Σημειώσεις
Θεωρίας Πιθανοτήτων

(βασισμένες στο Introduction to Probability Theory των Hoel, Port και Stone)

Πανεπιστημιο Κρητης
Τμημα Μαθηματικων
Φθινοπωρο 1999

Περιεχόμενα

Κεφάλαιο 1

Χώροι πιθανότητας

Η θεωρία πιθανοτήτων είναι ο κλάδος των μαθηματικών που ασχολείται με τα τυχαία φαινόμενα. Η μελέτη της κέρδισε πολλούς μαθηματικούς, τόσο για το θεωρητικό της ενδιαφέρον όσο και για τις επιτυχημένες εφαρμογές της σε πολλές περιοχές των φυσικών, βιολογικών και κοινωνικών επιστημών, στη μηχανική και στον επιχειρηματικό κόσμο.

Πολλά φαινόμενα έχουν την ιδιότητα η επανειλημμένη παρατήρησή τους κάτω από ένα δεδομένο σύστημα συνθηκών να οδηγεί πάντα στο ίδιο αποτέλεσμα. Για παράδειγμα, αν αφήνουμε μία μπάλα που ήταν αρχικά ακίνητη να πέσει από ύψος d μέτρων μέσα σε έναν κύλινδρο χωρίς αέρα, θα φτάνει στο έδαφος πάντα μετά από $t = \sqrt{2d/g}$ δευτερόλεπτα, όπου g η σταθερή επιτάχυνση της βαρύτητας σε $\mathrm{m}/\mathrm{sec}^2$. Υπάρχουν άλλα φαινόμενα των οποίων η επανειλημμένη παρατήρηση κάτω από ένα δεδομένο σύστημα συνθηκών δεν οδηγεί πάντα στο ίδιο αποτέλεσμα. Ένα οικείο παράδειγμα αυτού του είδους είναι η ρίψη ενός νομίσματος. Αν ρίξουμε ένα νόμισμα 1000 φορές, οι εμφανίσεις γραμμάτων ή κορώνας εναλλάσσονται με έναν φαινομενικά ακανόνιστο και απρόβλεπτο τρόπο. Αυτά τα φαινόμενα σκεφτόμαστε σαν τυχαία, και αυτά θα είναι το αντικείμενο της μελέτης μας.

Με την πρώτη ματιά φαίνεται ίσως αδύνατο να διατυπώσουμε αξιόλογα συμπεράσματα για τέτοια τυχαία φαινόμενα, όμως δεν είναι καθόλου έτσι. Η εμπειρία έχει δείξει ότι πολλά μη-ντετερμινιστικά φαινόμενα παρουσιάζουν μία *στατιστική κανονικότητα* που αξίζει να μελετηθεί. Σαν παράδειγμα, ας πάρουμε πάλι την ρίψη του νομίσματος. Για κάθε δεδομένη ρίψη του νομίσματος δεν μπορούμε να κάνουμε καμμία μη τετριμμένη πρόβλεψη, οι παρατηρήσεις όμως δείχνουν ότι για μεγάλο αριθμό δοκιμών το ποσοστό των εμφανίσεων γραμμάτων μοιάζει να κινείται γύρω από έναν σταθερό αριθμό p μεταξύ του 0 και του 1 (και ο p είναι πολύ κοντά στο $1/2$, εκτός αν το νόμισμα είναι τελείως ελαττωματικό). Φαίνεται ότι το ποσοστό που καταγράφουμε στις n δοκιμές τείνει στο p αν αφήσουμε το n να πάει στο άπειρο. Σκεφτόμαστε λοιπόν αυτό το οριακό ποσοστό p σαν την «πιθανότητα» να εμφανιστούν γράμματα σε μία και μοναδική ρίψη.

Πιό γενικά, ο ισχυρισμός ότι κάποιο πειραματικό αποτέλεσμα έχει πιθανότητα p

έχει την έννοια ότι για ένα μεγάλο πλήθος επαναλήψεων του πειράματος, το συγκεκριμένο αποτέλεσμα θα παρατηρηθεί για περίπου $100p\%$ από αυτές. Αυτή είναι η λεγόμενη ερμηνεία της πιθανότητας ως σχετικής συχνότητας. Είναι πολύ φυσιολογική για πολλές εφαρμογές της θεωρίας πιθανοτήτων σε προβλήματα των φυσικών επιστημών, συχνά όμως μοιάζει πολύ εξεζητημένη. Πώς, για παράδειγμα, θα μπορούσαμε να ερμηνεύσουμε σαν σχετική συχνότητα την πιθανότητα ότι ένα νεογέννητο μωρό θα ζήσει τουλάχιστον 70 χρόνια; Έχουν γίνει διάφορες προσπάθειες, από τις οποίες καμμία δεν είναι καθολικά αποδεκτή, για να δοθεί εναλλακτική ερμηνεία τέτοιων πιθανοθεωρητικών προτάσεων.

Για τη μαθηματική θεωρία των πιθανοτήτων, ο τρόπος ερμηνείας των πιθανοτήτων είναι αδιάφορος, όπως ακριβώς στη γεωμετρία ο τρόπος ερμηνείας των σημείων, των ευθειών και των επιπέδων είναι αδιάφορος. Θα χρησιμοποιούμε την ερμηνεία των πιθανοτήτων με βάση την σχετική συχνότητα μόνο σαν πλαίσιο για την διαισθητική κατανόηση των ορισμών και των θεωρημάτων που θα αναπτύξουμε στη συνέχεια.

1.1 Παραδείγματα τυχαίων φαινομένων

Σε αυτήν την παράγραφο θα συζητήσουμε δύο απλά παραδείγματα τυχαίων φαινομένων, που θα μας βοηθήσουν να εισαγάγουμε ομαλά την τυπική δομή της θεωρίας.

Παράδειγμα 1. Ένα δοχείο περιέχει s βώλους, που φέρουν τους αριθμούς $1, 2, \ldots, s$ αλλά είναι εντελώς όμοιες κατά τα άλλα. Θεωρούμε το εξής πείραμα. Ανακατεύουμε καλά τους βώλους στο δοχείο και κάποιος πλησιάζει το δοχείο και επιλέγει έναν βώλο. Σημειώνουμε τον αριθμό του βώλου και τον επανατοποθετούμε στο δοχείο. Το αποτέλεσμα του πειράματος είναι ο αριθμός του βώλου που επιλέχθηκε. Για το πείραμα αυτό δεν μπορεί να γίνει καμμία μη τετριμμένη πρόβλεψη.

Ας υποθέσουμε ότι επαναλαμβάνουμε αυτό το πείραμα n φορές. Συμβολίζουμε με $N_n(k)$ το πλήθος εκείνων από τις n δοκιμές στις οποίες επιλέχθηκε ο βώλος με τον αριθμό k. Έστω ότι είχαμε, ας πούμε, $s = 3$ βώλους και $n = 20$ δοκιμές. Τα αποτελέσματα αυτών των 20 δοκιμών θα μπορούσαν να περιγραφούν με την παράθεση των αριθμών που εμφανίστηκαν κατά τη σειρά που παρατηρήθηκαν. Ένα τυπικό αποτέλεσμα θα μπορούσε να είναι το

$$1, 1, 3, 2, 1, 2, 2, 3, 2, 3, 3, 2, 1, 2, 3, 3, 1, 3, 2, 2,$$

και στην περίπτωση αυτή,

$$N_{20}(1) = 5, \quad N_{20}(2) = 8, \quad N_{20}(3) = 7.$$

Οι σχετικές συχνότητες (δηλαδή, το ποσοστό των εμφανίσεων) των αποτελεσμάτων 1,2, και 3 είναι λοιπόν

$$\frac{N_{20}(1)}{20} = 0.25, \quad \frac{N_{20}(2)}{20} = 0.40, \quad \frac{N_{20}(3)}{20} = 0.35.$$

Καθώς το πλήθος των δοκιμών αυξάνει, θα περιμέναμε οι σχετικές συχνότητες $N_n(1)/n, \ldots, N_n(s)/n$ να σταθεροποιούνται σε κάποιους σταθερούς αριθμούς p_1, $p_2, \ldots, p_s$ (οι οποίοι σύμφωνα με τη διαίσθησή μας θα έπρεπε να είναι όλοι ίσοι με $1/s$ στην περίπτωση αυτή).

Σύμφωνα με την ερμηνεία της πιθανότητας με βάση την σχετική συχνότητα, ο αριθμός p_i πρέπει να δίνει την πιθανότητα επιλογής του βώλου με αριθμό i αν το πείραμα εκτελεστεί μία μόνο φορά ($i = 1, 2, \ldots, s$).

Θα κατασκευάσουμε τώρα ένα μαθηματικό μοντέλο για το πείραμα της επιλογής ενός βώλου από το δοχείο. Για το σκοπό αυτό, παίρνουμε πρώτα ένα σύνολο Ω που έχει s σημεία, τα οποία φέρνουμε σε ένα προς ένα αντιστοιχία με τα δυνατά αποτελέσματα του πειράματος. Μέσω αυτής της αντιστοιχίας, ακριβώς ένα σημείο του Ω σχετίζεται με το αποτέλεσμα να επιλεγεί ο βώλος με αριθμό k. Ονομάζουμε αυτό το σημείο ω_k. Στο σημείο ω_k αντιστοιχίζουμε τον αριθμό $p_k = 1/s$ και τον ονομάζουμε πιθανότητα του ω_k. Παρατηρούμε αμέσως ότι $0 \leq p_k \leq 1$ και ότι $p_1 + \ldots + p_k = 1$.

Υποθέτουμε τώρα ότι οι βώλοι, εκτός από το να φέρουν τους αριθμούς από 1 ως s, είναι χρωματισμένοι, οι πρώτοι r με κόκκινο χρώμα και οι υπόλοιποι $s - r$ με μαύρο χρώμα. Εκτελούμε το πείραμα όπως πρίν, τώρα όμως μας ενδιαφέρει μόνο το χρώμα του βώλου που επιλέχθηκε, και όχι ο αριθμός του. Με λίγη σκέψη βλέπουμε ότι η σχετική συχνότητα της επιλογής κόκκινου βώλου σε n επαναλήψεις του πειράματος είναι απλώς το άθροισμα των σχετικών συχνοτήτων $N_n(k)/n$ για εκείνες τις τιμές του k που αντιστοιχούν σε κόκκινο βώλο. Θα περιμέναμε, και η εμπειρία το αποδεικνύει, ότι για μεγάλα n αυτή η σχετική συχνότητα σταθεροποιείται σε κάποιον αριθμό. Αφού για μεγάλα n οι σχετικές συχνότητες $N_n(k)/n$ αναμένεται να βρίσκονται κοντά στον $p_k = 1/s$, είναι τελείως φυσικό να πούμε ότι η σχετική συχνότητα της επιλογής κόκκινου βώλου θα είναι κοντά στον r/s. Πάλι, η εμπειρία το επιβεβαιώνει. Σύμφωνα με την ερμηνεία της σχετικής συχνότητας, λέμε ότι η πιθανότητα επιλογής κόκκινου βώλου ισούται με r/s.

Ας δούμε με ποιό τρόπο αντανακλάται στο μοντέλο μας αυτός ο συλλογισμός. Έστω A το υποσύνολο του Ω που αποτελείται από εκείνα τα σημεία ω_k για τα οποία ο βώλος k είναι κόκκινος. Τότε το A έχει ακριβώς r σημεία. Λέμε ότι το A είναι ένα ενδεχόμενο. Πιό γενικά, στο μοντέλο που συζητάμε θα λέμε ενδεχόμενο κάθε υποσύνολο του Ω. Όταν λέμε ότι το ενδεχόμενο B συμβαίνει εννοούμε ότι το αποτέλεσμα του πειράματος αναπαρίσταται από κάποιο σημείο του B.

Έστω A και B δύο ενδεχόμενα. Θυμηθείτε ότι η ένωση $A \cup B$ των A και B είναι το σύνολο όλων των σημείων $\omega \in \Omega$ για τα οποία είτε $\omega \in A$ είτε $\omega \in B$. Τα σημεία του Ω βρίσκονται σε αντιστοιχία με τα αποτελέσματα του πειράματός μας. Το ενδεχόμενο A συμβαίνει αν το πείραμα δίνει ένα αποτέλεσμα που αναπαρίσταται από κάποιο σημείο του A, και όμοια το ενδεχόμενο B συμβαίνει αν το αποτέλεσμα του πειράματος αναπαρίσταται από κάποιο σημείο του B. Το σύνολο $A \cup B$ αναπαριστά λοιπόν το γεγονός ότι είτε το ενδεχόμενο A συμβαίνει είτε το ενδεχόμενο B συμβαίνει. Όμοια, η τομή $A \cap B$ των A και B αποτελείται από όλα τα σημεία που ανήκουν και στο A και στο B. Έτσι, αν $\omega \in A \cap B$ τότε $\omega \in A$ και $\omega \in B$, δηλαδή το $A \cap B$ αναπαριστά το γεγονός ότι τα ενδεχόμενα A και B συμβαίνουν και τα δύο. Το συμπλήρωμα A^c (ή A') του A είναι το σύνολο των σημείων του Ω που δεν ανήκουν στο A. Το ενδεχόμενο A δεν συμβαίνει αν το πείραμα δίνει αποτέλεσμα που αναπαρίσταται από σημείο του A^c.

Σχηματικά, αν τα A και B αναπαρίστανται από τα χωρία στο Σχήμα 1α, τότε τα $A \cup B$, $A \cap B$, και A^c αναπαρίστανται από τα σκιασμένα χωρία στα Σχήματα 1β, 1γ και 1δ, αντίστοιχα.

Για να εξοικειωθούμε με αυτές τις έννοιες, ας πάρουμε σαν A το ενδεχόμενο να «επιλέχθηκε κόκκινος βώλος» και σαν B το ενδεχόμενο να «επιλέχθηκε βώλος με άρτιο αριθμό». Τότε, η ένωση $A \cup B$ είναι το ενδεχόμενο να επιλέχθηκε είτε κόκκινος βώλος είτε βώλος με άρτιο αριθμό. Η τομή $A \cap B$ είναι το ενδεχόμενο να επιλέχθηκε κόκκινος βώλος με άρτιο αριθμό. Το ενδεχόμενο A^c συμβαίνει αν δεν επιλέχθηκε κόκκινος βώλος.

Θέλουμε τώρα να αντιστοιχίσουμε πιθανότητες στα ενδεχόμενα. Από μαθηματική άποψη, αυτό σημαίνει απλώς ότι αντιστοιχίζουμε σε κάθε σύνολο B έναν πραγματικό αριθμό. Κατ' αρχήν θα μπορούσαμε να το κάνουμε με αυθαίρετο τρόπο. Αν όμως θέλουμε αυτές οι πιθανότητες να περιγράφουν σωστά το πείραμα που προσπαθούμε να μοντελοποιήσουμε, τότε μπαίνουν διάφοροι περιορισμοί. Πώς πρέπει να γίνει η αντιστοίχιση; Έχουμε ήδη αντιστοιχίσει σε κάθε σημείο τον αριθμό $1/s$. Έτσι, σε κάθε μονοσύνολο $\{\omega\}$ πρέπει να αντιστοιχεί ο αριθμός $1/s$. Από τη συζήτηση που κάναμε για τη σχετική συχνότητα του ενδεχομένου «επιλογή κόκκινου βώλου», φαίνεται ότι θα πρέπει στο ενδεχόμενο A να αντιστοιχίσουμε την πιθανότητα $P(A) = r/s$. Πιό γενικά, αν B είναι οποιοδήποτε ενδεχόμενο, ορίζουμε την $P(B)$ θέτοντας $P(B) = j/s$ αν το B έχει ακριβώς j σημεία. Παρατηρούμε τότε ότι

$$P(B) = \sum_{\omega_k \in B} p_k,$$

όπου $\sum_{\omega_k \in B} p_k$ σημαίνει ότι αθροίζουμε τους αριθμούς p_k για εκείνες τις τιμές του k που αντιστοιχούν σε κάποιο $\omega_k \in B$. Από τον ορισμό της $P(B)$ βλέπουμε εύκολα ότι ισχύουν οι ακόλουθες προτάσεις (η επαλήθευσή τους αφήνεται στον αναγνώστη).

Έστω $\emptyset$ το κενό σύνολο. Τότε $P(\emptyset) = 0$ και $P(\Omega) = 1$. Αν A και B είναι δύο ξένα σύνολα, δηλαδή $A \cap B = \emptyset$, τότε

$$P(A \cup B) = P(A) + P(B).$$

Παράδειγμα 2. Από φυσικά πειράματα είναι γνωστό ότι ένα ισότοπο κάποιου ραδιενεργού στοιχείου είναι ασταθές. Με την πάροδο του χρόνου αποβάλλει νετρόνια και έρχεται σε σταθερή κατάσταση. Ενδιαφερόμαστε για το χρόνο που χρειάζεται για να διασπαστεί ένα άτομο του ισοτόπου. Σύμφωνα με τους νόμους της φυσικής είναι αδύνατο να πούμε με βεβαιότητα πότε κάποιο συγκεκριμένο άτομο του ισοτόπου θα διασπαστεί, αν όμως παρατηρήσουμε ένα μεγάλο πλήθος N ατόμων, τότε μπορούμε να κάνουμε κάποιες ακριβείς προβλέψεις για το πλήθος $N(t)$ των ατόμων που δεν έχουν διασπαστεί ως την χρονική στιγμή t. Με άλλα λόγια μπορούμε να προβλέψουμε με ακρίβεια τον λόγο των ατόμων $N(t)/N$ που δεν έχουν διασπαστεί ως την χρονική στιγμή t, αλλά δεν μπορούμε να πούμε ποιά είναι αυτά τα άτομα. Αφού όλα τα άτομα είναι ίδια, η ταυτόχρονη παρατήρηση N ατόμων πρέπει να θεωρηθεί ισοδύναμη με N επαναλήψεις του ίδιου πειράματος, όπου, σε αυτήν την περίπτωση, το πείραμα είναι η παρατήρηση του χρόνου που απαιτείται για την διάσπαση ενός ατόμου.

Κατά προσέγγιση (η οποία μάλιστα είναι πολύ ακριβής), ο ρυθμός με τον οποίο διασπάται το ισότοπο τη χρονική στιγμή t είναι ανάλογος με το πλήθος των ατόμων που υπάρχουν τη χρονική στιγμή t, άρα το $N(t)$ δίνεται προσεγγιστικά από την λύση της διαφορικής εξίσωσης

$$\frac{df}{dt} = -\lambda f(t), \quad f(0) = N,$$

όπου $\lambda > 0$ είναι μία θετική σταθερά αναλογίας. Η μοναδική λύση αυτής της εξίσωσης είναι η $f(t) = Ne^{-\lambda t}$, επομένως ο λόγος των ατόμων που δεν έχουν διασπαστεί ως την στιγμή t δίνεται κατά προσέγγιση από το $N(t)/N = e^{-\lambda t}$. Αν $0 \leq t_0 \leq t_1$, ο λόγος των ατόμων που διασπώνται στο χρονικό διάστημα $[t_0, t_1]$ είναι $(e^{-\lambda t_0} - e^{-\lambda t_1})$. Επομένως, σύμφωνα με την ερμηνεία της πιθανότητας ως σχετικής συχνότητας παίρνουμε το $(e^{-\lambda t_0} - e^{-\lambda t_1})$ σαν την πιθανότητα να διασπαστεί ένα άτομο ανάμεσα στις χρονικές στιγμές t_0 και t_1.

Για να κατασκευάσουμε ένα μαθηματικό μοντέλο για το πείραμα, μπορούμε να προχωρήσουμε όπως στο προηγούμενο παράδειγμα. Πρώτα επιλέγουμε ένα σύνολο Ω που μπορεί να τεθεί σε ένα προς ένα αντιστοιχία με τα δυνατά αποτελέσματα του πειράματος. Αποτέλεσμα σε αυτήν την περίπτωση είναι ο χρόνος που απαιτείται για τη σταθεροποίηση ενός ατόμου. Αυτός μπορεί να είναι οποιοσδήποτε θετικός πραγματικός αριθμός, παίρνουμε λοιπόν σαν Ω το διάστημα $[0, \infty)$ στην πραγματική ευθεία. Από τη συζήτηση που προηγήθηκε φαίνεται λογικό να αντιστοιχίσουμε στο διάστημα $[t_0, t_1]$ την πιθανότητα $(e^{-\lambda t_0} - e^{-\lambda t_1})$. Ειδικότερα, αν $t_0 = t_1 = t$, τότε το διάστημα εκφυλίζεται στο σύνολο $\{t\}$ και η πιθανότητα που αντιστοιχεί σε αυτό το σύνολο είναι 0.

Στο προηγούμενο παράδειγμά μας, το Ω είχε πεπερασμένα το πλήθος σημεία. Εδώ όμως, το Ω έχει υπεραριθμήσιμα άπειρα το πλήθος σημεία και κάθε σημείο έχει πιθανότητα 0. Για μία ακόμα φορά παρατηρούμε ότι $P(\Omega) = 1$ και $P(\emptyset) = 0$. Ας υποθέσουμε ότι A και B είναι δύο ξένα διαστήματα. Τότε ο λόγος των ατόμων που διασπώνται στο χρονικό διάστημα $A \cup B$ είναι το άθροισμα του λόγου αυτών που διασπώνται στο χρονικό διάστημα A και του λόγου αυτών που διασπώνται στο χρονικό διάστημα B. Παρατηρώντας αυτήν την προσθετικότητα, ζητάμε στο μαθηματικό μας μοντέλο το $A \cup B$ να έχει πιθανότητα $P(A) + P(B)$. Με άλλα λόγια, στο μαθηματικό μας μοντέλο θέλουμε να ισχύει η

$$P(A \cup B) = P(A) + P(B)$$

αν A και B είναι δύο ξένα διαστήματα.

1.2 Χώροι πιθανότητας

Σκοπός μας σε αυτήν την παράγραφο είναι να ορίσουμε την τυπική μαθηματική δομή, που λέγεται χώρος πιθανότητας, η οποία θα αποτελέσει τη βάση για τη μαθηματική μελέτη των τυχαίων φαινομένων.

Φανταστείτε κάποιο πραγματικό ή φανταστικό πείραμα το οποίο προσπαθούμε να μοντελοποιήσουμε. Το πρώτο πράγμα που έχουμε να κάνουμε είναι να αποφασίσουμε τα δυνατά αποτελέσματα του πειράματος. Δεν έχει ιδιαίτερη σημασία αν στη μελέτη μας θα αποδεχτούμε περισσότερα πράγματα από αυτά που μπορεί πραγματικά να συμβούν, θέλουμε όμως να εξασφαλίσουμε ότι δεν θα αποκλειστούν πράγματα που θα μπορούσε να συμβούν. Από τη στιγμή που θα αποφασίσουμε για τα δυνατά αποτελέσματα, επιλέγουμε ένα σύνολο Ω του οποίου τα σημεία ω σχετίζονται με αυτά τα αποτελέσματα. Από την καθαρά μαθηματική όμως άποψη, το Ω είναι απλώς ένα αφηρημένο σύνολο σημείων.

Στη συνέχεια παίρνουμε μία μη κενή συλλογή $\mathcal{A}$ υποσυνόλων του Ω η οποία πρόκειται να αναπαραστήσει την συλλογή των «ενδεχομένων» στα οποία επιθυμούμε να αντιστοιχίσουμε πιθανότητες. Εξ ορισμού τώρα, με τον όρο *ενδεχόμενο* εννοούμε ένα σύνολο A στην $\mathcal{A}$. Η πρόταση *το ενδεχόμενο A συμβαίνει* σημαίνει ότι το αποτέλεσμα του πειράματός μας αναπαρίσταται από κάποιο σημείο $\omega \in A$. Πάλι, από την καθαρά μαθηματική άποψη, το $\mathcal{A}$ είναι απλώς μια καθορισμένη συλλογή υποσυνόλων του Ω. Μόνο σε σύνολα $A \in \mathcal{A}$, δηλαδή σε ενδεχόμενα, αντιστοιχίζονται πιθανότητες. Στο μοντέλο του Παραδείγματος 1, η $\mathcal{A}$ αποτελείται από όλα τα υποσύνολα του Ω. Στη γενική περίπτωση που το Ω δεν έχει πεπερασμένα το πλήθος σημεία, όπως στο Παράδειγμα 2, μπορεί να μην είναι δυνατόν να επιλέξουμε την $\mathcal{A}$ κατ' αυτόν τον τρόπο.

Το επόμενο ερώτημα είναι, ποιά πρέπει να είναι η συλλογή $\mathcal{A}$. Είναι πολύ λογικό να απαιτήσουμε η $\mathcal{A}$ να είναι κλειστή ως προς πεπερασμένες ενώσεις και πεπερασμένες τομές, καθώς και ως προς συμπληρώματα. Για παράδειγμα, αν A και B είναι δύο ενδεχόμενα, τότε το $A \cup B$ συμβαίνει αν το αποτέλεσμα του πειράματός μας αναπαρίσταται είτε από σημείο στο A είτε από σημείο στο B. Προφανώς, τότε, αν έχει νόημα να συζητάμε για τις πιθανότητες να συμβαίνουν τα A και B, πρέπει να έχει νόημα να συζητάμε και για την πιθανότητα να συμβαίνει είτε το A είτε το B, δηλαδή να συμβαίνει το ενδεχόμενο $A \cup B$. Αφού μιλάμε μόνο για τις πιθανότητες των συνόλων που ανήκουν στην $\mathcal{A}$, είμαστε υποχρεωμένοι να απαιτήσουμε ότι το $A \cup B$ ανήκει στην $\mathcal{A}$ αν τα A και B είναι στοιχεία της $\mathcal{A}$. Τώρα, το $A \cap B$ συμβαίνει αν το αποτέλεσμα του πειράματός μας αναπαρίσταται από κάποιο σημείο που ανήκει τόσο στο A όσο και στο B. Ανάλογος συλλογισμός με αυτόν που χρησιμοποιήσαμε για το $A \cup B$ μας πείθει ότι πρέπει να έχουμε $A \cap B \in \mathcal{A}$ οποτεδήποτε $A, B \in \mathcal{A}$. Τέλος, λέγοντας ότι το ενδεχόμενο A δεν συμβαίνει εννοούμε ότι το αποτέλεσμα του πειράματός μας δεν αναπαρίσταται από σημείο στο A, επομένως αναπαρίσταται από κάποιο σημείο στο A^c. Θα ήταν αρκετά περίεργο να λέμε ότι επιτρέπεται να μιλάμε για την πιθανότητα του A αλλά όχι του A^c. Θα ζητάμε λοιπόν αν το A είναι στην $\mathcal{A}$ να είναι και το A^c.

Έχουμε έτσι καταλήξει στο συμπέρασμα ότι η $\mathcal{A}$ πρέπει να είναι μία μη κενή συλλογή υποσυνόλων του Ω που έχει τις εξής ιδιότητες:

(i) Αν το A ανήκει στην $\mathcal{A}$, το ίδιο ισχύει για το A^c.

(ii) Αν τα A και B ανήκουν στην $\mathcal{A}$, το ίδιο ισχύει για τα $A \cup B$ και $A \cap B$.

Ένα απλό επαγωγικό επιχείρημα δείχνει ότι αν $A_1, A_2, \ldots, A_n$ είναι σύνολα στην $\mathcal{A}$, τότε το ίδιο ισχύει και για τα $\bigcup_{i=1}^n A_i$ και $\bigcap_{i=1}^n A_i$. Εδώ χρησιμοποιούμε τις συντμήσεις

$$\bigcup_{i=1}^n A_i = A_1 \cup A_2 \cup \ldots \cup A_n$$

και

$$\bigcap_{i=1}^n A_i = A_1 \cap A_2 \cap \ldots \cap A_n.$$

Επίσης, αφού $A \cap A^c = \emptyset$ και $A \cup A^c = \Omega$, βλέπουμε ότι τόσο το κενό σύνολο $\emptyset$ όσο και το σύνολο Ω πρέπει να ανήκουν στην $\mathcal{A}$.

Μία μη κενή συλλογή υποσυνόλων ενός συνόλου Ω που είναι κλειστή ως προς τις πεπερασμένες συνολοθεωρητικές πράξεις λέγεται *άλγεβρα υποσυνόλων* του Ω. Μοιάζει λοιπόν λογικό να απαιτήσουμε ότι η $\mathcal{A}$ είναι μια άλγεβρα υποσυνόλων του Ω. Για μαθηματικούς όμως λόγους, το να απαιτήσουμε η $\mathcal{A}$ να είναι απλώς άλγεβρα υποσυνόλων του Ω είναι ανεπαρκές. Αυτό που θα απαιτήσουμε από την οικογένεια $\mathcal{A}$ είναι ισχυρότερο. Θα ζητήσουμε η $\mathcal{A}$ να είναι κλειστή όχι μόνο ως προς πεπερασμένες συνολοθεωρητικές πράξεις αλλά και ως προς άπειρες αριθμήσιμες συνολοθεωρητικές πράξεις. Με άλλα λόγια, αν $\{A_n\}$, $n \geq 1$, είναι μια ακολουθία συνόλων στην $\mathcal{A}$, απαιτούμε να ισχύουν οι

$$\bigcup_{n=1}^\infty A_n \in \mathcal{A} \quad , \quad \bigcap_{n=1}^\infty A_n \in \mathcal{A}.$$

Εδώ χρησιμοποιούμε τη σύντμηση

$$\bigcup_{n=1}^\infty A_n = A_1 \cup A_2 \cup \ldots$$

για την ένωση όλων των συνόλων της ακολουθίας, και την

$$\bigcap_{n=1}^{\infty} A_n = A_1 \cap A_2 \cap \dots$$

για την τομή όλων των συνόλων της ακολουθίας. Μία συλλογή υποσυνόλων ενός συνόλου Ω που είναι κλειστή ως προς αριθμήσιμες πράξεις συνόλων λέγεται σ-άλγεβρα υποσυνόλων του Ω. (Το γράμμα σ μπαίνει μπροστά για να διακρίνει μία τέτοια συλλογή από την άλγεβρα υποσυνόλων.) Πιό τυπικά έχουμε τον εξής ορισμό:

Ορισμός 1. *Μία μη κενή συλλογή $\mathcal{A}$ υποσυνόλων ενός συνόλου Ω λέγεται σ-άλγεβρα υποσυνόλων του Ω αν ισχύουν οι παρακάτω δύο ιδιότητες:*

(i) *Αν το A ανήκει στην $\mathcal{A}$, τότε το A^c είναι στην $\mathcal{A}$.*

(ii) *Αν το A_n ανήκει στην $\mathcal{A}$, $n = 1, 2, \dots$, τότε τα $\bigcup_{n=1}^{\infty} A_n$ και $\bigcap_{n=1}^{\infty} A_n$ είναι και τα δύο στην $\mathcal{A}$.*

Ερχόμαστε τώρα στην απόδοση πιθανοτήτων στα ενδεχόμενα. Όπως έγινε φανερό από τα παραδείγματα της προηγούμενης παραγράφου, η πιθανότητα ενός ενδεχομένου είναι ένας μη αρνητικός πραγματικός αριθμός. Για κάθε ενδεχόμενο A, θα συμβολίζουμε αυτόν τον αριθμό με $P(A)$. Τότε $0 \leq P(A) \leq 1$. Στο σύνολο Ω το οποίο αναπαριστά κάθε δυνατό αποτέλεσμα, θα πρέπει προφανώς να αποδοθεί ο αριθμός 1, δηλαδή $P(\Omega) = 1$. Στη συζήτηση του Παραδείγματος 1 δείξαμε ότι η πιθανότητα των ενδεχομένων έχει την ιδιότητα αν A και B είναι δύο ξένα ενδεχόμενα τότε $P(A \cup B) = P(A) + P(B)$. Όμοια, στο Παράδειγμα 2 δείξαμε ότι αν A και B είναι δύο ξένα διαστήματα, τότε θα έπρεπε πάλι να ζητήσουμε την

$$P(A \cup B) = P(A) + P(B).$$

Μοιάζει λοιπόν λογικό να απαιτήσουμε γενικά ότι αν A και B είναι ξένα ενδεχόμενα τότε $P(A \cup B) = P(A) + P(B)$. Με επαγωγή, θα προέκυπτε τότε ότι αν $A_1, A_2, \dots, A_n$ είναι n ξένα ανά δύο σύνολα (δηλαδή, αν $A_i \cap A_j = \emptyset$ οποτεδήποτε $i \neq j$), τότε

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i).$$

Για την ακρίβεια, πάλι για μαθηματικούς λόγους, θα απαιτήσουμε αυτή η προσθετική ιδιότητα να ισχύει για αριθμήσιμες συλλογές ξένων ενδεχομένων.

Ορισμός 2. *Ένα μέτρο πιθανότητας P σε μία σ-άλγεβρα $\mathcal{A}$ υποσυνόλων ενός συνόλου Ω είναι μια πραγματική συνάρτηση με πεδίο ορισμού την $\mathcal{A}$, που έχει τις εξής ιδιότητες:*

(i) $P(\Omega) = 1$.

(ii) $P(A) \geq 0$ *για κάθε* $A \in \mathcal{A}$.

(iii) *Αν* A_n, $n = 1, 2, 3, \ldots$, *είναι ξένα ανά δύο σύνολα στην* $\mathcal{A}$, *τότε*

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n).$$

Ένας χώρος πιθανότητας $(\Omega, \mathcal{A}, P)$ *είναι μία τριάδα που αποτελείται από ένα σύνολο* Ω, *μία* σ*-άλγεβρα* $\mathcal{A}$ *υποσυνόλων του* Ω, *και ένα μέτρο πιθανότητας ορισμένο στην* $\mathcal{A}$.

Είναι πολύ εύκολο να βρούμε έναν χώρο πιθανότητας που να αντιστοιχεί στο πείραμα της επιλογής ενός βώλου από ένα δοχείο. Ουσιαστικά τον περιγράψαμε όταν συζητούσαμε αυτό το πείραμα. Απλώς παίρνουμε σαν Ω ένα πεπερασμένο σύνολο με s στοιχεία, σαν $\mathcal{A}$ τη συλλογή όλων των υποσυνόλων του Ω, και σαν P το μέτρο πιθανότητας που αντιστοιχεί στο A την πιθανότητα $P(A) = j/s$ αν το A έχει ακριβώς j στοιχεία.

Ας δούμε τώρα τον χώρο πιθανότητας που αντιστοιχεί στο πείραμα της διάσπασης του ισοτόπου (Παράδειγμα 2). Εδώ είναι βέβαια φανερό ότι $\Omega = [0, \infty)$, δεν είναι όμως καθόλου προφανές ποιά πρέπει να είναι τα $\mathcal{A}$ και P. Πράγματι, όπως θα δούμε πιό κάτω, αυτό δεν είναι καθόλου τετριμμένο πρόβλημα, και η πλήρης επίλυσή του απαιτεί κάποια βαθύτερα αποτελέσματα της θεωρίας συνόλων που είναι έξω από τους σκοπούς αυτού του βιβλίου.

Ένα όμως πράγμα είναι σαφές: όπως κι αν επιλέξουμε τα $\mathcal{A}$ και P, η $\mathcal{A}$ πρέπει να περιέχει όλα τα διαστήματα, και το P πρέπει να αντιστοιχίζει πιθανότητα $(e^{-\lambda t_0} - e^{-\lambda t_1})$ στο διάστημα $[t_0, t_1]$ αν θέλουμε ο χώρος πιθανότητας που κατασκευάζουμε να αντικατοπτρίζει την φυσική κατάσταση. Το πρόβλημα της κατασκευής του χώρου διατυπώνεται λοιπόν αυστηρά ως εξής. Υπάρχουν σ-άλγεβρα $\mathcal{A}$ που περιέχει όλα τα διαστήματα και μέτρο πιθανότητας P ορισμένο στην $\mathcal{A}$ που σε κάθε διάστημα A να αντιστοιχίζει την παραπάνω πιθανότητα $P(A)$; Ερωτήματα αυτού του είδους ανήκουν στην επικράτεια ενός κλάδου των ανώτερων μαθηματικών που λέγεται θεωρία μέτρου, και δεν μπορούμε να τα αντιμετωπίσουμε επαρκώς μένοντας στο επίπεδο αυτού του βιβλίου. Αποτελέσματα από τη θεωρία μέτρου δείχνουν ότι η απάντηση στο συγκεκριμένο ερώτημα, καθώς και σε άλλα παρόμοιας υφής, είναι ναί, οπότε δεχόμαστε ότι κατασκευές τέτοιου είδους είναι πάντα δυνατές.

Γενικά, δεν θα εμβαθύνουμε στην κατασκευή χώρων πιθανότητας. Η μαθηματική θεωρία πιθανοτήτων αρχίζει με έναν αφηρημένο χώρο πιθανότητας και αναπτύσσει τη θεωρία χρησιμοποιώντας τον χώρο πιθανότητας σαν πλαίσιο λειτουργίας. Αν όμως εξαιρέσουμε τον ρόλο του σαν πλαίσιο για τον αυστηρό ορισμό των άλλων εννοιών της θεωρίας, ο χώρος πιθανότητας από μόνος του δεν παίζει μεγάλο ρόλο στην παραπέρα ανάπτυξη του αντικειμένου. Βοηθητικές ποσότητες (ειδικά οι τυχαίες μεταβλητές, μια έννοια που θα συναντήσουμε στο Κεφάλαιο 3) γίνονται πολύ σύντομα το κυρίαρχο θέμα, και ο χώρος πιθανότητας περνάει στο περιθώριο.

Θα ολοκληρώσουμε τη συζήτηση πάνω στους χώρους πιθανότητας, κατασκευάζοντας μία σημαντική κλάση χώρων πιθανότητας, τους *ομοιόμορφους χώρους πιθανότητας*.

Μερικά από τα πιό παλιά προβλήματα των πιθανοτήτων αφορούν την «τυχαία» επιλογή ενός σημείου από ένα σύνολο S. Η διαίσθησή μας υπαγορεύει ότι αν A και B

είναι δύο σύνολα που έχουν το ίδιο «μέγεθος» τότε η πιθανότητα να επιλέξουμε σημείο του A είναι η ίδια με την πιθανότητα να επιλέξουμε σημείο του B. Αν το S έχει πεπερασμένα το πλήθος σημεία, μπορούμε να μετρήσουμε το «μέγεθος» ενός συνόλου με τον πληθάριθμό του. Δύο σύνολα έχουν λοιπόν το ίδιο «μέγεθος» αν έχουν το ίδιο πλήθος σημείων. Είναι πολύ εύκολο να φτιάξουμε έναν χώρο πιθανότητας που αντιστοιχεί στο πείραμα της τυχαίας επιλογής ενός σημείου από ένα σύνολο S με s στοιχεία. Παίρνουμε $\Omega = S$, $\mathcal{A}$ όλα τα υποσύνολα του S, και αντιστοιχίζουμε στο σύνολο A την πιθανότητα $P(A) = j/s$ αν το A έχει ακριβώς j σημεία. Ένας τέτοιος χώρος πιθανότητας λέγεται *συμμετρικός χώρος πιθανότητας* γιατί κάθε μονοσύνολο έχει την ίδια πιθανότητα s^{-1}. Στο Κεφάλαιο 2 θα επιστρέψουμε στη μελέτη αυτών των χώρων.

Ας υποθέσουμε τώρα ότι S είναι το διάστημα $[a, b]$ στην πραγματική ευθεία, όπου $-\infty < a < b < +\infty$. Στην περίπτωση αυτή φαίνεται λογικό να μετρήσουμε το «μέγεθος» ενός υποσυνόλου A του $[a, b]$ με βάση το μήκος του. Δύο σύνολα έχουν το ίδιο μέγεθος αν έχουν το ίδιο μήκος. Θα συμβολίζουμε το μήκος ενός συνόλου A με $|A|$.

Για να κατασκευάσουμε έναν χώρο πιθανότητας για το πείραμα της «τυχαίας επιλογής ενός σημείου από το S», προχωράμε όπως στην περίπτωση του πειράματος του ισοτόπου. Παίρνουμε $\Omega = S$, και χρησιμοποιώντας αποτελέσματα της θεωρίας μέτρου βλέπουμε ότι υπάρχουν μία σ-άλγεβρα $\mathcal{A}$ υποσυνόλων του S, και ένα μέτρο πιθανότητας P ορισμένο στην $\mathcal{A}$ ώστε $P(A) = |A|/|S|$ αν το A είναι διάστημα.

Πιό γενικά, έστω S ένα υποσύνολο του r-διάστατου Ευκλείδειου χώρου, που έχει πεπερασμένο, μη μηδενικό r-διάστατο όγκο. Αν A είναι ένα υποσύνολο του S, συμβολίζουμε τον όγκο του A με $|A|$. Υπάρχουν τότε μία σ-άλγεβρα $\mathcal{A}$ υποσυνόλων του S που περιέχει όλα τα υποσύνολα του S των οποίων ορίζεται ο όγκος (στο πλαίσιο του απειροστικού λογισμού), και ένα μέτρο πιθανότητας P ορισμένο στην $\mathcal{A}$ ώστε $P(A) = |A|/|S|$ για κάθε τέτοιο σύνολο A. Ένας χώρος πιθανότητας αυτού του είδους συμβολίζεται με $(S, \mathcal{A}, P)$, και λέγεται *ομοιόμορφος χώρος πιθανότητας*.

1.3 Ιδιότητες των πιθανοτήτων

Σε αυτήν την παράγραφο, αποδεικνύουμε κάποιες επιπλέον ιδιότητες του μέτρου πιθανότητας P που προκύπτουν από τον ορισμό του μέτρου πιθανότητας. Οι ιδιότητες αυτές θα χρησιμοποιούνται συνεχώς στο υπόλοιπο αυτού του βιβλίου. Υποθέτουμε ότι μας έχει δοθεί ένας χώρος πιθανότητας $(\Omega, \mathcal{A}, P)$ και ότι όλα τα σύνολα που συζητάμε είναι ενδεχόμενα, δηλαδή στοιχεία της $\mathcal{A}$.

Για κάθε σύνολο A ισχύει $A \cup A^c = \Omega$, άρα για τυχόντα σύνολα A και B έχουμε τη διάσπαση του B

$$B = \Omega \cap B = (A \cup A^c) \cap B = (A \cap B) \cup (A^c \cap B). \tag{1}$$

Αφού τα $A \cap B$ και $A^c \cap B$ είναι ξένα, από την (iii) του Ορισμού 2 βλέπουμε ότι

$$P(B) = P(A \cap B) + P(A^c \cap B). \tag{2}$$

Θέτοντας $B = \Omega$ και χρησιμοποιώντας το γεγονός ότι $P(\Omega) = 1$, συμπεραίνουμε από την (2) ότι

$$P(A^c) = 1 - P(A). \tag{3}$$

Ειδικότερα $P(\emptyset) = 1 - P(\Omega)$, άρα

$$P(\emptyset) = 0. \tag{4}$$

Για μία δεύτερη εφαρμογή της (2) ας υποθέσουμε ότι $A \subset B$. Τότε $A \cap B = A$, επομένως

$$P(B) = P(A) + P(A^c \cap B), \qquad \alpha\nu\ A \subset B. \tag{5}$$

Αφού $P(A^c \cap B) \geq 0$ από την (ii), βλέπουμε από την (5) ότι

$$P(B) \geq P(A) \quad \alpha\nu \quad A \subset B. \tag{6}$$

Οι νόμοι του De Morgan ισχυρίζονται ότι αν $\{A_n\}$, $n \geq 1$, είναι μια ακολουθία συνόλων, τότε

$$\left(\bigcup_n A_n\right)^c = \left(\bigcap_n A_n^c\right) \tag{7}$$

και

$$\left(\bigcap_n A_n\right)^c = \left(\bigcup_n A_n^c\right). \tag{8}$$

Για να δούμε ότι ισχύει η (7), παρατηρούμε ότι $\omega \in \left(\bigcup_{n\geq 1} A_n\right)^c$ αν και μόνο αν $\omega \notin A_n$ για κάθε n. Δηλαδή, αν $\omega \in A_n^c$ για κάθε $n \geq 1$, ή ισοδύναμα, αν $\omega \in \bigcap_n A_n^c$. Για να αποδείξουμε την (8) εφαρμόζουμε την (7) για την $\{A_n^c\}$ και βλέπουμε ότι

$$\left(\bigcup_n A_n^c\right)^c = \bigcap_n A_n,$$

οπότε, παίρνοντας συμπληρώματα, καταλήγουμε στην

$$\bigcup_n A_n^c = \left(\bigcap_n A_n\right)^c.$$

Μία χρήσιμη σχέση που έπεται από τις (7) και (3) είναι η

$$P\left(\bigcup_n A_n\right) = 1 - P\left(\bigcap_n A_n^c\right). \tag{9}$$

Τώρα, $\bigcup_n A_n$ είναι το ενδεχόμενο να συμβαίνει τουλάχιστον ένα από τα ενδεχόμενα A_n, ενώ $\bigcap_n A_n^c$ είναι το ενδεχόμενο να μην συμβαίνει κανένα από αυτά. Με λόγια, η (9) ισχυρίζεται ότι η πιθανότητα να συμβαίνει τουλάχιστον ένα από τα ενδεχόμενα A_n είναι 1 μείον την πιθανότητα να μην συμβαίνει κανένα από τα ενδεχόμενα A_n. Το πλεονέκτημα της (9) είναι ότι σε ορισμένες περιπτώσεις είναι ευκολότερο να υπολογίσουμε την $P\left(\bigcap_n A_n^c\right)$ παρά να υπολογίσουμε την $P\left(\bigcup_n A_n\right)$. [Παρατηρήστε ότι αφού τα ενδεχόμενα A_n δεν είναι αναγκαστικά ξένα, δεν ισχύει ότι $P\left(\bigcup_n A_n\right) = \sum_n P(A_n)$.] Η χρήση της (9) περιγράφεται υποδειγματικά στο παράδειγμα που ακολουθεί.

Παράδειγμα 3. Ρίχνουμε τρία αμερόληπτα, όμοια νομίσματα. Να βρεθεί η πιθανότητα να εμφανιστούν τουλάχιστον μία φορά γράμματα.

Υπάρχουν οκτώ δυνατά αποτελέσματα για το πείραμα:

Η διαίσθησή μας υποδεικνύει ότι καθένα από αυτά τα οκτώ αποτελέσματα πρέπει να έχει πιθανότητα 1/8. Έστω A_1 το ενδεχόμενο το πρώτο νόμισμα να εμφανίσει γράμματα, A_2 το ενδεχόμενο το δεύτερο νόμισμα να εμφανίσει γράμματα, και A_3 το ενδεχόμενο το τρίτο νόμισμα να εμφανίσει γράμματα. Το πρόβλημα μας ζητάει να υπολογίσουμε την $P(A_1 \cup A_2 \cup A_3)$. Όμως, $A_1^c \cap A_2^c \cap A_3^c = \{(K, K, K)\}$ άρα

$$P(A_1^c \cap A_2^c \cap A_3^c) = 1/8.$$

Από την (9) συμπεραίνουμε ότι

$$P(A_1 \cup A_2 \cup A_3) = 1 - P(A_1^c \cap A_2^c \cap A_3^c) = 7/8.$$

Το αξίωμα (iii) των μέτρων πιθανότητας μάς λέει ότι αν τα σύνολα A και B είναι ξένα, τότε $P(A \cup B) = P(A) + P(B)$. Αν τα A και B δεν είναι αναγκαστικά ξένα, τότε

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{10}$$

επομένως

$$P(A \cup B) \leq P(A) + P(B). \tag{11}$$

Για να δούμε ότι ισχύει η (10) παρατηρούμε ότι τα σύνολα $A \cap B^c, A \cap B$, και $A^c \cap B$ είναι ξένα ανά δύο και η ένωσή τους είναι το $A \cup B$ (δείτε το Σχήμα 2). Άρα,

$$P(A \cup B) = P(A \cap B^c) + P(A^c \cap B) + P(A \cap B). \tag{12}$$

Όμως, από την (2),

$$P(A \cap B^c) = P(A) - P(A \cap B)$$

και

$$P(A^c \cap B) = P(B) - P(A \cap B).$$

Αντικαθιστώντας στην (12), παίρνουμε την (10).

Οι ισότητες (10) και (11) επεκτείνονται για αυθαίρετο πεπερασμένο πλήθος συνόλων. Η γενίκευση του τύπου (10) είναι αρκετά πολύπλοκη και θα την συζητήσουμε στο Κεφάλαιο 2. Η ανισότητα (11), όμως, επεκτείνεται εύκολα με επαγωγή και έχουμε

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) \leq \sum_{i=1}^{n} P(A_i). \tag{13}$$

Για να το αποδείξουμε, παρατηρούμε ότι αν $n \geq 2$, τότε από την (11)

$$\begin{aligned} P(A_1 \cup \ldots \cup A_n) &= P((A_1 \cup \ldots \cup A_{n-1}) \cup A_n) \\ &\leq P(A_1 \cup \ldots \cup A_{n-1}) + P(A_n). \end{aligned}$$

Επομένως, αν ισχύει η (13) για $n-1$ σύνολα, τότε ισχύει για n σύνολα. Αφού η (13) προφανώς ισχύει για $n = 1$, το αποτέλεσμα αποδεικνύεται με επαγωγή.

Μέχρι τώρα χρησιμοποιήσαμε μόνο το γεγονός ότι ένα μέτρο πιθανότητας είναι πεπερασμένα προσθετικό. Στο επόμενο αποτέλεσμά μας θα χρησιμοποιήσουμε την αριθμήσιμη προσθετικότητα.

Θεώρημα 1. *Θεωρούμε τα ενδεχόμενα A_n, $n \geq 1$.*

(i) *Αν $A_1 \subset A_2 \subset \ldots$ και $A = \bigcup_{n=1}^{\infty} A_n$, τότε*

$$\lim_{n\to\infty} P(A_n) = P(A). \tag{14}$$

(ii) *Αν $A_1 \supset A_2 \supset \ldots$ και $A = \bigcap_{n=1}^{\infty} A_n$, τότε ισχύει πάλι η (14).*

Απόδειξη του (i). Υποθέτουμε ότι $A_1 \subset A_2 \subset \ldots$ και $A = \bigcup_{n=1}^{\infty} A_n$. Θέτουμε $B_1 = A_1$, και για κάθε $n \geq 2$, συμβολίζουμε με B_n το σύνολο των σημείων που ανήκουν στο A_n αλλά όχι στο A_{n-1}, δηλαδή $B_n = A_n \cap A_{n-1}^c$. Ένα σημείο ω ανήκει στο B_n αν και μόνο αν το ω είναι στο A και το A_n είναι το πρώτο σύνολο της ακολουθίας $A_1, A_2, \ldots$ που περιέχει το ω. Από τον ορισμό τους, τα σύνολα B_n είναι ξένα,

$$A_n = \bigcup_{i=1}^{n} B_i,$$

και

$$A = \bigcup_{i=1}^{\infty} B_i.$$

Επομένως,

$$P(A_n) = \sum_{i=1}^{n} P(B_i)$$

και

$$P(A) = \sum_{i=1}^{\infty} P(B_i).$$

Όμως

$$\lim_{n\to\infty} \sum_{i=1}^{n} P(B_i) = \sum_{i=1}^{\infty} P(B_i)$$

από τον ορισμό του αθροίσματος άπειρης σειράς. Από την (15) έπεται ότι

$$\begin{aligned} \lim_{n\to\infty} P(A_n) &= \lim_{n\to\infty} \sum_{i=1}^{n} P(B_i) \\ &= \sum_{i=1}^{\infty} P(B_i) = P(A), \end{aligned}$$

δηλαδή ισχύει η (14).

Απόδειξη του (ii). Υποθέτουμε ότι $A_1 \supset A_2 \supset \ldots$ και $A = \bigcap_{n=1}^{\infty} A_n$. Τότε $A_1^c \subset A_2^c \subset \ldots$ και από την (8)

$$A^c = \bigcup_{n=1}^{\infty} A_n^c.$$

Έτσι, από το (i) του θεωρήματος

$$\lim_{n\to\infty} P(A_n^c) = P(A^c). \tag{16}$$

Αφού $P(A_n^c) = 1 - P(A_n)$ και $P(A^c) = 1 - P(A)$, από την (16) έπεται ότι

$$\begin{aligned} \lim_{n\to\infty} P(A_n) &= \lim_{n\to\infty} (1 - P(A_n^c)) \\ &= 1 - \lim_{n\to\infty} P(A_n^c) \\ &= 1 - P(A^c) = P(A), \end{aligned}$$

οπότε η (14) ισχύει πάλι. □

1.4 Δεσμευμένη πιθανότητα

Ένα δοχείο περιέχει r κόκκινους βώλους που φέρουν τους αριθμούς $1, 2, \ldots, r$ και b μαύρους βώλους που φέρουν τους αριθμούς $1, 2, \ldots, b$. Υποθέτουμε ότι η πιθανότητα επιλογής κάθε βώλου είναι $(b+r)^{-1}$. Αν ξέρουμε ότι ο βώλος που επιλέξαμε είναι κόκκινος, ποιά είναι η πιθανότητα να είναι ο κόκκινος βώλος που φέρει τον αριθμό 1; Ένας άλλος τρόπος να διατυπώσουμε το πρόβλημα είναι ο εξής. Έστω A το ενδεχόμενο να επιλέξουμε κόκκινο βώλο και B το ενδεχόμενο να επιλέξουμε τον κόκκινο βώλο με αριθμό 1. Το πρόβλημα είναι τότε να προσδιορίσουμε την «δεσμευμένη» πιθανότητα να έχει συμβεί το B, δεδομένου ότι συνέβη το ενδεχόμενο A. Στο πρόβλημα αυτό δεν μπορούμε να δώσουμε απάντηση αν δεν έχουμε προηγουμένως δώσει έναν ακριβή ορισμό της δεσμευμένης πιθανότητας ενός ενδεχομένου με δεδομένο κάποιο άλλο. Ο ακριβής αυτός ορισμός έχει ως εξής:

Ορισμός 3. *Έστω A και B δύο ενδεχόμενα τέτοια ώστε $P(A) > 0$. Τότε η δεσμευμένη πιθανότητα του B με δεδομένο το A, η οποία συμβολίζεται με $P(B|A)$, ορίζεται από την*

$$P(B|A) = \frac{P(B \cap A)}{P(A)}. \tag{17}$$

Αν $P(A) = 0$, τότε η δεσμευμένη πιθανότητα του B με δεδομένο το A δεν ορίζεται.

Ο παραπάνω ορισμός δικαιολογείται διαισθητικά αν ερμηνεύσουμε τις πιθανότητες σαν σχετικές συχνότητες. Θεωρούμε ένα πείραμα που επαναλαμβάνεται πολλές φορές. Συμβολίζουμε με $N_n(A), N_n(B)$, και $N_n(A \cap B)$, το πλήθος εκείνων από τις n δοκιμές στις οποίες συμβαίνουν τα ενδεχόμενα A, B και $A \cap B$, αντίστοιχα. Για μεγάλα n περιμένουμε ότι οι αριθμοί $N_n(A)/n, N_n(B)/n$ και $N_n(A \cap B)$ πρέπει να είναι κοντά στους $P(A), P(B)$ και $P(A \cap B)$, αντίστοιχα. Αν τώρα καταγράψουμε μόνο εκείνα τα πειράματα στα οποία συμβαίνει το A, έχουμε $N_n(A)$ δοκιμές στις οποίες το ενδεχόμενο B συμβαίνει $N_n(A \cap B)$ φορές. Έτσι, το ποσοστό των δοκιμών στις οποίες συμβαίνει το B αν περιοριστούμε σε αυτά τα $N_n(A)$ πειράματα είναι $N_n(A \cap B)/N_n(A)$. Αλλά

$$\frac{N_n(A \cap B)}{N_n(A)} = \frac{N_n(A \cap B)/n}{N_n(A)/n},$$

άρα για μεγάλες τιμές του n αυτό το κλάσμα πρέπει να είναι κοντά στο

$$P(A \cap B)/P(A).$$

Θα δώσουμε ένα πρώτο παράδειγμα εφαρμογής της (17) λύνοντας το πρόβλημα που τέθηκε στην αρχή αυτής της παραγράφου. Αφού το σύνολο Ω έχει $b + r$ σημεία καθένα από τα οποία έχει πιθανότητα $(b+r)^{-1}$, βλέπουμε ότι $P(A) = r(b+r)^{-1}$ και $P(A \cap B) = (b+r)^{-1}$. Επομένως

$$P(B|A) = \frac{1}{r}.$$

Συγκρίνετε με την «αδέσμευτη» πιθανότητα του B, που είναι ίση με $P(B) = (b+r)^{-1}$.

Παράδειγμα 4. Υποθέτουμε ότι ρίχνουμε μία φορά δύο όμοια, αμερόληπτα νομίσματα.

(α) Να βρεθεί η δεσμευμένη πιθανότητα να εμφανιστούν γράμματα και στα δύο νομίσματα, με δεδομένο ότι στο πρώτο νόμισμα εμφανίστηκαν γράμματα.

(β) Να βρεθεί η δεσμευμένη πιθανότητα να εμφανιστούν γράμματα και στα δύο νομίσματα, με δεδομένο ότι εμφανίστηκαν γράμματα σε τουλάχιστον ένα από αυτά.

Για να απαντήσουμε αυτά τα ερωτήματα παίρνουμε τον χώρο πιθανότητας Ω να αποτελείται από τα τέσσερα σημεία ΓΓ, ΓΚ, ΚΓ, ΚΚ, καθένα δε από αυτά να έχει πιθανότητα $1/4$. Έστω A το ενδεχόμενο το πρώτο νόμισμα να δείξει γράμματα, και B το ενδεχόμενο το δεύτερο νόμισμα να δείξει γράμματα. Για το (α) κάνουμε τον υπολογισμό

$$P(A \cap B | A) = P(A \cap B)/P(A) = (1/4)/(1/2) = 1/2.$$

Για το (β) κάνουμε τον υπολογισμό

$$P(A \cap B | A \cup B) = P(A \cap B)/P(A \cup B) = (1/4)/(3/4) = 1/3.$$

Στα προηγούμενα δύο παραδείγματα ο χώρος πιθανότητας ήταν γνωστός, και χρησιμοποιώντας την (17) υπολογίσαμε διάφορες δεσμευμένες πιθανότητες. Σε πολλά προβλήματα όμως, αυτό που κάνουμε είναι στην αντίθετη κατεύθυνση. Μάς δίνουν εκ των προτέρων κάποιες δεσμευμένες πιθανότητες και χρησιμοποιώντας αυτήν την πληροφορία μαθαίνουμε ποιό είναι το μέτρο πιθανότητας στο Ω. Η κατάσταση αυτή περιγράφεται στο παρακάτω τυπικό παράδειγμα.

Παράδειγμα 5. Υποθέτουμε ότι ο πληθυσμός κάποιας πόλης είναι 40% αρσενικός και 60% θηλυκός. Υποθέτουμε ακόμα ότι 50% του αρσενικού και 30% του θηλυκού πληθυσμού είναι καπνιστές. Να βρεθεί η πιθανότητα ένας καπνιστής να είναι γένους αρσενικού.

Συμβολίζουμε με A το ενδεχόμενο να επιλέξουμε άτομο αρσενικού γένους, και με Θ το ενδεχόμενο να επιλέξουμε άτομο θηλυκού γένους. Επίσης, συμβολίζουμε με K το ενδεχόμενο να επιλέξουμε καπνιστή, και με Λ το ενδεχόμενο να επιλέξουμε μη καπνιστή. Η πληροφορία που μας έχει δοθεί είναι ότι $P(K|A) = 0.5$, $P(K|\Theta) = 0.3$, $P(A) = 0.4$, και $P(\Theta) = 0.6$. Το πρόβλημα είναι να υπολογιστεί η $P(A|K)$. Από την (17),

$$P(A|K) = \frac{P(A \cap K)}{P(K)}.$$

Αλλά $P(A \cap K) = P(A)P(K|A) = (0.4)(0.5) = 0.20$, άρα ο αριθμητής υπολογίζεται με βάση τις πιθανότητες που μας δόθηκαν. Το K είναι η ένωση των ξένων συνόλων $K \cap A$ και $K \cap \Theta$, άρα

$$P(K) = P(K \cap A) + P(K \cap \Theta).$$

Αφού
$$P(K \cap \Theta) = P(\Theta)P(K|\Theta) = (0.6)(0.3) = 0.18,$$
βλέπουμε ότι
$$P(K) = 0.20 + 0.18 = 0.38.$$
Επομένως,
$$P(A|K) = \frac{0.20}{0.38} \approx 0.53.$$

Ο αναγνώστης θα παρατηρήσει ότι ο χώρος πιθανότητας δεν αναφέρθηκε ποτέ σαφώς. Το πρόβλημα αυτό και άλλα του ιδίου τύπου, λύνονται απλώς με βάση τα δεδομένα και τους κανόνες υπολογισμού πιθανοτήτων που δώσαμε στην Παράγραφο 3.

Είναι πολύ εύκολο να κατασκευάσουμε έναν χώρο πιθανότητας για το παραπάνω παράδειγμα. Παίρνουμε το σύνολο Ω να αποτελείται από τέσσερα σημεία, τα $KA, K\Theta, \Lambda A$, και $\Lambda\Theta$ που είναι, αντίστοιχα, τα μοναδικά σημεία στα σύνολα $K \cap A, K \cap \Theta, \Lambda \cap A$, και $\Lambda \cap \Theta$. Οι πιθανότητες που δίνουμε στα τέσσερα αυτά σημεία δεν ορίζονται απευθείας, θέλουμε όμως να τις προσδιορίσουμε έτσι ώστε τα ενδεχόμενα $K|A$, $K|\Theta$, A, και Θ να έχουν τις πιθανότητες που μας δόθηκαν. Βρήκαμε ήδη ότι $P(K \cap A) = 0.20$ και $P(K \cap \Theta) = 0.18$. Αφήνουμε σαν άσκηση τον υπολογισμό των πιθανοτήτων που αντιστοιχούν στα άλλα δύο σημεία.

Το πρόβλημα που συζητήσαμε σε αυτό το παράδειγμα είναι ειδική περίπτωση της ακόλουθης γενικής κατάστασης. Υποθέτουμε ότι $A_1, A_2, \ldots, A_n$ είναι n ξένα ανά δύο ενδεχόμενα με ένωση το Ω. Έστω B ένα ενδεχόμενο ώστε $P(B) > 0$ και ας υποθέσουμε ότι δίνονται οι $P(B|A_k)$ και $P(A_k)$ για $1 \leq k \leq n$. Ποιά είναι η $P(A_i|B)$; Για να λύσουμε αυτό το πρόβλημα παρατηρούμε ότι τα A_k είναι ξένα σύνολα με ένωση το Ω, επομένως
$$B = B \cap \left(\bigcup_{k=1}^{n} A_k \right) = \bigcup_{k=1}^{n} (B \cap A_k).$$
Άρα
$$P(B) = \sum_{k=1}^{n} P(B \cap A_k).$$
Όμως
$$P(B \cap A_k) = P(A_k)P(B|A_k),$$
οπότε μπορούμε να γράψουμε
$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i)P(B|A_i)}{\sum_{k=1}^{n} P(A_k)P(B|A_k)}. \tag{18}$$

Αυτός ο τύπος, που λέγεται τύπος του Bayes, εφαρμόζεται συχνά. Ένας τρόπος να σκεφτόμαστε το αποτέλεσμα της (18) είναι ο εξής. Ας υποθέσουμε ότι σκεφτόμαστε τα ενδεχόμενα A_k σαν τις πιθανές «αιτίες» ενός παρατηρήσιμου ενδεχομένου B. Τότε η $P(A_i|B)$ είναι η πιθανότητα το A_i να «προκάλεσε» το B με δεδομένο ότι το

B συμβαίνει. Ο νόμος του Bayes αποτελεί την βάση μίας στατιστικής μεθόδου που λέγεται «διαδικασίες Bayes» και θα συζητηθεί στον δεύτερο τόμο αυτής της σειράς (Εισαγωγή στην Στατιστική Θεωρία).

Σαν δείγμα εφαρμογής του νόμου του Bayes θεωρούμε το εξής (κάπως κλασικό) πρόβλημα.

Παράδειγμα 6. Υποθέτουμε ότι έχουμε τρία ερμάρια και καθένα από αυτά έχει τρία συρτάρια. Το πρώτο ερμάριο περιέχει ένα χρυσό νόμισμα σε κάθε συρτάρι, το δεύτερο περιέχει ένα χρυσό νόμισμα στο ένα συρτάρι κι ένα αργυρό νόμισμα στο άλλο συρτάρι, και το τρίτο περιέχει ένα αργυρό νόμισμα σε κάθε συρτάρι. Επιλέγουμε τυχαία ένα ερμάριο και ανοίγουμε ένα συρτάρι. Αν το συρτάρι περιέχει χρυσό νόμισμα, ποιά είναι η πιθανότητα το δεύτερο συρτάρι του να περιέχει κι αυτό χρυσό νόμισμα; Προτείνουμε στον αναγνώστη να σταματήσει και να μαντέψει την απάντηση πριν διαβάσει τη λύση. Συχνά δίνεται η λανθασμένη απάντηση $1/2$.

Το πρόβλημα λύνεται εύκολα και σωστά αν χρησιμοποιήσουμε τον τύπο του Bayes, αρκεί να δώσουμε κατάλληλη περιγραφή των δεδομένων. Μπορούμε να θεωρήσουμε έναν χώρο πιθανότητας στον οποίο τα ενδεχόμενα A_1, A_2 και A_3 αντιστοιχούν στην επιλογή του πρώτου, δεύτερου και τρίτου ερμαρίου, αντίστοιχα. Τα ενδεχόμενα αυτά είναι ξένα και η ένωσή τους είναι ολόκληρος ο χώρος Ω αφού επιλέγουμε ακριβώς ένα ερμάριο. Επιπλέον, μπορεί κανείς να υποθέσει ότι είναι εξίσου πιθανό να επιλέξουμε οποιοδήποτε από τα τρία ερμάρια, δηλαδή $P(A_i) = 1/3, i = 1, 2, 3$. Έστω B το ενδεχόμενο να βρήκαμε χρυσό νόμισμα. Τότε, από τη σύνθεση των ερμαρίων είναι φανερό ότι

$$P(B|A_1) = 1, \quad P(B|A_2) = 1/2, \quad P(B|A_3) = 0.$$

Το πρόβλημα ζητάει την πιθανότητα το δεύτερο συρτάρι να περιέχει χρυσό νόμισμα δεδομένου ότι στο πρώτο βρήκαμε χρυσό νόμισμα. Αυτό μπορεί να συμβεί μόνο αν επιλέξαμε το πρώτο ερμάριο, δηλαδή το πρόβλημα μάς ζητάει την $P(A_1|B)$. Μπορούμε τώρα να εφαρμόσουμε τον νόμο του Bayes (18) και να βρούμε την απάντηση, που είναι $2/3$. Αφήνουμε σαν άσκηση για τον αναγνώστη τον υπολογισμό της πιθανότητας στο δεύτερο συρτάρι να περιέχεται αργυρό νόμισμα δεδομένου ότι στο πρώτο συρτάρι βρήκαμε χρυσό νόμισμα.

Στο επόμενο παράδειγμά μας θεωρούμε ένα απλό πιθανοθεωρητικό σχήμα που οφείλεται στον Pólya.

Παράδειγμα 7. *Το δοχείο του Pólya.* Υποθέτουμε ότι ένα δοχείο περιέχει r κόκκινους και b μαύρους βώλους. Επιλέγουμε έναν βώλο και καταγράφουμε το χρώμα του. Τότε αυτός μαζί με $c > 0$ ακόμα βώλους του ίδιου χρώματος προστίθενται στο δοχείο. Η διαδικασία επαναλαμβάνεται $n - 1$ ακόμα φορές, δηλαδή το πλήθος των επιλογών που κάνουμε από το δοχείο είναι n.

Με R_j, $1 \leq j \leq n$, συμβολίζουμε το ενδεχόμενο ο j-στός βώλος που επιλέγουμε να είναι κόκκινος και με B_j, $1 \leq j \leq n$, συμβολίζουμε το ενδεχόμενο ο j-στός βώλος που επιλέγουμε να είναι μαύρος. Φυσικά, για κάθε j, τα R_j και B_j είναι ξένα. Κατά την k-στή επιλογή υπάρχουν $b + r + (k - 1)c$ βώλοι στο δοχείο, και υποθέτουμε ότι η πιθανότητα επιλογής οποιουδήποτε βώλου είναι $(b+r+(k-1)c)^{-1}$. Για να υπολογίσουμε την $P(R_1 \cap R_2)$ γράφουμε

$$P(R_1 \cap R_2) = P(R_1)P(R_2|R_1).$$

Τώρα

$$P(R_1) = \frac{r}{b+r}, \quad P(R_2|R_1) = \frac{r+c}{b+r+c},$$

επομένως,

$$P(R_1 \cap R_2) = \left(\frac{r}{b+r}\right)\left(\frac{r+c}{b+r+c}\right).$$

Όμοια

$$P(B_1 \cap R_2) = \left(\frac{b}{b+r}\right)\left(\frac{r}{b+r+c}\right)$$

επομένως

$$\begin{aligned} P(R_2) &= P(R_1 \cap R_2) + P(B_1 \cap R_2) \\ &= \left(\frac{r}{b+r}\right)\left(\frac{r+c}{b+r+c}\right) + \left(\frac{b}{b+r}\right)\left(\frac{r}{b+r+c}\right) \\ &= \frac{r}{b+r}. \end{aligned}$$

Συμπεραίνουμε ότι $P(R_2) = P(R_1)$. Αφού

$$P(B_2) = 1 - P(R_2) = \frac{b}{b+r},$$

βλέπουμε ότι $P(B_2) = P(B_1)$. Στις ασκήσεις θα δούμε κι άλλες ιδιότητες του πειράματος του Pólya.

1.5 Ανεξαρτησία

Θεωρούμε ένα δοχείο που περιέχει τέσσερις διαφορετικούς βώλους και εκτελούμε το πείραμα της επιλογής ενός βώλου από το δοχείο. Υποθέτουμε ότι είναι εξίσου πιθανό να επιλέξουμε οποιονδήποτε από τους βώλους. Θέτουμε $\Omega = \{1,2,3,4\}$ και δίνουμε πιθανότητα $1/4$ σε κάθε σημείο.

Έστω A και B δύο ενδεχόμενα. Για κάποιες επιλογές των A και B, αν γνωρίζουμε ότι συμβαίνει το A αυξάνεται η πιθανότητα να συμβαίνει το B. Για παράδειγμα, αν $A = \{1,2\}$ και $B = \{1\}$, τότε $P(A) = 1/2, P(B) = 1/4$, και $P(A \cap B) = 1/4$. Επομένως, $P(B|A) = 1/2$, που είναι μεγαλύτερη από την $P(B)$. Από την άλλη πλευρά, για άλλες επιλογές των A και B, αν γνωρίζουμε ότι συμβαίνει το A μειώνεται η πιθανότητα να συμβεί το B. Για παράδειγμα, αν $A = \{1,2,3\}$ και $B = \{1,2,4\}$, τότε $P(A) = 3/4, P(B) = 3/4$ και $P(A \cap B) = 1/2$. Άρα $P(B|A) = 2/3$, που είναι μικρότερη από την $P(B)$.

Μεγάλο ενδιαφέρον παρουσιάζει η περίπτωση κατά την οποία αν γνωρίζουμε ότι συμβαίνει το A δεν αλλάζει η πιθανότητα να συμβαίνει το B. Σαν παράδειγμα πάρτε $A = \{1,2\}$ και $B = \{1,3\}$. Τότε $P(A) = 1/2, P(B) = 1/2, P(A \cap B) = 1/4$, επομένως $P(B|A) = 1/2$. Ενδεχόμενα σαν κι αυτά, για τα οποία η δεσμευμένη πιθανότητα είναι η ίδια με την αδέσμευτη πιθανότητα, λέγονται ανεξάρτητα.

Έστω τώρα A και B δύο ενδεχόμενα σε έναν γενικό χώρο πιθανότητας και ας υποθέσουμε ότι $P(A) \neq 0$. Θέλουμε να λέμε ότι τα A και B είναι ανεξάρτητα αν $P(B|A) = P(B)$. Αφού $P(B|A) = P(B \cap A)/P(A)$ βλέπουμε ότι αν τα A και B είναι ανεξάρτητα τότε

$$P(A \cap B) = P(A)P(B). \tag{19}$$

Δεδομένου ότι η (19) έχει έννοια ακόμα κι αν $P(A) = 0$ και είναι συμμετρική ως προς τα γράμματα A και B, μας προσφέρει έναν προτιμότερο ορισμό της ανεξαρτησίας.

