) \\ &= \frac{1}{4} \cdot 0 + \frac{1}{4} \cdot 2 + \frac{1}{4} \cdot 5 + \frac{1}{4} \cdot 9 + 2 \cdot \frac{n}{4} + \frac{3n}{16} + \frac{n^2}{16} \\ &= 4 + \frac{11n}{16} + \frac{n^2}{16}. \end{aligned}$$

(β') Από το προηγούμενο ερώτημα και την ανισότητα Markov έχουμε,

$$\Pr(T \geq n^2) \leq \frac{E(T)}{n^2} = \frac{1}{16} + \frac{11}{16n} + \frac{4}{n^2}.$$

(γ') Ήδη γνωρίζουμε πως $E(T) = 4 + \frac{11n}{16} + \frac{n^2}{16}$, πως $E(Y) = \frac{1}{4} \cdot 0 + \frac{1}{4} \cdot 2 + \frac{1}{4} \cdot 5 + \frac{1}{4} \cdot 9 = 4$, και παρομοίως μπορούμε να υπολογίσουμε ότι,

$$E(Y^2) = \frac{1}{4} \cdot 0^2 + \frac{1}{4} \cdot 2^2 + \frac{1}{4} \cdot 5^2 + \frac{1}{4} \cdot 9^2 = 110.$$

Επιπλέον έχουμε $E(Z) = n/4$ και $E(Z^2) = (3n + n^2)/16$. Οπότε, εφαρμόζοντας την εναλλακτική έκφραση για τη συνδιακύμανση (9.2),

$$\begin{aligned} \text{Cov}(T, Y) &= E(TY) - E(T)E(Y) \\ &= E[(Y + 2Z + Z^2)Y] - E(T)E(Y) \\ &\stackrel{(a)}{=} E(Y^2) + 2E(YZ) + E(YZ^2) - E(T)E(Y) \\ &\stackrel{(b)}{=} E(Y^2) + 2E(Y)E(Z) + E(Y)E(Z^2) - E(T)E(Y), \end{aligned}$$

όπου στο βήμα (a) εφαρμόσαμε την πρώτη ιδιότητα του Θεωρήματος 6.1 και στο βήμα (b) την τρίτη ιδιότητα, αφού οι Τ.Μ. Y και Z είναι ανεξάρτητες. Αντικαθιστώντας τις γνωστές τιμές στην παραπάνω έκφραση,

$$\text{Cov}(T, Y) = 110 + 2n + \frac{3n}{4} + \frac{n^2}{4} - 15 - \frac{11n}{4} - \frac{n^2}{4} = 105.$$

Παρατηρούμε κατ' αρχάς ότι η συνδιακύμανση είναι θετική, γεγονός που δεν μας εκπλήσσει λόγω του ορισμού του T ως « Y + κάτι ανεξάρτητο». Άρα, όταν μεγαλώνει ή

μικραίνει το Y , αντιστοίχως θα τείνει να μεγαλώνει ή να μικραίνει και το T . Επίσης παρατηρούμε πως το αποτέλεσμα δεν εξαρτάται από το n , πράγμα που και πάλι βρίσκουμε λογικό: Εφόσον η συσχέτιση μεταξύ του T και του Y δεν περιλαμβάνει άμεσα κάποιον όρο που να εξαρτάται από το n , αναμενόμενο είναι και η συνδιακύμανση μεταξύ τους να είναι ανεξάρτητη του n .

5. Μέγιστο και ελάχιστο δύο ζαριών.

- (α) Η τιμές της από κοινού πυκνότητας $P(x_1, x_2)$ των X_1, X_2 εύκολα υπολογίζονται μέσω της ανεξαρτησίας: Για κάθε $x_1, x_2 = 1, 2, 3, 4, 5, 6$:

$$P(x_1, x_2) = \Pr(X_1 = x_1, X_2 = x_2) = \Pr(X_1 = x_1) \Pr(X_2 = x_2) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}.$$

Η δεύτερη ισότητα προέκυψε λόγω ανεξαρτησίας, και η τρίτη λόγω του ότι το ζάρι είναι δίκαιο.

- (β) Έχουμε ένα πείραμα με $6 \times 6 = 36$ διαφορετικά αποτελέσματα. Επιπλέον, η από κοινού πυκνότητα $P_{X,Y}(x, y)$ των X, Y πρέπει να υπολογιστεί για $6 \times 6 = 36$ ζεύγη τιμών (x, y) . Αρχικά παρατηρούμε πως κάποια από αυτά είναι αδύνατον να εμφανιστούν, π.χ., δεν μπορεί το ελάχιστο από τις δύο ζαριές να είναι 5 και το μέγιστο 2. Συνεπώς, $P_{X,Y}(5, 2) = 0$, και γενικά $P_{X,Y}(x, y) = 0$ αν $x > y$.

Για να υπολογίσουμε καθεμία από τις υπόλοιπες τιμές πρέπει να βρούμε τα αποτελέσματα που της αντιστοιχούν. Για παράδειγμα,

$$P_{X,Y}(1, 1) = \Pr(X_1 = 1, X_2 = 1) = \frac{1}{36},$$

$$P_{X,Y}(1, 2) = \Pr(X_1 = 1, X_2 = 2) + \Pr(X_1 = 2, X_2 = 1) = \frac{2}{36}.$$

Συνεχίζοντας με τον ίδιο τρόπο, συμπληρώνουμε τις τιμές της $P_{X,Y}(x, y)$ στον ακόλουθο πίνακα:

Y	1	2	3	4	5	6	
X							$P_X(x)$
1	1/36	2/36	2/36	2/36	2/36	2/36	11/36
2	0	1/36	2/36	2/36	2/36	2/36	9/36
3	0	0	1/36	2/36	2/36	2/36	7/36
4	0	0	0	1/36	2/36	2/36	5/36
5	0	0	0	0	1/36	2/36	3/36
6	0	0	0	0	0	1/36	1/36
$P_Y(y)$	1/36	3/36	5/36	7/36	9/36	11/36	

(γ') Οι περιθώριες πυκνότητες $P_X(x)$ και $P_Y(y)$ υπολογίζονται στο περιθώριο του πιο πάνω πίνακα, αθροίζοντας τις τιμές της από κοινού πυκνότητας στην αντίστοιχη στήλη ή γραμμή, όπως μας λέει η δεύτερη βασική ιδιότητα της από κοινού πυκνότητας. Προφανώς, όλες μας οι απαντήσεις, παρότι υπολογισμένες με διαφορετικό τρόπο, συμπίπτουν ακριβώς με τα αντίστοιχα αποτελέσματα στην Άσκηση 7 του Κεφαλαίου 6.

6. Συναρτήσεις τυχαίων μεταβλητών.

(α') Έστω πως X, Y είναι ανεξάρτητες. Και οι δύο Τ.Μ. X^2, Y^2 έχουν σύνολο τιμών το $S = \{0, 1, 4, 9, 16, \dots\}$. Για οποιαδήποτε $m, n \geq 1$, από την ανεξαρτησία των X, Y και τον κανόνα συνολικής πιθανότητας έχουμε πως η πιθανότητα $\Pr(X^2 = m^2, Y^2 = n^2)$ ισούται με,

$$\begin{aligned} & \Pr(\{X = m, Y = n\} \cup \{X = m, Y = -n\} \\ & \quad \cup \{X = -m, Y = n\} \cup \{X = -m, Y = -n\}) \\ &= \Pr(X = m, Y = n) + \Pr(X = m, Y = -n) \\ & \quad + \Pr(X = -m, Y = n) + \Pr(X = -m, Y = -n) \\ &= \Pr(X = m) \Pr(Y = n) + \Pr(X = m) \Pr(Y = -n) \\ & \quad + \Pr(X = -m) \Pr(Y = n) + \Pr(X = -m) \Pr(Y = -n) \\ &= \Pr(X = m) [\Pr(Y = n) + \Pr(Y = -n)] \\ & \quad + \Pr(X = -m) [\Pr(Y = n) + \Pr(Y = -n)] \\ &= \Pr(X = m) \Pr(Y^2 = n^2) + \Pr(X = -m) \Pr(Y^2 = n^2) \\ &= [\Pr(X = m) + \Pr(X = -m)] \Pr(Y^2 = n^2) \\ &= \Pr(X^2 = m^2) \Pr(Y^2 = n^2), \end{aligned}$$

δηλαδή, για $m \geq 1$ και $n \geq 1$,

$$\Pr(X^2 = m^2, Y^2 = n^2) = \Pr(X^2 = m^2) \Pr(Y^2 = n^2).$$

Με παρόμοιο τρόπο προκύπτει πως το ίδιο ισχύει και αν το ένα από τα δύο (ή και τα δύο) m, n είναι ίσο με μηδέν. Συνεπώς οι X^2, Y^2 είναι ανεξάρτητες Τ.Μ.

(β') Μία από τις πολλές δυνατές επιλογές είναι η ακόλουθη από κοινού κατανομή των X, Y ,

X	-1	0	1	$P_Y(y)$
-1	0	1/9	0	1/9
0	1/9	1/9	1/9	1/3
1	0	1/9	4/9	5/9
$P_X(x)$	1/9	1/3	5/9	

από την οποία εύκολα βρίσκουμε και την από κοινού κατανομή των X^2, Y^2 :

X^2	0	1	
Y^2			$P_{Y^2}(b)$
0	1/9	2/9	1/3
1	2/9	4/9	2/3
$P_{X^2}(a)$	1/3	2/3	

Παρατηρούμε πως,

$$0 = P_{X,Y}(-1, -1) \neq \frac{1}{9} \cdot \frac{1}{9} = P_X(-1)P_Y(-1) = \frac{1}{9},$$

άρα οι X, Y δεν είναι ανεξάρτητες. Απ' την άλλη μεριά, χρησιμοποιώντας τις τιμές του δεύτερου πίνακα εύκολα μπορούμε να επιβεβαιώσουμε πως ισχύει,

$$\Pr(X^2 = a, Y^2 = b) = \Pr(X^2 = a) \Pr(Y^2 = b),$$

για κάθε ζευγάρι τιμών (a, b) με $a, b = 0$ ή 1 , και κατά συνέπεια πως τα X^2, Y^2 είναι ανεξάρτητα.

- (γ') Το αντίστροφο μέρος του ερωτήματος είναι προφανές: Αν οι $f(X)$ και $g(Y)$ είναι ανεξάρτητες για οποιεσδήποτε συναρτήσεις f και g , τότε επιλέγοντας $f(x) = x$ και $g(y) = y$, συμπεραίνουμε πως και οι X, Y είναι ανεξάρτητες.

Για το ορθό μέρος, έστω μια Τ.Μ. X με πυκνότητα $P_X(x)$ στο σύνολο τιμών S_X , και έστω μια συνάρτηση $f: S_X \rightarrow S'_X$. Τότε η νέα Τ.Μ. $f(X)$ έχει σύνολο τιμών το S'_X και η πυκνότητά της $P_{f(X)}(z)$ έχει την εξής ιδιότητα: Για κάθε $z \in S'_X$ υπάρχει ένα ή περισσότερα x τέτοια ώστε $f(x) = z$. Έστω A_z το σύνολο αυτών των x :

$$A_z = \{x \in S_X : f(x) = z\}.$$

Τώρα η πυκνότητα $P_{f(X)}(z)$ μπορεί να εκφραστεί ως,

$$P_{f(X)}(z) = \Pr(f(X) = z) = \Pr(X \in A_z). \quad (\text{A'.13})$$

Με ακριβώς το ίδιο σκεπτικό, η $g(Y)$ έχει σύνολο τιμών το S'_Y και πυκνότητα $P_{g(Y)}(v)$ τέτοια ώστε,

$$P_{g(Y)}(v) = \Pr(g(Y) = v) = \Pr(Y \in B_v), \quad (\text{A'.14})$$

όπου τα σύνολα B_v ορίζονται ως,

$$B_v = \{y \in S_Y : g(y) = v\}.$$

Η ανεξαρτησία των $f(X), g(Y)$ επιβεβαιώνεται εύκολα ως εξής: Για οποιοσδήποτε τιμές $z \in S'_X, v \in S'_Y$,

$$\begin{aligned} \Pr(f(X) = z, g(Y) = v) &= \Pr(X \in A_z, Y \in B_v) \\ &\stackrel{(a)}{=} \Pr(X \in A_z) \Pr(Y \in B_v) \\ &\stackrel{(b)}{=} P_{f(X)}(z) P_{g(Y)}(v), \end{aligned}$$

όπου η ισότητα (a) προκύπτει από την Ιδιότητα 5 του Θεωρήματος 6.1, και η (b) από τις σχέσεις (A.13) και (A.14) πιο πάνω. Συνεπώς οι τυχαίες μεταβλητές $f(X), g(Y)$ είναι πράγματι ανεξάρτητες.

Αυτό το αποτέλεσμα εκ πρώτης όψεως φαίνεται να αντικρούει το αποτέλεσμα του προηγούμενου σκέλους: Αφού οι X^2 και Y^2 είναι ανεξάρτητες, δεν θα έπρεπε και οι X, Y να είναι ανεξάρτητες; Η απάντηση είναι «όχι απαραίτητα»: Όταν τα X, Y παίρνουν και θετικές και αρνητικές τιμές, το X δεν είναι απαραίτητα συνάρτηση του X^2 (και αντίστοιχα για το Y), μια που γνωρίζοντας, π.χ., ότι $X^2 = 1$, το μόνο που μπορούμε να συμπεράνουμε είναι ότι το X ισούται με $+1$ ή με -1 , αλλά δεν ξέρουμε απαραίτητα ποια είναι η τιμή του.

7. ★ Το φράγμα του Chernoff.

(α') Κατ' αρχάς, εφόσον το λ είναι θετικό, έχουμε,

$$\Pr(X \geq c) = \Pr(\lambda X \geq \lambda c) = \Pr(e^{\lambda X} \geq e^{\lambda c}) = \Pr(Y \geq e^{\lambda c}),$$

όπου έχουμε ορίσει τη νέα Τ.Μ. $Y = e^{\lambda X}$. Εφόσον η Y παίρνει πάντα τιμές $Y \geq 0$, χρησιμοποιώντας την ανισότητα του Markov για την Y , έχουμε:

$$\Pr(X \geq c) = \Pr(Y \geq e^{\lambda c}) \leq \frac{E(Y)}{e^{\lambda c}} = e^{-\lambda c} E(e^{\lambda X}),$$

που είναι ακριβώς η ζητούμενη σχέση.

(β') Έστω τώρα πως $X \sim \text{Διων}(n, 1/2)$. Τότε, όπως είδαμε στην Ενότητα 7.1, η X μπορεί να εκφραστεί ως το άθροισμα $X = \sum_{i=1}^n Y_i$, όπου οι Y_1, Y_2, \dots, Y_n είναι ανεξάρτητες $\text{Bern}(1/2)$ Τ.Μ. Άρα, εφαρμόζοντας τη σχέση (9.4) του προηγούμενου σκέλους στη X , με τη σταθερά nc στη θέση της c , βρίσκουμε πως,

$$\Pr(X \geq nc) \leq e^{-\lambda nc} E(e^{\lambda X}) = [e^{-\lambda c}]^n E(e^{\lambda \sum_{i=1}^n Y_i}) = [e^{-\lambda c}]^n E\left(\prod_{i=1}^n e^{\lambda Y_i}\right).$$

Από το τρίτο σκέλος της Άσκησης 6 παραπάνω γνωρίζουμε πως, εφόσον οι Τ.Μ. Y_i είναι ανεξάρτητες, θα είναι και οι Τ.Μ. $e^{\lambda Y_i}$ ανεξάρτητες. Συνεπώς, χρησιμοποιώντας

την Ιδιότητα 3 του Θεωρήματος 6.1, έχουμε,

$$\Pr(X \geq nc) \leq \left[e^{-\lambda c} \right]^n \prod_{i=1}^n E \left(e^{\lambda Y_i} \right).$$

Επιπλέον, αφού όλα τα Y_i έχουν την ίδια κατανομή, όλοι οι όροι του πιο πάνω γινομένου είναι ίσοι μεταξύ τους και όλοι ίσοι με $E(e^{\lambda Y_1})$. Άρα,

$$\Pr(X \geq nc) \leq \left[e^{-\lambda c} \right]^n \left[E \left(e^{\lambda Y_1} \right) \right]^n = \left\{ e^{-\lambda c} \left[\frac{1}{2} e^0 + \frac{1}{2} e^\lambda \right] \right\}^n = \left\{ \frac{1}{2} e^{-\lambda c} [1 + e^\lambda] \right\}^n.$$

(γ') Έστω τώρα πως $X \sim \text{Διων}(10, 1/2)$.

i. Για $c = 0.9$, το φράγμα (9.5) που μόλις αποδείξαμε μας δίνει,

$$\Pr(X \geq 9) \leq \left\{ \frac{1}{2} e^{-0.9\lambda} [1 + e^\lambda] \right\}^{10}.$$

Για να βρούμε το καλύτερο δυνατό φράγμα αρκεί να υπολογίσουμε την τιμή του λ που ελαχιστοποιεί το δεξί μέρος της παραπάνω σχέσης. Για να βρούμε αυτή την τιμή, αρκεί να ελαχιστοποιήσουμε τη συνάρτηση $f(\lambda) = e^{-\lambda c}(1 + e^\lambda)$ ως προς λ . Παίρνοντας παραγώγους, εύκολα βρίσκουμε πως,

$$f'(\lambda) = e^{-\lambda c} [(1 - c)e^\lambda - c],$$

το οποίο ισούται με μηδέν αν και μόνο αν $(1 - c)e^\lambda = c$, δηλαδή για την τιμή,

$$\lambda^* = \log \left(\frac{c}{1 - c} \right) = \log 9.$$

Επιπλέον, η δεύτερη παράγωγος ισούται με

$$f''(\lambda) = e^{-\lambda c} [(1 - c)^2 e^\lambda + c^2],$$

το οποίο είναι προφανώς πάντα θετικό, άρα η τιμή λ^* πράγματι ελαχιστοποιεί την $f(\lambda)$. Αντικαθιστώντας αυτή την τιμή στο φράγμα μας,

$$\Pr(X \geq 9) \leq \left\{ \frac{1}{2} e^{-0.9 \log 9} [1 + e^{\log 9}] \right\}^{10} \approx 0.0252.$$

ii. Η Τ.Μ. X έχει μέση τιμή $E(X) = 10 \times (1/2) = 5$, όποτε η ανισότητα του Markov μάς δίνει,

$$\Pr(X \geq 9) \leq \frac{E(X)}{9} = \frac{5}{9} \approx 0.5556,$$

το οποίο είναι σαφώς σημαντικά ασθενέστερο.

iii. Τέλος, η ζητούμενη πιθανότητα $\Pr(X \geq 9)$ μπορεί εύκολα να υπολογιστεί ακριβώς. Εφόσον το X έχει σύνολο τιμών το $\{0, 1, \dots, 10\}$, από τον τύπο της πυκνότητας της διωνυμικής κατανομής έχουμε,

$$\Pr(X \geq 9) = P(9) + P(10) = \binom{10}{9} \left(\frac{1}{2}\right)^9 \left(1 - \frac{1}{2}\right)^1 + \binom{10}{10} \left(\frac{1}{2}\right)^{10} \left(1 - \frac{1}{2}\right)^0,$$

το οποίο ισούται με, $\Pr(X \geq 9) = 11 \times \left(\frac{1}{2}\right)^{10} \approx 0.0107$. Προφανώς το φράγμα του Chernoff εδώ μας δίνει ένα αποτέλεσμα πολύ πιο ισχυρό και πιο κοντά στην πραγματική τιμή από ό,τι η ανισότητα του Markov.

8. Συμπλήρωση πίνακα.

(α) Από τις απλές ιδιότητες της από κοινού κατανομής που είδαμε στην Ενότητα 9.2, εύκολα βρίσκουμε:

X	-5	-2	2	5	
Y					$P_Y(y)$
0	0.01	0.01	0.01	0.09	0.12
1	0.05	0.09	0.05	0	0.19
2	0.07	0.07	0.09	0.07	0.30
3	0.15	0.09	0	0.15	0.39
$P_X(x)$	0.28	0.26	0.15	0.31	

(β) Χρησιμοποιώντας τις σχετικές τιμές του πίνακα διαπιστώνουμε πως,

$$\Pr(X = 5, Y = 2) = 0.07 \neq 0.093 = \Pr(X = 5) \Pr(Y = 2),$$

άρα οι X, Y , δεν είναι ανεξάρτητες.

(γ) Από τον ορισμό της δεσμευμένης πιθανότητας και τις τιμές του παραπάνω πίνακα, έχουμε,

$$\Pr(Y = 1|X = 2) = \frac{\Pr(Y = 1, X = 2)}{\Pr(X = 2)} = \frac{0.05}{0.15} = \frac{1}{3}.$$

9. Ενεργοποιημένες συνδέσεις. Σύμφωνα με την ανισότητα του Chebychev έχουμε ότι:

$$\Pr(|X - \mu| \geq s) \leq \frac{\sigma^2}{s^2}.$$

Θέλουμε να βρούμε μια τιμή για το s ώστε το ενδεχόμενο να πέσει το δίκτυο να έχει πιθανότητα το πολύ 1%, δηλαδή θέλουμε

$$\Pr(|X - \mu| \geq s) \leq 0.01.$$

Επομένως, θέτοντας $\sigma^2/s^2 = 0.01$ πετυχαίνουμε το επιθυμητό αποτέλεσμα, και η ζητούμενη τιμή του s είναι,

$$s^2 = \frac{\sigma^2}{0.01} = 100 \times \sigma^2 = 100 \times 500^2,$$

δηλαδή $s = 5000$.

10. *Συντελεστής συσχέτισης.* Κατ' αρχάς, από την εναλλακτική έκφραση για τη συνδιακύμανση (9.2), έχουμε,

$$\begin{aligned} \text{Cov}(X, Y) &= E(XY) - E(X)E(Y) \\ &= E(X(aX + b)) - E(X)E(aX + b) \\ &= E(aX^2 + bX) - E(X)(aE(X) + b) \\ &= aE(X^2) + bE(X) - aE(X)^2 - bE(X) \\ &= a\text{Var}(X), \end{aligned}$$

όπου χρησιμοποιήσαμε την πρώτη ιδιότητα του Θεωρήματος 6.1, και την εναλλακτική έκφραση για τη διασπορά (6.8). Επίσης, πάλι από το Θεώρημα 6.1, έχουμε,

$$\text{Var}(Y) = \text{Var}(aX + b) = a^2\text{Var}(X),$$

συνεπώς,

$$\begin{aligned} \rho_{X,Y} &= \frac{\text{Cov}(X, Y)}{[\text{Var}(X)\text{Var}(Y)]^{\frac{1}{2}}} \\ &= \frac{a\text{Var}(X)}{[\text{Var}(X)a^2\text{Var}(X)]^{\frac{1}{2}}} \\ &= \frac{a}{\sqrt{a^2}} = \frac{a}{|a|}, \end{aligned}$$

το οποίο φυσικά ισούται με $+1$ αν το $a > 0$ και με -1 αν το $a < 0$.

11. *Ασυσχέτιστες αλλά όχι ανεξάρτητες Τ.Μ.* Η Τ.Μ. V έχει σύνολο τιμών το $\{0, 1, 2\}$ και η W έχει σύνολο τιμών το $\{0, 1\}$. Αρχικά υπολογίζουμε τις τιμές της κοινού πυκνότητας $P_{V,W}(v, w)$ των V, W :

$$\begin{aligned} P_{V,W}(0, 0) &= \Pr(X = 0, Y = 0) = \Pr(X = 0) \Pr(Y = 0) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}, \\ P_{V,W}(1, 1) &= \Pr(X = 1, Y = 0) + \Pr(X = 0, Y = 1) \\ &= \Pr(X = 1) \Pr(Y = 0) + \Pr(X = 0) \Pr(Y = 1) \\ &= \frac{1}{4} + \frac{1}{4} = \frac{1}{2}, \\ P_{V,W}(2, 0) &= \Pr(X = 1, Y = 1) = \Pr(X = 1) \Pr(Y = 1) = \frac{1}{4}. \end{aligned}$$

Καθώς είναι αδύνατον να έχουμε $V = 0, W = 1$, η πιθανότητα $P_{V,W}(0,1) = 0$, και παρομοίως, $P_{V,W}(1,0) = P_{V,W}(2,1) = 0$.

Έχοντας υπολογίσει τις τιμές της από κοινού πυκνότητας, μπορούμε να υπολογίσουμε και τις περιθώριες πυκνότητες:

$$\begin{aligned} P_V(0) &= \frac{1}{4}, & P_V(1) &= \frac{1}{2}, & P_V(2) &= \frac{1}{4}, \\ P_W(0) &= \frac{1}{2}, & P_W(1) &= \frac{1}{2}. \end{aligned}$$

Από τα παραπάνω εύκολα προκύπτει πως οι V, W δεν είναι ανεξάρτητες. Για παράδειγμα,

$$0 = P_{V,W}(1,0) \neq P_V(1)P_W(0) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}.$$

Τέλος, για να δείξουμε ότι έχουν μηδενική συνδιακύμανση, υπολογίζουμε τις μέσες τιμές,

$$\begin{aligned} E(V) &= E(X + Y) = E(X) + E(Y) = \frac{1}{2} + \frac{1}{2} = 1, \\ E(W) &= 0 \times P_W(0) + 1 \times P_W(1) = \frac{1}{2}, \\ E(VW) &= (0 \times 0) \times P_{V,W}(0,0) + (1 \times 0) \times P_{V,W}(1,0) + (2 \times 0) \times P_{V,W}(2,0) \\ &\quad + (0 \times 1) \times P_{V,W}(0,1) + (1 \times 1) \times P_{V,W}(1,1) + (2 \times 1) \times P_{V,W}(2,1) \\ &= \frac{1}{2}. \end{aligned}$$

Και από την εναλλακτική έκφραση (9.2) βρίσκουμε πως η συνδιακύμανση,

$$\text{Cov}(V, W) = E(VW) - E(V)E(W) = \frac{1}{2} - 1 \cdot \frac{1}{2} = 0,$$

είναι πράγματι ίση με μηδέν.

12. Σύγκλιση κατά πιθανότητα.

(α') Εφόσον η κάθε $X_n \sim \text{Bern}(1/n)$, για αυθαίρετο $\epsilon > 0$ έχουμε,

$$\Pr(|X_n - 0| < \epsilon) = \Pr(X_n < \epsilon).$$

Αν το $\epsilon \geq 1$, τότε η πιθανότητα είναι ίση με 1 για κάθε n , ενώ αν $\epsilon < 1$ τότε,

$$\Pr(|X_n - 0| < \epsilon) = \Pr(X_n < \epsilon) = \Pr(X_n = 0) = 1 - \frac{1}{n},$$

το οποίο προφανώς τείνει στο 1 καθώς το $n \rightarrow \infty$. Άρα, έχουμε $X_n \rightarrow 0$ κατά πιθανότητα καθώς $n \rightarrow \infty$.

(β') Εδώ έχουμε μια σημαντική παρατήρηση: Ο ορισμός της σύγκλισης κατά πιθανότητα σε μια σταθερά εξαρτάται μόνο από πιθανότητες της μορφής $\Pr(|X_n - c| < \epsilon)$, οι οποίες αφορούν μόνο μία μεταβλητή και συνεπώς εξαρτώνται μόνο από τις περιθώριες κατανομές των $\{X_n\}$. Άρα, το πιο πάνω αποτέλεσμα ισχύει ανεξαρτήτως του αν οι Τ.Μ. $\{X_n\}$ είναι ανεξάρτητες ή όχι.

(γ') Αντίθετα με το προηγούμενο ερώτημα, για το αποτέλεσμα του Ν.Μ.Α. είναι πολύ σημαντική η υπόθεση ότι οι $\{X_n\}$ είναι ανεξάρτητες, και δεν μπορεί να παραλειφθεί εντελώς. Για παράδειγμα, έστω ότι τα $\{X_n\}$ έχουν όλα κατανομή $\text{Bern}(1/2)$. Αν είναι ανεξάρτητα, τότε ο εμπειρικός τους μέσος όρος τείνει στο $1/2$ κατά πιθανότητα, όπως μας λέει ο Ν.Μ.Α. Αντίθετα, αν όλα τα X_n είναι ίσα με το X_1 , τότε και ο εμπειρικός τους μέσος όρος είναι ίσος με το X_1 για κάθε n , το οποίο προφανώς δεν συγκλίνει στο $1/2!$

Ο λόγος για τον οποίο η ανεξαρτησία εδώ είναι καίρια, είναι διότι, αν αλλάξει η από κοινού κατανομή των $\{X_n\}$, τότε αλλάζει και η περιθώρια κατανομή του εμπειρικού μέσου όρου \bar{X}_n .

13. Εκτίμηση με θόρυβο.

(α') Εφόσον τα Z_i έχουν μέση τιμή ν , η κάθε Τ.Μ. $Z_i - \nu$ έχει μέση τιμή μηδέν. Και επιπλέον, αφού όλες οι Τ.Μ. $\{X_n, Z_n\}$ είναι ανεξάρτητες, τότε και οι νέες τυχαίες μεταβλητές $W_i = Y_i - \nu = X_i + Z_i - \nu$ σχηματίζουν μια ακολουθία ανεξάρτητων Τ.Μ. όπου όλες έχουν την ίδια κατανομή και συνεπώς την ίδια μέση τιμή,

$$E(W_i) = E(X_i + Z_i - \nu) = E(X_i) + E(Z_i) - \nu = \mu.$$

Άρα, αν αποφασίσουμε να εκτιμήσουμε το μ μέσω του εμπειρικού μέσου όρου,

$$\bar{W}_N = \frac{1}{N} \sum_{i=1}^N W_i = \frac{1}{N} \sum_{i=1}^N X_i + \frac{1}{N} \sum_{i=1}^N Z_i - \nu,$$

ο Ν.Μ.Α. μάς εγγυάται πως θα έχουμε $\bar{W}_N \rightarrow \mu$ κατά πιθανότητα, καθώς $N \rightarrow \infty$.

(β') Παρομοίως και εδώ ορίζουμε τις νέες Τ.Μ. $V_i = Y_i' - \nu = X + Z_i - \nu$, και ως εκτιμητρία του X χρησιμοποιούμε τον εμπειρικό τους μέσο όρο,

$$\bar{V}_N = \frac{1}{N} \sum_{i=1}^n V_i = X + \frac{1}{N} \sum_{i=1}^N (Z_i - \nu). \quad (\text{A'.15})$$

Εφόσον οι Τ.Μ. $Z_i - \nu$ είναι ανεξάρτητες, έχουν την ίδια κατανομή και έχουν μέση τιμή ίση με μηδέν, ο Ν.Μ.Α. μάς λέει πως ο δεύτερος όρος στο δεξί μέρος της (A'.15)

τείνει στο μηδέν, κατά πιθανότητα. Άρα διαισθητικά περιμένουμε πως θα έχουμε και $\bar{V}_N \rightarrow X$ κατά πιθανότητα, καθώς $N \rightarrow \infty$. Πράγματι, για αυθαίρετο $\epsilon > 0$,

$$\begin{aligned} \Pr(|\bar{V}_N - X| < \epsilon) &= \Pr\left(\left|X + \frac{1}{N} \sum_{i=1}^N (Z_i - \nu) - X\right| < \epsilon\right) \\ &= \Pr\left(\left|\frac{1}{N} \sum_{i=1}^N Z_i - \nu\right| < \epsilon\right), \end{aligned}$$

όπου η τελευταία πιθανότητα πιο πάνω τείνει στο 1 καθώς $N \rightarrow \infty$, από τον Ν.Μ.Α. για τον εμπειρικό μέσο όρο των Τ.Μ. $\{Z_i\}$. Άρα, πράγματι η εκτίμησή μας είναι «ασυμπτωτικά συνεπής», δηλαδή $\bar{V}_N \rightarrow X$ κατά πιθανότητα, καθώς $N \rightarrow \infty$.

14. ★ Σύγκλιση με πιθανότητα 1.

(α) Αν για κάποιο $\omega \in \Omega$ ισχύει ότι $|X_n(\omega) - X(\omega)| < \epsilon$ για όλα τα $n \geq m$, προφανώς το ίδιο θα ισχύει και για όλα τα $n \geq m + 1$. Άρα κάθε ω που ανήκει στο T_m , ανήκει και στο T_{m+1} , συνεπώς η ακολουθία των ενδεχομένων $\{T_m\}$ είναι αύξουσα.