Ορισμός 4. *Δύο ενδεχόμενα A και B λέγονται ανεξάρτητα αν και μόνο αν*

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

Μπορούμε να θεωρήσουμε το ίδιο πρόβλημα για τρία σύνολα A, B και C. Παίρνουμε $\Omega = \{1,2,3,4\}$ και δίνουμε πιθανότητα $1/4$ σε κάθε σημείο. Έστω $A = \{1,2\}$, $B = \{1,3\}$ και $C = \{1,4\}$. Αφήνουμε σαν άσκηση να δειχθεί ότι τα ζεύγη ενδεχομένων A και B, A και C, B και C είναι ανεξάρτητα. Λέμε ότι τα ενδεχόμενα A, B και C είναι *ανά δύο* ανεξάρτητα. Από την άλλη πλευρά, $P(C) = 1/2$ και

$$P(C|A \cap B) = 1.$$

Έτσι, αν γνωρίζουμε ότι συμβαίνει το ενδεχόμενο $A \cap B$ αυξάνεται η πιθανότητα να συμβεί το C. Με αυτήν την έννοια, τα ενδεχόμενα A, B και C *δεν* είναι ανεξάρτητα. Γενικά, τρία ενδεχόμενα A, B, και C λέγονται ανεξάρτητα αν είναι ανά δύο ανεξάρτητα και αν

$$P(A \cap B \cap C) = P(A)P(B)P(C).$$

Αφήνουμε σαν άσκηση να δειχθεί ότι αν τα A, B και C είναι ανεξάρτητα και αν $P(A \cap B) \neq 0$, τότε $P(C|A \cap B) = P(C)$.

Πιό γενικά λέμε ότι $n \geq 3$ ενδεχόμενα $A_1, A_2, \ldots, A_n$ είναι ανεξάρτητα αν

$$P(A_1 \cap \ldots \cap A_n) = P(A_1) \ldots P(A_n)$$

και οποιοδήποτε υποσύνολο που περιέχει τουλάχιστον δύο αλλά λιγότερα από n ενδεχόμενα, αποτελείται από ανεξάρτητα ενδεχόμενα.

Παράδειγμα 8. Έστω S το τετράγωνο $0 \leq x \leq 1, 0 \leq y \leq 1$ στο επίπεδο. Θεωρούμε τον ομοιόμορφο χώρο πιθανότητας στο τετράγωνο και ορίζουμε A το ενδεχόμενο

$$\{(x,y) : 0 \leq x \leq 1/2, 0 \leq y \leq 1\}$$

και B το ενδεχόμενο

$$\{(x,y) : 0 \leq x \leq 1, 0 \leq y \leq 1/4\}.$$

Δείξτε ότι τα A και B είναι ανεξάρτητα ενδεχόμενα.

Για να το δείξουμε, υπολογίζουμε τις $P(A), P(B)$ και $P(A \cap B)$, και δείχνουμε ότι $P(A \cap B) = P(A)P(B)$. Το A είναι ένα υποορθογώνιο του τετραγώνου S με εμβαδόν $1/2$ και το B είναι ένα υποορθογώνιο του S με εμβαδόν $1/4$, άρα $P(A) = 1/2$ και $P(B) = 1/4$. Αφού το

$$A \cap B = \{(x,y) : 0 \leq 1/2, 0 \leq y \leq 1/4\}$$

είναι κι αυτό υποορθογώνιο του S με εμβαδόν $1/8$, έχουμε $P(A \cap B) = 1/8$, το οποίο δείχνει ότι τα A και B είναι ανεξάρτητα ενδεχόμενα.

Η έννοια της ανεξαρτησίας χρησιμοποιείται συχνά για την κατασκευή χώρων πιθανότητας που αντιστοιχούν σε επαναλήψεις του ιδίου πειράματος. Θα ασχοληθούμε με αυτό το θέμα λεπτομερέστερα στο Κεφάλαιο 3. Εδώ θα αρκεστούμε στη μελέτη της απλούστερης περίπτωσης, που αφορά πειράματα σαν την ρίψη ενός νομίσματος, που έχουν δύο μόνο πιθανά αποτελέσματα - επιτυχία ή αποτυχία.

Ρίχνουμε ένα νόμισμα n φορές, με πιθανότητα επιτυχίας ή αποτυχίας p και $1-p$ αντίστοιχα. Διαισθητικά πιστεύουμε ότι το αποτέλεσμα της i-στής ρίψης δεν επηρεάζει το αποτέλεσμα των υπολοίπων ρίψεων. Θέλουμε τώρα να κατασκευάσουμε έναν χώρο πιθανότητας που θα αντιστοιχεί στο σύνθετο πείραμα της n-πλής επανάληψης του απλού μας πειράματος και θα συμβαδίζει με την διαίσθησή μας.

Αφού καθεμία από τις n δοκιμές καταλήγει είτε σε επιτυχία ή σε αποτυχία, έχουμε συνολικά 2^n δυνατά αποτελέσματα για το σύνθετο πείραμα. Τα αποτελέσματα αυτά αναπαρίστανται από n-άδες $(x_1, \ldots, x_n)$, όπου $x_i = 1$ ή 0 ανάλογα με την κατάληξη της i-στής δοκιμής σε επιτυχία ή αποτυχία. Ορίζουμε Ω να είναι το σύνολο όλων αυτών των n-άδων. Σαν σ-άλγεβρα παίρνουμε τη συλλογή όλων των υποσυνόλων του Ω.

Ερχόμαστε τώρα στον ορισμό του μέτρου πιθανότητας. Για να το ορίσουμε, αρκεί να αποφασίσουμε ποιά θα είναι η πιθανότητα για καθένα από τα 2^n μονοσύνολα $\{(x_1, \ldots, x_n)\}$. Υποθέτουμε ότι η n-άδα $(x_1, \ldots, x_n)$ είναι τέτοια που ακριβώς k από τα x_i να έχουν την τιμή 1. Για απλότητα, ας πούμε ότι $x_1 = x_2 = \ldots = x_k = 1$ και τα υπόλοιπα x_i έχουν την τιμή 0. Τότε, αν με A_i συμβολίζουμε το ενδεχόμενο η i-στή δοκιμή να είναι επιτυχημένη, $1 \leq i \leq n$, βλέπουμε ότι

$$\{(\underbrace{1, 1, \ldots, 1}_{k}, \underbrace{0, \ldots, 0}_{n-k})\} = A_1 \cap \ldots \cap A_k \cap A_{k+1}^c \cap \ldots \cap A_n^c.$$

Σύμφωνα με τη διαίσθησή μας, τα ενδεχόμενα $A_1, \ldots, A_k, A_{k+1}^c, \ldots, A_n^c$ πρέπει να είναι ανεξάρτητα και $P(A_i) = p$, $1 \leq i \leq n$. Πρέπει λοιπόν να ορίσουμε το P έτσι ώστε

$$\begin{aligned} P(\{(1, 1, \ldots, 1, 0, \ldots, 0)\}) &= P(A_1) \ldots P(A_k) P(A_{k+1}^c) \ldots P(A_n^c) \\ &= p^k (1-p)^{n-k}. \end{aligned}$$

Με τον ίδιο ακριβώς συλλογισμό βλέπουμε ότι για κάθε n-άδα $(x_1, \ldots, x_n)$ στην οποία ακριβώς k από τα x_i έχουν την τιμή 1, το μέτρο P πρέπει να είναι τέτοιο ώστε

$$P(\{(x_1, \ldots, x_n)\}) = p^k(1-p)^{n-k}.$$

Ας υπολογίσουμε τώρα την πιθανότητα ακριβώς k από τις n δοκιμές να καταλήγουν σε επιτυχία. Παρατηρήστε ότι δεν είναι η ίδια με την πιθανότητα k συγκεκριμένες δοκιμές να είναι επιτυχημένες και οι υπόλοιπες $n-k$ αποτυχημένες. Έστω B_k το ενδεχόμενο ακριβώς k από τις n δοκιμές να είναι επιτυχημένες. Αφού η επιλογή οποιασδήποτε τέτοιας ακολουθίας έχει πιθανότητα $p^k(1-p)^{n-k}$, το ενδεχόμενο B_k έχει πιθανότητα $P(B_k) = C(k,n)p^k(1-p)^{n-k}$, όπου $C(k,n)$ είναι το πλήθος των ακολουθιών $(x_1, \ldots, x_n)$ στίς οποίες ακριβώς k από τα x_i έχουν την τιμή 1. Ο υπολογισμός του $C(k,n)$ είναι ένα απλό συνδυαστικό πρόβλημα το οποίο θα λύσουμε στην Παράγραφο 2.4. Εκεί θα δείξουμε ότι

$$C(k,n) = \frac{n!}{k!(n-k)!}, \quad 0 \le k \le n. \tag{20}$$

Υπενθυμίζουμε ότι $0! = 1$ και, για κάθε θετικό ακέραιο m,

$$m! = m(m-1)\ldots 1.$$

Η ποσότητα $n!/k!(n-k)!$ συνήθως συμβολίζεται με $\binom{n}{k}$ (ο διωνυμικός συντελεστής). Άρα,

$$P(B_k) = \binom{n}{k} p^k(1-p)^{n-k}. \tag{21}$$

Με βάση το μοντέλο των ανεξάρτητων δοκιμών επιτυχίας-αποτυχίας μπορούμε να μελετήσουμε πλήθος εφαρμοσμένα προβλήματα. Τυπικό είναι το εξής.

Παράδειγμα 9. Υποθέτουμε ότι μία μηχανή παράγει λυχνίες, 10% από τις οποίες είναι ελαττωματικές. Να βρεθεί η πιθανότητα ένα κουτί με τρείς λυχνίες να περιέχει το πολύ μία ελαττωματική.

Για να λύσουμε το πρόβλημα υποθέτουμε ότι η παραγωγή των λυχνιών συνίσταται από επαναλήψεις ανεξάρτητων δοκιμών επιτυχίας-αποτυχίας με επιτυχία την παραγωγή ελαττωματικής λυχνίας. Η πιθανότητα επιτυχίας εδώ είναι 0.1. Έστω B_0 το ενδεχόμενο καμμία από τις τρείς λυχνίες να μην είναι ελαττωματική, και B_1 το ενδεχόμενο ακριβώς μία από τις λυχνίες να είναι ελαττωματική. Τότε, $B_0 \cup B_1$ είναι το ενδεχόμενο τουλάχιστον μία λυχνία να είναι ελαττωματική. Αφού τα ενδεχόμενα B_0 και B_1 είναι προφανώς ξένα, έπεται ότι

$$\begin{aligned} P(B_0 \cup B_1) &= P(B_0) + P(B_1) \\ &= \binom{3}{0}(0.1)^0(0.9)^3 + \binom{3}{1}(0.1)^1(0.9)^2 \\ &= (0.9)^3 + 3(0.1)(0.9)^2 \\ &= 0.972. \end{aligned}$$

1.6 Ασκήσεις

1 Έστω $(\Omega, \mathcal{A}, P)$ χώρος πιθανότητας, στον οποίο $\mathcal{A}$ είναι η σ-άλγεβρα όλων των υποσυνόλων του Ω και P είναι ένα μέτρο πιθανότητας που δίνει πιθανότητα $p > 0$ σε κάθε μονοσύνολο του Ω.

(α) Δείξτε ότι το Ω έχει πεπερασμένα το πλήθος σημεία. *Υπόδειξη:* Δείξτε ότι το Ω δεν μπορεί να έχει περισσότερα από p^{-1} σημεία.

(β) Δείξτε ότι αν n είναι το πλήθος των σημείων του Ω, τότε το p πρέπει να είναι ίσο με n^{-1}.

2 Κατασκευάζουμε ένα μοντέλο για μία ρουλέτα θεωρώντας έναν ομοιόμορφο χώρο πιθανότητας στην περιφέρεια ενός δίσκου ακτίνας 1, έτσι ώστε η πιθανότητα η ρουλέτα να δείξει ένα τόξο μήκους s να είναι $s/2\pi$. Υποθέτουμε ότι ο δίσκος διαιρείται σε 37 κυκλικούς τομείς αριθμημένους από 1 ως 37. Υπολογίστε την πιθανότητα η ρουλέτα να σταματήσει σε κυκλικό τομέα με άρτιο αριθμό.

3 Επιλέγουμε τυχαία ένα σημείο στο μοναδιαίο τετράγωνο. Υπολογίστε την πιθανότητα να βρίσκεται στο τρίγωνο που ορίζεται από τις $x = 0$, $y = 0$ και $x + y = 1$.

4 Επιλέγουμε τυχαία ένα σημείο στο δίσκο ακτίνας 1. Βρείτε την πιθανότητα να ανήκει στον κυκλικό τομέα που ορίζεται από τα 0 και τα $\pi/4$ ακτίνια.

5 Στο Παράδειγμα 2 υπολογίστε τις ακόλουθες πιθανότητες:

(α) Καμμία διάσπαση δεν συμβαίνει πριν τη χρονική στιγμή 10.

(β) Συμβαίνει μία διάσπαση είτε πριν τη χρονική στιγμή 2 ή μεταξύ των χρονικών στιγμών 3 και 5.

6 Ένα δοχείο περιέχει 10 βώλους, αριθμημένους από 1 ως 10. Επιλέγουμε τυχαία έναν βώλο από το δοχείο. Υπολογίστε την πιθανότητα ο αριθμός πάνω στον βώλο να είναι ένας από τους 3,4, ή 5.

7 Ρίχνουμε δύο ζάρια μία φορά και υποθέτουμε ότι τα 36 δυνατά αποτελέσματα είναι ισοπίθανα. Βρείτε την πιθανότητα το άθροισμα των αριθμών στις δύο έδρες να είναι άρτιος αριθμός.

8 Υποθέτουμε ότι τα ενδεχόμενα A και B ικανοποιούν τις $P(A) = 2/5$, $P(B) = 2/5$ και $P(A \cup B) = 1/2$. Βρείτε την $P(A \cap B)$.

9 Αν $P(A) = 1/3$, $P(A \cup B) = 1/2$ και $P(A \cap B) = 1/4$, βρείτε την $P(B)$.

10 Επιλέγουμε τυχαία ένα σημείο στο μοναδιαίο τετράγωνο. Έστω A το ενδεχόμενο να ανήκει στο τρίγωνο που ορίζεται από τις $y = 0$, $x = 1$ και $x = y$, και B το ενδεχόμενο να ανήκει στο ορθογώνιο με κορυφές τα σημεία $(0, 0)$, $(1, 0)$, $(1, 1/2)$ και $(0, 1/2)$. Υπολογίστε τις $P(A \cup B)$ και $P(A \cap B)$.

11 Ενα δοχείο περιέχει 10 βώλους, αριθμημένους από 1 ως 10. Επιλέγουμε τυχαία έναν βώλο και στη συνέχεια επιλέγουμε τυχαία έναν από τις υπόλοιπους εννέα βώλους. Βρείτε την πιθανότητα οι αριθμοί στους δύο βώλους που επιλέξαμε να διαφέρουν κατά δύο ή περισσότερες μονάδες.

12 Γνωρίζουμε ότι ένα σημείο που επιλέχθηκε τυχαία στο μοναδιαίο τετράγωνο ανήκει στο τρίγωνο που ορίζεται από τις $x = 0$, $y = 0$ και $x + y = 1$. Βρείτε την πιθανότητα να ανήκει και στο τρίγωνο που ορίζεται από τις $y = 0$, $x = 1$ και $x = y$.

13 Έχουμε τέσσερα ερμάρια, καθένα από τα οποία έχει δύο συρτάρια. Το πρώτο και το δεύτερο ερμάριο περιέχουν ένα χρυσό νόμισμα στο ένα συρτάρι και ένα αργυρό νόμισμα στο άλλο συρτάρι. Το τρίτο ερμάριο περιέχει δύο χρυσά νομίσματα και το τέταρτο περιέχει δύο αργυρά νομίσματα. Επιλέγουμε τυχαία ένα ερμάριο και ανοίγουμε ένα συρτάρι. Μέσα σε αυτό βρίσκουμε ένα χρυσό νόμισμα. Βρείτε την πιθανότητα το άλλο συρτάρι να περιέχει
(α) ένα αργυρό νόμισμα·
(β) ένα χρυσό νόμισμα.

14 Ένα δοχείο περιέχει 10 βώλους, από τους οποίους 6 είναι μαύροι και 4 άσπροι. Βγάζουμε τρείς βώλους από το δοχείο, χωρίς να παρατηρήσουμε το χρώμα τους. Βρείτε την πιθανότητα, αν βγάλουμε και τέταρτο βώλο από το δοχείο, αυτός να είναι άσπρος. Υποθέτουμε ότι είναι εξίσου πιθανό να επιλέξουμε οποιονδήποτε από τους 10 βώλους.

15 Με την ίδια σύνθεση δοχείου όπως στην Άσκηση 14, βρείτε την πιθανότητα οι τρείς βώλοι που επιλέξαμε να είναι όλοι μαύροι αν ξέρουμε ότι τουλάχιστον ένας από αυτούς είναι μαύρος.

16 Υποθέτουμε ότι ένα εργοστάσιο έχει δύο μηχανήματα A και B τα οποία κατασκευάζουν 60% και 40% της συνολικής παραγωγής, αντίστοιχα. Το ποσοστό των ελαττωματικών κομματιών είναι 3% για το μηχάνημα A και 5% για το μηχάνημα B. Βρείτε την πιθανότητα ένα ελαττωματικό κομμάτι της παραγωγής να κατασκευάστηκε από το μηχάνημα B.

17 Δείξτε με επαγωγή ως προς n ότι η πιθανότητα να επιλέξουμε κόκκινο βώλο στο n-στό βήμα του πειράματος του Pólya (Παράδειγμα 7) είναι $r(b+r)^{-1}$.

18 Ένας φοιτητής δίνει ένα διαγώνισμα πολλαπλών επιλογών, στο οποίο κάθε ερώτημα έχει 5 δυνατές απαντήσεις, από τις οποίες ακριβώς μία είναι σωστή. Αν ο φοιτητής γνωρίζει την απάντηση, επιλέγει τη σωστή απάντηση. Αλλιώς, επιλέγει τυχαία μία από τις 5 δυνατές απαντήσεις. Υποθέτουμε ότι ο φοιτητής γνωρίζει την απάντηση για το 70% των ερωτημάτων.
(α) Ποιά είναι η πιθανότητα ο φοιτητής να δώσει την σωστή απάντηση σε δεδομένο ερώτημα;
(β) Αν ο φοιτητής δώσει την σωστή απάντηση σε κάποιο ερώτημα, ποιά είναι η πιθανότητα να γνώριζε την απάντηση;

19 Επιλέγουμε τυχαία ένα σημείο στο μοναδιαίο τετράγωνο. Αν γνωρίζουμε ότι το σημείο ανήκει στο ορθογώνιο που ορίζεται από τις $y = 0$, $y = 1$, $x = 0$ και $x = 1/2$, ποιά είναι η πιθανότητα το σημείο να ανήκει στο τρίγωνο που ορίζεται από τις $y = 0$, $x = 1/2$ και $x + y = 1$;

20 Υποθέτουμε ότι ένα δοχείο περιέχει r κόκκινους και b μαύρους βώλους. Επιλέγουμε τυχαία έναν βώλο από το δοχείο, και στη συνέχεια επιλέγουμε τυχαία έναν

δεύτερο βώλο από αυτούς που είχαν απομείνει στο δοχείο. Βρείτε την πιθανότητα των ενδεχομένων:
(α) και οι δύο βώλοι είναι κόκκινοι·
(β) η πρώτος βώλος είναι κόκκινος και ο δεύτερος μαύρος·
(γ) ο πρώτος βώλος είναι μαύρος και ο δεύτερος κόκκινος·
(δ) και οι δύο βώλοι είναι μαύροι.

21 Ένα δοχείο περιέχει 10 κόκκινους και 5 μαύρους βώλους. Επιλέγουμε τυχαία έναν βώλο από το δοχείο. Αν ο βώλος είναι κόκκινος, τον επιστρέφουμε στο δοχείο. Αν ο βώλος είναι μαύρος, αυτός και 2 ακόμα μαύροι βώλοι προστίθενται στο δοχείο. Βρείτε την πιθανότητα, αν επιλέξουμε έναν δεύτερο βώλο από το δοχείο, αυτός να είναι
(α) κόκκινος· (β) μαύρος.

22 Δύο βώλοι επιλέγονται, με επανατοποθέτηση του βώλου που επιλέχθηκε πρώτος, από ένα δοχείο που περιέχει 3 άσπρους και 2 μαύρους βώλους.
(α) Κατασκευάστε δειγματικό χώρο γι' αυτό το πείραμα, τα σημεία του οποίου να είναι ισοπίθανα.
(β) Υπολογίστε την πιθανότητα οι δύο βώλοι που επιλέγουμε να έχουν το ίδιο χρώμα.
(γ) Υπολογίστε την πιθανότητα τουλάχιστον ένας από τους βώλους που επιλέγουμε να είναι άσπρος.

23 Δουλέψτε την Άσκηση 22, υποθέτοντας τώρα ότι ο πρώτος βώλος δεν επανατοποθετείται στο δοχείο.

24 Δουλέψτε την Άσκηση 22, κατασκευάζοντας δειγματικό χώρο με τέσσερα σημεία που αντιστοιχούν σε άσπρο και μαύρο για κάθε επιλογή.

25 Το δοχείο I περιέχει 2 άσπρους βώλους και 2 μαύρους βώλους, το δοχείο II περιέχει 2 άσπρους βώλους και 1 μαύρο βώλο, και το δοχείο III περιέχει 1 άσπρο βώλο και 3 μαύρους βώλους.
(α) Επιλέγουμε έναν βώλο από κάθε δοχείο. Υπολογίστε την πιθανότητα όλοι οι βώλοι να είναι άσπροι.
(β) Επιλέγουμε τυχαία ένα δοχείο, και στη συνέχεια έναν βώλο από αυτό. Υπολογίστε την πιθανότητα να είναι άσπρος.
(γ) Στο (β), υπολογίστε την πιθανότητα να είχε επιλεγεί το πρώτο δοχείο δεδομένου ότι ο βώλος που επιλέξαμε είναι άσπρος.

26 Ένα δοχείο περιέχει 3 άσπρους και 2 μαύρους βώλους. Επιλέγουμε δύο βώλους από αυτό, χωρίς επανατοποθέτηση.
(α) Υπολογίστε την πιθανότητα ο δεύτερος βώλος να είναι μαύρος δεδομένου ότι ο πρώτος είναι μαύρος.
(β) Υπολογίστε την πιθανότητα ο δεύτερος βώλος να έχει το ίδιο χρώμα με τον πρώτο.
(γ) Υπολογίστε την πιθανότητα ο πρώτος βώλος να είναι άσπρος δεδομένου ότι ο δεύτερος είναι άσπρος.

27 Σε ένα κολλέγιο, 70% των φοιτητών είναι άνδρες και 30% είναι γυναίκες. Είναι γνωστό ότι 40% των ανδρών και 60% των γυναικών είναι καπνιστές. Ποιά είναι η πιθανότητα ένας φοιτητής που εθεάθη να καπνίζει να είναι άνδρας;

28 Υποθέτουμε ότι τα αυτοκίνητα κατασκευάζονται με την ίδια πιθανότητα τη Δευτέρα, την Τρίτη, την Τετάρτη, την Πέμπτη ή την Παρασκευή. Τα αυτοκίνητα που κατασκευάζονται τη Δευτέρα είναι σε ποσοστό 4% ελαττωματικά, τα αυτοκίνητα που κατασκευάζονται την Τρίτη, την Τετάρτη ή την Πέμπτη είναι ελαττωματικά σε ποσοστό 1%, και τα αυτοκίνητα που κατασκευάζονται την Παρασκευή είναι ελαττωματικά σε ποσοστό 2%. Αν πήρατε ένα ελαττωματικό αυτοκίνητο, ποιά είναι η πιθανότητα να είχε κατασκευαστεί κάποια Δευτέρα;

29 Υποθέτουμε ότι υπάρχει κάποια εξέταση για τον καρκίνο η οποία δίνει θετικό αποτέλεσμα στο 90% των εξεταζομένων που έχουν καρκίνο και στο 5% αυτών που δεν έχουν καρκίνο. Έστω ότι 1% των ασθενών σε ένα νοσοκομείο έχουν καρκίνο. Ποιά είναι η πιθανότητα να έχει καρκίνο ένας ασθενής που επιλέγεται τυχαία και που η εξέτασή του δίνει θετικό αποτέλεσμα;

30 Στο πρόβλημα με τα τρία ερμάρια που συζητήσαμε στο Παράδειγμα 6, υπολογίστε την πιθανότητα το δεύτερο συρτάρι να περιέχει αργυρό νόμισμα δεδομένου ότι το πρώτο συρτάρι περιέχει χρυσό νόμισμα.

31 Στο πείραμα του Pólya (Παράδειγμα 7), με δεδομένο ότι ο δεύτερος βώλος που επιλέχθηκε ήταν κόκκινος, υπολογίστε την πιθανότητα των ενδεχομένων:
(α) ο πρώτος βώλος που επιλέχθηκε ήταν κόκκινος·
(β) ο πρώτος βώλος που επιλέχθηκε ήταν μαύρος.

32 Ρίχνουμε μαζί τρία όμοια, αμερόληπτα νομίσματα. Έστω A_i το ενδεχόμενο να φέρει γράμματα το i-στό νόμισμα. Δείξτε ότι τα ενδεχόμενα A_1, A_2 και A_3 είναι ανεξάρτητα.

33 Υποθέτουμε ότι τα έξι δυνατά αποτελέσματα της ρίψης ενός ζαριού είναι ισοπίθανα και ότι οι διαδοχικές ρίψεις του ζαριού είναι ανεξάρτητες. Κατασκευάστε έναν χώρο πιθανότητας για το σύνθετο πείραμα της ρίψης του ζαριού τρείς διαδοχικές φορές.

34 Έστω A και B ανεξάρτητα ενδεχόμενα. Δείξτε ότι τα A και B^c, τα A^c και B, τα A^c και B^c είναι επίσης ανεξάρτητα.

35 Θέτουμε $\Omega = \{1, 2, 3, 4\}$ και υποθέτουμε ότι κάθε σημείο έχει πιθανότητα $1/4$. Θεωρούμε τα $A = \{1, 2\}$, $B = \{1, 3\}$ και $C = \{1, 4\}$. Δείξτε ότι τα ζεύγη ενδεχομένων A και B, A και C, B και C είναι ανεξάρτητα.

36 Υποθέτουμε ότι A, B και C είναι ανεξάρτητα ενδεχόμενα και $P(A \cap B) \neq 0$. Δείξτε ότι $P(C|A \cap B) = P(C)$.

37 Η εμπειρία δείχνει ότι 20% των ανθρώπων που κάνουν κράτηση τραπεζιού σε κάποιο εστιατόριο δεν εμφανίζονται ποτέ. Αν το εστιατόριο διαθέτει 50 τραπέζια και έχει δεχτεί 52 κρατήσεις, ποιά είναι η πιθανότητα να καταφέρει να τακτοποιήσει όλους τους πελάτες του;

38 Ένας κυκλικός στόχος με μοναδιαία ακτίνα χωρίζεται σε τέσσερις ζώνες με εξωτερικές ακτίνες $1/4$, $1/2$, $3/4$ και 1, αντίστοιχα. Επιχειρούμε 10 τυχαίες και ανεξάρτητες βολές που βρίσκουν τον στόχο.
(α) Υπολογίστε την πιθανότητα το πολύ τρείς από τις βολές να πετύχουν τη ζώνη που ορίζεται από τους κύκλους με ακτίνες $1/2$ και 1.
(β) Αν 5 βολές πετυχαίνουν τον δίσκο ακτίνας $1/2$, βρείτε την πιθανότητα τουλάχιστον μία από αυτές να πέσει στον δίσκο με ακτίνα $1/4$.

39 Ένα μηχάνημα αποτελείται από τέσσερις συνιστώσες που συνδέονται παράλληλα έτσι ώστε το μηχάνημα να μην λειτουργεί μόνο αν και οι τέσσερις συνιστώσες του δεν λειτουργούν. Υποθέτουμε ότι τα ενδεχόμενα βλάβης των συνιστωσών είναι ανεξάρτητα. Αν οι συνιστώσες δεν λειτουργούν με πιθανότητες 0.1, 0.2, 0.3 και 0.4 όταν το μηχάνημα τεθεί σε λειτουργία, ποιά είναι η πιθανότητα να δουλέψει το μηχάνημα όταν του ζητηθεί;

40 Κάποιο εξάρτημα της μηχανής ενός πυραύλου αποτυγχάνει να λειτουργήσει 5% των φορών που ανάβει η μηχανή. Για να πετύχουμε μεγαλύτερη αξιοπιστία στη λειτουργία της μηχανής, φτιάχνουμε n αντίτυπα αυτού του εξαρτήματος. Τότε, η μηχανή δεν λειτουργεί μόνο αν και τα n αντίτυπα του εξαρτήματος δεν λειτουργήσουν. Υποθέτουμε ότι οι αποτυχίες των n εξαρτημάτων είναι ανεξάρτητες. Ποιά είναι η μικρότερη τιμή του n που μπορούμε να χρησιμοποιήσουμε για να εγγυηθούμε ότι η μηχανή θα δουλέψει στο 99% των δοκιμών;

41 Ρίχνουμε τρείς φορές ένα αμερόληπτο ζάρι. Αν ξέρουμε ότι η έδρα 1 εμφανίστηκε τουλάχιστον μία φορά, ποιά είναι η πιθανότητα να εμφανίστηκε ακριβώς μία φορά;

42 Σε μία τράπουλα των 52 χαρτιών, υπάρχουν 4 ρηγάδες. Επιλέγουμε τυχαία ένα χαρτί από την τράπουλα και σημειώνουμε την τιμή όψης του. Στη συνέχεια επιστρέφουμε το χαρτί στην τράπουλα. Η διαδικασία επαναλαμβάνεται 4 φορές. Υπολογίστε την πιθανότητα να προκύψουν ακριβώς 2 ρηγάδες στα 4 χαρτιά που επιλέξαμε, αν γνωρίζουμε ότι υπάρχει τουλάχιστον ένας ρήγας ανάμεσά τους.

43 Δείξτε ότι αν A, B και C είναι τρία ενδεχόμενα που ικανοποιούν τις $P(A \cap B \cap C) \neq 0$ και $P(C|A \cap B) = P(C|B)$, τότε $P(A|B \cap C) = P(A|B)$.

44 Ένας σκοπευτής επιχειρεί 12 ανεξάρτητες βολές προς ένα στόχο. Ποιά είναι η πιθανότητα να πετύχει τον στόχο τουλάχιστον μία φορά αν η πιθανότητά του να πετύχει τον στόχο είναι $9/10$ για κάθε δεδομένη βολή;

45 Ρίχνουμε ένα ζάρι 12 φορές. Υπολογίστε την πιθανότητα να φέρουμε
(α) 2 εξάρια·
(β) το πολύ δύο εξάρια.

46 Υποθέτουμε ότι η πιθανότητα να πετύχουμε ένα στόχο είναι $1/4$. Αν επιχειρήσουμε οκτώ βολές προς τον στόχο, ποιά είναι η πιθανότητα να τον πετύχουμε τουλάχιστον δύο φορές;

47 Στην Άσκηση 44, ποιά είναι η πιθανότητα ο στόχος να χτυπήθηκε τουλάχιστον δύο φορές αν είναι γνωστό ότι χτυπήθηκε τουλάχιστον μία φορά;

Κεφάλαιο 2

Συνδυαστική Ανάλυση

Στην Παράγραφο 1.2 είδαμε ότι ο συμμετρικός χώρος πιθανότητας με s σημεία είναι το μοντέλο για την τυχαία επιλογή ενός σημείου από ένα σύνολο S που έχει s σημεία. Στη συνέχεια, όταν θα μιλάμε για τυχαία επιλογή ενός σημείου από ένα πεπερασμένο σύνολο S, θα εννοούμε ότι η πιθανότητα κάθε μονοσυνόλου είναι s^{-1}, επομένως η πιθανότητα που δίνουμε σε ένα σύνολο με j σημεία είναι j/s.

Έστω $N(A)$ το πλήθος των σημείων του A. Αφού $P(A) = N(A)/s$, το πρόβλημα του υπολογισμού της $P(A)$ είναι ισοδύναμο με αυτό του υπολογισμού του $N(A)$. Η διαδικασία που ακολουθούμε για να βρούμε την $P(A)$ είναι να «μετρήσουμε» το πλήθος των σημείων του A και να διαιρέσουμε με το συνολικό πλήθος σημείων s. Μερικές φορές όμως η διαδικασία αντιστρέφεται. Αν για κάποιο λόγο γνωρίζουμε την $P(A)$, τότε μπορούμε να βρούμε το $N(A)$ χρησιμοποιώντας τον τύπο $N(A) = sP(A)$. Αυτή η αντίστροφη διαδικασία θα χρησιμοποιηθεί αρκετές φορές στη συνέχεια.

Ο υπολογισμός του $N(A)$ είναι εύκολος αν το A έχει λίγα μόνο σημεία, γιατί σε αυτήν την περίπτωση μπορούμε απλώς να απαριθμήσουμε όλα τα σημεία του A. Όμως, ακόμα κι αν το A έχει ένα μέτριο πλήθος σημείων, η μέθοδος της ευθείας απαρίθμησης είναι ανεφάρμοστη και η ανάγκη για απλούς κανόνες μετρήματος γίνεται επιτακτική. Σκοπός μας σε αυτό το κεφάλαιο είναι να δώσουμε μία όσο γίνεται λιγότερο τεχνική αλλά συστηματική παρουσίαση τεχνικών που είναι στοιχειώδεις και έχουν ευρύ φάσμα εφαρμογών. Το αντικείμενο αυτό πολύ σύντομα γίνεται δύσκολο, θα περιοριστούμε λοιπόν σε εκείνα τα ζητήματα που είναι τα πλέον χρήσιμα στην θεωρία πιθανοτήτων. Οι πρώτες τέσσερις παράγραφοι αυτού του κεφαλαίου περιέχουν το απαραίτητο υλικό, ενώ οι τελευταίες τέσσερις παράγραφοι αναπτύσσουν προαιρετικό και κάπως δυσκολότερο υλικό.

2.1 Διατεταγμένα δείγματα

Ας υποθέσουμε ότι έχουμε δύο σύνολα S και T. Αν το S έχει m σημεία $s_1, s_2, \ldots, s_m$ και το T έχει n σημεία $t_1, t_2, \ldots, t_n$, τότε το πλήθος των ζευγών (s_i, t_j) που μπορούμε να σχηματίσουμε παίρνοντας ένα σημείο από το σύνολο S και ένα σημείο από το σύνολο T είναι mn. Αυτό είναι φανερό αφού κάθε δοσμένο στοιχείο του συνόλου S μπορεί να σχηματίσει ζεύγος με οποιοδήποτε από τα n στοιχεία του συνόλου T.

Παράδειγμα 1. Αν $S = \{1, 2\}$ και $T = \{1, 2, 3\}$, τότε υπάρχουν έξη ζεύγη: τα $(1,1), (1,2), (1,3), (2,1), (2,2), (2,3)$. Προσέξτε ότι το ζεύγος $(1,2)$ είναι διαφορετικό από το ζεύγος $(2,1)$.

Πιό γενικά, ας υποθέσουμε ότι έχουμε n σύνολα $S_1, S_2, \ldots, S_n$ με πλήθος σημείων $s_1, s_2, \ldots, s_n$ αντίστοιχα. Τότε το πλήθος των n-άδων $(x_1, x_2, \ldots, x_n)$ που σχηματίζονται αν το x_1 είναι στοιχείο του S_1, το x_2 είναι στοιχείο του S_2, ..., το x_n είναι στοιχείο του S_n, είναι $s_1 s_2 \ldots s_n$. Ο ισχυρισμός αυτός είναι μία μάλλον προφανής επέκταση της περίπτωσης $n = 2$ που συζητήσαμε πιό πάνω. (Η τυπική απόδειξη του ότι το πλήθος των n-άδων είναι $s_1 s_2 \ldots s_n$ μπορεί να γίνει με επαγωγή ως προς n).

Σημαντική ειδική περίπτωση παίρνουμε αν καθένα από τα σύνολα $S_i, 1 \leq i \leq n$, είναι το ίδιο πάντα σύνολο S που έχει s στοιχεία. Τότε υπάρχουν s^n n-άδες $(x_1, x_2, \ldots, x_n)$ για τις οποίες κάθε x_i είναι ένα από τα σημεία του S.

Παράδειγμα 2. $S = \{1, 2\}$ και $n = 3$. Τότε υπάρχουν οκτώ n-άδες: $(1,1,1)$, $(1,1,2), (1,2,1), (1,2,2), (2,1,1), (2,1,2), (2,2,1), (2,2,2)$.

Η ειδική περίπτωση κατά την οποία τα σύνολα $S_i, 1 \leq i \leq n$ συμπίπτουν προσεγγίζεται και από μία άλλη οπτική γωνία. Υποθέτουμε ότι ένα δοχείο περιέχει s διακεκριμένους βώλους που φέρουν τους αριθμούς $1, 2, \ldots, s$. Επιλέγουμε έναν βώλο από το δοχείο, σημειώνουμε τον αριθμό του και τον επιστρέφουμε στο δοχείο. Η διαδικασία επαναλαμβάνεται n φορές. Καθεμία από τις n επιλογές δίνει έναν αριθμό από 1 ως s. Το αποτέλεσμα των n επιλογών περιγράφεται από την n-άδα $(x_1, x_2, \ldots, x_n)$, όπου x_1 είναι ο αριθμός πάνω στον πρώτη βώλο που επιλέξαμε, x_2 ο αριθμός πάνω στον δεύτερο, κλπ. Συνολικά, υπάρχουν s^n δυνατές n-άδες. Η διαδικασία αυτή λέγεται *δειγματοληψία με επανατοποθέτηση* από έναν πληθυσμό s διαφορετικών αντικειμένων. Το αποτέλεσμα $(x_1, x_2, \ldots, x_n)$ ονομάζεται δείγμα μεγέθους n από έναν πληθυσμό s αντικειμένων με επανατοποθέτηση. Μιλώντας για *τυχαία δειγματοληψία με επανατοποθέτηση*, υποθέτουμε ότι όλα τα s^n δυνατά δείγματα έχουν την ίδια πιθανότητα (είναι *ισοπίθανα*, στην παραδοσιακή γλώσσα).

Παράδειγμα 3. Ρίχνουμε n φορές ένα αμερόληπτο νόμισμα. Βρείτε την πιθανότητα να εμφανιστούν γράμματα τουλάχιστον μία φορά.

Η υπόθεση ότι το νόμισμα είναι αμερόληπτο πρακτικά σημαίνει ότι η πιθανότητα να εμφανιστούν γράμματα σε μία δεδομένη ρίψη είναι $1/2$. Αν είναι έτσι και αν υποθέσουμε ότι το να ρίξουμε το νόμισμα n φορές ισοδυναμεί με την επιλογή τυχαίου δείγματος μεγέθους n από τον πληθυσμό $\{\Gamma, K\}$, τότε καθένα από τα 2^n δυνατά αποτελέσματα έχει την ίδια πιθανότητα. Έστω A το ενδεχόμενο να εμφανιστούν γράμματα τουλάχιστον μία φορά, και A_i το ενδεχόμενο να εμφανιστούν γράμματα στην i-στή δοκιμή. Τότε $A = \bigcup_{i=1}^n A_i$. Όμως

$$\begin{aligned} P(A) &= 1 - P(A^c) \\ &= 1 - P\left(\left(\bigcup_{i=1}^n A_i\right)^c\right) \\ &= 1 - P\left(\bigcap_{i=1}^n A_i^c\right) \end{aligned}$$

και το $\bigcap_{i=1}^n A_i^c$ συμβαίνει αν και μόνο αν όλες οι n ρίψεις καταλήγουν σε κορώνα. Άρα, $P\left(\bigcap_{i=1}^n A_i^c\right) = 2^{-n}$, δηλαδή $P(A) = 1 - 2^{-n}$.

Έστω S ένα σύνολο με s αντικείμενα. Επιλέγουμε ένα αντικείμενο του S και σημειώνουμε ποιό είναι, ας υποθέσουμε όμως τώρα ότι δεν το επανατοποθετούμε στο σύνολο. Αν επαναλάβουμε αυτήν τη διαδικασία θα έχουμε να επιλέξουμε κάποιο από τα υπόλοιπα $(s-1)$ αντικείμενα. Ας υποθέσουμε ότι η διαδικασία επαναλαμβάνεται άλλες $n-1$ φορές, δηλαδή συνολικά επιλέγονται n αντικείμενα από το σύνολο. (Προφανώς πρέπει να έχουμε $n \leq s$ σε αυτήν την περίπτωση). Μπορούμε πάλι να περιγράψουμε το αποτέλεσμα σαν μία n-άδα $(x_1, x_2, \ldots, x_n)$, αυτή τη φορά όμως οι αριθμοί $x_1, x_2, \ldots, x_n$ πρέπει να είναι διακεκριμένοι. Δεν μπορούμε να έχουμε διπλές εμφανίσεις στο δείγμα μας. Το πρώτο αντικείμενο που επιλέξαμε μπορεί να είναι οποιοδήποτε από τα s, το δεύτερο μπορεί να είναι οποιοδήποτε από τα υπόλοιπα $s-1$, το τρίτο μπορεί να είναι οποιοδήποτε από τα υπόλοιπα $s-2$, κλπ., άρα υπάρχουν $(s)_n = s(s-1)\ldots(s-n+1)$ διαφορετικά δυνατά αποτελέσματα για το πείραμα. Η διαδικασία ονομάζεται *δειγματοληψία n αντικειμένων χωρίς επανατοποθέτηση* από έναν πληθυσμό s διακεκριμένων αντικειμένων. Μιλάμε για ένα *τυχαίο δείγμα μεγέθους n μέσα από έναν πληθυσμό s αντικειμένων χωρίς επανατοποθέτηση* αν υποθέτουμε ότι τα $(s)_n$ αποτελέσματα είναι ισοπίθανα.

Συμβολίσαμε το γινόμενο $s(s-1)\ldots(s-n+1)$ με $(s)_n$. Ειδικότερα, $(s)_s = s(s-1)\ldots 1 = s!$. Η επιλογή ενός δείγματος μεγέθους s από έναν πληθυσμό s αντικειμένων είναι ισοδύναμη με το να γράψουμε τους αριθμούς $1, 2, \ldots, s$ με κάποια σειρά. Δηλαδή ο $s!$ περιγράφει το πλήθος των διαφορετικών διατάξεων (ή μεταθέσεων) s αντικειμένων.

Ας υποθέσουμε ότι επιλέγουμε τυχαίο δείγμα μεγέθους n από ένα σύνολο s αντικειμένων με επανατοποθέτηση. Ζητάμε την πιθανότητα του ενδεχομένου A κανένα σημείο να μην εμφανίζεται δύο φορές στο δείγμα. Η λύση του προβλήματος είναι απλή. Το πλήθος των δειγμάτων μεγέθους n με επανατοποθέτηση είναι s^n. Από αυτά τα s^n τυχαία δείγματα, το πλήθος εκείνων στα οποία κανένα σημείο δεν εμφανίζεται δύο φορές είναι το ίδιο με το πλήθος των δειγμάτων μεγέθους n που επιλέγονται από s αντικείμενα *χωρίς επανατοποθέτηση*, δηλ. $(s)_n$. Αφού τα s^n δείγματα είναι ισοπίθανα, βλέπουμε ότι η ζητούμενη πιθανότητα είναι

$$(1) \qquad \frac{(s)_n}{s^n} = \frac{s(s-1)\ldots(s-n+1)}{s^n}$$

$$= \left(1-\frac{1}{s}\right)\left(1-\frac{2}{s}\right)\cdots\left(1-\frac{n-1}{s}\right).$$

Παράδειγμα 4. Μία νέα και μάλλον απρόσμενη εφαρμογή της (1) συναντάμε στο λεγόμενο πρόβλημα των γενεθλίων. Υποθέτουμε ότι η γενέθλια ημερομηνία ενός ανθρώπου είναι εξίσου πιθανό να συμπέσει με οποιαδήποτε από τις 365 μέρες του χρόνου. (Εδώ αγνοούμε τα δίσεκτα χρόνια και το γεγονός ότι οι ρυθμοί γεννήσεων δεν είναι ακριβώς ομοιόμορφοι στην διάρκεια του χρόνου). Βρείτε την πιθανότητα p σε μία ομάδα n ατόμων να μην υπάρχουν δύο που έχουν γενέθλια την ίδια μέρα.

Σε αυτό το πρόβλημα $s = 365$, οπότε εφαρμόζοντας την (1) βλέπουμε ότι

$$p = \left(1-\frac{1}{365}\right)\left(1-\frac{2}{365}\right)\cdots\left(1-\frac{n-1}{365}\right).$$

Οι αριθμητικές συνέπειες είναι τελείως απρόβλεπτες. Ακόμα και για n τόσο μικρό όσο το 23, έχουμε $p < 1/2$, ενώ για $n = 56$, είναι $p = 0.01$. Δηλαδή, σε μία ομάδα 23 ατόμων η πιθανότητα να συμπίπτουν τα γενέθλια τουλάχιστον δύο από αυτά ξεπερνάει το 1/2. Σε μία ομάδα 56 ατόμων, είναι σχεδόν βέβαιο ότι δύο έχουν κοινά γενέθλια.

Αν έχουμε έναν πληθυσμό s αντικειμένων, μπορούμε να επιλέξουμε s^n δείγματα μεγέθους n με επανατοποθέτηση και $(s)_n$ δείγματα μεγέθους n χωρίς επανατοποθέτηση. Αν το s είναι μεγάλο σε σχέση με το n, τότε η διαφορά ανάμεσα στις δύο μεθόδους τυχαίας δειγματοληψίας είναι μικρή. Πράγματι, από την (1) βλέπουμε ότι για κάθε σταθερό n,

$$\lim_{s\to\infty}\frac{(s)_n}{s^n} = \lim_{s\to\infty}\left(1-\frac{1}{s}\right)\left(1-\frac{2}{s}\right)\cdots\left(1-\frac{n-1}{s}\right) = 1.$$

(Για πιό ακριβείς εκτιμήσεις δείτε την Ασκηση 12.)

2.2 Μεταθέσεις

Ας υποθέσουμε ότι έχουμε n δοχεία και n βώλους. Ο συνολικός αριθμός των τρόπων με τους οποίους μπορούμε να κατανείμουμε τους n βώλους στα n δοχεία έτσι ώστε *κάθε δοχείο να περιέχει ακριβώς έναν βώλο* είναι $n!$. Λέγοντας ότι οι n βώλοι κατανέμονται τυχαία στα n δοχεία με έναν βώλο σε κάθε δοχείο, εννοούμε ότι δίνουμε πιθανότητα $1/n!$ σε καθέναν από τους δυνατούς τρόπους. Ποιά είναι η πιθανότητα ένας συγκεκριμένη βώλος, ας πούμε ο i-στός βώλος, να βρεθεί σε ένα συγκεκριμένο δοχείο, ας πούμε στο j-στό δοχείο; Αν ο i-στός βώλος είναι στο j-στό δοχείο, μένουν $(n-1)$ δοχεία και $(n-1)$ βώλοι που πρέπει να κατανεμηθούν σε αυτά έτσι ώστε ακριβώς ένας βώλος να βρεθεί σε κάθε δοχείο. Αυτό μπορεί να γίνει με $(n-1)!$ τρόπους, άρα η ζητούμενη πιθανότητα είναι $(n-1)!/n! = 1/n$.

Ένας άλλος τρόπος για να δούμε αυτό το αποτέλεσμα είναι ο εξής. Αν έχουμε n διακεκριμένα αντικείμενα και κάνουμε τυχαία μετάθεσή τους, τότε η πιθανότητα ένα συγκεκριμένο αντικείμενο να βρεθεί σε μία συγκεκριμένη θέση είναι $1/n$. Πράγματι, μπορούμε να ταυτίσουμε τις θέσεις με τα δοχεία και τα αντικείμενα με τους βώλους.

Ο παραπάνω συλλογισμός γενικεύεται εύκολα από 1 σε $k \geq 1$ αντικείμενα. Αν μεταθέσουμε τυχαία n αντικείμενα, η πιθανότητα k συγκεκριμένα αντικείμενα να βρεθούν σε k συγκεκριμένες θέσεις είναι $(n-k)!/n!$. Αφήνουμε την απόδειξη σαν άσκηση για τον αναγνώστη.

Τα προβλήματα που αφορούν τυχαίες μεταθέσεις μπορούν να πάρουν διάφορες μορφές αν διατυπωθούν με λόγια. Δίνουμε δύο παραδείγματα:

(α) Ανακατεύουμε μία τράπουλα με τα χαρτιά $1, 2, \dots, n$, και στη συνέχεια τραβάμε ένα-ένα τα χαρτιά. Ποιά είναι η πιθανότητα, για κάποιο συγκεκριμένο i, το i-στό χαρτί που τραβήξαμε να φέρει τον αριθμό i;

(β) Ας υποθέσουμε ότι 10 ζευγάρια πηγαίνουν σε μιά γιορτή. Τα αγόρια και τα κορίτσια σχηματίζουν τυχαία ζευγάρια. Ποιά είναι πιθανότητα ακριβώς k δεδομένα αγόρια να βρεθούν με τα κορίτσια τους;

Ένα πιό πολύπλοκο πρόβλημα τυχαίων μεταθέσεων είναι να βρεθεί η πιθανότητα να υπάρχουν ακριβώς k «συμπτώσεις». Για να χρησιμοποιήσουμε το συνηθισμένο μας παράδειγμα με τους βώλους και τα δοχεία, το πρόβλημα είναι να βρεθεί η πιθανότητα ο i-στός βώλος να βρεθεί στο i-στό δοχείο για ακριβώς k διαφορετικές τιμές του i.

Το πρόβλημα των συμπτώσεων λύνεται με πολλούς διαφορετικούς τρόπους. Αναβάλλουμε την συζήτηση αυτού του θέματος για την Παράγραφο 2.6.

2.3 Συνδυασμοί (μη διατεταγμένα δείγματα)

Ένα χέρι του πόκερ αποτελείται από πέντε χαρτιά που επιλέγονται από μία τράπουλα των 52 χαρτιών. Από την οπτική γωνία της προηγούμενης παραγράφου, υπάρχουν $(52)_5$ τέτοια χέρια. Αν κάνουμε έτσι τον υπολογισμό, διαφορετικές διατάξεις των ίδιων πέντε χαρτιών θεωρούνται διαφορετικά χέρια. Δηλαδή, η πεντάδα 2,3,4,5,6 σπαθιά με αυτήν τη διάταξη, θεωρείται διαφορετική από την πεντάδα 2,4,3,5,6 σπαθιά με αυτήν τη διάταξη. Όταν όμως παίζουμε μία παρτίδα πόκερ, αυτά τα δύο χέρια είναι ίδια. Για την ακρίβεια, όλες οι $5!$ μεταθέσεις των ίδιων πέντε χαρτιών είναι ισοδύναμες. Από τα $(52)_5$ δυνατά χέρια, ακριβώς $5!$ από αυτά είναι απλώς μεταθέσεις αυτών των ίδιων πέντε χαρτιών. Όμοια, για κάθε δοσμένο σύνολο πέντε χαρτιών υπάρχουν $5!$ διαφορετικές μεταθέσεις. Άρα το συνολικό πλήθος χεριών του πόκερ, *αν αγνοήσουμε την σειρά με την οποία εμφανίζονται τα χαρτιά*, είναι $(52)_5/5!$. Με αυτό το νέο μέτρημα δύο χέρια θεωρούνται διαφορετικά αν και μόνο αν διαφέρουν σαν σύνολα αντικειμένων, δηλαδή, αν έχουν τουλάχιστον ένα διαφορετικό στοιχείο. Για παράδειγμα, ανάμεσα στα $(52)_5/5!$ χέρια του πόκερ, τα χέρια $(2, 3, 4, 5, 6)$ σπαθιά και $(3, 2, 4, 5, 6)$ σπαθιά συμπίπτουν, αλλά τα χέρια $(2, 3, 4, 5, 7)$ σπαθιά και $(2, 3, 4, 5, 6)$ σπαθιά είναι διαφορετικά.

Πιό γενικά, ας υποθέσουμε ότι έχουμε ένα σύνολο S από s διαφορετικά αντικείμενα. Τότε, όπως εξηγήσαμε προηγουμένως, υπάρχουν $(s)_r$ διαφορετικά δείγματα μεγέθους r που μπορούμε να επιλέξουμε από το S χωρίς επανατοποθέτηση. Κάθε διακεκριμένο υποσύνολο $\{x_1, \ldots, x_r\}$ από r στοιχεία του S μπορεί να διαταχθεί με $r!$ διαφορετικούς τρόπους. Αν αποφασίσουμε να αγνοήσουμε τη σειρά με την οποία τα αντικείμενα εμφανίζονται στο δείγμα, τότε αυτές οι $r!$ αναδιατάξεις των $x_1, \ldots, x_r$ πρέπει να θεωρηθούν ταυτόσημες. Υπάρχουν λοιπόν $(s)_r/r!$ *διαφορετικά δείγματα μεγέθους r που μπορούμε να επιλέξουμε χωρίς επανατοποθέτηση και χωρίς να μάς ενδιαφέρει η διάταξη από ένα σύνολο s διακεκριμένων αντικειμένων.*

Η ποσότητα $(s)_r/r!$ γράφεται συνήθως με τη βοήθεια του συμβόλου του διωνυμικού συντελεστή:

$$\frac{(s)_r}{r!} = \binom{s}{r}.$$

Παρατηρήστε ότι για $r = 0, 1, 2, \ldots, s$

$$\binom{s}{r} = \frac{(s)_r}{r!} = \frac{s!}{r!(s-r)!}.$$

Σημειώνουμε εδώ για μελλοντική χρήση ότι το σύμβολο $\binom{a}{r}$ είναι καλά ορισμένο για κάθε πραγματικό αριθμό a και μη αρνητικό ακέραιο r, από την

$$\binom{a}{r} = \frac{(a)_r}{r!} = \frac{a(a-1)\ldots(a-r+1)}{r!}, \tag{3}$$

όπου τα $0!$ και $(a)_0$ είναι και τα δύο εξ ορισμού ίσα με 1.

Παράδειγμα 5.

$$\begin{aligned}\binom{-\pi}{3} &= \frac{(-\pi)(-\pi-1)(-\pi-2)}{3!} \\ &= -\frac{\pi(\pi+1)(\pi+2)}{3!}.\end{aligned}$$

Παρατηρούμε ότι για κάθε θετικό ακέραιο a έχουμε $\binom{a}{r} = 0$ αν $r > a$. Αν ο r είναι αρνητικός ακέραιος, υιοθετούμε την σύμβαση $\binom{a}{r} = 0$. Έτσι, το σύμβολο $\binom{a}{r}$ είναι ορισμένο για κάθε πραγματικό αριθμό a και ακέραιο r.

Όπως έχουμε ήδη παρατηρήσει, αν s είναι ένας θετικός ακέραιος και r είναι ένας μη αρνητικός ακέραιος, είναι χρήσιμο να σκεφτόμαστε τον $\binom{s}{r}$ σαν το πλήθος των τρόπων με τους οποίους μπορούμε να επιλέξουμε δείγμα μεγέθους r από έναν πληθυσμό s διακεκριμένων αντικειμένων χωρίς επανατοποθέτηση και χωρίς να μάς ενδιαφέρει η σειρά με την οποία επιλέγονται αυτά τα r αντικείμενα.

Παράδειγμα 6. Θεωρούμε το σύνολο αριθμών $\{1, 2, \ldots, n\}$. Τότε, αν $1 \leq r \leq n$, υπάρχουν ακριβώς $\binom{n}{r}$ επιλογές αριθμών $i_1, i_2, \ldots, i_r$ με $1 \leq i_1 < i_2 < \ldots < i_r \leq n$. Πράγματι, καθεμία από τις $(n)_r$ επιλογές r *διακεκριμένων* αριθμών από 1 ως n έχει $r!$ αναδιατάξεις από τις οποίες ακριβώς μία ικανοποιεί τις απαιτήσεις μας. Άρα, το πλήθος των διαφορετικών επιλογών αριθμών που ικανοποιούν τις απαιτήσεις μας είναι το ίδιο με το πλήθος των διαφορετικών υποσυνόλων μεγέθους r που μπορούν να επιλεγούν από το σύνολο $\{1, 2, \ldots, n\}$.

Παράδειγμα 7. *Σύνθεση επιτροπής.* Το τμήμα μαθηματικών αποτελείται από 25 καθηγητές πρώτης βαθμίδας, 15 αναπληρωτές καθηγητές, και 35 επίκουρους καθηγητές. Μία εξαμελής επιτροπή επιλέγεται τυχαία από τα μέλη του τμήματος. Βρείτε την πιθανότητα όλα τα μέλη της επιτροπής να είναι επίκουροι καθηγητές.

Συνολικά, το τμήμα έχει 75 μέλη. Η επιτροπή των 6 μπορεί να επιλεγεί από τους 75 με $\binom{75}{6}$ τρόπους. Υπάρχουν 35 επίκουροι καθηγητές, και οι 6 της επιτροπής μπορούν να επιλεγούν από τους 35 με $\binom{35}{6}$ τρόπους. Άρα η ζητούμενη πιθανότητα είναι $\binom{35}{6}/\binom{75}{6}$. Ένας υπολογισμός δείχνει ότι η προσεγγιστική τιμή είναι 0.01. Άρα, το μόνιμο προσωπικό (αναπληρωτές και καθηγητές πρώτης βαθμίδας) δεν χρειάζεται να ανησυχεί και τόσο για το ενδεχόμενο έλλειψης αντιπροσώπευσης.

Παράδειγμα 8. Θεωρούμε ένα χέρι του πόκερ με πέντε χαρτιά. Βρείτε την πιθανότητα να περιέχει τέσσερα χαρτιά με την ίδια τιμή όψης, αν υποθέσουμε ότι τα πέντε χαρτιά επιλέγονται τυχαία.

Μπορούμε να λύσουμε το πρόβλημα ως εξής.

Υπάρχουν $\binom{52}{5}$ διαφορετικά χέρια, που θέλουμε να είναι ισοπίθανα. Άρα ο Ω θα έχει $\binom{52}{5}$ σημεία. Για να συμβεί το ενδεχόμενο που μάς ενδιαφέρει, πρέπει να έχουμε τέσσερα χαρτιά με την ίδια τιμή όψης. Υπάρχουν 13 διαφορετικές επιλογές για την τιμή που θα έχει αυτή η τετράδα, οι $2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A$. Για κάθε τέτοια επιλογή (που προσδιορίζει τέσσερα από τα πέντε χαρτιά στο χέρι) υπάρχουν 48 ακόμα χαρτιά από τα οποία μπορούμε να επιλέξουμε το πέμπτο χαρτί στο χέρι. Αφού καθεμία από τις 13 επιλογές της τετράδας συνδυάζεται με οποιαδήποτε από τις 48 επιλογές για το πέμπτο χαρτί, υπάρχουν συνολικά (13)(48) δυνατοί τρόποι να πάρουμε χέρι του πόκερ με τετράδα. Επομένως η ζητούμενη πιθανότητα είναι

$$\frac{(13)(48)}{\binom{52}{5}} \approx 2.40 \times 10^{-4}.$$

Παράδειγμα 9. Υποθέτουμε ότι n βώλοι κατανέμονται σε n δοχεία έτσι ώστε όλες οι n^n δυνατές κατανομές να είναι ισοπίθανες. Βρείτε την πιθανότητα να μείνει άδειο μόνο το δοχείο 1.

Ο χώρος πιθανότητας σε αυτήν την περίπτωση αποτελείται από n^n ισοπίθανα σημεία. Έστω A το ενδεχόμενο να μείνει άδειο μόνο το δοχείο 1. Αυτό μπορεί να συμβεί μόνο αν οι n βώλοι τοποθετηθούν στα υπόλοιπα $n-1$ δοχεία με τέτοιο τρόπο ώστε κανένα από αυτά να μην μείνει άδειο. Έτσι, ακριβώς ένα από αυτά τα $(n-1)$ δοχεία πρέπει να περιέχει δύο βώλους, και τα υπόλοιπα $(n-2)$ δοχεία πρέπει να περιέχουν ακριβώς έναν βώλο το καθένα. Έστω B_j το ενδεχόμενο το δοχείο j, $j = 2, 3, \ldots, n$, να περιέχει δύο βώλους, το δοχείο 1 να είναι άδειο, και τα υπόλοιπα $(n-2)$ δοχεία να περιέχουν ακριβώς έναν βώλο το καθένα. Τότε τα B_j είναι ξένα και $A = \bigcup_{j=2}^{n} B_j$. Για να υπολογίσουμε την $P(B_j)$ παρατηρούμε ότι οι δύο βώλοι στο δοχείο j επιλέγονται από τους n βώλους με $\binom{n}{2}$ τρόπους. Οι $(n-2)$ βώλοι που απομένουν τακτοποιούνται στα υπόλοιπα $(n-2)$ δοχεία με $(n-2)!$ τρόπους. Άρα το πλήθος των διαφορετικών τρόπων με τους οποίους μπορούμε να τοποθετήσουμε δύο βώλους στο δοχείο j, κανέναν βώλο στο δοχείο 1, και ακριβώς έναν βώλο σε καθένα από τα υπόλοιπα δοχεία είναι $\binom{n}{2}(n-2)!$. Επομένως

$$P(B_j) = \frac{\binom{n}{2}(n-2)!}{n^n}$$

και τελικά

$$P(A) = \frac{(n-1)\binom{n}{2}(n-2)!}{n^n} = \frac{\binom{n}{2}(n-1)!}{n^n}.$$

2.4 Διαμερίσεις

Πολλά συνδυαστικά προβλήματα που αφορούν μη διατεταγμένα δείγματα είναι ουσιαστικά ισοδύναμα με το εξής. Ένα δοχείο περιέχει r κόκκινους βώλους και b μαύρους βώλους. Επιλέγουμε τυχαίο δείγμα μεγέθους n από το δοχείο χωρίς επανατοποθέτηση. Ποιά είναι η πιθανότητα στο δείγμα να περιέχονται ακριβώς k κόκκινοι βώλοι (άρα και $n-k$ μαύροι βώλοι);

Για να λύσουμε το πρόβλημα σκεφτόμαστε ως εξής. Μάς ενδιαφέρει μόνο πόσοι συνολικά είναι οι κόκκινοι και οι μαύροι βώλοι στο δείγμα, και όχι η σειρά με την οποία επιλέχθηκαν. Δηλαδή, το αντικείμενό μας είναι δειγματοληψία χωρίς επανατοποθέτηση και χωρίς να παίζει ρόλο η διάταξη. Μπορούμε λοιπόν να πάρουμε σαν χώρο πιθανότητας του προβλήματος την συλλογή όλων των $\binom{b+r}{n}$ δειγμάτων μεγέθους n που μπορεί κανείς να επιλέξει με αυτόν τον τρόπο από τους $b+r$ συνολικά βώλους. Σε καθένα από αυτά τα $\binom{b+r}{n}$ δείγματα αντιστοιχεί η ίδια πιθανότητα $\binom{b+r}{n}^{-1}$. Πρέπει τώρα να υπολογίσουμε το πλήθος των τρόπων με τους οποίους μπορεί να επιλεγεί δείγμα μεγέθους n έτσι ώστε να περιέχει ακριβώς k κόκκινους βώλους. Οι k κόκκινοι βώλοι επιλέγονται από τους r κόκκινους βώλους με $\binom{r}{k}$ τρόπους αν δεν μας ενδιαφέρει η διάταξη, και οι $n-k$ μαύροι βώλοι επιλέγονται από τους b μαύρους βώλους με $\binom{b}{n-k}$ τρόπους αν δεν μας ενδιαφέρει η διάταξη. Αφού κάθε επιλογή από k κόκκινους βώλους μπορεί να συνδυαστεί με οποιαδήποτε επιλογή από $n-k$ μαύρους βώλους, υπάρχουν συνολικά $\binom{r}{k}\binom{b}{n-k}$ δυνατές επιλογές. Επομένως η ζητούμενη πιθανότητα είναι

$$\frac{\binom{r}{k}\binom{b}{n-k}}{\binom{r+b}{n}}.$$

Η ουσία των προβλημάτων αυτού του τύπου είναι ότι ο πληθυσμός (σε αυτήν την περίπτωση οι βώλοι) διαμερίζεται σε δύο κλάσεις (κόκκινους και μαύρους βώλους). Παίρνουμε τυχαίο δείγμα κάποιου μεγέθους και ζητάμε την πιθανότητα το δείγμα να περιέχει συγκεκριμένο αριθμό αντικειμένων από καθεμία από τις δύο κλάσεις.

Σε μερικά προβλήματα αυτού του είδους οι δύο κλάσεις δεν προσδιορίζονται σαφώς, μπορεί όμως κανείς να τις αναγνωρίσει αν αναλύσει τη γλώσσα του προβλήματος.

Παράδειγμα 10. Ένα χέρι του πόκερ αποτελείται από πέντε χαρτιά που επιλέγονται από μία συνήθη τράπουλα των 52 χαρτιών. Βρείτε την πιθανότητα σε ένα χέρι του πόκερ να περιέχονται ακριβώς 2 ρηγάδες.

Για να λύσουμε το πρόβλημα παρατηρούμε ότι υπάρχουν $\binom{52}{5}$ χέρια του πόκερ. Στην τράπουλα υπάρχουν 4 ρηγάδες και 48 άλλα χαρτιά. Αυτό διαμερίζει την τράπουλα σε δύο κλάσεις, τους ρηγάδες και τα άλλα χαρτιά, που περιέχουν 4 και 48 αντικείμενα αντίστοιχα. Το χέρι του πόκερ είναι ένα δείγμα μεγέθους 5 που επιλέγεται χωρίς επανατοποθέτηση και χωρίς να μάς ενδιαφέρει η διάταξη από τα 52 χαρτιά. Το ερώτημα λοιπόν είναι να βρούμε την πιθανότητα το δείγμα να έχει 2 στοιχεία από την πρώτη κλάση και 3 στοιχεία από την δεύτερη κλάση. Άρα η ζητούμενη πιθανότητα είναι

$$\frac{\binom{4}{2}\binom{48}{3}}{\binom{52}{5}} \approx 3.99 \times 10^{-2}.$$

Παράδειγμα 11. Μία τράπουλα αποτελείται από 4 σειρές των 13 χαρτιών, μπαστούνια, καρό, κούπες και σπαθιά.

(α) Ποιά είναι η πιθανότητα ένα χέρι των πέντε χαρτιών να περιέχει ακριβώς 3 μπαστούνια;
(β) Ποιά είναι η πιθανότητα ένα χέρι των πέντε χαρτιών να περιέχει ακριβώς 3 χαρτιά από την ίδια σειρά;

Για να λύσουμε το πρόβλημα (α) παρατηρούμε ότι οι συνθήκες του προβλήματος διαμερίζουν την τράπουλα των 52 χαρτιών σε 2 κλάσεις. Η πρώτη είναι τα «μπαστούνια» και έχει 13 μέλη, ενώ η δεύτερη είναι τα «υπόλοιπα χαρτιά» και έχει 39 μέλη. Το χέρι των 5 χαρτιών είναι ένα δείγμα μεγέθους 5 από τον πληθυσμό των 52 χαρτιών, και το πρόβλημα ζητάει 3 από τα 5 να προέρχονται από την πρώτη κλάση. Αρα η ζητούμενη πιθανότητα είναι

$$p = \frac{\binom{13}{3}\binom{39}{2}}{\binom{52}{5}} \approx 8.15 \times 10^{-2}.$$

Για να λύσουμε το (β) θέτουμε A_1 το ενδεχόμενο ακριβώς τρία από τα χαρτιά να είναι μπαστούνια, A_2 το ενδεχόμενο ακριβώς τρία από τα χαρτιά να είναι καρό, A_3 το ενδεχόμενο ακριβώς τρία από τα χαρτιά να είναι κούπες, και A_4 το ενδεχόμενο ακριβώς τρία από τα χαρτιά να είναι σπαθιά. Τότε, αφού κάθε χέρι αποτελείται από 5 μόνο χαρτιά, τα ενδεχόμενα A_1, A_2, A_3, A_4 είναι ξένα. Η ένωσή τους $A_1 \cup A_2 \cup A_3 \cup A_4$ είναι το ενδεχόμενο ακριβώς 3 από τα 5 χαρτιά να ανήκουν στην ίδια σειρά. Αρα η ζητούμενη πιθανότητα είναι $4p$.

Παράδειγμα 12. Θεωρούμε πάλι ένα χέρι του πόκερ με 5 χαρτιά. Ποιά είναι η πιθανότητα να έχουμε φούλ (δηλαδή, δύο χαρτιά με την ίδια τιμή όψης και μιά τριάδα χαρτιών με την ίδια τιμή όψης), αν υποθέσουμε ότι τα χαρτιά επιλέγονται τυχαία από την τράπουλα;

Για να λύσουμε το πρόβλημα παρατηρούμε πάλι ότι υπάρχουν $\binom{52}{5}$ ισοπίθανα χέρια του πόκερ. Από αυτά πρέπει τώρα να υπολογίσουμε το πλήθος των τρόπων με τους οποίους μπορούμε να έχουμε ένα ζευγάρι και μία τριάδα. Ας δούμε πρώτα με πόσους τρόπους μπορούμε να επιλέξουμε μία *συγκεκριμένη τριάδα*, ας πούμε 3 άσους, και ένα *συγκεκριμένο ζευγάρι*, ας πούμε 2 ρηγάδες. Η τριάδα αποτελείται από 3 χαρτιά που πρέπει να επιλεγούν από τους τέσσερις άσους χωρίς να μάς ενδιαφέρει η διάταξη, και αυτό μπορεί να γίνει με $\binom{4}{3}$ τρόπους. Το ζευγάρι αποτελείται από δύο χαρτιά που πρέπει να επιλεγούν από τους τέσσερις ρηγάδες χωρίς να μάς ενδιαφέρει η διάταξη. Αυτό μπορεί να γίνει με $\binom{4}{2}$ τρόπους. Το πλήθος των τρόπων με τους οποίους μπορούμε να πάρουμε μία τριάδα από άσους και ένα ζευγάρι από ρηγάδες είναι επομένως $\binom{4}{3}\binom{4}{2}$. Άρα η πιθανότητα να πάρουμε ένα χέρι του πόκερ με τρείς άσους και δύο ρηγάδες είναι $\binom{4}{3}\binom{4}{2}/\binom{52}{5} = p$. Φυσικά, η πιθανότητα αυτή είναι η ίδια για κάθε συγκεκριμένο ζευγάρι και κάθε συγκεκριμένη τριάδα. Αλλά η τιμή όψης των χαρτιών της τριάδας μπορεί να είναι οποιαδήποτε από τις υπάρχουσες 13, και η τιμή όψης των χαρτιών του ζευγαριού μπορεί να είναι οποιαδήποτε από τις υπόλοιπες 12 τιμές όψης. Αφού καθεμία από τις 13 τιμές για την τριάδα μπορεί να σχετιστεί με οποιαδήποτε από τις 12 τιμές για το ζευγάρι, υπάρχουν $(13)(12)$ τέτοιες επιλογές. Όλες αυτές οι επιλογές συνιστούν ξένα ενδεχόμενα που έχουν την ίδια πιθανότητα p, άρα η ζητούμενη πιθανότητα είναι

$$(13)(12)p = \frac{(13)(12)(4)(6)}{\binom{52}{5}} \approx 1.44 \times 10^{-3}.$$

Παράδειγμα 13. Σε ένα χέρι του πόκερ, ποιά είναι η πιθανότητα να πάρουμε ακριβώς δύο ζευγάρια; Εδώ, ένα χέρι της μορφής $(2, 2, 2, 2, x)$ δεν μετράει σαν δύο ζευγάρια αλλά σαν μία τετράδα.

Για να λύσουμε το πρόβλημα παρατηρούμε ότι αν το χέρι έχει δύο ζευγάρια, τότε δύο από τα χαρτιά έχουν την ίδια τιμή όψης x, δύο από τα χαρτιά έχουν την ίδια τιμή όψης $y \neq x$, και το πέμπτο χαρτί έχει τιμή όψης διαφορετική από τις x και y. Τώρα, υπάρχουν 13 διαφορετικές τιμές όψης. Οι τιμές των δύο ζευγαριών μπορούν να επιλεγούν από αυτές με $\binom{13}{2}$ τρόπους. Το άλλο χαρτί μπορεί να έχει οποιαδήποτε από τις υπόλοιπες 11 τιμές. Τα δύο χαρτιά με την τιμή x επιλέγονται από τα τέσσερα που έχουν αυτήν την τιμή με $\binom{4}{2}$ τρόπους, και όμοια για τα δύο με την τιμή y. Το χαρτί που απομένει, με τιμή z, επιλέγεται από τα τέσσερα που έχουν αυτήν την τιμή με $\binom{4}{1} = 4$ τρόπους. Αρα το πλήθος των επιλογών είναι $\binom{13}{2}(11)\binom{4}{2}\binom{4}{2}(4)$, και η ζητούμενη πιθανότητα είναι

$$\frac{\binom{13}{2}(11)(6)(6)(4)}{\binom{52}{5}} \approx 4.75 \times 10^{-2}.$$

Σε κάποια προβλήματα που αφορούν διαμερίσεις, μπορούμε να φανταζόμαστε τις κλάσεις όπως στο εξής.

Παράδειγμα 14. Υποθέτουμε ότι έχουμε ένα δοχείο που περιέχει r βώλους που φέρουν τους αριθμούς $1, 2, \ldots, r$. Ένα τυχαίο δείγμα μεγέθους n επιλέγεται χωρίς επανατοποθέτηση, και σημειώνουμε τους αριθμούς από τους βώλους του δείγματος.

Κατόπιν οι βώλοι επανατοποθετούνται στο δοχείο, και ένα δεύτερο δείγμα μεγέθους m επιλέγεται χωρίς επανατοποθέτηση. Βρείτε την πιθανότητα τα δύο δείγματα να έχουν ακριβώς k κοινούς βώλους.

Για να λύσουμε αυτό το πρόβλημα σκεφτόμαστε ως εξής. Το αποτέλεσμα του πρώτου δείγματος είναι να διαμερίσει τους βώλους σε δύο κλάσεις, εκείνους τους n που επιλέχθηκαν και εκείνους τους $r-n$ που δεν επιλέχθηκαν. (Μπορούμε να φανταζόμαστε ότι οι n βώλοι που επιλέχθηκαν στο δείγμα βάφονται κόκκινοι πριν επιστραφούν στο δοχείο). Το πρόβλημα λοιπόν είναι να βρεθεί η πιθανότητα το δείγμα μεγέθους m να περιέχει ακριβώς k βώλους από την πρώτη κλάση, οπότε η ζητούμενη πιθανότητα είναι

$$\frac{\binom{n}{k}\binom{r-n}{m-k}}{\binom{r}{m}}.$$

Αν στο επιχείρημα αντιστρέφαμε τους ρόλους των δειγμάτων, και σκεφτόμαστε ότι οι βώλοι σημειώνονται με βάση το δεύτερο δείγμα, τότε θα βρίσκαμε ότι η πιθανότητα είναι

$$\frac{\binom{m}{k}\binom{r-m}{n-k}}{\binom{r}{n}}.$$

Αφήνουμε σαν άσκηση να επαληθεύσετε ότι αυτοί οι δύο αριθμοί είναι ίσοι.

Μπορούμε εύκολα να επεκτείνουμε τη μελέτη της διαμέρισης ενός πληθυσμού σε δύο κλάσεις, σε αυτήν της διαμέρισής του σε $m \geq 2$ κλάσεις. Ας υποθέσουμε ότι έχουμε ένα σύνολο r αντικειμένων έτσι ώστε καθένα από τα αντικείμενα να ανήκει σε ένα από m είδη. Ο πληθυσμός αποτελείται από r_1 αντικείμενα τύπου 1, r_2 αντικείμενα τύπου 2, …, r_m αντικείμενα τύπου m, όπου $r_1 + r_2 + \ldots + r_m = r$. Αν επιλέξουμε τυχαίο δείγμα μεγέθους n χωρίς επανατοποθέτηση από τον πληθυσμό αυτών των r αντικειμένων, ποιά είναι η πιθανότητα στο δείγμα να περιέχονται ακριβώς k_1 αντικείμενα τύπου 1, …, k_m αντικείμενα τύπου m, όπου $k_1 + \ldots + k_m = n$;

Για μία ακόμα φορά ο χώρος πιθανότητας είναι η συλλογή όλων των $\binom{r}{n}$ ισοπίθανων δειγμάτων μεγέθους n που μπορούμε να επιλέξουμε χωρίς επανατοποθέτηση και χωρίς να μάς ενδιαφέρει η διάταξη από τα r αντικείμενα του πληθυσμού. Τα k_i αντικείμενα τύπου i στο δείγμα επιλέγονται από τα r_i αντικείμενα αυτού του τύπου με $\binom{r_i}{k_i}$ τρόπους αν δεν μάς ενδιαφέρει η διάταξη. Άρα η πιθανότητα να επιλέξουμε δείγμα που να ικανοποιεί τους παραπάνω περιορισμούς είναι

$$\frac{\binom{r_1}{k_1}\binom{r_2}{k_2}\cdots\binom{r_m}{k_m}}{\binom{r}{n}}.$$

Παράδειγμα 15. Σε ένα χέρι 13 χαρτιών που επιλέγεται τυχαία από μιά συνηθισμένη τράπουλα, βρείτε την πιθανότητα να έχουμε ακριβώς 3 μπαστούνια, 4 καρό, 4 κούπες, και 2 σπαθιά.

Σε αυτό το πρόβλημα $r = 52, n = 13$. Η κλάση 1 είναι τα μπαστούνια, η κλάση 2 τα καρό, η κλάση 3 οι κούπες, και η κλάση 4 τα σπαθιά. Άρα, $m = 4, k_1 = 3, k_2 =$

$4, k_3 = 4$, και $k_4 = 2$, οπότε η πιθανότητα που ζητάμε είναι

$$\frac{\binom{13}{3}\binom{13}{4}\binom{13}{4}\binom{13}{2}}{\binom{52}{13}}.$$

Παράδειγμα 16. *Το πρόβλημα της επιτροπής.* Στο πρόβλημα της επιτροπής που συζητήσαμε πιό πάνω, βρείτε την πιθανότητα η επιτροπή των 6 να αποτελείται από 2 καθηγητές πρώτης βαθμίδας, 3 αναπληρωτές καθηγητές, και 1 επίκουρο καθηγητή.

Χρησιμοποιώντας την ίδια μέθοδο, βλέπουμε ότι η απάντηση είναι

$$\frac{\binom{25}{2}\binom{15}{3}\binom{35}{1}}{\binom{75}{6}}.$$

2.5 Ένωση ενδεχομένων*

Ας δούμε πάλι την τυχαία μετάθεση n διακεκριμένων αντικειμένων. Λέμε ότι έχουμε σύμπτωση στην i-στή θέση αν το i-στό αντικείμενο βρίσκεται στην i-στή θέση. Έστω A_i το ενδεχόμενο να συμβεί σύμπτωση στην i-στή θέση. Τότε το $A = \bigcup_{i=1}^n A_i$ είναι το ενδεχόμενο να συμβεί τουλάχιστον μία σύμπτωση. Μπορούμε να υπολογίσουμε την $P(\bigcup_{i=1}^n A_i)$ για $n = 2$ από την Εξίσωση (10) του Κεφαλαίου 1, που μάς λέει ότι

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

Χρησιμοποιώντας αυτόν τον τύπο μπορούμε να αποδείξουμε ανάλογο τύπο για $n = 3$. Έστω A_1, A_2 και A_3 τρία ενδεχόμενα και ας θέσουμε $B = A_1 \cup A_2$. Τότε

$$P(A_1 \cup A_2 \cup A_3) = P(B \cup A_3) = P(B) + P(A_3) - P(B \cap A_3).$$

Τώρα,

$$P(B) = P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2). \tag{4}$$

Αφού $B \cap A_3 = (A_1 \cup A_2) \cap A_3 = (A_1 \cap A_3) \cup (A_2 \cap A_3)$, έπεται ότι

$$P(B \cap A_3) = P(A_1 \cap A_3) + P(A_2 \cap A_3) - P(A_1 \cap A_2 \cap A_3). \tag{5}$$

Αντικαθιστώντας τις (4) και (5) στην παράσταση για την $P(A_1 \cup A_2 \cup A_3)$, βλέπουμε ότι

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) &= [P(A_1) + P(A_2) - P(A_1 \cap A_2)] + P(A_3) \\ &\quad -[P(A_1 \cap A_3) + P(A_2 \cap A_3) - P(A_1 \cap A_2 \cap A_3)] \\ &= [P(A_1) + P(A_2) + P(A_3)] \\ &\quad -[P(A_1 \cap A_2) + P(A_1 \cap A_3) + P(A_2 \cap A_3)] \\ &\quad +P(A_1 \cap A_2 \cap A_3). \end{aligned}$$

Προκειμένου να δώσουμε μία πιό εύχρηστη μορφή στον παραπάνω τύπο, θέτουμε

$$\begin{aligned} S_1 &= P(A_1) + P(A_2) + P(A_3), \\ S_2 &= P(A_1 \cap A_2) + P(A_1 \cap A_3) + P(A_2 \cap A_3), \end{aligned}$$

και

$$S_3 = P(A_1 \cap A_2 \cap A_3).$$

Τότε

$$P(A_1 \cup A_2 \cup A_3) = S_1 - S_2 + S_3. \tag{6}$$

Υπάρχει μία γενίκευση της (6) που ισχύει για όλους τους φυσικούς n. Θεωρούμε τα ενδεχόμενα $A_1, A_2, \ldots, A_n$. Ορίζουμε n αριθμούς $S_r, 1 \leq r \leq n$, θέτοντας

$$S_r = \sum_{1 \leq i_1 < \ldots < i_r \leq n} P(A_{i_1} \cap \ldots \cap A_{i_r}).$$

Ειδικότερα,

$$\begin{aligned} S_1 &= P(A_1) + P(A_2) + \ldots + P(A_n), \\ S_2 &= \sum_{i=1}^{n} \sum_{j=i+1}^{n} P(A_i \cap A_j), \end{aligned}$$

και

$$S_n = P(A_1 \cap \ldots \cap A_n).$$

Ο τύπος που δίνει την $P(\bigcup_{i=1}^n A_i)$ είναι ο εξής:

$$\begin{aligned} P\left(\bigcup_{i=1}^{n} A_i\right) &= \sum_{i=1}^{n} (-1)^{r-1} S_r \\ &= S_1 - S_2 + \ldots + (-1)^{n-1} S_n. \end{aligned} \tag{7}$$

Ο αναγνώστης μπορεί εύκολα να ελέγξει ότι αυτός ο τύπος συμφωνεί με την (6) αν $n = 3$ και με την Εξίσωση (10) του Κεφαλαίου 1 αν $n = 2$. Η απόδειξη της (7) γίνεται με επαγωγή, κατά τα άλλα όμως είναι όμοια με την απόδειξη της (6). Παραλείπουμε τις λεπτομέρειες.

Το άθροισμα S_1 έχει n όρους, το άθροισμα S_2 έχει $\binom{n}{2}$ όρους, και γενικά το άθροισμα S_r έχει $\binom{n}{r}$ όρους. Για να το δούμε, παρατηρούμε ότι το r-στό άθροισμα είναι απλώς το άθροισμα των αριθμών $P(A_{i_1} \cap \ldots \cap A_{i_r})$ πάνω από όλες τις τιμές των δεικτών $i_1, i_2, \ldots, i_r$ που ικανοποιούν την $i_1 < i_2 < \ldots < i_r$. Οι δείκτες παίρνουν τιμές από 1 ως n. Επομένως το πλήθος των διαφορετικών τιμών που μπορούν να πάρουν αυτοί οι δείκτες είναι το ίδιο με το πλήθος των τρόπων με τους οποίους μπορούμε να επιλέξουμε r διαφορετικούς αριθμούς από n αριθμούς χωρίς επανατοποθέτηση και χωρίς να μάς ενδιαφέρει η διάταξη.

2.6 Προβλήματα συμπτώσεων*

Μπορούμε τώρα εύκολα να λύσουμε το πρόβλημα του αριθμού των συμπτώσεων. Έστω A_i το ενδεχόμενο να συμβεί σύμπτωση στην i-στή θέση και έστω p_n η πιθανότητα να μην υπάρχουν συμπτώσεις. Για να υπολογίσουμε την $1-p_n = P\left(\bigcup_{i=1}^n A_i\right)$, χρειάζεται να υπολογίσουμε την $P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_r})$ όπου $i_1, i_2, \ldots, i_r$ είναι r διακεκριμένοι αριθμοί από το $\{1, 2, \ldots, n\}$. Αλλά αυτή η πιθανότητα είναι απλώς η πιθανότητα σύμπτωσης στις θέσεις $i_1, i_2, \ldots, i_r$, και έχουμε ήδη βρεί ότι η πιθανότητα να συμβαίνει κάτι τέτοιο είναι $(n-r)!/n!$. Αφού το r-στό άθροισμα S_r έχει ακριβώς $\binom{n}{r}$ όρους, βλέπουμε ότι

$$\begin{aligned}
P(A_1 \cup \ldots \cup A_n) &= \sum_{r=1}^{n} \binom{n}{r} \frac{(n-r)!}{n!} (-1)^{r-1} \\
&= \sum_{r=1}^{n} (-1)^{r-1} \frac{n!}{r!(n-r)!} \frac{(n-r)!}{n!} \\
&= \sum_{r=1}^{n} \frac{(-1)^{r-1}}{r!},
\end{aligned}$$

δηλαδή,

$$(1-p_n) = 1 - \frac{1}{2!} + \frac{1}{3!} - \ldots + \frac{(-1)^{n-1}}{n!}. \tag{8}$$

Χρησιμοποιώντας την (8) βλέπουμε ότι η πιθανότητα p_n να μην υπάρχουν συμπτώσεις είναι

$$p_n = 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \ldots + \frac{(-1)^n}{n!} = \sum_{k=0}^{n} \frac{(-1)^k}{k!}. \tag{9}$$

Το δεξιό μέλος της (9) είναι οι πρώτοι $n+1$ όροι του αναπτύγματος Taylor του e^{-1}. Επομένως, μπορούμε να προσεγγίσουμε την p_n με τον e^{-1} και να πάρουμε τον $1-e^{-1} = 0.6321\ldots$ σαν προσέγγιση του $(1-p_n)$. Όπως προκύπτει, αυτή η προσέγγιση είναι *εξαιρετικά καλή* ακόμα και για μικρές τιμές του n. Στον πίνακα που ακολουθεί υπολογίζουμε τις τιμές του $(1-p_n)$ για διάφορες τιμές του n.

Έχουμε έτσι το αξιοσημείωτο αποτέλεσμα ότι η πιθανότητα να συμβεί τουλάχιστον μία σύμπτωση κατά την τυχαία μετάθεση n αντικειμένων είναι πρακτικά ανεξάρτητη από το n.

Το πρόβλημα των συμπτώσεων μπορεί να διατυπωθεί με πολλούς διαφορετικούς τρόπους. Ένας από τους πιό διάσημους είναι ο εξής.

Ανακατεύουμε καλά δύο όμοιες τράπουλες και εξετάζουμε τον αριθμό των συμπτώσεων. Ποιά είναι η πιθανότητα να έχουμε τουλάχιστον μία σύμπτωση;

Για να λύσουμε το πρόβλημα αρκεί να παρατηρήσουμε ότι η πρώτη τράπουλα μπορεί να χρησιμοποιηθεί για τον προσδιορισμό των θέσεων (δοχείων). Χωρίς περιορισμό της γενικότητας μπορούμε λοιπόν να υποθέσουμε ότι τα χαρτιά της πρώτης τράπουλας είναι διατεταγμένα με την διάταξη $1, 2, \ldots, n$. Τα χαρτιά της δεύτερης τράπουλας (οι βώλοι) συγκρίνονται με τις θέσεις που προσδιορίστηκαν από την πρώτη τράπουλα. Έχουμε σύμπτωση στην θέση i αν και μόνο αν το i-στό χαρτί που πήραμε από την δεύτερη τράπουλα είναι το χαρτί με αριθμό i.

Τώρα που ξέρουμε πώς να υπολογίζουμε την πιθανότητα p_n να μην υπάρχουν συμπτώσεις, μπορούμε εύκολα να βρούμε την πιθανότητα $\beta_n(r)$ να υπάρχουν ακριβώς r συμπτώσεις. Για να λύσουμε το πρόβλημα υπολογίζουμε πρώτα την πιθανότητα να υπάρχουν ακριβώς r συμπτώσεις και να συμβαίνουν στις πρώτες r θέσεις. Αυτό μπορεί να συμβεί μόνο αν δεν υπάρχουν συμπτώσεις στις υπόλοιπες $(n-r)$ θέσεις. Η πιθανότητα να μην υπάρχουν συμπτώσεις κατά την τυχαία μετάθεση j αντικειμένων είναι p_j. Άρα, $j!p_j$ είναι το πλήθος των τρόπων με τους οποίους μπορούμε να μεταθέσουμε j αντικείμενα έτσι ώστε να μην υπάρχουν συμπτώσεις. (Γιατί;) Αφού υπάρχει μόνο ένας τρόπος να έχουμε r συμπτώσεις στις πρώτες r θέσεις, το πλήθος των τρόπων με τους οποίους μπορούμε να έχουμε ακριβώς r συμπτώσεις στις πρώτες r θέσεις και καμμία σύμπτωση στις υπόλοιπες $(n-r)$ θέσεις είναι $(n-r)!p_{n-r}$. Άρα η ζητούμενη πιθανότητα είναι

$$\alpha_r = \frac{(n-r)!}{n!} p_{n-r}.$$

Η πιθανότητα να υπάρχουν ακριβώς r συμπτώσεις και αυτές να συμβαίνουν σε οποιεσδήποτε προκαθορισμένες r θέσεις είναι πάντα η ίδια, δηλαδή α_r.

Για να λύσουμε το πρόβλημα να υπάρχουν ακριβώς r συμπτώσεις, αυτό που απομένει είναι να καταλάβουμε ότι τα ενδεχόμενα «ακριβώς r συμπτώσεις συμβαίνουν στις θέσεις $i_1, i_2, \ldots, i_r$», είναι ξένα ενδεχόμενα για τις διάφορες επιλογές των $i_1, i_2, \ldots, i_r$. Το πλήθος αυτών των επιλογών είναι $\binom{n}{r}$. Επομένως η ζητούμενη πιθανότητα είναι $\binom{n}{r}\alpha_r$. Άρα, αν $\beta_n(r)$ είναι η πιθανότητα να έχουμε ακριβώς r συμπτώσεις κατά την τυχαία μετάθεση n αντικειμένων, βλέπουμε ότι

$$\begin{aligned} \beta_n(r) &= \frac{n!\alpha_r}{r!(n-r)!} \\ &= \frac{n!}{r!(n-r)!}\frac{(n-r)!p_{n-r}}{n!} \\ &= \frac{p_{n-r}}{r!} \\ &= \frac{1}{r!}\left[1 - 1 + \frac{1}{2!} + \ldots + \frac{(-1)^{n-r}}{(n-r)!}\right]. \end{aligned} \tag{10}$$

Χρησιμοποιώντας το γεγονός ότι η p_{n-r} είναι κατά προσέγγιση ίση με e^{-1} (η οποία είναι πολύ καλή ακόμα κι αν το $n-r$ είναι μετρίου μεγέθους) βλέπουμε ότι

$$\beta_n(r) \approx \frac{e^{-1}}{r!}. \tag{11}$$

Σαν τελευταίο παράδειγμα εφαρμογής αυτών των ιδεών, δίνουμε τον υπολογισμό της πιθανότητας να υπάρχει σύμπτωση στην j-στή θέση, με δεδομένο ότι υπάρχουν ακριβώς r συμπτώσεις.

Για να λύσουμε αυτό το πρόβλημα θέτουμε A_j το ενδεχόμενο να συμβεί σύμπτωση στην j-στή θέση, και B_r το ενδεχόμενο να υπάρχουν ακριβώς r συμπτώσεις. Η πιθανότητα που ζητάμε είναι η $P(A_j|B_r)$. Από την (10), $P(B_r) = p_{n-r}/r!$, άρα χρειαζόμαστε την $P(A_j \cap B_r)$. Αλλά το ενδεχόμενο $A_j \cap B_r$ συμβαίνει αν και μόνο αν υπάρχει σύμπτωση στην j-στή θέση και υπάρχουν ακριβώς $(r-1)$ συμπτώσεις στις υπόλοιπες $(n-1)$ θέσεις. Το πλήθος των τρόπων με τους οποίους μπορούμε να έχουμε ακριβώς $(r-1)$ συμπτώσεις στις υπόλοιπες $(n-1)$ θέσεις είναι $(n-1)!\beta_{n-1}(r-1)$. Άρα,

$$\begin{aligned} P(A_j \cap B_r) &= \frac{(n-1)!\beta_{n-1}(r-1)}{n!} \\ &= \frac{p_{n-r}}{(r-1)!n}. \end{aligned}$$

Επομένως,

$$P(A_j|B_r) = \frac{p_{n-r}}{n(r-1)!}\frac{r!}{p_{n-r}} = \frac{r}{n}.$$

2.7 Προβλήματα κατοχής*

Πολλά προβλήματα των συνδυαστικών πιθανοτήτων είναι ισοδύναμα με το πρόβλημα του να κατανεμηθούν n διακεκριμένοι βώλοι σε r διακεκριμένα δοχεία. Αφού καθένας από τους n βώλους μπορεί να πάει σε οποιοδήποτε από τα r δοχεία, υπάρχουν συνολικά r^n διαφορετικοί τρόποι με τους οποίους μπορούμε να κατανείμουμε τους βώλους στα δοχεία. Αν υποθέσουμε ότι οι βώλοι κατανέμονται τυχαία στα δοχεία, καθένας από τους r^n τρόπους έχει πιθανότητα r^{-n}. Ο χώρος πιθανότητας Ω του προβλήματος έχει επομένως r^n ισοπίθανα σημεία. Στα προβλήματα που αφορούν τον τρόπο με τον οποίο κατανέμονται οι βώλοι, επιβάλλουμε διάφορους περιορισμούς για το τι θα περιέχουν τα δοχεία και ζητάμε την πιθανότητα να συμβαίνει η κατάσταση που μάς ενδιαφέρει. Σαν πρώτο παράδειγμα θεωρούμε το εξής πρόβλημα.

Αν n βώλοι κατανεμηθούν τυχαία σε r δοχεία, ποιά είναι η πιθανότητα κανένα δοχείο να μην περιέχει περισσότερους από έναν βώλους;

Για να λύσουμε το πρόβλημα, παρατηρούμε πρώτα ότι η ζητούμενη πιθανότητα είναι 0 αν $n > r$, οπότε υποθέτουμε ότι $n \le r$. Τότε (σκεφτείτε ότι μοιράζουμε τους βώλους έναν-έναν) ο πρώτος βώλος μπορεί να πάει σε οποιοδήποτε από τα r δοχεία, ο δεύτερος σε οποιοδήποτε από τα υπόλοιπα $(r-1)$ δοχεία, κλπ., άρα υπάρχουν συνολικά $(r)_n$ διαφορετικοί τρόποι. Η ζητούμενη πιθανότητα είναι λοιπόν $(r)_n/r^n$.

Η πιθανότητα αυτή είναι ακριβώς η ίδια με την πιθανότητα επιλογής δείγματος μεγέθους n με επανατοποθέτηση από έναν πληθυσμό r αντικειμένων, έτσι ώστε όλα τα στοιχεία στο δείγμα να είναι διαφορετικά. Παρατηρήστε ακόμα ότι r^n είναι το πλήθος των δειγμάτων μεγέθους n από έναν πληθυσμό r διακεκριμένων αντικειμένων. Αυτό δεν είναι συμπτωματικό. Τυχαία δειγματοληψία n αντικειμένων με επανατοποθέτηση είναι τυπικά το ίδιο πράγμα με την τυχαία κατανομή n βώλων σε r δοχεία. Για να το δείτε, σκεφτείτε ότι μοιράζουμε τους βώλους στα δοχεία ως εξής. Πρώτα επιλέγουμε τυχαίο δείγμα μεγέθους n από ένα σύνολο r αντικειμένων, και αν το i-στό στοιχείο στο δείγμα ήταν το j-στό αντικείμενο τοποθετούμε το βώλο i στο δοχείο j. Μερικές φορές είναι χρήσιμο να σκεφτόμαστε έτσι την τυχαία δειγματοληψία με επανατοποθέτηση, δηλ., σαν την τυχαία κατανομή βώλων σε δοχεία (δείτε το πρόβλημα με τα κουπόνια στο τέλος του κεφαλαίου).

Θεωρούμε την τυχαία κατανομή n βώλων σε r δοχεία. Ποιά είναι η πιθανότητα κάποιος συγκεκριμένος βώλος, ας πούμε ο βώλος j, να βρεθεί σε ένα συγκεκριμμένο δοχείο, ας πούμε το δοχείο i; Αν ο βώλος j είναι στο δοχείο i, τότε μάς μένουν $(n-1)$ βώλοι να μοιράσουμε στα r δοχεία χωρίς περιορισμούς για το πού μπορούν να πάνε. Ο βώλος j μπορεί να τοποθετηθεί στο δοχείο i με έναν μόνο τρόπο, και οι υπόλοιποι $(n-1)$ βώλοι μπορούν να τοποθετηθούν στα r δοχεία με r^{n-1} τρόπους. Άρα η ζητούμενη πιθανότητα είναι $r^{n-1}/r^n = 1/r$.

Μεταφράζοντας στην γλώσσα της τυχαίας δειγματοληψίας βλέπουμε ότι σε ένα τυχαίο δείγμα μεγέθους n που επιλέγεται με επανατοποθέτηση από έναν πληθυσμό r αντικειμένων, είναι εξίσου πιθανό το j-στό στοιχείο του δείγματος να είναι οποιοδήποτε από τα r αντικείμενα.

Οι παραπάνω συλλογισμοί επεκτείνονται εύκολα από το ένα προκαθορισμένο δοχείο στα k δοχεία, $1 \leq k \leq r$. Αφήνουμε σαν άσκηση να δείξετε ότι η πιθανότητα k προκαθορισμένοι βώλοι να βρεθούν σε k προκαθορισμένα δοχεία είναι απλώς r^{-k}. Στη γλώσσα της τυχαίας δειγματοληψίας αυτό λέει ότι αν ένα δείγμα μεγέθους n επιλεγεί με επανατοποθέτηση από έναν πληθυσμό r αντικειμένων, τότε η πιθανότητα τα j_1-στό, j_2-στό, ..., j_k-στό στοιχεία στο δείγμα να είναι οποιαδήποτε k προκαθορισμένα αντικείμενα είναι r^{-k}.

Έστω $A_j(i)$ το ενδεχόμενο το j-στό στοιχείο στο δείγμα να είναι το i-στό αντικείμενο. Τότε έχουμε μόλις δεί ότι για κάθε επιλογή $j_1 < j_2 < \ldots < j_k, 1 \leq k \leq n$, στοιχείων του δείγματος (δηλ. βώλων) και κάθε επιλογή $i_1, i_2, \ldots, i_k$ αντικειμένων (δηλ. δοχείων),

$$P(A_{j_1}(i_1) \cap A_{j_2}(i_2) \cap \ldots \cap A_{j_k}(i_k)) = r^{-k}.$$

Αφού $P(A_j(i)) = r^{-1}$ για κάθε j και i, βλέπουμε ότι

$$P(A_{j_1}(i_1) \cap \ldots \cap A_{j_k}(i_k)) = P(A_{j_1}(i_1)) \ldots P(A_{j_k}(i_k)). \tag{12}$$

Αφού αυτό ισχύει για όλα τα k και όλες τις επιλογές των $j_1, \ldots, j_k$, βλέπουμε ότι για κάθε $i_1, i_2, \ldots, i_n$, *τα ενδεχόμενα* $A_{j_1}(i_1), \ldots, A_{j_n}(i_n)$ *είναι ανεξάρτητα.*

Αν σκεφτόμαστε την επιλογή τυχαίου δείγματος μεγέθους n από ένα σύνολο r αντικειμένων σαν μία n-πλή επανάληψη του πειράματος της τυχαίας επιλογής ενός αντικειμένου από αυτό το σύνολο των r αντικειμένων, τότε βλέπουμε πως ο ισχυρισμός ότι τα ενδεχόμενα $A_1(i_1), \ldots, A_n(i_n)$ είναι ανεξάρτητα μάς λέει ότι το αποτέλεσμα ενός πειράματος δεν επηρεάζει το αποτέλεσμα των άλλων πειραμάτων. Αυτό, φυσικά, συμφωνεί απόλυτα με την διαισθητική μας αντίληψη για την τυχαία δειγματοληψία.

Παράδειγμα 17. Υποθέτουμε ότι n βώλοι κατανέμονται τυχαία σε r δοχεία. Βρείτε την πιθανότητα να υπάρχουν ακριβώς k βώλοι στα πρώτα r_1 δοχεία.

Για να λύσουμε το πρόβλημα παρατηρούμε ότι η πιθανότητα δεδομένος βώλος να βρεθεί σε κάποιο από τα πρώτα r_1 δοχεία είναι r_1/r. Σκεφτόμαστε την κατανομή των n βώλων σαν μία n-πλή επανάληψη του πειράματος της τοποθέτησης ενός βώλου σε ένα από τα r δοχεία. Θεωρούμε το πείραμα επιτυχημένο αν ο βώλος τοποθετηθεί σε ένα από τα πρώτα r_1 δοχεία, αλλιώς το ονομάζουμε αποτυχία. Τότε, από τα αποτελέσματα της Παραγράφου 1.5, βλέπουμε ότι η πιθανότητα τα πρώτα r_1 δοχεία να περιέχουν ακριβώς k βώλους είναι

$$\binom{n}{k}\left(\frac{r_1}{r}\right)^k\left(1-\frac{r_1}{r}\right)^{n-k}.$$

2.8 Το πλήθος των άδειων δοχείων*

Επιστρέφουμε στο παράδειγμα όπου n βώλοι κατανέμονται τυχαία σε r δοχεία και ψάχνουμε την πιθανότητα $p_k(r,n)$ ακριβώς k από τα δοχεία να μείνουν άδεια.

Για να προχωρήσουμε στη λύση του προβλήματος, θέτουμε A_i το ενδεχόμενο το i-στό δοχείο να είναι άδειο. Για να συμβαίνει αυτό το ενδεχόμενο, πρέπει όλοι οι n βώλοι να βρίσκονται στα υπόλοιπα $(r-1)$ δοχεία, και αυτό μπορεί να συμβεί με $(r-1)^n$ τρόπους. Αρα $P(A_i)=(r-1)^n/r^n=(1-1/r)^n$.

Όμοια, αν $1\le i_1<i_2<\ldots<i_k\le r$, τότε το ενδεχόμενο $A_{i_1}\cap A_{i_2}\cap\ldots\cap A_{i_k}$ συμβαίνει αν και μόνο αν όλοι οι βώλοι βρίσκονται στα υπόλοιπα $r-k$ δοχεία. Επομένως, $P(A_{i_1}\cap\ldots\cap A_{i_k})=(r-k)^n/r^n=(1-k/r)^n$. Μπορούμε τώρα να εφαρμόσουμε την (7) για να υπολογίσουμε την πιθανότητα του $A_1\cup\ldots\cup A_n$, που είναι το ενδεχόμενο τουλάχιστον ένα από τα δοχεία να μείνει άδειο. Στην κατάσταση που μελετάμε, έχουμε $S_k=\binom{r}{k}(1-k/r)^n$, οπότε χρησιμοποιώντας την (7) βλέπουμε ότι

$$\begin{aligned}
P(A_1\cup\ldots\cup A_r) &= \sum_{k=1}^{r}(-1)^{k-1}S_k \\
&= \sum_{k=1}^{r}(-1)^{k-1}\binom{r}{k}\left(1-\frac{k}{r}\right)^n.
\end{aligned}$$

Έτσι, η πιθανότητα $p_0(r,n)$ να είναι κατειλημμένα όλα τα δοχεία δίνεται από την

$$\begin{aligned}
\text{(13)}\qquad p_0(r,n) &= 1-P(A_1\cup\ldots\cup A_r) \\
&= 1-\sum_{j=1}^{r}(-1)^{j-1}\binom{r}{j}\left(1-\frac{j}{r}\right)^n
\end{aligned}$$

$$= \sum_{j=0}^{r} (-1)^j \binom{r}{j} \left(1 - \frac{j}{r}\right)^n .$$

Σαν επόμενο βήμα, θα υπολογίσουμε την πιθανότητα $\alpha_k(r,n)$ ακριβώς k *εκ των προτέρων καθορισμένα* δοχεία (ας πούμε, τα πρώτα k) να μείνουν άδεια. Το ενδεχόμενο αυτό μπορεί να συμβεί μόνο αν οι n βώλοι βρίσκονται στα υπόλοιπα $r-k$ δοχεία και αν κανένα από αυτά τα $r-k$ δοχεία δεν είναι άδειο. Το πλήθος των τρόπων με τους οποίους μπορούμε να κατανείμουμε n βώλους σε $r-k$ δοχεία έτσι ώστε κανένα να μην μείνει άδειο είναι $(r-k)^n p_0(r-k,n)$. Άρα, η ζητούμενη πιθανότητα είναι

$$(14) \qquad \alpha_k(r,n) = \frac{(r-k)^n p_0(r,n)}{r^n} = \left(1 - \frac{k}{r}\right)^n p_0(r-k,n).$$

Μπορούμε τώρα να υπολογίσουμε εύκολα τις πιθανότητες $p_k(r,n)$. Για κάθε επιλογή k διαφορετικών αριθμών $i_1, i_2, \ldots, i_k$ από το σύνολο αριθμών $\{1, 2, \ldots, n\}$, το ενδεχόμενο «ακριβώς τα δοχεία $i_1, i_2, \ldots, i_k$ είναι άδεια» έχει πιθανότητα $\alpha_k(r,n)$ και όλα αυτά τα ενδεχόμενα είναι ξένα ανά δύο. Υπάρχουν $\binom{r}{k}$ τέτοια ενδεχόμενα και η ένωσή τους είναι το ενδεχόμενο «ακριβώς k δοχεία είναι άδεια». Άρα,

$$(15) \qquad p_k(r,n) = \binom{r}{k} \left(1 - \frac{k}{r}\right)^n p_0(r-k,n).$$

Χρησιμοποιώντας την παράσταση που δίνει την $p_0(r,n)$ στην (13), βλέπουμε ότι

$$(16) \qquad p_k(r,n) = \binom{r}{k} \sum_{j=0}^{r-k} (-1)^j \binom{r-k}{j} \left(1 - \frac{j+k}{r}\right)^n .$$

Όπως τα προβλήματα συμπτώσεων, έτσι και τα προβλήματα κατοχής επιδέχονται διάφορες αναδιατυπώσεις. Παρακάτω αναφέρουμε μία από τις πιό διάσημες.

Το πρόβλημα των κουπονιών. Κουπόνια ή, στις μέρες μας, παιχνίδια τοποθετούνται στα κουτιά με τα δημητριακά (κορνφλέικς) για να δελεάσουν τους νεαρούς καταναλωτές. Ας υποθέσουμε ότι υπάρχουν r διαφορετικά είδη κουπονιών ή παιχνιδιών και ότι μία δοσμένη συσκευασία έχει την ίδια πιθανότητα να περιέχει οποιοδήποτε από αυτά. Αν αγοράσουμε n κουτιά, βρείτε την πιθανότητα να

(α) έχουμε συγκεντρώσει τουλάχιστον ένα από κάθε είδος.

(β) μάς λείπουν ακριβώς k από τα n είδη.

2.9 Ασκήσεις

1 Ο γενετικός κώδικας προσδιορίζει ένα αμινοξύ με βάση μία ακολουθία τριών νουκλεοτιδίων. Κάθε νουκλεοτίδιο μπορεί να ανήκει σε ένα από τα τέσσερα είδη T, A, C

ή G και οι επαναλήψεις επιτρέπονται. Πόσα αμινοξέα κωδικοποιούνται με αυτόν τον τρόπο;

2 Ο κώδικας του Morse χρησιμοποιεί ακολουθίες από τελείες και παύλες στις οποίες επιτρέπονται οι επαναλήψεις.
(α) Πόσα γράμματα μπορούμε να κωδικοποιήσουμε χρησιμοποιώντας ακριβώς n σύμβολα;
(β) Πόσα γράμματα μπορούμε να κωδικοποιήσουμε χρησιμοποιώντας n ή λιγότερα σύμβολα;

3 Ένας άνθρωπος έχει n κλειδιά, από τα οποία ακριβώς ένα ταιριάζει στην κλειδαριά. Δοκιμάζει τα κλειδιά ένα κάθε φορά, επιλέγοντας σε κάθε δοκιμή τυχαία κάποιο από τα κλειδιά που δεν είχε δοκιμάσει νωρίτερα. Βρείτε την πιθανότητα να πετύχει το σωστό κλειδί στην r-στή δοκιμή.

4 Ένα λεωφορείο ξεκινάει με 6 επιβάτες και σταματάει σε 10 διαφορετικές στάσεις. Υποθέτοντας ότι οι επιβάτες είναι εξίσου πιθανό να αποβιβαστούν σε οποιαδήποτε στάση, βρείτε την πιθανότητα να μην αποβιβαστούν δύο επιβάτες στην ίδια στάση.

5 Έστω ότι έχουμε r δοχεία. Τοποθετούμε τυχαία έναν βώλο κάθε φορά στα δοχεία, μέχρι τη στιγμή που για πρώτη φορά κάποιο από τα δοχεία περιέχει δύο βώλους. Βρείτε την πιθανότητα αυτό να συμβεί κατά την n-στή δοκιμή.

6 Ένα δοχείο περιέχει r βώλους που φέρουν τους αριθμούς $1, 2, \ldots, r$. Επιλέγουμε τυχαία N βώλους (όπου $N \leq r$) από το δοχείο, σημειώνουμε τους αριθμούς τους, και επιστρέφουμε τους N βώλους στο δοχείο. Αν αυτή η διαδικασία επαναληφθεί r φορές, ποιά είναι η πιθανότητα κανένας από τους N αρχικούς βώλους να μην επιλεγεί για δεύτερη φορά;

7 Ο Θωμάς και ο Πέτρος ανήκουν σε μία ομάδα n ατόμων που διατάσσονται τυχαία σε μία γραμμή. Ποιά είναι η πιθανότητα να βρίσκονται ανάμεσά τους ακριβώς k άτομα;

8 Ντόμινο είναι ένα ορθογώνιο χωρισμένο σε δύο ίσα υποορθογώνια όπως φαίνεται στο σχήμα.

Κάθε υποορθογώνιο φέρει έναν αριθμό· έστω x και y αυτοί οι (όχι αναγκαστικά διακεκριμένοι) αριθμοί. Αφού το ορθογώνιο είναι συμμετρικό, το ντόμινο (x, y) ταυτίζεται με το (y, x). Πόσα διαφορετικά ορθογώνια ντόμινο μπορούμε να φτιάξουμε χρησιμοποιώντας n διαφορετικούς αριθμούς;

9 Θεωρούμε ένα πρόβλημα σύμπτωσης με n αντικείμενα, και συμβολίζουμε με i και r δύο διαφορετικές συγκεκριμένες θέσεις.
(α) Ποιά είναι η πιθανότητα να έχουμε σύμπτωση στη θέση i αλλά όχι στη θέση r;
(β) Με δεδομένο ότι δεν υπάρχει σύμπτωση στη θέση r, ποιά είναι η πιθανότητα να έχουμε σύμπτωση στη θέση i;

10 Υποθέτουμε ότι n βώλοι κατανέμονται σε n δοχεία.

(α) Ποιά είναι η πιθανότητα ακριβώς ένα δοχείο να μείνει άδειο; *Υπόδειξη:* Χρησιμοποιήστε το αποτέλεσμα του Παραδείγματος 9.
(β) Με δεδομένο ότι το πρώτο δοχείο είναι άδειο, ποιά είναι η πιθανότητα να μείνει άδειο ακριβώς ένα δοχείο;
(γ) Με δεδομένο ότι ακριβώς ένα δοχείο μένει άδειο, ποιά είναι η πιθανότητα να μείνει άδειο το πρώτο δοχείο;

11 Αν κατανείμουμε τυχαία n βώλους σε r δοχεία, ποιά είναι η πιθανότητα το πρώτο δοχείο να περιέχει ακριβώς j βώλους, $0 \leq j \leq n$;

12 Δείξτε ότι
$$\left(1-\frac{n-1}{s}\right)^{n-1} \leq \frac{(s)_n}{s^n} \leq \left(1-\frac{1}{s}\right)^{n-1}.$$

13 Ένα δοχείο περιέχει b μαύρους και r κόκκινους βώλους. Επιλέγουμε έναν βώλο κάθε φορά από το δοχείο, χωρίς επανατοποθέτηση. Βρείτε την πιθανότητα η πρώτη επιλογή μαύρου βώλου να συμβεί κατά την n-στή δοκιμή.

Το επόμενο πρόβλημα αφορά χέρια του πόκερ. Μία τράπουλα έχει 52 χαρτιά. Αυτά τα χαρτιά σχηματίζουν τέσσερις σειρές που λέγονται μπαστούνια, καρό, κούπες, και σπαθιά. Κάθε σειρά έχει 13 χαρτιά με τιμές όψης $2, 3, \ldots, 10, J, Q, K, A$. Ένα χέρι του πόκερ αποτελείται από 5 χαρτιά που επιλέγονται από την τράπουλα χωρίς επανατοποθέτηση και χωρίς να έχει σημασία η διάταξη. Λέμε ότι ένα χέρι του πόκερ είναι κέντα αν είναι της μορφής: $A, 2, 3, 4, 5$ ή $2, 3, 4, 5, 6$ ή $\ldots$ ή $10, J, Q, K, A$.

14 Υπολογίστε την πιθανότητα να εμφανιστεί καθένα από τα εξής χέρια του πόκερ:
(α) Φλός ρουαγιάλ ($(10, J, Q, K, A)$ από την ίδια σειρά)·
(β) Φλός (ακολουθία πέντε χαρτιών της ίδιας σειράς)·
(γ) Καρέ (τιμές όψης της μορφής (x, x, x, x, y) με τα x και y διαφορετικά)·
(δ) Φουλ (τιμές όψης της μορφής (x, x, x, y, y) με τα x και y διαφορετικά)·
(ε) Χρώμα (πέντε χαρτιά της ίδιας σειράς)·
(στ) Κέντα (πέντε συνεχόμενα χαρτιά, χωρίς να μάς ενδιαφέρει η σειρά)·
(ζ) Τριάδα (τιμές όψης της μορφής (x, x, x, y, z) με τα x, y και z διαφορετικά)·
(η) Ζευγάρια (τιμές όψης της μορφής (x, x, y, y, z) με τα x, y και z διαφορετικά)·
(ϑ) Ζευγάρι (τιμές όψης της μορφής (w, w, x, y, z) με τα w, x, y και z διαφορετικά).

15 Ένα δοχείο περιέχει 10 βώλους που φέρουν τους αριθμούς $1, 2, \ldots, 10$. Επιλέγουμε τυχαίο δείγμα μεγέθους 3. Βρείτε την πιθανότητα οι βώλοι 1 και 6 να βρίσκονται ανάμεσα στους τρείς βώλους που επιλέξαμε.

16 Ανοίγουμε ένα-ένα τα χαρτιά μίας συνηθισμένης τράπουλας μέχρι να εμφανιστεί για πρώτη φορά ρήγας. Βρείτε την πιθανότητα αυτό να συμβεί όταν ανοίγουμε το n-στό χαρτί.

17 Έστω ότι επιλέγουμε τυχαίο δείγμα μεγέθους n από έναν πληθυσμό r ατόμων. Βρείτε την πιθανότητα να μην υπάρχει στο δείγμα κανένα από k προκαθορισμένα άτομα, αν η μέθοδος που χρησιμοποιούμε είναι
(α) δειγματοληψία χωρίς επανατοποθέτηση·

(β) δειγματοληψία με επανατοποθέτηση.

18 Υποθέτουμε ότι τυχαίο δείγμα μεγέθους n επιλέγεται χωρίς επανατοποθέτηση από έναν πληθυσμό r ατόμων. Βρείτε την πιθανότητα να βρεθούν στο δείγμα k προκαθορισμένα άτομα.

19 Υποθέτουμε ότι n άτομα μετατίθενται τυχαία. Δείξτε ότι η πιθανότητα k συγκεκριμένα άτομα να καταλάβουν k συγκεκριμένες θέσεις είναι $(n-k)!/n!$.

20 Δείξτε ότι (όπως αναφέρθηκε στο Παράδειγμα 20)

$$\frac{\binom{n}{k}\binom{r-n}{m-k}}{\binom{r}{m}} = \frac{\binom{m}{k}\binom{r-m}{n-k}}{\binom{r}{n}}.$$

21 Ένα δοχείο περιέχει 40 καλές και 10 ελαττωματικές ασφάλειες. Αν επιλέξουμε 10 ασφάλειες, ποιά είναι η πιθανότητα να είναι όλες καλές;

22 Ποιά είναι η πιθανότητα στα χαρτιά του βορρά και του νότου μίας παρτίδας μπρίτζ (26, συνολικά, χαρτιά) να περιέχονται ακριβώς 3 άσοι;

23 Ποιά είναι η πιθανότητα, αν επιλέξουμε 4 χαρτιά από μία τράπουλα, 2 από αυτά να είναι μαύρα και τα άλλα 2 κόκκινα;

24 Βρείτε την πιθανότητα ένα χέρι του πόκερ με 5 χαρτιά να μην περιέχει κανένα χαρτί με αριθμό μικρότερο από 7, δεδομένου ότι περιέχει τουλάχιστον ένα χαρτί μεγαλύτερο από 10, όπου οι άσοι θεωρούνται υψηλά χαρτιά.

25 Αν κρατάτε 3 λαχνούς για μία κλήρωση για την οποία έχουν πουληθεί n λαχνοί και θα δοθούν 5 έπαθλα, ποιά είναι η πιθανότητα να κερδίσετε τουλάχιστον ένα έπαθλο;

26 Ένα δοχείο με 100 βίδες περιέχει 5 ελαττωματικές. Ποιά είναι η πιθανότητα, αν επιλέξουμε δύο βίδες τυχαία και χωρίς επανατοποθέτηση από το δοχείο, να είναι και οι δύο καλές;

27 Δύο δοχεία περιέχουν r βώλους το καθένα, αριθμημένους από 1 ως r. Επιλέγουμε τυχαίο δείγμα μεγέθους $n \leq r$ από κάθε δοχείο, χωρίς επανατοποθέτηση. Βρείτε την πιθανότητα τα δείγματα να περιέχουν ακριβώς k βώλους που φέρουν τους ίδιους αριθμούς.

Κεφάλαιο 3

Διακριτές τυχαίες μεταβλητές

Θεωρούμε ένα πείραμα στο οποίο ρίχνουμε ένα νόμισμα τρείς φορές, με πιθανότητα p να εμφανιστούν γράμματα σε κάθε ρίψη. Υποθέτουμε ότι αν σε μία ρίψη εμφανιστούν γράμματα κερδίζουμε 1 δραχμή, ενώ αν εμφανιστεί κορώνα χάνουμε 1 δραχμή. Προφανώς, η ποσότητα που μας ενδιαφέρει σε αυτήν την περίπτωση είναι το συνολικό μας κέρδος. Συμβολίζουμε αυτήν την ποσότητα με X. Είναι φανερό ότι η X μπορεί να πάρει μόνο μία από τις τιμές $3, 1, -1$, και -3. Δεν μπορούμε να πούμε με βεβαιότητα ποιά από αυτές τις τιμές θα πάρει η X, αφού η τιμή της εξαρτάται από το αποτέλεσμα του τυχαίου μας πειράματος. Αν για παράδειγμα το αποτέλεσμα είναι ΓΓΓ, τότε η X παίρνει την τιμή 3, ενώ για το αποτέλεσμα ΓΚΓ, η X παίρνει την τιμή 1. Στον πίνακα που ακολουθεί καταγράφουμε τις τιμές της X (σε δραχμές) που αντιστοιχούν σε καθένα από τα οκτώ δυνατά αποτελέσματα.

Μπορούμε να σκεφτόμαστε την X σαν μία πραγματική συνάρτηση ορισμένη στον χώρο πιθανότητας που αντιστοιχεί στο πείραμα. Για κάθε $\omega \in \Omega$, ο $X(\omega)$ είναι ένας από τους αριθμούς $3, 1, -1, -3$. Πάρτε για παράδειγμα το ενδεχόμενο $\{\omega : X(\omega) = 1\}$. Αυτό το σύνολο περιέχει τα τρία σημεία ω_2, ω_3 και ω_4 που αντιστοιχούν στα αποτελέσματα ΓΓΚ, ΓΚΓ και ΚΓΓ, αντίστοιχα. Η τελευταία στήλη του πίνακα δίνει τις πιθανότητες που αντιστοιχούν στα οκτώ δυνατά αποτελέσματα του πειράματος. Από τον πίνακα βλέπουμε ότι το ενδεχόμενο $\{\omega : X(\omega) = 1\}$ έχει πιθανότητα $3p^2(1-p)$. Για συντομία λέμε συνήθως ότι το $\{X = 1\}$ έχει πιθανότητα $3p^2(1-p)$. Ανάλογα μπορούμε φυσικά να χειριστούμε και τις άλλες τιμές της X. Βλέπουμε λοιπόν ότι, για κάθε πιθανή τιμή της X, η πιθανότητα με την οποία η X παίρνει αυτήν την τιμή είναι πλήρως καθορισμένη. Όπως θα δούμε στην επόμενη παράγραφο, η ποσότητα X είναι ένα παράδειγμα *διακριτής τυχαίας μεταβλητής*.

3.1 Ορισμοί

Έστω $(\Omega, \mathcal{A}, P)$ ένας χώρος πιθανότητας, και X μία πραγματική συνάρτηση στο Ω που παίρνει πεπερασμένες ή άπειρες αριθμήσιμες το πλήθος τιμές $x_1, x_2, \ldots$. Όπως και στο παράδειγμα που μόλις δώσαμε, θα θέλαμε βέβαια να μπορούμε να μιλάμε για την πιθανότητα με την οποία η X παίρνει την τιμή x_i, για κάθε i. Για το σκοπό αυτό πρέπει να ξέρουμε ότι για κάθε i, το σύνολο $\{\omega \in \Omega : X(\omega) = x_i\}$ είναι ενδεχόμενο, δηλαδή ότι ανήκει στην $\mathcal{A}$. Αν, όπως στο προηγούμενο παράδειγμα, η $\mathcal{A}$ είναι η σ-άλγεβρα όλων των υποσυνόλων του Ω, τότε αυτό ισχύει πάντα. Γιατί σε αυτήν την περίπτωση, όποιο κι αν είναι το x_i, το $\{\omega \in \Omega : X(\omega) = x_i\}$ είναι ένα υποσύνολο του Ω άρα στοιχείο της $\mathcal{A}$, αφού η $\mathcal{A}$ περιέχει κάθε υποσύνολο του Ω. Όπως όμως τονίσαμε στην Παράγραφο 1.2, γενικά η $\mathcal{A}$ δεν περιέχει όλα τα υποσύνολα του Ω, άρα πρέπει να εξασφαλίσουμε ότι το $\{\omega \in \Omega : X(\omega) = x_i\}$ ανήκει στην $\mathcal{A}$. Ο μόνος λογικός τρόπος για να το πετύχουμε είναι να *υποθέσουμε ρητά* ότι η X είναι μία συνάρτηση στο Ω που έχει αυτήν την επιθυμητή ιδιότητα. Οδηγούμαστε έτσι στον ακόλουθο ορισμό.

Ορισμός 1. *Διακριτή πραγματική τυχαία μεταβλητή X σε έναν χώρο πιθανότητας $(\Omega, \mathcal{A}, P)$ είναι μία συνάρτηση X με πεδίο ορισμού το Ω και πεδίο τιμών ένα πεπερασμένο ή άπειρο αριθμήσιμο υποσύνολο $\{x_1, x_2, \ldots\}$ των πραγματικών αριθμών, η οποία έχει την ιδιότητα το $\{\omega : X(\omega) = x_i\}$ να είναι ενδεχόμενο για κάθε i.*

Εξ ορισμού λοιπόν, το $\{\omega : X(\omega) = x_i\}$ είναι ενδεχόμενο, άρα μπορούμε να μιλάμε για την πιθανότητά του. Για συντομία συμβολίζουμε το ενδεχόμενο $\{\omega : X(\omega) = x_i\}$ με $\{X = x_i\}$ και την πιθανότητα $P(\{\omega : X(\omega) = x_i\})$ με $P(X = x_i)$.

Έστω X μία διακριτή πραγματική τυχαία μεταβλητή. Τότε για κάθε πραγματικό αριθμό x, το $\{\omega : X(\omega) = x_i\}$ είναι ενδεχόμενο. Πράγματι, αν $x_1, x_2, \ldots$ είναι οι τιμές που μπορεί να πάρει η X, τότε το $\{\omega : X(\omega) = x_i\}$ είναι ενδεχόμενο από τον ορισμό της διακριτής πραγματικής τυχαίας μεταβλητής. Αν ο x δεν είναι κάποιος από αυτούς τους αριθμούς, τότε $\{\omega : X(\omega) = x\} = \emptyset$, το οποίο είναι επίσης ενδεχόμενο.

Αν οι πιθανές τιμές μίας διακριτής τυχαίας μεταβλητής X αποτελούνται μόνο από ακεραίους ή μη αρνητικούς ακεραίους, τότε λέμε ότι η X είναι μία τυχαία μεταβλητή με ακέραιες ή μη αρνητικές ακέραιες τιμές αντίστοιχα. Οι περισσότερες διακριτές τυχαίες μεταβλητές που εμφανίζονται στις εφαρμογές έχουν μη αρνητικές ακέραιες τιμές.

Ορισμός 2. *Η πραγματική συνάρτηση f που ορίζεται στο $\mathbb{R}$ από την $f(x) = P(X = x)$ λέγεται διακριτή συνάρτηση πυκνότητας της X. Ένας αριθμός x λέγεται πιθανή τιμή της X αν $f(x) > 0$.*

Όταν είναι απαραίτητο, θα συμβολίζουμε τη συνάρτηση πυκνότητας της X με f_X για να δείξουμε ότι είναι η συνάρτηση πυκνότητας της τυχαίας μεταβλητής X.

Παράδειγμα 1. Έστω X η τυχαία μεταβλητή που ορίσαμε στην αρχή αυτού του κεφαλαίου, όταν συζητούσαμε τις τρείς διαδοχικές ρίψεις ενός ζαριού με π.χ. $p = 0.4$. Τότε η X έχει την διακριτή πυκνότητα που ορίζεται από τις

$$f(-3) = 0.216, \quad f(-1) = 0.432, \quad f(1) = 0.288, \quad f(3) = 0.064,$$

και $f(x) = 0$ αν $x \neq -3, -1, 1, 3$. Η πυκνότητα αυτή αναπαρίσταται από το διάγραμμα στο Σχήμα 1.

Παράδειγμα 2. *Διωνυμική κατανομή.* Θεωρούμε n ανεξάρτητες επαναλήψεις του απλού πειράματος επιτυχίας-αποτυχίας που συζητήσαμε στην Παράγραφο 1.5. Με S_n συμβολίζουμε το πλήθος των επιτυχιών στις n δοκιμές. Τότε η S_n είναι μία τυχαία μεταβλητή που παίρνει μόνο τις τιμές $0, 1, 2, \ldots, n$. Στο Κεφάλαιο 1 δείξαμε ότι για κάθε ακέραιο k με $0 \leq k \leq n$,

$$P(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k},$$

άρα η πυκνότητα f της S_n δίνεται από την

$$f(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & , x = 0, 1, 2, \ldots, n, \\ 0 & , \text{αλλιώς}. \end{cases}$$

Αυτή η πυκνότητα, που είναι μία από τις πιό σημαντικές πυκνότητες που εμφανίζονται στη θεωρία πιθανοτήτων, λέγεται *διωνυμική πυκνότητα* με παραμέτρους n

και p. Η πυκνότητα στο Παράδειγμα 1 είναι διωνυμική πυκνότητα με παραμέτρους $n = 3$ και $p = 0.4$.

Για μία τυχαία μεταβλητή X που έχει διωνυμική πυκνότητα λέμε συχνά ότι ακολουθεί τη *διωνυμική κατανομή* (με παραμέτρους n και p αν θέλουμε να είμαστε πιό ακριβείς). Ανάλογη φρασεολογία χρησιμοποιείται και για άλλες τυχαίες μεταβλητές που η πυκνότητά τους έχει κάποιο ειδικό όνομα.

Όπως εξηγήσαμε στο Κεφάλαιο 2, η διωνυμική κατανομή εμφανίζεται κατά την τυχαία δειγματοληψία με επανατοποθέτηση. Για την τυχαία δειγματοληψία χωρίς επανατοποθέτηση έχουμε την ακόλουθη.

Παράδειγμα 3. *Υπεργεωμετρική κατανομή.* Θεωρούμε έναν πληθυσμό r ατόμων, από τα οποία r_1 είναι τύπου I και $r_2 = r - r_1$ είναι τύπου II. Υποθέτουμε ότι ένα τυχαίο δείγμα μεγέθους $n \leq r$ επιλέγεται από τον πληθυσμό. Έστω X το πλήθος των αντικειμένων τύπου I στο δείγμα. Τότε, η X είναι μία τυχαία μεταβλητή με πιθανές τιμές τις $0, 1, 2, \ldots, n$. Από τα αποτελέσματα της Παραγράφου 2.4 ξέρουμε ότι

$$P(X = x) = \frac{\binom{r_1}{x}\binom{r-r_1}{n-x}}{\binom{r}{n}}, \qquad x = 0, 1, 2, \ldots, n.$$

Μπορούμε να γράψουμε

$$\begin{aligned}
\frac{\binom{r_1}{x}\binom{r-r_1}{n-x}}{\binom{r}{n}} &= \frac{(r_1)_x (r-r_1)_{n-x}}{x!(n-x)!} \frac{n!}{(r)_n} \\
&= \binom{n}{x} \frac{(r_1)_x (r-r_1)_{n-x}}{(r)_n}.
\end{aligned}$$

Άρα η πυκνότητα f της X γράφεται με δύο τρόπους:

$$f(x) = \begin{cases} \frac{\binom{r_1}{x}\binom{r-r_1}{n-x}}{\binom{r}{n}} & , x = 0, 1, 2, \ldots, n, \\ 0 & , \text{αλλιώς}, \end{cases}$$

ή

$$f(x) = \begin{cases} \binom{n}{x} \frac{(r_1)_x (r-r_1)_{n-x}}{(r)_n} & , x = 0, 1, 2, \ldots, n, \\ 0 & , \text{αλλιώς}. \end{cases}$$

Η πυκνότητα αυτή λέγεται *υπεργεωμετρική πυκνότητα*.

Συνεχίζουμε με μερικά ακόμα παραδείγματα τυχαίων μεταβλητών.

Παράδειγμα 4. *Σταθερή τυχαία μεταβλητή.* Έστω c ένας πραγματικός αριθμός. Τότε η συνάρτηση X που ορίζεται από την $X(\omega) = c$ για κάθε ω είναι μία διακριτή τυχαία μεταβλητή, γιατί το σύνολο $\{\omega : X(\omega) = c\}$ είναι ολόκληρος ο χώρος Ω, και το Ω είναι ενδεχόμενο. Προφανώς, $P(X = c) = 1$, άρα η πυκνότητα f της X είναι απλώς η $f(c) = 1$ και $f(x) = 0$, $x \neq c$. Μία τέτοια τυχαία μεταβλητή λέγεται σταθερή τυχαία μεταβλητή. Κατ' αυτόν τον τρόπο, μπορούμε να βλέπουμε μία αριθμητική σταθερά σαν τυχαία μεταβλητή.

Παράδειγμα 5. *Δείκτρια τυχαία μεταβλητή.* Έστω A ένα ενδεχόμενο. Θέτουμε $X(\omega) = 1$ αν $\omega \in A$ και $X(\omega) = 0$ αν $\omega \notin A$. Τότε το ενδεχόμενο A συμβαίνει αν και μόνο αν $X = 1$. Η τυχαία αυτή μεταβλητή λέγεται δείκτρια τυχαία μεταβλητή του A γιατί η τιμή της X μάς λέει αν το ενδεχόμενο A συμβαίνει ή όχι. Αντίστροφα, αν X είναι μία τυχαία μεταβλητή στον χώρο πιθανότητας $(\Omega, \mathcal{A}, P)$ που παίρνει τις τιμές 1 ή 0, τότε η X είναι η δείκτρια τυχαία μεταβλητή του ενδεχομένου

$$A = \{\omega : X(\omega) = 1\}.$$

Θέτουμε $p = P(X = 1)$. Τότε, η πυκνότητα f της X δίνεται από τις

$$f(0) = 1 - p, \quad f(1) = p, \quad \text{και} \quad f(x) = 0, \quad x \neq 0 \text{ ή } 1.$$

Παράδειγμα 6. Θεωρούμε το εξής τυχερό παιχνίδι. Έχουμε έναν κυκλικό στόχο ακτίνας 1 που χωρίζεται σε ζώνες με n ομόκεντρους κύκλους ακτίνων $1/n, 2/n, \ldots, n/n = 1$, όπως φαίνεται στο Σχήμα 2 για $n = 5$. Ρίχνουμε ένα βέλος τυχαία στο στόχο, και αν πέσει μέσα στη ζώνη που ορίζεται από τους κύκλους με ακτίνες i/n και $(i+1)/n$, τότε κερδίζουμε ποσό $n - i$, $i = 0, 1, 2, \ldots, n-1$. Συμβολίζουμε με X το ποσό που κερδίζουμε. Βρείτε την πυκνότητα της X.

Σαν χώρο πιθανότητας για το παραπάνω πείραμα επιλέγουμε τον ομοιόμορφο χώρο πιθανότητας στο δίσκο ακτίνας 1. Προφανώς η X είναι μία διακριτή τυχαία μεταβλητή σε αυτόν το χώρο, με πιθανές τιμές τις $1, 2, \ldots, n$. Το ενδεχόμενο $A = \{X = n - i\}$ συμβαίνει αν και μόνο αν το βέλος πέσει στο χωρίο που ορίζεται από τους κύκλους ακτίνων i/n και $(i+1)/n$. Σύμφωνα με την συζήτηση στην Παράγραφο 1.2, η πιθανότητα του A είναι το εμβαδόν του A διαιρεμένο με το εμβαδόν του μοναδιαίου δίσκου. Έτσι, για κάθε $i = 0, 1, 2, \ldots, n-1$

$$\begin{aligned} P(X = n - i) &= P(A) \\ &= \frac{\pi\left[\left(\frac{i+1}{n}\right)^2 - \left(\frac{i}{n}\right)^2\right]}{\pi} = \frac{2i+1}{n^2}. \end{aligned}$$

Θέτοντας $n - i = x$ βλέπουμε ότι η πυκνότητα της X είναι η

$$f(x) = \begin{cases} \frac{2(n-x)+1}{n^2} & , x = 1, 2, \ldots, n, \\ 0 & , \text{αλλιώς.} \end{cases}$$

Η πυκνότητα f μίας διακριτής τυχαίας μεταβλητής X έχει τις ακόλουθες τρείς σημαντικές ιδιότητες:

(i) $f(x) \geq 0, \ x \in \mathbb{R}$.

(ii) *Το* $\{x : f(x) \neq 0\}$ *είναι ένα πεπερασμένο ή άπειρο αριθμήσιμο υποσύνολο του* $\mathbb{R}$. *Έστω* $\{x_1, x_2, \ldots\}$ *αυτό το σύνολο. Τότε,*

(iii) $\sum_i f(x_i) = 1$.

Οι ιδιότητες (i) και (ii) έπονται άμεσα από τον ορισμό της διακριτής συνάρτησης πυκνότητας της X. Για να δούμε ότι ισχύει η (iii), παρατηρούμε ότι τα ενδεχόμενα $\{\omega : X(\omega) = x_i\}$ είναι ξένα και η ένωσή τους είναι το Ω. Άρα

$$\begin{aligned}\sum_i f(x_i) &= \sum_i P(X = x_i) \\ &= P\left(\bigcup_i \{X = x_i\}\right) = P(\Omega) = 1.\end{aligned}$$

Ορισμός 3. *Μία πραγματική συνάρτηση f ορισμένη στο $\mathbb{R}$ λέγεται διακριτή συνάρτηση πυκνότητας αν ικανοποιεί τις ιδιότητες (i), (ii), και (iii) που διατυπώσαμε πιό πάνω.*

Είναι εύκολο να δούμε ότι κάθε διακριτή συνάρτηση πυκνότητας f είναι η συνάρτηση πυκνότητας κάποιας τυχαίας μεταβλητής X. Με άλλα λόγια, αν μας δώσουν την f μπορούμε να κατασκευάσουμε ένα χώρο πιθανότητας $(\Omega, \mathcal{A}, P)$ και μία τυχαία μεταβλητή X ορισμένη στο Ω της οποίας η διακριτή πυκνότητα είναι η f. Πράγματι, έστω ότι δίνεται η f και ας υποθέσουμε ότι $\{x_1, x_2, \ldots\}$ είναι το σύνολο των τιμών για τις οποίες $f(x) \neq 0$. Ορίζουμε $\Omega = \{x_1, x_2, \ldots\}$, $\mathcal{A}$ το δυναμοσύνολο του Ω, και P το μέτρο πιθανότητας που ορίζεται στο $\mathcal{A}$ από την $P(\{\omega\}) = f(x_i)$ αν $\omega = x_i$. Η τυχαία μεταβλητή X που ορίζεται από την $X(\omega) = x_i$ αν $\omega = x_i$ είναι μία τυχαία μεταβλητή που ικανοποιεί τις απαιτήσεις μας. Για να το δείτε, παρατηρήστε ότι $\{\omega : X(\omega) = x_i\} = \{x_i\}$ άρα

$$P(X = x_i) = P(\{x_i\}) = f(x_i).$$

Το αποτέλεσμα αυτό μας επιτρέπει να χρησιμοποιούμε εκφράσεις όπως «Έστω X μία τυχαία μεταβλητή με διακριτή πυκνότητα την f» χωρίς να διευκρινίζουμε τον χώρο πιθανότητας στον οποίο ορίζεται η X. Για συντομία, στο υπόλοιπο αυτού του κεφαλαίου θα χρησιμοποιούμε τον όρο πυκνότητα αντί του όρου διακριτή πυκνότητα.

Η έννοια της διακριτής τυχαίας μεταβλητής μας δίνει έναν βολικό τρόπο για να περιγράφουμε ένα τυχαίο πείραμα που έχει πεπερασμένα ή άπειρα αριθμήσιμα το πλήθος δυνατά αποτελέσματα. Δεν θα μας απασχολεί η εισαγωγή κατάλληλου χώρου πιθανότητας για το πείραμα. Αντί γι' αυτό, απλώς ορίζουμε μία τυχαία μεταβλητή X που παίρνει τιμές $x_1, x_2, \ldots$ έτσι ώστε να ισχύει $X = x_i$ αν και μόνο αν το πείραμα δίνει το i-στό αποτέλεσμα. Έτσι, για παράδειγμα, όταν τραβάμε ένα χαρτί από μία δεσμίδα με n χαρτιά, θέτουμε $X = i$ αν επιλεγεί το i-στό χαρτί. Τότε $P(X =$

$i) = n^{-1}$, άρα μπορούμε να περιγράψουμε το πείραμα λέγοντας ότι παρατηρούμε μία τυχαία μεταβλητή X που παίρνει τις ακέραιες τιμές $1, 2, \ldots, n$ και έχει συνάρτηση πυκνότητας την συνάρτηση με $f(x) = n^{-1}$ αν $x = 1, 2, \ldots, n$ και $f(x) = 0$ αλλιώς.

Γενικά, κάθε πείραμα που έχει πεπερασμένα ή άπειρα αριθμήσιμα το πλήθος δυνατά αποτελέσματα μπορεί να περιγραφεί σαν η παρατήρηση της τιμής μίας διακριτής τυχαίας μεταβλητής. Για την ακρίβεια, πολλές φορές το πείραμα μάς δίνεται ήδη με αυτήν τη μορφή, και είναι συχνά ευκολότερο να το σκεφτόμαστε σε αυτό το πλαίσιο παρά στο πλαίσιο ενός χώρου πιθανότητας.

Ορίζουμε τώρα δύο ακόμα διακριτές πυκνότητες που είναι πολύ χρήσιμες για την επίλυση κάποιων κλάσεων προβλημάτων που η σημασία τους θα γίνει φανερή αργότερα.

Παράδειγμα 7. *Γεωμετρικές πυκνότητες.* Έστω $0 < p < 1$. Τότε η πραγματική συνάρτηση f που ορίζεται στο $\mathbb{R}$ από την

$$f(x) = \begin{cases} p(1-p)^x & , x = 0, 1, 2, \ldots, \\ 0 & , \text{αλλιώς}, \end{cases}$$

είναι μία διακριτή συνάρτηση πυκνότητας και λέγεται *γεωμετρική πυκνότητα* με παράμετρο p.

Για να δούμε ότι η f είναι πυκνότητα, το μόνο που χρειάζεται να ελέγξουμε είναι ότι ισχύει η συνθήκη (iii), γιατί οι συνθήκες (i) και (ii) ικανοποιούνται προφανώς. Όμως η (iii) προκύπτει από το γεγονός ότι το άθροισμα της γεωμετρικής σειράς $\sum_{x=0}^{\infty}(1-p)^x$ ισούται με p^{-1}.

Παράδειγμα 8. *Αρνητικές διωνυμικές πυκνότητες.* Έστω α ένας θετικός πραγματικός αριθμός, και $0 < p < 1$. Μία πυκνότητα που συνδέεται στενά με την γεωμετρική είναι η *αρνητική διωνυμική πυκνότητα* με παραμέτρους α και p, η οποία ορίζεται από την

$$(1) \qquad f(x) = \begin{cases} p^\alpha \binom{-\alpha}{x}(-1)^x(1-p)^x & , x = 0, 1, 2, \ldots, \\ 0 & , \text{αλλιώς}. \end{cases}$$

Για να δείξουμε ότι είναι πυκνότητα πρέπει να επαληθεύσουμε τις ιδιότητες (i)-(iii). Η ιδιότητα (ii) ισχύει προφανώς. Μπορούμε να δούμε ότι ισχύει η (i) ως εξής. Για κάθε μη αρνητικό ακέραιο x,

$$\begin{aligned} \binom{-\alpha}{x} &= \frac{(-\alpha)_x}{x!} \\ &= \frac{(-\alpha)(-\alpha-1)\ldots(-\alpha-x+1)}{x!} \\ &= \frac{(-1)^x\alpha(\alpha+1)\ldots(\alpha+x-1)}{x!} \\ &= (-1)^x\frac{(\alpha+x-1)_x}{x!} \end{aligned}$$

$$= \quad (-1)^x \binom{\alpha + x - 1}{x}.$$

Άρα,

$$p^\alpha \binom{-\alpha}{x} (-1)^x (1-p)^x = p^\alpha \binom{\alpha + x - 1}{x} (1-p)^x. \tag{2}$$

Το δεξιό μέλος της (2) είναι προφανώς μη αρνητικό, οπότε βλέπουμε ότι η (i) ισχύει. Για να επαληθεύσουμε την (iii), χρησιμοποιούμε το γεγονός ότι η σειρά Taylor της $(1-t)^{-\alpha}$ για $-1 < t < 1$ είναι η

$$(1-t)^{-\alpha} = \sum_{x=0}^{\infty} \binom{-\alpha}{x} (-t)^x. \tag{3}$$

Από την (3) με $t = 1 - p$, βλέπουμε ότι

$$p^{-\alpha} = \sum_{x=0}^{\infty} \binom{-\alpha}{x} (-1)^x (1-p)^x,$$

επομένως $\sum_x f(x) = 1$.

Από την (2) βλέπουμε ότι μπορούμε να γράψουμε την αρνητική διωνυμική πυκνότητα στη μορφή

$$f(x) = \begin{cases} p^\alpha \binom{\alpha+x-1}{x} (1-p)^x & , x = 0, 1, 2, \ldots, \\ 0 & , \text{αλλιώς}. \end{cases} \tag{4}$$

Σε κάποιες περιπτώσεις αυτή η μορφή είναι πιό χρήσιμη από την (1). Παρατηρήστε ότι η γεωμετρική πυκνότητα με παράμετρο p είναι μία αρνητική διωνυμική πυκνότητα με παραμέτρους $\alpha = 1$ και p.

Παράδειγμα 9. *Πυκνότητες Poisson.* Έστω λ ένας θετικός αριθμός. Η πυκνότητα Poisson με παράμετρο λ ορίζεται από την

$$f(x) = \begin{cases} \frac{\lambda^x e^{-\lambda}}{x!} & , x = 0, 1, 2, \ldots, \\ 0 & , \text{αλλιώς}. \end{cases}$$

Είναι φανερό ότι αυτή η συνάρτηση ικανοποιεί τις (i) και (ii) του ορισμού της διακριτής συνάρτησης πυκνότητας. Η ιδιότητα (iii) προκύπτει άμεσα από το ανάπτυγμα Taylor της εκθετικής συνάρτησης

$$e^\lambda = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!}.$$

Η εμπειρία δείχνει ότι πολλά τυχαία φαινόμενα που έχουν σχέση με το μέτρημα ακολουθούν κατά προσέγγιση την κατανομή Poisson. Παραδείγματα τέτοιων φαινομένων είναι το πλήθος των ατόμων μιάς ραδιενεργού ουσίας που αποσυντίθενται στη μονάδα του χρόνου, το πλήθος των κλήσεων που δέχεται ένα τηλεφωνικό κέντρο στη μονάδα του χρόνου, το πλήθος των τυπογραφικών λαθών σε μία σελίδα ενός βιβλίου. Για την πλήρη μελέτη αυτών των μοντέλων απαιτείται η έννοια της ανέλιξης Poisson, την οποία θα συζητήσουμε στο Κεφάλαιο 9.