Επιπλέον, αν για κάποιο συγκεκριμένο ω έχουμε $\lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)$, τότε από τον κλασικό ορισμό του ορίου, για κάθε $\epsilon > 0$ υπάρχει κάποιο $m \geq 1$ τέτοιο ώστε $|X_n(\omega) - X(\omega)| < \epsilon$ για όλα τα $n \geq m$. Με άλλα λόγια, αν κάποιο ω ανήκει στο Σ , τότε σίγουρα θα ανήκει και σε κάποιο από τα T_m . Άρα έχουμε ότι ισχύει η ζητούμενη σχέση (9.6).

(β) Από τον δεύτερο κανόνα πιθανότητας και τη σχέση (9.6) έχουμε πως,

$$1 = \mathbb{P}(\Sigma) \leq \mathbb{P}\left(\bigcup_{m=1}^{\infty} T_m\right),$$

όπου το γεγονός ότι $\mathbb{P}(\Sigma) = 1$ προκύπτει από την υπόθεση ότι οι $\{X_n\}$ συγκλίνουν στη X με πιθανότητα 1. Άρα η πιθανότητα της παραπάνω ένωσης ισούται κι αυτή με 1, και, δεδομένου του προηγούμενου σκέλους, από την Άσκηση 9 (α) του Κεφαλαίου 3 συμπεραίνουμε πως,

$$\lim_{n \rightarrow \infty} \mathbb{P}(T_m) = \mathbb{P}\left(\bigcup_{m=1}^{\infty} T_m\right) = 1.$$

(γ) Τέλος, παρατηρούμε πως, για κάθε m ,

$$T_m = \{|X_n - X| < \epsilon \text{ για κάθε } n \geq m\} \subset \{|X_m - X| < \epsilon\},$$

άρα, χρησιμοποιώντας τον δεύτερο κανόνα πιθανότητας στο αποτέλεσμα του πιο πάνω ερωτήματος έχουμε,

$$\Pr(|X_m - X| < \epsilon) \geq \Pr(T_m) \rightarrow 1, \text{ καθώς } m \rightarrow \infty,$$

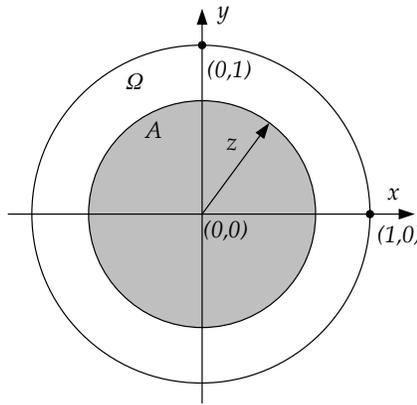
δηλαδή,

$$\lim_{n \rightarrow \infty} \Pr(|X_n - X| < \epsilon) = 1.$$

Και αφού το $\epsilon > 0$ που επιλέξαμε στην αρχή ήταν αυθαίρετο, έχουμε πράγματι αποδείξει ότι οι $\{X_n\}$ τείνουν στη X κατά πιθανότητα.

A'.10 Ασκήσεις Κεφαλαίου 10

1. Κυκλικός δίσκος. Προφανώς οι δυνατές τιμές για το Z ανήκουν στο διάστημα $[0, 1]$, οπότε $S_Z = [0, 1]$. Για τη συνάρτηση κατανομής, θα έχουμε $Z \leq z$ αν και μόνο αν το επιλεγμένο σημείο είναι εντός του κυκλικού δίσκου A με ακτίνα z που εμφανίζεται σκιασμένος στο Σχήμα A'.29.



Σχήμα A'.29: Το ενδεχόμενο $\{Z \leq z\}$ στην Άσκηση 1.

Εφόσον το σημείο επιλέγεται ομοιόμορφα ανάμεσα σε όλα τα σημεία του Ω , η πιθανότητα να έχουμε $Z \leq z$ ισούται με το εμβαδόν του σκιασμένου δίσκου προς το εμβαδόν όλου του Ω . Άρα, για $z \in [0, 1]$, έχουμε,

$$F(z) = \Pr(Z \leq z) = \frac{\pi z^2}{\pi} = z^2,$$

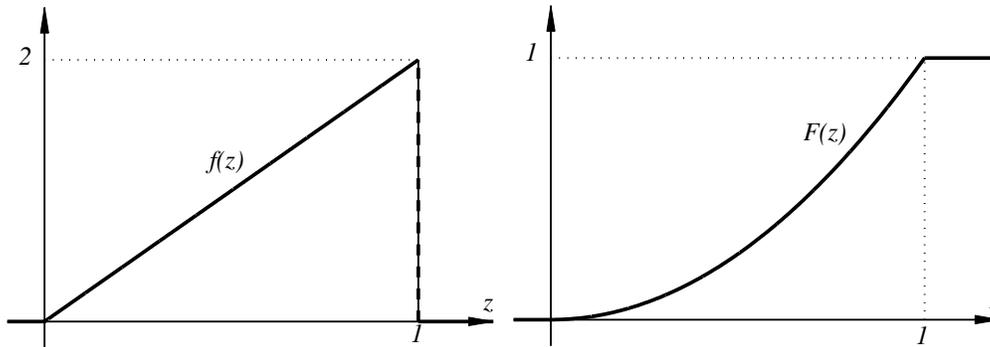
και γενικά,

$$F(z) = \begin{cases} 0, & z < 0, \\ z^2, & 0 \leq z \leq 1, \\ 1, & z > 1. \end{cases}$$

Τέλος για την πυκνότητα, από τη σχέση (10.7) έχουμε,

$$f(z) = F'(z) = \begin{cases} 2z, & z \in [0, 1], \\ 0, & z \notin [0, 1]. \end{cases}$$

Οι $F(z)$, $f(z)$, έχουν σχεδιαστεί στο Σχήμα A'.30.



Σχήμα Α΄.30: Η πυκνότητα και η συνάρτηση κατανομής της Τ.Μ. Z στην Άσκηση 1.

2. Προσδιορισμός παραμέτρων.

(α) Από την τρίτη ιδιότητα της συνάρτησης κατανομής, θα πρέπει $\lim_{x \rightarrow \infty} F(x) = 1$, οπότε θα πρέπει η A να ισούται με μηδέν γιατί διαφορετικά $\lim_{x \rightarrow \infty} F(x) = \pm\infty$. Άρα, για $x \geq 4$, $F(x) = B - 4/x$, και το όριο της $F(x)$ καθώς το $x \rightarrow \infty$ είναι, $\lim_{x \rightarrow \infty} (B - 4/x) = B$. Συνεπώς πρέπει να έχουμε και $B = 1$, δηλαδή,

$$F(x) = \begin{cases} 0, & x < 4, \\ 1 - \frac{4}{x}, & x \geq 4. \end{cases}$$

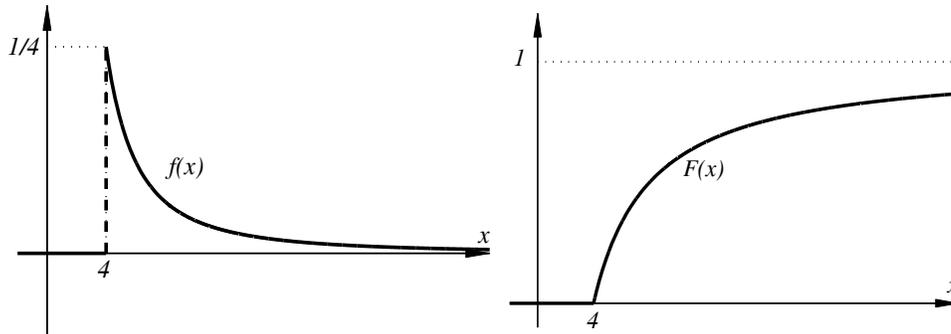
(β) Για την πυκνότητα, από τη σχέση (10.7) βρίσκουμε,

$$f(x) = F'(x) = \begin{cases} \frac{4}{x^2}, & x \geq 4, \\ 0, & x < 4. \end{cases}$$

Η γραφική αναπαράσταση των $f(x)$ και $F(x)$ δίνεται στο Σχήμα Α΄.31.

(γ) Από τον ορισμό της δεσμευμένης πιθανότητας, και χρησιμοποιώντας τη συνάρτηση κατανομής $F(x)$,

$$\begin{aligned} \Pr(X < 5 | X < 6) &= \frac{\Pr(\{X < 5\} \cap \{X < 6\})}{\Pr(X < 6)} \\ &= \frac{\Pr(X < 5)}{\Pr(X < 6)} \\ &= \frac{F(5)}{F(6)} \\ &= \frac{1 - 4/5}{1 - 4/6} = \frac{1/5}{1/3} = 0.6. \end{aligned}$$



Σχήμα Α'.31: Γραφική αναπαράσταση της πυκνότητας $f(x)$ και της συνάρτησης κατανομής $F(x)$ στην Άσκηση 2.

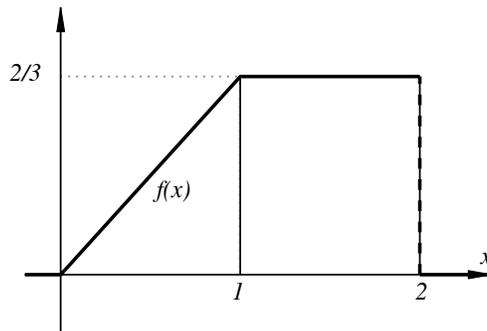
3. Μια απλή πυκνότητα.

(α) Το ολοκλήρωμα της πυκνότητας από το $-\infty$ ως το $+\infty$ πρέπει να ισούται με 1, αλλά,

$$\int_{-\infty}^{\infty} f(x) dx = \int_0^1 cx dx + \int_1^2 c dx = \left[c \frac{x^2}{2} \right]_0^1 + [cx]_0^1 = \frac{1}{2}c + c = \frac{3}{2}c,$$

οπότε πρέπει να έχουμε $c = 2/3$. Άρα η πυκνότητα $f(x)$, όπως αναπαρίσταται και στο Σχήμα Α'.32, είναι:

$$f(x) = \begin{cases} 0, & x < 0, \\ \frac{2}{3}x, & 0 \leq x \leq 1, \\ \frac{2}{3}, & 1 \leq x \leq 2, \\ 0, & x > 2. \end{cases}$$



Σχήμα Α'.32: Η πυκνότητα $f(x)$ στην Άσκηση 3.

(β') Η ζητούμενη πιθανότητα εύκολα υπολογίζεται από την πυκνότητα,

$$\begin{aligned}
 \Pr(\{X > 1.5\} \cup \{X < 0.5\}) &= \Pr(X > 1.5) + \Pr(X < 0.5) \\
 &= \int_{1.5}^{\infty} f(x) dx + \int_{-\infty}^{0.5} f(x) dx \\
 &= \int_{1.5}^2 \frac{2}{3} dx + \int_0^{0.5} \frac{2}{3} x dx \\
 &= \left[\frac{2}{3} x \right]_{1.5}^2 + \left[\frac{2}{3} \cdot \frac{x^2}{2} \right]_0^{0.5} \\
 &= \frac{4}{3} - 1 + \frac{1}{12} - 0 \\
 &= \frac{5}{12}.
 \end{aligned}$$

(γ') Από τον ορισμό της μέσης τιμής εύκολα βρίσκουμε,

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_0^1 \frac{2x^2}{3} dx + \int_1^2 \frac{2x}{3} dx = \left[\frac{2x^3}{9} \right]_0^1 + \left[\frac{2x^2}{6} \right]_1^2 = \frac{11}{9}.$$

4. Μέση τιμή συναρτήσεων T.M.

(α') Από τον τύπο (10.10) έχουμε,

$$\begin{aligned}
 E[g(X)] &= \int_{-\infty}^{\infty} g(x) f(x) dx \\
 &= \int_0^1 (-1) \cdot e^{-x} dx + \int_1^5 0 \cdot e^{-x} dx + \int_5^{\infty} 1 \cdot e^{-x} dx \\
 &= [e^{-x}]_0^1 + 0 + [-e^{-x}]_5^{\infty} \\
 &= e^{-1} + e^{-5} - 1 \approx -0.6254.
 \end{aligned}$$

(β') Η διακριτή T.M. $Y = g(X)$ έχει σύνολο τιμών $S_Y = \{-1, 0, 1\}$ και οι τιμές της πυκνότητάς της υπολογίζονται εύκολα ως εξής,

$$\begin{aligned}
 P(-1) &= \Pr(Y = -1) = \Pr(g(X) = -1) = \Pr(0 \leq X < 1) \\
 &= \int_0^1 e^{-x} dx = [-e^{-x}]_0^1 = 1 - e^{-1},
 \end{aligned}$$

και παρομοίως,

$$\begin{aligned}
 P(0) &= \Pr(Y = 0) = \Pr(g(X) = 0) = \Pr(1 \leq X \leq 5) \\
 &= \int_1^5 e^{-x} dx = [-e^{-x}]_1^5 = e^{-1} - e^{-5},
 \end{aligned}$$

και,

$$\begin{aligned} P(1) &= \Pr(Y = 1) = \Pr(g(X) = 1) = \Pr(X > 5) \\ &= \int_5^{\infty} e^{-x} dx = [-e^{-x}]_5^{\infty} = e^{-5}. \end{aligned}$$

Εφαρμόζοντας τώρα τον ορισμό της μέσης τιμής μιας διακριτής Τ.Μ., βρίσκουμε,

$$\begin{aligned} E(Y) &= (-1) \cdot P(-1) + 0 \cdot P(0) + 1 \cdot P(1) \\ &= (-1) \cdot (1 - e^{-1}) + e^{-5} \\ &= e^{-1} + e^{-5} - 1, \end{aligned}$$

το οποίο φυσικά ισούται ακριβώς με την τιμή της $E[g(X)]$ που υπολογίσαμε παραπάνω.

Στην περίπτωση που η X είναι μια οποιαδήποτε διακριτή Τ.Μ., στην Άσκηση 1 του Κεφαλαίου 6 αποδείξαμε το αντίστοιχο αποτέλεσμα για τη μέση τιμή $E[g(X)]$ μιας αυθαίρετης συνάρτησης της X .

5. Άπειρη μέση τιμή.

(α') Από τον Ορισμό 10.1 γνωρίζουμε πως το ολοκλήρωμα της πυκνότητας $f(x)$ για όλα τα $x \in \mathbb{R}$ πρέπει να ισούται με 1, συνεπώς,

$$1 = \int_{-\infty}^{\infty} f(x) dx = \int_2^{\infty} \frac{C}{x^2} dx = \left[-\frac{C}{x} \right]_2^{\infty} = \frac{C}{2},$$

άρα έχουμε $C = 2$.

(β') Από τον ορισμό της μέσης τιμής,

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_2^{\infty} \frac{2}{x} dx = 2 \int_2^{\infty} \frac{1}{x} dx = \infty,$$

μια που, ως γνωστόν, το καταχρηστικό ολοκλήρωμα $\int_a^{\infty} \frac{1}{x} dx = \infty$ για κάθε $a > 0$:

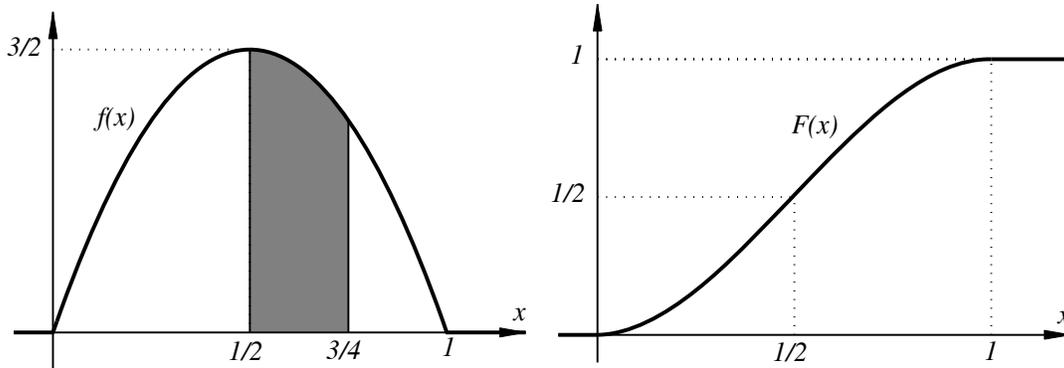
$$\int_a^{\infty} \frac{1}{x} dx = \lim_{b \rightarrow \infty} \int_a^b \frac{1}{x} dx = \lim_{b \rightarrow \infty} [\log x]_a^b = \lim_{b \rightarrow \infty} (\log b - \log a) = \infty.$$

6. Κατανομή Βήτα.

(α') Η σταθερά c πρέπει να έχει τιμή τέτοια ώστε το ολοκλήρωμα της πυκνότητας να ισούται με τη μονάδα,

$$\int_{-\infty}^{\infty} f(x) dx = \int_0^1 cx(1-x) dx = c \int_0^1 x - x^2 dx = c \left[\frac{x^2}{2} - \frac{x^3}{3} \right]_0^1 = c \left[\frac{1}{2} - \frac{1}{3} \right] = \frac{c}{6},$$

Συνεπώς $c = 6$. Η πυκνότητα έχει σχεδιαστεί στο Σχήμα Α'.33.



Σχήμα Α΄.33: Η πυκνότητα και η συνάρτηση κατανομής της τυχαίας μεταβλητής X στην Άσκηση 6.

(β΄) Παρομοίως, η ζητούμενη πιθανότητα είναι,

$$\begin{aligned}
 \Pr\left(\frac{1}{2} \leq X \leq \frac{3}{4}\right) &= \int_{\frac{1}{2}}^{\frac{3}{4}} f(x) dx \\
 &= 6 \int_{\frac{1}{2}}^{\frac{3}{4}} x(1-x) dx \\
 &= 6 \left[\frac{x^2}{2} - \frac{x^3}{3} \right]_{\frac{1}{2}}^{\frac{3}{4}} \\
 &= 6 \left[\frac{9}{32} - \frac{1}{8} - \frac{9}{64} + \frac{1}{24} \right] \\
 &= \frac{11}{32}.
 \end{aligned}$$

Το αντίστοιχο εμβαδόν εμφανίζεται σκιασμένο στο Σχήμα Α΄.33.

(γ΄) Προφανώς για $x < 0$ έχουμε $F(x) = 0$, και για $x > 1$ έχουμε $F(x) = 1$. Για την περίπτωση $0 \leq x \leq 1$, βρίσκουμε:

$$F(x) = \int_{-\infty}^x f(t) dt = \int_0^x 6t(1-t) dt = 3x^2 - 2x^3.$$

Συνεπώς:

$$F(x) = \begin{cases} 0, & x < 0, \\ 3x^2 - 2x^3, & 0 \leq x \leq 1, \\ 1, & x > 1. \end{cases}$$

Η συνάρτηση κατανομής έχει σχεδιαστεί στο Σχήμα Α΄.33.

7. Υποψήφιες πυκνότητες. Σε καθεμία από τις παρακάτω περιπτώσεις, για να είναι η $f(x)$ πυκνότητα, θα πρέπει να ισχύει ότι $f(x) \geq 0$ για όλα τα x και $\int_{-\infty}^{\infty} f(x) dx = 1$. Επομένως έχουμε:

(α') Για τη συνάρτηση $f(x)$,

$$\int_{-\infty}^{\infty} f(x) dx = \int_{-1}^1 \frac{3}{4}(1-x^2) dx = \left[\frac{3}{4}x - \frac{x^3}{4} \right]_{-1}^1 = \frac{3}{4} - \frac{1}{4} + \frac{3}{4} - \frac{1}{4} = 1.$$

Επιπλέον, προφανώς $f(x) \geq 0$ για κάθε x . Άρα η $f(x)$ είναι πυκνότητα. Το γράφημά της δίνεται στο Σχήμα Α'.34.

Σχετικά με τη μέση τιμή και τη διασπορά, έχουμε:

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} xf(x) dx = \int_{-1}^1 \frac{3}{4}x(1-x^2) dx \\ &= \int_{-1}^1 \frac{3}{4}(x-x^3) dx = \frac{3}{4} \left[\frac{x^2}{2} - \frac{x^4}{4} \right]_{-1}^1 = 0, \\ E(X^2) &= \int_{-1}^1 x^2 f(x) dx = \int_{-1}^1 \frac{3}{4}x^2(1-x^2) dx \\ &= \int_{-1}^1 \frac{3}{4}(x^2-x^4) dx = \frac{3}{4} \left[\frac{x^3}{3} - \frac{x^5}{5} \right]_{-1}^1 \\ &= \frac{3}{4} \times 2 \times \left[\frac{1}{3} - \frac{1}{5} \right] = \frac{1}{5}, \\ \text{Var}(X) &= E(X^2) - (E(X))^2 = \frac{1}{5}. \end{aligned}$$

Το ότι η μέση τιμή είναι 0 είναι αναμενόμενο, λόγω του ότι η πυκνότητα είναι συμμετρική γύρω από το $x = 0$. Και για τη ζητούμενη πιθανότητα,

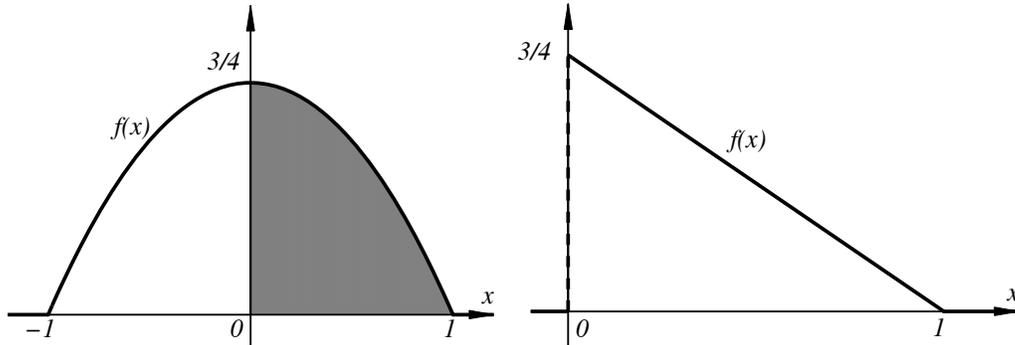
$$\begin{aligned} \Pr(X > 0) &= \int_0^{\infty} f(x) dx = \int_0^1 \frac{3}{4}(1-x^2) dx \\ &= \frac{3}{4} \left[x - \frac{x^3}{3} \right]_0^1 = \frac{3}{4} \left(1 - \frac{1}{3} \right) = \frac{1}{2}, \end{aligned}$$

όπου και αυτό είναι αναμενόμενο λόγω της συμμετρίας της $f(x)$. Βλ. Σχήμα Α'.34.

(β') Για τη συνάρτηση $f(x)$ αυτού του υποερωτήματος,

$$\int_{-\infty}^{\infty} f(x) dx = \int_0^1 \frac{3}{4}(1-x) dx = \frac{3}{8} \left[x - \frac{x^2}{2} \right]_0^1 = \frac{3}{8} \left(1 - \frac{1}{2} \right) = \frac{3}{8}.$$

Άρα η $f(x)$ δεν είναι πυκνότητα. Το γράφημά της φαίνεται στο Σχήμα Α'.34.



Σχήμα Α΄.34: Αριστερά, η πυκνότητα $f(x)$ της Άσκησης 7α΄, και το εμβαδόν που αντιστοιχεί στην πιθανότητα $\Pr(X > E(X))$. Δεξιά, η συνάρτηση $f(x)$ στην Άσκηση 7β΄.

(γ) Εδώ, το ολοκλήρωμα της $f(x)$ είναι,

$$\int_{-\infty}^{\infty} f(x) dx = \int_{-\frac{2}{3}}^{\frac{2}{3}} \frac{3}{4}(1-x) dx = \frac{3}{4} \left[x - \frac{3x^2}{2} \right]_{-\frac{2}{3}}^{\frac{2}{3}} = \frac{3}{4} \left(\frac{2}{3} - \frac{4}{18} + \frac{2}{3} + \frac{4}{18} \right) = 1,$$

και άρα, αφού επιπλέον $f(x) \geq 0$ παντού, η $f(x)$ είναι πυκνότητα. Το γράφημά της δίνεται στο Σχήμα Α΄.35. Σχετικά με τη μέση τιμή και τη διασπορά, βρίσκουμε:

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x f(x) dx = \int_{-\frac{2}{3}}^{\frac{2}{3}} x \frac{3}{4}(1-x) dx = \frac{3}{4} \int_{-\frac{2}{3}}^{\frac{2}{3}} (x - x^2) dx \\ &= \frac{3}{4} \left[\frac{x^2}{2} - \frac{x^3}{3} \right]_{-\frac{2}{3}}^{\frac{2}{3}} = -\frac{4}{27}, \end{aligned}$$

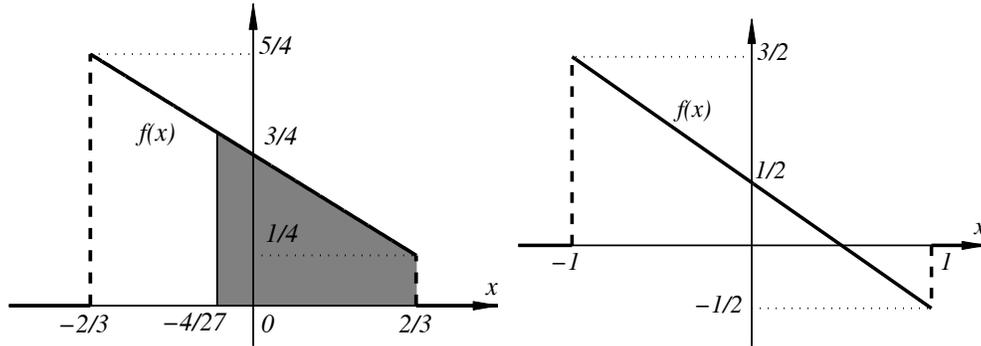
$$\begin{aligned} E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x) dx = \int_{-\frac{2}{3}}^{\frac{2}{3}} x^2 \frac{3}{4}(1-x) dx \\ &= \frac{3}{4} \int_{-\frac{2}{3}}^{\frac{2}{3}} (x^2 - x^3) dx = \frac{3}{4} \left[\frac{x^3}{3} - \frac{x^4}{4} \right]_{-\frac{2}{3}}^{\frac{2}{3}} = \frac{4}{27}, \end{aligned}$$

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{4}{27} - \left(-\frac{4}{27} \right)^2 = \frac{92}{729}.$$

Και για τη ζητούμενη πιθανότητα,

$$\begin{aligned} \Pr\left(X > -\frac{4}{27}\right) &= \int_{-\frac{4}{27}}^1 f(x) dx = \int_{-\frac{4}{27}}^1 \frac{3}{4}(1-x) dx \\ &= \frac{3}{4} \int_{-\frac{4}{27}}^1 \left(x - \frac{x^2}{2} \right)' dx = \frac{3}{4} \left[1 - \frac{1}{2} + \frac{4}{27} + \frac{1}{2} \left(-\frac{4}{27} \right)^2 \right] = \frac{961}{1458}. \end{aligned}$$

Παρατηρούμε ότι η πιθανότητα δεν ισούται με $\frac{1}{2}$. Το αντίστοιχο εμβαδόν εμφανίζεται σκιασμένο στο Σχήμα A'.35.



Σχήμα A'.35: Αριστερά: Η πυκνότητα $f(x)$ της Άσκησης 7γ', και το εμβαδόν που αντιστοιχεί στην πιθανότητα $\Pr(X > E(X))$. Δεξιά: Η συνάρτηση $f(x)$ στην Άσκηση 7δ'.

(δ') Η $f(x)$ εδώ δεν είναι πυκνότητα, διότι παίρνει αρνητικές τιμές όταν το x είναι μεγαλύτερο από $1/2$. Το γράφημά της φαίνεται στο Σχήμα A'.35.

8. Μια απλή πυκνότητα στο \mathbb{R} .

(α) Η τιμή της c θα προσδιοριστεί χρησιμοποιώντας τη βασική ιδιότητα της πυκνότητας που μας λέει πως το ολοκλήρωμά της στο \mathbb{R} πρέπει να ισούται με τη μονάδα. Υπολογίζουμε, λοιπόν,

$$\int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^0 ce^{4x} dx + \int_0^{\infty} ce^{-4x} dx \stackrel{(a)}{=} 2c \int_0^{\infty} e^{-4x} dx = 2c \left[-\frac{e^{-4x}}{4} \right]_0^{\infty} = \frac{c}{2},$$

όπου το βήμα (a) προκύπτει από την αλλαγή μεταβλητής $x \mapsto -x$. Συνεπώς θα πρέπει να έχουμε $c = 2$.

(β') Από τον ορισμό της $E(X)$, εύκολα υπολογίζουμε,

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} xf(x) dx \\ &= \int_{-\infty}^0 2xe^{4x} dx + \int_0^{\infty} 2xe^{-4x} dx \\ &\stackrel{(b)}{=} - \int_0^{\infty} 2xe^{-4x} dx + \int_0^{\infty} 2xe^{-4x} dx = 0, \end{aligned}$$

όπου το βήμα (b) προκύπτει από την αλλαγή μεταβλητής $x \mapsto -x$.

(γ') Παρομοίως με το προηγούμενο ερώτημα, από τον ορισμό της διασποράς βρίσκουμε,

$$\begin{aligned}
 \text{Var}(X) &= E[(X - E(X))^2] \\
 &= E(X^2) \\
 &= \int_{-\infty}^{\infty} x^2 f(x) dx \\
 &= \int_{-\infty}^0 2x^2 e^{4x} dx + \int_0^{\infty} 2x^2 e^{-4x} dx \\
 &\stackrel{(c)}{=} 4 \int_0^{\infty} x^2 e^{-4x} dx \\
 &= - \int_0^{\infty} x^2 (e^{-4x})' dx \\
 &= - [x^2 e^{-4x}]_0^{\infty} + 2 \int_0^{\infty} x e^{-4x} dx \\
 &= - \frac{1}{2} \int_0^{\infty} x (e^{-4x})' dx \\
 &= - \frac{1}{2} [x e^{-4x}]_0^{\infty} + \frac{1}{2} \int_0^{\infty} e^{-4x} dx \\
 &= - \frac{1}{8} \int_0^{\infty} (e^{-4x})' dx \\
 &= \frac{1}{8},
 \end{aligned}$$

όπου η ισότητα (c) προκύπτει με αλλαγή μεταβλητής $x \mapsto -x$, και τα όρια που εμφανίζονται στα παραπάνω καταχρηστικά ολοκληρώματα είναι όλα ίσα με μηδέν, όπως προκύπτει με απλή εφαρμογή του κανόνα του L'Hôpital.

(δ') Τέλος, από τον ορισμό της πυκνότητας, εύκολα βρίσκουμε πως,

$$\begin{aligned}
 \Pr(|X| \geq 1/2) &= \Pr(X \geq 1/2) + \Pr(X \leq -1/2) \\
 &= \int_{\frac{1}{2}}^{\infty} 2e^{-4x} dx + \int_{-\infty}^{-\frac{1}{2}} 2e^{4x} dx \\
 &= 4 \int_{\frac{1}{2}}^{\infty} e^{-4x} dx \\
 &= \int_{\frac{1}{2}}^{\infty} (-e^{-4x})' dx \\
 &= e^{-2},
 \end{aligned}$$

όπου η τρίτη ισότητα και πάλι προκύπτει από την αλλαγή μεταβλητής $x \mapsto -x$.

9. Βαρουφάκης εναντίον Merkel. Έστω Z η συνολική διάρκεια του Eurogroup, οπότε $Z = X$ ή $Z = Y$, με πιθανότητα $1/2$ και για τις δύο περιπτώσεις.

Για τη ζητούμενη πιθανότητα, από τον κανόνα συνολικής πιθανότητας έχουμε,

$$\begin{aligned} \Pr(Z \leq 1/6) &= \Pr(Z \leq 1/6 | Z = X) \Pr(Z = X) + \Pr(Z \leq 1/6 | Z = Y) \Pr(Z = X) \\ &= \frac{1}{2} \Pr(X \leq 1/6) + \frac{1}{2} \Pr(Y \leq 1/6) \\ &= \frac{1}{2} \int_0^{1/6} \frac{1}{3} dx + \frac{1}{2} \int_0^{1/6} \frac{1}{\sqrt{y}} dy \\ &= \frac{1}{2} \cdot \frac{1}{6} \cdot \frac{1}{3} + \frac{1}{2} [2\sqrt{y}]_0^{1/6} \\ &= \frac{1}{36} + \frac{1}{\sqrt{6}} \approx 0.436. \end{aligned}$$

10. Συνεχείς και διακριτές τυχαίες μεταβλητές.