3.2 Υπολογισμοί με πυκνότητες

Μέχρι τώρα περιορίσαμε την προσοχή μας στον υπολογισμό της $P(X = x)$. Συχνά ενδιαφερόμαστε να υπολογίσουμε την πιθανότητα του ενδεχομένου $\{\omega : X(\omega) \in A\}$, όπου A είναι ένα υποσύνολο του $\mathbb{R}$ που δεν είναι μονοσύνολο.

Έστω A υποσύνολο του $\mathbb{R}$, και X διακριτή τυχαία μεταβλητή με πιθανές τιμές τις $x_1, x_2, \ldots$. Τότε το $\{\omega : X(\omega) \in A\}$ είναι ενδεχόμενο. Για να το δείτε, παρατηρήστε ότι

$$\{\omega \mid X(\omega) \in A\} = \bigcup_{x_i \in A} \{\omega \mid X(\omega) = x_i\}, \tag{5}$$

όπου με $\bigcup_{x_i \in A}$ εννοούμε την ένωση πάνω από όλα τα i για τα οποία $x_i \in A$. Για συντομία συνήθως συμβολίζουμε το ενδεχόμενο $\{\omega : X(\omega) \in A\}$ με $\{X \in A\}$, και την πιθανότητά του με $P(X \in A)$. Αν $-\infty \leq a < b \leq \infty$ και το A είναι ένα διάστημα με άκρα τα a και b, ας πούμε το $A = (a, b]$, τότε συνήθως γράφουμε $P(a < X \leq b)$ αντί του $P(X \in (a, b])$. Ανάλογος συμβολισμός χρησιμοποιείται για τα υπόλοιπα διαστήματα που έχουν αυτά τα άκρα.

Ανάλογες συντμήσεις χρησιμοποιούνται και στο συμβολισμό των δεσμευμένων πιθανοτήτων. Έτσι, για παράδειγμα, αν A και B είναι δύο υποσύνολα του $\mathbb{R}$, γράφουμε $P(X \in A | X \in B)$ για τη δεσμευμένη πιθανότητα του ενδεχομένου $\{X \in A\}$ δοθέντος του ενδεχομένου $\{X \in B\}$.

Έστω f η πυκνότητα της X. Μπορούμε να υπολογίσουμε την $P(X \in A)$ απευθείας από την πυκνότητα f με τη βοήθεια του τύπου

$$P(X \in A) = \sum_{x_i \in A} f(x_i), \tag{6}$$

όπου $\sum_{x_i \in A}$ σημαίνει το άθροισμα πάνω από όλα τα i για τα οποία $x_i \in A$. Αυτός ο τύπος είναι άμεση συνέπεια της (5) αφού τα ενδεχόμενα $\{\omega | X(\omega) = x_i\}$, $i = 1, 2, \ldots$, είναι ξένα. Το δεξιό μέλος της (6) συνήθως γράφεται στη μορφή $\sum_{x \in A} f(x)$. Σύμφωνα με αυτόν το συμβολισμό η (6) γίνεται

$$P(X \in A) = \sum_{x \in A} f(x). \tag{7}$$

Η συνάρτηση $F(t), -\infty < t < \infty$, που ορίζεται από την

$$F(t) = P(X \leq t) = \sum_{x \leq t} f(x), \qquad -\infty < t < \infty,$$

λέγεται *συνάρτηση κατανομής* της τυχαίας μεταβλητής X ή της πυκνότητας f. Από τον ορισμό της συνάρτησης κατανομής προκύπτει άμεσα ότι

$$P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a).$$

Αν X είναι μία τυχαία μεταβλητή με ακέραιες τιμές, τότε

$$F(t) = \sum_{x=-\infty}^{[t]} f(x),$$

όπου $[t]$ είναι το ακέραιο μέρος του t (ο μεγαλύτερος ακέραιος που δεν ξεπερνάει τον t, π.χ., $[2.6]=[2]=2$). Βλέπουμε ότι η F είναι μη φθίνουσα συνάρτηση και ότι, για κάθε ακέραιο x, η F παρουσιάζει άλμα μεγέθους $f(x)$ στο x και είναι σταθερή στο διάστημα $[x, x+1)$. Στο Κεφάλαιο 5 θα αποδείξουμε κι άλλες ιδιότητες των συναρτήσεων κατανομής, δουλεύοντας σε ένα γενικότερο πλαίσιο.

Παράδειγμα 10. Θέτουμε $S = \{1, 2, \ldots, 10\}$ και θεωρούμε την X ομοιόμορφα κατανεμημένη στο S. Τότε $f(x) = 1/10$ για $x = 1, 2, \ldots, 10$ και $f(x) = 0$ αλλιώς. Η συνάρτηση κατανομής της X είναι η $F(t) = 0$ αν $t < 1$, $F(t) = 1$ αν $t > 10$ και

$$F(t) = \sum_{x=1}^{[t]} f(x) = \frac{[t]}{10}, \qquad 1 \leq x \leq 10.$$

Στο Σχήμα 3 βλέπετε τη γραφική παράσταση αυτής της συνάρτησης κατανομής. Μπορούμε να υπολογίσουμε την πιθανότητα $P(3 < X \leq 5)$ γράφοντας είτε

$$P(3 < X \leq 5) = f(4) + f(5) = 2/10$$

ή

$$P(3 < X \leq 5) = F(5) - F(3) = 5/10 - 3/10 = 2/10.$$

Όμοια παίρνουμε την $P(3 \leq X \leq 5)$ γράφοντας είτε

$$P(3 \leq X \leq 5) = f(3) + f(4) + f(5) = 3/10$$

ή

$$P(3 \leq X \leq 5) = P(2 < X \leq 5) = F(5) - F(2) = 5/10 - 2/10 = 3/10.$$

Παράδειγμα 11. Υποθέτουμε ότι η X ακολουθεί γεωμετρική κατανομή με παράμετρο p. Βρείτε τη συνάρτηση κατανομής της X και την $P(X \geq x)$ για κάθε μη αρνητικό ακέραιο x.

Η πυκνότητα της X, σύμφωνα με το Παράδειγμα 7, είναι η

$$f(x) = \begin{cases} p(1-p)^x & , x = 0, 1, 2, \ldots, \\ \\ 0 & , \text{αλλιώς}. \end{cases}$$

Άρα, $F(t) = 0$ αν $t < 0$ και

$$F(t) = \sum_{x=0}^{[t]} p(1-p)^x, \quad t \geq 0.$$

Χρησιμοποιώντας τον τύπο για το άθροισμα πεπερασμένης γεωμετρικής προόδου συμπεραίνουμε ότι

$$F(t) = 1 - (1-p)^{[t]+1}, \quad t \geq 0.$$

Ειδικότερα, αν x είναι ένας μη αρνητικός ακέραιος, τότε $F(x-1) = 1 - (1-p)^x$, επομένως

$$\begin{aligned} P(X \geq x) &= 1 - P(X < x) = 1 - P(X \leq x-1) \\ &= 1 - F(x-1) = (1-p)^x. \end{aligned}$$

Τυχαίες μεταβλητές που ακολουθούν τη γεωμετρική κατανομή εμφανίζονται συχνά στις εφαρμογές. Ας υποθέσουμε ότι έχουμε μία ηλεκτρική συσκευή που δεν φθείρεται ούτε βελτιώνεται με την πάροδο του χρόνου αλλά ενδέχεται να μην λειτουργήσει εξαιτίας σποραδικών τυχαίων γεγονότων που συμβαίνουν ομογενώς στον χρόνο. Έστω ότι παρατηρούμε τη συσκευή κατά τακτά χρονικά διαστήματα, π.χ. μέρες ή ώρες, και έστω X το πλήθος των χρονικών μονάδων έως και τη στιγμή της πρώτης αποτυχίας, με την υπόθεση ότι η συσκευή είναι καινούργια την χρονική στιγμή 0. Η X είναι μία διακριτή τυχαία μεταβλητή με πιθανές τιμές τους ακεραίους $1, 2, 3, \ldots$. Το ενδεχόμενο $\{X = n\}$ συμβαίνει αν και μόνο αν η συσκευή αποτυγχάνει να λειτουργήσει για πρώτη φορά κατά το n-στό χρονικό διάστημα. Η υπόθεσή μας ότι η συσκευή δεν φθείρεται ούτε βελτιώνεται με την πάροδο του χρόνου διατυπώνεται αυστηρά ως εξής. Αν ξέρουμε ότι η συσκευή λειτουργούσε συνεχώς μέχρι και τη χρονική στιγμή n, δηλαδή αν η πρώτη αποτυχία είναι μετά τη χρονική στιγμή n $(X > n)$, τότε η πιθανότητα να μην αποτύχει μέχρι και τη χρονική στιγμή $n + m$, δηλαδή η $P(X > n+m | X > n)$, πρέπει να είναι η ίδια με την πιθανότητα να ξεκινάμε με μία συσκευή που είναι καινούργια τη χρονική στιγμή n και να παρατηρούμε ότι δεν αποτυγχάνει μέχρι και τη χρονική στιγμή $n + m$. Η υπόθεσή μας ότι οι αιτίες αποτυχίας παρουσιάζονται ομογενώς στον χρόνο μπορεί να θεωρηθεί ότι σημαίνει πως η τελευταία πιθανότητα εξαρτάται μόνο από το πλήθος των χρονικών μονάδων που παρεμβάλλονται μεταξύ n και $n + m$, δηλαδή m, αλλά όχι από το n. Άρα, η $P(X > n)$ πρέπει να ικανοποιεί την

$$P(X > n + m | X > n) = P(X > m). \tag{8}$$

Αφού

$$P(X > n+m|X > n) = \frac{P(X > n+m)}{P(X > n)},$$

μπορούμε να ξαναγράψουμε την (8) στην μορφή

$$(9) \qquad P(X > n+m) = P(X > n)P(X > m), \quad n, m = 0, 1, 2, \ldots.$$

Θέτοντας $n = m = 0$ βλέπουμε ότι $P(X > 0) = P(X > 0)^2$, άρα η $P(X > 0)$ ισούται με 1 ή 0. Αν $P(X > 0) = 0$, τότε $P(X = 0) = 1$, κάτι που δεν μπορεί να συμβεί στην περίπτωσή μας γιατί η X παίρνει μόνο θετικές ακέραιες τιμές. Επομένως, $P(X > 0) = 1$.

Θέτουμε $p = P(X = 1)$. Τότε $P(X > 1) = 1 - p$ και από την (9) βλέπουμε ότι

$$P(X > n+1) = (1-p)P(X > n).$$

Συνεχίζοντας με τον ίδιο τρόπο βλέπουμε ότι $P(X > n) = (1-p)^n$. Άρα, για $n = 1, 2, \ldots$,

$$(10) \qquad \begin{aligned} P(X = n) &= P(X > n-1) - P(X > n) \\ &= (1-p)^{n-1} - (1-p)^n = p(1-p)^{n-1}. \end{aligned}$$

Αν $p = 0$ τότε $P(X = n) = 0$ για κάθε $n = 1, 2, \ldots$, άρα $P(X = \infty) = 1$, δηλαδή η συσκευή δεν παθαίνει ποτέ βλάβη. Θα εξαιρέσουμε αυτήν την περίπτωση από την συζήτησή μας. Όμοια, η τιμή $p = 1$ εξαιρείται γιατί τότε $P(X = 1) = 1$, δηλαδή η συσκευή δεν λειτουργεί ποτέ.

Θέτουμε $Y = X - 1$. Τότε η Y παίρνει τις τιμές $0, 1, 2, \ldots$ με πιθανότητες $P(Y = n) = p(1-p)^n$. Βλέπουμε λοιπόν ότι η Y ακολουθεί γεωμετρική κατανομή με παράμετρο p.

Όπως μόλις είδαμε, η τυχαία μεταβλητή $Y = X - 1$ ακολουθεί τη γεωμετρική κατανομή. Το παράδειγμα αυτό είναι τυπικό με την έννοια ότι γεωμετρικά κατανεμημένες τυχαίες μεταβλητές συνήθως περιγράφουν τον χρόνο αναμονής μέχρις ότου συμβεί κάποιο ενδεχόμενο. Θα επανέλθουμε σε αυτό το θέμα με περισότερες λεπτομέρειες αφού μιλήσουμε για ανεξάρτητες δοκιμές στην Παράγραφο 3.4.

3.3 Διακριτά τυχαία διανύσματα

Πολύ συχνά ενδιαφερόμαστε να μελετήσουμε την σχέση ανάμεσα σε δύο ή περισσότερες τυχαίες μεταβλητές. Αν για παράδειγμα κάνουμε τυχαία επιλογή δείγματος μεγέθους n από ένα δοχείο που περιέχει r βώλους αριθμημένους με τους αριθμούς $1, 2, \ldots, r$, μπορεί να ενδιαφερόμαστε να μάθουμε τον μεγαλύτερο αριθμό Y που εμφανίζεται σε βώλο του δείγματος καθώς και τον μικρότερο τέτοιο αριθμό Z.

Έστω $(\Omega, \mathcal{A}, P)$ χώρος πιθανότητας, και $X_1, \ldots, X_r$ διακριτές τυχαίες μεταβλητές ορισμένες σε αυτόν τον χώρο. Τότε για κάθε σημείο $\omega \in \Omega$ καθεμία από τις τυχαίες μεταβλητές $X_1, \ldots, X_r$ παίρνει μία από τις πιθανές τιμές της, γεγονός στο οποίο θα αναφερόμαστε γράφοντας

$$X_1(\omega) = x_1, X_2(\omega) = x_2, \ldots, X_r(\omega) = x_r.$$

Αντί να σκεφτόμαστε ότι παρατηρούμε r πραγματικούς αριθμούς $x_1, x_2, \ldots, x_r$ μπορούμε να σκεφτόμαστε ότι παρατηρούμε την r-άδα $\mathbf{x} = (x_1, x_2, \ldots, x_r)$, όπου για κάθε δείκτη i, ο αριθμός x_i είναι μία από τις πεπερασμένες ή άπειρες αριθμήσιμες το πλήθος τιμές που παίρνει η τυχαία μεταβλητή X_i.

Συμβολίζουμε με $\mathbb{R}^r$ το σύνολο όλων των r-άδων πραγματικών αριθμών. Ένα σημείο $\mathbf{x} = (x_1, \ldots, x_r)$ του $\mathbb{R}^r$ ονομάζεται συνήθως *r-διάστατο διάνυσμα*. Έτσι, για κάθε $\omega \in \Omega$, οι r τιμές $X_1(\omega), \ldots, X_r(\omega)$ ορίζουν ένα σημείο

$$\mathbf{X}(\omega) = (X_1(\omega), X_2(\omega), \ldots, X_r(\omega))$$

στον $\mathbb{R}^r$. Με αυτόν τον τρόπο ορίζεται μία r-διάστατη διανυσματική συνάρτηση $\mathbf{X} : \Omega \to \mathbb{R}^r$ στον Ω, η οποία συνήθως συμβολίζεται με $\mathbf{X} = (X_1, X_2, \ldots, X_r)$. Λέμε ότι η συνάρτηση $\mathbf{X}$ είναι ένα *διακριτό r-διάστατο τυχαίο διάνυσμα*.

Μόλις ορίσαμε την έννοια του r-διάστατου τυχαίου διανύσματος που σχετίζεται με r το πλήθος πραγματικές τυχαίες μεταβλητές. Εναλλακτικά, μπορούμε να ορίσουμε το r-διάστατο τυχαίο διάνυσμα απευθείας σαν μία συνάρτηση $\mathbf{X} : \Omega \to \mathbb{R}^r$ επεκτείνοντας τον ορισμό της πραγματικής τυχαίας μεταβλητής κατά προφανή τρόπο.

Ορισμός 4. *Ένα διακριτό r-διάστατο τυχαίο διάνυσμα $\mathbf{X}$ είναι μία συνάρτηση $\mathbf{X}$ από το Ω στον $\mathbb{R}^r$ που παίρνει πεπερασμένες ή άπειρες αριθμήσιμες το πλήθος τιμές $\mathbf{x}_1, \mathbf{x}_2, \ldots$, τέτοιες ώστε το σύνολο*

$$\{\omega : \mathbf{X}(\omega) = \mathbf{x}_i\}$$

να είναι ενδεχόμενο για κάθε i.

Η διακριτή συνάρτηση πυκνότητας f του τυχαίου διανύσματος $\mathbf{X}$ ορίζεται από την

$$f(x_1, \ldots, x_r) = P(X_1 = x_1, \ldots, X_r = x_r)$$

ή, ισοδύναμα, από την

$$f(\mathbf{x}) = P(\mathbf{X} = \mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^r.$$

Για να βρούμε την πιθανότητα το $\mathbf{X}$ να ανήκει σε κάποιο υποσύνολο του $\mathbb{R}^r$ χρησιμοποιούμε το ανάλογο της (7), δηλαδή,

$$P(\mathbf{X} \in A) = \sum_{\mathbf{x} \in A} f(\mathbf{x}).$$

Όπως και στην μονοδιάστατη περίπτωση, αυτή η συνάρτηση f έχει τις ακόλουθες τρείς ιδιότητες:

(i) $f(\mathbf{x}) \geq 0, \mathbf{x} \in \mathbb{R}^r$.

(ii) *Το σύνολο $\{\mathbf{x} : f(\mathbf{x}) \neq 0\}$ είναι ένα πεπερασμένο ή άπειρο αριθμήσιμο υποσύνολο του $\mathbb{R}^r$, το οποίο θα συμβολίζουμε με $\{\mathbf{x}_1, \mathbf{x}_2, \ldots\}$.*

(iii) $\sum_i f(\mathbf{x}_i) = 1$.

Κάθε πραγματική συνάρτηση f που ορίζεται στον $\mathbb{R}^r$ και έχει αυτές τις τρείς ιδιότητες, θα λέγεται διακριτή r-διάστατη συνάρτηση πυκνότητας. Το επιχείρημα που δώσαμε στη μονοδιάστατη περίπτωση εφαρμόζεται χωρίς ουσιαστικές αλλαγές και δείχνει ότι κάθε r-διάστατη διακριτή συνάρτηση πυκνότητας είναι η συνάρτηση πυκνότητας κάποιου r-διάστατου τυχαίου διανύσματος.

Τα τυχαία διανύσματα και οι συναρτήσεις πυκνότητάς τους συνοδεύονται από κάποια ποσότητα παραδοσιακής ορολογίας. Έστω $\mathbf{X} = (X_1, X_2, \ldots, X_r)$ ένα r-διάστατο τυχαίο διάνυσμα με πυκνότητα f. Η συνάρτηση f ονομάζεται συνήθως *από κοινού πυκνότητα* των τυχαίων μεταβλητών $X_1, X_2, \ldots, X_r$. Η συνάρτηση πυκνότητας της τυχαίας μεταβλητής X_i είναι η *i-στή περιθώρια πυκνότητα* του $\mathbf{X}$ ή της f.

Έστω X και Y δύο διακριτές τυχαίες μεταβλητές. Για κάθε επιλογή πραγματικών αριθμών x και y, το σύνολο $\{\omega \mid X(\omega) = x \ kai \ Y(\omega) = y\}$ είναι ένα ενδεχόμενο, το οποίο θα συμβολίζουμε συνήθως με $\{X = x, Y = y\}$. Υποθέτουμε ότι οι διακεκριμένες πιθανές τιμές της X είναι οι $x_1, x_2, \ldots$, και ότι οι διακεκριμένες πιθανές τιμές της Y είναι οι $y_1, y_2, \ldots$. Για κάθε x, τα ενδεχόμενα $\{X = x, Y = y\}$, $j = 1, 2, \ldots$, είναι ξένα και η ένωσή τους είναι το ενδεχόμενο $\{X = x\}$. Άρα,

$$\begin{aligned} P(X = x) &= P\left(\bigcup_j \{X = x, Y = y_j\}\right) \\ &= \sum_j P(X = x, Y = y_j) = \sum_y P(X = x, Y = y). \end{aligned}$$

Η τελευταία παράσταση προκύπτει αν ακολουθήσουμε το συμβολισμό που εγκαινιάσαμε για τις τυχαίες μεταβλητές στην Παράγραφο 3.2. Όμοια,

$$\begin{aligned} P(Y = y) &= P\left(\bigcup_i \{X = x_i, Y = y\}\right) \\ &= \sum_i P(X = x_i, Y = y) = \sum_x P(X = x, Y = y). \end{aligned}$$

Με άλλα λόγια, αν ξέρουμε την από κοινού πυκνότητα δύο διακριτών τυχαίων μεταβλητών X και Y τότε μπορούμε να υπολογίσουμε την πυκνότητα f_X της X αθροίζοντας ως προς y και την πυκνότητα f_Y της Y αθροίζοντας ως προς x. Δηλαδή, με βάση τις πυκνότητες, αν f είναι η από κοινού πυκνότητα των X και Y, τότε

$$f_X(x) = \sum_y f(x, y) \tag{11}$$

και

$$f_Y(y) = \sum_x f(x, y). \tag{12}$$

Παράδειγμα 12. Ας υποθέσουμε ότι δύο χαρτιά επιλέγονται τυχαία χωρίς επανατοποθέτηση από μία δεσμίδα τριών χαρτιών που φέρουν τους αριθμούς 1,2,3. Έστω X η ένδειξη στο πρώτο χαρτί και Y η ένδειξη στο δεύτερο. Τότε, η από κοινού πυκνότητα f των X και Y δίνεται από τις $f(1,2) = f(1,3) = f(2,1) = f(2,3) = f(3,1) = f(3,2) = 1/6$ και $f(x,y) = 0$ αλλιώς. Η πρώτη περιθώρια πυκνότητα, δηλαδή η πυκνότητα της X, δίνεται από την

$$f_X(1) = f(1,1) + f(1,2) + f(1,3) = 0 + 1/6 + 1/6 = 2/6 = 1/3$$

και όμοια για $x = 2$ και 3. Επομένως $f_X(x) = 1/3, x = 1,2,3,$ και $f_X(x) = 0$ αλλιώς, όπως ακριβώς θα έπρεπε.

Παράδειγμα 13. Υποθέτουμε ότι X και Y είναι τυχαίες μεταβλητές που παίρνουν τις τιμές x και y, όπου $x = 1$ ή 2 και $y = 1,2,3,4$, με τις πιθανότητες που δίνονται στον παρακάτω πίνακα.

Τότε, $f_X(1) = \sum_{y=1}^{4} f(1,y) = 1/4 + 1/8 + 1/16 + 1/16 = 1/2$, και $f_X(2) = 1 - f_X(1) = 1/2$, δηλαδή η X έχει την ομοιόμορφη κατανομή στο $\{1,2\}$. Όμοια,

$$f_Y(1) = 1/4 + 1/16 = 5/16, \quad f_Y(2) = 3/16, \quad f_Y(3) = 5/16, \quad f_Y(4) = 3/16.$$

3.4 Ανεξάρτητες τυχαίες μεταβλητές

Θεωρούμε ένα πείραμα στο οποίο ρίχνουμε ένα νόμισμα και ένα ζάρι. Διαισθητικά, πιστεύουμε ότι όποιο κι αν είναι το αποτέλεσμα της ρίψης του νομίσματος, δεν θα έπρεπε να έχει καμμία επίδραση στο αποτέλεσμα της ρίψης του ζαριού, και αντίστροφα. Θέλουμε τώρα να κατασκευάσουμε ένα πιθανοθεωρητικό μοντέλο που να συμφωνεί με τη διαίσθησή μας. Έστω X η τυχαία μεταβλητή που ισούται με 1 ή 0 ανάλογα με το αν το νόμισμα δείχνει κορώνα ή γράμματα: δηλαδή, το ενδεχόμενο $\{X = 1\}$ αντιστοιχεί στην περίπτωση που το νόμισμα δείχνει κορώνα, και το ενδεχόμενο $\{X = 0\}$ αντιστοιχεί στην περίπτωση που το νόμισμα δείχνει γράμματα. Με ανάλογο τρόπο, στην ρίψη του ζαριού αντιστοιχεί η τυχαία μεταβλητή Y που παίρνει την τιμή $1, 2, \ldots,$ ή 6 ανάλογα με το αν η άνω έδρα του ζαριού φέρει τον αριθμό $1, 2, \ldots$ ή 6 αντίστοιχα. Το αποτέλεσμα του διπλού πειράματος περιγράφεται λοιπόν από το τυχαίο διάνυσμα (X, Y). Το διαισθητικό μας συμπέρασμα ότι τα αποτελέσματα των ρίψεων του νομίσματος και του ζαριού δεν έχουν καμμία επίδραση το ένα στο άλλο, μπορεί να διατυπωθεί αυστηρά αν πούμε ότι αν x είναι ένας από τους αριθμούς 0 ή 1 και y είναι ένας από τους αριθμούς $1, 2, \ldots, 6$, τότε τα ενδεχόμενα $\{X = x\}$ και $\{Y = y\}$ πρέπει να είναι ανεξάρτητα. Δηλαδή, το τυχαίο διάνυσμα (X, Y) πρέπει να έχει σαν από κοινού πυκνότητα την

$$f(x,y) = \begin{cases} P(X = x)P(Y = y) & , x = 0, 1, \ \ y = 1, 2, \ldots, 6, \\ \\ 0 & , \text{αλλιώς}. \end{cases}$$

Με άλλα λόγια, η από κοινού πυκνότητα f των X και Y πρέπει να δίνεται από την

$$f(x, y) = f_X(x) f_Y(y).$$

Ορισμός 5. *Έστω $X_1, X_2, \ldots, X_r$ διακριτές τυχαίες μεταβλητές με πυκνότητες $f_1, f_2, \ldots, f_r$ αντίστοιχα. Λέμε ότι αυτές οι τυχαίες μεταβλητές είναι ανεξάρτητες αν η από κοινού συνάρτηση πυκνότητάς τους f δίνεται από την*

$$f(x_1, x_2, \ldots, x_r) = f_1(x_1) f_2(x_2) \ldots f_r(x_r). \tag{13}$$

Λέμε ότι οι τυχαίες μεταβλητές είναι εξαρτημένες αν δεν είναι ανεξάρτητες. Όπως ακριβώς στην περίπτωση του συνδυασμένου πειράματος της ρίψης ενός νομίσματος και ενός ζαριού, η έννοια της ανεξαρτησίας τυχαίων μεταβλητών μάς προσφέρει έναν εύχρηστο τρόπο για να διατυπώσουμε αυστηρά τη διαισθητική μας αντίληψη ότι κάποια πειράματα είναι ανεξάρτητα το ένα από το άλλο.

Θεωρούμε δύο ανεξάρτητες διακριτές τυχαίες μεταβλητές με πυκνότητες f_X και f_Y αντίστοιχα. Τότε, αν A και B είναι τυχόντα υποσύνολα του $\mathbb{R}$,

$$P(X \in A, Y \in B) = P(X \in A) P(Y \in B). \tag{14}$$

Για να το δούμε, παρατηρούμε ότι

$$\begin{aligned} P(X \in A, Y \in B) &= \sum_{x \in A} \sum_{y \in B} f_{X,Y}(x, y) \\ &= \sum_{x \in A} \sum_{y \in B} f_X(x) f_Y(y) \\ &= \left[\sum_{x \in A} f_X(x) \right] \left[\sum_{y \in B} f_Y(y) \right] \\ &= P(X \in A) P(Y \in B). \end{aligned}$$

Η σχέση (14) επεκτείνεται εύκολα από τις 2 στις r ανεξάρτητες τυχαίες μεταβλητές. Αν $A_1, A_2, \ldots, A_r$ είναι τυχόντα r υποσύνολα του $\mathbb{R}$, τότε

$$P(X_1 \in A_1, \ldots, X_r \in A_r) = P(X_1 \in A_1) \ldots P(X_r \in A_r). \tag{15}$$

Παράδειγμα 14. Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν γεωμετρική κατανομή με παράμετρο p.

(α) Βρείτε την κατανομή της $\min(X, Y)$.

(β) Βρείτε την $P(\min(X, Y) = X) = P(Y \geq X)$.

(γ) Βρείτε την κατανομή της $X + Y$.

(δ) Βρείτε την $P(Y = y \mid X + Y = z)$ για $y = 0, 1, \ldots, z$.

Για το (α) παρατηρούμε ότι για κάθε μη αρνητικό ακέραιο z

$$P(\min(X, Y) \geq z) = P(X \geq z, Y \geq z) = P(X \geq z)P(Y \geq z),$$

άρα, από το Παράδειγμα 11,

$$P(\min(X, Y) \geq z) = (1-p)^z(1-p)^z = (1-p)^{2z}.$$

Από το Παράδειγμα 11 έπεται ότι η $\min(X, Y)$ ακολουθεί γεωμετρική κατανομή με παράμετρο

$$1 - (1-p)^2 = 2p - p^2.$$

Για το (β) παρατηρούμε ότι

$$\begin{aligned}
P(Y \geq X) &= \sum_{x=0}^{\infty} P(X = x, Y \geq X) \\
&= \sum_{x=0}^{\infty} P(X = x, Y \geq x) \\
&= \sum_{x=0}^{\infty} P(X = x)P(Y \geq x) \\
&= \sum_{x=0}^{\infty} p(1-p)^x(1-p)^x \\
&= p\sum_{x=0}^{\infty}(1-p)^{2x} \\
&= p/(1-(1-p)^2) = p/(2p-p^2).
\end{aligned}$$

Για το (γ) θεωρούμε έναν μη αρνητικό ακέραιο z. Τότε

$$\begin{aligned}
P(X + Y = z) &= \sum_{x=0}^{z} P(X = x, X + Y = z) \\
&= \sum_{x=0}^{z} P(X = x, Y = z - x) \\
&= \sum_{x=0}^{z} P(X = x)P(Y = z - x) \\
&= \sum_{x=0}^{z} p(1-p)^x p(1-p)^{z-x} \\
&= (z+1)p^2(1-p)^z.
\end{aligned}$$

Η λύση για το (δ) δίνεται από την

$$\begin{aligned} P(Y=y \mid X+Y=z) &= \frac{P(Y=y, X+Y=z)}{P(X+Y=z)} \\ &= \frac{P(X=z-y, Y=y)}{P(X+Y=z)} \\ &= \frac{P(X=z-y)P(Y=y)}{P(X+Y=z)} \\ &= \frac{p(1-p)^{z-y}p(1-p)^y}{(z+1)p^2(1-p)^z} \\ &= \frac{1}{z+1}. \end{aligned}$$

Θεωρούμε ένα πείραμα (για παράδειγμα, την ρίψη ενός ζαριού) που έχει πεπερασμένα ή άπειρα το πλήθος δυνατά αποτελέσματα. Τότε, όπως έχουμε ήδη εξηγήσει, μπορούμε να σκεφτόμαστε το πείραμα σαν την παρατήρηση της τιμής μιάς διακριτής τυχαίας μεταβλητής X. Ας υποθέσουμε ότι το πείραμα επαναλαμβάνεται n φορές. Το συνδυασμένο πείραμα περιγράφεται σαν η παρατήρηση των τιμών των τυχαίων μεταβλητών $X_1, X_2, \ldots, X_n$, όπου X_i είναι το αποτέλεσμα του i-στού πειράματος. Αν τα πειράματα επαναλαμβάνονται κάτω από τις ίδιες ακριβώς συνθήκες, ο μηχανισμός της τύχης παραμένει υποθετικά ο ίδιος, οπότε πρέπει να υποθέσουμε ότι οι n τυχαίες μεταβλητές έχουν όλες την ίδια πυκνότητα. Η διαίσθηση μάς υπαγορεύει ότι τα επαναλαμβανόμενα πειράματα δεν επηρεάζουν το ένα το άλλο, κάτι που διατυπώνουμε τυπικά απαιτώντας οι $X_1, X_2, \ldots, X_n$ να είναι ανεξάρτητες. Συνοψίζοντας, n ανεξάρτητες τυχαίες μεταβλητές $X_1, \ldots, X_n$ που έχουν την ίδια πυκνότητα f περιγράφουν την n-απλή ανεξάρτητη επανάληψη ενός πειράματος με πεπερασμένα ή άπειρα αριθμήσιμα δυνατά αποτελέσματα.

Τα απλούστερα τυχαία πειράματα είναι αυτά που έχουν δύο μόνο δυνατά αποτελέσματα, τα οποία ονομάζουμε επιτυχία και αποτυχία. Όταν για παράδειγμα ρίχνουμε ένα νόμισμα, μπορούμε να σκεφτόμαστε ότι αν εμφανιστούν γράμματα έχουμε επιτυχία, ή αν τραβάμε ένα χαρτί από μία δέσμη r χαρτιών μπορεί να θεωρούμε επιτυχία την επιλογή άσου. Ας υποθέσουμε ότι επαναλαμβάνουμε ανεξάρτητα το απλό μας πείραμα n φορές. Μπορούμε τότε να περιγράψουμε την κατάσταση θεωρώντας n ανεξάρτητες δείκτριες τυχαίες μεταβλητές $X_1, X_2, \ldots, X_n$ με $X_i = 1$ ή 0 ανάλογα με το αν η i-στή δοκιμή του πειράματος καταλήγει σε επιτυχία ή αποτυχία. Στη βιβλιογραφία, τέτοιες δοκιμές λέγονται *δοκιμές Bernoulli*, και για να περιγράψουμε καταστάσεις όπως η πιό πάνω λέμε ότι εκτελούμε n δοκιμές Bernoulli με την ίδια πιθανότητα επιτυχίας $p = P(X_i = 1)$. Σε αυτό το πλαίσιο, λέμε ότι μία τυχαία μεταβλητή που παίρνει τις τιμές 1 και 0 με πιθανότητες p και $1-p$ αντίστοιχα, έχει *πυκνότητα Bernoulli* με παράμετρο p.

Το αποτέλεσμα n δοκιμών Bernoulli περιγράφεται από το τυχαίο διάνυσμα $\mathbf{X} = (X_1, X_2, \ldots, X_n)$. Η πληροφορία που περιέχει αυτό το διάνυσμα μάς λέει ακριβώς ποιές δοκιμές ήταν επιτυχημένες και ποιές αποτυχημένες. Πολύ συχνά, μία τόσο ακριβής πληροφορία δεν είναι απαραίτητη, και το μόνο που θέλουμε να ξέρουμε είναι το πλήθος S_n των επιτυχημένων δοκιμών ανάμεσα στις n δοκιμές. Στο Παράδειγμα 2 δείξαμε ότι η S_n ακολουθεί διωνυμική κατανομή με παραμέτρους n και p. Παρατηρήστε ότι $S_n = X_1 + \ldots + X_n$. Κάθε τυχαία μεταβλητή Y που ακολουθεί την διωνυμική κατανομή μ¨αυτές τις παραμέτρους μπορεί να θεωρηθεί σαν το άθροισμα n ανεξάρτητων τυχαίων μεταβλητών Bernoulli $X_1, \ldots, X_n$ με παράμετρο p.

Στη συνέχεια θα μελετήσουμε ανεξάρτητες επαναλήψεις ενός πειράματος που έχει $r \geq 2$ δυνατά αποτελέσματα.

3.4.1 Η πολυωνυμική κατανομή

Θεωρούμε ένα πείραμα, π.χ. την ρίψη ενός ζαριού, το οποίο έχει r το πλήθος διαφορετικά δυνατά αποτελέσματα. Μπορούμε να αναπαραστήσουμε αυτό το πείραμα λέγοντας ότι παρατηρούμε μία τυχαία μεταβλητή Y που παίρνει τις τιμές $1, 2, \ldots, r$, δηλαδή το ενδεχόμενο $\{Y = i\}$ αντιστοιχεί στο ενδεχόμενο να προκύψει το i-στό αποτέλεσμα κατά το πείραμα. Θέτουμε $p_i = P(Y = i)$. Αν εκτελέσουμε n ανεξάρτητες επαναλήψεις του πειράματος, μπορούμε να παραστήσουμε το αποτέλεσμα αυτών των n δοκιμών με ένα n-διάστατο τυχαίο διάνυσμα $(Y_1, \ldots, Y_n)$, όπου η τυχαία μεταβλητή Y_j αντιστοιχεί στην j-στή δοκιμή. Οι τυχαίες μεταβλητές $Y_1, \ldots, Y_n$ είναι ανεξάρτητες και $P(Y_j = i) = p_i$.

Το τυχαίο διάνυσμα $(Y_1, \ldots, Y_n)$ μάς δίνει τα αποτελέσματα των n δοκιμών. Όπως και στην περίπτωση των $r = 2$ αποτελεσμάτων, τις περισότερες φορές δεν ενδιαφερόμαστε για μία τόσο λεπτομερή περιγραφή. Θα θέλαμε μόνο να ξέρουμε πόσες από τις n δοκιμές κατέληξαν σε καθένα από τα διάφορα πιθανά αποτελέσματα. Έστω $X_i, i = 1, 2, \ldots, r$, το πλήθος των δοκιμών που έδωσαν το i-στό αποτέλεσμα. Τότε $X_i = x_i$ αν και μόνο αν ακριβώς x_i από τις n τυχαίες μεταβλητές $Y_1, \ldots, Y_n$ παίρνουν την τιμή i, δηλαδή, ακριβώς x_i από τις n δοκιμές δίνουν το i-στό αποτέλεσμα.

Για παράδειγμα, για $r = 3$, $n = 5$, αν

$$Y_1 = 2,\ Y_2 = 3,\ Y_3 = 3,\ Y_4 = 2 \text{ και } Y_5 = 2,$$

τότε

$$X_1 = 0,\ X_2 = 3 \text{ και } X_3 = 2.$$

Θα υπολογίσουμε τώρα την από κοινού πυκνότητα των $X_1, \ldots, X_r$. Για το σκοπό αυτό, θεωρούμε r μη αρνητικούς ακεραίους $x_1, x_2, \ldots, x_r$ με άθροισμα $x_1 + \ldots + x_r = n$. Λίγη σκέψη δείχνει ότι, αφού οι τυχαίες μεταβλητές $Y_1, Y_2, \ldots, Y_n$ είναι ανεξάρτητες και έχουν την ίδια πυκνότητα, κάθε επιλογή για την οποία x_1 από αυτές παίρνουν την τιμή 1, x_2 από αυτές παίρνουν την τιμή 2, …, x_r από αυτές παίρνουν την τιμή r, έχει την ίδια πιθανότητα

$$p_1^{x_1} p_2^{x_2} \ldots p_r^{x_r}.$$

Αν λοιπόν συμβολίσουμε με $C(n; x_1, \ldots, x_r)$ το πλήθος αυτών των επιλογών, βλέπουμε ότι

$$P(X_1 = x_1, \ldots, X_r = x_r) = C(n; x_1, \ldots, x_r) p_1^{x_1} \ldots p_r^{x_r}.$$

Ο υπολογισμός του $C(n; x_1, \ldots, x_r)$ είναι ένα πρόβλημα συνδυαστικής ανάλυσης που λύνεται εύκολα με τις μεθόδους του Κεφαλαίου 2. Ο απλούστερος τρόπος για να το κάνουμε είναι να σκεφτούμε τις r τιμές $1, 2, \ldots, r$ σαν r δοχεία και τις n δοκιμές σαν n βώλους. Τότε, $C(n; x_1, \ldots, x_r)$ είναι το πλήθος των τρόπων με τους οποίους μπορούμε να τοποθετήσουμε τους n βώλους στα r δοχεία με τέτοιο τρόπο ώστε να έχουμε ακριβώς x_1 βώλους στο δοχείο 1, ..., ακριβώς x_r βώλους στο δοχείο r. Αν αυτό συμβαίνει, το δοχείο 1 περιέχει x_1 βώλους. Αυτοί οι x_1 βώλοι μπορούν να επιλεγούν από τους n βώλους με $\binom{n}{x_1}$ τρόπους. Οι υπόλοιποι $n - x_1$ βώλοι πρέπει να τοποθετηθούν στα $r-1$ δοχεία $2, \ldots, r$ με τέτοιο τρόπο ώστε να έχουμε x_2 βώλους στο δοχείο 2, ..., x_r βώλους στο δοχείο r. Άρα,

$$C(n; x_1, \ldots, x_r) = \binom{n}{x_1} C(n - x_1; x_2, \ldots, x_r). \tag{16}$$

Με επαγωγή ως προς r βλέπουμε ότι

$$C(n; x_1, \ldots, x_r) = \frac{n!}{(x_1!)(x_2!)\ldots(x_r!)}. \tag{17}$$

Πράγματι, αν $r = 1$ δεν έχουμε τίποτα να δείξουμε. Υποθέτουμε ότι η (17) ισχύει για $r-1$ δοχεία. Τότε από την (16) βλέπουμε ότι

$$\begin{aligned} C(n; x_1, \ldots, x_r) &= \frac{n!}{(x_1!)((n-x_1)!)} \frac{(n-x_1)!}{(x_2!)\ldots(x_r!)} \\ &= \frac{n!}{(x_1!)(x_2!)\ldots(x_r!)} \end{aligned}$$

όπως θέλαμε.

Η από κοινού πυκνότητα f των $X_1, \ldots, X_r$ δίνεται λοιπόν από την

$$f(x_1, \ldots, x_r) = \frac{n!}{(x_1!)(x_2!)\ldots(x_r!)} p_1^{x_1} \ldots p_r^{x_r} \tag{18}$$

αν x_i είναι μη αρνητικοί ακέραιοι με $x_1 + \ldots + x_r = n$, και 0 αλλιώς. Η πυκνότητα αυτή λέγεται *πολυωνυμική πυκνότητα* με παραμέτρους n και $p_1, \ldots, p_r$.

Παρατηρούμε αμέσως ότι οι r τυχαίες μεταβλητές $X_1, \ldots, X_r$ *δεν* είναι ανεξάρτητες. Για παράδειγμα, αφού $X_1 + \ldots + X_r = n$, οποιεσδήποτε $r-1$ από αυτές προσδιορίζουν την r-στή. Αυτή η παρατήρηση, μαζί με το γεγονός ότι $p_1 + \ldots + p_r = 1$, χρησιμοποιείται μερικές φορές για να εκφράσουμε την πολυωνυμική κατανομή σε κάπως διαφορετική μορφή. Έστω $x_1, x_2, \ldots, x_{r-1}$ μη αρνητικοί ακέραιοι τέτοιοι ώστε $x_1 + \ldots + x_{r-1} \leq n$. Τότε

$$\begin{aligned} P(X_1 = x_1, \ldots, X_{r-1} = x_{r-1}) &= \frac{n!}{(x_1!)\ldots(x_{r-1}!)((n - x_1 - \ldots - x_{r-1})!)} \\ &\quad \times p_1^{x_1} \ldots p_{r-1}^{x_{r-1}} (1 - p_1 - \ldots - p_{r-1})^{n - x_1 - \ldots - x_{r-1}}. \end{aligned} \tag{19}$$

Αυτή η μορφή είναι βολική όταν ενδιαφερόμαστε για τα πρώτα $r-1$ αποτελέσματα και σκεφτόμαστε το r-στό αποτέλεσμα σαν εκείνο το αποτέλεσμα που δεν είναι «κάποιο από τα $r-1$ αποτελέσματα». Ετσι, αν ρίχνουμε ένα ζάρι, μπορεί να μας ενδιαφέρει μόνο αν εμφανίστηκε κάποιος από τους αριθμούς $2, 4$, και 6. Το πείραμα έχει τότε για μας τέσσερα πιθανά αποτελέσματα, τα «2», «4», «6», και «όχι 2,4,6».

Έστω k ένας μη αρνητικός ακέραιος, $k \leq r$. Ενα απλό πιθανοθεωρητικό επιχείρημα δείχνει ότι αν $x_1, x_2, \ldots, x_k$ είναι μη αρνητικοί ακέραιοι με $x_1 + \ldots + x_k \leq n$, τότε

$$P(X_1 = x_1, \ldots, X_k = x_k) = \frac{n!}{(x_1!) \ldots (x_k!)[(n - (x_1 + \ldots + x_k))!]} \times p_1^{x_1} \ldots p_k^{x_k}(1 - (p_1 + \ldots + p_k))^{n-(x_1+\ldots+x_k)}. \tag{20}$$

Για να το δείτε, παρατηρήστε ότι όταν εκτελούμε τις n δοκιμές ενδιαφερόμαστε μόνο για τα $k+1$ αποτελέσματα «1», «2»,...«k», και «όχι $(1, 2, \ldots, k)$». Αυτό λοιπόν που ουσιαστικά έχουμε, είναι n δοκιμές ενός πειράματος με $k+1$ αποτελέσματα, με X_i το πλήθος των δοκιμών στις οποίες εμφανίζεται το i-στό αποτέλεσμα, $i = 1, 2, \ldots, k$. Τότε, η (20) προκύπτει από την (19) αν θέσουμε $r - 1 = k$.

3.4.2 Προσέγγιση της διωνυμικής κατανομής από την Poisson

Υπάρχει πολύ σημαντική σχέση ανάμεσα στην διωνυμική κατανομή και την κατανομή Poisson. Ας υποθέσουμε, για παράδειγμα, ότι εκτελούμε n δοκιμές Bernoulli με πιθανότητα επιτυχίας $p_n = \lambda/n$ σε κάθε δοκιμή. Τότε η πιθανότητα να έχουμε $S_n = k$ επιτυχίες στις n δοκιμές ισούται με

$$\begin{aligned} P(S_n = k) &= \binom{n}{k}(p_n)^k(1 - p_n)^{n-k} \\ &= \frac{\lambda^k}{k!}\frac{(n)_k}{n^k}\left(1 - \frac{\lambda}{n}\right)^n\left(1 - \frac{\lambda}{n}\right)^{-k}. \end{aligned}$$

Τώρα καθώς $n \to \infty$, $(n)_k/n^k \to 1$, $(1 - \lambda/n)^n \to e^{-\lambda}$, και $(1 - \lambda/n)^{-k} \to 1$. Επομένως,

$$\lim_{n\to\infty} \binom{n}{k}(p_n)^k(1 - p_n)^{n-k} = \frac{\lambda^k}{k!}e^{-\lambda}. \tag{21}$$

Στην απόδειξη της (21) χρησιμοποιήσαμε την $np_n = \lambda$. Για την ακρίβεια, η (21) ισχύει αν $np_n \to \lambda$ καθώς $n \to \infty$.

Στις εφαρμογές χρησιμοποιούμε την (21) για να προσεγγίσουμε τη διωνυμική κατανομή με κατανομή Poisson, όταν η πιθανότητα επιτυχίας p είναι μικρή και το n μεγάλο. Στην πράξη, προσεγγίζουμε τη διωνυμική πιθανότητα $P(S_n = x)$ με την $f(x)$ όπου f είναι η πυκνότητα Poisson με $\lambda = np$. Η προσέγγιση είναι πολύ καλή αν το np^2 είναι μικρό. Το παράδειγμα που ακολουθεί περιγράφει τη χρήση αυτής της τεχνικής.

Παράδειγμα 15. Μία μηχανή παράγει βίδες, 1% από τις οποίες είναι ελαττωματικές. Βρείτε την πιθανότητα να μην υπάρχουν ελαττωματικές βίδες σε ένα κουτί των 200 τεμαχίων.

Εδώ έχουμε $n = 200$ δοκιμές Bernoulli, με πιθανότητα επιτυχίας $p = 0.01$. Η πιθανότητα να μην υπάρχουν ελατωματικές βίδες ισούται με

$$(1 - 0.01)^{200} = (0.99)^{200} = 0.1340.$$

Η προσέγγιση που μας δίνει η κατανομή Poisson είναι

$$e^{-200(0.01)} = e^{-2} = 0.1353.$$

Το γεγονός ότι η κατανομή Poisson προκύπτει σαν όριο διωνυμικών κατανομών έχει σημαντικές θεωρητικές συνέπειες. Είναι ένας από τους βασικούς λόγους για τους οποίους αναπτύσσουμε μοντέλα βασισμένα στις ανελίξεις Poisson, τις οποίες θα συζητήσουμε στο Κεφάλαιο 9. Η χρήση της προσέγγισης Poisson για να εξοικονομήσουμε κόπο στον υπολογισμό διωνυμικών πιθανοτήτων είναι δευτερεύουσας σημασίας, αφού οι διωνυμικές πιθανότητες υπολογίζονται επίσης εύκολα.

3.5 Άπειρες ακολουθίες δοκιμών Bernoulli

Έστω ότι εκτελούμε συνεχώς ένα πείραμα επιτυχίας-αποτυχίας που έχει πιθανότητα επιτυχίας p, μέχρις ότου εμφανιστεί η πρώτη επιτυχία. Αν μας δοθεί οσοδήποτε μεγάλο πλήθος δοκιμών n, υπάρχει θετική πιθανότητα $(1-p)^n$ να μην εμφανιστεί καμμία επιτυχία. Επομένως, όταν θεωρούμε το πλήθος των δοκιμών μέχρι την πρώτη επιτυχία, δεν μπορούμε να περιοριστούμε σε συγκεκριμένο πλήθος δοκιμών. Αντίθετα, πρέπει να θεωρήσουμε μία άπειρη ακολουθία δοκιμών.

Ένα δοσμένο πεπερασμένο πλήθος n δοκιμών συνίσταται από n δοκιμές Bernoulli, που αναπαρίστανται από n ανεξάρτητες τυχαίες μεταβλητές Bernoulli $X_1, \ldots, X_n$. Για να παραστήσουμε μία άπειρη ακολουθία δοκιμών Bernoulli θεωρούμε μία άπειρη ακολουθία $\{X_n\}$, $n \geq 1$, ανεξάρτητων τυχαίων μεταβλητών Bernoulli που έχουν την ίδια παράμετρο p.

Γενικά, λέμε ότι οι τυχαίες μεταβλητές $X_1, X_2, \ldots$ είναι ανεξάρτητες αν για κάθε θετικό ακέραιο n, οι τυχαίες μεταβλητές $X_1, \ldots, X_n$ είναι ανεξάρτητες. Αποδεικνύεται ότι, για κάθε δοσμένη πυκνότητα f, υπάρχει ένας χώρος πιθανότητας $(\Omega, \mathcal{A}, P)$ στον οποίο ορίζονται ανεξάρτητες τυχαίες μεταβλητές $X_1, X_2, \ldots$ που έχουν όλες πυκνότητα την f.

Σαν μοντέλο μας για την εκτέλεση άπειρης ακολουθίας δοκιμών Bernoulli, παίρνουμε λοιπόν μία άπειρη ακολουθία $\{X_n\}$, $n \geq 1$, ανεξάρτητων τυχαίων μεταβλητών Bernoulli με $P(X_n = 1) = p$, $n \geq 1$. Γράφοντας $X_n = 1$ εννοούμε ότι το αποτέλεσμα της n-στής δοκιμής είναι επιτυχία, και γράφοντας $X_n = 0$ εννοούμε ότι το αποτέλεσμά της είναι αποτυχία.

Θεωρούμε το πλήθος W_1 των δοκιμών μέχρι την πρώτη επιτυχία. Η τυχαία μεταβλητή W_1 παίρνει μόνο τις ακέραιες τιμές $1, 2, \ldots$. Το ενδεχόμενο $\{W_1 = n\}$ συμβαίνει αν και μόνο αν οι πρώτες $n-1$ δοκιμές είναι αποτυχημένες και η n-στή δοκιμή επιτυχημένη. Επομένως,

$$\{W_1 = n\} = \{X_1 = 0, \ldots, X_{n-1} = 0, X_n = 1\}.$$

Επεται ότι

$$\begin{aligned} P(W_1 = n) &= P(X_1 = 0, \dots, X_{n-1} = 0, X_n = 1) \\ &= P(X_1 = 0) \dots P(X_{n-1} = 0) P(X_n = 1) \\ &= (1-p)^{n-1} p. \end{aligned}$$

Άρα,

$$P(W_1 - 1 = n) = p(1-p)^n. \tag{22}$$

Δηλαδή, η $W_1 - 1$ ακολουθεί γεωμετρική κατανομή με παράμετρο p.

Έστω $r \geq 1$ ένας ακέραιος, και T_r το πλήθος των δοκιμών μέχρι την r-στή επιτυχία (δηλαδή η r-στή επιτυχία συμβαίνει κατά την δοκιμή T_r). Τότε η T_r είναι μία τυχαία μεταβλητή που μπορεί να πάρει μόνο τις ακέραιες τιμές $r, r+1, \dots$. Το ενδεχόμενο $\{T_r = n\}$ συμβαίνει αν και μόνο αν έχουμε επιτυχία στην n-στή δοκιμή και στις πρώτες $n-1$ δοκιμές έχουμε ακριβώς $r-1$ επιτυχίες. Άρα,

$$\{T_r = n\} = \{X_1 + \dots + X_{n-1} = r-1\} \cap \{X_n = 1\}.$$

Αφού τα δύο ενδεχόμενα στο δεξιό μέλος είναι ανεξάρτητα και η $X_1 + \dots + X_{n-1}$ ακολουθεί διωνυμική κατανομή με παραμέτρους $n-1$ και p, βλέπουμε ότι για $n = r, r+1, \dots$

$$\begin{aligned} P(T_r = n) &= P(X_1 + \dots + X_{n-1} = r-1) P(X_n = 1) \\ &= \binom{n-1}{r-1} p^{r-1} (1-p)^{n-r} p \\ &= \binom{n-1}{r-1} p^r (1-p)^{n-r}. \end{aligned}$$

Επομένως

$$P(T_r - r = n) = \binom{r+n-1}{r-1} p^r (1-p)^n. \tag{23}$$

Από τις (4) και (23) βλέπουμε ότι η $T_r - r$ ακολουθεί αρνητική διωνυμική κατανομή με παραμέτρους $\alpha = r$ και p.

Θέτουμε $T_0 = 0$ και για κάθε ακέραιο $r \geq 1$ θεωρούμε την T_r όπως παραπάνω. Ορίζουμε $W_i = T_i - T_{i-1}$, $i = 1, 2, \dots$. Τότε, W_i είναι το πλήθος των δοκιμών μετά την $(i-1)$-στή επιτυχία και μέχρι την i-στή επιτυχία. Θα δείξουμε τώρα ότι για κάθε ακέραιο $r \geq 1$ οι τυχαίες μεταβλητές $W_1 - 1, W_2 - 1, \dots, W_r - 1$ είναι ανεξάρτητες και έχουν την ίδια γεωμετρική πυκνότητα με παράμετρο p.

Για να το δούμε, θεωρούμε r θετικούς ακεραίους $n_1, n_2, \dots, n_r$. Τότε το ενδεχόμενο $\{W_1 = n_1, \dots, W_r = n_r\}$ συμβαίνει αν και μόνο αν οι πρώτες $n_1 + \dots + n_r$ είναι όλες αποτυχημένες, εκτός από τις δοκιμές

$$n_1, n_1 + n_2, \dots, n_1 + \dots + n_r,$$

που είναι επιτυχημένες. Αφού οι δοκιμές είναι ανεξάρτητες και με πιθανότητα επιτυχίας p, βλέπουμε ότι

$$\begin{aligned} P(W_1 = n_1, \ldots, W_r = n_r) &= (1-p)^{n_1-1}p(1-p)^{n_2-1}p \ldots (1-p)^{n_r-1}p \\ &= \prod_{i=1}^{r} [p(1-p)^{n_i-1}]. \end{aligned}$$

Δηλαδή, οι τυχαίες μεταβλητές $W_1 - 1, \ldots, W_r - 1$ είναι ανεξάρτητες, γεωμετρικά κατανεμημένες με παράμετρο p.

Προφανώς $T_r - r = (W_1 - 1) + \ldots + (W_r - 1)$, οπότε βλέπουμε ότι η $T_r - r$ είναι το άθροισμα r ανεξάρτητων, γεωμετρικά κατανεμημένων τυχαίων μεταβλητών. Είδαμε προηγουμένως ότι η $T_r - r$ ακολουθεί αρνητική διωνυμική κατανομή με παραμέτρους r και p. Αποδείξαμε λοιπόν την ενδιαφέρουσα και σημαντική πρόταση ότι *η κατανομή του αθροίσματος r ανεξάρτητων ισοκατανεμημένων γεωμετρικών τυχαίων μεταβλητών με παράμετρο p είναι αρνητική διωνυμική με παραμέτρους r και p.*

Στις ασκήσεις θα συναντήσετε περισσότερες ιδιότητες των άπειρων ακολουθιών δοκιμών Bernoulli.

3.6 Αθροίσματα ανεξάρτητων τυχαίων μεταβλητών

Σε αυτήν την παράγραφο αναπτύσσουμε μεθόδους για την εύρεση της κατανομής του αθροίσματος πεπερασμένων το πλήθος ανεξάρτητων διακριτών τυχαίων μεταβλητών. Ας αρχίσουμε θεωρώντας δύο τέτοιες μεταβλητές X και Y.

Υποθέτουμε λοιπόν ότι X και Y είναι ανεξάρτητες διακριτές τυχαίες μεταβλητές. Έστω $x_1, x_2, \ldots$ οι πιθανές τιμές της X. Για κάθε z, το ενδεχόμενο $\{X + Y = z\}$ ταυτίζεται με το ενδεχόμενο

$$\bigcup_i \{X = x_i, Y = z - x_i\}.$$

Αφού τα ενδεχόμενα $\{X = x_i, Y = z - x_i\}$ είναι ξένα για διαφορετικές τιμές του i, έπεται ότι

$$\begin{aligned} P(X + Y = z) &= \sum_i P(X = x_i, Y = z - x_i) \\ &= \sum_i P(X = x_i)P(Y = z - x_i) \\ &= \sum_i f_X(x_i) f_Y(z - x_i). \end{aligned}$$

Με άλλα λόγια

$$f_{X+Y}(z) = \sum_x f_X(x) f_Y(z - x). \tag{24}$$

Αν οι τυχαίες μεταβλητές X και Y παίρνουν ακέραιες τιμές, τότε η $X+Y$ παίρνει κι αυτή ακέραιες τιμές. Σε αυτήν την περίπτωση, η (24) ισχύει για κάθε ακέραια τιμή του z και η μεταβλητή x στο δεξιό μέλος της (24) διατρέχει το σύνολο των ακεραίων. Εξειδικεύοντας ακόμα περισσότερο παίρνουμε μία πολύ χρήσιμη περίπτωση. Έστω ότι οι X και Y παίρνουν μόνο μη αρνητικές ακέραιες τιμές. Τότε η $X+Y$ παίρνει κι αυτή μόνο μη αρνητικές ακέραιες τιμές. Αν z είναι ένας μη αρνητικός ακέραιος, τότε $f_X(x)f_Y(z-x) = 0$ εκτός αν η x είναι μία από τις τιμές $0, 1, \ldots, z$. Επομένως, με αυτές τις υποθέσεις η (24) γράφεται στη μορφή

$$f_{X+Y}(z) = \sum_{x=0}^{z} f_X(x) f_Y(z-x). \tag{25}$$

Η (25) είναι χρήσιμη για τον υπολογισμό της πυκνότητας της $X+Y$, συνήθως όμως είναι προτιμότερη η χρήση των γεννητριών συναρτήσεων πιθανότητας. Θα περιγράψουμε αυτές τις συναρτήσεις στη συνέχεια, και θα δώσουμε αρκετές σημαντικές εφαρμογές τους στον υπολογισμό της πυκνότητας του αθροίσματος ανεξάρτητων τυχαίων μεταβλητών.

Ορισμός 6. *Έστω X τυχαία μεταβλητή με μη αρνητικές ακέραιες τιμές. Η γεννήτρια συνάρτηση πιθανότητας Φ_X της X ορίζεται από την*

$$\Phi_X(t) = \sum_{x=0}^{\infty} P(X=x)t^x = \sum_{x=0}^{\infty} f_X(x)t^x, \qquad -1 \le t \le 1.$$

Θα υπολογίσουμε την $\Phi_X(t)$ σε τρείς ειδικές περιπτώσεις.

Παράδειγμα 16. *Διωνυμική κατανομή.* Υποθέτουμε ότι η X ακολουθεί διωνυμική κατανομή με παραμέτρους n και p. Τότε

$$P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}$$

άρα

$$\Phi_X(t) = \sum_{x=0}^{n} P(X=x)t^x = \sum_{x=0}^{n} \binom{n}{x} (pt)^x (1-p)^{n-x}.$$

Από τον τύπο του διωνυμικού αναπτύγματος

$$(a+b)^n = \sum_{x=0}^{n} \binom{n}{x} a^x b^{n-x},$$

συμπεραίνουμε ότι

$$\Phi_X(t) = (pt+1-p)^n. \tag{26}$$

Παράδειγμα 17. *Αρνητική διωνυμική κατανομή.* Υποθέτουμε ότι η X ακολουθεί αρνητική διωνυμική κατανομή με παραμέτρους α και p. Τότε

$$P(X=x) = p^\alpha \binom{-\alpha}{x} (-1)^x (1-p)^x$$

άρα

$$\begin{aligned}\Phi_X(t) &= \sum_{x=0}^{\infty} p^{\alpha} \binom{-\alpha}{x} (-1)^x (1-p)^x t^x \\ &= p^{\alpha} \sum_{x=0}^{\infty} \binom{-\alpha}{x} (-t(1-p))^x.\end{aligned}$$

Από το ανάπτυγμα Taylor

$$(1+s)^{-\alpha} = \sum_{x=0}^{\infty} \binom{-\alpha}{x} s^x,$$

θέτοντας $s = -t(1-p)$ παίρνουμε

$$\Phi_X(t) = \left(\frac{p}{1-t(1-p)}\right)^{\alpha}. \tag{27}$$

Παράδειγμα 18. *Κατανομή Poisson.* Υποθέτουμε ότι η X ακολουθεί κατανομή Poisson με παράμετρο λ. Τότε

$$P(X=x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

άρα

$$\Phi_X(t) = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda t)^x}{x!}.$$

Θέτοντας $s = \lambda t$ στο ανάπτυγμα Taylor

$$e^s = \sum_{x=0}^{\infty} \frac{s^x}{x!},$$

βλέπουμε ότι

$$\Phi_X(t) = e^{\lambda t} e^{-\lambda} = e^{\lambda(t-1)}. \tag{28}$$

Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές με μη αρνητικές ακέραιες τιμές. Τότε

$$\Phi_{X+Y}(t) = \Phi_X(t)\Phi_Y(t). \tag{29}$$

Για την απόδειξη, παρατηρούμε ότι από την (25)

$$\Phi_{X+Y}(t) = \sum_{z=0}^{\infty} f_Z(z) t^z$$

$$
\begin{aligned}
&= \sum_{z=0}^{\infty} t^z \sum_{x=0}^{z} f_X(x) f_Y(z-x) \\
&= \sum_{x=0}^{\infty} f_X(x) t^x \sum_{z=x}^{\infty} f_Y(z-x) t^{z-x} \\
&= \sum_{x=0}^{\infty} f_X(x) t^x \sum_{y=0}^{\infty} f_Y(y) t^y \\
&= \Phi_X(t)\Phi_Y(t),
\end{aligned}
$$

δηλαδή το ζητούμενο.

Από την (29) έπεται εύκολα με επαγωγή ότι αν $X_1, \ldots, X_r$ είναι ανεξάρτητες τυχαίες μεταβλητές με μη αρνητικές ακέραιες τιμές, τότε

$$
\Phi_{X_1+\ldots+X_r}(t) = \Phi_{X_1}(t)\ldots\Phi_{X_r}(t). \tag{30}
$$

Οι ισχυρισμοί του θεωρήματος που ακολουθεί αποδεικνύονται πιό εύκολα με την «τεχνική της γεννήτριας συνάρτησης», που βασίζεται στο γεγονός ότι αν

$$
\sum_{x=0}^{\infty} a_x t^x = \sum_{x=0}^{\infty} b_x t^x, \qquad -1 < t < 1,
$$

τότε μπορούμε να εξισώσουμε τους συντελεστές του t^x στις δύο δυναμοσειρές, και να συμπεράνουμε ότι $a_x = b_x, x = 0, 1, 2 \ldots$. Αυτό δείχνει ότι αν δύο τυχαίες μεταβλητές με μη αρνητικές ακέραιες τιμές έχουν την ίδια γεννήτρια συνάρτηση πιθανότητας, τότε πρέπει να έχουν την ίδια κατανομή. Με άλλα λόγια, η γεννήτρια συνάρτηση πιθανότητας μίας τυχαίας μεταβλητής με μη αρνητικές ακέραιες τιμές προσδιορίζει μονοσήμαντα την κατανομή της.

Θεώρημα 1. *Έστω $X_1, \ldots, X_r$ ανεξάρτητες τυχαίες μεταβλητές.*

(i) *Αν η X_i ακολουθεί διωνυμική κατανομή με παραμέτρους n_i και p, τότε η $X_1 + \ldots + X_r$ ακολουθεί διωνυμική κατανομή με παραμέτρους $n_1 + \ldots + n_r$ και p.*

(ii) *Αν η X_i ακολουθεί αρνητική διωνυμική κατανομή με παραμέτρους α_i και p, τότε η $X_1 + \ldots + X_r$ ακολουθεί αρνητική διωνυμική κατανομή με παραμέτρους $\alpha_1 + \ldots + \alpha_r$ και p.*

(iii) *Αν η X_i ακολουθεί κατανομή Poisson με παράμετρο λ_i, τότε η $X_1 + \ldots + X_r$ ακολουθεί κατανομή Poisson με παράμετρο $\lambda_1 + \ldots + \lambda_r$.*

Απόδειξη της (i). Αν οι X_i είναι όπως στην (i), τότε από το Παράδειγμα 16

$$
\begin{aligned}
\Phi_{X_1+\ldots+X_r}(t) &= \Phi_{X_1}(t)\ldots\Phi_{X_r}(t) \\
&= (pt+1-p)^{n_1}\ldots(pt+1-p)^{n_r} \\
&= (pt+1-p)^{n_1+\ldots+n_r}.
\end{aligned}
$$

Άρα η γεννήτρια συνάρτηση πιθανότητας της $X_1 + \ldots + X_r$ συμπίπτει με αυτήν μίας τυχαίας μεταβλητής που ακολουθεί διωνυμική κατανομή με παραμέτρους $n_1 + \ldots + n_r$

και p. Αυτό σημαίνει ότι η $X_1 + \ldots + X_r$ πρέπει να ακολουθεί αυτήν τη διωνυμική κατανομή. Γιατί, αν συμβολίσουμε με

$$a_x = \binom{n_1 + \ldots + n_r}{x} p^x (1-p)^{n_1+\ldots+n_r-x}$$

τις αντίστοιχες διωνυμικές πιθανότητες, τότε

$$\begin{aligned}
\sum_{x=0}^{\infty} P(X_1 + \ldots + X_r = x) t^x &= \Phi_{X_1+\ldots+X_r}(t) \\
&= (pt + 1 - p)^{n_1+\ldots+n_r} \\
&= \sum_{x=0}^{\infty} a_x t^x.
\end{aligned}$$

Εξισώνοντας συντελεστές βλέπουμε ότι

$$P(X_1 + \ldots + X_r = x) = a_x$$

επομένως η $X_1 + \ldots + X_r$ είναι διωνυμικά κατανεμημένη όπως ισχυριστήκαμε στην (i).

Απόδειξη της (ii). Αν οι X_i είναι όπως στην (ii), τότε από το Παράδειγμα 17

$$\begin{aligned}
\Phi_{X_1+\ldots+X_r}(t) &= \Phi_{X_1}(t) \ldots \Phi_{X_r}(t) \\
&= \left(\frac{p}{1-t(1-p)}\right)^{\alpha_1} \ldots \left(\frac{p}{1-t(1-p)}\right)^{\alpha_r} \\
&= \left(\frac{p}{1-t(1-p)}\right)^{\alpha_1+\ldots+\alpha_r}.
\end{aligned}$$

Άρα η γεννήτρια συνάρτηση πιθανότητας της $X_1 + \ldots + X_r$ συμπίπτει με αυτήν μίας τυχαίας μεταβλητής που ακολουθεί αρνητική διωνυμική κατανομή με παραμέτρους $\alpha_1 + \ldots + \alpha_r$ και p. Με το επιχείρημα που χρησιμοποιήσαμε για να αποδείξουμε την (i) βλέπουμε ότι η $X_1 + \ldots + X_r$ ακολουθεί αυτήν την αρνητική διωνυμική κατανομή.

Η απόδειξη της (iii) είναι όμοια με αυτήν των (i) και (ii) και αφήνεται σαν άσκηση για τον αναγνώστη. □

Υποθέτουμε ότι $\alpha_1 = \ldots = \alpha_r = 1$ στον ισχυρισμό (ii) του Θεωρήματος 1. Τότε οι $X_1, \ldots, X_r$ ακολουθούν γεωμετρική κατανομή με παράμετρο p, και η (ii) μάς λέει ότι η $X_1 + \ldots + X_r$ ακολουθεί αρνητική διωνυμική κατανομή με παραμέτρους r και p. Παίρνουμε έτσι μία εναλλακτική απόδειξη του αποτελέσματος της Παραγράφου 3.5.

Το παράδειγμα που ακολουθεί παρουσιάζει την χρήση των δεσμευμένων πιθανοτήτων.

Παράδειγμα 19. Έστω $X_1, X_2, \ldots$ ανεξάρτητες τυχαίες μεταβλητές με μη αρνητικές ακέραιες τιμές, που έχουν την ίδια πυκνότητα. Θέτουμε $S_0 = 0$ και

$S_n = X_1 + \ldots + X_n$, $n \geq 1$. Θεωρούμε μία τυχαία μεταβλητή N με μη αρνητικές ακέραιες τιμές και υποθέτουμε ότι οι $N, X_1, X_2, \ldots$ είναι ανεξάρτητες. Τότε η $S_N = X_1 + \ldots + X_N$ είναι το άθροισμα τυχαίου πλήθους τυχαίων μεταβλητών. [Για την καλύτερη κατανόηση της S_N, υποθέστε ότι την χρονική στιγμή 0 ένα τυχαίο πλήθος N βακτηρίων εισέρχεται σε κάποιο σύστημα, και ότι την χρονική στιγμή 1 η αποικία που εγκαινιάστηκε από το i-στό βακτήριο αριθμεί X_i μέλη.] Δείξτε ότι η γεννήτρια συνάρτηση πιθανότητας της S_N δίνεται από την

$$\Phi_{S_N}(t) = \Phi_N(\Phi_{X_1}(t)), \qquad -1 \leq t \leq 1. \tag{31}$$

Για να επαληθεύσουμε την (31) παρατηρούμε πρώτα ότι

$$\begin{aligned} P(S_N = x) &= \sum_{n=0}^{\infty} P(S_N = x, N = n) \\ &= \sum_{n=0}^{\infty} P(S_n = x, N = n) \\ &= \sum_{n=0}^{\infty} P(N = n) P(S_n = x \mid N = n). \end{aligned}$$

Αφού η N είναι ανεξάρτητη από τις $X_1, \ldots, X_n$, είναι ανεξάρτητη και από την S_n, επομένως $P(S_n = x | N = n) = P(S_n = x)$. Άρα,

$$P(S_n = x) = \sum_{n=0}^{\infty} P(N = n) P(S_n = x). \tag{32}$$

Βλέπουμε λοιπόν ότι για κάθε $-1 \leq t \leq 1$

$$\begin{aligned} \Phi_{S_N}(t) &= \sum_{x=0}^{\infty} t^x P(S_N = x) \\ &= \sum_{x=0}^{\infty} t^x \sum_{n=0}^{\infty} P(N = n) P(S_n = x) \\ &= \sum_{n=0}^{\infty} P(N = n) \sum_{x=0}^{\infty} t^x P(S_n = x) \\ &= \sum_{n=0}^{\infty} P(N = n) \Phi_{S_n}(t) \\ &= \sum_{n=0}^{\infty} P(N = n) \left(\Phi_{X_1}(t)\right)^n = \Phi_N\left(\Phi_{X_1}(t)\right). \end{aligned}$$

3.7 Ασκήσεις

1 Κάθε σημείο του διαστήματος $[0,1)$ αναπαρίσταται από το δεκαδικό του ανάπτυγμα $0.x_1x_2\ldots$. Έστω ότι επιλέγουμε τυχαία ένα σημείο του διαστήματος $[0,1)$. Έστω X το πρώτο ψηφίο στο δεκαδικό ανάπτυγμα που αναπαριστά το σημείο. Υπολογίστε την πυκνότητα της X.

2 Έστω ότι η X έχει αρνητική διωνυμική πυκνότητα με παραμέτρους $\alpha = r$ (όπου r είναι ακέραιος) και p. Υπολογίστε την πυκνότητα της $X + r$.

3 Υποθέτουμε ότι ένα δοχείο περιέχει 6 κόκκινους και 4 μαύρους βώλους. Επιλέγουμε τυχαίο δείγμα μεγέθους n. Με X συμβολίζουμε το πλήθος των κόκκινων βώλων που επιλέχθηκαν. Υπολογίστε την πυκνότητα της X αν η δειγματοληψία γίνεται (α) χωρίς επανατοποθέτηση, (β) με επανατοποθέτηση.

4 Έστω N ένας θετικός ακέραιος και

$$f(x) = \begin{cases} c2^x & , x = 1, 2, \ldots, N, \\ \\ 0 & , \text{αλλιώς}. \end{cases}$$

Βρείτε την τιμή της c ώστε η f να είναι συνάρτηση πυκνότητας.

5 Υποθέτουμε ότι η πυκνότητα f μίας τυχαίας μεταβλητής X δίνεται από τον πίνακα

Υπολογίστε τις εξής πιθανότητες:
(α) Η X είναι αρνητική·
(β) Η X είναι άρτια·
(γ) Η τιμή της X είναι από 1 ως 8·
(δ) $P(X = -3 \mid X \leq 0)$·
(ε) $P(X \geq 3 \mid X > 0)$.