(α') Για την $\Pr(Y_1 \leq 1/2)$ χρησιμοποιούμε τον κανόνα συνολικής πιθανότητας:

$$\begin{aligned} \Pr(Y_1 \leq 1/2) &= \Pr(0 \leq Y_1 \leq 1/2 | X = 1) \Pr(X = 1) \\ &\quad + \Pr(0 \leq Y_1 \leq 1/2 | X = 2) \Pr(X = 2) \\ &\quad + \Pr(0 \leq Y_1 \leq 1/2 | X = 3) \Pr(X = 3) \\ &= \frac{1}{3} \int_0^{1/2} f_1(z) dz + \frac{1}{3} \int_0^{1/2} f_2(z) dz + \frac{1}{3} \int_0^{1/2} f_3(z) dz \\ &= \frac{1}{3} \int_0^{1/2} 1 dz + \frac{1}{3} \int_0^{1/2} \frac{1}{2} dz + \frac{1}{3} \int_0^{1/2} \frac{1}{3} dz \\ &= \frac{1}{6} + \frac{1}{12} + \frac{1}{18} = \frac{11}{36}. \end{aligned}$$

(β') Από τον Ορισμό 10.1, μια συνάρτηση $f(y)$ για να είναι η πυκνότητα του Y_1 πρέπει να ικανοποιεί τη σχέση, $\Pr(a \leq Y_1 \leq b) = \int_a^b f(y) dy$, για οποιοδήποτε διάστημα $[a, b]$. Με τον ίδιο συλλογισμό όπως στο πρώτο ερώτημα, η πιο πάνω πιθανότητα είναι:

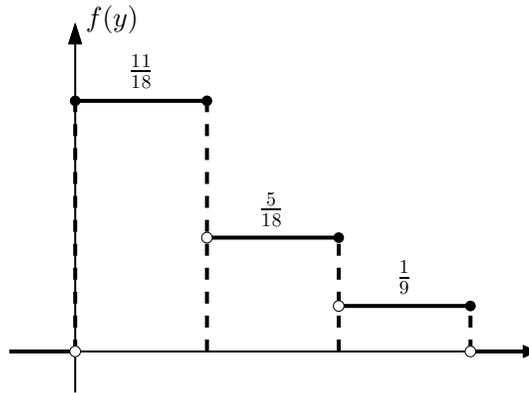
$$\begin{aligned} \Pr(a \leq Y_1 \leq b) &= \Pr(a \leq Y_1 \leq b | X = 1) \Pr(X = 1) \\ &\quad + \Pr(a \leq Y_1 \leq b | X = 2) \Pr(X = 2) \\ &\quad + \Pr(a \leq Y_1 \leq b | X = 3) \Pr(X = 3) \\ &= \frac{1}{3} \int_a^b f_1(z) dz + \frac{1}{3} \int_a^b f_2(z) dz + \frac{1}{3} \int_a^b f_3(z) dz \\ &= \int_a^b \frac{1}{3} [f_1(z) + f_2(z) + f_3(z)] dz. \end{aligned}$$

Συνεπώς, η συνάρτηση,

$$f(y) = \frac{1}{3}[f_1(y) + f_2(y) + f_3(y)] = \begin{cases} 0, & y < 0, \\ 11/18, & 0 \leq y \leq 1, \\ 5/18, & 1 < y \leq 2, \\ 1/9, & 2 < y \leq 3, \\ 0, & y > 3, \end{cases}$$

είναι η πυκνότητα του Y_1 , αφού προφανώς ικανοποιεί και τις δύο άλλες συνθήκες του ορισμού, δηλαδή, $f(y) \geq 0$ για κάθε y , και $\int_{-\infty}^{\infty} f(y) dy = 1$. Βλ. Σχήμα Α΄.36.

Επιπλέον, παρατηρούμε ότι προφανώς και η Y_2 έχει την ίδια πυκνότητα.



Σχήμα Α΄.36: Η πυκνότητα $f(y)$ της Άσκησης 10.

(γ΄) Αφού, δεδομένης της τιμής του X , τα Y_1 και Y_2 είναι ανεξάρτητα, έχουμε,

$$\Pr(Y_1 \leq 1/2, Y_2 \leq 1/2 | X = 1) = \Pr(Y_1 \leq 1/2 | X = 1) \Pr(Y_2 \leq 1/2 | X = 1),$$

όπου οι δύο παραπάνω πιθανότητες είναι και οι δύο ίσες με $\int_0^{1/2} 1 dy = 1/2$, οπότε,

$$\Pr(Y_1 \leq 1/2, Y_2 \leq 1/2 | X = 1) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}.$$

(δ΄) Για την τελευταία πιθανότητα, χρησιμοποιώντας τον κανόνα του Bayes βρίσκουμε ότι η $\Pr(X = 1 | Y_1 \leq 1/2, Y_2 \leq 1/2)$ ισούται με,

$$\frac{\Pr(Y_1 \leq \frac{1}{2}, Y_2 \leq \frac{1}{2} | X = 1) \Pr(X = 1)}{\sum_{i=1}^3 \Pr(Y_1 \leq \frac{1}{2}, Y_2 \leq \frac{1}{2} | X = i) \Pr(X = i)}.$$

Υπολογίζοντας αρχικά τους τρεις όρους του παρονομαστή όπως πριν,

$$\begin{aligned}\Pr\left(Y_1 \leq \frac{1}{2}, Y_2 \leq \frac{1}{2} \mid X = 1\right) \Pr(X = 1) &= \left(\int_0^{1/2} 1 \, dy\right)^2 \frac{1}{3} = \frac{1}{12}, \\ \Pr\left(Y_1 \leq \frac{1}{2}, Y_2 \leq \frac{1}{2} \mid X = 2\right) \Pr(X = 2) &= \left(\int_0^{1/2} \frac{1}{2} \, dy\right)^2 \frac{1}{3} = \frac{1}{48}, \\ \Pr\left(Y_1 \leq \frac{1}{2}, Y_2 \leq \frac{1}{2} \mid X = 3\right) \Pr(X = 3) &= \left(\int_0^{1/2} \frac{1}{3} \, dy\right)^2 \frac{1}{3} = \frac{1}{108},\end{aligned}$$

και αντικαθιστώντας, βρίσκουμε τελικά:

$$\Pr(X = 1 \mid Y_1 \leq 1/2, Y_2 \leq 1/2) = \frac{\frac{1}{12}}{\frac{1}{12} + \frac{1}{48} + \frac{1}{108}} = \frac{36}{49}.$$

11. Απόσταση χ^2 .

(α') Αναπτύσσοντας το τετράγωνο στο ορισμό της $d_{\chi^2}(f, g)$ βρίσκουμε,

$$\begin{aligned}d_{\chi^2}(f, g) &= \int_S \frac{f(x)^2 + g(x)^2 - 2f(x)g(x)}{g(x)} \, dx \\ &= \int_S \frac{f(x)^2}{g(x)} \, dx + \int_S g(x) \, dx - 2 \int_S f(x) \, dx \\ &= \int_S \frac{f(x)^2}{g(x)} \, dx - 1,\end{aligned}$$

όπου χρησιμοποιήσαμε τη βασική ιδιότητα του ορισμού μιας συνεχούς πυκνότητας που μας λέει πως το ολοκλήρωμά της σε ολόκληρο το \mathbb{R} (ή, ισοδύναμα, σε ολόκληρο το σύνολο τιμών της) ισούται με 1.

(β') Παρομοίως, για δύο διακριτές πυκνότητες $P(x)$ και $Q(x)$ στο ίδιο σύνολο τιμών S , έχουμε,

$$\begin{aligned}d_{\chi^2}(P, Q) &= \sum_{x \in S} \frac{(P(x) - Q(x))^2}{Q(x)} \\ &= \sum_{x \in S} \frac{P(x)^2 + Q(x)^2 - 2P(x)Q(x)}{Q(x)} \\ &= \sum_{x \in S} \frac{P(x)^2}{Q(x)} + \sum_{x \in S} Q(x) - 2 \sum_{x \in S} P(x),\end{aligned}$$

άρα,

$$d_{\chi^2}(P, Q) = \sum_{x \in S} \frac{P(x)^2}{Q(x)} - 1,$$

όπου παραπάνω χρησιμοποιήσαμε την αντίστοιχη ιδιότητα μιας διακριτής πυκνότητας, δηλαδή το ότι το άθροισμα όλων των τιμών της ισούται με 1.

Α'.11 Ασκήσεις Κεφαλαίου 11

1. Από την ομοιόμορφη στην εκθετική κατανομή. Θα υπολογίσουμε τη συνάρτηση κατανομής $F(y)$ της Y . Κατ' αρχάς, παρατηρούμε ότι, εξ ορισμού, η U έχει σύνολο τιμών το $[0, \infty)$, άρα $F(y) = 0$ για $y \leq 0$. Για $y > 0$,

$$\begin{aligned} F(y) = \Pr(Y \leq y) &= \Pr(-\lambda \log U \leq y) \\ &= \Pr(\log U \geq -y/\lambda) \\ &= \Pr(U \geq e^{-y/\lambda}) = \int_{e^{-y/\lambda}}^1 1 dx = 1 - e^{-y/\lambda}. \end{aligned}$$

Συνεπώς, η U έχει συνάρτηση κατανομής εκείνη της $\text{Eκθ}(\lambda)$ κατανομής, βλ. Θεώρημα 11.1, άρα έχουμε $Y \sim \text{Eκθ}(\lambda)$.

2. Χρόνος μετάδοσης.

(α) Η μέση τιμή του χρόνου T είναι:

$$\begin{aligned} E(T) = E\left(\frac{X(X+1)}{4}\right) &= \frac{1}{4}E(X^2 + X) \\ &\stackrel{(a)}{=} \frac{1}{4}[E(X^2) + E(X)] \\ &\stackrel{(b)}{=} \frac{1}{4}[\text{Var}(X) + (E(X))^2 + E(X)] \\ &\stackrel{(c)}{=} \frac{1}{4}[5^2 + 5^2 + 5] = 13.75, \end{aligned}$$

όπου στο βήμα (a) χρησιμοποιήσαμε την Ιδιότητα 1 του Θεωρήματος 11.5, στο βήμα (b) εφαρμόσαμε τον εναλλακτικό τύπο για τη διασπορά (10.13), και στο βήμα (c) αντικαταστήσαμε τις τιμές για τη μέση τιμή και τη διασπορά μιας εκθετικής Τ.Μ. από το Θεώρημα 11.1.

(β') Η πιθανότητα ο χρόνος εκτέλεσης να ξεπεράσει τα 10 λεπτά, δηλαδή $T > 10$ είναι,

$$\begin{aligned} \Pr(T > 10) &= \Pr\left(\frac{X(X+1)}{4} > 10\right) \\ &= \Pr(X(X+1) > 40) \\ &= \Pr(X^2 + X - 40 > 0). \end{aligned}$$

Αλλά το τριώνυμο $x^2 + x - 40$ παραγοντοποιείται ως $(x - \frac{-1-\sqrt{161}}{2})(x - \frac{-1+\sqrt{161}}{2})$, και εφόσον η Τ.Μ. X παίρνει μόνο θετικές τιμές, το πιο πάνω τριώνυμο είναι θετικό

αν και μόνο αν το x είναι μεγαλύτερο από τη θετική του ρίζα. Συνεπώς,

$$\begin{aligned} \Pr(T > 10) &= \Pr\left(X > \frac{-1 + \sqrt{161}}{2}\right) \\ &= 1 - \Pr\left(X \leq \frac{-1 + \sqrt{161}}{2}\right) \\ &\stackrel{(d)}{=} e^{-\frac{-1 + \sqrt{161}}{10}} \approx 0.3107, \end{aligned}$$

όπου στο βήμα (d) αντικαταστήσαμε τον τύπο της συνάρτησης κατανομής μιας εκθετικής Τ.Μ. από το Θεώρημα 11.1.

3. *Μηνυματάκια!* Έστω οι ακόλουθες τυχαίες μεταβλητές:

$$\begin{aligned} X_1 &= \text{η διάρκεια αποστολής του πρώτου SMS,} \\ X_2 &= \text{η διάρκεια αποστολής του δεύτερου SMS,} \\ Y &= \text{η διάρκεια αποστολής του MMS,} \\ Z &= \text{η συνολική διάρκεια αποστολής των δύο SMS και του MMS,} \end{aligned}$$

οπότε, $Z = X_1 + X_2 + Y$.

(α') Για τη μέση τιμή της συνολικής διάρκειας αποστολής, από την πρώτη ιδιότητα του Θεωρήματος 11.5,

$$E(Z) = E(X_1 + X_2 + Y) = E(X_1) + E(X_2) + E(Y) = \frac{3+1}{2} + \frac{3+1}{2} + 8 = 12,$$

αφού οι X_1, X_2 έχουν ομοιόμορφη κατανομή στο διάστημα $[1, 3]$ και η Y έχει εκθετική κατανομή με μέση τιμή 8.

(β') Εφόσον οι τρεις αποστολές είναι ανεξάρτητες μεταξύ τους, η πιθανότητα και τα 2 SMS να έχουν διάρκεια πάνω από 2 δευτερόλεπτα το καθένα είναι:

$$\Pr(X_1 > 2, X_2 > 2) = \Pr(X_1 > 2) \Pr(X_2 > 2) = \left[\int_2^3 \frac{1}{3-1} dx \right]^2 = \left(\frac{1}{2} \right)^2 = \frac{1}{4}.$$

(γ') Η πιθανότητα το MMS να έχει διάρκεια μεγαλύτερη από τη μέση τιμή της συνολικής διάρκειας των 2 SMS είναι,

$$\Pr(Y > E(X_1 + X_2)) = \Pr(Y > 2 + 2) = 1 - \Pr(Y \leq 4) = 1 - [1 - e^{-4/8}] \approx 0.6065,$$

όπου χρησιμοποιήσαμε τον τύπο της συνάρτησης κατανομής μιας εκθετικής Τ.Μ. από το Θεώρημα 11.1.

- (δ') Από τον ορισμό της δεσμευμένης πιθανότητας και τον τύπο της συνάρτησης κατανομής $F(y)$ της εκθετικής T.M. Y (βλ. Θεώρημα 11.1), η ζητούμενη δεσμευμένη πιθανότητα ισούται με:

$$\begin{aligned} \Pr(10 < Y < 20 | Y > 10) &= \frac{\Pr(10 < Y < 20, Y > 10)}{\Pr(Y > 10)} \\ &= \frac{\Pr(10 < Y < 20)}{\Pr(Y > 10)} \\ &= \frac{F(20) - F(10)}{1 - F_Y(10)} \\ &= \frac{1 - e^{-20/8} - 1 + e^{-10/8}}{1 - 1 + e^{-10/8}} \approx 0.7135. \end{aligned}$$

4. Ελάχιστο και μέγιστο δύο T.M.

- (α') Για τη συνάρτηση κατανομής της W , χρησιμοποιώντας το γεγονός ότι οι X, Y είναι ανεξάρτητες και παρατηρώντας πως το μέγιστο μεταξύ δύο αριθμών είναι μικρό αν και μόνο αν και οι δύο αριθμοί είναι μικροί, έχουμε,

$$\begin{aligned} F_W(w) &= \Pr(W \leq w) \\ &= \Pr(\{X \leq w\} \cap \{Y \leq w\}) \\ &= \Pr(X \leq w) \Pr(Y \leq w) \\ &= F_X(w)F_Y(w). \end{aligned}$$

Παρομοίως, για να βρούμε τη συνάρτηση κατανομής της V , παρατηρώντας πως το ελάχιστο μεταξύ δύο αριθμών είναι μεγάλο αν και μόνο αν και οι δύο αριθμοί είναι μεγάλοι, υπολογίζουμε,

$$\begin{aligned} F_V(v) &= \Pr(V \leq v) \\ &= 1 - \Pr(V > v) = 1 - \Pr(\{X > v\} \cap \{Y > v\}) \\ &= 1 - \Pr(X > v) \Pr(Y > v) \\ &= 1 - (1 - F_X(v))(1 - F_Y(v)). \end{aligned}$$

- (β') Εφόσον οι X και Y έχουν σύνολο τιμών το $S = [0, \infty)$, και η V θα έχει το ίδιο σύνολο τιμών, συνεπώς $F_V(v) = 0$ για $v < 0$. Και για $v \geq 0$, χρησιμοποιώντας το αποτέλεσμα του προηγούμενου ερωτήματος και τον τύπο για την εκθετική συνάρτηση κατανομής από το Θεώρημα 11.1,

$$F_V(v) = 1 - (1 - F_X(v))(1 - F_Y(v)) = 1 - e^{-v/\theta} e^{-v/\phi} = 1 - \exp\{-v(1/\theta + 1/\phi)\},$$

άρα και η V έχει εκθετική κατανομή, με παράμετρο, $(1/\theta + 1/\phi)^{-1}$.

Παρατηρούμε πως αυτό το αποτέλεσμα είναι ακριβώς ανάλογο του αντίστοιχου αποτελέσματος για τη γεωμετρική κατανομή στην Άσκηση 4 του Κεφαλαίου 7.

5. Ελάχιστο N Τ.Μ.

(α') Όπως και στην προηγούμενη άσκηση, παρατηρούμε πως το ελάχιστο μεταξύ N αριθμών είναι μεγάλο αν και μόνο αν και οι N αριθμοί είναι μεγάλοι. Άρα, χρησιμοποιώντας την ανεξαρτησία των X_i ,

$$\begin{aligned} F_V(v) &= \Pr(V \leq v) = 1 - \Pr(V > v) \\ &= 1 - \Pr(\{X_1 > v\} \cap \{X_2 > v\} \cap \dots \cap \{X_N > v\}) \\ &= 1 - \Pr(X_1 > v) \Pr(X_2 > v) \dots \Pr(X_N > v) \\ &= 1 - (1 - \Pr(X_1 \leq v))(1 - \Pr(X_2 \leq v)) \dots (1 - \Pr(X_N \leq v)) \\ &= 1 - (1 - F_X(v))^N. \end{aligned}$$

(β') Στην ειδική περίπτωση που η $F_X(x)$ είναι εκθετική με παράμετρο θ , δηλαδή

$$F_X(x) = \begin{cases} 1 - e^{-\frac{x}{\theta}}, & x \geq 0, \\ 0, & x < 0, \end{cases}$$

τότε από το προηγούμενο ερώτημα προκύπτει πως,

$$F_V(v) = \begin{cases} 1 - e^{-\frac{vN}{\theta}}, & v \geq 0, \\ 0, & v < 0, \end{cases}$$

δηλαδή η V είναι επίσης εκθετικά κατανομημένη, με παράμετρο θ/N .

6. Γραμμικός μετασχηματισμός. Επαναλαμβάνουμε την ίδια απόδειξη όπως για την Ιδιότητα 3 του Θεωρήματος 11.2, αυτήν τη φορά όμως υποθέτοντας πως η σταθερά a είναι αρνητική. Εξετάζουμε αρχικά τις συναρτήσεις κατανομής, έστω $F(x)$ και $G(y)$ των τυχαίων μεταβλητών X, Y αντίστοιχα. Για την $G(y)$ έχουμε,

$$G(y) = \Pr(Y \leq y) = \Pr(aX + b \leq y) = \Pr\left(X \geq \frac{y-b}{a}\right) = 1 - F\left(\frac{y-b}{a}\right),$$

όπου η ανισότητα αντιστρέφεται όταν πολλαπλασιάζουμε με την αρνητική ποσότητα $1/a$. Και τώρα υπολογίζουμε την πυκνότητα της Y ως την παράγωγο της συνάρτησης κατανομής της, από τη σχέση (10.7),

$$g(y) = G'(y) = \frac{d}{dy} \left[1 - F\left(\frac{y-b}{a}\right) \right] = -\frac{1}{a} F'\left(\frac{y-b}{a}\right),$$

άρα, όταν η a είναι αρνητική, η Y έχει πυκνότητα,

$$g(y) = -\frac{1}{a} f\left(\frac{y-b}{a}\right).$$

Παρατηρήστε πως μπορούμε να εκφράσουμε και τις δύο περιπτώσεις ($a < 0$ και $a > 0$) με ενιαίο τρόπο ως,

$$g(y) = \frac{1}{|a|} f\left(\frac{y-b}{a}\right).$$

7. *Κανονικοποίηση.* (α') Στην περίπτωση που η X είναι συνεχής, χρησιμοποιώντας την πρώτη ιδιότητα του Θεωρήματος 11.2,

$$E\left(\frac{X-\mu}{\sigma}\right) = E\left(\frac{1}{\sigma}X - \frac{\mu}{\sigma}\right) = \frac{1}{\sigma}E(X) + \frac{\mu}{\sigma} = 0,$$

και παρομοίως, εφαρμόζοντας τη δεύτερη ιδιότητα του Θεωρήματος 11.2,

$$\text{Var}\left(\frac{X-\mu}{\sigma}\right) = \text{Var}\left(\frac{1}{\sigma}X - \frac{\mu}{\sigma}\right) = \frac{1}{\sigma^2}\text{Var}(X) = 1.$$

(β') Η απόδειξη στη διακριτή περίπτωση είναι ακριβώς η ίδια, με τη μόνη διαφορά ότι αντί για την πρώτη και τη δεύτερη ιδιότητα του Θεωρήματος 11.2, επικαλούμαστε τις δύο πρώτες ιδιότητες του Θεωρήματος 6.1 αντίστοιχα.

8. *Το φράγμα του Chernoff για συνεχείς Τ.Μ.*

- (α') Η λύση είναι ακριβώς ίδια όπως για το αντίστοιχο ερώτημα της Άσκησης 7 του Κεφαλαίου 9.
- (β') Έστω τώρα πως $X \sim \text{Exp}(\theta)$. Από τον γενικό ορισμό της μέσης τιμής μιας συνάρτησης Τ.Μ., μαζί με τον ορισμό της πυκνότητας της εκθετικής κατανομής,

$$\begin{aligned} E(e^{\lambda X}) &= \int_{-\infty}^{\infty} e^{\lambda x} f(x) dx \\ &= \int_0^{\infty} e^{\lambda x} \frac{1}{\theta} e^{-x/\theta} dx \\ &= \frac{1}{\theta} \int_0^{\infty} e^{-x(1/\theta-\lambda)} dx \\ &= \frac{1}{\theta(1/\theta-\lambda)} \left[-e^{-x(1/\theta-\lambda)} \right]_0^{\infty} \\ &= \frac{1}{1-\lambda\theta}, \end{aligned}$$

όπου παρατηρούμε πως το παραπάνω καταχρηστικό ολοκλήρωμα υπάρχει, επειδή το όριο,

$$\lim_{x \rightarrow \infty} e^{-x(1/\theta-\lambda)},$$

υπάρχει και ισούται με μηδέν, κι αυτό λόγω του περιορισμού που έχουμε θέσει ότι $\lambda < 1/\theta$.

Αντικαθιστώντας αυτό το αποτέλεσμα στη σχέση 11.5, έχουμε,

$$\Pr(X \geq c) \leq \frac{e^{-\lambda c}}{1 - \lambda \theta}.$$

(γ') Έστω τώρα πως $X \sim \text{Exp}(1)$.

i. Για $c = 10$, το φράγμα (11.6), που μόλις αποδείξαμε, μας δίνει,

$$\Pr(X \geq 10) \leq \frac{e^{-10\lambda}}{1 - \lambda}.$$

Για να βρούμε το καλύτερο δυνατό φράγμα αρκεί να υπολογίσουμε την τιμή του λ που ελαχιστοποιεί το δεξί μέρος της παραπάνω σχέσης. Για να βρούμε αυτή την τιμή θα ελαχιστοποιήσουμε τη συνάρτηση $f(\lambda) = e^{-\lambda c}/(1 - \lambda)$ ως προς λ . Παίρνοντας παραγώγους, εύκολα βρίσκουμε πως,

$$f'(\lambda) = \frac{e^{-\lambda c} [1 - c(1 - \lambda)]}{(1 - \lambda)^2},$$

το οποίο ισούται με μηδέν αν και μόνο αν $1 - c(1 - \lambda) = 0$, δηλαδή για την τιμή,

$$\lambda^* = 1 - \frac{1}{c} = 0.9.$$

Επιπλέον, η δεύτερη παράγωγος ισούται με,

$$f''(\lambda) = \frac{e^{-\lambda c}}{(1 - \lambda)} \left[\left(c - \frac{1}{1 - \lambda} \right)^2 + \frac{1}{(1 - \lambda)^2} \right],$$

το οποίο είναι προφανώς πάντα θετικό για $c \geq 0$ και $\lambda \in (0, 1)$, άρα η τιμή λ^* πράγματι ελαχιστοποιεί την $f(\lambda)$. Αντικαθιστώντας αυτή την τιμή στο φράγμα μας,

$$\Pr(X \geq 10) \leq \frac{e^{-10\lambda^*}}{1 - \lambda^*} = 10e^{-9} \approx 0.00123.$$

ii. Η Τ.Μ. X έχει μέση τιμή $E(X) = 1$, όποτε η ανισότητα του Markov μάς δίνει,

$$\Pr(X \geq 10) \leq \frac{E(X)}{10} = \frac{1}{10} = 0.1,$$

το οποίο είναι προφανώς εξαιρετικά ασθενέστερο.

iii. Τέλος, η ζητούμενη πιθανότητα $\Pr(X \geq 10)$ μπορεί εύκολα να υπολογιστεί ακριβώς. Εφόσον το $X \sim \text{Exp}(1)$, γνωρίζουμε από το Θεώρημα 11.1 πως,

$$\Pr(X \geq 10) = 1 - \Pr(X \leq 10) = 1 - F(10) = e^{-10} \approx 0.0000454.$$

Προφανώς, λοιπόν, και πάλι το φράγμα του Chernoff μάς δίνει ένα αποτέλεσμα πολύ πιο ισχυρό και πιο κοντά στην πραγματική τιμή από ό,τι η ανισότητα του Markov.

9. Γραμμικοί συνδυασμοί. Κατ' αρχάς έχουμε,

$$E(X) = E(Y) = 2 \quad \text{και} \quad \text{Var}(X) = \text{Var}(Y) = 2^2 = 4.$$

Επίσης θυμίζουμε πως, από τη σχέση (10.13), για οποιαδήποτε Τ.Μ. X έχουμε,

$$E(X^2) = \text{Var}(X) + (E(X))^2.$$

Χρησιμοποιώντας αυτήν τη σχέση, και τις δύο ιδιότητες του Θεωρήματος 11.5, βρίσκουμε,

$$\begin{aligned} E(A) &= E(2X + Y) = 2E(X) + E(Y) = 6, \\ E(B) &= E(2X - Y) = 2E(X) - E(Y) = 2, \\ \text{Var}(A) &= 2^2\text{Var}(X) + \text{Var}(Y) = 20, \\ \text{Var}(B) &= 2^2\text{Var}(X) + (-1)^2\text{Var}(Y) = 20, \\ E(AB) &= E[(2X + Y)(2X - Y)] \\ &= E(4X^2 - Y^2) = 3E(X^2) = 3\text{Var}(X) + 3(E(X))^2 = 24. \end{aligned}$$

10. Συνάρτηση μιας εκθετικής Τ.Μ. Εφόσον η X έχει σύνολο τιμών το $[0, \infty)$, η Y προφανώς έχει σύνολο τιμών το $[5, \infty)$. Αν, τώρα, $F(x)$ και $f(x)$ είναι η συνάρτηση κατανομής και η πυκνότητα του X αντίστοιχα, και $G(y)$, $g(y)$ είναι η συνάρτηση κατανομής και η πυκνότητα του Y αντίστοιχα, τότε για $y < 5$ έχουμε $G(y) = 0$, ενώ για $y \geq 5$,

$$G(y) = \Pr(Y \leq y) = \Pr(X^3 + 5 \leq y) = \Pr(X^3 \leq y - 5) = \Pr(X \leq (y - 5)^{1/3}),$$

οπότε,

$$G(y) = \begin{cases} F((y - 5)^{1/3}), & y \geq 5, \\ 0, & y < 5. \end{cases}$$

Άρα η πυκνότητα της Y θα είναι,

$$\begin{aligned} g(y) &= G'(y) \\ &= \begin{cases} F'((y - 5)^{1/3})((y - 5)^{1/3})', & y \geq 5, \\ 0, & y < 5, \end{cases} \\ &= \begin{cases} f((y - 5)^{1/3}) \frac{1}{3}(y - 5)^{-2/3}, & y \geq 5, \\ 0, & y < 5, \end{cases} \\ &= \begin{cases} \frac{1}{60}(y - 5)^{-\frac{2}{3}} e^{-\frac{(y-5)^{1/3}}{20}}, & y \geq 5, \\ 0, & y < 5. \end{cases} \end{aligned}$$

11. *Ανισότητα Markov.* Ξεκινώντας από τον ορισμό της μέσης τιμής,

$$\mu = \int_0^{\infty} x f(x) dx = \int_0^c x f(x) dx + \int_c^{\infty} x f(x) dx.$$

Εφόσον όλες οι τιμές x στο πρώτο ολοκλήρωμα είναι μεγαλύτερες ή ίσες του μηδενός, το ολοκλήρωμα είναι κι αυτό μεγαλύτερο ή ίσο του μηδενός, συνεπώς,

$$\mu \geq \int_c^{\infty} x f(x) dx \geq \int_c^{\infty} c f(x) dx = c \int_c^{\infty} f(x) dx,$$

όπου χρησιμοποιήσαμε το γεγονός ότι στο πρώτο από τα παραπάνω ολοκληρώματα όλα τα x εκεί είναι μεγαλύτερα ή ίσα του c . Παρατηρώντας, τέλος, πως από τον ορισμό της πυκνότητας το τελευταίο από τα παραπάνω ολοκληρώματα ισούται με $\Pr(X \geq c)$, έχουμε,

$$\mu \geq c \Pr(X \geq c),$$

που είναι η ζητούμενη ανισότητα.

12. *Ανισότητα Chebychev.* Η απόδειξη της ανισότητας Chebychev στη συνεχή περίπτωση είναι ακριβώς η ίδια με την απόδειξη του αντίστοιχου αποτελέσματος για μια διακριτή Τ.Μ. στο Θεώρημα 9.2, με τη μόνη διαφορά ότι, αντί για τη διακριτή ανισότητα Markov στο Θεώρημα 9.1, επικαλούμαστε το ανάλογο συνεχές αποτέλεσμα, δηλαδή την ανισότητα Markov του Θεωρήματος 11.3.
13. *Πόσο ακριβής είναι η ανισότητα Chebychev;* Αν η X είναι ομοιόμορφα κατανομημένη στο $[-k, k]$, τότε $\mu = 0$ και $\sigma^2 = (2k)^2/12 = k^2/3$. Άρα η ανισότητα Chebychev μας δίνει,

$$\Pr(|X - \mu| \geq c) \leq \frac{\sigma^2}{c^2} = \frac{k^2}{3c^2},$$

ενώ στην πραγματικότητα η παραπάνω πιθανότητα είναι,

$$\Pr(|X - \mu| \geq c) = \Pr(|X| \geq c) = \begin{cases} 1 - \frac{c}{k}, & c \leq k, \\ 0, & c > k, \end{cases}$$

όπως προκύπτει απλά ολοκληρώνοντας την πυκνότητα της X . Παρατηρούμε ότι, καθώς αυξάνει το c , το φράγμα τείνει στο 0 σχετικά αργά, ενώ στην πραγματικότητα η πιθανότητα γίνεται ακριβώς 0 για $c \geq k$.

14. *Απόδειξη του N.M.A.* Στη συνεχή περίπτωση η απόδειξη του N.M.A. είναι ακριβώς η ίδια με την απόδειξη του αντίστοιχου αποτελέσματος για διακριτές Τ.Μ. στο Θεώρημα 9.3, με τις ακόλουθες διαφορές:

- Αντί για την Ιδιότητα 1 του Θεωρήματος 6.1, εδώ επικαλούμαστε τη σχέση (11.3) που γενικεύει την αντίστοιχη Ιδιότητα 1 του Θεωρήματος 11.5.
 - Αντί για την Ιδιότητα 2 του Θεωρήματος 6.1, εδώ επικαλούμαστε την αντίστοιχη Ιδιότητα 2 του Θεωρήματος 11.2.
 - Αντί για την Ιδιότητα 4 του Θεωρήματος 6.1, εδώ επικαλούμαστε τη σχέση (11.4) που γενικεύει την αντίστοιχη Ιδιότητα 2 του Θεωρήματος 11.5.
 - Αντί για τη διακριτή ανισότητα Chebyshev στο Θεώρημα 9.2, εδώ επικαλούμαστε το ανάλογο συνεχές αποτέλεσμα, δηλαδή την ανισότητα Chebyshev του Θεωρήματος 11.4.
15. *Συνεχείς και διακριτές T.M.* Θυμίζουμε πως στην Άσκηση 10 του Κεφαλαίου 10 δείξαμε ότι $\Pr(Y_1 \leq 1/2) = \Pr(Y_2 \leq 1/2) = \frac{11}{36}$, και επίσης υπολογίσαμε πως,

$$\begin{aligned} \Pr\left(Y_1 \leq \frac{1}{2}, Y_2 \leq \frac{1}{2}\right) &= \sum_{i=1}^3 \Pr\left(Y_1 \leq \frac{1}{2}, Y_2 \leq \frac{1}{2} \mid X = i\right) \Pr(X = i) \\ &= \frac{1}{12} + \frac{1}{48} + \frac{1}{108} = \frac{49}{432}. \end{aligned}$$

Εφόσον $\frac{49}{432} \neq \frac{11}{36} \times \frac{11}{36}$, έχουμε ότι,

$$\Pr\left(Y_1 \leq \frac{1}{2}, Y_2 \leq \frac{1}{2}\right) \neq \Pr(Y_1 \leq 1/2) \times \Pr(Y_2 \leq 1/2),$$

άρα οι Y_1 και Y_2 δεν είναι ανεξάρτητες.