6 Υποθέτουμε ότι η X έχει γεωμετρική κατανομή με $p = 0.8$. Υπολογίστε τις πιθανότητες των εξής ενδεχομένων:
(α) $X > 3$,
(β) $4 \leq X \leq 7$ ή $X > 9$,
(γ) $3 \leq X \leq 5$ ή $7 \leq X \leq 10$.

7 Έστω ότι η X είναι ομοιόμορφα κατανεμημένη στο $\{0, 1, \ldots, 99\}$. Υπολογίστε τις
(α) $P(X \leq 25)$,
(β) $P(2.6 < X < 12.2)$,
(γ) $P(8 < X \leq 10$ ή $30 < X \leq 32)$,
(δ) $P(25 \leq X \leq 30)$.

8 Υποθέτουμε ότι ένα δοχείο περιέχει 12 βώλους με τους αριθμούς $1, 2, \ldots, 12$. Κάνουμε δύο ανεξάρτητες δοκιμές του πειράματος της τυχαίας επιλογής ενός βώλου από το δοχείο. Έστω X η μεγαλύτερη από τις ενδείξεις στους βώλους που επιλέξαμε. Υπολογίστε την πυκνότητα της X.

9 Έστω ότι έχουμε την ίδια κατάσταση όπως στην Άσκηση 8, τώρα όμως επιλέγουμε τους δύο βώλους χωρίς επανατοποθέτηση. Υπολογίστε την πυκνότητα της X.

10 Υποθέτουμε ότι η τυχαία μεταβλητή X ακολουθεί γεωμετρική κατανομή με παράμετρο p. Θέτουμε $Y = X$ αν $X < M$ και $Y = M$ αν $X \geq M$. Δηλαδή, $Y = \min(X, M)$. Υπολογίστε την πυκνότητα της Y.

11 Έστω ότι η X ακολουθεί γεωμετρική κατανομή με παράμετρο p. Υπολογίστε την πυκνότητα της
(α) X^2·
(β) $X + 3$.

12 Υποθέτουμε ότι ένα δοχείο περιέχει r βώλους που φέρουν τους αριθμούς $1, \ldots, r$. Επιλέγουμε τυχαίο δείγμα μεγέθους n χωρίς επανατοποθέτηση. Έστω Y ο μικρότερος και Z ο μεγαλύτερος από τους αριθμούς στους βώλους που επιλέξαμε.
(α) Υπολογίστε την πιθανότητα $P(Y \leq y)$.
(β) Υπολογίστε την πιθανότητα $P(Z \geq z)$.

13 Έστω X και Y δύο τυχαίες μεταβλητές που η από κοινού πυκνότητά τους δίνεται από τον παρακάτω πίνακα.

Υπολογίστε την πιθανότητα των εξής ενδεχομένων:
(α) Ο Y είναι άρτιος·
(β) Ο XY είναι περιτός·
(γ) $X > 0$ και $Y \geq 0$.

14 Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές που έχουν την ομοιόμορφη πυκνότητα στο $\{0, 1, \ldots, N\}$. Βρείτε τις
(α) $P(X \geq Y)$·
(β) $P(X = Y)$.

15 Θεωρούμε τις X και Y της Άσκησης 14. Βρείτε τις πυκνότητες των
(α) $\min(X, Y)$·
(β) $\max(X, Y)$·
(γ) $|Y - X|$.

16 Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές που έχουν γεωμετρικές πυκνότητες με παραμέτρους p_1 και p_2 αντίστοιχα. Βρείτε τις
(α) $P(X \geq Y)$·
(β) $P(X = Y)$.

17 Θεωρούμε τις X και Y της Άσκησης 16. Βρείτε τις πυκνότητες των

(α) $\min(X,Y)$·
(β) $X+Y$.

18 Έστω X και Y διακριτές τυχαίες μεταβλητές, και έστω g και h συναρτήσεις για τις οποίες ισχύει η εξής ταυτότητα:

$$P(X=x, Y=y) = g(x)h(y).$$

(α) Εκφράστε την $P(X=x)$ συναρτήσει των g και h.
(β) Εκφράστε την $P(Y=y)$ συναρτήσει των g και h.
(γ) Δείξτε ότι $(\sum_x g(x))(\sum_y h(y)) = 1$.
(δ) Δείξτε ότι οι X και Y είναι ανεξάρτητες.

19 Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές που έχουν γεωμετρική πυκνότητα με παράμετρο p. Θέτουμε $Z = Y - X$ και $M = \min(X,Y)$.
(α) Δείξτε ότι, αν ο z είναι ακέραιος και ο m μη αρνητικός ακέραιος, τότε

$$P(M=m, Z=z) = \begin{cases} P(X=m-z)P(Y=m) & , z<0, \\ \\ P(X=m)P(Y=m+z) & , z\geq 0. \end{cases}$$

(β) Συμπεράνατε από το (α) ότι αν $z \in \mathbb{Z}$ και $m \geq 0$

$$P(M=m, Z=z) = p^2(1-p)^{2m}(1-p)^{|z|}.$$

(γ) Χρησιμοποιώντας το (β) και την Άσκηση 18 δείξτε ότι οι M και Z είναι ανεξάρτητες.

20 Ένας κυκλικός στόχος χωρίζεται σε τρείς ζώνες με τρείς ομόκεντρους κύκλους ακτίνων 1/3, 1/2, και 1, όπως φαίνεται στο σχήμα που ακολουθεί.

Αν επιχειρήσουμε τρείς τυχαίες βολές προς τον στόχο, ποιά είναι η πιθανότητα να καταλήξει ακριβώς μία βολή σε κάθε ζώνη;

21 Υποθέτουμε ότι $2r$ βώλοι κατανέμονται τυχαία σε r δοχεία. Με X_i συμβολίζουμε το πλήθος των βώλων στο δοχείο i.
(α) Βρείτε την από κοινού πυκνότητα των $X_1, \ldots, X_r$.
(β) Βρείτε την πιθανότητα κάθε δοχείο να περιέχει ακριβώς 2 βώλους.

22 Θεωρούμε ένα πείραμα που έχει τρία δυνατά αποτελέσματα τα οποία συμβαίνουν με πιθανότητες p_1, p_2, και p_3, αντίστοιχα. Εκτελούμε n ανεξάρτητες επαναλήψεις του

πειράματος, και συμβολίζουμε με X_i το πλήθος των δοκιμών στις οποίες εμφανίζεται το i-στό αποτέλεσμα.
(α) Ποιά είναι η πυκνότητα της $X_1 + X_2$;
(β) Βρείτε την $P(X_2 = y|X_1 + X_2 = x)$, $y = 0, 1, \ldots, z$.

23 Χρησιμοποιήστε την προσέγγιση Poisson για να υπολογίσετε την πιθανότητα το πολύ 2 σε ένα δείγμα 50 ατόμων να έχουν πλαστή άδεια οδήγησης, αν είναι γνωστό ότι αυτό ισχύει για το 5% των οδηγών.

24 Χρησιμοποιήστε την προσέγγιση Poisson για να υπολογίσετε την πιθανότητα ένα κουτί με 100 ασφάλειες να περιέχει το πολύ 2 ελαττωματικές, αν είναι γνωστό ότι το 3% των ασφαλειών που κατασκευάζονται είναι ελαττωματικές.

25 Ρίχνουμε ένα ζάρι μέχρι να εμφανιστεί 6.
(α) Ποιά είναι η πιθανότητα να χρειαστούν το πολύ έξι ρίψεις;
(β) Πόσες ρίψεις απαιτούνται ώστε η πιθανότητα να εμφανιστεί 6 να είναι τουλάχιστον 1/2;

Οι ασκήσεις 26-30 σχετίζονται μεταξύ τους, και αφορούν μία άπειρη ακολουθία δοκιμών Bernoulli (δείτε την Παράγραφο 3.5).

26 Έστω T_i το πλήθος των δοκιμών που απαιτούνται μέχρι και την i-στή επιτυχία. Θεωρούμε τους ακεραίους $0 \leq x_1 < \ldots < x_r$. Υπολογίστε την πιθανότητα

$$P(T_1 = x_1, T_2 = x_2, \ldots, T_r = x_r).$$

Υπόδειξη: Θέτουμε $W_r = T_r - T_{r-1}$, $r \geq 2$, και $W_1 = T_1$. Τότε,

$$P(T_1 = x_1, \ldots, T_r = x_r) = P(W_1 = x_1, W_2 = x_2 - x_1, \ldots, W_r = x_r - x_{r-1}).$$

Τώρα, χρησιμοποιήστε το γεγονός ότι οι $W_1 - 1, \ldots, W_r - 1$ είναι ανεξάρτητες τυχαίες μεταβλητές και ακολουθούν γεωμετρική κατανομή με παράμετρο p.

27 Έστω N_n το πλήθος των επιτυχιών στις πρώτες n δοκιμές. Δείξτε ότι

$$P(T_1 = x \mid N_n = 1) = \frac{1}{n}, \qquad x = 1, 2, \ldots, n.$$

28 Πιό γενικά, δείξτε ότι αν $0 < x_1 < x_2 < \ldots < x_r \leq n$, τότε

$$P(T - 1 = x_1, T_2 = x_2, \ldots, T_r = x_r \mid N_n = r) = \binom{n}{r}^{-1}.$$

Αυτό δείχνει ότι αν έχουμε r επιτυχίες στις πρώτες n δοκιμές, τότε οι δοκιμές στις οποίες εμφανίστηκαν αυτές οι επιτυχίες συνιστούν τυχαίο δείγμα μεγέθους r (χωρίς επανατοποθέτηση) από τον «πληθυσμό» των δυνατών θέσεων.

29 Έστω k ένας θετικός ακέραιος, $k \leq r$. Από την Άσκηση 28 βλέπουμε αμέσως ότι

$$P(T_k = x \mid N_n = r) = \frac{\binom{x-1}{k-1}\binom{n-x}{r-k}}{\binom{n}{r}}.$$

Πράγματι, αν $T_k = x$ τότε στην θέση x έχουμε την k-στή επιτυχία. Στις πρώτες $x-1$ θέσεις πρέπει να έχουμε ακριβώς $k-1$ επιτυχίες, και στις τελευταίες $n-x$ θέσεις πρέπει να έχουμε ακριβώς $r-k$ επιτυχίες. Αφού, δεδομένου ότι $N_n = r$, οι θέσεις των r επιτυχιών συνιστούν τυχαίο δείγμα μεγέθους r από τον «πληθυσμό» των n θέσεων, έπεται το αποτέλεσμα. Επαληθεύστε αυτόν τον ισχυρισμό υπολογίζοντας απευθείας την $P(T_k = x \mid N_n = r)$.

30 Έστω $1 \leq i < j \leq r$ μη αρνητικοί ακέραιοι. Υπολογίστε την

$$P(T_i = x, T_j = y \mid N_n = r)$$

για $0 < x < y \leq n$.

31 Υποθέτουμε ότι οι ανεξάρτητες τυχαίες μεταβλητές X και Y έχουν την ομοιόμορφη πυκνότητα στο $\{1, 2, \ldots, N\}$. Υπολογίστε την πυκνότητα της $X + Y$.

32 Υποθέτουμε ότι η X είναι ομοιόμορφα κατανεμημένη στο $\{0, 1, 2, \ldots, N\}$. Βρείτε την $\Phi_X(t)$.

33 Έστω X μία μη αρνητική τυχαία μεταβλητή με ακέραιες τιμές, που έχει γεννήτρια συνάρτηση πιθανότητας την $\Phi_X(t) = e^{\lambda(t^2-1)}$, όπου $\lambda > 0$. Βρείτε την f_X.

34 Αποδείξτε το (iii) του Θεωρήματος 1.

35 Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές που έχουν πυκνότητες Poisson με παραμέτρους λ_1 και λ_2 αντίστοιχα. Βρείτε την $P(Y = y \mid X + Y = z)$ για $y = 0, \ldots, z$. *Υπόδειξη:* Χρησιμοποιήστε το (iii) του Θεωρήματος 1.

36 Έστω X, Y και Z ανεξάρτητες τυχαίες μεταβλητές που έχουν πυκνότητες Poisson με παραμέτρους λ_1, λ_2 και λ_3 αντίστοιχα. Βρείτε την

$$P(X = x, Y = y, Z = z \mid X + Y + Z = x + y + z),$$

όπου x, y και z είναι μη αρνητικοί ακέραιοι. *Υπόδειξη:* Χρησιμοποιήστε το (iii) του Θεωρήματος 1.

37 Στο Παράδειγμα 19 υποθέτουμε ότι η X_1 παίρνει τις τιμές 1 και 0 με πιθανότητες p και $1-p$ αντίστοιχα, όπου $0 < p < 1$. Υποθέτουμε επίσης ότι η N έχει πυκνότητα Poisson με παράμετρο λ.
(α) Χρησιμοποιήστε την (31) για να βρείτε τη γεννήτρια συνάρτηση πιθανότητας της S_N.
(β) Χρησιμοποιήστε το (α) για να βρείτε την πυκνότητα της S_N.

Για μία ερμηνεία της S_N υποθέστε ότι τυχαίο πλήθος N καρκινογόνων κυττάρων εισάγεται την χρονική στιγμή 0 και ότι κάθε κύτταρο, ανεξάρτητα από τα άλλα κύτταρα και ανεξάρτητα από το N, έχει πιθανότητα p να επιβιώσει αν επιχειρήσουμε χημειοθεραπεία. Θέτουμε $X_i = 1$ αν το i-στό κύτταρο επιβιώνει και $X_i = 0$ αλλιώς. Τότε, S_N είναι το πλήθος των κυττάρων που επιβιώνουν κατά τη θεραπεία.

38 Λύστε το (β) της Άσκησης 37 χωρίς να χρησιμοποιήσετε γεννήτριες συναρτήσεις πιθανότητας, χρησιμοποιώντας στη θέση τους την (32) και το γεγονός ότι η $X_1 + \ldots + X_n$ έχει διωνυμική πυκνότητα.

Κεφάλαιο 4

Μέση τιμή διακριτών τυχαίων μεταβλητών

Ας υποθέσουμε ότι συμμετέχουμε σε κάποιο τυχερό παιχνίδι. Κάθε φορά που παίζουμε, πρέπει να πληρώσουμε ένα ποσό a, και εισπράττουμε ένα ποσό X, όπου X είναι μία τυχαία μεταβλητή με πιθανές τιμές τις $x_1, x_2, \ldots, x_r$. Το ερώτημα είναι αν συμφέρει να συμμετάσχουμε. Αν το παιχνίδι παίζεται μόνο μία φορά, τότε το ερώτημα είναι πολύ δύσκολο. Ας υποθέσουμε όμως ότι παίζουμε το παιχνίδι πολλές φορές. Μετά από n παρτίδες θα έχουμε πληρώσει ποσό na και εισπράξει ποσό $X_1 + \ldots + X_n$. Αν υποθέσουμε ότι οι διαφορετικές παρτίδες του παιχνιδιού συνιστούν ανεξάρτητες επαναλήψεις του ιδίου πειράματος (το οποίο παρατηρεί μία τιμή της X), τότε μπορούμε να θεωρήσουμε ότι οι τυχαίες μεταβλητές $X_1, X_2, \ldots, X_n$ είναι ανεξάρτητες και έχουν κοινή πυκνότητα την πυκνότητα f της X. Συμβολίζουμε με $N_n(x_i)$ το πλήθος των παρτίδων που έδωσαν την τιμή x_i, δηλαδή το πλήθος των X_i που παίρνουν την τιμή x_i. Τότε μπορούμε να γράφουμε

$$X_1 + \ldots + X_n = \sum_{i=1}^{r} x_i N_n(x_i).$$

Το μέσο ποσό που εισπράτουμε είναι

$$\frac{X_1 + \ldots + X_n}{n} = \sum_{i=1}^{r} x_i \left[\frac{N_n(x_i)}{n} \right].$$

Ερμηνεύοντας την πιθανότητα σαν σχετική συχνότητα, αν το n είναι μεγάλο, περιμένουμε οι αριθμοί $N_n(x_i)/n$ να είναι περίπου ίσοι με $f(x_i)$, επομένως το άθροισμα στο δεξιό μέλος θα πρέπει να είναι κατά προσέγγιση ίσο με $\mu = \sum_{i=1}^{r} x_i f(x_i)$. Φαίνεται λοιπόν λογικό να αναμένουμε καθαρό κέρδος από το παιχνίδι αν $\mu > a$ και καθαρή ζημία αν $\mu < a$. Αν $\mu = a$, θα αναμέναμε μηδενικό κέρδος ή ζημία.

Η ποσότητα $\sum_{i=1}^{r} x_i f(x_i)$ λέγεται μέση τιμή της τυχαίας μεταβλητής X. Πιο γενικά, έστω X μία διακριτή τυχαία μεταβλητή που παίρνει τις πεπερασμένες το πλήθος τιμές $x_1, \ldots, x_r$. Τότε η μέση τιμή της X, η οποία συμβολίζεται με $\mathbb{E}X$ ή μ, είναι ο αριθμός

$$\mathbb{E}X = \sum_{i=1}^{r} x_i f(x_i), \tag{1}$$

όπου f είναι η πυκνότητα της X.

Υποθέτουμε ότι η X έχει την ομοιόμορφη κατανομή στο σύνολο $\{x_1, \ldots, x_r\}$. Τότε $f(x_i) = P(X = x_i) = r^{-1}$, και από την (1) βλέπουμε ότι $\mathbb{E}X = (x_1 + \ldots + x_r)r^{-1}$, άρα σε αυτήν την περίπτωση η $\mathbb{E}X$ είναι απλώς ο αριθμητικός μέσος των πιθανών τιμών της X. Γενικά, η (1) δείχνει ότι η $\mathbb{E}X$ είναι ένας *μέσος όρος με βάρη* των πιθανών τιμών της X. Το βάρος που αντιστοιχεί στην i-στή τιμή x_i είναι η πιθανότητά της $f(x_i)$.

Η μέση τιμή $\mathbb{E}X$ λέγεται και μέσος της X (ή της πυκνότητας f της X) και συνήθως συμβολίζεται με μ. Ο μέσος μας δίνει έναν τρόπο να προσπαθήσουμε να συνοψίσουμε μία κατανομή πιθανότητας σε έναν μόνο αριθμό που υποτίθεται ότι αντιπροσωπεύει μία «τυπική τιμή» της X. Η επιτυχία αυτής της προσπάθειας εξαρτάται από τον βαθμό στον οποίο οι τιμές της X συγκεντρώνονται γύρω από την τιμή μ. Θα μελετήσουμε αυτό το πρόβλημα λεπτομερέστερα όταν θα συζητήσουμε την διασπορά της X στην Παράγραφο 4.3.

Παράδειγμα 1. *Διωνυμική κατανομή.* Υποθέτουμε ότι η X ακολουθεί διωνυμική κατανομή με παραμέτρους n και p. Βρείτε την $\mathbb{E}X$.

Για $n = 1$, η X παίρνει τις δύο τιμές 0 και 1 με πιθανότητες $(1-p)$ και p αντίστοιχα. Επομένως,

$$\mathbb{E}X = 0.P(X = 0) + 1.P(X = 1) = p.$$

Μία τυχαία μεταβλητή που έχει διωνυμική πυκνότητα με παραμέτρους 1 και p είναι απλώς μία δείκτρια τυχαία μεταβλητή, βλέπουμε λοιπόν ότι μπορούμε να βρούμε την πιθανότητα του ενδεχομένου $X = 1$ υπολογίζοντας την μέση τιμή της δείκτριας συνάρτησης.

Υπολογίζουμε τώρα την $\mathbb{E}X$ για κάθε $n \geq 1$. Σε αυτήν την περίπτωση η X παίρνει τις τιμές $0, 1, 2, \ldots, n$, και

$$\mathbb{E}X = \sum_{j=0}^{n} j \binom{n}{j} p^j (1-p)^{n-j}.$$

Για τον υπολογισμό αυτής της ποσότητας παρατηρούμε ότι

$$\begin{aligned} j\binom{n}{j} &= \frac{jn!}{j!(n-j)!} \\ &= \frac{n(n-1)!}{(j-1)![(n-1)-(j-1)]!} \\ &= n\binom{n-1}{j-1}. \end{aligned}$$

Άρα,

$$\mathbb{E}X = n\sum_{j=1}^{n}\binom{n-1}{j-1}p^j(1-p)^j.$$

Κάνοντας την αλλαγή μεταβλητής $i = j-1$ βλέπουμε ότι

$$\mathbb{E}X = np\sum_{i=0}^{n-1}\binom{n-1}{i}p^i(1-p)^{n-i-1}.$$

Από το διωνυμικό θεώρημα

$$\sum_{i=0}^{n-1}\binom{n-1}{i}p^i(1-p)^{n-i-1} = [p+(1-p)]^{n-1} = 1,$$

οπότε καταλήγουμε στην

$$\mathbb{E}X = np.$$

4.1 Ορισμός της μέσης τιμής

Υποθέτουμε τώρα ότι X είναι τυχούσα διακριτή τυχαία μεταβλητή με τιμές τις $x_1, x_2, \ldots$. Θα θέλαμε να ορίσουμε την μέση τιμή της X μέσω της

$$\mathbb{E}X = \sum_{j=1}^{\infty} x_j f(x_j). \tag{2}$$

Αν η X παίρνει μόνο πεπερασμένες το πλήθος τιμές $x_1, \ldots, x_r$, τότε η (2) συμφωνεί με τον προηγούμενο ορισμό μας. Στην γενική διακριτή περίπτωση, ο παραπάνω ορισμός έχει νόημα αν το άθροισμα $\sum_j x_j f(x_j)$ είναι καλά ορισμένο. Για να ικανοποιήσουμε αυτήν την απαίτηση, ζητάμε να ισχύει η ισχυρότερη συνθήκη $\sum_j |x_j| f(x_j) < \infty$. Οδηγούμαστε έτσι στον ακόλουθο ορισμό.

Ορισμός 1. *Έστω X διακριτή τυχαία μεταβλητή με πυκνότητα f. Αν ικανοποιείται η $\sum_j |x_j| f(x_j) < \infty$, τότε λέμε ότι η X έχει πεπερασμένη μέση τιμή και ορίζουμε την μέση τιμή της μέσω της (2). Αν $\sum_j |x_j| f(x_j) = \infty$, τότε λέμε ότι η X δεν έχει πεπερασμένη μέση τιμή, και η $\mathbb{E}X$ δεν ορίζεται.*

Αν X είναι μία μη αρνητική τυχαία μεταβλητή, γράφουμε $\mathbb{E}X < \infty$ για να συμβολίσουμε το γεγονός ότι η X έχει πεπερασμένη μέση τιμή.

Παράδειγμα 2. *Κατανομή Poisson.* Έστω X τυχαία μεταβλητή που ακολουθεί κατανομή Poisson με παράμετρο λ. Τότε,

$$\begin{aligned}\mathbb{E}X &= \sum_{j=1}^{\infty} j\frac{\lambda^j}{j!}e^{-\lambda} = \sum_{j=1}^{\infty}\frac{\lambda^j}{(j-1)!}e^{-\lambda}\\ &= \lambda e^{-\lambda}\sum_{j=0}^{\infty}\frac{\lambda^j}{j!} = \lambda e^{-\lambda}e^{\lambda} = \lambda.\end{aligned}$$

Παράδειγμα 3. *Γεωμετρική κατανομή.* Υποθέτουμε ότι η X ακολουθεί γεωμετρική κατανομή με παράμετρο p. Βρείτε την $\mathbb{E}X$.

Τώρα,

$$\begin{aligned}\mathbb{E}X &= \sum_{j=0}^{\infty} jp(1-p)^j \\ &= p(1-p)\sum_{j=0}^{\infty} j(1-p)^{j-1} \\ &= -p(1-p)\sum_{j=0}^{\infty} \frac{d}{dp}(1-p)^j.\end{aligned}$$

Παραγωγίζοντας τη δυναμοσειρά όρο προς όρο, συμπεραίνουμε ότι

$$\mathbb{E}X = -p(1-p)\frac{d}{dp}\sum_{j=0}^{\infty}(1-p)^j.$$

Χρησιμοποιώντας τον τύπο για το άθροισμα γεωμετρικής προόδου, βλέπουμε ότι

$$\mathbb{E}X = -p(1-p)\frac{d}{dp}\left(\frac{1}{p}\right) = -p(1-p)\left(-\frac{1}{p^2}\right).$$

Επομένως,

$$\mathbb{E}X = \frac{1-p}{p}.$$

Παράδειγμα 4. Θεωρούμε τη συνάρτηση f που ορίζεται στο $\mathbb{R}$ από την

$$f(x) = \begin{cases} \frac{1}{x(x+1)} & , x = 1, 2, \ldots, \\ 0 & , \text{αλλιώς}. \end{cases}$$

Η συνάρτηση f προφανώς ικανοποιεί τις ιδιότητες (i) και (ii) του ορισμού των συναρτήσεων πυκνότητας που δώσαμε στο Κεφάλαιο 3. Για να δούμε ότι η f ικανοποιεί την ιδιότητα (iii) παρατηρούμε ότι

$$\frac{1}{x(x+1)} = \frac{1}{x} - \frac{1}{x+1}$$

άρα

$$\begin{aligned}\sum_{x=1}^{\infty} f(x) &= \sum_{x=1}^{\infty}\left[\frac{1}{x} - \frac{1}{x+1}\right] \\ &= (1-1/2) + (1/2-1/3) + \ldots = 1.\end{aligned}$$

Άρα η (iii) ισχύει, και η f είναι πυκνότητα. Όμως η f δεν έχει πεπερασμένη μέση τιμή γιατί

$$\sum_{x=1}^{\infty} |x| f(x) = \sum_{x=1}^{\infty} \frac{1}{x+1}$$

και η αρμονική σειρά $\sum_{x=1}^{\infty} x^{-1}$ αποκλίνει.

4.2 Ιδιότητες της μέσης τιμής

Συχνά θέλουμε να υπολογίσουμε την μέση τιμή τυχαίων μεταβλητών όπως οι $Z = X_1 + X_2$ και $Z = X^2$, που είναι συναρτήσεις $\varphi(\mathbf{X})$ του τυχαίου διανύσματος $\mathbf{X}$. Φυσικά, αν γνωρίζουμε την πυκνότητα f_Z της Z, αυτό μπορεί να γίνει με απευθείας χρήση της (2). Πολύ συχνά όμως, η πυκνότητα της Z είναι άγνωστη ή ο υπολογισμός της $\mathbb{E}Z$ από την γνωστή πυκνότητα της Z μπορεί να είναι πολύ δύσκολος. Το αποτέλεσμα που ακολουθεί μας δίνει έναν τρόπο να αποφασίζουμε αν η Z έχει πεπερασμένη μέση τιμή και, αν έχει, να υπολογίζουμε την $\mathbb{E}Z$ απευθείας με βάση την πυκνότητα $f_{\mathbf{X}}$ και τη συνάρτηση φ.

Για τη διατύπωση του αποτελέσματος θα χρειαστούμε μία σύμβαση σχετικά με τον συμβολισμό. Έστω $\mathbf{X}$ ένα διακριτό r-διάστατο τυχαίο διάνυσμα που παίρνει τις τιμές $\mathbf{x}_1, \mathbf{x}_2, \ldots$ και έχει πυκνότητα f, και έστω φ μία συνάρτηση με πραγματικές τιμές, ορισμένη στον $\mathbb{R}^r$. Τότε, το $\sum_{\mathbf{x}} \varphi(\mathbf{x}) f(\mathbf{x})$ ορίζεται από την

$$\sum_{\mathbf{x}} \varphi(\mathbf{x}) f(\mathbf{x}) = \sum_{j} \varphi(\mathbf{x}_j) f(\mathbf{x}_j). \tag{3}$$

Θεώρημα 1. *Έστω* $\mathbf{X}$ *ένα διακριτό* r*-διάστατο τυχαίο διάνυσμα με πυκνότητα* f*, και έστω* φ *μία συνάρτηση με πραγματικές τιμές, ορισμένη στον* $\mathbb{R}^r$*. Τότε, η τυχαία μεταβλητή* $Z = \varphi(\mathbf{X})$ *έχει πεπερασμένη μέση τιμή αν και μόνο αν*

$$\sum_{\mathbf{x}} |\varphi(\mathbf{x})| f(\mathbf{x}) < \infty \tag{4}$$

και, αν ισχύει η (4)*,*

$$\mathbb{E}Z = \sum_{\mathbf{x}} \varphi(\mathbf{x}) f(\mathbf{x}) \tag{5}$$

Απόδειξη: Έστω $z_1, z_2, \ldots$ οι πιθανές τιμές της Z και $\mathbf{x}_1, \mathbf{x}_2, \ldots$ οι πιθανές τιμές του X. Για κάθε z_j υπάρχει τουλάχιστον ένα $\mathbf{x}_i$ τέτοιο ώστε $z_j = \varphi(\mathbf{x}_i)$, είναι όμως πιθανό να υπάρχουν περισσότερα από ένα τέτοια $\mathbf{x}_i$. Συμβολίζουμε με A_j το σύνολο όλων αυτών των $\mathbf{x}_i$, δηλαδή,

$$A_j = \{\mathbf{x}_i : \varphi(\mathbf{x}_i) = z_j\}.$$

Τότε, τα ενδεχόμενα $\{\mathbf{X} \in A_j\}$ και $\{Z = z_j\}$ ταυτίζονται. Άρα,

$$P(Z = z_j) = P(\mathbf{X} \in A_j) = \sum_{\mathbf{x} \in A_j} f_{\mathbf{X}}(\mathbf{x}).$$

Επομένως,

$$\begin{aligned}\sum_j |z_j| f_Z(z_j) &= \sum_j |z_j| P(Z = z_j) \\ &= \sum_j |z_j| \sum_{\mathbf{x} \in A_j} f_{\mathbf{X}}(\mathbf{x}) \\ &= \sum_j \sum_{\mathbf{x} \in A_j} |z_j| f_{\mathbf{X}}(\mathbf{x}).\end{aligned}$$

Αφού $\varphi(\mathbf{x}) = z_j$ για κάθε $\mathbf{x} \in A_j$, έπεται ότι

$$\sum_j |z_j| f_Z(z_j) = \sum_j \sum_{\mathbf{x} \in A_j} |\varphi(\mathbf{x})| f_{\mathbf{X}}(\mathbf{x}).$$

Από τον ορισμό τους, τα σύνολα A_j είναι ξένα ανά δύο και η ένωσή τους είναι το σύνολο όλων των πιθανών τιμών του $\mathbf{X}$. Άρα,

$$\sum_j |z_j| f_Z(z_j) = \sum_{\mathbf{x}} |\varphi(\mathbf{x})| f_{\mathbf{X}}(\mathbf{x}).$$

Αυτό δείχνει ότι η Z έχει πεπερασμένη μέση τιμή αν και μόνο αν ισχύει η (4).

Αν η Z έχει πεπερασμένη μέση τιμή, τότε επαναλαμβάνουμε το ίδιο επιχείρημα παραλείποντας τις απόλυτες τιμές, και καταλήγουμε στην (5). $\square$

Έστω X μία τυχαία μεταβλητή με πυκνότητα f, και έστω $\varphi(x) = |x|$. Από το Θεώρημα 1, η $|X|$ έχει πεπερασμένη μέση τιμή αν και μόνο αν $\sum_x |x| f(x) < \infty$. Όμως, σύμφωνα με τον ορισμό που δώσαμε για τη μέση τιμή, η X έχει πεπερασμένη μέση τιμή αν και μόνο αν η ίδια ακριβώς σειρά συγκλίνει. Βλέπουμε λοιπόν ότι η X έχει πεπερασμένη μέση τιμή αν και μόνο αν $\mathbb{E}|X| < \infty$.

Χρησιμοποιώντας το Θεώρημα 1 θα αποδείξουμε τις πιό σημαντικές ιδιότητες της μέσης τιμής.

Θεώρημα 2. *Έστω X και Y δύο τυχαίες μεταβλητές με πεπερασμένη μέση τιμή.*

(i) *Αν c είναι μία σταθερά και $P(X = c) = 1$, τότε $\mathbb{E}X = c$.*

(ii) *Αν c είναι μία σταθερά, τότε η cX έχει πεπερασμένη μέση τιμή και $\mathbb{E}(cX) = c\mathbb{E}X$.*

(iii) *Η $X + Y$ έχει πεπερασμένη μέση τιμή και*

$$\mathbb{E}(X + Y) = \mathbb{E}X + \mathbb{E}Y.$$

(iv) *Υποθέτουμε ότι* $P(X \geq Y) = 1$. *Τότε* $\mathbb{E}X \geq \mathbb{E}Y$. *Επιπλέον,* $\mathbb{E}X = \mathbb{E}Y$ *αν και μόνο αν* $P(X = Y) = 1$.

(v) $|\mathbb{E}X| \leq \mathbb{E}|X|$.

Απόδειξη: Η απόδειξη της (i) είναι πολύ απλή. Αν $P(X = c) = 1$, τότε η X έχει πυκνότητα $f_X(x) = 0$ αν $x \neq c$ και $f_X(c) = 1$. Από την (2) παίρνουμε

$$\mathbb{E}X = \sum_x x f_X(x) = c f_X(c) = c.$$

Για να αποδείξουμε την (ii) θέτουμε $\varphi(x) = cx$ και παρατηρούμε ότι

$$\sum_x |cx| f_X(x) = |c| \sum_x |x| f_X(x) < \infty,$$

οπότε η cX έχει πεπερασμένη μέση τιμή. Επομένως από την (5)

$$\mathbb{E}(cX) = \sum_x (cx) f_X(x) = c \sum_x x f_X(x) = c\mathbb{E}X.$$

Για να αποδείξουμε την (iii) θέτουμε $\varphi(x, y) = x + y$ και γράφουμε f για την από κοινού πυκνότητα των X και Y. Τότε

$$\begin{aligned}
\sum_{x,y} |x + y| f(x, y) &\leq \sum_{x,y} |x| f(x, y) + \sum_{x,y} |y| f(x, y) \\
&= \sum_x |x| \sum_y f(x, y) + \sum_y |y| \sum_x f(x, y) \\
&= \sum_x |x| f_X(x) + \sum_y |y| f_Y(y) < \infty
\end{aligned}$$

το οποίο αποδεικνύει ότι η $X + Y$ έχει πεπερασμένη μέση τιμή. Εφαρμόζοντας την (5) βλέπουμε ότι

$$\begin{aligned}
\mathbb{E}(X + Y) &= \sum_{x,y} (x + y) f(x, y) \\
&= \sum_{x,y} x f(x, y) + \sum_{x,y} y f(x, y) \\
&= \mathbb{E}X + \mathbb{E}Y.
\end{aligned}$$

Για την απόδειξη της (iv) παρατηρούμε ότι $Z = X - Y = X + (-Y)$, και από τις (ii) και (iii) βλέπουμε ότι

$$\mathbb{E}X - \mathbb{E}Y = \mathbb{E}(X - Y) = \mathbb{E}Z = \sum_z z f_Z(z).$$

Αφού $P(Z \geq 0) = P(X \geq Y) = 1$, οι τιμές z_i που παίρνει η $Z = X - Y$ είναι όλες μη αρνητικές. Επομένως $\sum_z z f_Z(z) \geq 0$, δηλαδή $\mathbb{E}X - \mathbb{E}Y \geq 0$. Αυτό αποδεικνύει το πρώτο μέρος της (iv). Αν $\mathbb{E}X = \mathbb{E}Y$, τότε $\mathbb{E}Z = 0$. Όμως τότε,

$$0 = \mathbb{E}Z = \sum_i z_i f_Z(z_i).$$

Αλλά το άθροισμα μη αρνητικών όρων μπορεί να ισούται με μηδέν μόνο αν όλοι οι όροι είναι ίσοι με μηδέν. Αφού $f_Z(z_i) > 0$ πρέπει να έχουμε $z_i = 0$. Δηλαδή η μοναδική δυνατή τιμή της Z είναι 0, το οποίο είναι ισοδύναμο με την $P(Z = 0) = 1$.

Τέλος, η (v) προκύπτει από τις (iv) και (ii), γιατί $-|X| \leq X \leq |X|$ άρα $-\mathbb{E}|X| \leq \mathbb{E}X \leq \mathbb{E}|X|$. Αυτό ολοκληρώνει την απόδειξη του θεωρήματος. □

Από τις (ii) και (iii) βλέπουμε εύκολα ότι αν $X_1, \ldots, X_n$ είναι οποιεσδήποτε n τυχαίες μεταβλητές με πεπερασμένη μέση τιμή, και $c_1, \ldots, c_n$ είναι n σταθερές, τότε

$$\mathbb{E}(c_1 X_1 + \ldots + c_n X_n) = c_1 \mathbb{E}X_1 + \ldots + \mathbb{E}X_n. \tag{6}$$

Μία χρήσιμη παρατήρηση είναι ότι κάθε φραγμένη τυχαία μεταβλητή έχει πεπερασμένη μέση τιμή. Ακριβέστερα,

Θεώρημα 3. *Έστω X μία τυχαία μεταβλητή που ικανοποιεί την $P(|X| \leq M) = 1$ για κάποια θετική σταθερά M. Τότε, η X έχει πεπερασμένη μέση τιμή και $|\mathbb{E}X| \leq M$.*

Απόδειξη: Έστω $x_1, x_2, \ldots$ οι πιθανές τιμές της X. Τότε $|x_i| \leq M$ για κάθε i. Πράγματι, αν $|x_i| > M$ για κάποια τιμή x_i, τότε

$$P(|X| > M) \geq P(|X| = |x_i|) > 0,$$

το οποίο αντιφάσκει προς την $P(|X| \leq M) = 1$. Επομένως,

$$\sum_i |x_i| f(x_i) \leq M \sum_i f(x_i) \leq M,$$

δηλαδή η X έχει πεπερασμένη μέση τιμή. Επιπλέον, από την (v) του Θεωρήματος 2,

$$|\mathbb{E}X| \leq \mathbb{E}|X| = \sum_i |x_i| f(x_i) \leq M.$$

Αυτό ολοκληρώνει την απόδειξη. □

Από το Θεώρημα 3 και την (iii) του Θεωρήματος 2 προκύπτει εύκολα ότι αν X και Y είναι δύο τυχαίες μεταβλητές τέτοιες ώστε η Y να έχει πεπερασμένη μέση τιμή και $P(|X - Y| \leq M) = 1$ για κάποια σταθερά M, τότε η X έχει επίσης πεπερασμένη μέση τιμή και $|\mathbb{E}X - \mathbb{E}Y| \leq M$. Αφήνουμε την απόδειξη αυτού του ισχυρισμού στον αναγνώστη.

Η μέση τιμή του αθροίσματος δύο τυχαίων μεταβλητών είναι το άθροισμα των μέσων τιμών τους, θα μπορούσε λοιπόν κανείς να υποθέσει ότι και η μέση τιμή ενός γινομένου είναι το γινόμενο των μέσων τιμών. Αυτό δεν είναι γενικά σωστό: σαν παράδειγμα μπορεί κανείς να πάρει την τυχαία μεταβλητή X που παίρνει τις τιμές 1 και -1 με πιθανότητα $1/2$ και να θέσει $Y = X$. Τότε $\mathbb{E}X = \mathbb{E}Y = 0$ αλλά $\mathbb{E}(XY) = \mathbb{E}(X^2) = 1$.

Υπάρχει όμως μία σημαντική περίπτωση στην οποία αυτός ο κανόνας του γινομένου ισχύει: όταν οι X και Y είναι ανεξάρτητες τυχαίες μεταβλητές. Η τυπική διατύπωση έχει ως εξής.

Θεώρημα 4. *Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές με πεπερασμένες μέσες τιμές. Τότε, η XY έχει πεπερασμένη μέση τιμή και*

$$\mathbb{E}(XY) = (\mathbb{E}X)(\mathbb{E}Y). \tag{7}$$

Απόδειξη: Παρατηρούμε ότι, αφού οι X και Y είναι ανεξάρτητες, η από κοινού πυκνότητα των X και Y είναι η $f_X(x)f_Y(y)$. Άρα,

$$\begin{aligned}\sum_{x,y}|xy|f(x,y) &= \sum_{x,y}|x||y|f_X(x)f_Y(y)\\ &= \left(\sum_x |x|f_X(x)\right)\left(\sum_y |y|f_Y(y)\right) < \infty,\end{aligned}$$

δηλαδή η XY έχει πεπερασμένη μέση τιμή. Χρησιμοποιώντας το Θεώρημα 1, συμπεραίνουμε ότι

$$\begin{aligned}\mathbb{E}(XY) &= \sum_{x,y}(xy)f_X(x)f_Y(y)\\ &= \left[\sum_x xf_X(x)\right]\left[\sum_y yf_Y(y)\right] = (\mathbb{E}X)(\mathbb{E}Y).\square\end{aligned}$$

Το αντίστροφο δεν ισχύει. Δύο τυχαίες μεταβλητές X και Y μπορεί να ικανοποιούν την $\mathbb{E}(XY) = (\mathbb{E}X)(\mathbb{E}Y)$ χωρίς να είναι ανεξάρτητες.

Παράδειγμα 5. Υποθέτουμε ότι το ζεύγος (X,Y) παίρνει τις τιμές $(1,0)$, $(0,1)$, $(-1,0)$ και $(0,-1)$ με ίσες πιθανότητες. Τότε, $\mathbb{E}X = \mathbb{E}Y = 0$. Αφού $XY = 0$, έπεται ότι $\mathbb{E}(XY) = 0$ άρα $\mathbb{E}(XY) = (\mathbb{E}X)(\mathbb{E}Y)$. Όμως οι X και Y δεν είναι ανεξάρτητες. Παρατηρήστε, για παράδειγμα, ότι $P(X=0) = P(Y=0) = 1/2$ ενώ $P(X=0, Y=0) = 0$. Δηλαδή,

$$P(X=0,Y=0) \neq P(X=0)P(Y=0).$$

Είναι συχνά ευκολότερο να υπολογίσουμε μέσες τιμές χρησιμοποιώντας τις ιδιότητες που μας δίνει το Θεώρημα 2, και όχι να δουλέψουμε απευθείας με τον ορισμό. Θα περιγράψουμε αυτήν την τεχνική με μία σειρά παραδειγμάτων.

Παράδειγμα 6. *Διωνυμική κατανομή.* Από το Παράδειγμα 1 γνωρίζουμε ήδη ότι ο μέσος της διωνυμικής κατανομής με παραμέτρους n και p είναι np. Μπορούμε να καταλήξουμε στο ίδιο συμπέρασμα με πολύ απλό τρόπο, χρησιμοποιώντας το γεγονός ότι η μέση τιμή ενός αθροίσματος είναι το άθροισμα των μέσων τιμών (την (iii) του Θεωρήματος 2). Για το σκοπό αυτό, θεωρούμε n ανεξάρτητες τυχαίες μεταβλητές Bernoulli $X_1, \ldots, X_n$ με παράμετρο p, και θέτουμε $S_n = X_1 + \ldots + X_n$. Τότε, η S_n ακολουθεί τη διωνυμική κατανομή με παραμέτρους n και p. Από το πρώτο μέρος του Παραδείγματος 1, $\mathbb{E}X_i = p, 1 \leq i \leq n$, επομένως

$$\mathbb{E}S_n = \mathbb{E}(X_1 + \ldots X_n) = \sum_{i=1}^n \mathbb{E}X_i = np.$$

Παράδειγμα 7. *Υπεργεωμετρική κατανομή.* Υποθέτουμε ότι έχουμε έναν πληθυσμό r ατόμων, r_1 από τα οποία είναι τύπου I και $r-r_1$ τύπου II. Από αυτόν τον πληθυσμό παίρνουμε ένα δείγμα μεγέθους n χωρίς επανατοποθέτηση. Συμβολίζουμε με S_n το πλήθος των ατόμων τύπου I στο δείγμα. Υπολογίστε την $\mathbb{E}S_n$.

Ξέρουμε ότι η S_n ακολουθεί την υπεργεωμετρική κατανομή, θα μπορούσαμε λοιπόν να υπολογίσουμε την $\mathbb{E}S_n$ χρησιμοποιώντας την (2). Είναι όμως πολύ απλούστερο να δουλέψουμε με τη βοήθεια δεικτριών τυχαίων μεταβλητών $X_1, \ldots, X_n$ ως εξής. Ζητάμε η τυχαία μεταβλητή X_i να παίρνει την τιμή 1 αν και μόνο αν το i-στό στοιχείο του δείγματος είναι τύπου I. Τότε,

$$\mathbb{E}X_i = P(X_i = 1) = \frac{r_1}{r}.$$

Όμως $S_n = X_1 + \ldots + X_n$, οπότε χρησιμοποιώντας την (iii) του Θεωρήματος 2 βλέπουμε ότι

$$\mathbb{E}S_n = \sum_{i=1}^{n} \mathbb{E}X_i = n\frac{r_1}{r}.$$

Παρατηρήστε ότι οι τυχαίες μεταβλητές $X_i, 1 \leq i \leq n$, *δεν* είναι ανεξάρτητες.

Παράδειγμα 8. Υποθέτουμε ότι έχουμε έναν πληθυσμό από r διακεκριμένα άτομα τα οποία ονομάζουμε $1, 2, \ldots, r$. Επιλέγουμε άτομα με επανατοποθέτηση μέχρι τη στιγμή κατά την οποία έχουν επιλεγεί ακριβώς $k \leq r$ διαφορετικά άτομα. Συμβολίζουμε με S_k το μέγεθος του δείγματος που απαιτείται. Υπολογίστε την $\mathbb{E}S_k$.

Είναι φανερό ότι $S_1 = 1$ επομένως $\mathbb{E}S_1 = 1$. Υποθέτουμε ότι $k \geq 2$ και θέτουμε $X_i = S_{i+1} - S_i$, $i = 1, 2, \ldots, k-1$. Τότε, $S_k = 1 + X_1 + \ldots + X_{k-1}$. Τώρα, X_i είναι το πλήθος των ατόμων που πρέπει να επιλέξουμε από τη στιγμή που το i-στό άτομο μπαίνει στο δείγμα μέχρι τη στιγμή που το $(i+1)$-στό νέο άτομο θα μπεί στο δείγμα. Με λίγη σκέψη βλέπουμε ότι το ενδεχόμενο $\{X_i = n\}$ συμβαίνει αν και μόνο αν τα πρώτα $n-1$ άτομα που επιλέγουμε μετά την εμφάνιση του i-στού ατόμου στο δείγμα συμπίπτουν με κάποιο από τα προηγούμενα i αντικείμενα, και το n-οστό άτομο είναι διαφορετικό από αυτά. Έτσι, αφού οι δοκιμές είναι ανεξάρτητες,

$$P(X_i = n) = \left(\frac{i}{r}\right)^{n-1}\left(1 - \frac{i}{r}\right) \quad , \quad n = 1, 2, \ldots.$$

Αυτό δείχνει ότι η τυχαία μεταβλητή $X_i - 1$ είναι γεωμετρική με παράμετρο $p_i = 1 - (i/r)$. Άρα, από το Παράδειγμα 3, $\mathbb{E}(X_i - 1) = p_i^{-1}(1-p_i)$ και

$$\mathbb{E}X_i = p_i^{-1}(1-p_i) + 1 = p_i^{-1} = (1 - i/r)^{-1} = r(r-i)^{-1}.$$

Επομένως,

$$\mathbb{E}S_k = 1 + \sum_{i=1}^{k}\left(\frac{r}{r-i}\right) \qquad (8)$$

$$
\begin{aligned}
&= \sum_{i=0}^{k-1} \left(\frac{r}{r-i} \right) \\
&= r\left(\frac{1}{r} + \frac{1}{r-1} + \ldots + \frac{1}{r-k+1} \right).
\end{aligned}
$$

Σημειώνουμε για μεταγενέστερη χρήση ότι από την κατασκευή τους οι X_i είναι προφανώς ανεξάρτητες τυχαίες μεταβλητές.

Στο προηγούμενο κεφάλαιο είδαμε ότι οι τυχαίες μεταβλητές που παίρνουν μη αρνητικές ακέραιες τιμές έχουν ιδιαίτερη σημασία. Το Θεώρημα που ακολουθεί εφαρμόζεται συχνά για να αποφασίσουμε αν μία τυχαία μεταβλητή X αυτού του είδους έχει πεπερασμένη μέση τιμή, αλλά και για να υπολογίσουμε τη μέση τιμή της X.

Θεώρημα 5. *Έστω X τυχαία μεταβλητή που παίρνει μη αρνητικές ακέραιες τιμές. Τότε η X έχει πεπερασμένη μέση τιμή αν και μόνο αν η σειρά $\sum_{x=1}^{\infty} P(X \geq x)$ συγκλίνει. Αν η σειρά συγκλίνει, τότε*

$$
\mathbb{E}X = \sum_{x=1}^{\infty} P(X \geq x). \tag{9}
$$

Απόδειξη: Θα δείξουμε ότι

$$
\sum_{x=1}^{\infty} xP(X = x) = \sum_{x=1}^{\infty} P(X \geq x), \tag{10}
$$

σχέση από την οποία το θεώρημα προκύπτει άμεσα. Για το σκοπό αυτό γράφουμε πρώτα το αριστερό μέλος της (10) στη μορφή

$$
\sum_{x=1}^{\infty} P(X = x) \sum_{y=1}^{x} 1.
$$

Αλλάζοντας την σειρά της άθροισης, ξαναγράφουμε την τελευταία ποσότητα στη μορφή

$$
\sum_{y=1}^{\infty} \sum_{x=y}^{\infty} P(X = x) = \sum_{y=1}^{\infty} P(X \geq y).
$$

Αντικαθιστώντας τη μεταβλητή y με τη μεταβλητή x στο δεξιό μέλος αυτής της ισότητας, παίρνουμε το δεξιό μέλος της (10). Αυτό αποδεικνύει την (10), όπως θέλαμε. □

Για μία στοιχειώδη εφαρμογή του θεωρήματος, ας υποθέσουμε ότι η τυχαία μεταβλητή X ακολουθεί γεωμετρική κατανομή με παράμετρο p. Τότε $P(X \geq x) = (1-p)^x$, άρα από το θεώρημα,

$$
\mathbb{E}X = \sum_{x=1}^{\infty} (1-p)^x = (1-p) + (1-p)^2 + \ldots = p^{-1}(1-p).
$$

Αυτό συμφωνεί με το αποτέλεσμα που βρήκαμε στο Παράδειγμα 3.

4.3 Ροπές

Έστω X διακριτή τυχαία μεταβλητή, και $r \geq 0$ ακέραιος. Λέμε ότι η X έχει ροπή τάξης r αν η X^r έχει πεπερασμένη μέση τιμή. Σε αυτήν την περίπτωση συμβολίζουμε την ροπή τάξης r της X με $\mathbb{E}X^r$. Αν η X έχει ροπή τάξης r, τότε η ροπή τάξης r της $X-\mu$, όπου μ είναι ο μέσος της X, λέγεται *κεντρική ροπή τάξης* r της X. Από το Θεώρημα 1 μπορούμε να υπολογίσουμε την ροπή τάξης r και την κεντρική ροπή τάξης r της X απευθείας από την πυκνότητα f με βάση τους τύπους

$$\mathbb{E}X^r = \sum_x x^r f(x) \tag{11}$$

και

$$\mathbb{E}(X-\mu)^r = \sum_x (x-\mu)^r f(x). \tag{12}$$

Σύμφωνα με τις (11) και (12), η ροπή τάξης r και η κεντρική ροπή τάξης r προσδιορίζονται πλήρως από την πυκνότητα f, μπορούμε λοιπόν εξίσου καλά να μιλάμε για την ροπή τάξης r και την κεντρική ροπή τάξης r αυτής της πυκνότητας.

Υποθέτουμε ότι η X έχει ροπή τάξης r. Τότε η X έχει ροπή τάξης k για κάθε $k \leq r$. Για να το δούμε, παρατηρούμε ότι αν $|x| \leq 1$, τότε

$$|x^k| = |x|^k \leq 1$$

ενώ αν $|x| > 1$,

$$|x|^k \leq |x|^r.$$

Επομένως σε κάθε περίπτωση ισχύει ότι

$$|x|^k \leq |x|^r + 1.$$

Άρα, από το κριτήριο σύγκρισης για την σύγκλιση σειρών, βλέπουμε ότι

$$\sum_x |x|^k f(x) \leq \sum_x [|x|^r + 1] f(x) = \mathbb{E}(|X|^r) + 1 < \infty,$$

δηλαδή η X^k έχει πεπερασμένη μέση τιμή.

Από την άλλη πλευρά, όπως είδαμε στο Παράδειγμα 4, μία τυχαία μεταβλητή X μπορεί να μην έχει ούτε καν πρώτη ροπή. Με απλή τροποποίηση αυτού του παραδείγματος βλέπουμε ότι μία τυχαία μεταβλητή μπορεί να έχει ροπή τάξης r αλλά καμμία ροπή ανώτερης τάξης. (Δείτε την Ασκηση 9.)

Η πρώτη ροπή ($r = 1$) είναι απλώς ο μέσος της X. Γενικά, όσο περισσότερες ροπές της X γνωρίζουμε τόσο περισσότερες πληροφορίες έχουμε αποκτήσει για την κατανομή της X. Στις εφαρμογές όμως, μεγαλύτερο ενδιαφέρον παρουσιάζουν συνήθως οι δύο πρώτες ροπές.

Από την ιδιότητα (iii) του Θεωρήματος 2 ξέρουμε ότι αν οι X και Y έχουν και οι δύο πεπερασμένη μέση τιμή, τότε το ίδιο ισχύει και για την $X+Y$. Θα δείξουμε τώρα ότι αντίστοιχη ιδιότητα ισχύει και για τις ροπές τάξης r.

Θεώρημα 6. *Αν οι τυχαίες μεταβλητές X και Y έχουν ροπές τάξης r, τότε και η $X+Y$ έχει ροπή τάξης r.*

Απόδειξη: Το θεώρημα βασίζεται στην ακόλουθη απλή ανισότητα. Για κάθε μη αρνητικό ακέραιο $j \le r$,

$$|x|^j|y|^{r-j} \le |x|^r + |y|^r, \quad x, y \in \mathbb{R}. \tag{13}$$

Για να την δείξουμε, παρατηρούμε ότι αν $|x| \le |y|$, τότε $|x|^j|y|^{r-j} \le |y|^j|y|^{r-j}||y|^r \le |x|^r + |y|^r$, ενώ αν $|x| \ge |y|$, τότε $|x|^j|y|^r \le |x|^r \le |x|^r + |y|^r$. Άρα, η (13) ισχύει. Χρησιμοποιώντας την (13) και το διωνυμικό θεώρημα βλέπουμε τώρα ότι

$$\begin{aligned} |x+y|^r &\le (|x|+|y|)^r \\ &= \sum_{j=0}^{r} \binom{r}{j} |x|^j|y|^{r-j} \\ &\le \sum_{j=0}^{r} \binom{r}{j} (|x|^r + |y|^r). \end{aligned}$$

Όμως,

$$\sum_{j=0}^{r} \binom{r}{j} = 2^r$$

γιατί

$$2^r = (1+1)^r = \sum_{j=0}^{r} \binom{r}{j} 1^j 1^{r-j} = \sum_{j=0}^{r} \binom{r}{j}.$$

Επομένως,

$$|x+y|^r \le 2^r(|x|^r + |y|^r).$$

Έστω f η από κοινού πυκνότητα των X και Y. Τότε,

$$\begin{aligned} \sum_{x,y} |x+y|^r f(x,y) &\le 2^r \sum_{x,y} (|x|^r + |y|^r) f(x,y) \\ &= 2^r \mathbb{E}(|X|^r + |Y|^r) \\ &= 2^r (\mathbb{E}|X|^r + \mathbb{E}|Y|^r) < \infty. \end{aligned}$$

Άρα από το Θεώρημα 1, η $(X+Y)^r$ έχει πεπερασμένη μέση τιμή. □

Έπεται εύκολα με επαγωγή ότι αν οι $X_1, X_2, \ldots, X_n$ έχουν όλες ροπή τάξης r, τότε το ίδιο ισχύει και για την $X_1 + \ldots + X_n$.

Έστω X μία τυχαία μεταβλητή με πεπερασμένη δεύτερη ροπή. Τότε η *διασπορά της* X, η οποία συμβολίζεται με $\mathrm{Var}X$ ή $V(X)$, ορίζεται από την

$$\mathrm{Var}X = \mathbb{E}[(X - \mathbb{E}X)^2].$$

Αναπτύσσοντας το δεξιό μέλος βλέπουμε ότι

$$\begin{aligned}\mathrm{Var}X &= \mathbb{E}[X^2 - (2X)(\mathbb{E}X) + (\mathbb{E}X)^2] \\ &= \mathbb{E}X^2 - 2(\mathbb{E}X)^2 + (\mathbb{E}X)^2.\end{aligned}$$

Με άλλα λόγια,

$$\mathrm{Var}X = \mathbb{E}X^2 - (\mathbb{E}X)^2. \tag{14}$$

Συχνά συμβολίζουμε την $\mathbb{E}X$ με μ και την $\mathrm{Var}X$ με σ^2. Ο μη αρνητικός αριθμός $\sigma = \sqrt{\mathrm{Var}X}$ λέγεται *τυπική απόκλιση* της X ή της f_X.

Σύμφωνα με τη συζήτηση που προηγήθηκε, ο μέσος μ είναι η μέση τιμή της τυχαίας μεταβλητής X. Η διασπορά δίνει ένα μέτρο της διάχυσης της κατανομής της X γύρω από τον μέσο της. Όσο περισσότερο αποκλίνει η X από τη μέση τιμή της μ, τόσο μεγαλύτερη είναι η $(X-\mu)^2$, άρα και η διασπορά.

Από την άλλη πλευρά, έχουμε $\mathrm{Var}X = 0$ αν και μόνο αν η X είναι σταθερή. Για να το δείτε, παρατηρήστε ότι αν $P(X = c) = 1$ για κάποια σταθερά c, τότε $\mathbb{E}X = c$ και $\mathrm{Var}X = 0$. Αντίστροφα, αν $\mathrm{Var}X = 0$, τότε $\mathbb{E}[(X - \mathbb{E}X)^2] = 0$, επομένως $P((X - \mathbb{E}X)^2 = 0) = 1$. Άρα, $P(X = \mathbb{E}X) = 1$.

Μία εναλλακτική παρουσίαση του μέσου και της διασποράς δίνεται από το ακόλουθο πρόβλημα, που παρουσιάζει ενδιαφέρον στη Στατιστική. Έστω X μία τυχαία μεταβλητή με πεπερασμένη δεύτερη ροπή, και ας υποθέσουμε ότι θέλουμε να επιλέξουμε την τιμή του a που ελαχιστοποιεί την $\mathbb{E}(X-a)^2$. Αυτή η τιμή δίνει τη σταθερά που ταιριάζει καλύτερα στην X αν το σφάλμα μετριέται με βάση τη μέση τετραγωνική απόκλιση.

Ένας τρόπος για να λύσουμε αυτό το πρόβλημα είναι χρησιμοποιώντας απειροστικό λογισμό. Παρατηρήστε ότι

$$\mathbb{E}(X-a)^2 = \mathbb{E}X^2 - 2a\mathbb{E}X + a^2.$$

Αν παραγωγίσουμε ως προς a και ζητήσουμε η παράγωγος να ισούται με μηδέν, βλέπουμε ότι $a = \mathbb{E}X$. Αφού η δεύτερη παράγωγος είναι θετική (για την ακρίβεια, είναι ίση με 2), το σημείο αυτό αντιστοιχεί σε ελάχιστο, και η ελάχιστη τιμή είναι η $\mathrm{Var}X$.

Υπάρχει κι ένας δεύτερος τρόπος για να λύσουμε αυτό το πρόβλημα, η κατανόηση του οποίου έχει επίσης σημασία. Παρατηρούμε ότι

$$\begin{aligned}(X-a)^2 &= [(X-\mu) + (\mu - a)]^2 \\ &= (X-\mu)^2 + 2(X-\mu)(\mu-a) + (\mu-a)^2.\end{aligned}$$

Αφού $\mathbb{E}(X-\mu) = 0$, ο μεσαίος όρος έχει μηδενική μέση τιμή, άρα

$$\mathbb{E}(X-a)^2 = \mathbb{E}(X-\mu)^2 + (\mu-a)^2 = \mathrm{Var}X + (\mu-a)^2. \tag{15}$$

Είναι τώρα φανερό από την (15) ότι η $\mathbb{E}(X-a)^2$ ελαχιστοποιείται όταν $\mu = a$, και ότι η ελάχιστη τιμή είναι η $\mathrm{Var}X$.

Πολύ συχνά υπολογίζουμε τις ροπές μιάς τυχαίας μεταβλητής X με μη αρνητικές ακέραιες τιμές παραγωγίζοντας την γεννήτρια συνάρτηση πιθανότητας Φ_X. Ας υποθέσουμε για απλότητα ότι

$$\sum_{x=0}^{\infty} f_X(x) t_0^x < \infty$$

για κάποιο $t_0 > 1$. Μπορούμε τότε να ορίσουμε την Φ_X στο $-t_0 < t < t_0$ από την

$$\Phi_X(t) = \sum_{x=0}^{\infty} f_X(x) t^x \ , \quad -t_0 < t < t_0.$$

Μπορούμε να παραγωγίσουμε την $\Phi_X(t)$ οσεσδήποτε φορές παραγωγίζοντας τη δυναμοσειρά όρο προς όρο. Ειδικότερα,

$$\Phi'_X(t) = \sum_{x=1}^{\infty} x f_X(x) t^{x-1} \ , \quad -t_0 < t < t_0,$$

και

$$\Phi''_X(t) = \sum_{x=2}^{\infty} x(x-1) f_X(x) t^{x-2} \ , \quad -t_0 < t < t_0.$$

Από τις υποθέσεις μας για το t_0, μπορούμε να θέσουμε $t = 1$ στους παραπάνω τύπους, οπότε παίρνουμε

$$\Phi'_X(1) = \sum_{x=1}^{\infty} x f_X(x) = \mathbb{E}X$$

και

$$\Phi''_X(1) = \sum_{x=2}^{\infty} x(x-1) f_X(x) = \mathbb{E}X(X-1).$$

Έτσι, ο μέσος και η διασπορά της X προσδιορίζονται από την Φ_X μέσω των τύπων

$$\mathbb{E}X = \Phi'_X(1)$$

και

$$\mathrm{Var}X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \Phi''_X(1) + \Phi'_X(1) - (\Phi'_X(1))^2.$$

Ανάλογοι τύποι, που χρησιμοποιούν παραγώγους ανώτερης τάξης της $\Phi_X(t)$ στο $t = 1$, δίνουν τις ροπές ανώτερης τάξης της X.

Παρουσιάζουμε τώρα κάποιες εφαρμογές αυτών των τύπων στα παραδείγματα που ακολουθούν.

Παράδειγμα 9. *Αρνητική διωνυμική κατανομή.* Έστω X τυχαία μεταβλητή που ακολουθεί αρνητική διωνυμική κατανομή με παραμέτρους α και p. Βρείτε τον μέσο και τη διασπορά της X.

Από το Παράδειγμα 17 του Κεφαλαίου 3, ξέρουμε ότι η γεννήτρια συνάρτηση της X δίνεται από την $\Phi_X(t) = p^\alpha[1 - t(1-p)]^{-\alpha}$. Επομένως,

$$\Phi'_X(t) = \alpha p^\alpha [1 - t(1-p)]^{-(\alpha+1)}(1-p)$$

και

$$\Phi''_X(t) = (\alpha+1)\alpha p^\alpha [1 - t(1-p)]^{-(\alpha+2)}(1-p)^2.$$

Άρα

$$\Phi'_X(1) = \alpha\left(\frac{1-p}{p}\right)$$

και

$$\Phi''_X(1) = (\alpha+1)\alpha\left(\frac{1-p}{p}\right)^2.$$

Συμπεραίνουμε λοιπόν ότι $\mathbb{E}X = \alpha p(1-p)^{-1}$ και

$$\begin{aligned}
\mathrm{Var}X &= (\alpha+1)\alpha\left(\frac{1-p}{p}\right)^2 + \alpha\left(\frac{1-p}{p}\right) - \alpha^2\left(\frac{1-p}{p}\right)^2 \\
&= \alpha\frac{1-p}{p}.
\end{aligned}$$

Ειδικότερα, αν η X ακολουθεί γεωμετρική κατανομή με παράμετρο p, τότε $\mathbb{E}X = p^{-1}(1-p)$ (όπως έχουμε ήδη δεί) και $\mathrm{Var}X = p^{-2}(1-p)$.

Παράδειγμα 10. *Κατανομή Poisson.* Υποθέτουμε ότι η X ακολουθεί κατανομή Poisson με παράμετρο λ. Βρείτε το μέσο και τη διασπορά της X.

Στο Παράδειγμα 18 του Κεφαλαίου 3 είδαμε ότι $\Phi_X(t) = e^{\lambda(t-1)}$. Άρα

$$\Phi'_X(t) = \lambda e^{\lambda(t-1)}$$

και

$$\Phi''_X(t) = \lambda^2 e^{\lambda(t-1)}.$$

Έπεται ότι $\Phi'_X(1) = \lambda$ και $\Phi''_X(1) = \lambda^2$. Βλέπουμε αμέσως ότι

$$\mathbb{E}X = \lambda,$$

ακριβώς όπως και στο Παράδειγμα 2, ενώ

$$\mathrm{Var}X = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Δηλαδή, αν η X ακολουθεί κατανομή Poisson με παράμετρο λ, τότε ο μέσος και η διασπορά της X ισούνται και οι δύο με λ.

4.4 Διασπορά αθροίσματος

Έστω X και Y τυχαίες μεταβλητές με πεπερασμένη δεύτερη ροπή. Τότε η $X+Y$ έχει πεπερασμένη δεύτερη ροπή, άρα και πεπερασμένη διασπορά. Έχουμε

$$\begin{aligned}
\mathrm{Var}(X+Y) &= \mathbb{E}[(X+Y)-\mathbb{E}(X+Y)]^2 \\
&= \mathbb{E}[(X-\mathbb{E}X)+(Y-\mathbb{E}Y)]^2 \\
&= \mathbb{E}(X-\mathbb{E}X)^2+\mathbb{E}(Y-\mathbb{E}Y)^2+2\mathbb{E}[(X-\mathbb{E}X)(Y-\mathbb{E}Y)] \\
&= \mathrm{Var}X+\mathrm{Var}Y+2\mathbb{E}[(X-\mathbb{E}X)(Y-\mathbb{E}Y)]\,.
\end{aligned}$$

Βλέπουμε ότι, σε αντίθεση με το μέσο, η διασπορά του αθροίσματος δύο τυχαίων μεταβλητών δεν είναι, γενικά, το άθροισμα των διασπορών τους. Η ποσότητα

$$\mathbb{E}[(X-\mathbb{E}X)(Y-\mathbb{E}Y)]$$

λέγεται *συνδιακύμανση* των X και Y και συμβολίζεται με $\mathrm{Cov}(X,Y)$. Με αυτόν τον συμβολισμό παίρνουμε το σημαντικό τύπο

$$\mathrm{Var}(X+Y)=\mathrm{Var}X+\mathrm{Var}Y+2\mathrm{Cov}(X,Y). \tag{16}$$

Τώρα

$$(X-\mathbb{E}X)(Y-\mathbb{E}Y)=XY-Y\mathbb{E}X-X\mathbb{E}Y+(\mathbb{E}X)(\mathbb{E}Y),$$

και παίρνοντας μέσες τιμές βλέπουμε ότι

$$\mathrm{Cov}(X,Y)=\mathbb{E}[(X-\mathbb{E}X)(Y-\mathbb{E}Y)]=\mathbb{E}(XY)-(\mathbb{E}X)(\mathbb{E}Y). \tag{17}$$

Από την σχέση αυτή γίνεται φανερό ότι $\mathrm{Cov}(X,Y)=0$ αν οι X και Y είναι ανεξάρτητες. (Το Παράδειγμα 5 δείχνει ότι το αντίστροφο δεν ισχύει.) Από την (16) βλέπουμε ότι αν οι X και Y είναι ανεξάρτητες τυχαίες μεταβλητές με πεπερασμένες δεύτερες ροπές, τότε $\mathrm{Var}(X,Y)=\mathrm{Var}X+\mathrm{Var}Y$.

Ειδικότερα, αν $P(Y=c)=1$ για κάποια σταθερά c, τότε οι X και Y είναι ανεξάρτητες και η διασπορά της Y ισούται με μηδέν. Επομένως,

$$\mathrm{Var}(X+c)=\mathrm{Var}X+\mathrm{Var}(c)=\mathrm{Var}X. \tag{18}$$

Γενικότερα, αν $X_1,\ldots,X_n$ είναι n τυχαίες μεταβλητές με πεπερασμένη δεύτερη ροπή, τότε

$$\mathrm{Var}\left(\sum_{i=1}^n X_i\right)=\sum_{i=1}^n \mathrm{Var}X_i+2\sum_{i=1}^{n-1}\sum_{j=i+1}^n \mathrm{Cov}(X_i,X_j), \tag{19}$$

και, αν οι $X_1,\ldots,X_n$ είναι ανεξάρτητες, τότε

$$\mathrm{Var}\left(\sum_{i=1}^n X_i\right)=\sum_{i=1}^n \mathrm{Var}X_i. \tag{20}$$

Μπορείτε να αποδείξετε αυτούς τους τύπους με απευθείας υπολογισμό, ανάλογο (αλλά πιό πολύπλοκο) αυτού που χρησιμοποιήσαμε για την περίπτωση $n = 2$, ή κάνοντας επαγωγή ως προς n βασιζόμενοι στην περίπτωση $n = 2$.

Ειδικότερα, αν $X_1, \ldots, X_n$ είναι ανεξάρτητες τυχαίες μεταβλητές με την ίδια διασπορά σ^2 (για παράδειγμα, αν έχουν όλες την ίδια πυκνότητα), τότε

$$\text{Var}(X_1 + \ldots + X_n) = n\text{Var}X_1 = n\sigma^2. \tag{21}$$

Μία άλλη στοιχειώδης, αλλά χρήσιμη, παρατήρηση είναι ότι $\text{Var}(aX) = a^2\text{Var}X$. Αφήνουμε την επαλήθευση αυτού του ισχυρισμού στον αναγνώστη.

Παράδειγμα 11. *Διωνυμική κατανομή.* Έστω $X_1, \ldots, X_n$ ανεξάρτητες τυχαίες μεταβλητές Bernoulli που παίρνουν την τιμή 1 με την ίδια πιθανότητα p. Τότε (δείτε το Παράδειγμα 6) το άθροισμα $S_n = X_1 + \ldots + X_n$ ακολουθεί τη διωνυμική κατανομή με παραμέτρους n και p. Εχουμε ήδη αποδείξει ότι $\mathbb{E}S_n = np$. Χρησιμοποιώντας την (21) βλέπουμε αμέσως ότι

$$\text{Var}S_n = n\text{Var}X_1.$$

Όμως $X_1^2 = X_1$ γιατί η X_1 παίρνει μόνο τις τιμές 0 και 1. Επομένως $\mathbb{E}X_1^2 = \mathbb{E}X_1 = p$, δηλαδή

$$\text{Var}X_1 = \mathbb{E}X_1^2 - (\mathbb{E}X_1)^2 = p - p^2 = p(1-p).$$

Έπεται ότι $\text{Var}S_n = np(1-p)$.

Συνοψίζοντας, ο μέσος μιάς διωνυμικά κατανεμημένης τυχαίας μεταβλητής είναι np και η διασπορά της $np(1-p)$.

Παράδειγμα 12. *Υπεργεωμετρική κατανομή.* Θεωρούμε την κατάσταση του Παραδείγματος 7. Θέλουμε τώρα να υπολογίσουμε την $\text{Var}S_n$, δηλαδή τη διασπορά μιάς υπεργεωμετρικής κατανομής. Για το σκοπό αυτό χρησιμοποιούμε την (19).