Ο λόγος είναι απλός: Δεδομένης της τιμής της Y , όταν δηλαδή γνωρίζουμε την κατανομή των Y_1, Y_2 , τότε είναι πράγματι ανεξάρτητες. Αλλά όταν δεν ξέρουμε το X , τότε βλέποντας την τιμή της Y_1 αποκτάμε κάποια ένδειξη για το X και άρα και για τις πιθανές τιμές της Y_2 . Π.χ., αν η $Y_1 = 2.8$, ξέρουμε σίγουρα ότι το $X = 3$ και ξέρουμε με βεβαιότητα την κατανομή της Y_2 . Ενώ αν η $Y_1 = 0.4$, έχουμε πολύ μεγαλύτερη αβεβαιότητα για την κατανομή της Y_2 .

16. Άλλος ένας γραμμικός μετασχηματισμός. Έστω $F_X(x)$ και $F_Y(y)$ οι συναρτήσεις κατανομής των X και Y αντίστοιχα. Παρατηρούμε πως, από τον ορισμό της Y ,

$$F_Y(y) = \Pr(Y \leq y) = \Pr(10X + 5 \leq y) = \Pr\left(X \leq \frac{y-5}{10}\right) = F_X\left(\frac{y-5}{10}\right).$$

Όμως, η $F_X(x)$ είναι η:

$$F_X(x) = \begin{cases} 0, & x < 0, \\ x, & 0 \leq x \leq 1, \\ 1, & x > 1. \end{cases}$$

Ακολουθώντας, παίρνουμε περιπτώσεις:

Αν το $y < 5$, τότε $(y - 5)/10 < 0$ και άρα $F_Y(y) = F_X((y - 5)/10) = 0$. Αν το $y > 15$, τότε $(y - 5)/10 > 1$ οπότε $F_Y(y) = F_X((y - 5)/10) = 1$. Τέλος, αν $5 \leq y \leq 15$, τότε $0 \leq (y - 5)/10 \leq 1$ και επομένως $F_Y(y) = F_X((y - 5)/10) = (y - 5)/10$. Συνεπώς,

$$F_Y(y) = \begin{cases} 0, & y < 0, \\ \frac{y-5}{10}, & 5 \leq y \leq 15, \\ 1, & y > 15, \end{cases}$$

δηλαδή η Y έχει ομοιόμορφη κατανομή στο διάστημα $[5, 15]$.

17. Συνδιακύμανση και συσχέτιση.

(α') Αν εφαρμόσουμε την ανισότητα Cauchy-Schwarz της Πρότασης 11.1 στις τυχαίες μεταβλητές $X' = X - E(X)$ και $Y' = Y - E(Y)$, βρίσκουμε πως,

$$(E[(X - E(X))(Y - E(Y))])^2 \leq E[(X - E(X))^2]E[(Y - E(Y))^2],$$

ή, ισοδύναμα, αντικαθιστώντας τους ορισμούς της συνδιακύμανσης και της διασποράς,

$$[\text{Cov}(X, Y)]^2 \leq \text{Var}(X)\text{Var}(Y), \quad (\text{A'.16})$$

το οποίο προφανώς συνεπάγεται και το ζητούμενο αποτέλεσμα.

(β') Η ανισότητα $|\rho_{X,Y}| \leq 1$, δηλαδή η,

$$\rho_{X,Y}^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X)\text{Var}(Y)} \leq 1,$$

είναι άμεση συνέπεια της (A'.16) πιο πάνω.

Επιπλέον, αν για κάποια T.M. X με μη μηδενική διασπορά θέσουμε $Y = X$, τότε $\text{Var}(Y) = \text{Var}(X)$ και,

$$\text{Cov}(X, Y) = E[(X - E(X))(Y - E(Y))] = E[(X - E(X))^2] = \text{Var}(X),$$

οπότε έχουμε,

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = 1.$$

Παρομοίως, αν θέσουμε $Y = -X$, θα έχουμε $\rho_{X,Y} = -1$.

Α'.12 Ασκήσεις Κεφαλαίου 12

1. *Τυχαία επιλογή Τ.Μ.* Ορίζουμε την Τ.Μ. B η οποία παίρνει την τιμή 1 αν φέρουμε κορώνα και την τιμή 0 αν φέρουμε Γράμματα. Επιπλέον, έστω T η τιμή που καταγράψαμε. Χρησιμοποιώντας τον κανόνα του Bayes, ή επαναλαμβάνοντας τα ίδια βήματα, η ζητούμενη πιθανότητα ισούται με:

$$\begin{aligned}
 \Pr(B = 1 | T > 4) &= \frac{\Pr(B = 1, T > 4)}{\Pr(T > 4)} \\
 &= \frac{\Pr(B = 1, X > 4)}{\Pr(T > 4 | B = 1) \Pr(B = 1) + \Pr(T > 4 | B = 0) \Pr(B = 0)} \\
 \stackrel{(a)}{=} &= \frac{\Pr(X > 4) \Pr(B = 1)}{\Pr(X > 4) \Pr(B = 1) + \Pr(Y > 4) \Pr(B = 0)} \\
 &= \frac{\frac{1}{2} \int_4^5 \frac{1}{8} dx}{\frac{1}{2} \int_4^5 \frac{1}{8} dx + \frac{1}{2} \Pr\left(\frac{Y-1}{\sqrt{10}} > \frac{4-1}{\sqrt{10}}\right)} \\
 &= \frac{\frac{1}{8}}{\frac{1}{8} + (1 - \Phi(\frac{3}{\sqrt{10}}))} \\
 &\approx 0.422,
 \end{aligned}$$

όπου στο βήμα (a) χρησιμοποιήσαμε την ανεξαρτησία των Τ.Μ. B, X , στο τέλος εφαρμόσαμε τη γενική μέθοδο υπολογισμού πιθανοτήτων για την κανονική κατανομή που είδαμε στην Ενότητα 12.1, και αντικαταστήσαμε την τιμή $\Phi(\frac{3}{\sqrt{10}}) \approx \Phi(0.95) \approx 0.8289$ από τον Πίνακα 12.2.

2. *Ύψη μαθητών.* Ορίζουμε τις Τ.Μ.,

$$\begin{aligned}
 X &= \text{ύψος ενός τυχαία επιλεγμένου αγοριού,} \\
 Y &= \text{ύψος ενός τυχαία επιλεγμένου κοριτσιού,}
 \end{aligned}$$

έτσι ώστε $X \sim N(1.7, 0.01)$ και $Y \sim N(1.6, 0.01)$. Επίσης ορίζουμε τα σχετικά ενδεχόμενα $A = \text{«Ο μαθητής είναι αγόρι»}$ και $B = \text{«ο μαθητής είναι πάνω από 1.80 m»}$. Τότε, από τον κανόνα συνολικής πιθανότητας,

$$\begin{aligned}
 \Pr(B) &= \Pr(B|A) \Pr(A) + \Pr(B|A') \Pr(A') \\
 &= \Pr(X > 1.8) \cdot \frac{1}{2} + \Pr(Y > 1.8) \cdot \frac{1}{2} \\
 &= \frac{1}{2}[1 - \Pr(X \leq 1.8)] + \frac{1}{2}[1 - \Pr(Y \leq 1.8)],
 \end{aligned}$$

και, χρησιμοποιώντας τη γενική μέθοδο υπολογισμού πιθανοτήτων για την κανονική κατανομή που είδαμε στην Ενότητα 12.1, έχουμε,

$$\begin{aligned}\Pr(B) &= \frac{1}{2} \left[1 - \Pr \left(Z \leq \frac{1.8 - 1.7}{0.1} \right) \right] + \frac{1}{2} \left[1 - \Pr \left(Z < \frac{1.8 - 1.6}{0.1} \right) \right] \\ &= 1 - \frac{1}{2} \Phi(1) - \frac{1}{2} \Phi(2) \approx 0.09075,\end{aligned}$$

όπου η Τ.Μ. $Z \sim N(0, 1)$ και αντικαταστήσαμε τις σχετικές τιμές της $\Phi(z)$ από τον Πίνακα 12.2.

3. *Αριθμητικοί υπολογισμοί με την κανονική κατανομή.* Σε πολλά από τα παρακάτω ερωτήματα θα χρησιμοποιήσουμε τη γενική μέθοδο υπολογισμού πιθανοτήτων για την κανονική κατανομή που είδαμε στην Ενότητα 12.1. Έστω $Z \sim N(0, 1)$.

(α') Αντικαθιστώντας τις σχετικές τιμές της $\Phi(z)$ από τους Πίνακες 12.1 και 12.2,

$$\begin{aligned}\Pr(X < +1.22) &= \Pr(Z < 1.22) = \Phi(1.22) \approx 0.8888, \\ \Pr(X > -1.22) &= \Pr(Z > -1.22) = 1 - \Pr(Z \leq -1.22) = 1 - \Phi(-1.22) \\ &\approx 0.8888.\end{aligned}$$

(β') Παρομοίως έχουμε,

$$\Pr(X > 2.7) = \Pr \left(Z > \frac{2.7 - 1}{1} \right) = 1 - \Phi \left(\frac{2.7 - 1}{1} \right) = 1 - \Phi(1.7) \approx 0.0446,$$

και,

$$\begin{aligned}\Pr(X < -4.7 \text{ ή } X > 2.7) &= \Pr(X < -4.7) + \Pr(X > 2.7) \\ &= \Pr \left(Z < \frac{-4.7 - 1}{1} \right) + 1 - \Pr \left(Z \leq \frac{2.7 - 1}{1} \right) \\ &= \Phi(-5.7) + 1 - \Phi(1.7) \\ &\approx 0 + 1 - 0.9554 = 0.0446.\end{aligned}$$

(γ') Όπως και στο προηγούμενο ερώτημα,

$$\begin{aligned}\Pr(X > 2.1 \text{ ή } -1 < X < 1) &= \Pr(X > 2.1) + \Pr(-1 < X < 1) \\ &= 1 - \Pr(X \leq 2.1) + \Pr(X < 1) - \Pr(X \leq -1) \\ &= 1 - \Pr \left(Z \leq \frac{2.1 - 1}{1} \right) + \Pr \left(Z < \frac{1 - 1}{1} \right) - \Pr \left(Z \leq \frac{-1 - 1}{1} \right) \\ &= 1 - \Phi(1.1) + \Phi(0) - \Phi(-2) \\ &\approx 1 - 0.8643 + 0.5 - 0.0228 \\ &= 0.6129.\end{aligned}$$

(δ') Εδώ έχουμε,

$$\Pr(X > 3) = 1 - \Pr(X \leq 3) = 1 - \Pr\left(Z \leq \frac{3 - (-1)}{2}\right) = 1 - \Phi(2) \approx 0.0228,$$

και,

$$\Pr(X > 2) = 1 - \Pr(X \leq 2) = 1 - \Pr\left(Z \leq \frac{2 - (-1)}{2}\right) = 1 - \Phi(1.5),$$

οπότε,

$$\begin{aligned} \Pr(X > 3 | X > 2) &= \frac{\Pr(X > 3, X > 2)}{\Pr(X > 2)} = \frac{\Pr(X > 3)}{\Pr(X > 2)} \\ &= \frac{1 - \Phi(2)}{1 - \Phi(1.5)} \\ &\approx \frac{1 - 0.9772}{1 - 0.9332} \\ &\approx 0.3413. \end{aligned}$$

(ε') Εφόσον για οποιαδήποτε Τ.Μ. έχουμε $\text{Var}(X) = E(X^2) - [E(X)]^2$,

$$E(Y) = E(1 - X^2) = 1 - E(X^2) = 1 - (\sigma^2 + \mu^2) = 1 - 3.5 - 4 = -6.5,$$

και για τη $V = (X + 2)^{15}$ έχουμε,

$$E(V) = E((X + 2)^{15}) = E(W^{15}),$$

όπου $W = X + 2$ είναι μια Τ.Μ. με κατανομή $N(0, 3.5)$. Αν συμβολίσουμε την πυκνότητα της W με $f(x)$,

$$E(Y) = E(W^{15}) = \int_{-\infty}^{\infty} w^{15} f(w) dw = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} w^{15} e^{-w^2/2\sigma^2} dw,$$

και παρατηρούμε ότι έχουμε το ολοκλήρωμα από το $-\infty$ ως το $+\infty$ της περιττής συνάρτησης $g(w) = \frac{1}{\sqrt{2\pi\sigma^2}} w^{15} e^{-w^2/2\sigma^2}$ [δηλαδή η $g(w)$ ικανοποιεί $g(-w) = -g(w)$], άρα το ολοκλήρωμα ισούται με μηδέν!

(ς') Κατ' αρχάς έχουμε,

$$\Pr(X < 0) = \Pr\left(\frac{X - 2}{\sigma} < \frac{-2}{\sigma}\right) = \Pr(Z < -2/\sigma) = \Phi(-2/\sigma),$$

και από τον Πίνακα 12.2 βρίσκουμε πως η τιμή z για την οποία το $\Phi(z)$ είναι όσο το δυνατόν πιο κοντά στο $1/3$, είναι το $z = -0.43$. Άρα για να ισχύει η ζητούμενη σχέση πρέπει να έχουμε, $-2/\sigma \approx -0.43$, δηλαδή, $\sigma^2 \approx 21.63$.

4. Τετράγωνο κανονικής Τ.Μ. Υπολογίζουμε πρώτα τη συνάρτηση κατανομής. Για την περίπτωση $y \leq 0$, προφανώς $G(y) = 0$, ενώ για $y > 0$,

$$\begin{aligned}
 G(y) &= \Pr(Y \leq y) \\
 &= \Pr(aX^2 \leq y) \\
 &= \Pr\left(|X| \leq \sqrt{\frac{y}{a}}\right) \\
 &= \Pr\left(X \leq \sqrt{\frac{y}{a}}\right) - \Pr\left(X < \sqrt{\frac{y}{a}}\right) \\
 &= \Pr\left(Z \leq \sqrt{\frac{y}{a\sigma^2}}\right) - \Pr\left(Z < \sqrt{\frac{y}{a\sigma^2}}\right) \\
 &= \Phi\left(\sqrt{\frac{y}{a\sigma^2}}\right) - \Phi\left(-\sqrt{\frac{y}{a\sigma^2}}\right),
 \end{aligned}$$

όπου η Τ.Μ. $Z \sim N(0, 1)$, άρα τελικά,

$$G(y) = \begin{cases} \Phi\left(\sqrt{\frac{y}{a\sigma^2}}\right) - \Phi\left(-\sqrt{\frac{y}{a\sigma^2}}\right), & y > 0, \\ 0, & y \leq 0. \end{cases} \quad (\text{A.17})$$

Τώρα η πυκνότητα $g(y)$ μπορεί εύκολα να υπολογιστεί παίρνοντας παραγώγους. Για $y \leq 0$ έχουμε $g(y) = 0$, ενώ για $y > 0$,

$$\begin{aligned}
 g(y) &= \left[\Phi\left(\sqrt{\frac{y}{a\sigma^2}}\right)\right]' - \left[\Phi\left(-\sqrt{\frac{y}{a\sigma^2}}\right)\right]' \\
 &= \Phi'\left(\sqrt{\frac{y}{a\sigma^2}}\right) \left(\sqrt{\frac{y}{a\sigma^2}}\right)' - \Phi'\left(-\sqrt{\frac{y}{a\sigma^2}}\right) \left(-\sqrt{\frac{y}{a\sigma^2}}\right)' \\
 &= \frac{1}{2\sqrt{a\sigma^2 y}} \left[\phi\left(\sqrt{\frac{y}{a\sigma^2}}\right) - \phi\left(-\sqrt{\frac{y}{a\sigma^2}}\right)\right] \\
 &= \frac{1}{\sqrt{2\pi a\sigma^2 y}} \exp\left[-\frac{y}{2a\sigma^2}\right],
 \end{aligned}$$

όπου χρησιμοποιήσαμε τη βασική ιδιότητα της πυκνότητας ως παράγωγο της συνάρτησης κατανομής, δηλαδή $\phi(z) = \Phi'(z)$ για την τυπική κανονική κατανομή, και τον τύπο της $N(0, 1)$ πυκνότητας $\phi(z)$. Τελικά, λοιπόν, προκύπτει ότι,

$$g(y) = \begin{cases} \frac{1}{\sqrt{2\pi a\sigma^2 y}} \exp\left[-\frac{y}{2a\sigma^2}\right], & y > 0, \\ 0, & y \leq 0. \end{cases} \quad (\text{A.18})$$

5. Η κατανομή χ^2 .

(α) Έστω $\phi(z)$ η πυκνότητα της $Z_1 \sim N(0, 1)$ για την οποία, όπως παρατηρήσαμε στην απόδειξη του Θεωρήματος 12.1, ισχύει ότι $-\phi'(z) = z\phi(z)$. Για τη ζητούμενη μέση τιμή, ολοκληρώνοντας κατά παράγοντες βρίσκουμε,

$$E(Z_1^4) = \int_{-\infty}^{\infty} z^4 \phi(z) dz = \int_{-\infty}^{\infty} z^3 (z\phi(z)) dz = [-z^3 \phi(z)]_{-\infty}^{\infty} + 3 \int_{-\infty}^{\infty} z^2 \phi(z) dz,$$

όπου θέσαμε $u = z^3$ και $dv = z\phi(z)dz$, έτσι ώστε $du = 3z^2 dz$ και $v = -\phi(z)$. Και συνεχίζοντας τον παραπάνω υπολογισμό, έχουμε,

$$E(Z_1^4) = 3E(Z_1^2) = 3\text{Var}(Z_1) = 3.$$

(β') Χρησιμοποιώντας το γεγονός ότι $E(Z_i^2) = \text{Var}(Z_i) = 1$ για κάθε i και εφαρμόζοντας το Θεώρημα 11.5, η μέση τιμή της $Y \sim \chi^2(k)$ είναι,

$$E(Y) = E\left(\sum_{i=1}^k Z_i^2\right) = \sum_{i=1}^k E(Z_i^2) = k.$$

(γ') Για $k = 1$, χρησιμοποιώντας το αποτέλεσμα του πρώτου σκέλους, η διασπορά της $Y = Z_1^2 \sim \chi^2(1)$ υπολογίζεται εύκολα ως,

$$\text{Var}(Y) = \text{Var}(Z_1^2) = E(Z_1^4) - [E(Z_1^2)]^2 = 2.$$

Η πυκνότητα και η συνάρτηση κατανομής της $Y = Z_1^2$ έχουν ήδη υπολογιστεί στις σχέσεις (A'.18) και (A'.17) της Άσκησης 4, στην ειδική περίπτωση $\sigma^2 = 1$ και $a = 1$.

6. Λογαριθμοκανονική κατανομή. Και πάλι θα ξεκινήσουμε από τη συνάρτηση κατανομής. Προφανώς για $y \leq 0$ έχουμε $G(y) = 0$, ενώ για $y > 0$,

$$G(y) = \Pr(Y \leq y) = \Pr(e^X \leq y) = \Pr(X \leq \log y) = F(\log y).$$

Και παίρνοντας παραγώγους βρίσκουμε ότι, για $y > 0$,

$$g(y) = \frac{1}{y} F'(\log y),$$

ενώ προφανώς έχουμε $g(y) = 0$ όταν $y \leq 0$.

Με αντικατάσταση της κανονικής συνάρτησης κατανομής και πυκνότητας στις παραπάνω εκφράσεις, προκύπτει ότι,

$$G(y) = \begin{cases} \Phi\left(\frac{\log y - \mu}{\sigma}\right), & y > 0, \\ 0, & y \leq 0, \end{cases}$$

$$g(y) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma y}} \exp\left[-\frac{(\log y - \mu)^2}{2\sigma^2}\right], & y > 0, \\ 0, & y \leq 0. \end{cases}$$

Παρατηρήστε ότι η πυκνότητα $g(y) \rightarrow \infty$ καθώς το $y \rightarrow 0^+$.

7. *Γραμμικός μετασχηματισμός.* Έστω ότι η X έχει κατανομή $N(\mu, \sigma^2)$ με πυκνότητα $f_X(x)$. Ο ευκολότερος τρόπος να βρούμε την κατανομή της $Y = aX + b$ είναι παρατηρώντας, από το Θεώρημα 11.2 και την Άσκηση 6 του Κεφαλαίου 11, ότι (για $a \neq 0$), η πυκνότητα $f_Y(y)$ της Y είναι,

$$\begin{aligned} f_Y(y) &= \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right) \\ &= \frac{1}{|a|} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2} \left(\frac{y-b}{a} - \mu\right)^2\right\} \\ &= \frac{1}{\sqrt{2\pi a^2 \sigma^2}} \exp\left\{-\frac{1}{2a^2 \sigma^2} (y - (a\mu + b))^2\right\}, \end{aligned}$$

την οποία αναγνωρίζουμε φυσικά ως την πυκνότητα της κατανομής $N(a\mu + b, a^2\sigma^2)$.

Η περίπτωση $a = 0$ είναι, προφανώς, τετριμμένη. (Γιατί;)

8. Όλα στο 13! Κατ' αρχάς ορίζουμε τις εξής Τ.Μ.:

$$\begin{aligned} Z_1 &= \text{συνολικά χρήματα του πρώτου παίκτη μετά το παιχνίδι,} \\ Z_2 &= \text{συνολικά χρήματα του δεύτερου παίκτη μετά από όλα τα παιχνίδια,} \\ X_i &= \text{χρήματα που παίρνει πίσω ο δεύτερος παίκτης μετά το παιχνίδι } i, \\ &\quad \text{για } i = 1, 2, \dots, N = 120. \end{aligned}$$

- (α') Η Z_1 παίρνει τις τιμές 36×120 και 0 , με αντίστοιχες πιθανότητες $1/37$ και $36/37$, άρα,

$$E(Z_1) = \frac{1}{37} \times 36 \times 120 + 0 \times \frac{36}{37} = 120 \times \frac{36}{37} \approx 116.76.$$

Η Z_2 ισούται με το άθροισμα των X_i , όπου η κάθε X_i παίρνει τις τιμές 36 και 0 με αντίστοιχες πιθανότητες $1/37$ και $36/37$, οπότε,

$$E(X_i) = \frac{1}{37} \times 36 + 0 \times \frac{36}{37} = \frac{36}{37},$$

συνεπώς η μέση τιμή της Z_2 είναι,

$$E(Z_2) = E\left(\sum_{i=1}^{120} X_i\right) = \sum_{i=1}^{120} E(X_i) = 120E(X_1) = 120 \times \frac{36}{37} \approx 116.76,$$

η οποία είναι ακριβώς ίδια με την $E(Z_1)$.

(β') Παρομοίως υπολογίζουμε και τις αντίστοιχες διασπορές. Για τον πρώτο παίκτη,

$$\begin{aligned} \text{Var}(Z_1) &= E(Z_1^2) - (E(Z_1))^2 \\ &= \left(\frac{1}{37} \times (36 \times 120)^2 + \frac{36}{37} \times 0^2 \right) - \left(120 \times \frac{36}{37} \right)^2 \\ &= 120^2 \times 35 \times \left(\frac{36}{37} \right)^2. \end{aligned}$$

Για κάθε γύρο του παιχνιδιού του δεύτερου παίκτη έχουμε,

$$\text{Var}(X_i) = E(X_i^2) - (E(X_i))^2 = \frac{1}{37} \times 36^2 + \frac{36}{37} \times 0^2 - \left(\frac{36}{37} \right)^2 = 35 \times \left(\frac{36}{37} \right)^2,$$

και επειδή τα διαδοχικά του παιχνίδια είναι ανεξάρτητα,

$$\text{Var}(Z_2) = \text{Var}\left(\sum_{i=1}^{120} X_i\right) = \sum_{i=1}^{120} \text{Var}(X_i) = 120 \times 35 \times \left(\frac{36}{37} \right)^2.$$

Παρατηρούμε πως, παρότι οι δύο παίκτες έχουν το ίδιο αναμενόμενο κέρδος, η διασπορά των χρημάτων του πρώτου παίκτη είναι 120 φορές μεγαλύτερη από του δεύτερου! Αυτό είναι αναμενόμενο, μια και ο πρώτος παίκτης ή τα χάνει όλα ή έχει τεράστιο κέρδος, ενώ για τον δεύτερό υπάρχουν και πολλά «ενδιάμεσα» ενδεχόμενα.

(γ') Η Z_1 παίρνει μόνο 2 τιμές, $Z_1 = 0$ με πιθανότητα $36/37$, και $Z_1 = 36 \times 120 = 4320$ με πιθανότητα $1/37$. Επομένως,

$$\Pr(Z_1 > 100) = \Pr(Z_1 = 4320) = \frac{1}{37}.$$

Η Z_2 , από την άλλη, ισούται με το άθροισμα $N = 120$ ανεξάρτητων τυχαίων μεταβλητών με ίδια κατανομή, άρα μπορούμε να προσεγγίσουμε την κατανομή της με βάση το Κ.Ο.Θ. Κατ' αρχάς έχουμε,

$$\Pr(Z_2 > 100) = \Pr\left(\sum_{i=1}^{120} X_i > 100\right) = 1 - \Pr\left(\sum_{i=1}^{120} X_i \leq 100\right),$$

και σημειώνοντας πως πιο πάνω βρήκαμε για τις Τ.Μ. X_i ότι $\mu = E(X_i) = 36/37$ και $\sigma^2 = \text{Var}(X_i) = 35(36/37)^2$, κανονικοποιώντας το πιο πάνω άθροισμα και χρησιμοποιώντας το Κ.Ο.Θ.:

$$\begin{aligned} \Pr(Z_2 > 100) &= 1 - \Pr\left(\frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N (X_i - \mu) \leq \frac{100 - 120 \times \frac{36}{37}}{\sqrt{35} \times \frac{36}{37} \times \sqrt{120}}\right) \\ &\approx 1 - \Pr(\bar{S}_N \leq -0.27) \\ &\approx 1 - \Phi(-0.27) \\ &\approx 0.6064. \end{aligned}$$

Άρα, ο πρώτος παίκτης που τα ποντάρει όλα μαζί σε ένα παιχνίδι είναι πολύ πιο απίθανο να βγει «λίγο χαμένος», κατά μόνο 20 ευρώ, από τον πιο συντηρητικό δεύτερο παίκτη, ο οποίος ποντάρει σιγά σιγά και γι' αυτόν τον λόγο τα χρήματά του έχουν μικρότερες τυχαίες διακυμάνσεις.

9. *Κόκκινο-μαύρο.* Έστω $N = 150$ ανεξάρτητες Τ.Μ. X_1, X_2, \dots, X_N όπου η κάθε X_i ισούται με $+1$ με πιθανότητα $16/33$, ή με -1 με πιθανότητα $17/33$. Έτσι, αν ορίσουμε την Τ.Μ. $Y = X_1 + X_2 + \dots + X_N$, η Y περιγράφει το κέρδος του παίκτη μετά από N γύρους, και ο παίκτης τελικά δεν θα έχει βγει χαμένος αν $Y \geq 0$.

Θα προσεγγίσουμε την κατανομή του Y μέσω του Κ.Ο.Θ., οπότε θα χρειαστούμε τη μέση τιμή και τη διασπορά των X_i ,

$$\begin{aligned}\mu &= E(X_i) = (+1) \times \frac{16}{33} + (-1) \times \frac{17}{33} = -\frac{1}{33}, \\ \sigma^2 &= \text{Var}(X_i) = E(X_i^2) - \mu^2 = 1 - \left(\frac{1}{33}\right)^2 = \frac{1088}{1089},\end{aligned}$$

όπου χρησιμοποίησαμε το γεγονός ότι πάντοτε $X_i^2 = 1$. Τώρα μπορούμε εύκολα να εφαρμόσουμε το Κ.Ο.Θ. με $\sigma = \sqrt{1088}/33 \approx 0.9995408$, για να υπολογίσουμε τη ζητούμενη πιθανότητα,

$$\Pr(Y \geq 0) = \Pr(Y \geq 1/2) = \Pr(\bar{X}_N \geq 1/300),$$

όπου $\bar{X}_N = Y/N$ είναι ο εμπειρικός μέσος όρος των X_i , και όπως έχουμε εξηγήσει, επειδή η Τ.Μ. Y παίρνει μόνο ακέραιες τιμές, προσθέσαμε ένα επιπλέον $1/2$. Συνεχίζοντας τον παραπάνω υπολογισμό, η $\Pr(Y \geq 0)$ ισούται με,

$$\Pr\left(\frac{1}{N} \sum_{i=1}^N (X_i - \mu) > \frac{1}{300} + \frac{1}{33}\right) = \Pr\left(\frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N (X_i - p) > \left(\frac{1}{300} + \frac{1}{33}\right) \frac{\sqrt{150}}{\sigma}\right),$$

άρα τελικά από το Κ.Ο.Θ. παίρνουμε την προσέγγιση,

$$\Pr(Y \geq 0) \approx \Pr(\bar{S}_N > 0.412) \approx 1 - \Phi(0.41) \approx 34.1\%.$$

10. *Online game.*

(α') Από τον ορισμό και τις απλές ιδιότητες της μέσης τιμής και της διασποράς, εύκολα βρίσκουμε ότι:

$$\begin{aligned}\mu &= E(X) = 0 \times 0.3 + 20 \times 0.6 + 40 \times 0.1 = 16, \\ E(X^2) &= 0^2 \times 0.3 + 20^2 \times 0.6 + 40^2 \times 0.1 = 400, \\ \sigma^2 &= \text{Var}(X) = E(X^2) - (E(X))^2 = 144.\end{aligned}$$

(β') Έστω $N = 100$ ανεξάρτητες Τ.Μ. X_1, X_2, \dots, X_M , όπου η κάθε X_i ισούται με τις μονάδες ζωής που αφαιρεί η βολή i και έχει την ίδια κατανομή με την X του προηγούμενου σκέλους. Θέλουμε να υπολογίσουμε την πιθανότητα το άθροισμα όλων των X_i να ξεπερνάει το 1700. Θα την προσεγγίσουμε χρησιμοποιώντας το Κ.Ο.Θ.: Εφόσον οι X_i έχουν μέση τιμή $\mu = 16$ και τυπική απόκλιση $\sigma = 12$,

$$\begin{aligned} \Pr\left(\sum_{i=1}^{100} X_i > 1700\right) &= \Pr\left(\sum_{i=1}^{100} (X_i - \mu) > 1700 - 16 \times 100\right) \\ &= \Pr\left(\frac{1}{\sigma\sqrt{N}} \sum_{i=1}^{100} (X_i - \mu) > \frac{100}{12 \times 10}\right) \\ &= \Pr(\bar{S}_N > 5/6) \\ &\approx 1 - \Phi(5/6) \approx 0.2023. \end{aligned}$$

11. Χαμένος χρόνος. Έστω $N = 259$ ανεξάρτητες Τ.Μ. X_1, X_2, \dots, X_N όπου η κάθε X_i έχει $\text{Exp}(20)$ κατανομή, και περιγράφει πόσα λεπτά περιμένα το λεωφορείο τη μέρα i . Αν ορίσουμε την Τ.Μ. $Y = X_1 + X_2 + \dots + X_N$, τότε η Y περιγράφει συνολικά πόσο χρόνο περιμένα το λεωφορείο (σε λεπτά) σε ένα χρόνο. Μας ενδιαφέρει η πιθανότητα το Y να είναι μεγαλύτερο από $4 \times 24 \times 60 = 5760$ λεπτά.

Εφόσον η κάθε X_i έχει γνωστή μέση τιμή και τυπική απόκλιση $\mu = \sigma = 20$, χρησιμοποιώντας για ακόμα μια φορά το Κ.Ο.Θ.:

$$\begin{aligned} \Pr(Y > 5760) &= \Pr\left(\sum_{i=1}^N X_i > 5760\right) \\ &= \Pr\left(\sum_{i=1}^N (X_i - \mu) > 5760 - 259 \times 20\right) \\ &= \Pr\left(\frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N (X_i - \mu) > \frac{580}{20 \times \sqrt{259}}\right) \\ &= \Pr(\bar{S}_N > 1.802) \\ &\approx 1 - \Phi(1.8) \approx 3.6\%. \end{aligned}$$

12. Ζυγοβίστι Αρκαδίας.

(α') Παρατηρούμε ότι η X έχει ομοιόμορφη κατανομή στο διάστημα $[30, 130]$, επομένως, από τις γνωστές ιδιότητες της ομοιόμορφης κατανομής, έχουμε:

$$\mu = E(X) = \frac{30 + 130}{2} = 80, \quad \text{και} \quad \sigma^2 = \text{Var}(X) = \frac{(130 - 30)^2}{12} = \frac{2500}{3}.$$

(β') Από το Θεώρημα 11.5 απλώς έχουμε,

$$E(Y) = E(20X + 500) = 20E(X) + 500 = 20 \times 80 + 500 = 2100.$$

(γ') Θα χρησιμοποιήσουμε το Κ.Ο.Θ. Έστω $X_i, i = 1, 2, \dots, N$, η βροχόπτωση για κάθε ένα από τα $N = 80$ χρόνια. Η πιθανότητα να πετύχει η δενδροφύτευση ισούται με:

$$\begin{aligned} \Pr\left(\sum_{i=1}^N X_i < 5200\right) &= \Pr\left(\sum_{i=1}^N (X_i - \mu) < 7000 - 80 \times 80\right) \\ &= \Pr\left(\frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N (X_i - \mu) < \frac{600}{\sqrt{2500/3}\sqrt{80}}\right) \\ &= \Pr(\bar{S}_N < 2.324) \\ &\approx \Phi(2.32) \\ &\approx 98.98\%. \end{aligned}$$

13. *Αποσυγχρονισμός.* Για κάθε $i = 1, 2, \dots, N = 200$, έστω X_i η Τ.Μ. που περιγράφει κατά πόσα δέκατα του δευτερολέπτου αποσυγχρονίστηκε το βίντεο τη φορά i που το μετακίνησα. Ξέρουμε ότι οι Τ.Μ. X_i είναι ανεξάρτητες, και από την Άσκηση 8 του Κεφαλαίου 10 επίσης γνωρίζουμε ότι έχουν μέση τιμή $\mu = E(X_i) = 0$ και διασπορά $\sigma^2 = \text{Var}(X_i) = 1/8$.

Μας ενδιαφέρει η πιθανότητα το άθροισμα των X_i να ξεπερνά σε απόλυτη τιμή το ένα δευτερολέπτο. Χρησιμοποιώντας το Κ.Ο.Θ. μπορούμε εύκολα να την εκτιμήσουμε ως,

$$\begin{aligned} \Pr\left(\left|\sum_{i=1}^n X_i\right| > 10\right) &= \Pr\left(\left|\sum_{i=1}^n (X_i - \mu)\right| > 10\right) \\ &= \Pr\left(\left|\frac{1}{\sigma\sqrt{N}} \sum_{i=1}^n (X_i - \mu)\right| > \frac{10}{\sqrt{1/8}\sqrt{200}}\right) \\ &= \Pr(|\bar{S}_N| > 2) \\ &= \Pr(\bar{S}_N < -2) + 1 - \Pr(\bar{S}_N \leq 2) \\ &\approx \Phi(-2) + 1 - \Phi(2) \\ &\approx 0.0456. \end{aligned}$$

14. *Τηλεφωνικές κλήσεις.* Έστω μια Τ.Μ. $Y \sim \text{Exp}(85)$. Η πιθανότητα μια κλήση να είναι σύντομη είναι,

$$p = \Pr(Y \leq 60) = 1 - e^{-60/85} \approx 0.55.$$

Έστω, τώρα, ανεξάρτητες $\text{Bern}(p)$ Τ.Μ. X_i , όπου η κάθε $X_i = 1$ αν η κλήση i είναι σύντομη. Η ζητούμενη πιθανότητα είναι,

$$\begin{aligned} \Pr\left(\sum_{i=1}^N X_i < 120\right) &= \Pr\left(\sum_{i=1}^N X_i \leq 119.5\right) \\ &= \Pr\left(\sum_{i=1}^N (X_i - \mu) \leq 119.5 - 250 \times 0.55\right), \end{aligned}$$

όπου $\mu = p = 0.55$ είναι η μέση τιμή των X_i , $N = 250$ είναι το πλήθος τους, και αφαιρέσαμε $1/2$ με ακριβώς το ίδιο σκεπτικό όπως στο Παράδειγμα 12.5 και το Πόρισμα 12.1. Εφόσον τα X_i έχουν διασπορά $\sigma^2 = \text{Var}(X) = p(1-p) = 0.2465$, χρησιμοποιώντας το Κ.Ο.Θ. η παραπάνω πιθανότητα μπορεί να εκφραστεί ως,

$$\begin{aligned} \Pr\left(\frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N (X_i - \mu) \leq \frac{-18}{\sqrt{0.2465 \times 250}}\right) &\approx \Pr(Z \leq -2.29) \\ &= \Phi(-2.29) \\ &\approx 0.011. \end{aligned}$$

15. Εκτίμηση με θόρυβο.

(α) Εφαρμόζοντας όπως και στις προηγούμενες περιπτώσεις το Κ.Ο.Θ., για οποιοδήποτε μέγεθος δείγματος N έχουμε ότι η πιθανότητα που μας ενδιαφέρει μπορεί να προσεγγιστεί ως,

$$\begin{aligned} \Pr(|\bar{X}_N - \mu| \leq 0.2) &= \Pr\left(-0.2 \leq \frac{1}{N} \sum_{i=1}^N (X_i - \mu) \leq 0.2\right) \\ &= \Pr\left(-\frac{\sqrt{N}}{5\sigma} \leq \frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N (X_i - \mu) \leq \frac{\sqrt{N}}{5\sigma}\right) \\ &= \Pr\left(-\sqrt{N}/5 \leq \bar{S}_N \leq \sqrt{N}/5\right) \\ &= 1 - \Pr\left(\bar{S}_N < -\sqrt{N}/5\right) - \Pr\left(\bar{S}_N > \sqrt{N}/5\right) \\ &\approx 1 - \Pr\left(Z < -\sqrt{N}/5\right) - \Pr\left(Z > \sqrt{N}/5\right) \\ &= 1 - 2\Phi\left(-\frac{\sqrt{N}}{5}\right), \end{aligned}$$

όπου η $Z \sim N(0,1)$, και χρησιμοποιήσαμε το γεγονός ότι η πυκνότητα $\phi(z)$ της τυπικής κανονικής κατανομής είναι συμμετρική γύρω από το μηδέν, και άρα, για οποιοδήποτε z , $\Pr(Z < -z) = \Pr(Z > z)$.

Για να ισούται αυτή η πιθανότητα με 0.95, θα πρέπει να έχουμε $1 - 2\Phi(-\sqrt{N}/5) \approx 0.95$, δηλαδή $\Phi(-\sqrt{N}/5) \approx 0.025$, το οποίο, από τους πίνακες τιμών της $\Phi(z)$, συμβαίνει για $-\sqrt{N}/5 \approx -1.96$, δηλαδή για $N \approx 96$.

- (β') Όπως στη λύση της Άσκησης 13 του Κεφαλαίου 9, ορίζουμε τις νέες (ανεξάρτητες) Τ.Μ. $W_i = X_i + Z_i - \nu$, για $i = 1, 2, \dots, N$, και εκτιμούμε το μ μέσω του εμπειρικού μέσου όρου,

$$\bar{W}_N = \frac{1}{N} \sum_{i=1}^N W_i = \frac{1}{N} \sum_{i=1}^N (X_i + Z_i - \nu),$$

Όπως είδαμε εκεί, ο εμπειρικός μέσος όρος \bar{W}_N πράγματι συγκλίνει κατά πιθανότητα στη ζητούμενη τιμή μ .

Για να επιλέξουμε το μέγεθος του δείγματος N , κατ' αρχάς εύκολα υπολογίζουμε,

$$\begin{aligned} \xi &= E(W_i) = E(X_i + Z_i - \nu) = E(X_i) + E(Z_i) - \nu = \mu, \\ \rho^2 &= \text{Var}(W_i) = \text{Var}(X_i + Z_i - \nu) = \text{Var}(X_i) + \text{Var}(Z_i) = \sigma^2 + \tau^2 = 4, \end{aligned}$$

όπου χρησιμοποιήσαμε τις βασικές ιδιότητες της μέσης τιμής και της διασποράς. Οπότε, τώρα μπορούμε να επαναλάβουμε ακριβώς τα ίδια βήματα όπως στο προηγούμενο ερώτημα, αντικαθιστώντας το \bar{X}_N με το \bar{W}_N , και το κανονικοποιημένο άθροισμα \bar{S}_N με το αντίστοιχο,

$$\bar{T}_N = \frac{1}{\rho\sqrt{N}} \sum_{i=1}^N (W_i - \xi) = \frac{1}{2\sqrt{N}} \sum_{i=1}^N (W_i - \mu).$$

Έτσι, παρομοίως βρίσκουμε,

$$\begin{aligned} \Pr(|\bar{W}_N - \mu| \leq 0.2) &= \Pr\left(-0.2 \leq \frac{1}{N} \sum_{i=1}^N (W_i - \mu) \leq 0.2\right) \\ &= \Pr\left(-\frac{\sqrt{N}}{10} \leq \frac{1}{2\sqrt{N}} \sum_{i=1}^N (W_i - \mu) \leq \frac{\sqrt{N}}{10}\right) \\ &= \Pr\left(-\sqrt{N}/10 \leq \bar{T}_N \leq \sqrt{N}/10\right) \\ &\approx 1 - \Pr\left(Z < -\sqrt{N}/10\right) - \Pr\left(Z > \sqrt{N}/10\right) \\ &= 1 - 2\Phi\left(-\frac{\sqrt{N}}{10}\right). \end{aligned}$$

Και πάλι, για να είναι η πιθανότητα αυτή κατά προσέγγιση ίση με 0.95, θα πρέπει να έχουμε $\Phi(-\sqrt{N}/10) \approx 0.025$, το οποίο, από τους πίνακες τιμών της $\Phi(z)$, συμβαίνει για $-\sqrt{N}/10 \approx -1.96$, δηλαδή για $N \approx 384$.

Παρατηρούμε πως, αφού τετραπλασιάστηκε η διασπορά των δειγμάτων μας (αρχικά $\text{Var}(X_i) = 1$ ενώ $\text{Var}(W_i) = 4$), επίσης τετραπλασιάστηκε και το απαιτούμενο μέγεθος του δείγματος ώστε η εκτίμησή μας να διατηρήσει την ίδια ακρίβεια.

Α'.13 Ασκήσεις Κεφαλαίου 13

1. Σύγκλιση κατά πιθανότητα \Rightarrow σύγκλιση κατά κατανομή. Για κάθε n , συμβολίζουμε με $F_n(x)$ τη συνάρτηση κατανομής της X_n και αντίστοιχα με $F(x)$ τη συνάρτηση κατανομής της X . Έστω ένα οποιοδήποτε $x \in \mathbb{R}$ στο οποίο η $F(x)$ είναι συνεχής. Θα δείξουμε το ζητούμενο αποτέλεσμα σε δύο βήματα.

Πρώτα, για οποιοδήποτε $n \geq 1$ και κάθε $\epsilon > 0$, έχουμε,

$$\begin{aligned}
 F_n(x) &= \Pr(X_n \leq x) \\
 &= \Pr((X_n - X) + X \leq x) \\
 &\stackrel{(a)}{=} \Pr\left(\{(X_n - X) + X \leq x\} \cap \{|X_n - X| < \epsilon\}\right) \\
 &\quad + \Pr\left(\{(X_n - X) + X \leq x\} \cap \{|X_n - X| \geq \epsilon\}\right) \\
 &\stackrel{(b)}{\leq} \Pr\left(\{X \leq x - (X_n - X)\} \cap \{|X_n - X| < \epsilon\}\right) + \Pr(|X_n - X| \geq \epsilon) \\
 &\stackrel{(c)}{\leq} \Pr(X \leq x + \epsilon) + \Pr(|X_n - X| \geq \epsilon) \\
 &= F(x + \epsilon) + \Pr(|X_n - X| \geq \epsilon),
 \end{aligned}$$

όπου στο βήμα (a) χρησιμοποιήσαμε τον κανόνα συνολικής πιθανότητας, στο (b) το προφανές γεγονός ότι η τομή δύο ενδεχομένων είναι υποσύνολο του καθενός από τα δύο, σε συνδυασμό με τον δεύτερο κανόνα πιθανότητας, και στο (c) πάλι εφαρμόσαμε τον δεύτερο κανόνα πιθανότητας, αφού, για οποιοδήποτε δ με $|\delta| < \epsilon$, $\{X \leq x - \delta\} \subseteq \{X \leq x + \epsilon\}$.

Τώρα, παίρνοντας το όριο $n \rightarrow \infty$ στην παραπάνω ανισότητα, εφόσον $\Pr(|X_n - X| \geq \epsilon) \rightarrow 0$ από την υπόθεση ότι τα $\{X_n\}$ συγκλίνουν κατά πιθανότητα, έχουμε ότι για κάθε $\epsilon > 0$:

$$\limsup_{n \rightarrow \infty} F_n(x) \leq F(x + \epsilon).$$

Και παίρνοντας το όριο $\epsilon \rightarrow 0^+$, από τη συνέχεια της $F(x)$ στο x συμπεραίνουμε πως,

$$\limsup_{n \rightarrow \infty} F_n(x) \leq F(x). \tag{A'.19}$$

Παρομοίως, για οποιοδήποτε $n \geq 1$ και κάθε $\epsilon > 0$, προς την αντίθετη κατεύθυνση έχουμε,

$$\begin{aligned}
 F_n(x) &= \Pr\left(\{(X_n - X) + X \leq x\} \cap \{|X_n - X| < \epsilon\}\right) \\
 &\quad + \Pr\left(\{(X_n - X) + X \leq x\} \cap \{|X_n - X| \geq \epsilon\}\right) \\
 &\geq \Pr\left(\{X \leq x - (X_n - X)\} \cap \{|X_n - X| < \epsilon\}\right) \\
 &\stackrel{(d)}{\geq} \Pr(X \leq x - \epsilon) \\
 &= F(x - \epsilon),
 \end{aligned}$$

όπου το βήμα (d) προκύπτει από τον δεύτερο κανόνα πιθανότητας, αφού, για οποιοδήποτε δ με $|\delta| < \epsilon$, $\{X \leq x - \epsilon\} \subseteq \{X \leq x - \delta\}$. Παίρνοντας το όριο $n \rightarrow \infty$, αμέσως προκύπτει πως,

$$\liminf_{n \rightarrow \infty} F_n(x) \geq F(x - \epsilon),$$

για κάθε $\epsilon > 0$, και παίρνοντας τώρα το όριο $\epsilon \rightarrow 0^+$, από τη συνέχεια της $F(x)$ στο x , έχουμε,

$$\liminf_{n \rightarrow \infty} F_n(x) \geq F(x). \quad (\text{A'.20})$$

Ο συνδυασμός των (A'.19) και (A'.20) συνεπάγεται πως $\lim_{n \rightarrow \infty} F_n(x) = F(x)$, και άρα οι $\{X_n\}$ συγκλίνουν στη X κατά κατανομή.

2. Σύγκλιση κατά κατανομή $\not\Rightarrow$ σύγκλιση κατά πιθανότητα. Θα ξεκινήσουμε αποδεικνύοντας το (γ'). Έχουμε, για αυθαίρετο $\epsilon > 0$,

$$\Pr(|X_n - Y| < \epsilon) = \Pr\left(\frac{1}{n}Z < \epsilon\right) = \Pr(Z < n\epsilon),$$

το οποίο όχι απλώς τείνει στο 1, αλλά είναι ίσο με 1 για κάθε $n > 1/\epsilon$. Συνεπώς οι $\{X_n\}$ τείνουν στην Y κατά πιθανότητα, αποδεικνύοντας το ζητούμενο του (γ'). Αυτό, σε συνδυασμό με το αποτέλεσμα της προηγούμενης άσκησης, επίσης συνεπάγεται ότι οι $\{X_n\}$ τείνουν στην Y κατά κατανομή, αποδεικνύοντας το (α').

Για το (β'), παρατηρούμε πως η συνάρτηση κατανομής $F(z)$ της Z είναι η

$$F(z) = \begin{cases} 0, & \text{για } z < 0, \\ \frac{1}{2}, & \text{για } 0 \leq z < 1, \\ 1, & \text{για } z \geq 1, \end{cases}$$

και πως η κάθε X_n παίρνει τέσσερις δυνατές τιμές, τις $\{0, 1/n, 1, 1 + 1/n\}$ με πιθανότητα $1/4$ για την καθεμία, συνεπώς η συνάρτηση κατανομής $F_n(x)$ της X_n είναι,

$$F_n(x) = \begin{cases} 0, & \text{για } x < 0, \\ \frac{1}{4}, & \text{για } 0 \leq x < 1/n, \\ \frac{1}{2}, & \text{για } 1/n \leq x < 1, \\ \frac{3}{4}, & \text{για } 1 \leq x < 1 + 1/n, \\ 1, & \text{για } x \geq 1 + 1/n. \end{cases}$$

Θα δείξουμε, για κάθε z στο οποίο η $F(z)$ είναι συνεχής, δηλαδή για κάθε πραγματικό $z \neq 0, 1$, ότι $F_n(z) \rightarrow F(z)$ καθώς $n \rightarrow \infty$. Για $z < 0$, η σύγκλιση είναι προφανής και τετριμμένη αφού $F(z) = F_n(z) = 0$ για κάθε n . Όταν $z > 1$, για όλα τα n που είναι

αρκετά μεγάλα ώστε να έχουμε $1 + 1/n \leq z$, $F_n(z) = 1$, η οποία προφανώς συγκλίνει στην $F(z) = 1$. Τέλος, όταν το $z \in (0, 1)$, για κάθε $n > 1/z$ έχουμε $F_n(z) = F(z) = 1/2$. Άρα έχουμε αποδείξει ότι $F_n(z) \rightarrow F(z)$ καθώς $n \rightarrow \infty$, για όλα τα $z \neq 0, 1$, συνεπώς οι $\{X_n\}$ συγκλίνουν στη Z κατά κατανομή, αποδεικνύοντας το (β').

Για να αποδείξουμε το (δ'), παρατηρούμε ότι η Τ.Μ. $W = X_n - Z = Y - (1 - \frac{1}{n})Z$ παίρνει τέσσερις δυνατές τιμές, τις $\{-1 + 1/n, 0, 1/n, 1\}$ με πιθανότητα $1/4$ για την καθεμία. Συνεπώς, για οποιοδήποτε $\epsilon < 1/4$ και για κάθε $n \geq 5$, η πιθανότητα,

$$\Pr(|X_n - Z| < \epsilon) = \Pr(|W| < \epsilon) = \Pr(W = 0 \text{ ή } W = 1/n) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2},$$

το οποίο φυσικά δεν τείνει στο 1, άρα οι $\{X_n\}$ δεν τείνουν στη Z κατά πιθανότητα.

3. Σύγκλιση κατά κατανομή και κατά πιθανότητα σε σταθερά. Έστω $F_n(x)$ η συνάρτηση κατανομής της X_n , για κάθε n . Η σταθερά c , ως τυχαία μεταβλητή, έχει συνάρτηση κατανομής,

$$F(x) = \begin{cases} 0, & \text{για } x < c, \\ 1, & \text{για } x \geq c. \end{cases}$$

Εφόσον οι $\{X_n\}$ τείνουν στην c κατά κατανομή, έχουμε ότι για κάθε $x \neq c$,

$$F_n(x) = \Pr(X_n \leq x) \rightarrow F(x), \text{ καθώς } n \rightarrow \infty.$$

Για να αποδείξουμε ότι η σύγκλιση ισχύει και κατά πιθανότητα, παρατηρούμε ότι, για αυθαίρετο $\epsilon > 0$,

$$\begin{aligned} \Pr(|X_n - c| < \epsilon) &= \Pr(-\epsilon < X_n - c < \epsilon) \\ &= \Pr(X_n - c < \epsilon) - \Pr(X_n - c \leq -\epsilon) \\ &= \Pr(X_n < c + \epsilon) - \Pr(X_n \leq c - \epsilon) \\ &\geq \Pr(X_n \leq c + \epsilon/2) - \Pr(X_n \leq c - \epsilon) \\ &= F_n(c + \epsilon/2) - F_n(c - \epsilon), \end{aligned}$$

όπου η ανισότητα προκύπτει από τον δεύτερο κανόνα πιθανότητας. Εφόσον, τώρα, καθώς το $n \rightarrow \infty$,

$$F_n(c + \epsilon/2) - F_n(c - \epsilon) \rightarrow F(c + \epsilon/2) - F(c - \epsilon) = 1 - 0 = 1,$$

συμπεραίνουμε πως και οι πιθανότητες $\Pr(|X_n - c| < \epsilon) \rightarrow 1$, άρα οι $\{X_n\}$ συγκλίνουν στη c και κατά πιθανότητα.

4. Η ουρά της μέσης τιμής μιας συνεχούς Τ.Μ. Έστω πως η Τ.Μ. Y έχει πυκνότητα $f(y)$ και σύνολο τιμών κάποιο $S \subset [0, \infty)$. Η απόδειξη σε αυτή την περίπτωση είναι απολύτως ανάλογη με εκείνη που είδαμε στην Άσκηση 3 του Κεφαλαίου 7 για διακριτές Τ.Μ.

Εδώ, η μέση τιμή που μας ενδιαφέρει, από τον ορισμό της μέσης τιμής μιας συνάρτησης κάποιας Τ.Μ., είναι,

$$E[YH_x(Y)] = \int_0^\infty yf(y)H_x(y) dy,$$

και παρατηρώντας πως $H_x(y) \neq 0$ αν και μόνο αν $y > x$, έχουμε,

$$E[YH_x(Y)] = \int_x^\infty yf(y)H_x(y) dy = \int_x^\infty yf(y) dy.$$

Αλλά, εφόσον μας δίνεται πως το καταχρηστικό ολοκλήρωμα $E(Y) = \int_0^\infty yf(y) dy$ είναι πεπερασμένο, απαραίτητα η «ουρά» του ολοκληρώματος θα τείνει στο μηδέν, δηλαδή,

$$E[YH_x(Y)] = \int_x^\infty yf(y) dy \rightarrow 0,$$

καθώς το $x \rightarrow \infty$, που είναι ακριβώς το ζητούμενο αποτέλεσμα.

5. Σύγκλιση ως προς την απόσταση χ^2 . Έστω πως οι αποστάσεις $d_{\chi^2}(P_n, P) \rightarrow 0$ καθώς $n \rightarrow \infty$. Από το αποτέλεσμα της Άσκησης 18 του Κεφαλαίου 6, έχουμε πως, καθώς $n \rightarrow \infty$,

$$0 \leq \sum_{x \in S} |P_n(x) - P(x)| \leq \sqrt{d_{\chi^2}(P_n, P)} \rightarrow 0,$$

το οποίο συνεπάγεται πως,

$$P_n(x) \rightarrow P(x), \quad \text{για κάθε } x \in S.$$

Αν η κάθε X_n έχει συνάρτηση κατανομής $F_n(x)$, και η X έχει συνάρτηση κατανομής $F(x)$, τότε, για οποιοδήποτε $y \in \mathbb{R}$, χρησιμοποιώντας τη σύγκλιση όλων των τιμών $P_n(x)$ των πυκνοτήτων P_n , έχουμε, καθώς $n \rightarrow \infty$,

$$F_n(y) = \Pr(X_n \leq y) = \sum_{x \in S: x \leq y} P_n(x) \rightarrow \sum_{x \in S: x \leq y} P(x) = \Pr(X \leq y) = F(y),$$

το οποίο μας λέει ακριβώς ότι οι $\{X_n\}$ τείνουν στη X κατά κατανομή.

6. Σύγκλιση της διωνυμικής στην Poisson. Παρατηρούμε κατ' αρχάς πως αρκεί να αποδείξουμε απευθείας το ερώτημα (β'). Έστω λοιπόν $\{X_n\}$ μια ακολουθία Τ.Μ. με αντίστοιχες πυκνότητες $P_n(k)$, συναρτήσεις κατανομής $F_n(x)$, και σύνολο τιμών $S = \{0, 1, 2, \dots\}$. Και έστω Z μια άλλη Τ.Μ. με το ίδιο σύνολο τιμών S , πυκνότητα $P(k)$ και συνάρτηση κατανομής $F(x)$. Υποθέτουμε ότι οι P_n συγκλίνουν στην P όπως στην σχέση (13.25).

Για $x < 0$ έχουμε, εξ ορισμού, $F_n(x) = F(x) = 0$, ενώ για οποιοδήποτε $x \geq 0$ η $F_n(x)$ μπορεί να εκφραστεί ως,

$$F_n(x) = \sum_{0 \leq k \leq x} P_n(k).$$

Αλλά από την (13.25) ξέρουμε πως ο κάθε όρος $P_n(k)$ σε αυτό το πεπερασμένο άθροισμα τείνει, καθώς το $n \rightarrow \infty$, στο αντίστοιχο $P(k)$, και συνεπώς,

$$F_n(x) = \sum_{0 \leq k \leq x} P_n(k) \rightarrow \sum_{0 \leq k \leq x} P(k) = F(x),$$

άρα οι X_n τείνουν στη Z κατά κατανομή.

Παρατηρούμε, τέλος, πως σ' αυτήν την περίπτωση η πιο πάνω σύγκλιση ισχύει για όλα τα $x \in \mathbb{R}$, όχι μόνο για εκείνα στα οποία η $F(x)$ είναι συνεχής.

Α'.14 Ασκήσεις Κεφαλαίου 14

1. *Διάρκειες τραγουδιών.* Έστω X_1, X_2, \dots, X_N οι διάρκειες των $N = 145$ τραγουδιών στην έρευνα της εταιρίας, όπου υποθέτουμε ότι οι Τ.Μ. X_i είναι ανεξάρτητες, με άγνωστο μέσο $\mu = E(X_i)$ και τυπική απόκλιση $\sigma = \sqrt{\text{Var}(X_i)} = 54/60 = 0.9$ λεπτά.

(α') Ακολουθούμε τη γενική μέθοδο της Ενότητας 14.1 στην περίπτωση γνωστής διασποράς. Έχοντας ήδη την τιμή του εμπειρικού μέσου $\bar{X}_N = 3.12$, για το επίπεδο εμπιστοσύνης $c = 98\%$ βρίσκουμε την τιμή του $z^* = 2.326$, και από τη σχέση (14.5) υπολογίζουμε το μέγεθος του στατιστικού σφάλματος $e = \sigma z^*/\sqrt{N} \approx 0.1738$ λεπτά ή $e \approx 10.4$ δευτερόλεπτα. Άρα, έχουμε το διάστημα εμπιστοσύνης $[2.9462, 3.2938]$ με επίπεδο εμπιστοσύνης 98%.

(β') Ακολουθώντας την ίδια μεθοδολογία, από τη σχέση (14.6) βρίσκουμε πως το ελάχιστο πλήθος δειγμάτων που απαιτούνται ώστε να έχουμε στατιστικό σφάλμα $e = 5$ δευτερόλεπτα, δηλαδή $e = 0.0833$ λεπτά, είναι $N = (\sigma z^*/e)^2 \approx 632$ δείγματα.

2. *Ελέφαντες.* Έχουμε τον εμπειρικό μέσο $\bar{X}_N = 7.2$ των $N = 83$ ανεξάρτητων Τ.Μ. X_1, X_2, \dots, X_N , οι οποίες έχουν άγνωστη μέση τιμή $\mu = E(X_i)$ και τυπική απόκλιση $\sigma = \sqrt{\text{Var}(X_i)}$ μεταξύ $5 \leq \sigma \leq \sigma_{\max} = 7$.

Εδώ εφαρμόζουμε τη γενική μέθοδο της Ενότητας 14.1 στην περίπτωση άγνωστης διασποράς. Έχοντας ήδη την εκτίμηση $\bar{X}_N = 7.2$, για το $c = 90\%$ βρίσκουμε $z^* = 1.645$, και από τη σχέση (14.7) υπολογίζουμε ότι το στατιστικό σφάλμα e είναι $e \leq \sigma_{\max} z^*/\sqrt{N} \approx 1.26$. Συνεπώς, βρίσκουμε ένα 90%-διάστημα εμπιστοσύνης όχι μεγαλύτερο από το $[5.94, 8.46]$.

3. *Ναύτες.* Έστω y_1, y_2, \dots, y_M τα ύψη των $M = 4000$ ναυτών στο στρατόπεδο, και έστω X_i το ύψος του ναύτη i στην τυχαία επιλογή μας, για $i = 1, 2, \dots, N = 108$. Η κάθε X_i έχει μέση τιμή,

$$\mu = E(X_i) = \frac{1}{M} \sum_{j=1}^M y_j,$$

και διασπορά,

$$\sigma^2 = \text{Var}(X_i) = \frac{1}{M} \sum_{j=1}^M (y_j - \mu)^2.$$

Θέλουμε να εκτιμήσουμε τον άγνωστο μέσο μ , ενώ υποθέτουμε ότι η τυπική απόκλιση των X_i είναι μεταξύ $0.03 \leq \sigma \leq \sigma_{\max} = 0.09$.

(α') Χρησιμοποιώντας και πάλι τη γενική μέθοδο της Ενότητας 14.1 στην περίπτωση άγνωστης διασποράς, έχουμε $\bar{X}_N = 1.77$ και για το $c = 98\%$ βρίσκουμε $z^* = 2.326$. Οπότε, από τη σχέση (14.7) υπολογίζουμε ότι το στατιστικό σφάλμα e ικανοποιεί $e \leq \sigma_{\max} z^*/\sqrt{N} \approx 0.02$ μέτρα και, άρα, με επίπεδο εμπιστοσύνης $c = 98\%$ έχουμε διάστημα εμπιστοσύνης όχι μεγαλύτερο από το $[1.75, 1.79]$.

- (β') Το διάστημα εμπιστοσύνης $[\bar{X}_N - e, \bar{X}_N + e]$ βασίζεται στο γεγονός ότι, από το Κ.Ο.Θ., γνωρίζουμε κατά προσέγγιση την κατανομή του εμπειρικού μέσου \bar{X}_N . Βάσει αυτής της προσέγγισης, υπολογίζουμε την πιθανότητα του διαστήματος εμπιστοσύνης:

$$\Pr(\mu \in [\bar{X}_N - e, \bar{X}_N + e]) = \Pr(\bar{X}_N \in [\mu - e, \mu + e]) = 98\%.$$

Αυτή η πιθανότητα προφανώς δεν είναι η ίδια με το ποσοστό των ναυτών που έχουν ύψος εντός του διαστήματος εμπιστοσύνης, το οποίο ποσοστό ισούται με:

$$\Pr(X_i \in [\bar{X}_N - e, \bar{X}_N + e]).$$

Η διαφορά μεταξύ αυτών των δύο πιθανοτήτων μπορεί να παρατηρηθεί πιο ανάγλυφα στο εξής παράδειγμα. Έστω πως 500 ναύτες έχουν ύψος 1.60, 3000 έχουν ύψος 1.77 και 250 έχουν ύψος 1.94. Τότε το μέσο ύψος είναι,

$$\mu = E(X_i) = \frac{1}{4000} [500 \times 1.6 + 3000 \times 1.77 + 500 \times 1.94] = 1.77,$$

και η διασπορά,

$$\begin{aligned} \sigma^2 &= \frac{1}{4000} [500 \times (1.6 - \mu)^2 + 3000 \times (1.77 - \mu)^2 + 500 \times (1.94 - \mu)^2] \\ &\approx 0.007225, \end{aligned}$$

οπότε η τυπική απόκλιση $\sigma \approx 0.085$. Αυτές οι τιμές συμφωνούν με τα δεδομένα του προβλήματος, και η εκτίμηση $\bar{X}_N = 1.77$ είναι επίσης απολύτως λογική. Αλλά παρότι έχουμε το 98%-διάστημα εμπιστοσύνης $[1.75, 1.79]$, βλέπουμε πως μόνο το 75% των ναυτών έχει ύψος εντός αυτού του διαστήματος.