Είδαμε ότι οι εξαρτημένες δείκτριες $X_1, \ldots, X_n$ ικανοποιούν την

$$P(X_i = 1) = \mathbb{E}X_i = \frac{r_1}{r}.$$

Αφού $X_i^2 = X_i$ βλέπουμε ότι

$$\begin{aligned} \text{Var}X_i &= \mathbb{E}X_i^2 - (\mathbb{E}X_i)^2 = \left(\frac{r_1}{r}\right) - \left(\frac{r_1}{r}\right)^2 \\ &= \left(\frac{r_1}{r}\right)\left(1 - \frac{r_1}{r}\right). \end{aligned}$$

Πρέπει τώρα να υπολογίσουμε τις συνδιακυμάνσεις. Υποθέτουμε ότι $1 \le i < j \le n$. Εχουμε $X_iX_j = 0$ εκτός αν οι X_i και X_j παίρνουν και οι δύο την τιμή 1, άρα

$$\mathbb{E}X_iX_j = P(X_i = 1, X_j = 1) = \left(\frac{r_1}{r}\right)\left(\frac{r_1 - 1}{r - 1}\right).$$

Κάνοντας πράξεις παίρνουμε

$$\begin{aligned}
\mathrm{Cov}(X_i, X_j) &= \mathbb{E}(X_i X_j) - (\mathbb{E}X_i)(\mathbb{E}X_j) \\
&= \frac{r_1(r_1-1)}{r(r_1)} - \left(\frac{r_1}{r}\right)^2 \\
&= \left(\frac{r_1}{r}\right)\left(\frac{r_1-1}{r-1} - \frac{r_1}{r}\right) \\
&= \left(\frac{r_1}{r}\right)\frac{r_1-r}{r(r-1)},
\end{aligned}$$

οπότε

$$\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \mathrm{Cov}(X_i, X_j) = \frac{n(n-1)}{2}\left(\frac{r_1}{r}\right)\frac{r_1-r}{r(r-1)}.$$

Χρησιμοποιώντας και την (19) βλέπουμε ότι

$$\begin{aligned}
\mathrm{Var}S_n &= n\frac{r_1(r-r_1)}{r^2} - n(n-1)\frac{r_1(r-r_1)}{r^2(r-1)} \\
&= n\left(\frac{r_1}{r}\right)\left(1-\frac{r_1}{r}\right)\left(1-\frac{n-1}{r-1}\right).
\end{aligned}$$

Παρουσιάζει ενδιαφέρον η σύγκριση του μέσου και της διασποράς της υπεργεωμετρικής κατανομής με τις αντίστοιχες τιμές για την διωνυμική κατανομή που έχει την ίδια πιθανότητα επιτυχίας $p = (r_1/r)$. Ας υποθέσουμε ότι έχουμε έναν πληθυσμό r ατόμων, r_1 από τα οποία είναι τύπου I και $r - r_1$ είναι τύπου II. Ενα τυχαίο δείγμα μεγέθους n επιλέγεται από τον πληθυσμό. Συμβολίζουμε με Y το πλήθος των ατόμων τύπου I στο δείγμα.

Αν η δειγματοληψία γίνεται *με επανατοποθέτηση*, τότε η Y ακολουθεί διωνυμική κατανομή με παραμέτρους n και $p = (r_1/r)$, επομένως

$$\mathbb{E}Y = n\left(\frac{r_1}{r}\right) \quad \textit{και} \quad \mathrm{Var}Y = n\left(\frac{r_1}{r}\right)\left(1-\frac{r_1}{r}\right).$$

Από την άλλη πλευρά, αν η δειγματοληψία γίνεται *χωρίς επανατοποθέτηση*, τότε η Y ακολουθεί υπεργεωμετρική κατανομή, με

$$\mathbb{E}Y = n\left(\frac{r_1}{r}\right) \quad \textit{και} \quad \mathrm{Var}Y = n\left(\frac{r_1}{r}\right)\left(1-\frac{r_1}{r}\right)\left(1-\frac{n-1}{r-1}\right).$$

Ο μέσος είναι ο ίδιος και στις δύο περιπτώσεις, η διασπορά όμως είναι μικρότερη όταν η δειγματοληψία γίνεται χωρίς επανατοποθέτηση. Διαισθητικά, όσο πιό κοντά βρίσκονται τα n και r τόσο πιό ντετερμινιστική είναι η Y για τη δειγματοληψία χωρίς επανατοποθέτηση. Πράγματι, αν $n = r$ τότε η διασπορά είναι μηδέν και $P(Y = r_1) = 1$. Αν όμως το r είναι μεγάλο σε σχέση με το n, οπότε ο λόγος (n/r) είναι κοντά στο μηδέν, τότε ο λόγος των διασπορών που παίρνουμε για τις δειγματοληψίες με και χωρίς επανατοποθέτηση είναι κοντά στο ένα. Αυτό είναι απολύτως φυσιολογικό, γιατί αν κρατήσουμε το n σταθερό και πάρουμε μεγάλο r υπάρχει μικρή διαφορά ανάμεσα στη δειγματοληψία με επανατοποθέτηση και τη δειγματοληψία χωρίς επανατοποθέτηση.

4.5 Συντελεστής συσχέτισης

Έστω X και Y δύο τυχαίες μεταβλητές με πεπερασμένες μη μηδενικές διασπορές. Ο συντελεστής συσχέτισης $\rho(X,Y)$ που ορίζεται από την

$$\rho = \rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{(\mathrm{Var}X)(\mathrm{Var}Y)}}, \tag{22}$$

μάς δίνει ένα μέτρο για τον βαθμό εξάρτησης ανάμεσα στις δύο τυχαίες μεταβλητές. Λέμε ότι οι τυχαίες μεταβλητές είναι ασυσχέτιστες αν $\rho = 0$. Αφού $\mathrm{Cov}(X,Y) = 0$ όταν οι X και Y είναι ανεξάρτητες, βλέπουμε αμέσως ότι δύο ανεξάρτητες τυχαίες μεταβλητές είναι πάντα ασυσχέτιστες. Είναι δυνατόν δύο εξαρτημένες τυχαίες μεταβλητές να είναι ασυσχέτιστες, κάτι που μπορείτε να διαπιστώσετε στο Παράδειγμα 5.

Για τις εφαρμογές στη Στατιστική είναι σημαντικό να ξέρουμε ότι ο συντελεστής ρ παίρνει πάντα τιμές μεταξύ -1 και 1, και ότι $|\rho| = 1$ αν και μόνο αν $P(X = aY) = 1$ για κάποια σταθερά a. Τα παραπάνω είναι απλές συνέπειες της ακόλουθης βασικής ανισότητας, που λέγεται *ανισότητα του Schwarz.*

Θεώρημα 7. *Η ανισότητα του Schwarz. Υποθέτουμε ότι οι X και Y έχουν πεπερασμένες δεύτερες ροπές. Τότε*

$$[\mathbb{E}(XY)]^2 \leq (\mathbb{E}X^2)(\mathbb{E}Y^2). \tag{23}$$

Επιπλέον, στην (23) ισχύει ισότητα αν και μόνο αν είτε $P(Y = 0) = 1$ ή $P(X = aY) = 1$ για κάποια σταθερά a.

Απόδειξη: Αν $P(Y=0) = 1$, τότε $P(XY = 0) = 1$, $\mathbb{E}(XY) = 0$, και $\mathbb{E}Y^2 = 0$. Έτσι, σε αυτήν την περίπτωση η (23) ισχύει σαν ισότητα. Επίσης, αν $P(X = aY) = 1$, ένας απλός υπολογισμός δείχνει ότι και τα δύο μέλη της (23) είναι ίσα με $(a^2\mathbb{E}Y^2)^2$.

Δείχνουμε τώρα ότι η (23) ισχύει πάντα. Σύμφωνα με τη συζήτηση που προηγήθηκε, μπορούμε να υποθέσουμε ότι $P(Y = 0) < 1$, άρα $\mathbb{E}Y^2 > 0$. Η απόδειξη θα βασιστεί σε ένα απλό αλλά έξυπνο τέχνασμα. Παρατηρούμε ότι για κάθε πραγματικό αριθμό λ

$$0 \leq \mathbb{E}(X - \lambda Y)^2 = \lambda^2\mathbb{E}Y^2 - 2\lambda\mathbb{E}(XY) + \mathbb{E}X^2.$$

Αυτή είναι μία τετραγωνική συνάρτηση του λ. Αφού ο συντελεστής $\mathbb{E}Y^2$ του λ^2 είναι θετικός, το ελάχιστό της πιάνεται για κάποια τιμή του λ, ας πούμε $\lambda = a$, η οποία υπολογίζεται αν θέσουμε την παράγωγο ίση με μηδέν και επιλύσουμε. Η απάντηση είναι $a = [\mathbb{E}(XY)][\mathbb{E}Y^2]^{-1}$. Η αντίστοιχη τιμή της συνάρτησης είναι

$$0 \leq \mathbb{E}(X - aY)^2 = \mathbb{E}X^2 - \frac{[\mathbb{E}(XY)]^2}{\mathbb{E}Y^2}, \tag{24}$$

και αυτό αποδεικνύει την (23). Αν έχουμε ισότητα στην ανισότητα του Schwarz (23), τότε από την (24) βλέπουμε ότι $\mathbb{E}(X - aY)^2 = 0$, άρα

$$P[(X - aY) = 0] = 1.$$

Αυτό ολοκληρώνει την απόδειξη. □

Εφαρμόζοντας την ανισότητα του Schwarz στις τυχαίες μεταβλητές $(X - \mathbb{E}X)$ και $(Y - \mathbb{E}Y)$ βλέπουμε ότι

$$(\mathbb{E}[(X - \mathbb{E}X)(Y - \mathbb{E}Y)])^2 \leq [\mathbb{E}(X - \mathbb{E}X)^2][\mathbb{E}(Y - \mathbb{E}Y)^2],$$

δηλαδή,

$$[\mathrm{Cov}(X, Y)]^2 \leq (\mathrm{Var}X)(\mathrm{Var}Y).$$

Σύμφωνα με τον ορισμό του ρ

$$|\rho(X, Y)| \leq 1.$$

Βλέπουμε ακόμα από το Θεώρημα 7 ότι $|\rho| = 1$ αν και μόνο αν $P(X = aY) = 1$ για κάποια σταθερά a.

Ο συντελεστής συσχέτισης έχει περιορισμένες εφαρμογές στην θεωρία πιθανοτήτων. Εμφανίζεται κυρίως στη Στατιστική.

4.6 Η ανισότητα του Chebyshev

Έστω X μία μη αρνητική τυχαία μεταβλητή με πεπερασμένη μέση τιμή, και t ένας θετικός πραγματικός αριθμός. Ορίζουμε μία τυχαία μεταβλητή Y θέτοντας $Y = 0$ αν $X < t$ και $Y = t$ αν $X \geq t$. Η Y είναι διακριτή τυχαία μεταβλητή, και παίρνει τις τιμές 0 και t με πιθανότητες $P(Y = 0) = P(X < t)$ και $P(Y = t) = P(X \geq t)$ αντίστοιχα. Άρα,

$$\mathbb{E}Y = tP(Y = t) + 0.P(Y = 0) = tP(Y = t) = tP(X \geq t).$$

Είναι φανερό ότι $X \geq Y$, επομένως $\mathbb{E}X \geq \mathbb{E}Y$. Άρα,

$$\mathbb{E}X \geq \mathbb{E}Y = tP(X \geq t)$$

ή

$$P(X \geq t) \leq \frac{\mathbb{E}X}{t}. \tag{25}$$

Πολλές χρήσιμες ανισότητες προκύπτουν από την (25). Η πιό σημαντική από αυτές είναι η ανισότητα του Chebyshev.

Ανισότητα του Chebyshev. Έστω X μία τυχαία μεταβλητή με μέσο μ και πεπερασμένη διασπορά σ^2. Τότε για κάθε πραγματικό αριθμό $t > 0$

$$P(|X-\mu| \geq t) \leq \frac{\sigma^2}{t^2}. \tag{26}$$

Για να αποδείξουμε την (26), εφαρμόζουμε την (25) για τη μη αρνητική τυχαία μεταβλητή $(X-\mu)^2$ και τον αριθμό t^2. Συμπεραίνουμε ότι

$$P((X-\mu)^2 \geq t^2) \leq \frac{\mathbb{E}(X-\mu)^2}{t^2} = \frac{\sigma^2}{t^2}.$$

Αφού $(X-\mu)^2 \geq t^2$ αν και μόνο αν $|X-\mu| \geq t$, βλέπουμε ότι ισχύει η (26).

Η ανισότητα του Chebyshev δίνει ένα φράγμα, συναρτήσει των $\mathrm{Var}X$ και t, για την πιθανότητα η X να αποκλίνει από το μέσο της περισσότερο από t μονάδες. Η ισχύς της βρίσκεται στη μεγάλη της γενικότητα. Δεν υποθέτει τίποτε άλλο για την κατανομή της X παρά μόνο το ότι έχει πεπερασμένη διασπορά. Η ανισότητα αυτή είναι το σημείο αφετηρίας για πολλά θεωρητικά αποτελέσματα. Για τις περισσότερες κατανομές που εμφανίζονται στην πράξη, υπάρχουν πολύ ακριβέστερα φράγματα για την $P(|X-\mu| \geq t)$ από αυτό που δίνει η ανισότητα του Chebyshev. Υπάρχουν όμως παραδείγματα που δείχνουν ότι, σε πλήρη γενικότητα, το φράγμα που δίνει η ανισότητα του Chebyshev δεν βελτιώνεται (δείτε την Ασκηση 26).

Έστω $X_1, \ldots, X_n$ ανεξάρτητες τυχαίες μεταβλητές που έχουν την ίδια κατανομή. Μπορείτε να σκέφτεστε αυτές τις τυχαίες μεταβλητές σαν n ανεξάρτητες μετρήσεις κάποιας ποσότητας που ακολουθεί την κοινή τους κατανομή. Με αυτήν την έννοια, συχνά λέμε ότι οι τυχαίες μεταβλητές $X_1, \ldots, X_n$ συνιστούν ένα τυχαίο δείγμα μεγέθους n από αυτήν την κατανομή.

Υποθέτουμε ότι η κοινή κατανομή αυτών των τυχαίων μεταβλητών έχει πεπερασμένο μέσο μ. Τότε, για αρκετά μεγάλο n θα περιμέναμε ο αριθμητικός τους μέσος $S_n/n = (X_1 + \ldots + X_n)/n$ να βρίσκεται κοντά στον μ. Αν υποθέσουμε ότι οι X_i έχουν και πεπερασμένη διασπορά, τότε

$$\mathrm{Var}\left(\frac{S_n}{n}\right) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

δηλαδή $\mathrm{Var}(S_n/n) \to 0$ καθώς $n \to \infty$. Όπως είδαμε στην Παράγραφο 4.3, αυτό σημαίνει ότι καθώς το n μεγαλώνει η κατανομή της S_n/n συγκεντρώνεται όλο και περισσότερο γύρω από το μέσο της μ. Ακριβέστερα, εφαρμόζοντας την ανισότητα του Chebyshev για την S_n/n παίρνουμε την ανισότητα

$$P\left(|\frac{S_n}{n} - \mu| \geq \delta\right) \leq \frac{\mathrm{Var}(S_n/n)}{\delta^2} = \frac{\sigma^2}{n\delta^2}. \tag{27}$$

Ειδικότερα, από την (27) έπεται ότι για κάθε $\delta > 0$

$$\lim_{n\to\infty} P\left(|\frac{S_n}{n} - \mu| \geq \delta\right) = 0. \tag{28}$$

Μπορούμε να ερμηνεύσουμε την (28) με τον εξής τρόπο. Ο αριθμός δ δίνει την ακρίβεια που ζητάμε στην προσέγγιση του μ από την S_n/n. Η (28) μας βεβαιώνει ότι, όσο μικρό κι αν επιλέξουμε το δ, η πιθανότητα η S_n/n να προσεγγίζει τον μ με αυτήν την ακρίβεια, δηλαδή η $P(|(S_n/n) - \mu| < \delta)$, συγκλίνει στο 1 καθώς το πλήθος των παρατηρήσεων μεγαλώνει. Το γεγονός αυτό λέγεται Ασθενής Νόμος των Μεγάλων Αριθμών. Έχουμε αποδείξει τον ασθενή νόμο με μόνη υπόθεση για τις X_i να έχουν πεπερασμένη διασπορά. Για την ακρίβεια, η υπόθεση αυτή δεν είναι απαραίτητη. Το μόνο που χρειάζεται είναι οι X_i να έχουν πεπερασμένο μέσο. Διατυπώνουμε αυτό το γενικότερο αποτέλεσμα στο θεώρημα που ακολουθεί. Η απόδειξη θα δοθεί στο Κεφάλαιο 8.

Θεώρημα 8. *Ασθενής Νόμος των Μεγάλων Αριθμών. Έστω $X_1, X_2, \ldots, X_n$ ανεξάρτητες τυχαίες μεταβλητές που έχουν κοινή κατανομή με μέσο μ. Θέτουμε $S_n = X_1 + \ldots + X_n$. Τότε για κάθε $\delta > 0$*

$$\lim_{n\to\infty} P\left(\left|\frac{S_n}{n} - \mu\right| \geq \delta\right) = 0.$$

Αν οι X_i έχουν πεπερασμένο μέσο, τότε ο ασθενής νόμος ισχύει. Αν όμως έχουν και πεπερασμένη διασπορά, τότε ισχύει η (27). Αυτή είναι πολύ ακριβέστερη πρόταση, γιατί μας δίνει άνω φράγμα για την $P\left(\left|\frac{S_n}{n} - \mu\right| \geq \delta\right)$ μέσω του n. Θα δώσουμε ένα παράδειγμα εφαρμογής της (27) σε διωνυμικά κατανεμημένες τυχαίες μεταβλητές.

Έστω $X_1, X_2, \ldots, X_n$ ανεξάρτητες τυχαίες μεταβλητές Bernoulli που παίρνουν την τιμή 1 με την ίδια πιθανότητα p. Τότε $\mu = p$ και $\sigma^2 = p(1-p)$. Έτσι, η (27) δείχνει ότι

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \delta\right) \leq \frac{p(1-p)}{n\delta^2}. \tag{29}$$

Αφού $p(1-p) \leq 1/4$ αν $0 < p < 1$ (με στοιχειώδη απειροστικό λογισμό βλέπουμε ότι η $p(1-p)$ παίρνει την μέγιστη τιμή της όταν $p = 1/2$), έπεται ότι ανεξάρτητα από την τιμή του p,

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \delta\right) \leq \frac{1}{4n\delta^2}. \tag{30}$$

Η (29) είναι χρήσιμη όταν ξέρουμε την τιμή του p, ενώ η (30) μας δίνει φράγμα για την $P\left(\left|\frac{S_n}{n} - \mu\right| \geq \delta\right)$ που ισχύει για κάθε τιμή του p. Αν το p είναι κοντά στο $1/2$, τότε οι (29) και (30) δεν διαφέρουν πολύ, αν όμως το p είναι μακριά από το $1/2$ τότε η εκτίμηση που δίνει η (29) μπορεί να είναι πολύ καλύτερη. (Για την ακρίβεια, ακόμα και τα φράγματα που δίνει η (29) είναι πολύ φτωχά. Στο Κεφάλαιο 7 θα συζητήσουμε μία άλλη μέθοδο που δίνει πολύ καλύτερες εκτιμήσεις.)

Έστω ότι μάς δίνουν δ και $\varepsilon > 0$. Μπορούμε να χρησιμοποιήσουμε τις (29) και (30) για να βρούμε ένα κάτω φράγμα για το πλήθος των δοκιμών που αρκούν για να μάς εξασφαλίσουν ότι

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \delta\right) \leq \varepsilon.$$

Πράγματι, από την (29) βλέπουμε ότι αυτό ισχύει αν $p(1-p)/n\delta^2 \leq \varepsilon$. Λύνοντας ως προς n βρίσκουμε ότι $n \geq p(1-p)/\varepsilon\delta^2$. Αν χρησιμοποιήσουμε την (30), τότε βλέπουμε ότι $n \geq (4\varepsilon\delta^2)^{-1}$ δοκιμές αρκούν. Επαναλαμβάνουμε ότι αυτά τα φράγματα που δίνει η ανισότητα του Chebyshev για το n είναι φτωχά, και ότι ένας πολύ μικρότερος αριθμός δοκιμών μπορεί να αρκεί.

Για να δώσουμε ένα παράδειγμα της διαφοράς ανάμεσα στις δύο εκτιμήσεις για το n, επιλέγουμε $\delta = 0.1$ και $\varepsilon = 0.01$. Τότε $\delta^2\varepsilon = 10^{-4}$ και από την (30) βλέπουμε ότι για να εξασφαλίσουμε την

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq 0.1\right) \leq 0.01$$

θα χρειαζόμαστε $n = 10^4/4 = 2500$ παρατηρήσεις. Ας υποθέσουμε όμως ότι ξέραμε πως $p = 0.1$. Τότε, αφού $p(1-p) = 0.09$ βλέπουμε από την (29) ότι $n \geq 0.09 \times 10^4 = 900$ παρατηρήσεις αρκούν. Αν $p = 1/2$, τότε η (29) δίνει την ίδια εκτίμηση με την (30), δηλαδή 2500.

Για να δείξουμε ότι τα φράγματα της ανισότητας του Chebyshev είναι πραγματικά φτωχά στην περίπτωση της διωνυμικής κατανομής, υποθέτουμε ότι $n = 100$ και $p = 1/2$. Από την (29) παίρνουμε

$$P\left(\left|\frac{S_n}{n} - 0.5\right| \geq 0.1\right) \leq 0.25.$$

Συγκρίνετε με την ακριβή τιμή αυτής της πιθανότητας που είναι 0.038.

4.7 Ασκήσεις

1 Έστω N θετικός ακέραιος, και f η συνάρτηση που ορίζεται από την

$$f(x) = \begin{cases} \frac{2x}{N(N+1)} & , x = 1, 2, \ldots, N, \\ 0 & , \text{αλλιώς}. \end{cases}$$

Δείξτε ότι η f είναι διακριτή πυκνότητα και βρείτε τον μέσο της. *Υπόδειξη:*

$$\sum_{x=1}^{N} x = \frac{N(N+1)}{2} \quad , \quad \sum_{x=1}^{N} x^2 = \frac{N(N+1)(2N+1)}{6}.$$

2 Υποθέτουμε ότι η X έχει διωνυμική πυκνότητα με παραμέτρους $n = 4$ και p. Βρείτε την $\mathbb{E}[\sin(\pi X/2)]$.

3 Υποθέτουμε ότι η X είναι Poisson με παράμετρο λ. Υπολογίστε το μέσο της $(1+X)^{-1}$.

4 Αν η X έχει μέσο 1 και η Y έχει μέσο 3, ποιός είναι ο μέσος της $2X + 5Y$;

5 Υποθέτουμε ότι X και Y είναι δύο τυχαίες μεταβλητές με

$$P(|X - Y| \leq M) = 1$$

για κάποια σταθερά M. Δείξτε ότι αν η Y έχει πεπερασμένη μέση τιμή, τότε η X έχει πεπερασμένη μέση τιμή και $|\mathbb{E}X - \mathbb{E}Y| \leq M$.

6 Έστω X γεωμετρικά κατανεμημένη τυχαία μεταβλητή και $M > 0$ ακέραιος. Θέτουμε $Z = \min(X, M)$. Υπολογίστε το μέσο της Z. *Υπόδειξη:* Χρησιμοποιήστε το Θεώρημα 5.

7 Έστω X γεωμετρικά κατανεμημένη τυχαία μεταβλητή και $M > 0$ ακέραιος. Θέτουμε $Y = \max(X, M)$. Υπολογίστε το μέσο της Y. *Υπόδειξη:* Υπολογίστε την $P(Y < y)$ και στη συνέχεια χρησιμοποιήστε το Θεώρημα 5.

8 Υποθέτουμε ότι η X είναι ομοιόμορφα κατανεμημένη στο $\{0, 1, \ldots, N\}$. Βρείτε το μέσο και τη διασπορά της X χρησιμοποιώντας την υπόδειξη της Άσκησης 1.

9 Κατασκευάστε παράδειγμα πυκνότητας που έχει πεπερασμένη ροπή τάξης r αλλά δεν έχει ροπές υψηλότερης τάξης. *Υπόδειξη:* Θεωρήστε την σειρά $\sum_{x=1}^{\infty} x^{-(r+2)}$ και, χρησιμοποιώντας την, ορίστε κατάλληλη πυκνότητα.

10 Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές με $\mathbb{E}X^4 = 2$, $\mathbb{E}Y^2 = 1$, $\mathbb{E}X^2 = 1$, και $\mathbb{E}Y = 0$. Υπολογίστε την $\mathrm{Var}(X^2 Y)$.

11 Δείξτε ότι $\mathrm{Var}(aX) = a^2 \mathrm{Var} X$.

12 Έστω ότι η X ακολουθεί διωνυμική κατανομή με παραμέτρους n και p. Χρησιμοποιήστε τη γεννήτρια συνάρτηση πιθανότητας της X για να υπολογίσετε τη μέση τιμή και τη διασπορά της.

13 Έστω X μη αρνητική τυχαία μεταβλητή με ακέραιες τιμές.
(α) Δείξτε ότι

$$\Phi_X(t) = \mathbb{E}[t^X], \qquad -1 \leq t \leq 1,$$
$$\Phi'_X(t) = \mathbb{E}[X t^{X-1}], \qquad -1 < t < 1,$$

και

$$\Phi''_X(t) = \mathbb{E}[X(X-1)t^{X-2}], \qquad -1 < t < 1.$$

(β) Χρησιμοποιώντας το Θεώρημα 4 δώστε δεύτερη απόδειξη του ότι, αν X και Y είναι ανεξάρτητες μη αρνητικές τυχαίες μεταβλητές με ακέραιες τιμές, τότε

$$\Phi_{X+Y}(t) = \Phi_X(t)\Phi_Y(t), \qquad -1 \leq t \leq 1.$$

14 Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές με πεπερασμένες δεύτερες ροπές. Υπολογίστε τη μέση τιμή και τη διασπορά της $2X + 3Y$ με βάση αυτές των X και Y.

15 Έστω $X_1, \ldots, X_n$ ανεξάρτητες τυχαίες μεταβλητές που έχουν την ίδια πυκνότητα, μέση τιμή μ, και διασπορά σ^2. Θέτουμε $\overline{X} = (X_1 + \ldots + X_n)/n$.

(α) Γράφοντας $X_k - \overline{X} = (X_k - \mu) - (\overline{X} - \mu)$, δείξτε ότι

$$\sum_{k=1}^{n} (X_k - \overline{X})^2 = \sum_{k=1}^{n} (X_k - \mu)^2 - n(\overline{X} - \mu)^2.$$

(β) Από το (α) συμπεράνατε ότι

$$\mathbb{E}\left(\sum_{k=1}^{n} (X_k - \overline{X})^2 \right) = (n-1)\sigma^2.$$

16 Υποθέτουμε ότι n βώλοι κατανέμονται τυχαία σε r δοχεία. Θέτουμε $X_i = 1$ αν το δοχείο i είναι άδειο, και $X_i = 0$ αλλιώς.
(α) Υπολογίστε την $\mathbb{E}X_i$.
(β) Αν $i \neq j$, υπολογίστε την $\mathbb{E}(X_i X_j)$.
(γ) Έστω S_r το πλήθος των άδειων δοχείων. Γράψτε $S_r = X_1 + \ldots + X_r$, και χρησιμοποιήστε το (α) για να υπολογίσετε την $\mathbb{E}S_r$.
(δ) Χρησιμοποιώντας τα (α) και (β) υπολογίστε την $\mathrm{Var}S_r$.

17 Έστω ότι έχουμε δύο δεσμίδες των n χαρτιών, κάθε μία αριθμημένη από 1 ως n. Ανακατεύουμε τις δύο δεσμίδες και αντιστοιχίζουμε ένα προς ένα τα χαρτιά τους. Λέμε ότι έχουμε σύμπτωση στην θέση i αν το i-στό χαρτί κάθε δεσμίδας έχει τον ίδιο αριθμό. Έστω S_n το πλήθος των συμπτώσεων.
(α) Υπολογίστε την $\mathbb{E}S_n$.
(β) Υπολογίστε την $\mathrm{Var}S_n$.
Υπόδειξη: Θέτουμε $X_i = 1$ αν έχουμε σύμπτωση στην θέση i, και $X_i = 0$ αλλιώς. Τότε $S_n = X_1 + \ldots + X_n$. Από τα αποτελέσματα του Κεφαλαίου 2 γνωρίζουμε ότι $P(X_i = 1) = 1/n$ και ότι αν $i \neq j$,

$$P(X_i = 1, X_j = 1) = \frac{1}{n(n-1)}.$$

18 Θεωρούμε την τυχαία μεταβλητή S_k που συζητήσαμε στο Παράδειγμα 8. Υπολογίστε την $\mathrm{Var}S_k$.

19 Αποδείξτε τις ακόλουθες ιδιότητες της συνδιακύμανσης:

(α) $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$.

(β) $\mathrm{Cov}\left(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{n} b_j Y_j \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i b_j \mathrm{Cov}(X_i, Y_j)$.

20 Έστω X_1, X_2, και X_3 ανεξάρτητες τυχαίες μεταβλητές που έχουν πεπερασμένες θετικές διασπορές σ_1^2, σ_2^2, και σ_3^2 αντίστοιχα. Βρείτε τον συντελεστή συσχέτισης των $X_1 - X_2$ και $X_2 + X_3$.

21 Υποθέτουμε ότι X και Y είναι δύο τυχαίες μεταβλητές με $\rho(X, Y) = 1/2$, $\mathrm{Var}X = 1$, και $\mathrm{Var}Y = 2$. Υπολογίστε την $\mathrm{Var}(X - 2Y)$.

22 Ένα δοχείο περιέχει 3 κόκκινους και 2 μαύρους βώλους. Επιλέγουμε τυχαίο δείγμα μεγέθους 2, χωρίς επανατοποθέτηση. Έστω U το πλήθος των κόκκινων και V το πλήθος των μαύρων, από τους βώλους που επιλέχθηκαν. Υπολογίστε τον $\rho(U, V)$.

23 Υποθέτουμε ότι ένα δοχείο περιέχει 3 βώλους, που φέρουν τους αριθμούς 1,2, και 3. Επιλέγουμε τυχαία δύο βώλους από το δοχείο χωρίς επανατοποθέτηση. Έστω X ο αριθμός στον πρώτο βώλο και Y ο αριθμός στον δεύτερο βώλο. Υπολογίστε την $\mathrm{Cov}(X, Y)$ και τον $\rho(X, Y)$.

24 Εκτελούμε n φορές ένα πείραμα που έχει r δυνατά αποτελέσματα $1, 2, \ldots, r$ τα οποία συμβαίνουν με πιθανότητες $p_1, \ldots, p_n$. Έστω X το πλήθος των δοκιμών στις οποίες συμβαίνει το πρώτο αποτέλεσμα, και Y το πλήθος των δοκιμών στις οποίες συμβαίνει το δεύτερο αποτέλεσμα. Δείξτε ότι

$$\rho(X, Y) = -\sqrt{\frac{p_1 p_2}{(1-p_1)(1-p_2)}},$$

ακολουθώντας τα εξής βήματα. Ορίστε $I_i = 1$ αν η i-στή δοκιμή δώσει το αποτέλεσμα 1, και $I_i = 0$ αλλιώς. Όμοια, ορίστε $J_i = 1$ αν η i-στή δοκιμή δώσει το αποτέλεσμα 2, και $J_i = 0$ αλλιώς. Τότε, $X = I_1 + \ldots + I_n$ και $Y = J_1 + \ldots + J_n$. Τώρα δείξτε τα εξής:
(α) $\mathbb{E}(I_i J_i) = 0$.
(β) Αν $i \neq j$, τότε $\mathbb{E}(I_i J_j) = p_1 p_2$.
(γ) $\mathbb{E}(XY) = \mathbb{E}\left(\sum_{i=1}^n I_i J_i\right) + \mathbb{E}\left(\sum_{i=1}^n \sum_{j \neq i} I_i J_j\right) = n(n-1) p_1 p_2$.
(δ) $\mathrm{Cov}(X, Y) = -n p_1 p_2$.
(ε) $\rho(X, Y) = -\sqrt{\frac{p_1 p_2}{(1-p_1)(1-p_2)}}$.

25 Υποθέτουμε ότι ένας πληθυσμός r αντικειμένων αποτελείται από r_1 αντικείμενα τύπου 1, r_2 αντικείμενα τύπου 2, και r_3 αντικείμενα τύπου 3, όπου $r_1 + r_2 + r_3 = r$. Επιλέγουμε χωρίς επανατοποθέτηση τυχαίο δείγμα μεγέθους $n \leq r$ από αυτόν τον πληθυσμό. Έστω X το πλήθος των αντικειμένων τύπου 1 στο δείγμα, και Y το πλήθος των αντικειμένων τύπου 2. Υπολογίστε τον $\rho(X, Y)$ κάνοντας το εξής. Ορίστε $I_i = 1$ ή 0 ανάλογα με το αν το i-στό στοιχείο του δείγματος είναι τύπου 1 ή όχι, και $J_i = 1$ ή 0 ανάλογα με το αν το i-στό στοιχείο του δείγματος είναι τύπου 2 ή όχι.
(α) Δείξτε ότι $\mathbb{E} I_i = r_1 / r$ και $\mathbb{E} J_i = r_2 / r$.
(β) Δείξτε ότι αν $i \neq j$,

$$\mathbb{E}(I_i J_j) = \frac{r_1 r_2}{r(r-1)}$$

και ότι $\mathbb{E}(I_i J_i) = 0$.
(γ) Ορίστε $X = I_1 + \ldots + I_n$ και $Y = J_1 + \ldots + J_n$, και χρησιμοποιώντας τα (α) και (β) υπολογίστε τις $\mathbb{E}(XY)$, $\mathrm{Var} X$, και $\mathrm{Var} Y$.
(δ) Χρησιμοποιήστε το (γ) για να υπολογίσετε τον $\rho(X, Y)$. Συγκρίνατε με τον αντίστοιχο συντελεστή συσχέτισης στην Άσκηση 24, αν $p_1 = r_1 / r$ και $p_2 = r_2 / r$.

26 Έστω X τυχαία μεταβλητή με πυκνότητα f που δίνεται από την

$$f(x) = \begin{cases} 1/18 & , x = 1, 3, \\ \\ 16/18 & , x = 2. \end{cases}$$

Δείξτε ότι υπάρχει τιμή του δ με την ιδιότητα $P(|X - \mu| \geq \delta) = \mathrm{Var}(X/\delta^2)$, άρα, σε πλήρη γενικότητα, το φράγμα που δίνει η ανισότητα του Chebyshev δεν βελτιώνεται.

27 Κάποιος φτιάχνει βίδες και γνωρίζει ότι το 5% της παραγωγής του είναι ελαττωματικό. Στέλνει παρτίδες των 10000 τεμαχίων, δίνοντας την εγγύηση ότι αν περισσότερες από a βίδες είναι ελαττωματικές θα επιστρέψει την αμοιβή του. Πόσο μικρό a μπορεί να επιλέξει ο παραγωγός ώστε να είναι βέβαιος ότι δεν θα χρειαστεί να επιστρέψει την αμοιβή του για περισσότερες από το 1% των παρτίδων;

28 Υποθέτουμε ότι η X έχει πυκνότητα Poisson με παράμετρο λ. Χρησιμοποιώντας την ανισότητα του Chebyshev αποδείξτε τις ανισότητες:

(α) $P\left(X \leq \frac{\lambda}{2}\right) \leq \frac{4}{\lambda}$· (β) $P(X \geq 2\lambda) \leq \frac{1}{\lambda}$.

29 Έστω X μη αρνητική τυχαία μεταβλητή με ακέραιες τιμές, της οποίας η γεννήτρια συνάρτηση πιθανότητας $\Phi_X(t) = \mathbb{E}t^X$ είναι πεπερασμένη για κάθε t, και έστω x_0 ένας θετικός αριθμός. Μιμούμενοι την απόδειξη της ανισότητας του Chebyshev, αποδείξτε τις ανισότητες:

(α) $P(X \leq x_0) \leq \frac{\Phi_X(t)}{t^{x_0}}, \quad 0 \leq t \leq 1$·

(β) $P(X \geq x_0) \leq \frac{\Phi_X(t)}{t^{x_0}}, \quad t \geq 1$.

30 Υποθέτουμε ότι η X έχει πυκνότητα Poisson με παράμετρο λ. Αποδείξτε τις ανισότητες:

(α) $P\left(X \leq \frac{\lambda}{2}\right) \leq \left(\frac{2}{e}\right)^{\lambda/2}$· (β) $P(X \geq 2\lambda) \leq \left(\frac{e}{4}\right)^{\lambda}$.

Υπόδειξη: Με χρήση απειροστικού λογισμού, ελαχιστοποιήστε τα δεξιά μέλη των ανισοτήτων στην Άσκηση 29. Οι ανισότητες αυτές είναι πολύ ακριβέστερες, ειδικά για μεγάλες τιμές του λ, από εκείνες που δόθηκαν στην Άσκηση 28.

Στις Ασκήσεις 31-36 αναπτύσσουμε και εφαρμόζουμε τις έννοιες της δεσμευμένης πυκνότητας και της δεσμευμένης μέσης τιμής.

Έστω X και Y διακριτές τυχαίες μεταβλητές. Ορίζουμε τη *δεσμευμένη πυκνότητα $f_{Y|X}(y|x)$ της Y δεδομένου ότι $X = x$* μέσω της

$$f_{Y|X}(y|x) = \begin{cases} P(Y = y | X = x) & , \text{αν } P(X = x) > 0, \\ \\ 0 & , \text{αλλιώς}. \end{cases}$$

Για κάθε x που ικανοποιεί την $P(X = x) > 0$ έπεται ότι η $f_{Y|X}(y|x)$ είναι πυκνότητα ως προς y. Το Παράδειγμα 14(δ) του Κεφαλαίου 3 ερμηνεύεται τώρα ως εξής: αν οι X και Y είναι ανεξάρτητες, γεωμετρικά κατανεμημένες με παράμετρο p, τότε για κάθε $z \geq 0$, η δεσμευμένη πυκνότητα της Y δεδομένου ότι $X + Y = z$ είναι η ομοιόμορφη πυκνότητα στο $\{0, 1, \ldots, z\}$.

Έστω ότι η Y έχει πεπερασμένη μέση τιμή. Η *δεσμευμένη μέση τιμή της Y δεδομένου ότι $X = x$* ορίζεται να είναι η μέση τιμή της δεσμευμένης πυκνότητας της Y δεδομένου ότι $X = x$, δηλαδή

$$\mathbb{E}[Y|X = x] = \sum_y y f_{Y|X}(y|x).$$

31 Επαληθεύστε τις ακόλουθες ιδιότητες της δεσμευμένης πυκνότητας και της δεσμευμένης μέσης τιμής:
(α) $f_Y(y) = \sum_x f_X(x) f_{Y|X}(y|x)$·
(β) $\mathbb{E}Y = \sum_x f_X(x) \mathbb{E}[Y|X = x]$.

32 Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές που έχουν γεωμετρική πυκνότητα με παράμετρο p. Βρείτε την $\mathbb{E}[Y|X + Y = z]$ για κάθε μη αρνητικό ακέραιο z. *Υπόδειξη:* Χρησιμοποιήστε το Παράδειγμα 14(δ) και την Άσκηση 8.

33 Έστω X και Y δύο ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν κατανομή Poisson με παραμέτρους λ_1 και λ_2 αντίστοιχα. Υπολογίστε την $\mathbb{E}[Y|X + Y = z]$ για κάθε μη αρνητικό ακέραιο z. *Υπόδειξη:* Χρησιμοποιήστε το αποτέλεσμα της Άσκησης 35 του Κεφαλαίου 3.

34 Έστω N μία μη αρνητική τυχαία μεταβλητή με ακέραιες τιμές. Υποθέτουμε ότι $\{Y_n\}$, $n \geq 0$, είναι τυχαίες μεταβλητές με πεπερασμένες μέσες τιμές, ανεξάρτητες από την N. Δείξτε ότι

$$\mathbb{E}[Y_N|N = n] = \mathbb{E}Y_n.$$

35 Έστω $\{X_n\}$, $n \geq 1$, ανεξάρτητες τυχαίες μεταβλητές που έχουν την ίδια πεπερασμένη μέση τιμή μ και διασπορά σ^2. Θέτουμε $S_0 = 0$ και $S_n = X_1 + \ldots + X_n$, $n \geq 1$. Έστω N μία μη αρνητική τυχαία μεταβλητή με πεπερασμένη μέση τιμή και διασπορά, και ας υποθέσουμε ότι η N είναι ανεξάρτητη από όλες τις τυχαίες μεταβλητές που ορίζονται με βάση τις $\{X_n\}$, $n \geq 1$. Τότε η S_N έχει πεπερασμένη μέση τιμή και διασπορά. Δείξτε ότι

$$\mathbb{E}S_N = \mu \mathbb{E}N, \qquad \mathbb{E}S_N^2 = \sigma^2 \mathbb{E}N + \mu^2 \mathbb{E}N^2,$$

και

$$\mathrm{Var} S_N = \sigma^2 \mathbb{E}N + \mu^2 \mathrm{Var} N.$$

Υπόδειξη: Χρησιμοποιήστε τις Ασκήσεις 31(β) και 34.

36 Αποδείξτε τα αποτελέσματα της Άσκησης 35 παραγωγίζοντας τη γεννήτρια συνάρτηση πιθανότητας της S_N, η οποία υπολογίστηκε στο Παράδειγμα 19 του Κεφαλαίου 3, και θέτοντας $t = 1$.

Κεφάλαιο 5

Συνεχείς τυχαίες μεταβλητές

Στο Κεφάλαιο 3 μελετήσαμε διακριτές τυχαίες μεταβλητές και τις πυκνότητές τους (για παράδειγμα, τη διωνυμική, την υπεργεωμετρική, και την Poisson). Στις εφαρμογές, αυτές οι τυχαίες μεταβλητές συνήθως περιγράφουν το πλήθος των αντικειμένων κάποιου συγκεκριμένου τύπου, όπως το πλήθος των κόκκινων βώλων σε ένα τυχαίο δείγμα μεγέθους n με ή χωρίς επανατοποθέτηση, ή το πλήθος των τηλεφωνημάτων που δέχεται ένα τηλεφωνικό κέντρο σε ένα λεπτό.

Στη θεωρία, αλλά και στην πράξη, εμφανίζονται συχνά καταστάσεις στις οποίες οι φυσιολογικές τυχαίες μεταβλητές που πρέπει να χρησιμοποιήσουμε είναι «συνεχείς» και όχι διακριτές. Αυτό που έχουμε στο μυαλό μας όταν μιλάμε για μία συνεχή τυχαία μεταβλητή X σε έναν χώρο πιθανότητας Ω είναι μία συνάρτηση $X(\omega), \omega \in \Omega$, με την ιδιότητα

$$P(\{\omega \in \Omega | X(\omega) = x\}) = 0, \qquad -\infty < x < \infty,$$

θεωρούμε δηλαδή ότι η X παίρνει οποιαδήποτε συγκεκριμένη τιμή με πιθανότητα μηδέν.

Δεν είναι δύσκολο να σκεφτούμε παραδείγματα συνεχών τυχαίων μεταβλητών. Σαν ένα πρώτο παράδειγμα, θεωρούμε ένα πιθανοθεωρητικό μοντέλο για το χρόνο διάσπασης ενός πεπερασμένου αριθμού ραδιενεργών σωματιδίων. Έστω T η τυχαία μεταβλητή που δίνει το χρόνο που απαιτείται μέχρι τη διάσπαση του πρώτου σωματιδίου. Τότε η T πρέπει να είναι συνεχής, γιατί η πιθανότητα η πρώτη διάσπαση να συμβεί ακριβώς σε μία καθορισμένη χρονική στιγμή (π.χ. μετά από 2000 δευτερόλεπτα) είναι μηδενική. Σαν ένα δεύτερο παράδειγμα, ας πάρουμε το πείραμα της τυχαίας επιλογής ενός σημείου από ένα υποσύνολο S του n-διάστατου Ευκλείδειου χώρου που έχει πεπερασμένο μη μηδενικό n-διάστατο όγκο (θυμηθείτε τη συζήτηση που έγινε στο Κεφάλαιο 1). Έστω X η τυχαία μεταβλητή που δίνει την πρώτη συντεταγμένη του σημείου που επιλέξαμε. Είναι φανερό ότι η X παίρνει κάθε συγκεκριμένη τιμή με μηδενική πιθανότητα. Ας υποθέσουμε, για παράδειγμα, ότι $n = 2$ και το S είναι ένας δίσκος στο επίπεδο που έχει κέντρο την αρχή των αξόνων και μοναδιαία ακτίνα. Τότε, το σύνολο των σημείων του S που έχουν μηδενική πρώτη συντεταγμένη είναι ένα ευθύγραμμο τμήμα στο επίπεδο. Κάθε τέτοιο ευθύγραμμο τμήμα έχει μηδενικό *εμβαδόν*, άρα μηδενική πιθανότητα.

Με λίγα λόγια, οι τυχαίες μεταβλητές που αφορούν μετρήσεις φυσικών ποσοτήτων όπως οι συντεταγμένες στον χώρο, το βάρος, ο χρόνος, η θερμοκρασία, η τάση του ρεύματος κλπ, περιγράφονται καλύτερα σαν συνεχείς τυχαίες μεταβλητές. Οι τυχαίες μεταβλητές που μετρούν αντικείμενα ή ενδεχόμενα είναι σαφώς παραδείγματα διακριτών τυχαίων μεταβλητών.

Υπάρχουν όμως περιπτώσεις τις οποίες θα μπορούσαν να περιγράψουν τόσο οι διακριτές όσο και οι συνεχείς τυχαίες μεταβλητές. Για παράδειγμα, η μέτρηση του μήκους φυσιολογικά περιγράφεται από μία συνεχή τυχαία μεταβλητή, θα μπορούσαμε όμως να κάνουμε στρογγύλευση των μετρήσεων μέχρι κάποιο δεκαδικό ψηφίο και να χρησιμοποιήσουμε μία διακριτή τυχαία μεταβλητή για τη μελέτη μας.

5.1 Τυχαίες μεταβλητές και οι συναρτήσεις κατανομής τους

Στις εφαρμογές, τυχαία μεταβλητή είναι μία αριθμητική ποσότητα που ορίζεται με βάση το αποτέλεσμα ενός τυχαίου πειράματος. Από τη μαθηματική όμως άποψη, μία τυχαία μεταβλητή X είναι μία πραγματική συνάρτηση που ορίζεται σε έναν χώρο πιθανότητας. Φυσικά, θέλουμε η $P(X \leq x)$ να είναι καλά ορισμένη για κάθε πραγματικό αριθμό x. Με άλλα λόγια, αν $(\Omega, \mathcal{A}, P)$ είναι ο χώρος πιθανότητας στον οποίο ορίζεται η X, θέλουμε το

$$\{\omega \big| \, X(\omega) \leq x\}$$

να είναι ενδεχόμενο (δηλαδή, στοιχείο της $\mathcal{A}$). Οδηγούμαστε έτσι στους εξής ορισμούς.

Ορισμός 1 *Τυχαία μεταβλητή X σε έναν χώρο πιθανότητας $(\Omega, \mathcal{A}, P)$ είναι μία πραγματική συνάρτηση $X(\omega)$, $\omega \in \Omega$, τέτοια ώστε για κάθε $-\infty < x < \infty$, το $\{\omega \mid X(\omega) \leq x\}$ να είναι ενδεχόμενο.*

Ορισμός 2 *Η συνάρτηση κατανομής F μιάς τυχαίας μεταβλητής X είναι η συνάρτηση*

$$F(x) = P(X \leq x), \qquad -\infty < x < \infty.$$

Η συνάρτηση κατανομής είναι χρήσιμη για τον υπολογισμό διαφόρων πιθανοτήτων που σχετίζονται με την τυχαία μεταβλητή X. Ένα παράδειγμα δίνεται από τον τύπο

$$P(a < X \leq b) = F(b) - F(a), \qquad a \leq b. \tag{1}$$

Για να επαληθεύσουμε την (1), θέτουμε $A = \{\omega \mid X(\omega) \leq a\}$ και $B = \{\omega \mid X(\omega) \leq b\}$. Τότε $A \subseteq B$ και, από τον ορισμό της τυχαίας μεταβλητής, τα σύνολα A και B

είναι ενδεχόμενα. Επομένως, το $\{\omega \mid a < X \leq b\} = B \cap A^c$ είναι ενδεχόμενο και η (1) είναι ειδική περίπτωση της ιδιότητας που αποδείξαμε στην Παράγραφο 1.3: αν $A \subseteq B$, τότε

$$P(B \cap A^c) = P(B) - P(A).$$

Παράδειγμα 1 Επιλέγουμε τυχαία ένα σημείο από τον δίσκο με κέντρο την αρχή των αξόνων και ακτίνα R στο επίπεδο. Για να φανεί πιό ενδιαφέρον το πείραμα, μπορούμε να σκεφτόμαστε την επιλογή του σημείου σαν το αποτέλεσμα μιάς βολής πρός έναν κυκλικό στόχο. Ο χώρος πιθανότητας που αντιστοιχεί σε αυτό το πείραμα είναι ο ομοιόμορφος χώρος πιθανότητας που περιγράψαμε στην Παράγραφο 1.2. Έστω X η τυχαία μεταβλητή που περιγράφει την απόσταση του σημείου που επιλέχθηκε από το κέντρο του δίσκου. Η συνάρτηση κατανομής της X υπολογίζεται εύκολα. Αν $0 \leq x \leq R$, το ενδεχόμενο $\{\omega \mid X(\omega) \leq x\}$ είναι ένας δίσκος με κέντρο την αρχή των αξόνων και ακτίνα x. Το εμβαδόν του είναι πx^2. Έτσι, από τον ορισμό του ομοιόμορφου χώρου πιθανότητας,

$$P(X \leq x) = \frac{\pi x^2}{\pi R^2} = \frac{x^2}{R^2}, \quad 0 \leq x \leq R.$$

Αν $x < 0$, τότε $P(X \leq x) = 0$. Αν $x > R$, τότε $P(X \leq x) = 1$. Άρα, η συνάρτηση κατανομής F της τυχαίας μεταβλητής X δίνεται από την

$$(2) \qquad F(x) = \begin{cases} 0 & , x < 0, \\ \frac{x^2}{R^2} & , 0 \leq x \leq R, \\ 1 & , x > R. \end{cases}$$

Η γραφική παράσταση της F δίνεται στο Σχήμα 1. Από τις σχέσεις (1) και (2) έπεται ότι αν $0 \leq a \leq b \leq R$, τότε

$$P(a < X \leq b) = F(b) - F(a) = \frac{b^2 - a^2}{R^2}.$$

Παράδειγμα 2. Θεωρούμε ένα πιθανοθεωρητικό μοντέλο για τον χρόνο διάσπασης ενός πεπερασμένου αριθμού ραδιενεργών σωματιδίων. Έστω X ο χρόνος που απαιτείται για ένα συγκεκριμένο σωματίδιο. Βρείτε τη συνάρτηση κατανομής της X.

Όπως είδαμε στην Παράγραφο 1.1, υπάρχει κατάλληλη θετική τιμή της παραμέτρου λ, τέτοια ώστε

$$P(a < X \leq b) = e^{-\lambda a} - e^{-\lambda b}, \qquad 0 \leq a \leq b < \infty.$$

Αφού η X παίρνει μόνο θετικές τιμές, έχουμε $P(X \leq x) = 0$ αν $x \leq 0$ και, ειδικότερα, $P(X \leq 0) = 0$. Αν $0 < x < \infty$,

$$\begin{aligned} P(X \leq x) &= P(X \leq 0) + P(0 < X \leq x) \\ &= P(0 < X \leq x) \\ &= 1 - e^{-\lambda x}. \end{aligned}$$

Επομένως, η συνάρτηση κατανομής F της X δίνεται από την

$$(3) \qquad F(x) = \begin{cases} 0 & , x \leq 0, \\ 1 - e^{-\lambda x} & , x > 0. \end{cases}$$

Οι διακριτές τυχαίες μεταβλητές έχουν φυσικά κι αυτές συναρτήσεις κατανομής, και στα Παραδείγματα 10 και 11 του Κεφαλαίου 3 υπολογίσαμε δύο από αυτές.

Παράδειγμα 3. Υποθέτουμε ότι η X ακολουθεί διωνυμική κατανομή με παραμέτρους $n = 2$ και $p = 1/2$. Τότε $f(0) = 1/4$, $f(1) = 1/2$, και $f(2) = 1/4$. Επομένως,

$$F(x) = \begin{cases} 0 & , x < 0, \\ 1/4 & , 0 \leq x \leq 1, \\ 3/4 & , 1 \leq x < 2, \\ 1 & , 2 \leq x. \end{cases}$$

Η γραφική παράσταση αυτής της συνάρτησης κατανομής δίνεται στο Σχήμα 2.

5.1.1 Ιδιότητες των συναρτήσεων κατανομής

Οι συναρτήσεις κατανομής ικανοποιούν κάποιες συνθήκες (δεν μπορούν όλες οι συναρτήσεις να παίξουν το ρόλο της συνάρτησης κατανομής). Έστω X μία τυχαία μεταβλητή και F η συνάρτηση κατανομής της. Τότε

(i) $0 \leq F(x) \leq 1$ για κάθε x.
(ii) Η F είναι αύξουσα συνάρτηση του x.

Η ιδιότητα (i) έπεται άμεσα από τον ορισμό: $F(x) = P(X \leq x)$. Για να δούμε ότι ισχύει η (ii) αρκεί να παρατηρήσουμε ότι αν $x < y$, τότε

$$F(y) - F(x) = P(x < X \leq y) \geq 0.$$

Λέμε ότι μία συνάρτηση f έχει όριο L από δεξιά (ή από αριστερά) στο x αν $f(x+h) \to L$ καθώς $h \to 0$, όπου το h περιορίζεται να παίρνει μόνο θετικές (αντίστοιχα, αρνητικές) τιμές. Τα όρια από δεξιά ή από αριστερά, όταν υπάρχουν, συμβολίζονται με $f(x+)$ και $f(x-)$ αντίστοιχα. Δεν είναι δύσκολο να δείτε ότι αν η f είναι φραγμένη και είτε αύξουσα ή φθίνουσα, τότε τα $f(x+)$ και $f(x-)$ υπάρχουν για κάθε x. Με τις ίδιες υποθέσεις, η f έχει όρια $f(-\infty)$ καθώς $x \to -\infty$ και $f(+\infty)$ καθώς $x \to +\infty$.

Από τις ιδιότητες (i) και (ii) και τη συζήτηση της προηγούμενης παραγράφου, έπεται ότι η συνάρτηση κατανομής F έχει όρια $F(x+)$ και $F(x-)$ για κάθε x, καθώς και τα όρια $F(-\infty)$ και $F(+\infty)$.

(iii) $F(-\infty) = 0$ και $F(+\infty) = 1$.
(iv) $F(x+) = F(x)$ για κάθε x.

Για να υπολογίσουμε τα $F(-\infty)$ και $F(+\infty)$ αρκεί να βρούμε το όριο της $F(n)$ καθώς $n \to -\infty$ και $n \to +\infty$ (αυτό, γιατί η F είναι αύξουσα). Ορίζουμε

$$B_n = \{\omega \mid X(\omega) \leq n\}.$$

Τότε $\ldots \subseteq B_{-2} \subseteq B_{-1} \subseteq B_0 \subseteq B_1 \subseteq B_2 \subseteq \ldots$. Επίσης,

$$\bigcap_{n=0}^{-\infty} B_n = \emptyset \quad , \quad \bigcup_{n=0}^{+\infty} B_n = \Omega.$$

Σύμφωνα με το Θεώρημα 1 του Κεφαλαίου 1, βλέπουμε ότι

$$\lim_{n\to-\infty} P(B_n) = P(\emptyset) = 0 \quad , \quad \lim_{n\to+\infty} P(B_n) = P(\Omega) = 1.$$

Αφού $F(n) = P(X \leq n) = P(B_n)$, έχουμε ότι

$$F(-\infty) = \lim_{n\to-\infty} F(n) = \lim_{n\to-\infty} P(B_n) = 0$$

και όμοια ότι $F(+\infty) = 1$.

Η ιδιότητα (iv) μας λέει ότι η F είναι συνάρτηση συνεχής από δεξιά και

(4) $$F(x+) = P(X \leq x), \qquad -\infty < x < \infty.$$

Ένα αποτέλεσμα που συνδέεται στενά με αυτό είναι ότι

(5) $$F(x-) = P(X < x), \qquad -\infty < x < \infty.$$

Οι αποδείξεις των (4) και (5) είναι όμοιες με την απόδειξη της (iii). Για να δείξουμε την (4), για παράδειγμα, αρκεί να δείξουμε ότι $F(x + 1/n) \to P(X \leq x)$ καθώς $n \to +\infty$. Αυτό μπορεί να γίνει αν θέσουμε

$$B_n = \left\{\omega \mid X(\omega) \leq x + \frac{1}{n}\right\},$$

παρατηρήσουμε ότι $\bigcap_n B_n = \{\omega \mid X(\omega) \leq x\}$, και επαναλάβουμε το επιχείρημα της (iii).

Από τις (4) και (5) βλέπουμε αμέσως ότι

(6) $$F(x+) - F(x-) = P(X = x), \qquad -\infty < x < \infty.$$

Αυτός ο τύπος μας λέει ότι αν $P(X = x) > 0$, τότε η F έχει άλμα μεγέθους $P(X = x)$ στο x. Αν $P(X = x) = 0$, τότε η F είναι συνεχής στο x. Υπενθυμίζουμε την έννοια της συνεχούς τυχαίας μεταβλητής που αναφέραμε στην εισαγωγή αυτού του κεφαλαίου.

Ορισμός 3 *Μία τυχαία μεταβλητή X λέγεται συνεχής τυχαία μεταβλητή αν*

$$P(X = x) = 0, \qquad -\infty < x < \infty.$$

Βλέπουμε λοιπόν ότι η X είναι συνεχής τυχαία μεταβλητή αν και μόνο αν η συνάρτηση F κατανομής της είναι συνεχής σε κάθε x, δηλαδή, αν η F είναι συνεχής συνάρτηση. Αν η X είναι συνεχής τυχαία μεταβλητή, τότε εκτός από την (1) έχουμε ότι

$$\begin{aligned} (7) \qquad P(a < X < b) &= P(a \leq X \leq b) = P(a \leq X < b) \\ &= F(b) - F(a), \end{aligned}$$

οπότε τα σύμβολα $<$ και $\leq$ χρησιμοποιούνται χωρίς διάκριση σε αυτό το πλαίσιο. Οι διάφορες ιδιότητες μιάς συνάρτησης κατανομής περιγράφονται στο Σχήμα 3. (Παρατηρήστε ότι η τυχαία μεταβλητή που έχει τη συγκεκριμένη συνάρτηση κατανομής δεν μπορεί να είναι ούτε διακριτή ούτε συνεχής.)

Θεωρούμε την τυχαία μεταβλητή X που ορίστηκε στο Παράδειγμα 1. Από τη σχέση (2) ή το Σχήμα 1 βλέπουμε ότι η συνάρτηση κατανομής της είναι συνεχής. Άρα η X είναι συνεχής τυχαία μεταβλητή. Όμοια, από την (3) είναι φανερό ότι η τυχαία μεταβλητή του Παραδείγματος 2 είναι συνεχής τυχαία μεταβλητή.

Οι περισσότερες τυχαίες μεταβλητές που εμφανίζονται σε πρακτικές εφαρμογές είναι είτε διακριτές ή συνεχείς. Υπάρχουν κάποιες εξαιρέσεις. Πάρτε, ας πούμε, το Παράδειγμα 2. Σε αυτό το παράδειγμα, η X παριστάνει το χρόνο διάσπασης ενός συγκεκριμένου σωματιδίου. Αν το πείραμα έχει προκαθορισμένη διάρκεια, για παράδειγμα ως την χρονική στιγμή $t_0 > 0$, και αν το σωματίδιο δεν έχει υποστεί μεταβολή μέχρι εκείνη τη στιγμή, τότε ο πραγματικός χρόνος αποσύνθεσής του X δεν θα παρατηρηθεί. Ένας πιθανός τρόπος για να ξεφύγουμε από αυτή τη δυσκολία είναι να ορίσουμε μία νέα τυχαία μεταβλητή Y ως εξής:

$$Y(\omega) = \begin{cases} X(\omega) & , X(\omega) \leq t_0, \\ t_0 & , X(\omega) > t_0. \end{cases}$$

Έτσι, η Y δίνει τον χρόνο διάσπασης αν αυτός ο χρόνος παρατηρείται (δηλαδή, αν είναι μικρότερος ή ίσος του t_0), αλλιώς $Y = t_0$. Η συνάρτηση κατανομής F_Y της Y δίνεται από την

$$F_Y(y) = \begin{cases} 0 & , y < 0, \\ 1 - e^{-\lambda y} & , 0 \leq y < t_0, \\ 1 & , y \geq t_0. \end{cases}$$

Η συνάρτηση κατανομής έχει άλμα μεγέθους $e^{-\lambda t_0}$ στο $y = t_0$. Είναι λοιπόν φανερό ότι η τυχαία μεταβλητή Y που κατασκευάσαμε δεν είναι ούτε διακριτή ούτε συνεχής.

Ορίσαμε τις συναρτήσεις κατανομής με βάση τις τυχαίες μεταβλητές. Θα μπορούσαμε να δώσουμε και απευθείας ορισμό.

Ορισμός 4 *Συνάρτηση κατανομής είναι κάθε συνάρτηση F που έχει τις ιδιότητες* (i)-(iv). *Δηλαδή,*

(i) $0 \leq F(x) \leq 1$ *για κάθε* x,
(ii) *Η* F *είναι αύξουσα συνάρτηση του* x,
(iii) $F(-\infty) = 0$ *και* $F(+\infty) = 1$,
(iv) $F(x+) = F(x)$ *για κάθε* x.

Σε πιό προχωρημένα βιβλία αποδεικνύεται ότι *αν F είναι συνάρτηση κατανομής, τότε υπάρχουν χώρος πιθανότητας και τυχαία μεταβλητή X ορισμένη σε αυτό το χώρο, έτσι ώστε η F να είναι η συνάρτηση κατανομής της X.*

5.2 Πυκνότητες συνεχών τυχαίων μεταβλητών

Στην πράξη, οι συνεχείς συναρτήσεις κατανομής συνήθως ορίζονται με τη βοήθεια των συναρτήσεων πυκνότητας.

Ορισμός 5 *Συνάρτηση πυκνότητας (συνεχούς τύπου) είναι μία μη αρνητική συνάρτηση f που ικανοποιεί την*

$$\int_{-\infty}^{\infty} f(x)dx = 1.$$

Παρατηρήστε ότι αν η f είναι συνάρτηση πυκνότητας, τότε η συνάρτηση F που ορίζεται από την

$$F(x) = \int_{-\infty}^{x} f(y)dy, \qquad -\infty < x < \infty, \tag{8}$$

είναι συνεχής, και ικανοποιεί τις ιδιότητες (i)-(iv) της Παραγράφου 5.1.1. Δηλαδή, η (8) ορίζει μία συνεχή συνάρτηση κατανομής. Λέμε ότι αυτή η συνάρτηση κατανομής έχει πυκνότητα την f. Είναι δυνατόν, αλλά δύσκολο, να κατασκευαστούν παραδείγματα συνεχών συναρτήσεων κατανομής που δεν έχουν πυκνότητες. Αυτές που έχουν πυκνότητες λέγονται *απολύτως συνεχείς* συναρτήσεις κατανομής.

Αν X είναι μία συνεχής τυχαία μεταβλητή με συνάρτηση κατανομής την F, όπου η F δίνεται από την (8), τότε η f λέγεται και πυκνότητα της X. Στη συνέχεια θα χρησιμοποιούμε τον όρο «συνάρτηση πυκνότητας» τόσο για τις διακριτές όσο και για τις συναρτήσεις πυκνότητας συνεχούς τύπου. Από το πλαίσιο στο οποίο συζητάμε θα είναι κάθε φορά φανερό το είδος της πυκνότητας που μελετάμε. Για παράδειγμα, η φράση «έστω X μία συνεχής τυχαία μεταβλητή με πυκνότητα f» υποχρεωτικά σημαίνει ότι η f είναι συνάρτηση πυκνότητας συνεχούς τύπου.

Από τις (1) και (8) έπεται ότι αν X είναι μία συνεχής τυχαία μεταβλητή με πυκνότητα f, τότε

$$P(a \leq X \leq b) = \int_a^b f(x)dx, \quad a \leq b, \tag{9}$$

ή, κάπως πιό γενικά, ότι

$$P(X \in A) = \int_A f(x)dx \tag{10}$$

αν το A είναι μία πεπερασμένη ή άπειρη αριθμήσιμη ένωση ξένων διαστημάτων. Δηλαδή, η $P(X \in A)$ αναπαρίσταται από το εμβαδόν κάτω από το γράφημα της f καθώς το x παίρνει τιμές στο A (βλέπε Σχήμα 4).

Στις περισσότερες εφαρμογές, ο ευκολότερος τρόπος για να υπολογίσουμε την πυκνότητα μιάς συνεχούς τυχαίας μεταβλητής είναι να παραγωγίσουμε την (8), οπότε παίρνουμε

$$f(x) = F'(x), \quad -\infty < x < \infty. \tag{11}$$

Για την ακρίβεια, η (11) ισχύει σε κάθε σημείο x στο οποίο η f είναι συνεχής.

Παράδειγμα 4. Έστω X η τυχαία μεταβλητή του Παραδείγματος 1, της οποίας η συνάρτηση κατανομής F δίνεται από την (2). Τότε

$$(12) \qquad F'(x) = \begin{cases} 0 & , x < 0, \\ 2x/R^2 & , 0 \leq x < R, \\ 0 & , x > R. \end{cases}$$

Στο σημείο $x = R$ η συνάρτηση F δεν είναι παραγωγίσιμη. Αν όμως ορίσουμε την f θέτοντας $f(x) = F'(x), x \neq R$, και $f(R) = 0$, τότε αυτή η f θα είναι πυκνότητα για την F.

Σημειώνουμε ότι η (8) δεν προσδιορίζει την f μονοσήμαντα, γιατί μπορούμε πάντα να αλλάξουμε την τιμή μιάς συνάρτησης σε πεπερασμένα το πλήθος σημεία χωρίς να αλλάξει το ολοκλήρωμα της συνάρτησης στα διαστήματα. Ένας τυπικός τρόπος ορισμού της f είναι να θέσουμε $f(x) = F'(x)$ σε κάθε σημείο στο οποίο υπάρχει η $F'(x)$ και $f(x) = 0$ παντού αλλού. Με αυτόν τον τρόπο ορίζεται μία πυκνότητα της F, αρκεί η F να είναι παντού συνεχής και η F' να υπάρχει και να είναι συνεχής σε όλα εκτός από πεπερασμένα το πλήθος σημεία.

Υπάρχουν κι άλλοι τρόποι για να αποδείξουμε ή να επαληθεύσουμε τύπους για την πυκνότητα μιάς συνεχούς συνάρτησης κατανομής F. Αν μας δώσουν μία συνάρτηση πυκνότητας f μπορούμε να δείξουμε ότι η f είναι η συνάρτηση πυκνότητας της F επαληθεύοντας ότι η (8) ισχύει. Μπορούμε, εξίσου καλά, να αντιστρέψουμε τη διαδικασία και να δείξουμε ότι η F γράφεται στην μορφή (8) για κάποια μη αρνητική συνάρτηση f. Τότε, η f είναι αναγκαστικά συνάρτηση πυκνότητας της F. Αυτές οι μέθοδοι, που είναι ουσιαστικά ισοδύναμες, είναι συνήθως πιό πολύπλοκες από την παραγώγιση. Είναι όμως αυστηρές και αποφεύγουν την ξεχωριστή μελέτη των σημείων στα οποία δεν υπάρχει η $F'(x)$.

Θα περιγράψουμε τη χρήση αυτών των μεθόδων στο πρώτο παράδειγμα της επόμενης υποπαραγράφου.

5.2.1 Τύποι αλλαγής μεταβλητής

Έστω X μία συνεχής τυχαία μεταβλητή με πυκνότητα f. Θα συζητήσουμε μεθόδους για την εύρεση της πυκνότητας μιάς τυχαίας μεταβλητής Y που είναι συνάρτηση της X.

Παράδειγμα 5. Έστω X συνεχής τυχαία μεταβλητή με πυκνότητα f. Βρείτε την πυκνότητα της τυχαίας μεταβλητής $Y = X^2$.

Για να λύσουμε αυτό το πρόβλημα, συμβολίζουμε με F και G τις αντίστοιχες συναρτήσεις κατανομής των X και Y. Τότε $G(y) = 0$ αν $y \leq 0$. Αν $y > 0$,

$$\begin{aligned} G(y) &= P(Y \leq y) = P(X^2 \leq y) \\ &= P(-\sqrt{y} \leq X \leq \sqrt{y}) \\ &= F(\sqrt{y}) - F(-\sqrt{y}), \end{aligned}$$

και παραγωγίζοντας βλέπουμε ότι

$$G'(y) = \frac{1}{2\sqrt{y}} \left(F'(\sqrt{y}) + F'(-\sqrt{y}) \right)$$

$$= \frac{1}{2\sqrt{y}}\left(f(\sqrt{y}) + f(-\sqrt{y})\right).$$

Επομένως, η $Y = X^2$ έχει πυκνότητα g που δίνεται από την

$$(13) \qquad g(y) = \begin{cases} \frac{1}{2\sqrt{y}}\left(f(\sqrt{y}) + f(-\sqrt{y})\right) & , y > 0, \\ 0 & , y \le 0. \end{cases}$$

Αν και η (13) ισχύει γενικά, η απόδειξή μας βασίστηκε στην παραγώγιση, η οποία μπορεί να μην ισχύει σε όλα τα σημεία. Για να δώσουμε μία στοιχειώδη αλλά τελείως αυστηρή απόδειξη της (13), ορίζουμε την g όπως στο δεξιό μέλος της (13), και για κάθε $x > 0$ γράφουμε

$$\int_{-\infty}^{x} g(y)dy = \int_{0}^{x} \frac{1}{2\sqrt{y}}(f(\sqrt{y}) + f(-\sqrt{y}))dy.$$

Κάνοντας την αλλαγή μεταβλητής $z = \sqrt{y}$ (οπότε $dz = dy/2\sqrt{y}$), παίρνουμε

$$\begin{aligned} \int_{-\infty}^{x} g(y)dy &= \int_{0}^{\sqrt{x}} (f(z) + f(-z))dz \\ &= \int_{-\sqrt{x}}^{\sqrt{x}} f(z)dz \\ &= F(\sqrt{x}) - F(-\sqrt{x}) = G(x), \end{aligned}$$

δηλαδή η g είναι όντως πυκνότητα της G.

Από το σημείο αυτό και πέρα, θα χρησιμοποιούμε ελεύθερα τη μέθοδο της παραγώγισης για να αποδεικνύουμε τύπους σαν τον (13), γνωρίζοντας ότι, αν μας το ζητούσαν, θα μπορούσαμε να δώσουμε εναλλακτικές αυστηρές αποδείξεις με βάση την ολοκλήρωση.

Θα χρησιμοποιήσουμε τώρα την (13) για να βρούμε την πυκνότητα της X^2, όπου X είναι η τυχαία μεταβλητή του Παραδείγματος 1. Στο Παράδειγμα 4 είδαμε ότι η πυκνότητα της X είναι $f(x) = 2x/R^2$ αν $0 \le x < R$, και $f(x) = 0$ αλλού. Επομένως, από την (13), η X^2 έχει πυκνότητα g που δίνεται από την

$$g(y) = \frac{1}{2\sqrt{y}}\frac{2\sqrt{y}}{R^2}, \qquad 0 < y < R^2,$$

και $g(y) = 0$ αλλού. Η πυκνότητα αυτή είναι ομοιόμορφη πυκνότητα στο $(0, R^2)$ σύμφωνα με τον ορισμό που ακολουθεί.

Ορισμός 6. *Έστω a και b σταθερές με $a < b$. Η ομοιόμορφη πυκνότητα στο διάστημα (a, b) είναι η πυκνότητα f που ορίζεται από την*

$$(14) \qquad f(x) = \begin{cases} (b-a)^{-1} & , a < x < b, \\ 0 & , \text{αλλιώς}. \end{cases}$$

Η συνάρτηση κατανομής που αντιστοιχεί στην (14) δίνεται από την

$$(15) \qquad F(x) = \begin{cases} 0 & , x < a, \\ (x-a)/(b-a), & , a \le x \le b, \\ 1 & , x > b. \end{cases}$$

Δεν είναι καθόλου δύσκολο να δώσουμε κι άλλα παραδείγματα ομοιόμορφα κατανεμημένων τυχαίων μεταβλητών. Αν ένας καλά κατασκευασμένος τροχός τεθεί σε περιστροφή και ακινητοποιηθεί μετά από ένα μεγάλο πλήθος περιστροφών, είναι λογικό να υποθέσουμε ότι η γωνία (μετρημένη σε ακτίνια) που σχηματίζει με την κατακόρυφη ο τροχός όταν σταματάει είναι ομοιόμορφα κατανεμημένη στο $(-\pi, \pi)$ ή, ισοδύναμα, στο $(0, 2\pi)$. Στις εφαρμογές της θεωρίας πιθανοτήτων στην αριθμητική ανάλυση, υποθέτουμε συχνά ότι το σφάλμα στρογγύλευσης που προκαλείται αν αγνοήσουμε όλα τα δεκαδικά ψηφία ενός αριθμού από το $(n+1)$-στό και πέρα, είναι ομοιόμορφα κατανεμημένο στο $(0, 10^{-n})$.

Παράδειγμα 6. Υποθέτουμε ότι η X είναι ομοιόμορφα κατανεμημένη στο $(0,1)$. Βρείτε την πυκνότητα της $Y = -\lambda^{-1}\log(1-X)$, όπου $\lambda > 0$.