- (γ') Ακολουθώντας την ίδια γενική μέθοδο όπως στο πρώτο σκέλος, για το επίπεδο εμπιστοσύνης $c = 99\%$ έχουμε $z^* = 2.576$, και το επιθυμητό στατιστικό σφάλμα είναι $e = 0.02$. Άρα, το μέγεθος N του στατιστικού δείγματος που απαιτείται ώστε να έχουμε επίπεδο εμπιστοσύνης 99%, από τη σχέση (14.8) είναι $(\sigma_{\max} z^*/e)^2 \approx 134$.
- (δ') Με $N = 108$ δείγματα και στατιστικό σφάλμα $e = 0.01$, το μόνο που μπορούμε να μεταβάλουμε είναι το επίπεδο εμπιστοσύνης c του διαστήματος $[1.76, 1.78]$. Επιστρέφοντας στη βασική σχέση (14.4), έχουμε,

$$\frac{1 - c}{2} = \Phi\left(-\frac{e\sqrt{N}}{\sigma_{\max}}\right) \approx \Phi(-1.54) \approx 0.0618,$$

όπου και πάλι λάβαμε υπόψη μας την χειρότερη δυνατή περίπτωση της τυπικής απόκλισης, και αντικαταστήσαμε την τιμή του $\Phi(-1.54)$ από τον σχετικό πίνακα της Ενότητας 12.4. Λύνοντας ως προς c βρίσκουμε $c = 0.8764$, άρα το καλύτερο αποτέλεσμα που μπορούμε να δώσουμε στον Αρχιπλοίαρχο κ. Τζίφα είναι ότι το διάστημα $[1.76, 1.78]$ έχει επίπεδο εμπιστοσύνης 87.6%.

4. Μια πιο ρεαλιστική δημοσκόπηση. Έστω X_1, X_2, \dots, X_N ανεξάρτητες $\text{Bern}(p)$ Τ.Μ., όπου το μέγεθος του δείγματος $N = 1000$ και θέλουμε να εκτιμήσουμε την παράμετρο $p = E(X_i)$.

(α') Η διασπορά $\sigma^2 = \text{Var}(X_i) = p(1 - p)$ των X_i μάς είναι άγνωστη, αλλά παρατηρούμε ότι είναι μέγιστη στην περίπτωση $p = 1/2$, οπότε έχουμε ότι η τυπική απόκλιση $\sigma \leq \sigma_{\max} = \sqrt{1/4} = 0.5$.

Θα εφαρμόσουμε και εδώ τη γενική μέθοδο της Ενότητας 14.1 στην περίπτωση άγνωστης διασποράς. Έχοντας ήδη την εκτίμηση $\bar{X}_N = 0.305$, για το επίπεδο εμπιστοσύνης $c = 95\%$ βρίσκουμε $z^* = 1.96$, και από τη σχέση (14.7) υπολογίζουμε ότι το στατιστικό σφάλμα $e \leq \sigma_{\max} z^* / \sqrt{N} \approx 0.031 = 3.1\%$. Άρα, έχουμε 95%-διάστημα εμπιστοσύνης όχι μεγαλύτερο από το 27.4% έως 33.6%.

Παρατηρήστε πως το αποτέλεσμα που υπολογίσαμε εδώ είναι ακριβώς ίδιο με αυτό που ανακοίνωσε η εταιρία alco βασισμένη στα ίδια δεδομένα!

(β') Υποθέτοντας τώρα ότι το p είναι μεταξύ 0.2 και 0.35, η μέγιστη τιμή που μπορεί να πάρει η διασπορά $\sigma^2 = p(1 - p)$ αντιστοιχεί στο $p = 0.35$, οπότε έχουμε,

$$\sigma \leq \sigma_{\max} = \sqrt{0.2275} \approx 0.477.$$

Επαναλαμβάνοντας τον υπολογισμό του προηγούμενου σκέλους με τη νέα τιμή σ_{\max} , το στατιστικό σφάλμα είναι $e \leq \sigma_{\max} z^* / \sqrt{N} \approx 0.0296 = 2.96\%$, και συνεπώς μπορούμε να δώσουμε ένα λίγο μικρότερο 95%-διάστημα εμπιστοσύνης, το $[0.2754, 0.3346]$.

5. ★ Εκτίμηση διασποράς.

(α') Εφόσον ορίσαμε το $\hat{\mu} = \bar{X}_N$, χρησιμοποιώντας τις βασικές ιδιότητες της μέσης τιμής (Ιδιότητα 1 του Θεωρήματος 6.1 και Ιδιότητα 1 του Θεωρήματος 11.5) υπολογίζουμε,

$$\begin{aligned} E[(X_1 - \hat{\mu})^2] &= E \left[\left(\frac{1}{N} \sum_{i=1}^N X_1 - \frac{1}{N} \sum_{i=1}^N X_i \right)^2 \right] \\ &= E \left[\left(\frac{1}{N} \sum_{i=2}^N (X_1 - X_i) \right)^2 \right] \\ &= E \left[\left(\frac{1}{N} \sum_{i=2}^N (X_1 - X_i) \right) \left(\frac{1}{N} \sum_{j=2}^N (X_1 - X_j) \right) \right] \\ &= \frac{1}{N^2} E \left[\sum_{i,j=2}^N (X_1 - X_i)(X_1 - X_j) \right], \end{aligned}$$

το οποίο μας δίνει,

$$E[(X_1 - \hat{\mu})^2] = \frac{1}{N^2} \sum_{i,j=2}^N E(X_1^2 - X_1X_i - X_1X_j + X_iX_j). \quad (\text{A'.21})$$

Το παραπάνω διπλό άθροισμα έχει $(N-1)$ όρους, αυτούς που αντιστοιχούν στα ζεύγη $i = j \neq 1$, της μορφής,

$$\begin{aligned} E(X_1^2 - X_1X_i - X_1X_i + X_i^2) &= 2E(X_1^2) - 2E(X_1)E(X_i) \\ &= 2[E(X_1^2) - \mu^2] = 2\sigma^2, \end{aligned} \quad (\text{A'.22})$$

όπου χρησιμοποιήσαμε την ανεξαρτησία των X_i , την Ιδιότητα 3 του Θεωρήματος 6.1, την αντίστοιχη ιδιότητα για συνεχείς Τ.Μ. η οποία θα αποδειχθεί στο Θεώρημα 15.1 του Κεφαλαίου 15, και την εναλλακτική έκφραση για τη διασπορά από τις εκφράσεις (6.8) και (10.13). Εφαρμόζοντας πάλι τις ίδιες ιδιότητες, βρίσκουμε ότι οι υπόλοιποι $(N-1)(N-2)$ όροι του διπλού αθροίσματος στην έκφραση (A'.21), οι οποίοι αντιστοιχούν στα ζεύγη $i \neq j$, είναι της μορφής,

$$E(X_1^2 - X_1X_i - X_1X_j + X_iX_j) = E(X_1^2) - \mu^2 - \mu^2 + \mu^2 = \sigma^2. \quad (\text{A'.23})$$

Οπότε, αντικαθιστώντας τα αποτελέσματα (A'.22) και (A'.23) στη σχέση (A'.21), τελικά βρίσκουμε,

$$E[(X_1 - \hat{\mu})^2] = \frac{1}{N^2}(N-1)(2\sigma^2) + \frac{1}{N^2}(N-1)(N-2)\sigma^2 = \frac{(N-1)}{N}\sigma^2.$$

(β') Εδώ κατ' αρχάς παρατηρούμε πως, λόγω της συμμετρίας του προβλήματος, το αποτέλεσμα του προηγούμενου σκέλους παραμένει το ίδιο με οποιαδήποτε Τ.Μ. X_i στη θέση της X_1 , δηλαδή,

$$E[(X_1 - \hat{\mu})^2] = E[(X_i - \hat{\mu})^2] = \frac{(N-1)\sigma^2}{N},$$

για κάθε $i = 1, 2, \dots, N$. Χρησιμοποιώντας αυτή την έκφραση (και εφαρμόζοντας και πάλι την Ιδιότητα 1 του Θεωρήματος 6.1 και την αντίστοιχη Ιδιότητα 1 του Θεωρήματος 11.5), έχουμε,

$$E(\hat{\sigma}^2) = E\left[\frac{1}{N} \sum_{i=1}^N (X_i - \hat{\mu})^2\right] = \frac{1}{N} \sum_{i=1}^N E(X_i - \hat{\mu})^2 = \left(\frac{N-1}{N}\right) \sigma^2,$$

και παρομοίως,

$$E(\hat{s}^2) = E\left[\frac{1}{N-1} \sum_{i=1}^N (X_i - \hat{\mu})^2\right] = \frac{1}{N-1} \sum_{i=1}^N E(X_i - \hat{\mu})^2 = \sigma^2.$$

(γ') Ο ορισμός της $\hat{\sigma}^2$ είναι εκ πρώτης όψευς πιο φυσικός, υπό την έννοια ότι βασίζεται στον εμπειρικό μέσο όρο των $(X_i - \bar{X}_N)^2$ οπότε αντιστοιχεί ακριβώς στον ορισμό της διασποράς.

Από την άλλη μεριά, η s^2 έχει την επιθυμητή ιδιότητα η μέση τιμή της να είναι ίση με τη ζητούμενη ποσότητα σ^2 . Τέτοιου είδους εκτιμητρίες ονομάζονται «αμερόληπτες» και παίζουν σημαντικό ρόλο στην κλασική στατιστική.

6. *Τυχαίο; Νομίζω.* Εδώ έχουμε ένα παράδειγμα ελέγχου υπόθεσης παραμέτρου Bernoulli όπως στην Ενότητα 14.2.1, με $p^* = 1/6$ που είναι η πιθανότητα να έρθουν διπλές σε μία ζαριά. Ακολουθώντας την εκεί γενική μεθοδολογία, υπολογίζουμε τον εμπειρικό μέσο όρο $\bar{X}_N = 53/275 \approx 0.1927$ των $N = 275$ παρατηρήσεων και, για να είμαστε αρκετά επιεικείς, επιλέγουμε το επίπεδο σημαντικότητας $\alpha = 10\%$. Η τιμή της στατιστικής συνάρτησης εδώ ισούται με $\delta \approx 1.345$, και για την p -τιμή από τους πίνακες του Κεφαλαίου 12 βρίσκουμε ότι ισούται με $2\Phi(-1.16) \approx 0.246$. Εφόσον η p -τιμή είναι μεγαλύτερη του $\alpha = 0.1$, απαντάμε στο φίλο μας ότι, βάσει της επίδοσής του, δεν μπορούμε να απορρίψουμε την υπόθεση ότι η τύχη του είναι ίδια με την τύχη όλων μας, σε επίπεδο σημαντικότητας 10%!

7. *Στυτική δυσλειτουργία.* Το πρόβλημα είναι της ίδιας ακριβώς μορφής με αυτό του Παραδείγματος 14.10. Με $N = 66 + 59 = 125$ για το συνολικό πλήθος των ασθενών, $X_i = 1$ ή 0 ανάλογα με το αν ο ασθενής i αντιμετώπισε ή όχι στυτικά προβλήματα, και $Y_i = 1$ ή 0 αν πήρε το φάρμακο ή το placebo αντίστοιχα, για τις εμπειρικές κατανομές έχουμε,

$$\begin{aligned}
 P_{X,N}(1) &= 1 - P_{X,N}(0) = \frac{36 + 28}{125} = \frac{64}{125}, \\
 P_{Y,N}(1) &= 1 - P_{Y,N}(0) = \frac{66}{125}, \\
 P_{XY,N}(1, 1) &= \frac{36}{125}, \quad P_{XY,N}(0, 1) = \frac{30}{125}, \\
 P_{XY,N}(1, 0) &= \frac{28}{125} \quad \text{και} \quad P_{XY,N}(0, 0) = \frac{31}{125},
 \end{aligned}$$

από τις οποίες βρίσκουμε ότι την τιμή της στατιστικής συνάρτησης ισούται με,

$$\begin{aligned}
 \delta &= 125 \times \left[\frac{\left(\frac{36}{125} - \frac{64}{125} \cdot \frac{66}{125}\right)^2}{\frac{64}{125} \cdot \frac{66}{125}} + \frac{\left(\frac{30}{125} - \frac{61}{125} \cdot \frac{66}{125}\right)^2}{\frac{61}{125} \cdot \frac{66}{125}} \right. \\
 &\quad \left. + \frac{\left(\frac{28}{125} - \frac{64}{125} \cdot \frac{59}{125}\right)^2}{\frac{64}{125} \cdot \frac{59}{125}} + \frac{\left(\frac{31}{125} - \frac{61}{125} \cdot \frac{59}{125}\right)^2}{\frac{61}{125} \cdot \frac{59}{125}} \right],
 \end{aligned}$$

δηλαδή $\delta \approx 0.626$. Άρα η αντίστοιχη p -τιμή είναι $2\Phi(-0.791) = 0.43$, και συμπεραίνουμε ότι δεν μπορούμε να απορρίψουμε τη μηδενική υπόθεση (δηλαδή το ενδεχόμενο τα προβλήματα που παρουσιάστηκαν στους ασθενείς να είναι ανεξάρτητα του φαρμάκου) με επίπεδο εμπιστοσύνης 5%.

8. Γιωργάκη, θα πουντιάσεις!

- (α') Από την περιγραφή του προβλήματος έχουμε πως οι Y_i προήλθαν από μια ακολουθία ανεξάρτητων $\text{Bern}(q)$ Τ.Μ. Επιπλέον θεωρούμε πως, για κάθε i , δεδομένης της τιμής του Y_i , η αντίστοιχη X_i έχει κατανομή με αυθαίρετη πυκνότητα $P_0(x)$ ή $P_1(x)$ όταν Y_i ισούται με 0 ή 1 αντίστοιχα.
- (β') Αν συμβολίσουμε τις τιμές των δύο παραπάνω πυκνοτήτων $P_0(-1) = a$, $P_0(0) = b$, $P_0(1) = 1 - a - b$, και $P_1(-1) = c$, $P_1(0) = d$, $P_1(1) = 1 - c - d$, τότε το μοντέλο έχει παραμέτρους τις τιμές q, a, b, c και d , οι οποίες όλες ανήκουν στο $(0, 1)$ και ικανοποιούν τους περιορισμούς $a + b < 1$ και $c + d < 1$. Η μηδενική υπόθεση H_0 , η οποία περιγράφει την περίπτωση ανεξαρτησίας των X_i από τα Y_i , αντιστοιχεί στη συνθήκη $a = c$ και $b = d$. Για την εναλλακτική υπόθεση H_1 έχουμε ότι οι τιμές a, b, c, d είναι αυθαίρετες.
- (γ') Όπως στην πιο απλή περίπτωση του χ^2 ελέγχου ανεξαρτησίας στην Ενότητα 14.2.2, προκειμένου να απαντήσουμε στο αν οι X_i και οι Y_i είναι ανεξάρτητες, θα συγκρίνουμε την εμπειρική κατανομή των δεδομένων με την κατανομή που θα είχαν αν τα X_i και τα Y_i ήταν ανεξάρτητα. Στη γενικότερη αυτή περίπτωση ορίζουμε τις εμπειρικές πυκνότητες,

$$\begin{aligned}\hat{P}_{XY,N}(x, y) &= \frac{1}{N}[\text{πλήθος από ζεύγη } (x, y) \text{ μεταξύ των } (X_i, Y_i)], \\ \hat{P}_{X,N}(x) &= \frac{1}{N}[\text{πλήθος από } x \text{ μεταξύ των } X_i], \\ \hat{P}_{Y,N}(y) &= \frac{1}{N}[\text{πλήθος από } y \text{ μεταξύ των } Y_i],\end{aligned}$$

για κάθε $x \in \{-1, 0, 1\}$ και $y \in \{0, 1\}$. Και τώρα μπορούμε να ορίσουμε τη στατιστική συνάρτηση Δ_N ως N φορές την χ^2 -απόσταση μεταξύ της από κοινού εμπειρικής πυκνότητας $\hat{P}_{XY,N}(x, y)$ και της από κοινού πυκνότητας $\hat{P}_{X,N}(x)\hat{P}_{Y,N}(y)$ που θα είχαν τα (X_i, Y_i) αν ήταν ανεξάρτητα:

$$\begin{aligned}\Delta_N &= Nd_{\chi^2}(\hat{P}_{XY,N}, \hat{P}_{X,N}\hat{P}_{Y,N}) \\ &= N \sum_{x=-1,0,1, y=0,1} \frac{\left(\hat{P}_{XY,N}(x, y) - \hat{P}_{X,N}(x)\hat{P}_{Y,N}(y)\right)^2}{\hat{P}_{X,N}(x)\hat{P}_{Y,N}(y)}.\end{aligned}$$

- (δ') Μας δίνεται πως, κάτω από τη μηδενική υπόθεση H_0 , η στατιστική συνάρτηση, Δ_N συγκλίνει κατά κατανομή στη $\chi^2(2)$, καθώς $N \rightarrow \infty$. Επιπλέον, από το αποτέλεσμα της Άσκησης 2 του Κεφαλαίου 15 έχουμε ότι η κατανομή $\chi^2(2)$ είναι η $\text{Exp}(1)$, οπότε, καθώς $N \rightarrow \infty$,

$$\text{Pr}(\Delta_N \leq \delta) \rightarrow 1 - e^{-\delta}, \quad \text{για κάθε } \delta \geq 0.$$

Βάσει αυτού, προτείνουμε την εξής μέθοδο χ^2 -ελέγχου ανεξαρτησίας για το εδώ σενάριο:

- i. Από τα δεδομένα $(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)$, υπολογίζουμε τις εμπειρικές πυκνότητες $\hat{P}_{X,N}, \hat{P}_{Y,N}$ και $\hat{P}_{XY,N}$, και την τιμή δ της στατιστικής συνάρτησης όπως ορίστηκαν παραπάνω.
- ii. Επιλέγουμε το επιθυμητό **επίπεδο σημαντικότητας** $\alpha \in (0, 1)$.
- iii. Υπολογίζουμε την p -τιμή, δηλαδή την πιθανότητα, κάτω από τη μηδενική υπόθεση H_0 , η Δ_N να είναι τόσο ή μεγαλύτερη από την τιμή δ που υπολογίσαμε. Από το παραπάνω αποτέλεσμα:

$$p\text{-τιμή} = \Pr(\Delta_N \geq \delta | H_0) \approx e^{-\delta}.$$

- iv. Αν η p -τιμή $\leq \alpha$, η H_0 απορρίπτεται σε επίπεδο σημαντικότητας α .
 Αν η p -τιμή $> \alpha$, η H_0 δεν απορρίπτεται σε επίπεδο σημαντικότητας α .
- (ε') Από τον πίνακα της εκφώνησης εύκολα υπολογίζουμε τις τιμές της από κοινού εμπειρικής πυκνότητας:

ΑΠΟ ΚΟΙΝΟΥ ΕΜΠΕΙΡΙΚΗ ΚΑΤΑΝΟΜΗ $\hat{P}_{XY,N}$	$x = -1$	0	$+1$
$y = 1$	6/173	65/173	8/173
$y = 0$	13/173	71/173	10/173

Και για τις εμπειρικές κατανομές των X_i και Y_i αντίστοιχα έχουμε:

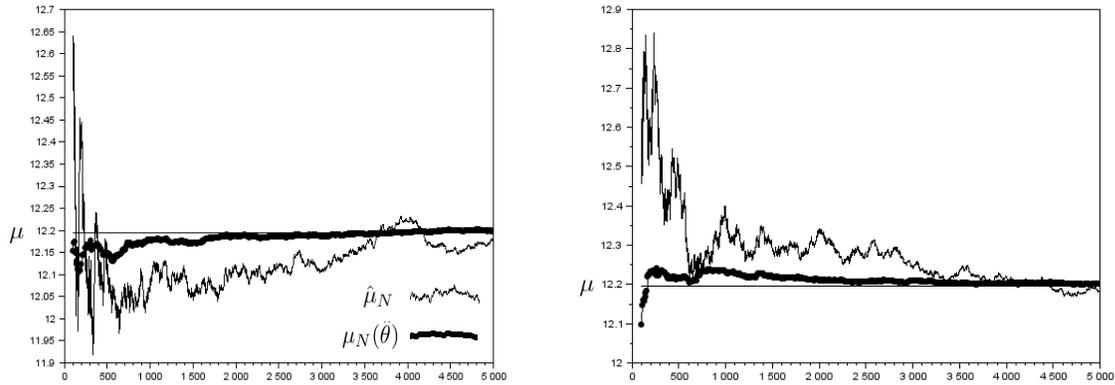
$$\begin{aligned} \hat{P}_{X,N}(-1) &= \frac{19}{173}, & \hat{P}_{X,N}(0) &= \frac{136}{173}, & \hat{P}_{X,N}(1) &= \frac{18}{173}, \\ \hat{P}_{Y,N}(0) &= \frac{94}{173}, & \hat{P}_{Y,N}(1) &= \frac{79}{173}. \end{aligned}$$

Αντικαθιστώντας αυτές τις τιμές στον ορισμό της Δ_N , η τιμή της στατιστικής συνάρτησης εδώ είναι,

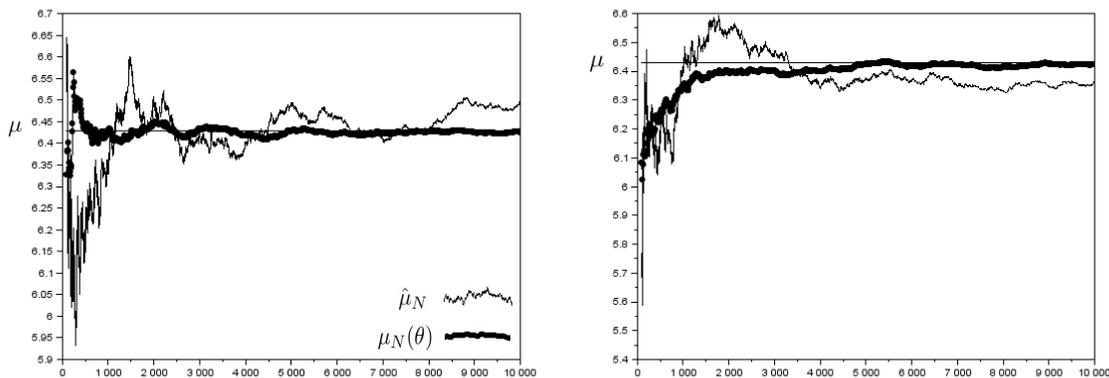
$$\delta = 173 \times \left[\frac{\left(\frac{6}{173} - \frac{19}{173} \cdot \frac{79}{173}\right)^2}{\frac{19}{173} \cdot \frac{79}{173}} + \frac{\left(\frac{65}{173} - \frac{136}{173} \cdot \frac{79}{173}\right)^2}{\frac{136}{173} \cdot \frac{79}{173}} + \frac{\left(\frac{8}{173} - \frac{18}{173} \cdot \frac{79}{173}\right)^2}{\frac{18}{173} \cdot \frac{79}{173}} + \frac{\left(\frac{13}{173} - \frac{19}{173} \cdot \frac{94}{173}\right)^2}{\frac{19}{173} \cdot \frac{94}{173}} + \frac{\left(\frac{71}{173} - \frac{136}{173} \cdot \frac{94}{173}\right)^2}{\frac{136}{173} \cdot \frac{94}{173}} + \frac{\left(\frac{10}{173} - \frac{18}{173} \cdot \frac{94}{173}\right)^2}{\frac{18}{173} \cdot \frac{94}{173}} \right],$$

που μας δίνει $\delta \approx 0.0103$, οπότε για την p -τιμή βρίσκουμε $e^{-\delta} \approx 0.9898$. Για ακόμα μια φορά, λοιπόν, η Ελληνίδα μάνα διαφεύδεται!

9. *Μεταβλητές ελέγχου.* Η άσκηση είναι καθαρά προγραμματιστική. Στα Σχήματα **A'.37** και **A'.38** δίνουμε δύο ακόμα αποτελέσματα προσομοίωσης για καθένα από τα δύο προβλήματα.



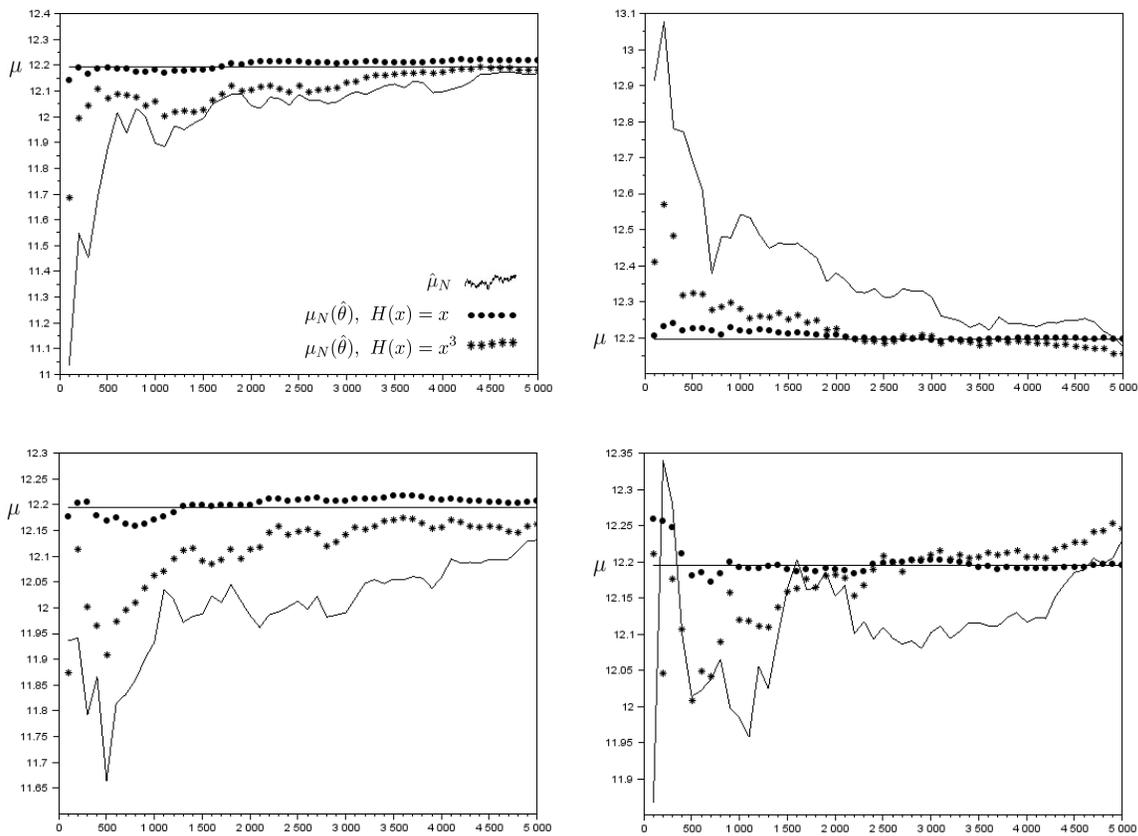
Σχήμα A'.37: Οι τιμές των εκτιμητριών $\mu_N(\hat{\theta})$ (έντονη γραμμή) και $\hat{\mu}_N$ (απλή γραμμή), σε δύο ακόμα πειράματα προσομοίωσης για τον υπολογισμό ενός ολοκληρώματος όπως στο Παράδειγμα 14.12. Παρατηρούμε ότι η προσαρμοστική εκτιμήτρια $\mu_N(\hat{\theta})$ είναι σημαντικά αποτελεσματικότερη από τον εμπειρικό μέσο $\hat{\mu}_N$.



Σχήμα A'.38: Οι τιμές των εκτιμητριών $\mu_N(\hat{\theta})$ (έντονη γραμμή) και $\hat{\mu}_N$ (απλή γραμμή), σε δύο ακόμα πειράματα προσομοίωσης για το πρόβλημα τιμολόγησης ενός παραγώγου όπως στο Παράδειγμα 14.13. Και εδώ βλέπουμε ότι η προσαρμοστική εκτιμήτρια $\mu_N(\hat{\theta})$ είναι σημαντικά αποτελεσματικότερη από τον εμπειρικό μέσο $\hat{\mu}_N$.

10. Διαφορετική μεταβλητή ελέγχου. Η μεθοδολογία που ακολουθούμε είναι ακριβώς ίδια με του Παραδείγματος 14.12, όπου χρησιμοποιούμε τους γενικούς τύπους (14.14), (14.15) και (14.16) για τον ορισμό της προσαρμοστικής εκτιμήτριας $\mu_N(\hat{\theta})$, με τη μόνη διαφορά ότι τώρα έχουμε $H(x) = x^3$ και $\nu = E[H(X)] = E(X^3) = 0$ (γιατί;).

Στο Σχήμα A.39 δίνονται τα αποτελέσματα τεσσάρων πειραμάτων για τις τρεις εκτιμήτριες $\hat{\mu}_N$, $\mu_N(\hat{\theta})$ με $H(x) = x$, και $\mu_N(\hat{\theta})$ με $H(x) = x^3$. Είναι εμφανές ότι η προσαρμοστική εκτιμήτρια με την μεταβλητή ελέγχου $H(X) = X^3$ είναι πιο αποτελεσματική από τον εμπειρικό μέσο, αλλά δεν συγκλίνει τόσο γρήγορα όσο στην περίπτωση της $H(X) = X$.



Σχήμα A.39: Αποτελέσματα πειραμάτων προσομοίωσης του προβλήματος στο Παράδειγμα 14.12 με $N = 5000$ δείγματα. Η τιμές του εμπειρικού μέσου εμφανίζονται ως μια απλή γραμμή, της προσαρμοστικής εκτιμήτριας $\mu_N(\hat{\theta})$ με μεταβλητή ελέγχου την $H(X) = X$ αναπαρίστανται ως έντονες κουκκίδες, και της $\mu_N(\hat{\theta})$ με μεταβλητή ελέγχου την $H(X) = X^3$ ως *. Για λόγους ευκρίνειας, οι τιμές των δύο τελευταίων εκτιμητριών σχεδιάστηκαν μόνο κάθε 100 βήματα.

11. Προεξοφλημένη τιμή μετοχής. Για τη ζητούμενη μέση τιμή, από τη σχέση (14.18) έχουμε,

$$\begin{aligned} E[e^{-rT} X(T)] &= E \left[e^{-rT} X(0) \exp \left\{ \left(r - \frac{1}{2} \sigma^2 \right) T + \sigma \sqrt{T} Z \right\} \right] \\ &= X(0) e^{-\sigma^2 T/2} E \left[e^{\sigma \sqrt{T} Z} \right], \end{aligned} \quad (\text{A'.24})$$

όπου η Z έχει τυπική κανονική κατανομή με πυκνότητα $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$. Οπότε η τελευταία μέση τιμή παραπάνω ισούται με,

$$\begin{aligned} \int_{-\infty}^{\infty} e^{\sigma \sqrt{T} z} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} [z^2 - 2\sigma \sqrt{T} z] \right\} dz \\ &= e^{\sigma^2 T/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} [(z - \sigma \sqrt{T})^2] \right\} dz, \end{aligned}$$

όπου αναγνωρίζουμε το τελευταίο παραπάνω ολοκλήρωμα ως το ολοκλήρωμα της πυκνότητας της $N(\sigma \sqrt{T}, 1)$ κατανομής σε όλο το \mathbb{R} , το οποίο φυσικά ισούται με 1. Συνεπώς,

$$E \left[e^{\sigma \sqrt{T} Z} \right] = e^{\sigma^2 T/2},$$

και αντικαθιστώντας αυτή την έκφραση στη (A'.24) προκύπτει ακριβώς το ζητούμενο αποτέλεσμα.

A'.15 Ασκήσεις Κεφαλαίου 15

1. Παραδείγματα ανεξαρτησίας.

(α') Οι Τ.Μ. X, Y του Παραδείγματος 15.1 έχουν από κοινού πυκνότητα $f_{XY}(x, y) = 1$ για $x, y \in [0, 1]$ και $f_{XY}(x, y) = 0$ για $x, y \notin [0, 1]$. Από την τέταρτη βασική ιδιότητα της από κοινού πυκνότητας εύκολα υπολογίζουμε πως η περιθώρια πυκνότητα του X είναι,

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy = \int_0^1 f_{XY}(x, y) dy = \begin{cases} 1, & \text{για } x \in [0, 1], \\ 0, & \text{για } x \notin [0, 1], \end{cases}$$

και παρομοίως ότι η $f_Y(y)$ είναι της ίδιας ακριβώς μορφής, δηλαδή πως $f_Y(y) = 1$ όταν $y \in [0, 1]$ και $f_Y(y) = 0$ όταν όχι. Είναι προφανές από τις παραπάνω εκφράσεις ότι έχουμε $f_{XY}(x, y) = f_X(x)f_Y(y)$ για κάθε $x, y \in \mathbb{R}$, άρα από το κριτήριο (15.7) προκύπτει ότι οι X και Y είναι ανεξάρτητες Τ.Μ. Επιπλέον παρατηρούμε ότι και οι δύο έχουν κατανομή $U[0, 1]$.

(β') Εδώ μπορούμε εύκολα να δείξουμε ότι οι Τ.Μ. X, Y του Παραδείγματος 15.2 δεν είναι ανεξάρτητες, χωρίς να χρειαστεί να κάνουμε λεπτομερείς υπολογισμούς. Από την τέταρτη ιδιότητα της από κοινού πυκνότητας και τον ορισμό της f_{XY} , αμέσως βλέπουμε πως οι περιθώριες πυκνότητες $f_X(x)$ και $f_Y(y)$ παίρνουν αυστηρά θετικές τιμές για όλα τα $x, y \in [-1, 1]$, οπότε και οι πιθανότητες,

$$\Pr(4/5 \leq X \leq 1) > 0 \quad \text{και} \quad \Pr(4/5 \leq Y \leq 1) > 0,$$

είναι θετικές. Αλλά το ορθογώνιο $R = [4/5, 1] \times [4/5, 1]$ είναι ξένο ως προς το σύνολο τιμών S του ζευγαριού (X, Y) , άρα η πιθανότητα,

$$\Pr((X, Y) \in R) = \Pr(4/5 \leq X \leq 1, 4/5 \leq Y \leq 1) = 0,$$

και συνεπώς δεν ισούται με το γινόμενο των $\Pr(4/5 \leq X \leq 1)$ και $\Pr(4/5 \leq Y \leq 1)$. Οπότε από τον ορισμό (15.6), οι X, Y δεν είναι ανεξάρτητες.

2. Περισσότερα για την κατανομή $\chi^2(k)$. Όπως στην Άσκηση 5 του Κεφαλαίου 12, για τις ανεξάρτητες Τ.Μ. Z_1, Z_2, \dots, Z_k όπου κάθε $Z_i \sim N(0, 1)$, ορίζουμε την τυχαία μεταβλητή $Y = Z_1^2 + Z_2^2 + \dots + Z_k^2 \sim \chi^2(k)$. Επίσης θυμίζουμε πως, εκεί, δείξαμε ότι η μέση τιμή της κατανομής $\chi^2(k)$ είναι $E(Y) = k$, και ότι $E(Z_i^4) = 3$ για κάθε i .

(α') Για τη διασπορά της Y πρώτα υπολογίζουμε,

$$E(Y^2) = E \left[\left(\sum_{i=1}^k Z_i^2 \right)^2 \right] = E \left[\left(\sum_{i=1}^k Z_i^2 \right) \left(\sum_{j=1}^k Z_j^2 \right) \right] = E \left[\sum_{i=1}^k \sum_{j=1}^k Z_i^2 Z_j^2 \right],$$

και από το Πόρισμα 15.1,

$$E(Y^2) = \sum_{i=1}^k \sum_{j=1}^k E[Z_i^2 Z_j^2].$$

Το παραπάνω διπλό άθροισμα έχει k όρους της μορφής $E(Z_i^4) = 3$ που αντιστοιχούν στις τιμές $i = j$, και $k(k-1)$ όρους της μορφής $E(Z_i^2 Z_j^2)$ για $i \neq j$, όπου, λόγω της ανεξαρτησίας των Z_i^2 , από το Θεώρημα 15.1 έχουμε $E(Z_i^2 Z_j^2) = E(Z_i^2)E(Z_j^2) = 1$. Οπότε η $E(Y^2) = 3k + k(k-1)$ και η ζητούμενη διασπορά είναι,

$$\text{Var}(Y) = E(Y^2) - [E(Y)]^2 = 3k + k(k-1) - k^2 = 2k.$$

(β') Για $k = 1$, η πυκνότητα της κατανομής $\chi^2(1)$ είναι η πυκνότητα $f_1(y)$ της Γ.Μ. $Y = Z_1^2$, για την οποία, από το αποτέλεσμα της Άσκησης 4 του Κεφαλαίου 12, έχουμε,

$$f_1(y) = \begin{cases} \frac{1}{\sqrt{2\pi y}} \exp\left(-\frac{y}{2}\right), & y > 0, \\ 0, & y \leq 0. \end{cases}$$

Για $k = 2$, η πυκνότητα $f_2(y)$ της $\chi^2(2)$ είναι η πυκνότητα της $Y = Z_1^2 + Z_2^2$, όπου οι Z_1^2, Z_2^2 είναι ανεξάρτητες (από την τέταρτη ιδιότητα του Θεωρήματος 15.1) και έχουν κοινή πυκνότητα την $f_1(y)$. Άρα, για $y \leq 0$ η $f_2(y) = 0$, ενώ για $y > 0$ σύμφωνα με το Θεώρημα 15.2 η $f_2(y)$ δίνεται από τη συνέλιξη,

$$\begin{aligned} f_2(y) &= \int_{-\infty}^{\infty} f_1(x)f_1(y-x) dx \\ &= \int_0^y \frac{1}{\sqrt{2\pi x}} \exp\left(-\frac{x}{2}\right) \frac{1}{\sqrt{2\pi(y-x)}} \exp\left(-\frac{y-x}{2}\right) dx \\ &= \frac{1}{2\pi} e^{-y/2} \int_0^y \frac{1}{\sqrt{x(y-x)}} dx. \end{aligned}$$

Και κάνοντας στο τελευταίο παραπάνω ολοκλήρωμα διαδοχικά τις αντικαταστάσεις $z = x/y$, $w = \sqrt{z}$, και $\theta = \arccos w$, βρίσκουμε πως,

$$\begin{aligned} f_2(y) &= \frac{1}{2\pi} e^{-y/2} \int_0^1 \frac{1}{\sqrt{z(1-z)}} dz \\ &= \frac{1}{\pi} e^{-y/2} \int_0^1 \frac{1}{\sqrt{1-w^2}} dw \\ &= \frac{1}{\pi} e^{-y/2} \int_0^{\pi/2} d\theta = \frac{1}{2} e^{-y/2}, \end{aligned}$$

την οποία αναγνωρίζουμε ως την πυκνότητα της $\text{Exp}(2)$ κατανομής! Άρα η $\chi^2(2)$ είναι η ίδια κατανομή με την $\text{Exp}(2)$.

Παρομοίως, για $k = 3$, η πυκνότητα $f_3(y)$ της $\chi^2(3)$ είναι η πυκνότητα της T.M. $Y = (Z_1^2 + Z_2^2) + Z_3^2$, όπου η $(Z_1^2 + Z_2^2)$ έχει πυκνότητα $f_2(y)$ και είναι ανεξάρτητη από τη Z_3^2 η οποία έχει πυκνότητα $f_1(y)$. Άρα, για $y \leq 0$ η $f_3(y) = 0$, και για $y > 0$ σύμφωνα με το Θεώρημα 15.2 η $f_3(y)$ μπορεί να υπολογιστεί ως η συνέλιξη,

$$\begin{aligned} f_3(y) &= \int_{-\infty}^{\infty} f_2(x)f_1(y-x) dx \\ &= \int_0^y \frac{1}{2} \exp\left(-\frac{x}{2}\right) \frac{1}{\sqrt{2\pi(y-x)}} \exp\left(-\frac{y-x}{2}\right) dx \\ &= \frac{1}{2\sqrt{2\pi}} e^{-y/2} \int_0^y \frac{1}{\sqrt{y-x}} dx. \end{aligned}$$

Κάνοντας τις ίδιες διαδοχικές αντικαταστάσεις όπως για την f_2 , έχουμε,

$$\begin{aligned} f_3(y) &= \frac{1}{2\sqrt{2\pi}} \sqrt{y} e^{-y/2} \int_0^1 \frac{1}{\sqrt{1-z}} dz \\ &= \frac{1}{\sqrt{2\pi}} \sqrt{y} e^{-y/2} \int_0^1 \frac{w}{\sqrt{1-w^2}} dw \\ &= \frac{1}{\sqrt{2\pi}} \sqrt{y} e^{-y/2} \int_0^{\pi/2} \cos \theta d\theta \\ &= \frac{1}{\sqrt{2\pi}} \sqrt{y} e^{-y/2} [\sin \theta]_0^{\pi/2}, \end{aligned}$$

οπότε, τελικά βρίσκουμε,

$$f_3(y) = \begin{cases} \frac{1}{\sqrt{2\pi}} \sqrt{y} e^{-y/2}, & y > 0, \\ 0, & y \leq 0. \end{cases}$$

3. Ασυσχετίστες αλλά όχι ανεξάρτητες συνεχείς T.M.

(α') Από την τέταρτη ιδιότητα της από κοινού πυκνότητας, η περιθώρια πυκνότητα της X είναι $f_X(x) = 0$ για $x \notin [-1, 1]$, και για $x \in [-1, 1]$,

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi} dy = \frac{2\sqrt{1-x^2}}{\pi}.$$

Λόγω συμμετρίας, η περιθώρια πυκνότητα του Y είναι ακριβώς ίδια, $f_Y(y) = f_X(y)$, για όλα τα $y \in \mathbb{R}$.

(β') Οι συναρτήσεις $f_X(x)$ και $f_Y(y)$ είναι άρτιες, οπότε οι αντίστοιχες συναρτήσεις $xf_X(x)$ και $yf_Y(y)$ είναι περιττές, άρα τα ολοκληρώματά τους σε όλο το \mathbb{R} , τα οποία είναι ίσα

με τις μέσες τιμές τους $E(X)$ και $E(Y)$, είναι ίσα με μηδέν. Επιπλέον, η μέση τιμή $E(XY)$ είναι και αυτή ίση με μηδέν όπως εύκολα διαπιστώνουμε λόγω της συμμετρίας του προβλήματος ή υπολογίζοντας αναλυτικά το διπλό ολοκλήρωμα,

$$E(XY) = \iint_S xy f_{XY}(x, y) dx dy = \frac{1}{\pi} \int_{-1}^1 x \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} y dy dx,$$

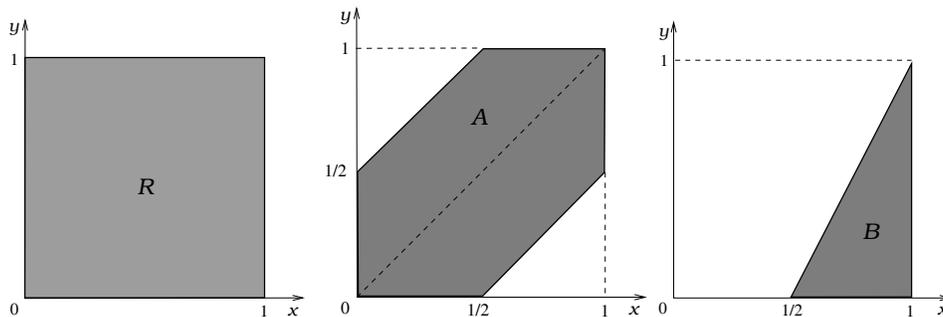
το οποίο ισούται με μηδέν αφού στο εσωτερικό ολοκλήρωμα (ως προς y) ολοκληρώνουμε μια περιττή συνάρτηση. Συνεπώς η συνδιακύμανση,

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0.$$

Το γεγονός ότι οι X, Y δεν είναι ανεξάρτητες προκύπτει με το ίδιο ακριβώς επιχείρημα όπως στο δεύτερο σκέλος της Άσκησης 1.

4. *Χρόνοι διεργασιών.* Εφόσον οι Τ.Μ. X, Y είναι ανεξάρτητες με κοινή κατανομή $U[0, 1]$, σύμφωνα με το κριτήριο (15.7) μπορούμε να θέσουμε για την από κοινού πυκνότητά τους $f_{XY}(x, y) = f_X(x)f_Y(y) = 1$ για $(x, y) \in R = [0, 1]^2$, και $f_{XY}(x, y) = 0$ αλλού. [Παρατηρήστε ότι η από κοινού κατανομή των (X, Y) είναι ίδια με αυτή του Παραδείγματος 15.1 και της Άσκησης 1.]

- (α) Η πιθανότητα $\Pr(|X - Y| < 1/2)$ ισούται με το ολοκλήρωμα της από κοινού πυκνότητας στο σύνολο $A = \{(x, y) : |x - y| < 1/2, 0 \leq x, y \leq 1\}$, δηλαδή με τον όγκο του στερεού που περικλείεται ανάμεσα στο A και το γράφημα της f_{XY} . Αλλά αφού η πυκνότητα ισούται με 1 σε όλα τα σημεία του A , η ζητούμενη πιθανότητα είναι ίση με το εμβαδόν του A το οποίο, όπως βλέπουμε και από το Σχήμα Α΄.40, ισούται με $3/4$. Άρα, $\Pr(|X - Y| < 1/2) = 3/4$.



Σχήμα Α΄.40: Άσκηση 4: Η από κοινού πυκνότητα είναι ίση με τη μονάδα εντός του τετραγώνου $R = \{(x, y) : 0 \leq x, y \leq 1\}$, και μηδενική εκτός.

(β') Η ζητούμενη μέση τιμή μπορεί να εκφραστεί ως το διπλό ολοκλήρωμα,

$$\begin{aligned} E(|X - Y|) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |x - y| f_{XY}(x, y) dx dy \\ &= \int_0^1 \int_0^1 |x - y| dx dy \\ &= \int_0^1 \int_0^x (x - y) dy dx + \int_0^1 \int_x^1 (y - x) dy dx, \end{aligned}$$

όπου το πρώτο πιο πάνω ολοκλήρωμα υπολογίζεται απλά ως,

$$\int_0^1 \int_0^x (x - y) dy dx = \int_0^1 \left[xy - \frac{y^2}{2} \right]_0^x dx = \int_0^1 \frac{x^2}{2} dx = \left[\frac{x^3}{6} \right]_0^1 = \frac{1}{6}.$$

Το δεύτερο ολοκλήρωμα μπορεί να υπολογιστεί με τον ίδιο τρόπο ή, εναλλακτικά, παρατηρώντας πως λόγω συμμετρίας πρέπει να ισούται με το πρώτο. Άρα, τελικά,

$$E(|X - Y|) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}.$$

(γ') Η μέση τιμή του κόστους είναι εύκολο να υπολογιστεί, βάσει της Πρότασης 15.1, και των ιδιοτήτων της ομοιόμορφης κατανομής:

$$E(Z) = E(20X + 10Y) = 20E(X) + 10E(Y) = 20 \times \frac{1}{2} + 10 \times \frac{1}{2} = 15.$$

(δ') Για τη ζητούμενη πιθανότητα κατ' αρχάς παρατηρούμε πως,

$$\begin{aligned} \Pr(Z > 20) &= \Pr(20X + 10Y > 20) \\ &= \Pr(2X + Y > 2) \\ &= \iint_B f_{XY}(x, y) dx dy, \end{aligned}$$

όπου το σύνολο $B \subset \mathbb{R}^2$ ορίζεται ως,

$$B = \{(x, y) : 2x + y > 2, 0 \leq x, y \leq 1\},$$

και έχει σχεδιαστεί στο Σχήμα A.40. Με το ίδιο σκεπτικό όπως στο πρώτο σκέλος της άσκησης, παρατηρούμε πως η ζητούμενη πιθανότητα ισούται με το εμβαδόν του B το οποίο, όπως βλέπουμε και από το Σχήμα A.40, ισούται με $1/4$, συνεπώς έχουμε, $\Pr(Z > 20) = 1/4$.

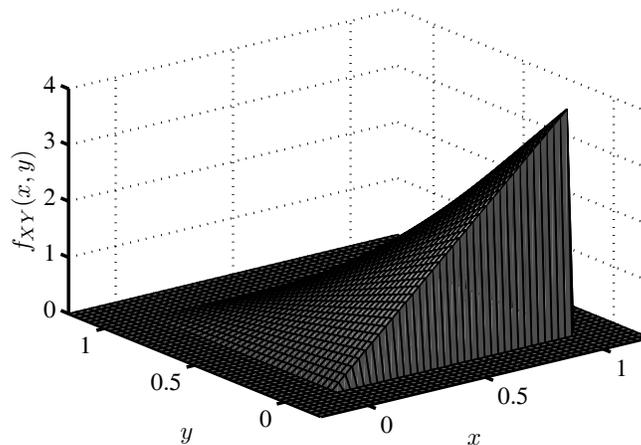
(ε') Αφού οι X, Y είναι ανεξάρτητες, από το Θεώρημα 15.1 γνωρίζουμε πως η συνδιακύμανση $\text{Cov}(X, Y) = 0$.

5. Θεραπείες.

(α') Από την πρώτη ιδιότητα της από κοινού πυκνότητας γνωρίζουμε ότι το ολοκλήρωμα της f_{XY} σε όλο το \mathbb{R}^2 πρέπει να ισούται με τη μονάδα:

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y) dx dy &= c \int_0^1 \int_0^1 x(1-y) dx dy \\ &= c \left(\int_0^1 x dx \right) \left(\int_0^1 (1-y) dy \right) \\ &= c \times \frac{1}{2} \times \frac{1}{2}, \end{aligned}$$

συνεπώς πρέπει να έχουμε $c = 4$. Η πυκνότητα έχει σχεδιαστεί στο Σχήμα Α΄.41.



Σχήμα Α΄.41: Η πυκνότητα $f_{XY}(x, y)$ της Άσκησης 5.

(β') Θα χρησιμοποιήσουμε την τέταρτη ιδιότητα της από κοινού πυκνότητας. Προφανώς για $x > 1$ ή $x < 0$, η περιθώρια πυκνότητα $f_X(x) = 0$, ενώ για $0 \leq x \leq 1$, έχουμε:

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy = \int_0^1 4x(1-y) dy = 4x \int_0^1 (1-y) dy = 2x.$$

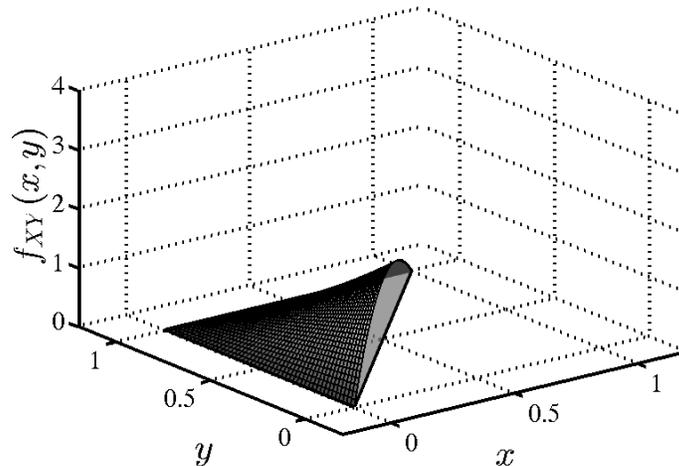
Παρομοίως, η περιθώρια πυκνότητα της Y είναι $f_Y(y) = 0$ όταν $y \notin [0, 1]$, ενώ για $y \in [0, 1]$:

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y) dx = \int_0^1 4x(1-y) dx = (1-y) \int_0^1 4x dx = 2(1-y).$$

- (γ') Για κάθε ζεύγος $(x, y) \in \mathbb{R} \times \mathbb{R}$, εύκολα διαπιστώνουμε ότι $f_{XY}(x, y) = f_X(x)f_Y(y)$, άρα βάσει του κριτηρίου (15.7) οι X, Y είναι ανεξάρτητες. Συνεπώς, από το Θεώρημα 15.1, η συνδιακύμανση $\text{Cov}(X, Y) = 0$.
- (δ') Τέλος, υπολογίζουμε τη ζητούμενη πιθανότητα:

$$\begin{aligned} \Pr(X \leq Y) &= \iint_{\{(x,y) : x \leq y\}} f_{XY}(x, y) \, dx \, dy \\ &= \int_0^1 \int_0^y 4x(1-y) \, dx \, dy \\ &= \int_0^1 4(1-y) \left[\frac{x^2}{2} \right]_0^y \, dy \\ &= \int_0^1 4(1-y) \frac{y^2}{2} \, dy \\ &= \left[\frac{4y^3}{6} - \frac{y^4}{8} \right]_0^1 \\ &= \frac{1}{6}. \end{aligned}$$

Στο Σχήμα A.42 έχουμε σχεδιάσει το στερεό του οποίου ο όγκος αντιστοιχεί στην παραπάνω πιθανότητα.



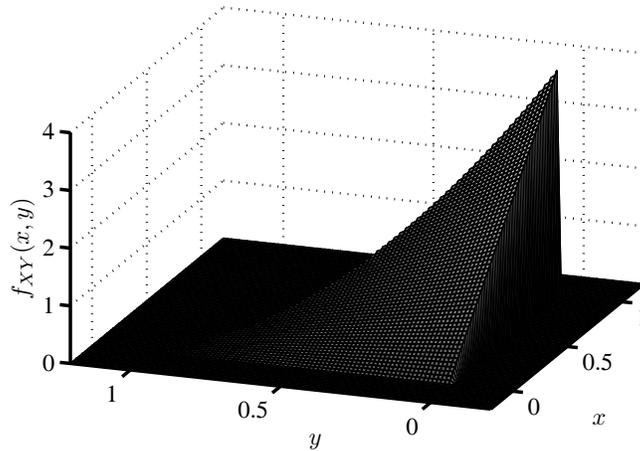
Σχήμα A.42: Ο όγκος του στερεού ισούται με την $\Pr(X \leq Y)$ της Άσκησης 5.

6. Εναλλακτικές θεραπείες.

(α') Και πάλι, πρέπει το ολοκλήρωμα της $f_{XY}(x, y)$ στο $\mathbb{R} \times \mathbb{R}$ να ισούται με τη μονάδα:

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y) dx dy &= \int_0^1 \int_0^{1-y} cx(1-y) dx dy \\ &= \int_0^1 c[x]_0^{1-y} (1-y) dx dy \\ &= \int_0^1 \frac{c}{2} (1-y)^3 dy \\ &= \left[\frac{c}{8} (1-y)^4 \right]_1^0 dy = \frac{c}{8}. \end{aligned}$$

Άρα, $c = 8$. Η πυκνότητα έχει σχεδιαστεί στο Σχήμα Α΄.43.



Σχήμα Α΄.43: Η πυκνότητα της Άσκησης 6.

(β') Από την τέταρτη ιδιότητα της από κοινού πυκνότητας, όταν $x \notin [0, 1]$ η $f_X(x) = 0$, ενώ για $x \in [0, 1]$,

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy = \int_0^{1-x} 8x(1-y) dy = \left[-8x \frac{(1-y)^2}{2} \right]_0^{1-x} = 4x(1-x^2).$$

Παρομοίως, αν $y \notin [0, 1]$ τότε $f_Y(y) = 0$, ενώ αν $y \in [0, 1]$,

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y) dx = \int_0^{1-y} 8x(1-y) dx = \left[4(1-y)x^2 \right]_0^{1-y} = 4(1-y)^3.$$

Οπότε, συγκεντρωτικά,

$$f_X(x) = \begin{cases} 4x(1-x^2), & x \in [0, 1], \\ 0, & x \notin [0, 1], \end{cases}$$

$$\text{και } f_Y(y) = \begin{cases} 4(1-y)^3, & y \in [0, 1], \\ 0, & y \notin [0, 1]. \end{cases}$$

(γ') Για τη συνδιακύμανση, υπολογίζουμε τη μέση τιμή των X, Y ,

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 4x^2(1-x^2) dx = 4 \left[\frac{1}{3}x^3 - \frac{1}{5}x^5 \right]_0^1 = \frac{8}{15},$$

$$E(Y) = \int_{-\infty}^{\infty} y f_Y(y) dy = 4 \int_0^1 y(1-y)^3 dy = 4 \int_0^1 (1-z)z^3 dz = \left[z^4 - \frac{4z^5}{5} \right]_0^1 = \frac{1}{5},$$

όπου για τη μέση τιμή του Y πραγματοποιήσαμε την αντικατάσταση $z = 1 - y$, και τη μέση τιμή του γινομένου τους,

$$\begin{aligned} E(XY) &= \int_{-\infty}^{\infty} xy f_{XY}(x, y) dx dy \\ &= \int_0^1 \int_0^{1-x} 8x^2(1-y)y dy dx \\ &= \int_0^1 8x^2 \left[\frac{y^2}{2} - \frac{y^3}{3} \right]_0^{1-x} dx \\ &= \int_0^1 8x^2 \left(\frac{1}{2}(1-x)^2 - \frac{1}{3}(1-x)^3 \right) dx \\ &= \int_0^1 \left(\frac{8}{3}x^5 - 4x^4 + \frac{4}{3}x^2 \right) dx \\ &= \left[\frac{4}{9}x^6 - \frac{4}{5}x^5 + \frac{4}{9}x^3 \right]_0^1 \\ &= \frac{4}{45}. \end{aligned}$$

Οπότε τελικά,

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = \frac{4}{45} - \frac{8}{15} \times \frac{1}{5} = -\frac{4}{225}.$$

Και εφόσον η συνδιακύμανση $\text{Cov}(X, Y)$ είναι μη μηδενική, από το Θεώρημα 15.1 έχουμε ότι οι X, Y δεν είναι ανεξάρτητες.

7. *Τριπλά ολοκληρώματα.* Καταρχάς ορίζουμε το σύνολο τιμών των (X, Y, Z) ως το σύνολο $S = [0, 3] \times [0, 2] \times [0, 1] \subset \mathbb{R}^3$ όπου η από κοινού πυκνότητα παίρνει θετικές τιμές.

(α') Θα υπολογίσουμε τη σταθερά k απαιτώντας το ολοκλήρωμα της $f_{XYZ}(x, y, z)$ σε όλο το \mathbb{R}^3 να είναι ίσο με 1. Γι' αυτό τον υπολογισμό (και όλους τους παρακάτω σε αυτήν την άσκηση), χρησιμοποιούμε την τρισδιάστατη εκδοχή του θεωρήματος του Fubini:

$$\begin{aligned} 1 &= \iiint_{\mathbb{R}^3} f_{XYZ}(x, y, z) dx dy dz \\ &= \iiint_S \frac{1}{k}(2x + 3y^2 + 4z^3) dx dy dz \\ &= \int_0^1 \int_0^2 \int_0^3 \frac{2x}{k} dx dy dz + \int_0^1 \int_0^3 \int_0^2 \frac{3y^2}{k} dy dx dz + \int_0^3 \int_0^2 \int_0^1 \frac{4z^3}{k} dz dy dx \\ &= \frac{2}{k} \left(\int_0^3 2x dx \right) + \frac{3}{k} \left(\int_0^2 3y^2 dy \right) + \frac{6}{k} \left(\int_0^1 4z^3 dz \right) = \frac{48}{k}, \end{aligned}$$

όπου έχουμε παραλείψει τον λεπτομερή υπολογισμό των τελευταίων τριών απλών ολοκληρωμάτων. Επομένως, $k = 48$.

(β') Με παρόμοιο τρόπο θα υπολογίσουμε και τη μέση τιμή $E(XYZ)$, ως,

$$\begin{aligned} &\iiint_{\mathbb{R}^3} xyz f_{XYZ}(x, y, z) dx dy dz \\ &= \frac{1}{48} \iiint_S xyz(2x + 3y^2 + 4z^3) dx dy dz \\ &= \frac{1}{48} \int_0^1 \int_0^2 \int_0^3 2x^2yz dx dy dz + \frac{1}{48} \int_0^1 \int_0^2 \int_0^3 3xy^3z dx dy dz \\ &\quad + \frac{1}{48} \int_0^1 \int_0^2 \int_0^3 4xyz^4 dx dy dz \\ &= \frac{1}{24} \left(\int_0^3 x^2 dx \right) \left(\int_0^2 y dy \right) \left(\int_0^1 z dz \right) \\ &\quad + \frac{1}{16} \left(\int_0^3 x dx \right) \left(\int_0^2 y^3 dy \right) \left(\int_0^1 z dz \right) \\ &\quad + \frac{1}{12} \left(\int_0^3 x dx \right) \left(\int_0^2 y dy \right) \left(\int_0^1 z^4 dz \right) = \frac{37}{80}, \end{aligned}$$

όπου και πάλι έχουμε παραλείψει τον λεπτομερή υπολογισμό των τελευταίων εννέα ολοκληρωμάτων.

(γ) Τέλος, ορίζοντας το σύνολο, $B = \{(x, y, z) : x + y + z \leq 1\} \subset \mathbb{R}^3$, έχουμε,

$$\begin{aligned} \Pr(X + Y + Z \leq 1) &= \iiint_B f_{XYZ}(x, y, z) dx dy dz \\ &= \frac{1}{48} \iiint_{S \cap B} (2x + 3y^2 + 4z^3) dx dy dz \\ &= \frac{1}{48} \int_0^1 \left(\int_0^{1-z} \left(\int_0^{1-y-z} (2x + 3y^2 + 4z^3) dx \right) dy \right) dz. \end{aligned}$$

Το εσωτερικό ολοκλήρωμα ως προς x είναι,

$$\begin{aligned} &\int_0^{1-y-z} (2x + 3y^2 + 4z^3) dx \\ &= [x^2 + (3y^2 + 4z^3)x]_0^{1-y-z} \\ &= [1 + z^2 - 2z + 4z^3 - 4z^4] + y[-2 + z - 4z^3] + y^2[4 - 3z] - 3y^3, \end{aligned}$$

και αντικαθιστώντας την παραπάνω έκφραση στο δεύτερο ολοκλήρωμα ως προς y έχουμε ότι ισούται με,

$$\begin{aligned} &\left[(1 + z^2 - 2z + 4z^3 - 4z^4)y + \left(-1 + \frac{1}{2}z - 2z^3\right)y^2 + \left(\frac{4}{3} - z\right)y^3 - \frac{3}{4}y^4 \right]_0^{1-z} \\ &= 2z^5 - \frac{15}{4}z^4 + \frac{1}{6}z^3 + \frac{7}{2}z^2 - \frac{5}{2}z + \frac{7}{12}, \end{aligned}$$

οπότε αντικαθιστώντας αυτή την έκφραση στο τρίτο ολοκλήρωμα ως προς z , βρίσκουμε,

$$\begin{aligned} \Pr(X + Y + Z \leq 1) &= \frac{1}{48} \int_0^1 \left[2z^5 - \frac{15}{4}z^4 + \frac{1}{6}z^3 + \frac{7}{2}z^2 - \frac{5}{2}z + \frac{7}{12} \right] dz \\ &= \frac{1}{48} \left[\frac{1}{3}z^6 - \frac{15}{20}z^5 + \frac{1}{24}z^4 + \frac{7}{6}z^3 - \frac{5}{4}z^2 + \frac{7}{12}z \right]_0^1 \\ &= \frac{1}{384}. \end{aligned}$$

8. *Περιθώρια και από κοινού πυκνότητα.* Η απόδειξη είναι παρόμοια με εκείνη της τέταρτης βασικής ιδιότητας της από κοινού πυκνότητας: Για οποιαδήποτε $a < b$ και $a' < b'$, η πιθανότητα $\Pr(a \leq X \leq b, a' \leq Y \leq b')$ μπορεί να εκφραστεί ως,

$$\Pr(a \leq X \leq b, a' \leq Y \leq b', -\infty < Y < \infty) = \int_a^b \int_{a'}^{b'} \left(\int_{-\infty}^{\infty} f_{XYZ}(x, y, z) dz \right) dy dx,$$

άρα, αν ορίσουμε τη συνάρτηση,

$$f_{XY}(x, y) = \int_{-\infty}^{\infty} f_{XYZ}(x, y, z) dz,$$

τότε ικανοποιεί τον ορισμό της από κοινού πυκνότητας 15.1 για δύο συνεχείς Τ.Μ. X, Y .

9. *Μίξη πυκνοτήτων.* Όταν η $Z = 1$ η $W = X$, ενώ όταν η $Z = 0$ η $W = Y$. Άρα μπορούμε να σκεφτούμε την W ως μια τυχαία επιλογή, με πιθανότητα $1/2$ για τα δύο αποτελέσματα, μεταξύ της X και της Y .

Για να βρούμε την πυκνότητα της W εφαρμόζουμε τον κανόνα συνολικής πιθανότητας: Για $a < b$ ανθαιρέτα,

$$\begin{aligned} \Pr(a \leq W \leq b) &= \Pr(a \leq W \leq b | Z = 1) \Pr(Z = 1) + \Pr(a \leq W \leq b | Z = 0) \Pr(Z = 0) \\ &= \frac{1}{2} \Pr(a \leq X \leq b) + \frac{1}{2} \Pr(a \leq Y \leq b) \\ &= \frac{1}{2} \int_a^b f_X(x) dx + \frac{1}{2} \int_a^b f_Y(y) dy \\ &= \int_a^b \left[\frac{1}{2} f_X(x) + \frac{1}{2} f_Y(x) \right] dx. \end{aligned}$$

Άρα, αντικαθιστώντας τις τιμές των πυκνοτήτων $f_X(x)$ και $f_Y(y)$, αν ορίσουμε τη συνάρτηση,

$$g(x) = \frac{1}{2} f_X(x) + \frac{1}{2} f_Y(x) = \begin{cases} \frac{1}{2} e^{-x}, & \text{για } 0 \leq x < 1 \text{ και } x > 2, \\ \frac{1}{2} e^{-x} + \frac{1}{2}, & \text{για } 1 \leq x \leq 2, \\ 0, & \text{για } x < 0, \end{cases}$$

τότε η $g(x)$ ικανοποιεί τον ορισμό της πυκνότητας της συνεχούς Τ.Μ. W όπως δίνεται στο Κεφάλαιο 10.