Έστω G η συνάρτηση κατανομής της Y. Παρατηρούμε πρώτα ότι η Y είναι θετική τυχαία μεταβλητή, επομένως $G(y) = 0$ αν $y \le 0$. Αν $y > 0$, έχουμε

$$\begin{aligned} G(y) = P(Y \le y) &= P(-\lambda^{-1}\log(1-X) \le y) \\ &= P(\log(1-X) \ge -\lambda y) \\ &= P(1 - X \ge e^{-\lambda y}) \\ &= P(X \le 1 - e^{-\lambda y}) \\ &= 1 - e^{-\lambda y}. \end{aligned}$$

Άρα, $G'(y) = \lambda e^{-\lambda y}$ αν $y > 0$, και $G'(y) = 0$ αν $y < 0$. Η πυκνότητα της Y δίνεται λοιπόν από την

$$(16) \qquad g(y) = \begin{cases} \lambda e^{-\lambda y} & , y > 0, \\ 0 & , y \le 0. \end{cases}$$

Αυτή η πυκνότητα λέγεται *εκθετική πυκνότητα με παράμετρο* λ και θα την συζητήσουμε εκτενέστερα στην επόμενη παράγραφο.

Το παράδειγμα που μελετήσαμε εντάσσεται σε μία κλάση προβλημάτων που λύνονται με τη βοήθεια του θεωρήματος που ακολουθεί.

Θεώρημα 1. *Έστω φ μία παραγωγίσιμη, γνησίως αύξουσα ή γνησίως φθίνουσα συνάρτηση σε ένα διάστημα I, και έστω $\varphi(I)$ το σύνολο τιμών της φ και φ^{-1} η αντίστροφη συνάρτηση της φ. Έστω X μία συνεχής τυχαία μεταβλητή με πυκνότητα f ώστε $f(x) = 0$ αν $x \notin I$. Τότε, η $Y = \varphi(X)$ έχει πυκνότητα g που δίνεται από την $g(y) = 0$ αν $y \notin \varphi(I)$ και*

$$(17) \qquad g(y) = f(\varphi^{-1}(y))\left|\frac{d}{dy}\varphi^{-1}(y)\right|, \qquad y \in \varphi(I).$$

Είναι συχνά προτιμότερο να γράφουμε την (17) στην ισοδύναμη μορφή

$$g(y) = f(x)\left|\frac{dx}{dy}\right|, \qquad y \in \varphi(I) \quad \textit{και} \quad x = \varphi^{-1}(y), \tag{18}$$

(ή αλλιώς, $g(y)|dy| = f(x)|dx|$).

Για να αποδείξουμε την (17), γράφουμε F και G για τις αντίστοιχες συναρτήσεις κατανομής των X και Y. Υποθέτουμε πρώτα ότι η φ είναι γνησίως αύξουσα (δηλαδή, αν $x_1, x_2 \in I$ και $x_1 < x_2$ τότε $\varphi(x_1) < \varphi(x_2)$). Τότε, η φ^{-1} είναι γνησίως αύξουσα στο $\varphi(I)$ και για κάθε $y \in \varphi(I)$,

$$\begin{aligned} G(y) &= P(Y \le y) \\ &= P(\varphi(X) \le y) \\ &= P(X \le \varphi^{-1}(y)) \\ &= F(\varphi^{-1}(y)). \end{aligned}$$

Άρα, από τον κανόνα της αλυσίδας για την παραγώγιση,

$$\begin{aligned} G'(y) &= \frac{d}{dy}F(\varphi^{-1}(y)) \\ &= F'(\varphi^{-1}(y))\frac{d}{dy}\varphi^{-1}(y) \\ &= f(\varphi^{-1}(y))\frac{d}{dy}\varphi^{-1}(y). \end{aligned}$$

Όμως,

$$\frac{d}{dy}\varphi^{-1}(y) = \left|\frac{d}{dy}\varphi^{-1}(y)\right|$$

γιατί η φ^{-1} είναι γνησίως αύξουσα, άρα η (17) ισχύει. Ας υποθέσουμε τώρα ότι η φ είναι γνησίως φθίνουσα στο I. Τότε η φ^{-1} είναι γνησίως φθίνουσα στο $\varphi(I)$, και για κάθε $y \in \varphi(I)$

$$\begin{aligned} G(y) &= P(Y \le y) \\ &= P(\varphi(X) \le y) \\ &= P(X \ge \varphi^{-1}(y)) \\ &= 1 - F(\varphi^{-1}(y)). \end{aligned}$$

Άρα,

$$\begin{aligned} G'(y) &= -F'(\varphi^{-1}(y))\frac{d}{dy}\varphi^{-1}(y) \\ &= f(\varphi^{-1}(y))\left(-\frac{d}{dy}\varphi^{-1}(y)\right). \end{aligned}$$

Όμως

$$-\frac{d}{dy}\varphi^{-1}(y) = \left|\frac{d}{dy}\varphi^{-1}(y)\right|$$

γιατί η φ^{-1} είναι γνησίως φθίνουσα. Βλέπουμε λοιπόν ότι, και στις δύο περιπτώσεις, η G έχει την πυκνότητα g που ορίζει η (17). $\square$

Παράδειγμα 7. Έστω X τυχαία μεταβλητή που έχει εκθετική πυκνότητα με παράμετρο λ. Βρείτε την πυκνότητα της $Y = X^{1/\beta}$, όπου $\beta \neq 0$.

Σύμφωνα με τον ορισμό που δώσαμε στο προηγούμενο παράδειγμα, η X έχει πυκνότητα f που δίνεται από την $f(x) = \lambda e^{-\lambda x}$ αν $x > 0$, και $f(x) = 0$ αν $x \leq 0$. Μπορούμε λοιπόν να εφαρμόσουμε το θεώρημα με $\varphi(x) = x^{1/\beta}$, $x > 0$. Η εξίσωση $y = x^{1/\beta}$ έχει λύση την $x = y^{\beta}$, οπότε $dx/dy = \beta y^{\beta-1}$. Σύμφωνα με την (18), η πυκνότητα g της Y δίνεται από την

$$g(y) = \begin{cases} |\beta|\lambda y^{\beta-1}e^{-\lambda y^{\beta}} & , y > 0, \\ 0 & , y \leq 0. \end{cases}$$

Παράδειγμα 8. Έστω X συνεχής τυχαία μεταβλητή με πυκνότητα f, και έστω a και b σταθερές με $b \neq 0$. Σύμφωνα με το Θεώρημα 1, η τυχαία μεταβλητή $Y = a + bX$ έχει πυκνότητα που δίνεται από την

$$g(y) = \frac{1}{|b|} f\left(\frac{y-a}{b}\right) \quad , \qquad -\infty < y < \infty. \tag{19}$$

Για μία εφαρμογή αυτού του τύπου, ας θεωρήσουμε την τυχαία μεταβλητή X του Παραδείγματος 1. Στο Παράδειγμα 4 είδαμε ότι η πυκνότητα f της X είναι ίση με $f(x) = 2x/R^2$ αν $0 < x < R$ και ίση με $f(x) = 0$ αλλιώς. Θεωρούμε την τυχαία μεταβλητή $Y = X/R$ και συμβολίζουμε την πυκνότητά της με g. Τότε, από τον τύπο (19) με $a = 0$ και $b = 1/R$,

$$g(y) = Rf(Ry) = 2y, \qquad 0 < y < 1,$$

και $g(y) = 0$ αλλιώς.

Ενδέχεται ο αναγνώστης να προτιμάει να αποδείξει τους τύπους των Παραδειγμάτων 7 και 8 χρησιμοποιώντας την ευθεία μέθοδο του Παραδείγματος 6 αντί για το Θεώρημα 1.

Όπως είδαμε στα παραδείγματα που προηγήθηκαν, μπορούμε να κατασκευάσουμε συναρτήσεις πυκνότητας θεωρώντας συναρτήσεις τυχαίων μεταβλητών. Υπάρχει κι ένας άλλος απλός τρόπος για να κατασκευάζουμε συναρτήσεις πυκνότητας. Έστω g μία μη αρνητική συνάρτηση με την ιδιότητα

$$0 < \int_{-\infty}^{\infty} g(x)dx < \infty.$$

Μπορούμε πάντα να κανονικοποιήσουμε την g ώστε να προκύψει μία συνάρτηση πυκνότητας $f = c^{-1}g$, όπου c είναι η σταθερά

$$c = \int_{-\infty}^{\infty} g(x)dx.$$

Τα παραδείγματα που ακολουθούν περιγράφουν αυτή τη μέθοδο.

Παράδειγμα 9. Έστω $g(x) = x(1-x), 0 \le x \le 1$, και $g(x) = 0$ παντού αλλού. Τότε,

$$c = \int_0^1 x(1-x)dx = \left(\frac{x^2}{2} - x^3 3\right)\Big|_0^1 = \frac{1}{6},$$

και η $f = c^{-1}g$ ορίζεται από την $f(x) = 6x(1-x)$, $0 \le x \le 1$, και $f(x) = 0$ αλλιώς. Η αντίστοιχη συνάρτηση κατανομής είναι η $F(x) = 0$ αν $x < 0$, $F(x) = 3x^2 - 2x^3$ αν $0 \le x \le 1$, και $F(x) = 1$ αν $x > 1$.

Παράδειγμα 10. Θεωρούμε την $g(x) = 1/(1+x^2)$, $-\infty < x < \infty$. Από τον απειροστικό λογισμό ξέρουμε ότι το αόριστο ολοκλήρωμα της $1/(1+x^2)$ είναι $\arctan x$. Άρα

$$c = \int_{-\infty}^{\infty} \frac{dx}{1+x^2} = \arctan x\Big|_{-\infty}^{\infty} = \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) = \pi.$$

Επομένως, η $f = c^{-1}g$ δίνεται από την

$$f(x) = \frac{1}{\pi(1+x^2)}, \qquad -\infty < x < \infty.$$

Η πυκνότητα αυτή είναι γνωστή σαν *πυκνότητα Cauchy.* Η αντίστοιχη συνάρτηση κατανομής δίνεται από την

$$F(x) = \frac{1}{2} + \frac{1}{\pi}\arctan x, \qquad -\infty < x < \infty.$$

Ας δούμε ένα παράδειγμα τυχαίας μεταβλητής που ακολουθεί κατανομή Cauchy:

Παράδειγμα 11. Έστω X η εφαπτομένη μιάς γωνίας (μετρημένης σε ακτίνια) που επιλέγεται τυχαία από το $(-\pi/2, \pi/2)$. Βρείτε την κατανομή της X.

Για τη λύση αυτού του προβλήματος θέτουμε Θ την τυχαία μεταβλητή που δίνει την γωνία που επιλέγουμε, μετρημένη σε ακτίνια. Τώρα, $X = \tan\Theta$, επομένως (δείτε το Σχήμα 5) αν $-\infty < x < \infty$,

$$\begin{aligned}
P(X \le x) &= P(\tan\Theta \le x) \\
&= P\left(-\frac{\pi}{2} < \Theta \le \arctan x\right) \\
&= \frac{1}{\pi}\left(\arctan x - \left(-\frac{\pi}{2}\right)\right) \\
&= \frac{1}{2} + \frac{1}{\pi}\arctan x.
\end{aligned}$$

Άρα η X ακολουθεί κατανομή Cauchy.

5.2.2 Συμμετρικές πυκνότητες

Θα κλείσουμε αυτήν την παράγραφο με λίγα λόγια για τις συμμετρικές πυκνότητες και τις συμμετρικές τυχαίες μεταβλητές. Μία συνάρτηση πυκνότητας f λέγεται συμμετρική αν $f(-x) = f(x)$ για κάθε x. Η πυκνότητα Cauchy και η ομοιόμορφη πυκνότητα στο $(-a, a)$ είναι και οι δύο συμμετρικές. Μία τυχαία μεταβλητή X λέγεται συμμετρική αν η X και η $-X$ έχουν την ίδια συνάρτηση κατανομής. Το αποτέλεσμα που ακολουθεί δείχνει ότι αυτές οι δύο έννοιες συμμετρίας είναι πολύ στενά συνδεδεμένες.

Θεώρημα 2 *Έστω X μία τυχαία μεταβλητή που έχει πυκνότητα. Τότε η f έχει συμμετρική πυκνότητα αν και μόνο αν η X είναι συμμετρική τυχαία μεταβλητή.*

Απόδειξη. Θα αποδείξουμε αυτόν τον ισχυρισμό για συνεχείς τυχαίες μεταβλητές. Η απόδειξη για διακριτές τυχαίες μεταβλητές είναι όμοια. Στην απόδειξή μας θα χρησιμοποιήσουμε το γεγονός ότι για κάθε ολοκληρώσιμη συνάρτηση f

$$\int_{-\infty}^{x} f(-y)dy = \int_{-x}^{\infty} f(y)dy, \qquad -\infty < x < \infty.$$

Υποθέτουμε πρώτα ότι η X έχει συμμετρική πυκνότητα f. Τότε

$$\begin{aligned} P(-X \leq x) &= P(X \geq -x) \\ &= \int_{-x}^{\infty} f(y)dy \\ &= \int_{-\infty}^{x} f(-y)dy \\ &= \int_{-\infty}^{x} f(y)dy \\ &= P(X \leq x), \end{aligned}$$

δηλαδή, οι X και $-X$ έχουν την ίδια συνάρτηση κατανομής.

Αντίστροφα, ας υποθέσουμε ότι οι X και $-X$ έχουν την ίδια πυκνότητα g. Ορίζουμε $f(x) = (g(x) + g(-x))/2$. Η f είναι προφανώς συμμετρική συνάρτηση πυκνότητας. Επίσης,

$$\begin{aligned}
\int_{-\infty}^{x} f(y)dy &= \frac{1}{2}\int_{-\infty}^{x} g(y)dy + \frac{1}{2}\int_{-\infty}^{x} g(-y)dy \\
&= \frac{1}{2}\int_{-\infty}^{x} g(y)dy + \frac{1}{2}\int_{-x}^{\infty} g(y)dy \\
&= \frac{1}{2}P(X \le x) + \frac{1}{2}P(-X \ge -x) \\
&= P(X \le x).
\end{aligned}$$

Άρα, η X έχει τη συμμετρική πυκνότητα f, όπως θέλαμε. □

Αν μία συνεχής συνάρτηση κατανομής F έχει συμμετρική πυκνότητα f, τότε $F(0) = 1/2$. Οι τιμές της F για αρνητικά x υπολογίζονται από τις τιμές της F για θετικά x. Γιατί

$$\begin{aligned}
F(-x) &= \int_{-\infty}^{-x} f(y)dy \\
&= \int_{x}^{\infty} f(-y)dy \\
&= \int_{x}^{\infty} f(y)dy \\
&= \int_{-\infty}^{\infty} f(y)dy - \int_{-\infty}^{x} f(y)dy
\end{aligned}$$

επομένως

$$F(-x) = 1 - F(x), \qquad -\infty < x < \infty. \tag{20}$$

Γιά το λόγο αυτό, όταν κατασκευάζουμε πίνακες για μία τέτοια συνάρτηση κατανομής, συνήθως παρουσιάζουμε μόνο τις μη αρνητικές τιμές του x.

5.3 Κανονικές, εκθετικές, και γάμμα πυκνότητες

Σε αυτήν την παράγραφο θα συζητήσουμε τρεις από τις πιό σημαντικές οικογένειες συναρτήσεων πυκνότητας που εμφανίζονται στη θεωρία πιθανοτήτων και τη στατιστική.

5.3.1 Κανονικές πυκνότητες

Έστω $g(x) = e^{-x^2/2}$, $-\infty < x < \infty$. Για να κανονικοποιήσουμε την g ώστε να την κάνουμε πυκνότητα, πρέπει να υπολογίσουμε τη σταθερά

$$c = \int_{-\infty}^{\infty} e^{-x^2/2} dx.$$

Δεν υπάρχει απλός τύπος για το αόριστο ολοκλήρωμα της $e^{-x^2/2}$. Ο ευκολότερος τρόπος για να υπολογίσουμε την c είναι με την βοήθεια ενός πολύ ειδικού τεχνάσματος: γράφουμε την c σαν ένα διπλό ολοκλήρωμα και εισάγουμε πολικές συντεταγμένες. Πιό συγκεκριμένα, γράφουμε

$$\begin{aligned}
c^2 &= \int_{-\infty}^{\infty} e^{-x^2/2}dx \int_{-\infty}^{\infty} e^{-y^2/2}dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)/2}dxdy \\
&= \int_0^{\infty}\left(\int_{-\pi}^{\pi} e^{-r^2/2}rd\theta\right)dr \\
&= 2\pi\int_0^{\infty} re^{-r^2/2}dr \\
&= -2\pi e^{-r^2/2}\Big|_0^{\infty} \\
&= 2\pi.
\end{aligned}$$

Άρα $c = \sqrt{2\pi}$, και η κανονικοποιημένη μορφή της g είναι

$$f(x) = (2\pi)^{-1/2}e^{-x^2/2}, \qquad -\infty < x < \infty.$$

Είναι καλό να θυμάται κανείς και τον τύπο

$$\int_{-\infty}^{\infty} e^{-x^2/2}dx = \sqrt{2\pi}. \tag{21}$$

Η πυκνότητα που μόλις υπολογίσαμε λέγεται *τυπική κανονική πυκνότητα* και συνήθως συμβολίζεται με φ, δηλαδή

$$\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}, \qquad -\infty < x < \infty. \tag{22}$$

Η τυπική κανονική πυκνότητα είναι προφανώς συμμετρική. Η συνάρτηση κατανομής της φ συμβολίζεται με Φ. Δεν υπάρχει απλός τύπος για την Φ, οι τιμές της λοιπόν υπολογίζονται με αριθμητικές μεθόδους. Υπάρχουν υπολογιστικά προγράμματα και πίνακες, όπως ο Πίνακας Ι στο τέλος αυτού του βιβλίου, που βοηθούν στον υπολογισμό της Φ. Αφού η φ είναι συμμετρική, εφαρμόζοντας την (20) παίρνουμε

$$\Phi(-x) = 1 - \Phi(x), \qquad -\infty < x < \infty. \tag{23}$$

Έστω X τυχαία μεταβλητή που έχει την τυπική κανονική πυκνότητα φ, και ας ορίσουμε $Y = \mu + \sigma X$, όπου $\sigma > 0$. Τότε από τον τύπο (19), η Y έχει πυκνότητα g που δίνεται από την

$$g(y) = \frac{1}{\sigma\sqrt{2\pi}}e^{-(y-\mu)^2/2\sigma^2}, \qquad -\infty < y < \infty.$$

Η πυκνότητα αυτή λέγεται κανονική πυκνότητα με μέσο μ και διασπορά σ^2, και συμβολίζεται με $n(\mu, \sigma^2)$ ή $n(y; \mu, \sigma^2)$, $-\infty < y < \infty$. Δηλαδή,

$$n(y;\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/2\sigma^2} = \varphi\left(\frac{y-\mu}{\sigma}\right), \quad -\infty < y < \infty. \tag{24}$$

Δεν έχουμε ακόμα ορίσει τις ροπές συνεχών τυχαίων μεταβλητών, πρέπει λοιπόν προσωρινά να σκεφτόμαστε τους μ και σ^2 σαν τις δύο παραμέτρους της οικογένειας των κανονικών πυκνοτήτων. Η αντίστοιχη συνάρτηση κατανομής υπολογίζεται με τη βοήθεια της Φ, αφού

$$\begin{aligned} P(Y \le y) &= P(\mu + \sigma X \le y) \\ &= P\left(X \le \frac{y-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{y-\mu}{\sigma}\right). \end{aligned}$$

Έπεται ότι αν η Y είναι κατανεμημένη όπως η $n(\mu, \sigma^2)$ και $a \le b$, τότε

$$P(a \le Y \le b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right). \tag{25}$$

Για παράδειγμα, ας υποθέσουμε ότι η Y ακολουθεί την κατανομή $n(1, 4)$, και έστω $a = 0$, $b = 3$. Από τον Πίνακα Ι βλέπουμε ότι

$$\begin{aligned} P(0 \le Y \le 3) &= \Phi(1) - \Phi(-1/2) = \Phi(1) - (1 - \Phi(1/2)) \\ &= 0.8413 - 0.3085 \\ &= 0.5328. \end{aligned}$$

Αν μία τυχαία μεταβλητή Y ακολουθεί την κατανομή $n(\mu, \sigma^2)$, τότε η τυχαία μεταβλητή $a + bY$, $b \ne 0$, ακολουθεί την κατανομή $n(a + b\mu, b^2\sigma^2)$. Αυτό είναι άμεση συνέπεια της (19). Για μία άλλη αιτιολόγηση, μπορούμε να γράψουμε $Y = \mu + \sigma X$, όπου η X έχει την τυπική κανονική κατανομή. Τότε

$$a + bY = a + b(\mu + \sigma X) = (a + b\mu) + b\sigma X,$$

η οποία ακολουθεί την κατανομή $n(a + b\mu, b^2\sigma^2)$.

Τυχαίες μεταβλητές που ακολουθούν την κανονική κατανομή εμφανίζονται συχνά στις εφαρμογές. Ο Νόμος του Maxwell στη Φυσική ισχυρίζεται ότι, κάτω από κατάλληλες συνθήκες, οι συνιστώσες της ταχύτητας ενός μορίου κάποιου αερίου κατανέμονται τυχαία σύμφωνα με μία κανονική πυκνότητα $n(0, \sigma^2)$, όπου η ποσότητα σ^2 εξαρτάται από κάποιες φυσικές παραμέτρους. Στις περισσότερες εφαρμογές όμως, οι τυχαίες μεταβλητές που μας ενδιαφέρουν έχουν συνάρτηση κατανομής που είναι μόνο κατά προσέγγιση κανονική. Για παράδειγμα, επιβεβαιώνεται εμπειρικά ότι οι μετρήσεις σφαλμάτων στα φυσικά πειράματα, η μεταβλητότητα των προϊόντων μιάς βιομηχανικής γραμμής παραγωγής, η βιολογική μεταβλητότητα (π.χ. ύψους ή βάρους) ακολουθούν κατά προσέγγιση κανονικές κατανομές. Έχει επίσης αποδειχθεί, τόσο θεωρητικά όσο και εμπειρικά, ότι τυχαίες διακυμάνσεις που προκύπτουν σαν συνδυασμός πολλών ασυσχέτιστων αιτίων, καθένα από τα οποία είναι από μόνο του ασήμαντο, τείνουν κατά προσέγγιση να ακολουθούν κανονική κατανομή. Τα θεωρητικά αποτελέσματα αυτού του είδους είναι γνωστά σαν «κεντρικά οριακά θεωρήματα» και συνιστούν ένα από τα σπουδαιότερα αντικείμενα έρευνας στη θεωρία πιθανοτήτων. Θα συζητήσουμε ένα από αυτά τα κεντρικά οριακά θεωρήματα στο Κεφάλαιο 7, και θα το αποδείξουμε στο Κεφάλαιο 8. Η σημασία των κανονικών κατανομών συνδέεται και με τις καλές θεωρητικές τους ιδιότητες. Ένα παράδειγμα τέτοιας ιδιότητας είναι ότι το άθροισμα ανεξάρτητων κανονικά κατανεμημένων τυχαίων μεταβλητών είναι κανονικά κατανεμημένο. Αυτό θα αποδειχθεί στο Κεφάλαιο 6. Στο δεύτερο τόμο αυτής της σειράς θα δούμε ότι οι κανονικές κατανομές παίζουν θεμελιώδη ρόλο στη θεωρητική και την εφαρμοσμένη στατιστική.

5.3.2 Εκθετικές πυκνότητες

Η εκθετική πυκνότητα με παράμετρο λ ορίστηκε στην Παράγραφο 5.2. Δίνεται από την

$$(26) \qquad f(x) = \begin{cases} \lambda e^{-\lambda x} & , x \geq 0, \\ 0 & , x < 0. \end{cases}$$

Η αντίστοιχη συνάρτηση κατανομής είναι η

$$(27) \qquad F(x) = \begin{cases} 1 - e^{-\lambda x} & , x \geq 0, \\ 0 & , x < 0. \end{cases}$$

Σύμφωνα με τη συζήτηση που έγινε στο Κεφάλαιο 1 και στο Παράδειγμα 2 αυτού του Κεφαλαίου βλέπουμε ότι οι τυχαίες μεταβλητές που ακολουθούν εκθετική κατανομή είναι χρήσιμες για τη μελέτη του χρόνου διάσπασης ραδιενεργών σωματιδίων.

Μία σημαντική ιδιότητα των εκθετικά κατανεμημένων τυχαίων μεταβλητών είναι ότι αν X είναι μια τέτοια τυχαία μεταβλητή, τότε

$$(28) \qquad P(X > a + b) = P(X > a)P(X > b), \quad a \geq 0 \ \textit{και} \ b \geq 0.$$

(Ο τύπος αυτός είναι όμοιος με εκείνον που αποδείξαμε στο Κεφάλαιο 3 για γεωμετρικά κατανεμημένες τυχαίες μεταβλητές). Για να δείξουμε ότι η (28) ισχύει, συμβολίζουμε με λ την παράμετρο της εκθετικής κατανομής της X. Τότε, από την (27),

$$\begin{aligned} P(X > a)P(X > b) &= e^{-\lambda a} e^{-\lambda b} \\ &= e^{-\lambda(a+b)} \\ &= P(X > a + b). \end{aligned}$$

Μία ενδιαφέρουσα ισοδύναμη μορφή της (28) είναι η

$$(29) \qquad P(X > a + b \big| X > a) = P(X > b), \quad a \geq 0 \ \textit{και} \ b \geq 0.$$

Σκεφτείτε την X σαν το χρόνο που απαιτείται για να χαλάσει ένα μηχάνημα, μετρημένο από την στιγμή που αρχίζει να λειτουργεί. Τότε η (29) μάς λέει ότι, αν το μηχάνημα δεν έχει χαλάσει μέχρι τη χρονική στιγμή a, η πιθανότητα να μην χαλάσει στις επόμενες b χρονικές μονάδες είναι ίση με την αδέσμευτη πιθανότητα το μηχάνημα να μην χαλάσει στις πρώτες b χρονικές μονάδες. Αυτό σημαίνει ότι η γήρανση του μηχανήματος δεν αυξάνει ούτε μειώνει την πιθανότητα να χαλάσει μέσα σε μία χρονική περίοδο καθορισμένου μήκους.

Το γεγονός ότι η (28) ή η (29) χαρακτηρίζει την οικογένεια των εκθετικών κατανομών φαίνεται από το αποτέλεσμα που ακολουθεί.

Θεώρημα 3 *Έστω X τυχαία μεταβλητή που ικανοποιεί την (28). Τότε, είτε $P(X > 0) = 0$ ή η X ακολουθεί εκθετική κατανομή.*

Απόδειξη: Αν $P(X > 0) = 0$, τότε η (28) ικανοποιείται κατά προφανή τρόπο. Ας υποθέσουμε ότι η (28) ισχύει και ότι $P(X > 0) \neq 0$. Τότε από την (28) με $a = b = 0$ βλέπουμε ότι $P(X > 0) = 1$, άρα η X είναι θετική τυχαία μεταβλητή. Γράφουμε F για την συνάρτηση κατανομής της X, και ορίζουμε G θέτοντας $G(x) = 1 - F(x) = P(X > x)$. Τότε η G είναι μία συνεχής από δεξιά, φθίνουσα συνάρτηση, $G(0) = 1$, $G(+\infty) = 0$, και από την (28)

$$G(a + b) = G(a)G(b), \qquad a > 0 \;\; \textit{και} \;\; b > 0.$$

Έπεται ότι αν $c > 0$ και οι m, n είναι θετικοί ακέραιοι, τότε

$$G(nc) = (G(c))^n \quad \textit{και} \quad G(c) = (G(c/m))^m. \tag{30}$$

Ισχυριζόμαστε τώρα ότι $0 < G(1) < 1$. Γιατί αν $G(1) = 1$, τότε $G(n) = (G(1))^n = 1$, το οποίο αντιφάσκει προς την $G(+\infty) = 0$. Αν $G(1) = 0$, τότε $G(1/m) = 0$ και η συνέχεια από δεξιά μάς δίνει $G(0) = 0$, το οποίο είναι πάλι αντίφαση.

Αφού $0 < G(1) < 1$, μπορούμε να γράψουμε $G(1) = e^{-\lambda}$ όπου $0 < \lambda < \infty$. Έπεται από την (30) ότι αν m είναι ένας θετικός ακέραιος, τότε $G(1/m) = e^{-\lambda/m}$. Εφαρμόζοντας πάλι την (30) βλέπουμε ότι αν m και n είναι θετικοί ακέραιοι, τότε $G(n/m) = e^{-\lambda n/m}$. Με άλλα λόγια, η $G(y) = e^{-\lambda y}$ ισχύει για κάθε θετικό ρητό αριθμό y. Η συνέχεια από δεξιά μάς εξασφαλίζει ότι $G(y) = e^{-\lambda y}$ για κάθε $y \geq 0$. Αυτό σημαίνει ότι η $F = 1 - G$ είναι η εκθετική συνάρτηση κατανομής με παράμετρο λ. □

5.3.3 Πυκνότητες Γάμμα

Πριν δώσουμε τον γενικό ορισμό των πυκνοτήτων γάμμα, ας δούμε ένα παράδειγμα στο οποίο εμφανίζονται φυσιολογικά.

Παράδειγμα 12. Έστω X τυχαία μεταβλητή με κανονική πυκνότητα $n(0, \sigma^2)$. Βρείτε την πυκνότητα της τυχαίας μεταβλητής $Y = X^2$.

Για να λύσουμε αυτό το πρόβλημα παρατηρούμε πρώτα ότι η πυκνότητα της X είναι η

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2}, \qquad -\infty < x < \infty.$$

Από τον τύπο (13), η Y έχει πυκνότητα g που δίνεται από την $g(y) = 0$ αν $y \leq 0$ και

$$g(y) = \frac{1}{2\sqrt{y}} \left(f(\sqrt{y}) + f(-\sqrt{y}) \right), \quad y > 0.$$

Αυτό έχει σαν συνέπεια την

$$g(y) = \frac{1}{\sigma\sqrt{2\pi y}} e^{-y^2/2\sigma^2}, \quad y > 0. \tag{31}$$

Για να δώσουμε τον γενικό ορισμό των πυκνοτήτων γάμμα, θεωρούμε πρώτα συναρτήσεις g της μορφής

$$g(x) = \begin{cases} x^{\alpha-1}e^{-\lambda x} & , x > 0, \\ 0 & , x \leq 0. \end{cases}$$

Ζητάμε να ικανοποιούνται οι $\alpha > 0$ και $\lambda > 0$ ώστε η g να είναι ολοκληρώσιμη. Η πυκνότητα στην (31) αντιστοιχεί στην ειδική περίπτωση $\alpha = 1/2$ και $\lambda = 1/2\sigma^2$. Προκειμένου να κανονικοποιήσουμε την g ώστε να γίνει πυκνότητα, πρέπει να υπολογίσουμε την

$$c = \int_0^\infty x^{\alpha-1}e^{-\lambda x} dx.$$

Κάνουμε την αλλαγή μεταβλητής $y = \lambda x$. Τότε

$$c = \frac{1}{\lambda^\alpha} \int_0^\infty y^{\alpha-1}e^{-y} dy.$$

Δεν υπάρχει απλός τύπος γιά το παραπάνω ολοκλήρωμα. Αντίθετα, χρησιμεύει σαν ορισμός μιάς συνάρτησης που λέγεται γάμμα συνάρτηση και συμβολίζεται με Γ. Δηλαδή

$$c = \frac{1}{\lambda^\alpha}\Gamma(\alpha),$$

όπου

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x} dx, \quad \alpha > 0. \tag{32}$$

Η κανονικοποιημένη συνάρτηση λέγεται πυκνότητα γάμμα με παραμέτρους α και λ, και συμβολίζεται με $\Gamma(\alpha, \lambda)$ ή $\Gamma(x; \alpha, \lambda)$. Βλέπουμε ότι

$$\Gamma(x; \alpha, \lambda) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} e^{-\lambda x} & , x > 0, \\ 0 & , x \leq 0. \end{cases} \tag{33}$$

Αξίζει να καταγράψουμε τον παρακάτω τύπο, που θα αποδειχθεί χρήσιμος στη συνέχεια:

$$\int_0^\infty x^{\alpha-1}e^{-\lambda x} dx = \frac{\Gamma(\alpha)}{\lambda^\alpha}. \tag{34}$$

Οι εκθετικές πυκνότητες είναι ειδικές περιπτώσεις των πυκνοτήτων γάμμα. Πιό συγκεκριμένα, η εκθετική πυκνότητα με παράμετρο λ συμπίπτει με την πυκνότητα γάμμα $\Gamma(1, \lambda)$. Είδαμε ότι η πυκνότητα στην (31) είναι επίσης πυκνότητα γάμμα με παραμέτρους $\alpha = 1/2$ και $\lambda = 1/2\sigma^2$. Μ' άλλα λόγια, *αν η X έχει την κανονική πυκνότητα $n(0, \sigma^2)$, τότε η X^2 έχει την πυκνότητα γάμμα* $\Gamma(1/2, 1/2\sigma^2)$. Εξισώνοντας τις (31) και (33) με $\alpha = 1/2$ και $\lambda = 1/2\sigma^2$ καταλήγουμε στην χρήσιμη παρατήρηση ότι

$$\Gamma(1/2) = \sqrt{\pi}. \tag{35}$$

Μία σημαντική ιδιότητα της συνάρτησης γάμμα είναι η

$$\Gamma(\alpha + 1) = \alpha\Gamma(\alpha), \qquad \alpha > 0. \tag{36}$$

Η σχέση αυτή έπεται από την (32) με απλή εφαρμογή της ολοκλήρωσης κατά μέρη. Γράφουμε

$$\begin{aligned} \Gamma(\alpha + 1) &= \int_0^\infty x^\alpha e^{-x} dx \\ &= -x^\alpha e^{-x} \Big|_0^\infty + \int_0^\infty \alpha x^{\alpha-1} e^{-x} dx \\ &= \alpha\Gamma(\alpha). \end{aligned}$$

Αφού $\Gamma(1) = 1$, από την (36) βλέπουμε ότι για κάθε θετικό ακέραιο n,

$$\Gamma(n) = (n-1)!. \tag{37}$$

Από τις (35), (36) και κάποιες απλοποιήσεις, έπεται ότι αν ο n είναι περιττός θετικός ακέραιος, τότε

$$\Gamma\left(\frac{n}{2}\right) = \frac{\sqrt{\pi}(n-1)!}{2^{n-1}\left(\frac{n-1}{2}\right)!}. \tag{38}$$

Δεν υπάρχουν απλοί τύποι για τη συνάρτηση κατανομής που αντιστοιχεί στην $\Gamma(\alpha, \lambda)$, εκτός αν ο $\alpha = m$ είναι θετικός ακέραιος. Σε αυτήν την περίπτωση μπορούμε, ολοκληρώνοντας κατά μέρη, να δούμε ότι για $x > 0$

$$\begin{aligned} \int_0^x \frac{\lambda^m y^{m-1} e^{-\lambda y}}{(m-1)!} dy &= \frac{-(\lambda y)^{m-1} e^{-\lambda y}}{(m-1)!}\Big|_0^x + \int_0^x \frac{\lambda^{m-1} y^{m-2} e^{-\lambda y}}{(m-2)!} dy \\ &= \int_0^x \frac{\lambda^{m-1} y^{m-2} e^{-\lambda y}}{(m-2)!} dy - \frac{(\lambda x)^{m-1} e^{-\lambda x}}{(m-1)!}, \end{aligned}$$

αρκεί να ισχύει $m \geq 2$. Αν ολοκληρώσουμε κατά μέρη $m-1$ φορές με τον ίδιο τρόπο, και παρατηρήσουμε ότι

$$\int_0^x \lambda e^{-\lambda y} dy = 1 - e^{-\lambda x},$$

καταλήγουμε στον τύπο

$$\int_0^x \frac{\lambda^m y^{m-1} e^{-\lambda y}}{(m-1)!} dy = 1 - \sum_{k=0}^{m-1} \frac{(\lambda x)^k e^{-\lambda x}}{k!}, \qquad x > 0. \tag{39}$$

Αυτός ο τύπος δίνει μία ενδιαφέρουσα σχέση ανάμεσα σε μία τυχαία μεταβλητή X που έχει την πυκνότητα γάμμα $\Gamma(m, \lambda)$ και μία τυχαία μεταβλητή Y που ακολουθεί κατανομή Poisson με παράμετρο λx. Συγκεκριμένα, η (39) μάς λέει ότι

$$P(X \le x) = P(Y \ge m). \tag{40}$$

Αυτή η σχέση έχει σημασία για τη θεωρία των ανελίξεων Poisson, όπως θα δούμε στο Κεφάλαιο 9.

Η ποιοτική συμπεριφορά της πυκνότητας γάμμα περιγράφεται στο Σχήμα 6, και ερμηνεύεται εύκολα με τις μεθόδους του απειροστικού λογισμού. Μία σημαντική ιδιότητα των πυκνοτήτων γάμμα είναι ότι αν X και Y είναι ανεξάρτητες τυχαίες μεταβλητές με πυκνότητες $\Gamma(\alpha_1, \lambda)$ και $\Gamma(\alpha_2, \lambda)$ αντίστοιχα, τότε η $X + Y$ έχει την πυκνότητα γάμμα $\Gamma(\alpha_1 + \alpha_2, \lambda)$. Το αποτέλεσμα αυτό θα αποδειχθεί στο Κεφάλαιο 6. Αυτή και άλλες ιδιότητες των πυκνοτήτων γάμμα τις καθιστούν πολύ εύχρηστο εργαλείο. Υπάρχουν πολλές περιπτώσεις εφαρμογών στις οποίες η πυκνότητα μιάς τυχαίας μεταβλητής X είναι άγνωστη. Μπορεί να είναι γνωστό ότι η X είναι θετική τυχαία μεταβλητή που η πυκνότητά της προσεγγίζεται αρκετά καλά από κάποια πυκνότητα γάμμα με κατάλληλες παραμέτρους. Σε τέτοιες περιπτώσεις, λύνοντας ένα πρόβλημα στο οποίο εμπλέκεται η X με την υπόθεση ότι η X έχει κάποια πυκνότητα γάμμα, θα πάρουμε μία προσέγγιση ή τουλάχιστον μία ιδέα για την πραγματική, αλλά άγνωστη, κατάσταση.

5.4 Αντίστροφες συναρτήσεις κατανομής*

Σημαντικές εφαρμογές των τύπων αλλαγής μεταβλητής της Παραγράφου 5.2.1 προκύπτουν αν η συνάρτηση φ συνδέεται με μία συνάρτηση κατανομής F.

Έστω X συνεχής τυχαία μεταβλητή με συνάρτηση κατανομής F και συνάρτηση πυκνότητας f. Θα εφαρμόσουμε τον τύπο αλλαγής μεταβλητής για τη συνάρτηση $\varphi = F$. Αν $y = F(x)$, τότε $dy/dx = F'(x) = f(x)$, επομένως $dx/dy = 1/f(x)$. Άρα, σύμφωνα με την (18), η τυχαία μεταβλητή $Y = F(X)$ έχει πυκνότητα την g, όπου

$$g(y) = \frac{f(x)}{f(x)} = 1, \qquad 0 < y < 1,$$

και $g(y) = 0$ αλλιώς. Με άλλα λόγια, η τυχαία μεταβλητή $Y = F(X)$ είναι ομοιόμορφα κατανεμημένη στο $(0,1)$. Το αποτέλεσμα αυτό ισχύει ακόμα κι αν η συνάρτηση $\varphi = F$ δεν ικανοποιεί όλες τις υποθέσεις του Θεωρήματος 1. Με απευθείας υπολογισμό μπορούμε να δείξουμε ότι *αν X είναι μία συνεχής τυχαία μεταβλητή με συνάρτηση κατανομής την F, τότε η $F(X)$ είναι ομοιόμορφα κατανεμημένη στο* $(0,1)$. (Αν η F είναι ασυνεχής σε κάποιο σημείο x_0, τότε $P(X = x_0) > 0$, άρα $P(F(X) = F(x_0)) > 0$, και η $F(X)$ δεν μπορεί να είναι ομοιόμορφα κατανεμημένη στο $(0,1)$.)

Μπορούμε να προχωρήσουμε και στην αντίστροφη κατεύθυνση. Έστω F μία συνεχής συνάρτηση κατανομής που είναι γνησίως αύξουσα σε κάποιο διάστημα I, και είναι ίση με μηδέν αριστερά από το I αν το I είναι κάτω φραγμένο, και ίση με 1 δεξιά από το I αν το I είναι άνω φραγμένο. Τότε αν $0 < y < 1$, από το θεώρημα ενδιάμεσης τιμής του απειροστικού λογισμού, υπάρχει μοναδική τιμή του x ώστε $y = F(x)$. Άρα, η $F^{-1}(y), 0 < y < 1$, είναι καλά ορισμένη. Με αυτές τις υποθέσεις, *αν η Y είναι ομοιόμορφα κατανεμημένη στο* $(0,1)$, *τότε η τυχαία μεταβλητή $F^{-1}(Y)$ έχει σαν συνάρτηση κατανομής την F.*

Για να περιγράψουμε αυτό το αποτέλεσμα μπορούμε να χρησιμοποιήσουμε δύο από τα παραδείγματα της Παραγράφου 5.2.1. Στο Παράδειγμα 6 πήραμε τις εκθετικά κατανεμημένες τυχαίες μεταβλητές σαν μετασχηματισμούς ομοιόμορφα κατανεμημένων τυχαίων μεταβλητών. Αφήνουμε τον αναγνώστη να ελέγξει ότι αυτοί οι μετασχηματισμοί προκύπτουν με τη μέθοδο που αναπτύξαμε στην προηγούμενη παράγραφο. Στο Παράδειγμα 11 δείξαμε ότι αν η Θ είναι ομοιόμορφα κατανεμημένη στο $(-\pi/2, \pi/2)$, τότε η $\tan\Theta$ ακολουθεί την κατανομή Cauchy. Έστω ότι η Y είναι ομοιόμορφα κατανεμημένη στο $(0,1)$. Τότε η $\Theta = \pi Y - \pi/2$ είναι ομοιόμορφα κατανεμημένη στο $(-\pi/2, \pi/2)$, άρα η

$$X = \tan\Theta = \tan\left(\pi Y - \frac{\pi}{2}\right)$$

ακολουθεί την κατανομή Cauchy. Το ίδιο ακριβώς θα παίρναμε αν χρησιμοποιούσαμε το αποτέλεσμα της προηγούμενης παραγράφου. Σύμφωνα με το Παράδειγμα 10, η συνάρτηση κατανομής Cauchy δίνεται από την

$$F(x) = \frac{1}{2} + \frac{1}{\pi}\arctan x, \qquad -\infty < x < \infty,$$

και η εξίσωση $y = F(x)$, δηλαδή η

$$y = \frac{1}{2} + \frac{1}{\pi}\arctan x,$$

έχει λύση την

$$x = F^{-1}(y) = \tan\left(\pi y - \frac{\pi}{2}\right).$$

Για διάφορους λόγους θέλουμε να μπορούμε να παράγουμε μία τυχαία μεταβλητή X που να έχει δεδομένη συνάρτηση κατανομής F. Ένας τρόπος για να το κάνουμε είναι να ορίσουμε πρώτα μία ομοιόμορφα κατανεμημένη τυχαία μεταβλητή Y και κατόπιν να θέσουμε $X = F^{-1}(Y)$. Η μέθοδος αυτή είναι εξαιρετικά χρήσιμη αν χρησιμοποιούμε υπολογιστή, αφού υπάρχουν πολύ ικανοποιητικές μέθοδοι «παραγωγής» ομοιόμορφα κατανεμημένων τυχαίων μεταβλητών με τη βοήθεια υπολογιστών. Ας υποθέσουμε για παράδειγμα ότι θέλουμε ένα πρόγραμμα που να παράγει μία τυχαία μεταβλητή που έχει την τυπική κανονική πυκνότητα $n(0,1)$. Θα χρησιμοποιούσαμε ένα υποπρόγραμμα για να παράγουμε μία τυχαία μεταβλητή Y ομοιόμορφα κατανεμημένη στο $(0,1)$, και ένα υποπρόγραμμα που να υπολογίζει τις τιμές της συνάρτησης Φ^{-1}, και μετά θα υπολογίζαμε την $X = \Phi^{-1}(Y)$. Για να παράγουμε μία τυχαία μεταβλητή X με πυκνότητα την $n(\mu, \sigma^2)$ θα θέταμε $X = \mu + \sigma\Phi^{-1}(Y)$.

Οι αντίστροφες συναρτήσεις κατανομής είναι χρήσιμες και για άλλους λόγους. Ας υποθέσουμε ότι η X έχει την κανονική πυκνότητα $n(\mu, \sigma^2)$. Στην Παράγραφο 5.3.1 είδαμε ότι

$$P(X \le b) = \Phi\left(\frac{b-\mu}{\sigma}\right).$$

Ας υποθέσουμε ότι θέλουμε να επιλέξουμε b ώστε $P(X \le b) = 0.9$. Θέλουμε να λύσουμε ως προς b την εξίσωση

$$\Phi\left(\frac{b-\mu}{\sigma}\right) = 0.9.$$

Η λύση δίνεται από την

$$\frac{b-\mu}{\sigma} = \Phi^{-1}(0.9)$$

δηλαδή

$$b = \mu + \sigma\Phi^{-1}(0.9).$$

Από τον Πίνακα I βλέπουμε ότι $\Phi^{-1}(0.9) = 1.28$. Άρα $b = \mu + 1.28\sigma$ και

$$P(X \le \mu + 1.28\sigma) = 0.9.$$

Στη στατιστική, ο αριθμός $b = \mu + 1.28\sigma$ λέγεται το άνω δέκατο της κατανομής $n(\mu, \sigma^2)$.

Έστω F μια συνάρτηση κατανομής που ικανοποιεί όλες τις απαιτήσεις που θέσαμε ώστε η $F^{-1}(y), 0 < y < 1$ να είναι καλά ορισμένη. Τότε, ο $m = F^{-1}(1/2)$ λέγεται *μέσος* της F, οι $F^{-1}(3/4)$ και $F^{-1}(1/4)$ λέγονται άνω και κάτω τεταρτημόρια της F, ο $F^{-1}(0.9)$ λέγεται άνω δεκατημόριο της F, και ο $F^{-1}(k/100)$ είναι το

k-ποσοστιαίο σημείο της F. Οι ορισμοί αυτοί ισχύουν, με κατάλληλη τροποποίηση, για κάθε συνάρτηση κατανομής (ακόμα και για τις διακριτές).

Αν η X έχει συμμετρική πυκνότητα, τότε προφανώς η X έχει μέσο $m = 0$. Για ένα πιό ενδιαφέρον παράδειγμα, ας υποθέσουμε ότι η X είναι εκθετικά κατανεμημένη με παράμετρο λ. Τότε ο μέσος της m δίνεται από την $1 - e^{-\lambda m} = 1/2$, που έχει λύση την $m = \lambda^{-1} \log 2$. Ας υποθέσουμε ότι η X παριστάνει το χρόνο διάσπασης ενός ραδιενεργού σωματιδίου. Τότε, αν έχουμε ένα πολύ μεγάλο πλήθος τέτοιων σωματιδίων περιμένουμε ότι μέχρι τη χρονική στιγμή m θα έχουν διασπαστεί τα μισά από τα σωματίδια. Στη φυσική αυτός ο χρόνος λέγεται *ημιζωή* του σωματιδίου. Αν μετρήσουμε την ημιζωή m μπορούμε να την χρησιμοποιήσουμε για να υπολογίσουμε τον ρυθμό διάσπασης λ, αφού $\lambda = m^{-1} \log 2$.

Για μία τελευταία εφαρμογή των αντίστροφων συναρτήσεων κατανομής, ας υποθέσουμε ότι η X έχει την κανονική πυκνότητα $n(\mu, \sigma^2)$ και ότι ζητάμε $a > 0$ με την ιδιότητα $P(\mu - a \le X \le \mu + a) = 0.9$. Από την (25) πρέπει να λύσουμε ως προς a την εξίσωση

$$\Phi\left(\frac{a}{\sigma}\right) - \Phi\left(-\frac{a}{\sigma}\right) = 0.9.$$

Αφού $\Phi(-x) = 1 - \Phi(x)$ για κάθε x, έχουμε

$$2\Phi\left(\frac{a}{\sigma}\right) - 1 = 0.9,$$

επομένως $a = \sigma\Phi^{-1}(0.95)$. Από τον Πίνακα I βλέπουμε ότι $\Phi^{-1}(0.95) = 1.645$. Με άλλα λόγια,

$$P(\mu - 0.675\sigma \le X \le \mu + 1.645\sigma) = 0.9.$$

Χρησιμοποιώντας την ίδια τεχνική παίρνουμε

$$P(\mu - 0.675\sigma \le X \le \mu + 0.675\sigma) = 0.5.$$

ή, ισοδύναμα,

$$P(|X - \mu| \le 0.675\sigma) = 0.5.$$

Αυτό σημαίνει ότι, αν η X έχει την κανονική πυκνότητα $n(\mu, \sigma^2)$, τότε η X θα διαφέρει από τον μ λιγότερο από 0.675σ με πιθανότητα ενός δευτέρου, και περισσότερο από 0.675σ με πιθανότητα ενός δευτέρου. Αν σκεφτόμαστε τον μ σαν μία πραγματική φυσική ποσότητα και την X σαν μία μέτρηση του μ, τότε η $|X - \mu|$ παριστάνει το σφάλμα μέτρησης. Γιά το λόγο αυτό, ο 0.675σ είναι γνωστός και σαν *πιθανό σφάλμα*.

5.5 Ασκήσεις

1 Έστω X τυχαία μεταβλητή με την ιδιότητα $P(|X - 1| = 2) = 0$. Εκφράστε την $P(|X - 1| \ge 2)$ συναρτήσει της συνάρτησης κατανομής F_X.

2 Επιλέγουμε τυχαία ένα σημείο από το εσωτερικό ενός δίσκου ακτίνας R στο επίπεδο. Έστω X το τετράγωνο της απόστασης του σημείου που επιλέγουμε από το κέντρο του δίσκου. Βρείτε τη συνάρτηση κατανομής της X.

3 Επιλέγουμε ένα σημείο ομοιόμορφα από μία μπάλα ακτίνας R στον τριδιάστατο χώρο. Συμβολίζουμε με X την απόσταση του σημείου που επιλέγουμε από το κέντρο της μπάλας. Βρείτε τη συνάρτηση κατανομής της X.

4 Επιλέγουμε ένα σημείο ομοιόμορφα από το διάστημα $[0, a]$. Συμβολίζουμε με X την απόσταση του σημείου που επιλέγουμε από το 0. Βρείτε τη συνάρτηση κατανομής της X.

5 Επιλέγουμε ένα σημείο ομοιόμορφα από το εσωτερικό τριγώνου που έχει βάση μήκους l και ύψος h προς την βάση. Ορίζουμε σαν X την απόσταση του σημείου που επιλέγουμε από την βάση. Βρείτε τη συνάρτηση κατανομής της X.

6 Θεωρούμε ένα ισόπλευρο τρίγωνο, οι πλευρές του οποίου έχουν μήκος s. Επιλέγουμε ομοιόμορφα ένα σημείο μιάς πλευράς του τριγώνου. Έστω X η τυχαία μεταβλητή που δίνει την απόσταση του σημείου από την απέναντι κορυφή. Βρείτε τη συνάρτηση κατανομής της X.

7 Υποθέτουμε ότι το σημείο (u, v) επιλέγεται ομοιόμορφα από το τετράγωνο $0 \leq u \leq 1, 0 \leq v \leq 1$. Έστω X η τυχαία μεταβλητή που αντιστοιχίζει στο σημείο (u, v) τον αριθμό $u + v$. Βρείτε τη συνάρτηση κατανομής της X.

8 Έστω F η συνάρτηση κατανομής που δίνεται από τον τύπο (3). Βρείτε έναν αριθμό m ώστε $F(m) = 1/2$.

9 Έστω X ο χρόνος διάσπασης ενός ραδιενεργού σωματιδίου. Υποθέτουμε ότι η συνάρτηση κατανομής της X δίνεται από τον τύπο (3). Υποθέτουμε ακόμα ότι το λ είναι τέτοιο ώστε $P(X \geq 0.01) = 1/2$. Βρείτε έναν αριθμό t με την ιδιότητα $P(X \geq t) = 0.9$.

10 Έστω X η τυχαία μεταβλητή της Άσκησης 4. Βρείτε τη συνάρτηση κατανομής της $Y = \min(X, a/2)$.

11 Έστω X μία τυχαία μεταβλητή με συνάρτηση κατανομής την

$$F(x) = \begin{cases} 0 & , x < 0, \\ \frac{x}{3} & , 0 \leq x < 1, \\ \frac{x}{2} & , 1 \leq x < 2, \\ 1 & , x \geq 2. \end{cases}$$

Υπολογίστε τις:
(α) $P(1/2 \leq X \leq 3/2)$,
(β) $P(1/2 \leq X \leq 1)$,
(γ) $P(1/2 \leq X < 1)$,
(δ) $P(1 \leq X \leq 3/2)$,
(ε) $P(1 < X < 2)$.

12 Υποθέτουμε ότι η έννοια της συνάρτησης κατανομής της X ορίζεται με έναν από τους παρακάτω τρόπους:

(α) $P(X < x)$,
(β) $P(X > x)$,
(γ) $P(X \geq x)$.
Σε κάθε περίπτωση, πώς θα τροποποιούνταν οι ιδιότητες (i)-(iv) της Παραγράφου 5.1.1;

13 Επιλέγουμε ομοιόμορφα ένα σημείο στο $(-10, 10)$. Έστω X η τυχαία μεταβλητή που ορίζεται έτσι ώστε να δίνει τη συντεταγμένη του σημείου αν το σημείο ανήκει στο $[-5, 5]$, $X = -5$ αν το σημείο ανήκει στο $(-10, -5)$, και $X = 5$ αν το σημείο ανήκει στο $(5, 10)$. Βρείτε τη συνάρτηση κατανομής της X.

14 Έστω X συνεχής τυχαία μεταβλητή με πυκνότητα f που δίνεται από την

$$f(x) = (1/2)e^{-|x|}, \quad -\infty < x < \infty.$$

Βρείτε την $P(1 \leq |X| \leq 2)$.

15 Έστω F η συνάρτηση κατανομής που ορίζεται από την

$$F(x) = \frac{1}{2} + \frac{x}{2(|x| + 1)}, \quad -\infty < x < \infty.$$

Βρείτε μία συνάρτηση πυκνότητας f για την F. Σε ποιά σημεία x ισχύει $F'(x) = f(x)$;

16 Βρείτε μία συνάρτηση πυκνότητας για την τυχαία μεταβλητή της Άσκησης 3.

17 Βρείτε μία συνάρτηση πυκνότητας για την τυχαία μεταβλητή της Άσκησης 7.

18 Έστω X συνεχής τυχαία μεταβλητή με πυκνότητα f. Βρείτε έναν τύπο για την πυκνότητα της $Y = |X|$.

19 Έστω X και $Y = X^2$ θετικές συνεχείς τυχαίες μεταβλητές με πυκνότητες f και g αντίστοιχα. Βρείτε την f συναρτήσει της g και την g συναρτήσει της f.

20 Υποθέτουμε ότι η X είναι ομοιόμορφα κατανεμημένη στο $(0, 1)$. Βρείτε την πυκνότητα της $Y = X^{1/\beta}$ όπου $\beta \neq 0$.

21 Έστω X θετική συνεχής τυχαία μεταβλητή με πυκνότητα f. Βρείτε έναν τύπο για την πυκνότητα της $Y = 1/(X + 1)$.

22 Έστω X τυχαία μεταβλητή, g συνάρτηση πυκνότητας συνεχούς τύπου, και φ μία παραγωγίσιμη γνησίως αύξουσα συνάρτηση στο $(-\infty, \infty)$. Υποθέτουμε ότι

$$P(X \leq x) = \int_{-\infty}^{\varphi(x)} g(z)dz, \quad -\infty < x < \infty.$$

Δείξτε ότι η τυχαία μεταβλητή $Y = \varphi(X)$ έχει πυκνότητα την g.

23 Υποθέτουμε ότι η τυχαία μεταβλητή X είναι ομοιόμορφα κατανεμημένη στο (a, b). Βρείτε μία γραμμική συνάρτηση φ ώστε η $Y = \varphi(X)$ να είναι ομοιόμορφα κατανεμημένη στο $(0, 1)$.

24 Έστω ότι η X έχει εκθετική πυκνότητα με παράμετρο λ. Βρείτε την πυκνότητα της $Y = cX$, όπου $c > 0$.

25 Έστω $g(x) = x(1-x)^2$, $0 \leq x \leq 1$, και $g(x) = 0$ αλλιώς. Κανονικοποιήστε την g ώστε να γίνει πυκνότητα.

26 Υποθέτουμε ότι η X έχει πυκνότητα Cauchy. Βρείτε την πυκνότητα της $Y = a + bX$, $b \neq 0$.

27 Συμβολίζουμε με X το ημίτονο μιάς γωνίας που επιλέγεται τυχαία από το $(-\pi/2, \pi/2)$. Βρείτε την πυκνότητα και τη συνάρτηση κατανομής της X.

28 Έστω X συνεχής τυχαία μεταβλητή που έχει συμμετρική πυκνότητα f. Υποθέτουμε επίσης ότι η X^2 έχει εκθετική πυκνότητα με παράμετρο λ. Βρείτε την f.

29 Έστω X συνεχής τυχαία μεταβλητή με συνάρτηση κατανομής F και συνάρτηση πυκνότητας f. Λέμε ότι η f είναι συμμετρική ως προς a αν $f(a+x) = f(a-x)$, $-\infty < x < \infty$. Βρείτε ισοδύναμες συνθήκες με βάση την τυχαία μεταβλητή X και με βάση τη συνάρτηση κατανομής F.

30 Η συνάρτηση σφάλματος ορίζεται από την

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-y^2} dy, \qquad -\infty < x < \infty.$$

Εκφράστε την Φ συναρτήσει της συνάρτησης σφάλματος.

31 Υποθέτουμε ότι η X έχει την κανονική πυκνότητα $n(0, \sigma^2)$. Βρείτε την πυκνότητα της $Y = |X|$.

32 Υποθέτουμε ότι η X έχει την κανονική πυκνότητα $n(\mu, \sigma^2)$. Βρείτε την πυκνότητα της $Y = e^X$. Η πυκνότητα αυτή λέγεται *λογαριθμική πυκνότητα.*

33 Υποθέτουμε ότι η X είναι κανονικά κατανεμημένη με παραμέτρους μ και σ^2. Βρείτε την $P(|X - \mu| \leq \sigma)$.

34 Υποθέτουμε ότι η X είναι κανονικά κατανεμημένη με παραμέτρους μ και σ^2. Βρείτε αριθμούς a και b ώστε η $a + bX$ να ακολουθεί την τυπική κανονική κατανομή.

35 Υποθέτουμε ότι η X ακολουθεί κανονική κατανομή με παραμέτρους $\mu = 0$ και $\sigma^2 = 4$. Με βάση την X ορίζουμε μία τυχαία μεταβλητή Y που παίρνει ακέραιες τιμές, ως εξής: θέτουμε $Y = m$ αν $m - 1/2 \leq X < m + 1/2$, όπου m είναι ένας ακέραιος με $-5 \leq m \leq 5$, $Y = -6$ αν $X < -5.5$, και $Y = 6$ αν $X \geq 5.5$. Βρείτε την f_Y και σχεδιάστε την.

36 Υποθέτουμε ότι το βάρος ενός ατόμου που επιλέγεται τυχαία από έναν πληθυσμό ακολουθεί κανονική κατανομή με παραμέτρους μ και σ. Υποθέτουμε ακόμα ότι $P(X \leq 160) = 1/2$ και $P(X \leq 140) = 1/4$. Βρείτε τα μ και σ, καθώς και την $P(X \geq 200)$. Ποιό είναι το ποσοστό των ατόμων που ζυγίζουν τουλάχιστον 100 κιλά, που το βάρος τους υπερβαίνει τα 110 κιλά;

37 Έστω t_p ο αριθμός για τον οποίο $\Phi(t_p) = p$, $0 < p < 1$. Υποθέτουμε ότι η X έχει την κανονική πυκνότητα $n(\mu, \sigma^2)$. Δείξτε ότι αν $0 < p_1 < p_2 < 1$, τότε

$$P(\mu + t_{p_1}\sigma \leq X \leq \mu + t_{p_2}\sigma) = p_2 - p_1.$$

38 Δίνεται μεγάλο πλήθος από όμοια ραδιενεργά σωματίδια, και υποθέτουμε ότι έχουν χρόνους διάσπασης που ακολουθούν εκθετική κατανομή με κάποια παράμετρο λ. Αν τα μισά σωματίδια διασπώνται στην διάρκεια του πρώτου δευτερολέπτου, πόσος χρόνος θα χρειαστεί για τη διάσπαση του 75% των σωματιδίων;

39 Υποθέτουμε ότι η X ακολουθεί εκθετική κατανομή με παράμετρο λ. Ορίζουμε μία τυχαία μεταβλητή Y με ακέραιες τιμές μέσω της X, θέτοντας $Y = m$ αν $m \leq X \leq m + 1$, όπου m είναι ένας μη αρνητικός ακέραιος. Ποιά είναι η κατανομή της Y;

40 Έστω T θετική συνεχής τυχαία μεταβλητή που δίνει το χρόνο αποτυχίας κάποιου συστήματος, έστω F η συνάρτηση κατανομής της T, και ας υποθέσουμε ότι $F(t) < 1$, $0 < t < \infty$. Τότε μπορούμε να γράψουμε $F(t) = 1 - e^{-G(t)}$, $t > 0$. Υποθέτουμε ότι η $G'(t) = g(t)$ υπάρχει για κάθε $t > 0$.
(α) Δείξτε ότι η T έχει πυκνότητα f που δίνεται από την

$$\frac{f(t)}{1 - F(t)} = g(t), \qquad 0 < t < \infty.$$

Η συνάρτηση g είναι γνωστή σαν «ρυθμός αποτυχίας», γιατί ευριστικά,

$$P(t \leq T \leq t + dt | T > t) = \frac{f(t)dt}{1 - F(t)} = g(t)dt.$$

(β) Δείξτε ότι αν $s > 0$ και $t > 0$,

$$P(T > t + s | T > t) = e^{-\int_t^{t+s} g(u)du}.$$

(γ) Δείξτε ότι το σύστημα βελτιώνεται με το χρόνο (δηλαδή, για σταθερό s η ποσότητα στο (β) αυξάνει ως προς t) αν η g είναι φθίνουσα συνάρτηση, ενώ το σύστημα παρακμάζει με το χρόνο αν η g είναι αύξουσα συνάρτηση.
(δ) Δείξτε ότι

$$\int_0^\infty g(u)du = \infty.$$

(ε) Πώς συμπεριφέρεται η g αν η T ακολουθεί εκθετική κατανομή;
(στ) Αν $G(t) = \lambda t^\alpha, t > 0$, για ποιές τιμές του α το σύστημα βελτιώνεται, παρακμάζει ή μένει σταθερό με το χρόνο;

41 Υποθέτουμε ότι η X έχει την πυκνότητα γάμμα $\Gamma(\alpha, \lambda)$. Βρείτε την πυκνότητα της $Y = cX$, όπου $c > 0$.

42 Δείξτε ότι αν $\alpha > 1$, η πυκνότητα γάμμα έχει μέγιστο στο $(\alpha - 1)/\lambda$.

43 Υποθέτουμε ότι η X έχει την πυκνότητα γάμμα $\Gamma(\alpha, \lambda)$. Βρείτε την πυκνότητα της $Y = \sqrt{X}$.

44 Υποθέτουμε ότι η Y είναι ομοιόμορφα κατανεμημένη στο $(0, 1)$. Βρείτε μία συνάρτηση φ τέτοια ώστε η $X = \varphi(Y)$ να έχει πυκνότητα την $f(x) = 2x, 0 \leq x \leq 1$, και $f(x) = 0$ αλλιώς.

45 Υποθέτουμε ότι η Y είναι ομοιόμορφα κατανεμημένη στο $(0, 1)$. Βρείτε μία συνάρτηση φ τέτοια ώστε η $\varphi(Y)$ να έχει την πυκνότητα γάμμα $\Gamma(1/2, 1/2)$. *Υπόδειξη:* Χρησιμοποιήστε το Παράδειγμα 12.

46 Βρείτε την $\Phi^{-1}(t)$ για $t = 0.1, 0.2, \ldots, 0.9$ και χρησιμοποιώντας αυτές τις τιμές σχεδιάστε την γραφική παράσταση της Φ^{-1}.

47 Έστω ότι η X έχει την κανονική πυκνότητα $n(\mu, \sigma^2)$. Βρείτε το άνω τεταρτημόριο της X.

48 Έστω ότι η X έχει την πυκνότητα Cauchy. Βρείτε το άνω τεταρτημόριο της X.

49 Υποθέτουμε ότι η X έχει κανονική πυκνότητα με παραμέτρους μ και $\sigma^2 = 0.25$. Βρείτε μία σταθερά c τέτοια ώστε

$$P(|X - \mu| \leq c) = 0.9.$$

50 Έστω X τυχαία μεταβλητή με ακέραιες τιμές και συνάρτηση κατανομής F, και έστω ότι η Y είναι ομοιόμορφα κατανεμημένη στο $(0, 1)$. Ορίζουμε μία τυχαία μεταβλητή Z με ακέραιες τιμές, θέτοντας $Z = m$ αν $F(m-1) < Y \leq F(m)$, $m \in \mathbb{Z}$. Δείξτε ότι η Z έχει την ίδια πυκνότητα με την X.

Κεφάλαιο 6

Πολυδιάστατες κατανομές

Στις πρώτες τρείς παραγράφους αυτού του Κεφαλαίου θα μελετήσουμε ένα ζεύγος συνεχών τυχαίων μεταβλητών X και Y, και κάποιες από τις ιδιότητές τους. Στο υπόλοιπο του Κεφαλαίου θα μελετήσουμε επεκτάσεις από τις δύο στις n τυχαίες μεταβλητές $X_1, X_2, \ldots, X_n$.

6.1 Ιδιότητες των διδιάστατων κατανομών

Έστω X και Y δύο τυχαίες μεταβλητές που ορίζονται στον ίδιο χώρο πιθανότητας. Η *από κοινού συνάρτηση κατανομής τους* F ορίζεται από την

$$F(x, y) = P(X \leq x, Y \leq y), \quad -\infty < x, y < \infty.$$

Για να δούμε ότι η F είναι καλά ορισμένη, παρατηρούμε ότι αφού οι X και Y είναι τυχαίες μεταβλητές, τα σύνολα $\{\omega | X(\omega) \leq x\}$ και $\{\omega | Y(\omega) \leq y\}$ είναι ενδεχόμενα. Η τομή τους $\{\omega | X(\omega) \leq x, Y(\omega) \leq y\}$ είναι επίσης ενδεχόμενο, άρα η πιθανότητά του είναι καλά ορισμένη.

Η από κοινού συνάρτηση κατανομής είναι χρήσιμη αν θέλουμε να υπολογίσουμε την πιθανότητα το ζεύγος (X, Y) να ανήκει σε ένα ορθογώνιο στο επίπεδο. Ας θεωρήσουμε το ορθογώνιο

$$R = \{(x, y) | a < x \leq b, c < y \leq d\},$$

όπου $a < b$ και $c < d$. Τότε,

$$\begin{aligned} (1) \quad P((X, Y) \in R) &= P(a < X \leq b, c < Y \leq d) \\ &= F(b, d) - F(a, d) - F(b, c) + F(a, c). \end{aligned}$$

Για να επαληθεύσουμε την (1) παρατηρούμε ότι

$$\begin{aligned} P(a < X \leq b, Y \leq d) &= P(X \leq b, Y \leq d) - P(X \leq a, Y \leq d) \\ &= F(b,d) - F(a,d). \end{aligned}$$

Όμοια,

$$P(a < X \leq b, Y \leq c) = F(b,c) - F(a,c).$$

Άρα,

$$\begin{aligned} P(a < X \leq b, c < Y \leq d) &= P(a < X \leq b, Y \leq d) - P(a < X \leq b, Y \leq c) \\ &= (F(b,d) - F(a,d)) - (F(b,c) - F(a,c)), \end{aligned}$$

και η (1) ισχύει όπως ισχυριστήκαμε.

Οι μονοδιάστατες συναρτήσεις κατανομής F_X και F_Y που ορίζονται από τις

$$F_X(x) = P(X \leq x) \quad \text{και} \quad F_Y(y) = P(Y \leq y)$$

λέγονται *περιθώριες συναρτήσεις κατανομής* των X και Y. Η σχέση τους με την από κοινού συνάρτηση κατανομής F δίνεται από τις

$$F_X(x) = F(x, \infty) = \lim_{y\to\infty} F(x,y)$$

και

$$F_Y(y) = F(\infty, y) = \lim_{x\to\infty} F(x,y).$$

Αν υπάρχει μία μη αρνητική συνάρτηση f ώστε

$$F(x,y) = \int_{-\infty}^{x} \left(\int_{-\infty}^{y} f(u,v)dv \right) du, \quad -\infty < x, y < \infty, \tag{2}$$

τότε η f λέγεται από κοινού συνάρτηση πυκνότητας (συνεχούς τύπου) για τη συνάρτηση κατανομής F ή το ζεύγος των τυχαίων μεταβλητών X, Y. Αν δεν αναφέρουμε ρητά κάτι άλλο, με τον όρο συναρτήσεις πυκνότητας σε αυτό το Κεφάλαιο θα εννοούμε συναρτήσεις πυκνότητας συνεχούς τύπου και όχι διακριτές συναρτήσεις πυκνότητας.

Αν η F έχει πυκνότητα την f, τότε μπορούμε με τη βοήθεια της f να ξαναγράψουμε την (1) στη μορφή

$$P(a < X \leq b, c < Y \leq d) = \int_{a}^{b} \left(\int_{c}^{d} f(x,y)dy \right) dx. \tag{3}$$

Χρησιμοποιώντας τις ιδιότητες του ολοκληρώματος και τον ορισμό του χώρου πιθανότητας, μπορούμε να δείξουμε ότι η σχέση

$$P((X,Y) \in A) = \int\int_A f(x,y)dxdy \tag{4}$$

ισχύει για όλα τα υποσύνολα A του επιπέδου που μελετάμε στον απειροστικό λογισμό. Παίρνοντας σαν A ολόκληρο το επίπεδο, από την (4) παίρνουμε

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1. \tag{5}$$

Από την (4) βλέπουμε ακόμα ότι

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x}\left(\int_{-\infty}^{\infty} f(u,y)dy\right)du,$$

επομένως η F_X έχει περιθώρια πυκνότητα f_X την

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$$

η οποία ικανοποιεί την

$$F_X(x) = \int_{-\infty}^{x} f_X(u)du.$$

Όμοια, η F_Y έχει περιθώρια πυκνότητα f_Y την

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx.$$

Όπως και στη μονοδιάστατη περίπτωση, η f δεν ορίζεται μονοσήμαντα από την (2). Μπορούμε να αλλάξουμε την τιμή της f σε πεπερασμένα το πλήθος σημεία ή και στην ένωση πεπερασμένων το πλήθος λείων καμπυλών του επιπέδου χωρίς να επηρεάσουμε το ολοκλήρωμα της f στα υποσύνολα του επιπέδου. Πάλι, όπως και στη μονοδιάστατη περίπτωση, η F καθορίζει την f στα σημεία συνέχειας της f. Η παρατήρηση αυτή είναι συνέπεια της (3).

Παραγωγίζοντας την (2) και εφαρμόζοντας τους κανόνες του απειροστικού λογισμού παίρνουμε

$$\begin{aligned}\frac{\partial}{\partial y}F(x,y) &= \int_{-\infty}^{x}\left(\frac{\partial}{\partial y}\int_{-\infty}^{y} f(u,v)dv\right)du \\ &= \int_{-\infty}^{x} f(u,y)du\end{aligned}$$

και

$$\frac{\partial^2}{\partial x \partial y}F(x,y) = f(x,y). \tag{6}$$

Με κάποιες επιπλέον υποθέσεις, μπορούμε να δικαιολογήσουμε αυτές τις πράξεις και να δείξουμε ότι η (6) ισχύει στα σημεία συνέχειας της f. Σε συγκεκριμένες περιπτώσεις, αντί να ελέγχουμε ότι ισχύουν τα βήματα που οδηγούν στην (6), είναι συνήθως απλούστερο να δείξουμε ότι η συνάρτηση f που ορίζεται από την (6) ικανοποιεί την (2).