10. ★ *Φτου φτου φτου.*

(α) Η περιθώρια πυκνότητα του X , από την τέταρτη ιδιότητα της από κοινού πυκνότητας είναι $f_X(x) = 0$ για $x \leq 0$, ενώ για $x > 0$ έχουμε,

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XYZ}(x, y, z) dy dz \\ &= \int_{-\infty}^{\infty} \int_{5/x}^{10/x} \frac{x}{50\sqrt{2\pi}} \exp \left\{ -\frac{x}{2} - \frac{(y-z)^2}{50} \right\} dy dz \\ &= \frac{x}{10} e^{-x/2} \int_{5/x}^{10/x} \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi \times 5^2}} \exp \left\{ -\frac{(y-z)^2}{2 \times 5^2} \right\} dz \right) dy, \end{aligned}$$

όπου παρατηρούμε ότι το εσωτερικό ολοκλήρωμα ως προς z είναι το ολοκλήρωμα της πυκνότητας της $N(y, 5^2)$ κατανομής σε όλο το \mathbb{R} , άρα ισούται με 1. Οπότε βρίσκουμε,

$$f_X(x) = \frac{x}{10} e^{-x/2} \int_{5/x}^{10/x} dy = \frac{1}{2} e^{-x/2},$$

δηλαδή η περιθώρια κατανομή του χρόνου των χασμουρητών X είναι $\text{Exp}(2)$. Συνεπώς, η madame Dery χασμουριέται κατά μέσο όρο για δύο λεπτά ανά ξεμάτιασμα.

(β') Από τον ορισμό της δεσμευμένης πιθανότητας, η ζητούμενη πιθανότητα είναι,

$$\Pr(Y \leq 10 | X < 1) = \frac{\Pr(X < 1, Y \leq 10)}{\Pr(X < 1)}, \quad (\text{A'.25})$$

όπου, από το πρώτο σκέλος, ήδη ξέρουμε πως $X \sim \text{Exp}(2)$, άρα, ο παρονομαστής είναι, $\Pr(X < 1) = 1 - e^{-1/2}$. Για τον αριθμητή παρατηρούμε ότι, από τους περιορισμούς των τιμών της πυκνότητας, η μόνη περίπτωση να έχουμε $Y \leq 10$ και $X < 1$ είναι αν το $X \geq 1/2$. Οπότε,

$$\begin{aligned} \Pr(X < 1, Y \leq 10) &= \Pr(X < 1, Y \leq 10, -\infty < Z < \infty) \\ &= \int_{-\infty}^{\infty} \int_0^{10} \int_{-\infty}^1 f_{XYZ}(x, y, z) dx dy dz \\ &= \int_{1/2}^1 \int_{5/x}^{10} \int_{-\infty}^{\infty} \frac{x}{50\sqrt{2\pi}} \exp\left\{-\frac{x}{2} - \frac{(y-z)^2}{50}\right\} dz dy dx, \end{aligned}$$

και προχωρώντας όπως και στο προηγούμενο ερώτημα, το παραπάνω τριπλό ολοκλήρωμα μπορεί να υπολογιστεί ως,

$$\int_{1/2}^1 \int_{5/x}^{10} \frac{x}{10} e^{-x/2} \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi} \times 5^2} \exp\left\{-\frac{(y-z)^2}{2 \times 5^2}\right\} dz \right) dy dx,$$

όπου το εσωτερικό ολοκλήρωμα ως προς z ισούται με ένα. Συνεπώς,

$$\begin{aligned} \Pr(X < 1, Y \leq 10) &= \int_{1/2}^1 \int_{5/x}^{10} \frac{x}{10} e^{-x/2} dy dx \\ &= \int_{1/2}^1 \left(10 - \frac{5}{x}\right) \frac{x}{10} e^{-x/2} dx \\ &= \int_{1/2}^1 \left(x - \frac{1}{2}\right) e^{-x/2} dx \\ &= \left[-(2x+3)e^{-x/2}\right]_{1/2}^1 \\ &= 4e^{-1/4} - 5e^{-1/2}. \end{aligned}$$

Τέλος, αντικαθιστώντας στη σχέση (Α΄.25), βρίσκουμε,

$$\Pr(Y \leq 10 | X < 1) = \frac{4e^{-1/4} - 5e^{-1/2}}{1 - e^{-1/2}} \approx 0.2098.$$

11. *Συσχέτιση, συνδιακύμανση, διασπορά.* Οι αποδείξεις είναι ακριβώς ίδιες με εκείνες που είδαμε στη διακριτή περίπτωση, στις λύσεις της Άσκησης 10 του Κεφαλαίου 9 και της Άσκησης 17 του Κεφαλαίου 11. Η μόνη διαφορά είναι ότι χρησιμοποιούμε τις αντίστοιχες ιδιότητες της μέσης τιμής, της διασποράς και της συνδιακύμανσης για συνεχείς αντί για διακριτές Τ.Μ. Συγκεκριμένα:

- Αντί για τις εναλλακτικές εκφράσεις (6.8) και (9.2) της διασποράς και της συνδιακύμανσης διακριτών Τ.Μ., εδώ χρησιμοποιούμε τη σχέση (10.13) και την Ιδιότητα 3 της Πρότασης 15.2 αντίστοιχα.
- Αντί για τη βασική ιδιότητα της μέσης τιμής από την πρώτη ιδιότητα του Θεωρήματος 6.1, εδώ χρησιμοποιούμε την Πρόταση 15.1.
- Για τη διασπορά του γινομένου μιας σταθεράς με μια Τ.Μ., αντί για την Ιδιότητα 2 του Θεωρήματος 6.1, χρησιμοποιούμε την Ιδιότητα 3 του Θεωρήματος 11.2.

12. * Η διμεταβλητή κανονική κατανομή.

(α΄) Από το Πρόσχημα 15.1 σε συνδυασμό με τις απλές ιδιότητες της κανονικής κατανομής στο Θεώρημα 12.1 έχουμε πως οι X, Y ακολουθούν κανονική κατανομή με,

$$\begin{aligned} E(X) &= \mu + \sigma E(Z) = \mu, \\ E(Y) &= \nu + \frac{c}{\sigma} E(Z) + \sqrt{\tau^2 - \frac{c^2}{\sigma^2}} E(W) = \nu, \\ \text{Var}(X) &= \sigma^2 \text{Var}(Z) = \sigma^2, \\ \text{και } \text{Var}(Y) &= \frac{c^2}{\sigma^2} \text{Var}(Z) + \left(\tau^2 - \frac{c^2}{\sigma^2} \right) \text{Var}(W) = \tau^2. \end{aligned}$$

(β΄) Χρησιμοποιώντας τις τιμές των $E(X)$ και $E(Y)$ που μόλις υπολογίσαμε, από τον ορισμό της συνδιακύμανσης και τους ορισμούς των X, Y , βρίσκουμε ότι,

$$\begin{aligned} \text{Cov}(X, Y) &= E[(X - E(X))(Y - E(Y))] \\ &= E \left[\sigma Z \left(\frac{c}{\sigma} Z + \sqrt{\tau^2 - \frac{c^2}{\sigma^2}} W \right) \right] \\ &\stackrel{(a)}{=} cE(Z^2) + \sigma \sqrt{\tau^2 - \frac{c^2}{\sigma^2}} E(ZW) \\ &\stackrel{(b)}{=} c + \sigma \sqrt{\tau^2 - \frac{c^2}{\sigma^2}} E(Z)E(W) \\ &\stackrel{(c)}{=} c, \end{aligned}$$

όπου στο βήμα (a) χρησιμοποιήσαμε τη βασική ιδιότητα της μέσης τιμής από την Πρόταση 15.1, στο (b) αντικαταστήσαμε την $E(Z^2) = \text{Var}(Z) + (E(Z))^2 = 1$ και χρησιμοποιήσαμε την Ιδιότητα 1 από το Θεώρημα 15.1 επειδή οι Z, W είναι ανεξάρτητες, και στο βήμα (c) απλά χρησιμοποιήσαμε το γεγονός ότι $E(Z) = E(W) = 0$.

(γ') Η Άσκηση 11 παραπάνω μάς λέει πως ο συντελεστής συσχέτισης $\rho_{X,Y}^2 \leq 1$, δηλαδή,

$$\rho_{X,Y}^2 = \frac{[\text{Cov}(X, Y)]^2}{\text{Var}(X)\text{Var}(Y)} = \frac{c^2}{\sigma^2\tau^2} \leq 1,$$

οπότε πάντοτε θα έχουμε $c^2 \leq \sigma^2\tau^2$.

Ο επιπλέον περιορισμός $c \neq \pm\sigma\tau$ απλώς αποκλείει την τετριμμένη περίπτωση όπου η Y μπορεί απλά να εκφραστεί ως γραμμικός μετασχηματισμός της X . Για παράδειγμα, αν $c = \sigma\tau$, τότε,

$$X = \mu + \sigma Z, \quad Y = \nu + \tau Z,$$

και κατά συνέπεια,

$$Y = \left(\nu - \frac{\tau\mu}{\sigma}\right) + \frac{\tau}{\sigma}X.$$

[Η περίπτωση $c = -\sigma\tau$ είναι παρόμοια.]

(δ') Αν και απαιτεί μακροσκελείς υπολογισμούς, η βασική ιδέα της απόδειξης του τελευταίου ερωτήματος είναι απλή: Θα εξετάσουμε την πιθανότητα $\Pr(a \leq X \leq b, a' \leq Y \leq b')$ και θα δείξουμε, μέσω απλών μετασχηματισμών, ότι μπορεί να εκφραστεί ως το διπλό ολοκλήρωμα της f_{XY} που ορίστηκε από την (15.8) στο ορθογώνιο $[a, b] \times [a', b']$.

Κατ' αρχάς, αντικαθιστώντας του ορισμούς των Γ.Μ. X, Y , η ζητούμενη πιθανότητα είναι,

$$\begin{aligned} \Pr(a \leq X \leq b, a' \leq Y \leq b') &= \Pr\left(a \leq \mu + \sigma Z \leq b, a' \leq \nu + \frac{c}{\sigma}Z + \sqrt{\tau^2 - \frac{c^2}{\sigma^2}}W \leq b'\right) \\ &= \Pr\left(a \leq \mu + \sigma Z \leq b, a' \leq \nu + \tau\left[\rho Z + \sqrt{1 - \rho^2}\right]W \leq b'\right), \end{aligned}$$

και εκφράζοντάς την ως προς τις Z, W αντί για τις X, Y γίνεται,

$$\Pr\left(\frac{a - \mu}{\sigma} \leq Z \leq \frac{b - \mu}{\sigma}, \frac{\frac{a' - \nu}{\tau} - \rho Z}{\sqrt{1 - \rho^2}} \leq W \leq \frac{\frac{b' - \nu}{\tau} - \rho Z}{\sqrt{1 - \rho^2}}\right).$$

Τώρα, εφόσον οι Z, W είναι ανεξάρτητες, το κριτήριο (15.7) μας λέει πως έχουν από κοινού πυκνότητα $\phi(z)\phi(w)$, όπου ϕ είναι η πυκνότητα της τυπικής κανονικής

κατανομής. Άρα, από το θεώρημα του Fubini, η πιο πάνω πιθανότητα ισούται με το διπλό ολοκλήρωμα,

$$\int_{(a-\mu)/\sigma}^{(b-\mu)/\sigma} \int_{\frac{a'-\nu-\rho z}{\tau}}^{\frac{b'-\nu-\rho z}{\tau}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \frac{1}{\sqrt{2\pi}} e^{-w^2/2} dw dz,$$

και κάνοντας στο εσωτερικό ολοκλήρωμα την αντικατάσταση,

$$y = \nu + \tau \left[\rho z + \sqrt{1 - \rho^2} w \right],$$

έτσι ώστε,

$$dy = \tau \sqrt{1 - \rho^2} dw \quad \text{και} \quad w = \frac{y - \nu}{\tau} - \frac{\rho z}{\sqrt{1 - \rho^2}},$$

έχουμε,

$$\begin{aligned} & \Pr(a \leq X \leq b, a' \leq Y \leq b') \\ &= \int_{(a-\mu)/\sigma}^{(b-\mu)/\sigma} \int_{a'}^{b'} \frac{1}{2\pi\tau\sqrt{1-\rho^2}} \exp \left\{ -\frac{z^2}{2} - \frac{1}{2(1-\rho^2)} \left(\frac{y-\nu}{\tau} - \rho z \right)^2 \right\} dy dz \\ &= \int_{(a-\mu)/\sigma}^{(b-\mu)/\sigma} \int_{a'}^{b'} \frac{1}{2\pi\tau\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[\frac{(y-\nu)^2}{\tau^2} + z^2 - 2\rho z \frac{y-\nu}{\tau} \right] \right\} dy dz. \end{aligned}$$

Τέλος, πραγματοποιώντας την αντικατάσταση $x = \sigma z + \mu$, έτσι ώστε $z = \frac{x-\mu}{\sigma}$ και $dx = \sigma dz$, βρίσκουμε,

$$\begin{aligned} & \Pr(a \leq X \leq b, a' \leq Y \leq b') \\ &= \int_a^b \int_{a'}^{b'} \frac{1}{2\pi\sigma\tau\sqrt{1-\rho^2}} \\ & \quad \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[\frac{(y-\nu)^2}{\tau^2} + \left(\frac{x-\mu}{\sigma} \right)^2 - 2\rho \left(\frac{x-\mu}{\sigma} \right) \left(\frac{y-\nu}{\tau} \right) \right] \right\} dy dx \\ &= \int_a^b \int_{a'}^{b'} f_{XY}(x, y) dy dx, \end{aligned}$$

άρα, σύμφωνα με τον Ορισμό 15.1, η $f_{XY}(x, y)$ είναι πράγματι η από κοινού πυκνότητα των X, Y .

13. Το Κ.Ο.Θ. χωρίς το «Ο». Έστω πως οι Τ.Μ. $\{X_n\}$ είναι ανεξάρτητες με κοινή κατανομή $X_i \sim N(\mu, \sigma^2)$. Τότε, από την Ιδιότητα 3 του Θεωρήματος 12.1, η κάθε $\frac{X_i - \mu}{\sigma}$ έχει κατανομή $N(0, 1)$, και από το Πόρισμα 15.1 το άθροισμα,

$$\frac{1}{\sigma} \sum_{i=1}^N (X_i - \mu) \sim N(0, N).$$

Οπότε, εφαρμόζοντας και πάλι την Ιδιότητα 3 του Θεωρήματος 12.1, το κανονικοποιημένο άθροισμα,

$$\bar{S}_N = \frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N (X_i - \mu),$$

έχει τυπική κανονική κατανομή για κάθε $N \geq 1$, όχι μόνο κατά προσέγγιση για μεγάλα N όπως μας λέει το Κ.Ο.Θ.

Βιβλιογραφία

Παραθέτουμε έναν συνοπτικό κατάλογο βιβλίων, στα ελληνικά και τα αγγλικά, που μπορούν να χρησιμοποιηθούν για περαιτέρω ή ταυτόχρονη μελέτη.

1. ΕΙΣΑΓΩΓΗ ΣΤΙΣ ΠΙΘΑΝΟΤΗΤΕΣ

- (α') *Εισαγωγή στις πιθανότητες*, D.P. Bertsekas, J.N. Tsitsiklis. Εκδόσεις Τζιόλα, 2010. [Πρωτότυπο στα αγγλικά: *Introduction to Probability*, D.P. Bertsekas, J.N. Tsitsiklis. 2η έκδοση, Athena Scientific, 2008.]
- (β') *Εισαγωγή στη θεωρία πιθανοτήτων*, P. Hoel, S. Port, C. Stone. Πανεπιστημιακές Εκδόσεις Κρήτης, 2001. [Πρωτότυπο στα αγγλικά: *Introduction to Probability Theory*, P.G. Hoel, S.C. Port, C.J. Stone. Brooks Cole, 2009.]
- (γ') *Βασικές αρχές θεωρίας πιθανοτήτων*, S. Ross. Εκδόσεις Κλειδάριθμος, 2012. [Πρωτότυπο στα αγγλικά: *First Course in Probability*, S. Ross. 8η έκδοση, Prentice Hall, 2009.]

2. ΠΕΡΑΙΤΕΡΩ ΘΕΩΡΙΑ ΚΑΙ ΕΦΑΡΜΟΓΕΣ ΤΩΝ ΠΙΘΑΝΟΤΗΤΩΝ

- (α') *Πιθανότητες, τυχαίες μεταβλητές και στοχαστικές διαδικασίες*, A. Papoulis, S.U. Pillai. 4η έκδοση, Εκδόσεις Τζιόλα, 2007. [Πρωτότυπο στα αγγλικά: *Probability, Random Variables and Stochastic Processes*, A. Papoulis, S.U. Pillai. 4η έκδοση, McGraw Hill, 2002.]
- (β') *Introduction to Probability Models*, S. Ross. 10η έκδοση, Academic Press, 2010.
- (γ') *Stochastic Processes*, S. Ross. 2η έκδοση, Wiley, 1995.

3. ΕΦΑΡΜΟΓΕΣ ΤΩΝ ΠΙΘΑΝΟΤΗΤΩΝ ΣΤΗΝ ΠΛΗΡΟΦΟΡΙΚΗ

- (α') *Probability and Algorithms*, J.M. Steele, D. Aldous, D.J. Bertsimas, E.G. Coffman, D. Hochbaum, M. Hofri, J.C. Lagarias, S.T. Weidman. U.S. National Research Council, 1992. Διαθέσιμο online:
<http://www.nap.edu/openbook.php?isbn=0309047765>

- (β') *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*, M. Mitzenmacher, E. Upfal. Cambridge University Press, 2005.
- (γ') *Randomized Algorithms*, R. Motwani, P. Raghavan. Cambridge University Press, 1995.
- (δ') *Average Case Analysis of Algorithms on Sequences*, W. Szpankowski. Wiley Series in Discrete Mathematics and Optimization, 2011.

4. ΣΤΑΤΙΣΤΙΚΗ ΚΑΙ ΠΡΟΣΟΜΟΙΩΣΗ

- (α') *Introduction to the Practice of Statistics*, D.S. Moore, G.P. McCabe, B.A. Craig. W.H. Freeman, 2014.
- (β') *Statistical Inference*, G. Casella, R.L. Berger. Duxbury Press, 2001.
- (γ') *Monte Carlo Statistical Methods*, C.P. Robert, G. Casella. Springer, 2004.
- (δ') *Monte Carlo Methods in Financial Engineering*, P. Glasserman. Springer, 2003.

5. ΘΕΩΡΙΑ ΠΙΘΑΝΟΤΗΤΩΝ ΒΑΣΕΙ ΤΗΣ ΘΕΩΡΙΑΣ ΜΕΤΡΟΥ

- (α') *An Introduction to Probability Theory and Its Applications, volumes I-II*, W. Feller. 3η έκδοση, Wiley, 1968.
- (β') *Probability and Measure*, P. Billingsley. 3η έκδοση, Wiley, 1995.
- (γ') *Probability: Theory and Examples*, R. Durrett. 4η έκδοση, Cambridge University Press, 2010.

6. ΜΑΘΗΜΑΤΙΚΟΣ ΛΟΓΙΣΜΟΣ

- (α') *Διαφορικός και ολοκληρωτικός λογισμός, τόμοι I-II*, T.M. Apostol. 1η έκδοση, Εκδόσεις Ατλαντίς, 1962. [Πρωτότυπο στα αγγλικά: *Calculus, volumes I-II*, T.M. Apostol. 2η έκδοση, Wiley, 1967.]
- (β') *Διαφορικός και ολοκληρωτικός λογισμός*, M. Spivak. 2η έκδοση, Πανεπιστημιακές Εκδόσεις Κρήτης, 2010. [Πρωτότυπο στα αγγλικά: *Calculus*, M. Spivak. 4η έκδοση, Publish or Perish, 2009.]
- (γ') *Απειροστικός λογισμός (ενιαίος τόμος)*, R.L. Finney, F.R. Giordano, M.D. Weir. Πανεπιστημιακές Εκδόσεις Κρήτης, 2012. [Πρωτότυπο στα αγγλικά: *Thomas' Calculus*, R.L. Finney, F.R. Giordano, M.D. Weir. Addison Wesley Longman, 2001.]

Ευρετήριο

- Bayes, κανόνας, 47, 55, 58
Bernoulli
 Daniel, 82
 κατανομή, 85
Cauchy-Schwarz, ανισότητα, 185
Chebychev, ανισότητα, 134, 184, 189, 231
Chernoff, φράγμα, 151, 188
Fermat, Pierre de, 2
Fourier, ανάλυση, 217
Fubini, θεώρημα, 267, 289
Gauss, Johann Carl Frederich, 2, 129
Kolmogorov, Andrei N., 5
Lindeberg, θεώρημα, 217, 224
Markov
 Andrei A., 4
 αλυσίδα, 4
 ανισότητα, 133, 183, 188
Merkel, Angela, 173
Monty Hall, παιχνίδι, 14, 18, 53
Pascal, Blaise, 2
Poisson
 κατανομή, 97
 προσέγγιση στη διωνυμική, 98, 203, 233
Stirling, τύπος, 96, 102, 129, 222
Taylor, ανάπτυγμα, 102, 110, 226, 354

cut set, 115, 130
de Moivre
 Abraham, 221
 de Moivre-Laplace, θεώρημα, 217, 221
min-cut, 115, 130
quicksort αλγόριθμος, 124, 132
randomized αλγόριθμος, 5, 111, 113, 115, 128, 129, 132
string matching, 121, 131

Βαρουφάκης, Γιάννης, 108, 173
Βήτα κατανομή, 172

Ζυγοβίστι Αρκαδίας, 214

Κεντρικό Οριακό Θεώρημα (Κ.Ο.Θ.), 2, 149, 155, 191, 198, 199, 217, 219, 221, 224, 237, 286
Κορδατέας Α.Ε., 261

Νόμος των Μεγάλων Αριθμών (N.M.A.), 133, 144, 147, 148, 175, 183, 186, 189, 191, 217, 219, 237

Τζίφας, Παν., Αρχιπλοίαρχος, 261

άθροισμα ανεξάρτητων κανονικών T.M., 280
αλγόριθμος του Karger, 115
αλυσίδα Markov, 4
ανάλυση Fourier, 217
ανάπτυγμα Taylor, 102, 110, 226, 354
ανεξαρτησία, 3, 41, 42, 73, 142, 277, 278
ανεξάρτητα ενδεχόμενα, 41, 42, 58, 73, 184
ανεξάρτητες T.M.
 διακριτές, 73, 137, 141, 142, 144, 147, 191, 198, 199, 219
 συνεχείς, 184, 186, 191, 198, 199, 219, 277–280

ανισότητα Cauchy-Schwarz, 185
ανισότητα του Chebychev, 134, 184, 189, 231
ανισότητα του Markov, 133, 183, 188
από κοινού κατανομή
 διακριτή, 135
 συνεχής, 265
από κοινού πυκνότητα
 διακριτή, 136
 διακριτή, ιδιότητες, 137, 150

- συνεχής, 266
 συνεχής, ιδιότητες, 269
 απόσταση ℓ_1 , 83
 απόσταση χ^2 , 83, 173, 232
 αρμονική σειρά, 127
- γεωμετρική κατανομή, 89
 ιδιότητες, 90
 γεωμετρική σειρά, 101
 γκαουσιανή κατανομή, 192
- δείκτρια συνάρτηση, 106, 225, 227, 232
 δείκτρια T.M., 86, 106, 237
 δεσμευμένη πιθανότητα, 41, 44, 46, 53, 55, 58, 59
 δεσμευμένο μέτρο πιθανότητας, 60
 δημοσκόπηση, 108, 145, 204, 207, 237, 238, 261
 δημοψήφισμα, 207
 διακριτή T.M., 64
 διασπορά
 διακριτής T.M., 70
 εκτίμηση, 262
 συνεχούς T.M., 163
- διάστημα εμπιστοσύνης, 235, 236, 240–242, 261
 διατάξεις, 31, 32, 35
 διμεταβλητή κανονική κατανομή, 285
 διπλό ολοκλήρωμα, 265, 287
 ιδιότητες, 287
 διωνυμική κατανομή, 85, 88, 97, 98, 203, 222, 233
 δυναμοσειρές, 102
- εκθετική κατανομή, 177
 ιδιότητες, 178
 εκθετική συνάρτηση, 98, 102, 110
 έλεγχος ανεξαρτησίας, 248, 250, 262
 έλεγχος παραμέτρου Bernoulli, 245, 247, 262
 έλεγχος υπόθεσης, 235, 243, 245, 247, 248, 250, 262
- εμπειρική κατανομή, 246
 εμπειρικός μέσος όρος, 144, 147, 155, 186, 191, 198, 199, 219, 237, 240–242, 246, 247, 249, 250, 253, 254, 257, 261, 263
- ενδεχόμενο, 7, 11, 12, 20, 24, 29, 41, 42
 επαλήθευση ισότητας πολυωνύμων, 112
 επιλογές, 31, 32, 35
 επιλογή με επανατοποθέτηση, 88, 93
 επιλογή χωρίς επανατοποθέτηση, 94
- επίπεδο εμπιστοσύνης, 240–242, 261
 επίπεδο σημαντικότητας, 247, 250, 262
 εύρεση ελαχιστιαίου cut set, 114
- θεώρημα de Moivre-Laplace, 217, 221
 θεώρημα του Fubini, 267, 289
 θεώρημα του Lindeberg, 217, 224
- ιδιότητα έλλειψης μνήμης
 γεωμετρική κατανομή, 90
 εκθετική κατανομή, 178
- ιδιότητες
 ανεξαρτησίας, 278
 γεωμετρικής κατανομής, 90
 διακριτής από κοινού πυκνότητας, 137, 150
 διακριτής πυκνότητας, 65
 διπλού ολοκληρώματος, 287
 εκθετικής κατανομής, 178
 κανονικής κατανομής, 193
 μέσης τιμής και διασποράς, 73, 184
 συνδιακύμανσης, 141, 276
 συνεχούς από κοινού πυκνότητας, 269
 συνεχούς πυκνότητας, 158
 συνεχούς συνάρτησης κατανομής, 160
- ισοπίθανα στοιχειώδη ενδεχόμενα, 23, 24, 29, 37
- κανόνας δεσμευμένης συνολικής πιθανότητας, 61
 κανόνας συνολικής πιθανότητας, 53, 58
 κανόνας του Bayes, 47, 55, 58
 κανόνες αρίθμησης, 37
 κανόνες πιθανότητας 1–5, 23
 κανόνες πιθανότητας 6–10, 58
 κανονική κατανομή, 149, 156, 192, 195, 199, 203, 217, 219, 228, 279, 285, 286
 διμεταβλητή, 285
 ιδιότητες, 193
- κανονική προσέγγιση στη διωνυμική, 203
 κανονικοποιημένο άθροισμα, 199, 221, 240, 241
 κανονικοποίηση, 182, 188, 193, 199
- κατανομή
 Bernoulli, 85
 Poisson, 97
 Βήτα, 172
 γεωμετρική, 89
 γκαουσιανή, 192
 διμεταβλητή κανονική, 285

- διωνυμική, 88
 εκθετική, 177
 κανονική, 192
 λογαριθμοκανονική, 213
 ομοιόμορφη, 175
 τυπική κανονική, 194
 υπεργεωμετρική, 94
 χ^2 , 213, 283
- κριτήριο ανεξαρτησίας, 142, 277, 278
- λογαριθμοκανονική κατανομή, 213
- μέθοδος της δεύτερης ροπής, 82
 μείωση διασποράς, 252, 254, 257
 μέση τιμή
 - διακριτής T.M., 70, 83
 - συνάρτησης διακριτής T.M., 70, 80
 - συνάρτησης πολλαπλών T.M., 274, 275
 - συνάρτησης συνεχούς T.M., 163
 - συνεχούς T.M., 163
- μεταβλητές ελέγχου, 252, 254, 257, 263
- μετασχηματισμός
 - γραμμικός, 181, 188, 189, 213
 - συνεχούς T.M., 181
- μετρησιμότητα, 77, 168, 281
- μέτρο πιθανότητας, 19, 24
 - δεσμευμένο, 60
 - συνέχεια, 27
- ξεμάτιασμα, 285
- ξένα ενδεχόμενα, 12, 19, 24, 53, 55
- ομοιόμορφη κατανομή, 175
- παιχνίδι Monty Hall, 14, 18, 53
- περιθώρια πυκνότητα
 - διακριτής T.M., 137
 - συνεχούς T.M., 269
- πιθανοκρατική ανάλυση, 5, 111, 121, 126, 131
- πράξεις συνόλων, 8, 13
- πρόβλημα γενεθλίων, 92
- πρόβλημα ταξινόμησης, 123
- προσαρμοστική εκτιμήτρια, 257, 263
- προσέγγιση Poisson στη διωνυμική, 98, 203, 233
- προσομοίωση, 131, 252, 254, 258, 263
- πυκνότητα
 - από κοινού, διακριτή, 136
 - από κοινού, συνεχής, 265
 - διακριτής T.M., 64
 - διακριτής T.M., περιθώρια, 137
 - συνεχούς T.M., 156
 - συνεχούς T.M., περιθώρια, 269
- στατιστική συνάρτηση, 246, 247, 249, 250, 262
- στατιστικό σφάλμα, 240–242, 261
- στοιχειώδες ενδεχόμενο, 12
- σύγκλιση κατά κατανομή, 199, 232, 233
- σύγκλιση κατά πιθανότητα, 146, 152, 219, 232
- σύγκλιση με πιθανότητα 1, 148, 153
- σύγκλιση ως προς την απόσταση χ^2 , 232
- συνάρτηση κατανομής, 65, 80, 199, 221
 - συνεχής, ιδιότητες, 160
- συνδιακύμανση, 140, 152, 189, 254, 256, 276, 285
 - ιδιότητες, 141, 276
- συνδυασμοί, 30–32, 35
- συνέλιξη
 - διακριτή, 81
 - συνεχής, 279
- συνεχής T.M., 156
- σύνολα, 8
- σύνολο τιμών τυχαίας μεταβλητής, 64
- συντελεστής συσχέτισης, 152, 189, 285
- ταυτότητα Vandermonde, 106
- τιμολόγηση δικαιώματος αγοράς, 258
- τυπική απόκλιση
 - διακριτής T.M., 70
 - συνεχούς T.M., 164
- τυπική κανονική κατανομή, 193–195, 199, 225, 286
 - τιμές συνάρτησης κατανομής, 211, 212
- τύπος του Stirling, 96, 102, 129, 222
- τυχαία μεταβλητή (T.M.), 63
 - δείτρια, 86, 106
 - διακριτή, 64
 - διακριτή, αυστηρός ορισμός, 77
 - ντετερμινιστική, 71, 80
 - συνεχής, 156
 - συνεχής, αυστηρός ορισμός, 169
- υπεργεωμετρική κατανομή, 94, 109
- φράγμα ένωσης, 51, 122, 131

φράγμα τομής, 131
φράγμα του Chernoff, 151, 188

χ^2

απόσταση, 83, 173, 232
έλεγχος, 245-248, 250, 262
κατανομή, 213, 283
χώρος πιθανότητας, 7, 12

Απόδοση αγγλικών όρων

- *average case analysis*: πιθανοκρατική ανάλυση
- *bits*: δυαδικά δεδομένα, δυαδικά ψηφία
- *cut set*: σύνολο αποκοπτούσών ακμών
- *divide and conquer*: διαίρει και βασίλευε
- *min-cut*: ελαχιστιαίο σύνολο αποκοπτούσών ακμών
- *multimedia*: πολυμεσικά δεδομένα
- *quicksort algorithm*: αλγόριθμος γρήγορης ταξινόμησης
- *randomized algorithm*: τυχαιοκρατικός αλγόριθμος
- *string matching*: ταίριασμα ακολουθιών
- *worst case analysis*: ανάλυση χειρότερης περίπτωσης

Αθήνα 2015

ISBN: 978-960-603-182-3