Παράδειγμα 1. Θα περιγράψουμε τους παραπάνω ορισμούς και τύπους κοιτάζοντας ξανά το Παράδειγμα 1 του Κεφαλαίου 5. Υπενθυμίζουμε ότι το παράδειγμα αφορούσε την ομοιόμορφη επιλογή ενός σημείου από δίσκο ακτίνας R. Συμφωνούμε ότι τα σημεία του επιπέδου προσδιορίζονται από τις συντεταγμένες τους (x, y). Τότε, ο δίσκος περιγράφεται από την

$$\{(x, y) \mid x^2 + y^2 \leq R^2\}.$$

Έστω X και Y οι τυχαίες μεταβλητές που δίνουν τις τυχαίες συντεταγμένες του σημείου που επιλέγουμε. Αφού υποθέσαμε ότι η επιλογή του σημείου γίνεται ομοιόμορφα, οι X και Y έχουν από κοινού πυκνότητα f που δίνεται από την

$$(7) \qquad f(x, y) = \begin{cases} \frac{1}{\pi R^2} & , x^2 + y^2 \leq R^2, \\ 0 & , \text{αλλιώς.} \end{cases}$$

Τότε, για κάθε λογικό υποσύνολο A του δίσκου,

$$\begin{aligned} P((X, Y) \in A) &= \int\int_A f(x, y)dxdy \\ &= \frac{E(A)}{\pi R^2}, \end{aligned}$$

όπου $E(A)$ το εμβαδόν του A, αφού η επιλογή γίνεται ομοιόμορφα. Η περιθώρια πυκνότητα f_X δίνεται από την

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy = \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \frac{1}{\pi R^2} dy = \frac{2\sqrt{R^2 - x^2}}{\pi R^2}$$

αν $-R < x < R$ και $f_X(x) = 0$ αλλιώς. Η περιθώρια πυκνότητα $f_Y(y)$ δίνεται από τον ίδιο τύπο, με αντικατάσταση του x από το y.

Οι μεταβλητές X και Y λέγονται *ανεξάρτητες τυχαίες μεταβλητές* αν για κάθε $a < b$ και $c < d$, ισχύει

$$(8) \qquad P(a < X \leq b, c < Y \leq d) = P(a < X \leq b)P(c < Y \leq d).$$

Θέτοντας $a = c = -\infty$, $b = x$, και $d = y$, βλέπουμε ότι αν οι X και Y είναι ανεξάρτητες, τότε

$$(9) \qquad F(x, y) = F_X(x)F_Y(y), \qquad -\infty < x, y < \infty.$$

Αντίστροφα, από την (9) έπεται ότι οι X και Y είναι ανεξάρτητες. Γιατί αν ισχύει η (9), τότε από την (1) το αριστερό μέλος της (8) είναι ίσο με

$$\begin{aligned} F(b, d) - F(a, d) &- F(b, c) + F(a, c) \\ &= F_X(b)F_Y(d) - F_X(a)F_Y(d) - F_X(b)F_Y(c) + F_X(a)F_Y(c) \\ &= (F_X(b) - F_X(a))(F_Y(d) - F_Y(c)) \\ &= P(a < X \leq b)P(c < Y \leq d). \end{aligned}$$

Πιό γενικά, αποδεικνύεται ότι αν οι X και Y είναι ανεξάρτητες και τα A, B είναι ενώσεις πεπερασμένων ή αριθμησίμων το πλήθος διαστημάτων, τότε

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$$

ή, αλλιώς, τα ενδεχόμενα

$$\{\omega \mid X(\omega) \in A\} \quad \text{και} \quad \{\omega \mid Y(\omega) \in B\}$$

είναι ανεξάρτητα ενδεχόμενα.

Έστω X και Y τυχαίες μεταβλητές με περιθώριες πυκνότητες τις f_X και f_Y. Τότε, οι X και Y είναι ανεξάρτητες αν και μόνο αν η συνάρτηση f που ορίζεται από την

$$f(x,y) = f_X(x)f_Y(y), \qquad -\infty < x, y < \infty,$$

είναι από κοινού πυκνότητα των X και Y. Αυτό προκύπτει από τον ορισμό της ανεξαρτησίας και τον τύπο

$$F_X(x)F_Y(y) = \int_{-\infty}^{x} \left(\int_{-\infty}^{y} f_X(u)f_Y(v)dv \right) du.$$

Σαν παράδειγμα εξαρτημένων τυχαίων μεταβλητών, θεωρούμε τις X και Y του Παραδείγματος 1. Τότε, αν $-R < x < R$ και $-R < y < R$,

$$f_X(x)f_Y(y) = \frac{4\sqrt{R^2 - x^2}\sqrt{R^2 - y^2}}{\pi^2 R^4}, \tag{10}$$

το οποίο δεν συμφωνεί με την από κοινού πυκνότητα αυτών των μεταβλητών όταν $x = 0, y = 0$. Αφού το $(0,0)$ είναι σημείο συνέχειας των συναρτήσεων που ορίζονται από τις (7) και (10), έπεται ότι οι X και Y είναι εξαρτημένες τυχαίες μεταβλητές. Αυτό συμφωνεί με τη διαίσθησή μας, γιατί αν η X είναι κοντά στο R, τότε η Y πρέπει να είναι κοντά στο μηδέν, δηλαδή η πληροφορία που έχουμε για την X μάς δίνει πληροφορία για την Y.

Μπορούμε να ορίσουμε τις συναρτήσεις πυκνότητας και απευθείας, όπως έχουμε δεί σε άλλες περιπτώσεις. Μία *διδιάστατη συνάρτηση πυκνότητας* f είναι μία μη αρνητική συνάρτηση στο $\mathbb{R}^2$ με την ιδιότητα

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)dxdy = 1.$$

Σε κάθε διδιάστατη συνάρτηση πυκνότητας f αντιστοιχούν ένας χώρος πιθανότητας και ένα ζεύγος τυχαίων μεταβλητών X και Y που ορίζονται σε αυτό το χώρο και έχουν από κοινού πυκνότητα την f.

Ο ευκολότερος τρόπος για να κατασκευάσουμε διδιάστατες συναρτήσεις πυκνότητας είναι να ξεκινήσουμε με δύο μονοδιάστατες πυκνότητες f_1 και f_2 και να ορίσουμε τη συνάρτηση f θέτοντας

$$f(x,y) = f_1(x)f_2(y), \qquad -\infty < x, y < \infty. \tag{11}$$

Τότε, η f είναι διδιάστατη συνάρτηση πυκνότητας αφού είναι προφανώς μη αρνητική, και ικανοποιεί την

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = \int_{-\infty}^{\infty} f_1(x)dx \int_{-\infty}^{\infty} f_2(y)dy = 1.$$

Αν οι τυχαίες μεταβλητές X και Y έχουν αυτήν την f σαν από κοινού πυκνότητα, τότε οι X και Y είναι ανεξάρτητες και έχουν περιθώριες πυκνότητες $f_X = f_1$ και $f_Y = f_2$.

Για ένα παράδειγμα εφαρμογής της (11), ας υποθέσουμε ότι f_1 και f_2 είναι η τυπική κανονική πυκνότητα $n(0,1)$. Τότε η f δίνεται από την

$$f(x,y) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}\frac{1}{\sqrt{2\pi}}e^{-y^2/2},$$

ή, ισοδύναμα,

$$f(x,y) = \frac{1}{2\pi}e^{-(x^2+y^2)/2}, \qquad -\infty < x,y < \infty. \tag{12}$$

Η πυκνότητα που δίνεται από την (12) λέγεται *τυπική διδιάστατη κανονική πυκνότητα*. Στο επόμενο παράδειγμά μας, θα τροποποιήσουμε ελαφρά το δεξιό μέλος της (12) για να πάρουμε μία από κοινού συνάρτηση πυκνότητας που αντιστοιχεί στην περίπτωση δύο εξαρτημένων τυχαίων μεταβλητών με κανονικές περιθώριες πυκνότητες.

Παράδειγμα 2. Έστω ότι οι X και Y έχουν από κοινού συνάρτηση πυκνότητας την

$$f(x,y) = ce^{-(x^2-xy+y^2)/2}, \qquad -\infty < x,y < \infty,$$

όπου c είναι μία θετική σταθερά την οποία θα προσδιορίσουμε στην πορεία της συζήτησης. Πρώτα «συμπληρώνουμε το τετράγωνο» ως προς τους όρους που περιέχουν το y, και ξαναγράφουμε την f στη μορφή

$$f(x,y) = ce^{-[(y-x/2)^2+3x^2/4]/2}, \qquad -\infty < x,y < \infty,$$

και παρατηρούμε ότι

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = ce^{-3x^2/8}\int_{-\infty}^{\infty} e^{-(y-x/2)^2/2}dy.$$

Κάνοντας την αλλαγή μεταβλητής $u = y - x/2$, βλέπουμε ότι

$$\int_{-\infty}^{\infty} e^{-(y-x/2)^2/2}dy = \int_{-\infty}^{\infty} e^{-u^2/2}du = \sqrt{2\pi}.$$

Επομένως,

$$f_X(x) = c\sqrt{2\pi}e^{-3x^2/8}.$$

Είναι τώρα σαφές ότι η f_X είναι η κανονική πυκνότητα $n(0,\sigma^2)$ με $\sigma^2 = 4/3$, οπότε

$$c\sqrt{2\pi} = \frac{1}{\sigma\sqrt{2\pi}} = \frac{\sqrt{3}}{2\sqrt{2\pi}}$$

δηλαδή $c = \sqrt{3}/4\pi$. Επομένως,

$$f(x,y) = \frac{\sqrt{3}}{4\pi} e^{-(x^2-xy+y^2)/2}, \qquad -\infty < x, y < \infty.$$

Οι παραπάνω υπολογισμοί δείχνουν τώρα ότι η f_X είναι η κανονική πυκνότητα $n(0, 4/3)$. Με ανάλογο τρόπο, μπορούμε να δείξουμε ότι και η f_Y είναι η $n(0, 4/3)$. Αφού $f(x,y) \neq f_X(x) f_Y(y)$, είναι φανερό ότι οι X και Y είναι εξαρτημένες.

6.2 Κατανομή αθροισμάτων και πηλίκων

Έστω X και Y τυχαίες μεταβλητές με από κοινού πυκνότητα την f. Πολύ συχνά έχουμε μία τυχαία μεταβλητή Z που ορίζεται με βάση τις X και Y και θέλουμε να υπολογίσουμε την πυκνότητα της Z. Ας υποθέσουμε ότι η Z δίνεται από την $Z = \phi(X,Y)$, όπου ϕ είναι μία πραγματική συνάρτηση που το πεδίο ορισμού της περιέχει το πεδίο τιμών του (X,Y). Για σταθερό z το ενδεχόμενο $\{Z \leq z\}$ συμπίπτει με το ενδεχόμενο $\{(X,Y) \in A_z\}$, όπου A_z είναι το υποσύνολο του $\mathbb{R}^2$ που ορίζεται από την

$$A_z = \{(x,y) \mid \phi(x,y) \leq z\}.$$

Άρα,

$$\begin{aligned} F_Z(z) &= P(Z \leq z) \\ &= P((X,Y) \in A_z) \\ &= \int\int_{A_z} f(x,y)\,dx\,dy. \end{aligned}$$

Αν μπορέσουμε να βρούμε μία μη αρνητική συνάρτηση g με την ιδιότητα

$$\int\int_{A_z} f(x,y)\,dx\,dy = \int_{-\infty}^{z} g(v)\,dv, \qquad -\infty < z < \infty,$$

τότε η g είναι υποχρεωτικά πυκνότητα για την Z. Θα χρησιμοποιήσουμε αυτή τη μέθοδο για να υπολογίσουμε πυκνότητες για τις $X+Y$ και Y/X.

6.2.1 Κατανομή αθροισμάτων

Θέτουμε $Z = X + Y$. Τότε, το

$$A_z = \{(x,y) \mid x + y \leq z\}$$

είναι το ημιεπίπεδο αριστερά της ευθείας $x + y = z$, όπως φαίνεται στο Σχήμα 1. Επομένως,

$$F_Z(z) = \int\int_{A_z} f(x,y)dxdy = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{z-x} f(x,y)dy \right) dx.$$

Κάνουμε την αλλαγή μεταβλητής $y = v - x$ στο εσωτερικό ολοκλήρωμα. Τότε

$$\begin{aligned} F_Z(z) &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{z} f(x, v - x)dv \right) dx \\ &= \int_{-\infty}^{z} \left(\int_{-\infty}^{\infty} f(x, v - x)dx \right) dv, \end{aligned}$$

αν εναλλάξουμε τη σειρά της ολοκλήρωσης. Επομένως, η πυκνότητα της $Z = X + Y$ δίνεται από την

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f(x, z - x)dx, \qquad -\infty < z < \infty. \tag{14}$$

Στις βασικές εφαρμογές της (14), οι X και Y είναι ανεξάρτητες και η (14) παίρνει τη μορφή

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z - x)dx, \qquad -\infty < z < \infty. \tag{15}$$

Αν X και Y είναι μη αρνητικές ανεξάρτητες τυχαίες μεταβλητές, τότε $f_{X+Y}(z) = 0$ αν $z \le 0$, και

$$f_{X+Y}(z) = \int_{0}^{z} f_X(x)f_Y(z - x)dx, \qquad 0 < z < \infty. \tag{16}$$

Το δεξιό μέλος της (15) μας προτείνει μία μέθοδο για να παίρνουμε νέες πυκνότητες. Αν μάς δώσουν δύο μονοδιάστατες πυκνότητες f και g, η συνάρτηση h που ορίζεται από την

$$h(z)=\int_{-\infty}^{\infty} f(x)g(z-x)dx, \quad -\infty<z<\infty,$$

είναι μία μονοδιάστατη συνάρτηση πυκνότητας, που λέγεται *συνέλιξη* των f και g. Δηλαδή, η πυκνότητα του αθροίσματος δύο ανεξάρτητων τυχαίων μεταβλητών είναι η συνέλιξη των πυκνοτήτων τους.

Παράδειγμα 3. Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν εκθετική κατανομή με παράμετρο λ. Βρείτε την κατανομή της $X+Y$.

Η πυκνότητα της X είναι η $f_X(x)=\lambda e^{-\lambda x}$ αν $x\geq 0$ και $f_X(x)=0$ αν $x<0$. Η πυκνότητα της Y είναι η ίδια. Άρα, $f_{X+Y}(z)=0$ αν $z\leq 0$, και από την (16), αν $z>0$ τότε

$$\begin{aligned} f_{X+Y}(z) &= \int_0^z \lambda e^{-\lambda x}\lambda e^{-\lambda(z-x)}dx \\ &= \lambda^2 e^{-\lambda z}\int_0^z dx=\lambda^2 z e^{-\lambda z}. \end{aligned}$$

Βλέπουμε λοιπόν ότι η $X+Y$ έχει την πυκνότητα γάμμα $\Gamma(2,\lambda)$.

Παράδειγμα 4. Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές, ομοιόμορφα κατανεμημένες στο $(0,1)$. Βρείτε την πυκνότητα της $X+Y$.

Η πυκνότητα της X είναι η $f_X(x)=1$ αν $0<x<1$ και $f_X(x)=0$ αλλιώς. Η πυκνότητα της Y είναι η ίδια. Άρα, $f_{X+Y}(z)=0$ αν $z\leq 0$. Στην περίπτωση $z>0$ εφαρμόζουμε την (16). Η προς ολοκλήρωση συνάρτηση $f_X(x)f_Y(z-x)$ παίρνει μόνο τις τιμές 0 και 1. Παίρνει την τιμή 1 αν τα x και z είναι τέτοια ώστε $0\leq x\leq 1$ και $0\leq z-x\leq 1$. Αν $0\leq z\leq 1$, η τιμή της συνάρτησης είναι 1 στο σύνολο $0\leq x\leq z$ και μηδέν αλλού. Επομένως, από την (16) βλέπουμε ότι

$$f_{X+Y}(z)=z, \quad 0\leq z\leq 1.$$

Αν $1<z\leq 2$, η τιμή της συνάρτησης είναι 1 στο σύνολο $z-1\leq x\leq 1$ και μηδέν αλλού. Επομένως, από την (16)

$$f_{X+Y}(z)=2-z, \quad 1<z\leq 2.$$

Αν $2<z<\infty$, η συνάρτηση στην (16) είναι ταυτοτικά μηδέν, άρα

$$f_{X+Y}(z)=0, \quad 2<z<\infty.$$

Για να συνοψίσουμε,

$$f_{X+Y}(z)=\begin{cases} z & ,\ 0\leq z\leq 1, \\ 2-z & ,\ 1<z\leq 2, \\ 0 & ,\ \text{αλλιώς}. \end{cases}$$

Το γράφημα της f δίνεται στο Σχήμα 2. Μπορούμε επίσης να βρούμε την πυκνότητα της $X+Y$ υπολογίζοντας το εμβαδόν του συνόλου

$$A_z = \{(x,y) \mid 0 \le x \le 1, 0 \le y \le 1, x+y \le z\}$$

(δείτε το Σχήμα 3) και παραγωγίζοντας το αποτέλεσμα ως προς z.

Το Παράδειγμα 3 έχει μία σημαντική γενίκευση, η οποία διατυπώνεται ως εξής.

Θεώρημα 1. *Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές. Υποθέτουμε ότι η X έχει την πυκνότητα γάμμα $\Gamma(\alpha_1, \lambda)$ και η Y έχει την πυκνότητα γάμμα $\Gamma(\alpha_2, \lambda)$. Τότε η $X+Y$ έχει την πυκνότητα γάμμα*

$$\Gamma(\alpha_1 + \alpha_2, \lambda).$$

Απόδειξη: Παρατηρούμε ότι οι X και Y είναι θετικές τυχαίες μεταβλητές και ότι

$$f_X(x) = \frac{\lambda^{\alpha_1} x^{\alpha_1 - 1} e^{-\lambda x}}{\Gamma(\alpha_1)}, \qquad x > 0,$$

και

$$f_Y(y) = \frac{\lambda^{\alpha_2} x^{\alpha_2 - 1} e^{-\lambda y}}{\Gamma(\alpha_2)}, \qquad y > 0.$$

Άρα $f_{X+Y}(z) = 0$ αν $z \le 0$, και, από την (16), αν $z > 0$

$$f_{X+Y}(z) = \frac{\lambda^{\alpha_1+\alpha_2} e^{-\lambda z}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \int_0^z x^{\alpha_1 - 1}(z-x)^{\alpha_2 - 1} dx.$$

Στο παραπάνω ολοκλήρωμα κάνουμε την αλλαγή μεταβλητής $x = zu$ (με $dx = zdu$) και παίρνουμε

$$(17) \qquad f_{X+Y}(z) = c\lambda^{\alpha_1+\alpha_2} z^{\alpha_1+\alpha_2-1} e^{-\lambda z}, \qquad z > 0,$$

όπου

$$(18) \qquad c = \frac{\int_0^1 u^{\alpha_1-1}(1-u)^{\alpha_2-1}du}{\Gamma(\alpha_1)\Gamma(\alpha_2)}.$$

Η σταθερά c προσδιορίζεται από το γεγονός ότι το ολοκλήρωμα της f_{X+Y} πρέπει να είναι ίσο με 1. Από την (17) και τον ορισμό των πυκνοτήτων γάμμα, είναι φανερό ότι η f_{X+Y} πρέπει να είναι η πυκνότητα γάμμα $\Gamma(\alpha_1 + \alpha_2, \lambda)$, όπως ισχυριστήκαμε. $\square$

Από την (17) και τον ορισμό της πυκνότητας γάμμα βλέπουμε επίσης ότι $c = 1/\Gamma(\alpha_1 + \alpha_2)$. Αυτό, μαζί με την (18), μάς επιτρέπει να υπολογίσουμε το ορισμένο ολοκλήρωμα που εμφανίζεται στην (18) με τη βοήθεια της συνάρτησης γάμμα:

$$(19) \qquad \int_0^1 u^{\alpha_1-1}(1-u)^{\alpha_2-1}du = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)}.$$

Αυτός ο τύπος μας επιτρέπει να ορίσουμε μία νέα οικογένεια πυκνοτήτων (με δύο παραμέτρους) που λέγονται πυκνότητες *Βήτα*. Η πυκνότητα Βήτα με παραμέτρους α_1 και α_2 ορίζεται από την

$$f(x) = \begin{cases} \frac{\Gamma(\alpha_1+\alpha_2)x^{\alpha_1-1}(1-x)^{\alpha_2-1}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} & , 0 < x < 1, \\ 0 & , \text{αλλιώς}. \end{cases}$$

Ο λόγος γι' αυτήν την ορολογία είναι ότι η συνάρτηση των α_1 και α_2 που ορίζεται από την

$$B(\alpha_1, \alpha_2) = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)}, \qquad 0 < \alpha_1, \alpha_2 < \infty,$$

λέγεται συνάρτηση Βήτα.

Δίνουμε τώρα μία τελευταία εφαρμογή του τύπου της συνέλιξης σε κανονικά κατανεμημένες τυχαίες μεταβλητές.

Θεώρημα 2. *Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές με κανονικές πυκνότητες, τις $n(\mu_1, \sigma_1^2)$ και $n(\mu_2, \sigma_2^2)$ αντίστοιχα. Τότε, η $X+Y$ έχει την κανονική πυκνότητα*

$$n(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

Απόδειξη: Υποθέτουμε πρώτα ότι $\mu_1 = \mu_2 = 0$. Τότε,

$$f_X(x) = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-x^2/2\sigma_1^2}, \qquad -\infty < x < \infty,$$

και

$$f_Y(y) = \frac{1}{\sigma_2\sqrt{2\pi}} e^{-y^2/2\sigma_2^2}, \qquad -\infty < y < \infty.$$

Επομένως, από την (15)

$$f_{X+Y}(z) = \frac{1}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_1^2} + \frac{(z-x)^2}{\sigma_2^2}\right)\right] dx.$$

Ο υπολογισμός αυτού του ολοκληρώματος δυστυχώς απαιτεί αρκετές πράξεις (τις οποίες δεν είναι εξίσου σημαντικό να συγκρατήσετε). Ένας τρόπος για να προχωρήσουμε είναι να κάνουμε πρώτα την αλλαγή μεταβλητής

$$u = \frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{\sigma_1\sigma_2} x.$$

Τότε, απλή άλγεβρα δείχνει ότι

$$f_{X+Y}(z) = \frac{1}{2\pi\sqrt{\sigma_1^2+\sigma_2^2}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\left(u^2 - \frac{2uz\sigma_1}{\sigma_2\sqrt{\sigma_1^2+\sigma_2^2}} + \frac{z^2}{\sigma_2^2}\right)\right] du.$$

Στη συνέχεια συμπληρώνουμε το τετράγωνο ως προς u και παρατηρούμε ότι

$$u^2 - \frac{2uz\sigma_1}{\sigma_2\sqrt{\sigma_1^2+\sigma_2^2}} + \frac{z^2}{\sigma_2^2} = \left(u - \frac{z\sigma_1}{\sigma_2\sqrt{\sigma_1^2+\sigma_2^2}}\right)^2 + \frac{z^2}{\sigma_1^2+\sigma_2^2}.$$

Κάνοντας μία δεύτερη αλλαγή μεταβλητής, την

$$v = u - \frac{z\sigma_1}{\sigma_2\sqrt{\sigma_1^2+\sigma_2^2}},$$

βλέπουμε ότι

$$\begin{aligned} f_{X+Y}(z) &= \frac{e^{-z^2/2(\sigma_1^2+\sigma_2^2)}}{\sqrt{2\pi}\sqrt{\sigma_1^2+\sigma_2^2}} \int_{-\infty}^{\infty} \frac{e^{-v^2/2}}{\sqrt{2\pi}} dv \\ &= \frac{e^{-z^2/2(\sigma_1^2+\sigma_2^2)}}{\sqrt{2\pi}\sqrt{\sigma_1^2+\sigma_2^2}}, \end{aligned}$$

που είναι απλώς η κανονική πυκνότητα $n(0, \sigma_1^2 + \sigma_2^2)$.

Γενικά, οι $X - \mu_1$ και $Y - \mu_2$ είναι ανεξάρτητες και έχουν πυκνότητες τις $n(0, \sigma_1^2)$ και $n(0, \sigma_2^2)$ αντίστοιχα. Άρα, από την παραπάνω ειδική περίπτωση, η $(X - \mu_1) + (Y - \mu_2) = X + Y - (\mu_1 + \mu_2)$ έχει την κανονική πυκνότητα $n(0, \sigma_1^2 + \sigma_2^2)$, άρα η $X + Y$ έχει την κανονική πυκνότητα

$$n(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

όπως ισχυριστήκαμε. □

Η απόδειξη που προηγήθηκε είναι στοιχειώδης αλλά πολύπλοκη. Στην Παράγραφο 8.3 θα δώσουμε μία απόδειξη με λιγότερους υπολογισμούς, η οποία όμως απαιτεί πιό προχωρημένες τεχνικές. Στην Άσκηση 36 αυτού του κεφαλαίου περιγράφουμε μία ακόμα απόδειξη.

Παράδειγμα 5. Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές που έχουν την κανονική πυκνότητα $n(0, \sigma^2)$. Βρείτε την πυκνότητα των $X+Y$ και X^2+Y^2.

Από το Θεώρημα 2 βλέπουμε αμέσως ότι η $X+Y$ έχει την κανονική πυκνότητα $n(0, 2\sigma^2)$. Από το Παράδειγμα 12 του Κεφαλαίου 5, καθεμία από τις X^2 και Y^2 έχει την πυκνότητα γάμμα $\Gamma(1/2, 1/2\sigma^2)$. Εύκολα ελέγχουμε ότι οι X^2 και Y^2 είναι ανεξάρτητες. Επομένως, από το Θεώρημα 1, η X^2+Y^2 έχει την πυκνότητα γάμμα $\Gamma(1, 1/2\sigma^2)$, η οποία συμπίπτει με την εκθετική πυκνότητα που έχει παράμετρο $1/2\sigma^2$.

6.2.2 Κατανομή πηλίκων*

Όπως πριν, με X και Y συμβολίζουμε τυχαίες μεταβλητές που έχουν από κοινού πυκνότητα την f. Θα αποδείξουμε τώρα έναν τύπο για την πυκνότητα της τυχαίας μεταβλητής $Z = Y/X$. Το σύνολο

$$A_z = \{(x, y) \mid y/x \le z\}$$

φαίνεται στο Σχήμα 4. Αν $x < 0$, τότε $y/x \le z$ αν και μόνο αν $y \ge xz$. Άρα,

$$A_z = \{(x, y) \mid x < 0, y \ge xz\} \cup \{(x, y) \mid x > 0, y \le xz\}.$$

Έπεται ότι

$$\begin{aligned} F_{Y/X}(z) &= \int\int_{A_z} f(x, y)dxdy \\ &= \int_{-\infty}^{0} \left(\int_{xz}^{\infty} f(x, y)dy \right) dx + \int_{0}^{\infty} \left(\int_{-\infty}^{xz} f(x, y)dy \right) dx. \end{aligned}$$

Στα εσωτερικά ολοκληρώματα, κάνουμε την αλλαγή μεταβλητής $y = xu$ (με $dy = xdv$) και παίρνουμε

$$
\begin{aligned}
F_{Y/X}(z) &= \int_{-\infty}^{0} \left(\int_{z}^{\infty} xf(x,xv)dv \right) dx + \int_{0}^{\infty} \left(\int_{-\infty}^{z} xf(x,xv)dv \right) dx \\
&= \int_{-\infty}^{0} \left(\int_{-\infty}^{z} (-x)f(x,xv)dv \right) dx + \int_{0}^{\infty} \left(\int_{-\infty}^{z} xf(x,xv)dv \right) dx \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{z} |x|f(x,xv)dv \right) dx.
\end{aligned}
$$

Αλλάζοντας τη σειρά της ολοκλήρωσης βλέπουμε ότι

$$
F_{Y/X}(z) = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{z} |x|f(x,xv)dv \right) dx. \tag{21}
$$

Έπεται από την (21) ότι η Y/X έχει πυκνότητα την $f_{Y/X}$ που δίνεται από την

$$
f_{Y/X}(z) = \int_{-\infty}^{\infty} |x|f(x,xz)dx, \qquad -\infty < z < \infty. \tag{22}
$$

Στην ειδική περίπτωση που οι X και Y είναι ανεξάρτητες θετικές τυχαίες μεταβλητές, η (22) γίνεται: $f_{Y/X}(z) = 0$ αν $z \leq 0$, και

$$
f_{Y/X}(z) = \int_{0}^{\infty} xf_X(x)f_Y(xz)dx, \qquad 0 < z < \infty. \tag{23}
$$

Το θεώρημα που ακολουθεί είναι άμεση εφαρμογή της (23).

Θεώρημα 3. *Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές με πυκνότητες τις $\Gamma(\alpha_1, \lambda)$ και $\Gamma(\alpha_2, \lambda)$ αντίστοιχα. Τότε η Y/X έχει πυκνότητα την: $f_{Y/X}(z) = 0$ αν $z \leq 0$, και*

$$f_{Y/X}(z) = \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \frac{z^{\alpha_2-1}}{(z+1)^{\alpha_1+\alpha_2}}, \qquad 0 < z < \infty. \tag{26}$$

Απόδειξη: Θυμηθείτε ότι

$$f_X(x) = \frac{\lambda^{\alpha_1} x^{\alpha_1-1} e^{-\lambda x}}{\Gamma(\alpha_1)}, \qquad x > 0,$$

και

$$f_Y(y) = \frac{\lambda^{\alpha_2} y^{\alpha_2-1} e^{-\lambda y}}{\Gamma(\alpha_2)}, \qquad y > 0.$$

Μπορούμε να εφαρμόσουμε τον τύπο (23), οπότε αν $0 < z < \infty$,

$$\begin{aligned} f_{Y/X}(z) &= \frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \int_0^\infty x x^{\alpha_1-1} e^{-\lambda x} (xz)^{\alpha_2-1} e^{-\lambda x z} dx \\ &= \frac{\lambda^{\alpha_1+\alpha_2} z^{\alpha_2-1}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \int_0^\infty x^{\alpha_1+\alpha_2-1} e^{-x\lambda(z+1)} dx. \end{aligned}$$

Από την ισότητα (34) του Κεφαλαίου 5

$$\int_0^\infty x^{\alpha_1+\alpha_2-1} e^{-x\lambda(z+1)} dx = \frac{\Gamma(\alpha_1+\alpha_2)}{(\lambda(z+1))^{\alpha_1+\alpha_2}}.$$

Επομένως η (24) ισχύει όπως ισχυριστήκαμε. □

Αφού η (24) ορίζει συνάρτηση πυκνότητας, βλέπουμε ότι για κάθε $\alpha_1, \alpha_2 > 0$

$$\int_0^\infty z^{\alpha_2-1}(z+1)^{-(\alpha_1+\alpha_2)} dz = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)}.$$

Παράδειγμα 6. Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές που έχουν την κανονική πυκνότητα $n(0, \sigma^2)$. Βρείτε την πυκνότητα της Y^2/X^2.

Οι τυχαίες μεταβλητές είναι ακριβώς οι ίδιες με αυτές του Παραδείγματος 5. Άρα, οι X^2 και Y^2 είναι πάλι ανεξάρτητες, και καθεμία έχει την πυκνότητα γάμμα $\Gamma(1/2, 1/2\sigma^2)$. Μπορούμε λοιπόν να εφαρμόσουμε το Θεώρημα 3, και βλέπουμε ότι η Y^2/X^2 έχει πυκνότητα την: $f_{Y^2/X^2}(z) = 0$ αν $z \leq 0$, και

$$\begin{aligned} f_{Y^2/X^2}(z) &= \frac{\Gamma(1)}{\Gamma(1/2)\Gamma(1/2)} \frac{z^{-1/2}}{(z+1)} \\ &= \frac{1}{\pi(z+1)\sqrt{z}}, \qquad 0 < z < \infty. \end{aligned}$$

(Θυμηθείτε εδώ ότι, από την ισότητα (35) του Κεφαλαίου 5, $\Gamma(1/2) = \sqrt{\pi}$.) Αφήνουμε σαν άσκηση στον αναγνώστη να δείξει ότι, με τις ίδιες υποθέσεις, τόσο η Y/X όσο και η $Y/|X|$ έχουν πυκνότητα Cauchy.

6.3 Δεσμευμένες πυκνότητες

Για να περάσουμε ομαλά στον ορισμό των δεσμευμένων πυκνοτήτων συνεχών τυχαίων μεταβλητών, θα συζητήσουμε πρώτα την περίπτωση των διακριτών τυχαίων μεταβλητών. Έστω X και Y διακριτές τυχαίες μεταβλητές που έχουν από κοινού πυκνότητα την f. Αν x είναι μία θετική τιμή της X, τότε

$$P(Y=y|X=x)=\frac{P(X=x,Y=y)}{P(X=x)}=\frac{f(x,y)}{f_X(x)}.$$

Η συνάρτηση $f_{Y|X}$ που ορίζεται από την

$$(25)\qquad f_{Y|X}(y|x)=\begin{cases}\frac{f(x,y)}{f_X(x)} & , f_X(x)\neq 0,\\ 0 & , f_X(x)=0.\end{cases}$$

λέγεται δεσμευμένη πυκνότητα της Y με δεδομένη την X. Για κάθε πιθανή τιμή x της X,

$$\sum_y f_{Y|X}(y|x)=\frac{\sum_y f(x,y)}{f_X(x)}=\frac{f_X(x)}{f_X(x)}=1,$$

δηλαδή για κάθε τέτοιο x, η $f_{Y|X}(y|x)$ ορίζει μία διακριτή συνάρτηση πυκνότητας ως προς y, που λέγεται δεσμευμένη πυκνότητα της Y δεδομένου ότι $X=x$. Στη διακριτή περίπτωση, η συζήτηση των δεσμευμένων πυκνοτήτων δεν απαιτεί πραγματικά νέες έννοιες.

Αν όμως η X είναι συνεχής τυχαία μεταβλητή, τότε $P(X=x)=0$ για κάθε x οπότε η $P(Y=y|X=x)$ είναι πάντα αόριστη. Σε αυτήν την περίπτωση οποιοσδήποτε ορισμός των δεσμευμένων πυκνοτήτων αναγκαστικά θα περιέχει μία νέα ιδέα. Ο απλούστερος τρόπος για να ορίσουμε δεσμευμένες πυκνότητες συνεχών τυχαίων μεταβλητών είναι κατ' αναλογία προς την (25) της διακριτής περίπτωσης.

Ορισμός 1 *Έστω X και Y συνεχείς τυχαίες μεταβλητές με από κοινού πυκνότητα την f. Η δεσμευμένη πυκνότητα $f_{Y|X}$ ορίζεται από την*

$$(26)\qquad f_{Y|X}(y|x)=\begin{cases}\frac{f(x,y)}{f_X(x)} & , 0<f_X(x)<\infty,\\ 0 & , \text{αλλιώς}.\end{cases}$$

Από τον ορισμό αυτό έπεται άμεσα ότι, σαν συνάρτηση του y, η $f_{Y|X}(y|x)$ είναι συνάρτηση πυκνότητας αν $0<f_X(x)<\infty$ (και λέγεται πάλι δεσμευμένη πυκνότητα της Y δεδομένου ότι $X=x$). Οι δεσμευμένες πυκνότητες χρησιμεύουν στον ορισμό των δεσμευμένων πιθανοτήτων. Ορίζουμε

$$(27)\qquad P(a\le Y\le b\mid X=x)=\int_a^b f_{Y|X}(y|x)dy,\qquad a\le b.$$

Εναλλακτικά, θα μπορούσαμε να επιχειρήσουμε να ορίσουμε τη δεσμευμένη πιθανότητα που εμφανίζεται στην (27) μέσω του εξής ορίου:

$$(28)\qquad P(a\le Y\le b\mid X=x)=\lim_{h\to 0^+}P(a\le Y\le b\mid x-h\le X\le x+h).$$

Το δεξιό μέλος της (28) ξαναγράφεται μέσω της f στην μορφή

$$\lim_{h\to 0^+} \frac{\int_{x-h}^{x+h}\left(\int_a^b f(u,y)dy\right)du}{\int_{x-h}^{x+h}\left(\int_{-\infty}^{\infty} f(u,y)dy\right)du} = \lim_{h\to 0^+} \frac{(1/2h)\int_{x-h}^{x+h}\left(\int_a^b f(u,y)dy\right)du}{(1/2h)\int_{x-h}^{x+h}\left(\int_{-\infty}^{\infty} f(u,y)dy\right)du}.$$

Αν το

$$\int_a^b f(u,y)dy$$

είναι συνεχής συνάρτηση του u στο $u = x$, τότε ο αριθμητής του τελευταίου ορίου συγκλίνει στο

$$\int_a^b f(x,y)dy$$

όταν $h \to 0^+$. Αν η f_X είναι συνεχής στο x, τότε ο παρονομαστής συγκλίνει στο $f_X(x)$ όταν $h \to 0^+$. Με την πρόσθετη υπόθεση ότι $f_X(x) \neq 0$, από την (28) οδηγούμαστε στην

$$P(a \leq Y \leq b \mid X = x) = \frac{\int_a^b f(x,y)dy}{f_X(x)},$$

κάτι που συμφωνεί με την (27). Για να συνοψίσουμε, ορίσαμε τις δεσμευμένες πυκνότητες και τις δεσμευμένες πιθανότητες στη συνεχή περίπτωση, κατ' αναλογία προς τη διακριτή περίπτωση. Παρατηρήσαμε επίσης, ότι με κάποιους επιπλέον περιορισμούς, μία διαδικασία περάσματος στο όριο θα οδηγούσε στον ίδιο ορισμό των δεσμευμένων πιθανοτήτων. Οι διαδικασίες αυτού του είδους είναι δύσχρηστες και δεν θα τις ξαναχρησιμοποιήσουμε.

Από τον ορισμό των δεσμευμένων συναρτήσεων πυκνότητας έπεται άμεσα ότι

(29) $$f(x,y) = f_X(x)f_{Y|X}(y|x), \quad -\infty < x,y < \infty.$$

Αν οι X και Y είναι ανεξάρτητες και

(30) $$f(x,y) = f_X(x)f_Y(y), \quad -\infty < x,y < \infty,$$

τότε

(31) $$f_{Y|X}(y|x) = f_Y(y), \quad 0 < f_X(x) < \infty \quad , \quad -\infty < y < \infty.$$

Αντίστροφα, αν ισχύει η (31), τότε από την (29) έπεται ότι η (30) ισχύει και οι X και Y είναι ανεξάρτητες. Δηλαδή, η (31) είναι ικανή και αναγκαία συνθήκη για την ανεξαρτησία δύο τυχαίων μεταβλητών X και Y που έχουν από κοινού πυκνότητα.

Παράδειγμα 7. Υποθέτουμε ότι οι X και Y έχουν τη διδιάστατη πυκνότητα που δίνει ο τύπος (13), δηλαδή

$$f(x,y) = \frac{\sqrt{3}}{4\pi}e^{-(x^2-xy+y^2)/2}, \quad -\infty < x,y < \infty.$$

Τότε, όπως είδαμε στο Παράδειγμα 2, η X έχει την κανονική πυκνότητα $n(0,4/3)$. Άρα, για κάθε $-\infty < x, y < \infty$

$$\begin{aligned} f_{Y|X}(y|x) &= \frac{\frac{\sqrt{3}}{4\pi}e^{-(x^2-xy+y^2)/2}}{\frac{\sqrt{3}}{2\sqrt{2\pi}}e^{-3x^2/8}} \\ &= \frac{1}{\sqrt{2\pi}}e^{-(y-x/2)^2/2}. \end{aligned}$$

Με άλλα λόγια, η δεσμευμένη πυκνότητα της Y δεδομένου ότι $X = x$ είναι η κανονική πυκνότητα $n(x/2, 1)$.

Ξεκινήσαμε με τις από κοινού πυκνότητες και τις χρησιμοποιήσαμε για να κατασκευάσουμε τις περιθώριες και τις δεσμευμένες πυκνότητες. Σε ορισμένες περιπτώσεις, η διαδικασία αντιστρέφεται: ξεκινώντας από τις περιθώριες και τις δεσμευμένες πυκνότητες, και χρησιμοποιώντας τις, κατασκευάζουμε από κοινού πυκνότητες.

Παράδειγμα 8. Έστω X τυχαία μεταβλητή ομοιόμορφα κατανεμημένη στο $(0,1)$, και έστω Y τυχαία μεταβλητή ομοιόμορφα κατανεμημένη στο $(0, X)$. Βρείτε την από κοινού πυκνότητα των X και Y, και την περιθώρια πυκνότητα της Y.

Από τη διατύπωση του προβλήματος, βλέπουμε ότι η περιθώρια πυκνότητα της X δίνεται από την

$$f_X(x) = \begin{cases} 1 & , 0 < x < 1, \\ 0 & , \text{αλλιώς}. \end{cases}$$

Η πυκνότητα της Y δεδομένου ότι $X = x$ είναι ομοιόμορφη στο $(0, x)$, άρα

$$f_{Y|X}(y|x) = \begin{cases} 1/x & , 0 < y < x < 1, \\ 0 & , \text{αλλιώς}. \end{cases}$$

Επομένως, η από κοινού πυκνότητα των X και Y δίνεται από την

$$f(x,y) = \begin{cases} 1/x & , 0 < y < x < 1, \\ 0 & , \text{αλλιώς}. \end{cases}$$

Η περιθώρια πυκνότητα της Y είναι η

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx = \int_y^1 \frac{1}{x}dx = -\log y, \qquad 0 < y < 1,$$

και $f_Y(y) = 0$ αλλιώς.

6.3.1 Ο τύπος του Bayes

Μπορούμε φυσικά να αντιστρέψουμε τους ρόλους των X και Y και να ορίσουμε τη δεσμευμένη πυκνότητα της X δεδομένου ότι $Y = y$ μέσω του τύπου

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}, \qquad 0 < f_Y(y) < \infty. \tag{32}$$

Αφού

$$f(x,y) = f_X(x) f_{Y|X}(y|x)$$

και

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx = \int_{-\infty}^{\infty} f_X(x) f_{Y|X}(y|x)dx,$$

μπορούμε να ξαναγράψουμε την (32) στη μορφή

$$f_{X|Y}(x|y) = \frac{f_X(x) f_{Y|X}(y|x)}{\int_{-\infty}^{\infty} f_X(x) f_{Y|X}(y|x)dx}. \tag{33}$$

Ο τύπος αυτός είναι το συνεχές ανάλογο του τύπου του Bayes που συζητήσαμε στο Κεφάλαιο 1.

Στα Κεφάλαια 3 και 4 θεωρήσαμε τυχαίες μεταβλητές X και Y που ήταν και οι δύο διακριτές. Μέχρι τώρα στο Κεφάλαιο 6 έχουμε συναντήσει κυρίως τυχαίες μεταβλητές X και Y που είναι και οι δύο συνεχείς. Υπάρχουν περιπτώσεις στις οποίες ενδιαφερόμαστε ταυτόχρονα για διακριτές και συνεχείς τυχαίες μεταβλητές. Ο αναγνώστης θα φαντάζεται με ποιόν τρόπο μπορούμε να τροποποιήσουμε τη συζήτηση ώστε να συμπεριλάβουμε κι αυτήν την περίπτωση. Μερικές από τις πιό ενδιαφέρουσες εφαρμογές του τύπου (33) είναι αυτού του είδους.

Παράδειγμα 9. Ας υποθέσουμε ότι το πλήθος των αυτοκινητιστικών ατυχημάτων στα οποία θα εμπλακεί ένας οδηγός στη διάρκεια ενός έτους είναι μία τυχαία μεταβλητή Y που ακολουθεί κατανομή Poisson με παράμετρο λ, όπου το λ εξαρτάται από τον οδηγό. Αν επιλέξουμε τυχαία κάποιον οδηγό από έναν πληθυσμό, μπορούμε να αφήσουμε το λ να ποικίλλει, και να ορίσουμε μία τυχαία μεταβλητή Λ με πυκνότητα f_Λ. Η δεσμευμένη πυκνότητα της Y δεδομένου ότι $\Lambda = \lambda$ είναι η πυκνότητα Poisson με παράμετρο λ που δίνεται από την

$$f_{Y|\Lambda}(y|\lambda) = \begin{cases} \frac{\lambda^y e^{-\lambda}}{y!} & , y = 0, 1, 2, \ldots, \\ 0 & , \text{αλλιώς}. \end{cases}$$

Επομένως, η από κοινού πυκνότητα των Λ και Y είναι η

$$f(\lambda, y) = \begin{cases} \frac{f_\Lambda(\lambda)\lambda^y e^{-\lambda}}{y!} & , y = 0, 1, 2, \ldots, \\ 0 & , \text{αλλιώς}. \end{cases}$$

Γενικά, δεν μπορούμε να βρούμε έναν ωραίο τύπο για την περθώρια πυκνότητα της Y ή τη δεσμευμένη πυκνότητα της Λ δεδομένου ότι $Y = y$, γιατί δεν μπορούμε να υπολογίσουμε τα απαιτούμενα ολοκληρώματα. Μπορούμε όμως να βρούμε απλούς τύπους στην ειδική περίπτωση που η f είναι μία πυκνότητα γάμμα $\Gamma(\alpha, \beta)$, δηλαδή αν

$$f_\Lambda(\lambda) = \begin{cases} \frac{\beta^\alpha \lambda^{\alpha-1} e^{-\lambda\beta}}{y!} & , \lambda > 0, \\ 0 & , \text{αλλιώς}. \end{cases}$$

Σε αυτήν την περίπτωση,

$$f_Y(y) = \int_{-\infty}^{\infty} f(\lambda, y)d\lambda$$

$$
\begin{aligned}
&= \int_0^\infty \frac{\beta^\alpha \lambda^{\alpha-1} e^{-\lambda\beta}}{\Gamma(\alpha)} \frac{\lambda^y e^{-\lambda}}{y!} d\lambda \\
&= \frac{\beta^\alpha}{y!\Gamma(\alpha)} \int_0^\infty \lambda^{\alpha+y-1} e^{-\lambda(\beta+1)} d\lambda \\
&= \frac{\Gamma(\alpha+y)\beta^\alpha}{y!\Gamma(\alpha)(\beta+1)^{\alpha+y}}.
\end{aligned}
$$

Η τιμή του τελευταίου ολοκληρώματος προκύπτει με τη βοήθεια του τύπου (34) από το Κεφάλαιο 5. Αφήνουμε σαν άσκηση στον αναγνώστη να δείξει ότι η f_Y είναι η αρνητική διωνυμική πυκνότητα με παραμέτρους α και $p = \beta/(1+\beta)$. Επίσης, αν $\lambda > 0$ και y είναι ένας μη αρνητικός ακέραιος, τότε

$$
\begin{aligned}
f_{\Lambda|Y}(\lambda|y) &= \frac{f(\lambda, y)}{f_Y(y)} \\
&= \frac{\beta^\alpha \lambda^{\alpha+y-1} e^{-\lambda(\beta+1)} y!\Gamma(\alpha)(\beta+1)^{\alpha+y}}{\Gamma(\alpha)y!\Gamma(\alpha+y)\beta^\alpha} \\
&= \frac{(\beta+1)^{\alpha+y}\lambda^{\alpha+y-1}e^{-\lambda(\beta+1)}}{\Gamma(\alpha+y)},
\end{aligned}
$$

κάτι που μας λέει ότι η δεσμευμένη πυκνότητα της Λ δεδομένου ότι $Y = y$ είναι η πυκνότητα γάμμα $\Gamma(\alpha+y, \beta+1)$. Αν κάποιο στέλεχος της ασφαλιστικής βιομηχανίας ήθελε να λύσει ένα πρόβλημα σαν κι αυτό, κατά πάσα πιθανότητα θα προσπαθούσε να προσεγγίσει την πραγματική πυκνότητα f_Λ με μία πυκνότητα γάμμα $\Gamma(\alpha, \beta)$, όπου τα α και β θα επιλέγονταν ώστε η προσέγγιση να γίνει όσο γίνεται καλύτερη.

6.4 Ιδιότητες των πολυδιάστατων κατανομών

Οι έννοιες που συζητήσαμε σε αυτό το κεφάλαιο για δύο τυχαίες μεταβλητές X και Y γενικεύονται άμεσα στις n τυχαίες μεταβλητές. Στην παράγραφο αυτή θα δείξουμε εν συντομία πώς μπορεί να γίνει κάτι τέτοιο.

Έστω $X_1, \ldots, X_n$ τυχαίες μεταβλητές που ορίζονται στον ίδιο χώρο πιθανότητας. Η *από κοινού συνάρτηση κατανομής* τους F ορίζεται από την

$$F(x_1, \ldots, x_n) = P(X_1 \le x_1, \ldots, X_n \le x_n), \quad -\infty < x_1, \ldots, x_n < \infty.$$

Οι περιθώριες συναρτήσεις κατανομής F_{x_m}, $m = 1, \ldots, n$, ορίζονται από τις

$$F_{X_m}(x_m) = P(X_m \le x_m), \quad -\infty < x_m < \infty.$$

Η τιμή $F_{X_m}(x_m)$ προκύπτει από την F αν αφήσουμε τα $x_1, \ldots, x_{m-1}, x_{m+1}, \ldots, x_n$ να πάνε στο $+\infty$.

Μία μη αρνητική συνάρτηση f λέγεται από κοινού συνάρτηση πυκνότητας (συνεχούς τύπου) για την από κοινού συνάρτηση κατανομής F, ή για τις τυχαίες μεταβλητές $X_1, \ldots, X_n$, αν

(34)
$$F(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(u_1, \ldots, u_n) du_1 \ldots du_n, \quad -\infty < x_1, \ldots, x_n < \infty.$$

Κάτω από κάποιες ήπιες πρόσθετες συνθήκες, η

$$f(x_1, \ldots, x_n) = \frac{\partial^n}{\partial x_1 \ldots \partial x_n} F(x_1, \ldots, x_n)$$

ισχύει στα σημεία συνέχειας της F. Αν η (34) ισχύει, και A είναι οποιοδήποτε λογικό χωρίο στον $\mathbb{R}^n$, τότε

$$P((X_1, \ldots, X_n) \in A) = \int \cdots \int_A f(x_1, \ldots, x_n) dx_1 \ldots dx_n.$$

Ειδικότερα

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_n) dx_1 \ldots dx_n = 1 \tag{35}$$

και αν $a_m \leq b_m$ για $m = 1, \ldots, n$, τότε

$$P(a_1 \leq X_1 \leq b_1, \ldots, a_n \leq X_n \leq b_n) = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x_1, \ldots, x_n) dx_1 \ldots dx_n.$$

Η τυχαία μεταβλητή X_m έχει περιθώρια πυκνότητα την f_{X_m}, που προκύπτει με ολοκλήρωση της f ως προς τις υπόλοιπες $n-1$ μεταβλητές. Για παράδειγμα,

$$f_{X_2}(x_2) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_n) dx_1 dx_3 \ldots dx_n.$$

Γενικά, οι τυχαίες μεταβλητές $X_1, \ldots, X_n$ λέγονται ανεξάρτητες αν, οποτεδήποτε $a_m \leq b_m$, $m = 1, \ldots, n$, τότε

$$P(a_1 < X_1 \leq b_1, \ldots, a_n < X_n \leq b_n) = P(a_1 < X_1 \leq b_1) \ldots P(a_n < X_n \leq b_n).$$

Ικανή και αναγκαία συνθήκη για την ανεξαρτησία είναι η

$$F(x_1, \ldots, x_n) = F_{X_1}(x_1) \ldots F_{X_n}(x_n), \quad -\infty < x_1, \ldots, x_n < \infty.$$

Είναι σαφές ότι η συνθήκη είναι αναγκαία, αν όμως $n > 2$, η απόδειξη του ότι είναι ικανή προκύπτει με τέχνασμα, και δεν θα δοθεί εδώ. Αν η F έχει πυκνότητα την f, τότε οι $X_1, \ldots, X_n$ είναι ανεξάρτητες αν και μόνο αν μπορούμε να επιλέξουμε την f ώστε να ισχύει η

$$f(x_1, \ldots, x_n) = f_{X_1}(x_1) \ldots f_{X_n}(x_n), \quad -\infty < x_1, \ldots, x_n < \infty.$$

Μπορούμε επίσης να ορίσουμε μία n-διάστατη πυκνότητα απευθείας σαν μία μη αρνητική συνάρτηση στον $\mathbb{R}^n$ που ικανοποιεί την (35). Ο απλούστερος τρόπος κατασκευής n-διάστατων πυκνοτήτων είναι να ξεκινήσουμε με n μονοδιάστατες πυκνότητες $f_1, \ldots, f_n$ και να ορίσουμε την f παίρνοντας

$$f(x_1, \ldots, x_n) = f_1(x_1) \ldots f_n(x_n), \quad -\infty < x_1, \ldots, x_n < \infty. \tag{36}$$

Αν $X_1, \ldots, X_n$ είναι τυχαίες μεταβλητές με από κοινού πυκνότητα την f που δίνεται από την (36), τότε οι $X_1, \ldots, X_n$ είναι ανεξάρτητες, και η X_m έχει περιθώρια πυκνότητα την f_m.

Παράδειγμα 10. Έστω $X_1, \ldots, X_n$ ανεξάρτητες τυχαίες μεταβλητές, καθεμία από τις οποίες έχει εκθετική πυκνότητα με παράμετρο λ. Βρείτε την από κοινού πυκνότητα των $X_1, \ldots, X_n$.

Η πυκνότητα της X_m δίνεται από την

$$f_{X_m}(x_m) = \begin{cases} \lambda e^{-\lambda x_m} & , 0 < x_m < \infty, \\ 0 & , \text{αλλιώς.} \end{cases}$$

Άρα η f δίνεται από την

$$f(x_1, \ldots, x_n) = \begin{cases} \lambda^n e^{-\lambda(x_1 + \ldots + x_n)} & , x_1, \ldots, x_n > 0, \\ 0 & , \text{αλλιώς.} \end{cases}$$

Για να υπολογίσουμε την πυκνότητα του αθροίσματος n ανεξάρτητων τυχαίων μεταβλητών, αλλά και για πολλές άλλες εφαρμογές, χρειαζόμαστε το εξής αποτέλεσμα.

Θεώρημα 4. *Έστω $X_1, \ldots, X_n$ ανεξάρτητες τυχαίες μεταβλητές. Έστω Y μία τυχαία μεταβλητή που ορίζεται με βάση τις $X_1, \ldots, X_n$, και έστω Z μία τυχαία μεταβλητή που ορίζεται με βάση τις $X_{m+1}, \ldots, X_n$ (όπου $1 \leq m < n$). Τότε, οι Y και Z είναι ανεξάρτητες.*

Η απόδειξη αυτού του θεωρήματος παραλείπεται, γιατί απαιτεί γνώσεις θεωρίας μέτρου.

Χρησιμοποιώντας αυτό το θεώρημα και ένα επιχείρημα βασισμένο στη μαθηματική επαγωγή, μπορούμε να επεκτείνουμε τα Θεωρήματα 1 και 2 σε αθροίσματα ανεξάρτητων τυχαίων μεταβλητών, ως εξής.

Θεώρημα 5. *Έστω $X_1, \ldots, X_n$ ανεξάρτητες τυχαίες μεταβλητές τέτοιες ώστε η X_m να έχει την πυκνότητα γάμμα $\Gamma(\alpha_m, \lambda)$, $m = 1, \ldots, n$. Τότε, η $X_1 + \ldots + X_n$ έχει την πυκνότητα γάμμα $\Gamma(\alpha, \lambda)$, όπου*

$$\alpha = \alpha_1 + \ldots + \alpha_n.$$

Θυμηθείτε ότι η εκθετική πυκνότητα με παράμετρο λ συμπίπτει με την πυκνότητα γάμμα $\Gamma(1, \lambda)$. Έτσι, σαν πόρισμα αυτού του θεωρήματος παίρνουμε το εξής: *Αν $X_1, \ldots, X_n$ είναι ανεξάρτητες τυχαίες μεταβλητές, καθεμία από τις οποίες έχει* *εκθετική πυκνότητα με παράμετρο λ, τότε η $X_1 + \ldots + X_n$ έχει την πυκνότητα γάμμα* $\Gamma(n, \lambda)$.

Θεώρημα 6. *Έστω $X_1, \ldots, X_n$ ανεξάρτητες τυχαίες μεταβλητές τέτοιες ώστε η X_m να έχει την κανονική πυκνότητα $n(\mu_m, \sigma_m^2)$, $m = 1, \ldots, n$. Τότε, η $X_1 + \ldots + X_n$ έχει την κανονική πυκνότητα $n(\mu, \sigma^2)$, όπου*

$$\mu = \mu_1 + \ldots + \mu_n \quad , \quad \sigma^2 = \sigma_1^2 + \ldots + \sigma_n^2.$$

Αν οι $X_1, \ldots, X_n$ έχουν από κοινού πυκνότητα την f, τότε κάθε υποοικογένεια αυτών των τυχαίων μεταβλητών έχει από κοινού πυκνότητα που υπολογίζεται με ολοκλήρωση ως προς τις υπόλοιπες μεταβλητές. Για παράδειγμα, αν $1 \leq m < n$, τότε

$$f_{X_1,\ldots,X_m}(x_1, \ldots, x_m) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_n) dx_{m+1} \ldots dx_n.$$

Η δεσμευμένη πυκνότητα μιας υποοικογένειας των $X_1, \ldots, X_n$ με δεδομένες τις υπόλοιπες μεταβλητές, μπορεί επίσης να οριστεί κατά προφανή τρόπο. Για παράδειγμα, η δεσμευμένη πυκνότητα των $X_{m+1}, \ldots, X_n$ με δεδομένες τις $X_1, \ldots, X_m$ ορίζεται από την

$$f_{X_{m+1},\ldots,X_n|X_1,\ldots,X_m}(x_{m+1}, \ldots, x_n | x_1, \ldots, x_m) = \frac{f(x_1, \ldots, x_n)}{f_{X_1,\ldots,X_m}(x_1, \ldots, x_m)},$$

όπου f είναι η από κοινού πυκνότητα των $X_1, \ldots, X_n$.

Οι δεσμευμένες πυκνότητες εκφράζονται συχνά με τη βοήθεια κάπως διαφορετικού συμβολισμού. Για παράδειγμα, έστω $n + 1$ τυχαίες μεταβλητές $X_1, \ldots, X_n, Y$ που έχουν από κοινού πυκνότητα την f. Τότε, η δεσμευμένη πυκνότητα της Y με δεδομένες τις $X_1, \ldots, X_n$ ορίζεται από την

$$f_{Y|X_1,\ldots,X_n}(y | x_1, \ldots, x_n) = \frac{f(x_1, \ldots, x_n, y)}{f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)}.$$

6.5 Διατεταγμένα δείγματα*

Έστω $U_1, \ldots, U_n$ ανεξάρτητες συνεχείς τυχαίες μεταβλητές, καθεμία από τις οποίες έχει συνάρτηση κατανομής την F και συνάρτηση πυκνότητας την f. Έστω $X_1, \ldots, X_n$ οι τυχαίες μεταβλητές που προκύπτουν αν σαν $X_1(\omega), \ldots, X_n(\omega)$ πάρουμε το σύνολο των $U_1(\omega), \ldots, U_n(\omega)$ μετά από μετάθεση ώστε να βρίσκεται σε αύξουσα διάταξη. Ειδικότερα, οι X_1 και X_n ορίζονται από τις σχέσεις

$$X_1(\omega) = \min(U_1(\omega), \ldots, U_n(\omega))$$

και

$$X_n(\omega) = \max(U_1(\omega), \ldots, U_n(\omega)).$$

Η τυχαία μεταβλητή X_k ονομάζεται *k-στή διατεταγμένη στατιστική συνάρτηση*. Μία άλλη σχετική τυχαία μεταβλητή που παρουσιάζει ενδιαφέρον είναι το εύρος R του δείγματος που ορίζεται από την

$$\begin{aligned} R(\omega) &= X_n(\omega) - X_1(\omega) \\ &= \max(U_1(\omega), \ldots, U_n(\omega)) - \min(U_1(\omega), \ldots, U_n(\omega)). \end{aligned}$$

Από τις υποθέσεις μας για τις $U_1, \ldots, U_n$ προκύπτει ότι, με πιθανότητα ένα, οι U_i παίρνουν διακεκριμένες τιμές, επομένως $X_1 < X_2 < \ldots < X_n$.

Για να δώσουμε ένα αριθμητικό παράδειγμα, ας υποθέσουμε ότι $U_1(\omega) = 4.8$, $U_2(\omega) = 3.5$ και $U_3(\omega) = 4.3$. Τότε, $X_1(\omega) = 4.3$, $X_3(\omega) = 4.8$ και $R(\omega) = 1.3$.

Παράδειγμα 11. Θεωρούμε ένα μηχάνημα που αποτελείται από n εξαρτήματα των οποίων οι χρόνοι αποτυχίας $U_1, \ldots, U_n$ ικανοποιούν τις υποθέσεις αυτής της παραγράφου. Τότε, X_k είναι ο χρόνος που απαιτείται μέχρι να υποστούν βλάβη k από τα εξαρτήματα του μηχανήματος. Αν το μηχάνημα παύει να λειτουργεί όταν κάποιο από τα εξαρτήματά του υποστεί βλάβη, τότε η $X_1 = \min(U_1, \ldots, U_n)$ δίνει τον χρόνο αποτυχίας του μηχανήματος. Αν το μηχάνημα συνεχίζει τη λειτουργία του μέχρι τη στιγμή που θα σταματήσουν να λειτουργούν όλα τα εξαρτήματά του, τότε η $X_n = \max(U_1, \ldots, U_n)$ δίνει τον χρόνο αποτυχίας του μηχανήματος.

Παράδειγμα 12. Ας υποθέσουμε ότι μία αλυσίδα παραγωγής επιδιώκει την παρασκευή n πανομοιότυπων αντικειμένων στη χρονική μονάδα της λειτουργίας της. Συμβολίζουμε με $U_1, \ldots, U_n$ τα μήκη αυτών των n αντικειμένων. Ένας επιθεωρητής ελέγχει το ελάχιστο μήκος X_1 και το μέγιστο μήκος X_n για να δεί αν βρίσκονται εντός ικανοποιητικών ορίων. Η αξιοπιστία της παραγωγικής μονάδας είναι υψηλότερη αν η διακύμανση μήκους κρατιέται σε χαμηλό επίπεδο. Το εύρος R του δείγματος είναι ένας από τους πιθανούς δείκτες αυτής της διακύμανσης.

Στη συνέχεια υπολογίζουμε τη συνάρτηση κατανομής της k-στής διατεταγμένης στατιστικής συνάρτησης X_k. Έστω $-\infty < x < \infty$. Η πιθανότητα ακριβώς j από τις U_j να βρίσκονται στο $(-\infty, x]$ και οι υπόλοιπες $(n-j)$ να βρίσκονται στο (x, ∞), είναι

$$\binom{n}{j} F^j(x)(1 - F(x))^{n-j},$$

κάτι που μπορείτε να δείτε χρησιμοποιώντας τη διωνυμική κατανομή με παραμέτρους n και $p = F(x)$. Το ενδεχόμενο $\{X_k \leq x\}$ συμβαίνει αν και μόνο αν k ή περισσότερες από τις U_i βρίσκονται στο $(-\infty, x]$. Άρα,

$$\begin{aligned} F_{X_k}(x) &= P(X_k \leq x) \\ &= \sum_{j=k}^{n} \binom{n}{j} F^j(x)(1 - F(x))^{n-j}, \qquad -\infty < x < \infty. \end{aligned} \tag{37}$$

Ειδικότερα, οι συναρτήσεις κατανομής των X_n και X_1 γράφονται πολύ απλά ως εξής:

$$F_{X_n}(x) = (F(x))^n, \qquad -\infty < x < \infty,$$

και

$$F_{X_1}(x) = 1 - (1 - F(x))^n, \qquad -\infty < x < \infty.$$

Για να υπολογίσουμε τις αντίστοιχες συναρτήσεις πυκνότητας, πρέπει να παραγωγίσουμε αυτές τις ποσότητες. Εύκολα βλέπουμε ότι

$$f_{X_n}(x) = nf(x)F^{n-1}(x), \qquad -\infty < x < \infty,$$

και

$$f_{X_1}(x) = nf(x)(1 - F(x))^{n-1}, \qquad -\infty < x < \infty.$$

Ο αντίστοιχος υπολογισμός για την X_k είναι κάπως πιό πολύπλοκος. Από την (37),

$$\begin{aligned}
f_{X_k}(x) &= \sum_{j=k}^{n} \frac{n!}{(j-1)!(n-j)!} f(x)F^{j-1}(x)(1-F(x))^{n-j} \\
&\quad - \sum_{j=k}^{n-1} \frac{n!}{j!(n-j-1)!} f(x)F^{j}(x)(1-F(x))^{n-j-1} \\
&= \sum_{j=k}^{n} \frac{n!}{(j-1)!(n-j)!} f(x)F^{j-1}(x)(1-F(x))^{n-j} \\
&\quad - \sum_{j=k+1}^{n} \frac{n!}{(j-1)!(n-j)!} f(x)F^{j-1}(x)(1-F(x))^{n-j},
\end{aligned}$$

από την οποία παίρνουμε

$$f_{X_k}(x) = \frac{n!}{(k-1)!(n-k)!} f(x)F^{k-1}(x)(1-F(x))^{n-k}, \quad -\infty < x < \infty. \tag{38}$$

Για να βρούμε την πυκνότητα του εύρους R, θα βρούμε πρώτα την από κοινού πυκνότητα των X_1 και X_n. Υποθέτουμε ότι $n \geq 2$ (αν $n = 1$, τότε $R = 0$). Έστω $x \leq y$. Τότε,

$$\begin{aligned}
P(X_1 > x\ ,\ X_n \leq y) &= P(x < U_1 \leq y, \ldots, x < U_n \leq y) \\
&= (F(y) - F(x))^n,
\end{aligned}$$

και, φυσικά,

$$P(X_n \leq y) = F^n(y).$$

Επομένως,

$$\begin{aligned}
F_{X_1,X_n}(x,y) &= P(X_1 \leq x\ ,\ X_n \leq y) \\
&= P(X_n \leq y) - P(X_1 > x\ ,\ X_n \leq y) \\
&= F^n(y) - (F(y) - F(x))^n.
\end{aligned}$$

Η από κοινού πυκνότητα δίνεται από την

$$\begin{aligned}
f_{X_1,X_n}(x,y) &= \frac{\partial^2}{\partial x \partial y} F_{X_1,X_n}(x,y) \\
&= n(n-1)f(x)f(y)(F(y) - F(x))^{n-2}, \qquad x \leq y.
\end{aligned}$$

Είναι εύκολο να αποδειχθεί ότι

$$f_{X_1,X_n}(x,y) = 0, \qquad x > y.$$

Με απλή παραλλαγή του επιχειρήματος που χρησιμοποιήσαμε στην Παράγραφο 6.2.1 για να βρούμε την πυκνότητα αθροίσματος τυχαίων μεταβλητών, βλέπουμε ότι η πυκνότητα της $R = X_n - X_1$ δίνεται από την

$$f_R(r) = \int_{-\infty}^{\infty} f_{X_1,X_n}(x, r+x)dx.$$

Με άλλα λόγια,

$$f_R(r) = \begin{cases} n(n-1)\int_{-\infty}^{\infty} f(x)f(r+x)(F(r+x)-F(x))^{n-2}dx & , r > 0 \\ 0 & , r < 0. \end{cases}$$

Όλοι αυτοί οι τύποι παίρνουν πολύ απλή μορφή στην περίπτωση που οι $U_1, \ldots, U_n$ είναι ανεξάρτητες και ομοιόμορφα κατανεμημένες στο $(0,1)$. Αυτό αφήνεται σαν άσκηση.

Υπάρχει ένας «ευριστικός» τρόπος για να αποδείξουμε τους παραπάνω τύπους, ο οποίος βοηθάει πολύ. Θα τον περιγράψουμε αποδεικνύοντας ξανά τον τύπο για την f_{X_k}. Έστω dx ένα απειροστό θετικό μέγεθος. Τότε, έχουμε την προσεγγιστική σχέση

$$f_{X_k}(x)dx \approx P(x \le X_k \le x+dx).$$

Ο πλέον πιθανός τρόπος για να συμβαίνει το ενδεχόμενο $\{x \le X_k \le x+dx\}$ είναι όταν $k-1$ από τις U_i βρίσκονται στο $(-\infty, x]$, μία από τις U_i βρίσκεται στο $(x, x+dx]$, και $n-k$ από τις U_i βρίσκονται στο $(x+dx, \infty)$ (δείτε το Σχήμα 5). Μπορούμε τώρα να χρησιμοποιήσουμε την πολυωνυμική κατανομή που συζητήσαμε στο Κεφάλαιο 3: η πιθανότητα να έχουμε το παραπάνω πλήθος από U_i στα κατάλληλα διαστήματα είναι

$$\begin{aligned} f_{X_k}(x)dx &\equiv \frac{n!}{(k-1)!(n-k)!} \\ &\quad \times \left(\int_{-\infty}^{x} f(u)du\right)^{k-1} \int_{x}^{x+dx} f(u)du \left(\int_{x+dx}^{\infty} f(u)du\right)^{n-k} \\ &\equiv \frac{n!}{(k-1)!(n-k)!} f(x)dxF^{k-1}(x)(1-F(x))^{n-k}, \end{aligned}$$

από την οποία παίρνουμε την (38). Δεν θα επιχειρήσουμε την αυστηρή θεμελίωση αυτής της μεθόδου.

6.6 Δειγματικές κατανομές*

Έστω $X_1, \ldots, X_n$ ανεξάρτητες τυχαίες μεταβλητές, που έχουν την κανονική πυκνότητα $n(0, \sigma^2)$. Σε αυτήν την παράγραφο θα βρούμε τις συναρτήσεις κατανομής διαφόρων τυχαίων μεταβλητών που ορίζονται με βάση τις X_i. Πέρα από το γεγονός ότι αυτή η μελέτη θα μάς δώσει την ευκαιρία να εφαρμόσουμε το υλικό των προηγουμένων παραγράφων, οι συναρτήσεις κατανομής που θα συζητήσουμε είναι θεμελιώδους σημασίας για τη στατιστική συμπερασματολογία, και θα τις χρειαστούμε στον Τόμο II.

Η σταθερά σ^2 είναι βολική αλλά επουσιώδης αφού οι $X_1/\sigma, \ldots, X_n/\sigma$ είναι ανεξάρτητες και ακολουθούν την τυπική κανονική κατανομή $n(0, 1)$. Μπορούμε λοιπόν πάντα να παίρνουμε $\sigma^2 = 1$ χωρίς περιορισμό της γενικότητας.

Από το Θεώρημα 6, η τυχαία μεταβλητή $X_1 + \ldots + X_n$ έχει κανονική πυκνότητα με παραμέτρους 0 και $n\sigma^2$. Διαιρώντας αυτό το άθροισμα με διάφορες σταθερές, παίρνουμε εναλλακτικές μορφές αυτού του αποτελέσματος. Για παράδειγμα, η

$$\frac{X_1 + \ldots + X_n}{n}$$

ακολουθεί κανονική κατανομή με παραμέτρους 0 και σ^2/n, ενώ η

$$\frac{X_1 + \ldots + X_n}{\sigma\sqrt{n}}$$

έχει την τυπική κανονική πυκνότητα $n(0, 1)$.

Αφού η X_1/σ έχει την τυπική κανονική πυκνότητα, από το Παράδειγμα 12 του Κεφαλαίου 5 έπεται ότι η X_1^2/σ^2 έχει την πυκνότητα γάμμα $\Gamma(1/2, 1/2)$. Το Θεώρημα 5 δείχνει ότι η

$$\frac{X_1^2 + \ldots + X_n^2}{\sigma^2}$$

έχει την πυκνότητα γάμμα $\Gamma(n/2, 1/2)$. Η συγκεκριμένη πυκνότητα γάμμα είναι πολύ σημαντική στη Στατιστική. Λέμε ότι η τυχαία μεταβλητή που αντιστοιχεί σε αυτήν έχει κατανομή χ^2 με n βαθμούς ελευθερίας, και συμβολίζεται με $\chi^2(n)$. Εφαρμόζοντας το Θεώρημα 5 καταλήγουμε στο εξής αποτέλεσμα για τις κατανομές χ^2.

Θεώρημα 7 *Έστω $Y_1, \ldots, Y_n$ ανεξάρτητες τυχαίες μεταβλητές τέτοιες ώστε η Y_m να έχει την κατανομή $\chi^2(k_m)$. Τότε, η $Y_1 + \ldots + Y_n$ έχει την κατανομή $\chi^2(k)$, όπου $k = k_1 + \ldots + k_n$.*

Απόδειξη: Από την υπόθεσή μας, η Y_m έχει την κατανομή γάμμα $\Gamma(k_m/2, 1/2)$. Το Θεώρημα 5 μάς εξασφαλίζει ότι η $Y_1 + \ldots + Y_n$ έχει την κατανομή γάμμα $\Gamma(k/2, 1/2)$, όπου $k = k_1 + \ldots + k_n$. Όμως αυτή είναι η κατανομή $\chi^2(k)$ εξ ορισμού. □

Μπορούμε επίσης να εφαρμόσουμε το Θεώρημα 3 για να βρούμε την κατανομή του λόγου δύο ανεξάρτητων τυχαίων μεταβλητών Y_1 και Y_2 που έχουν κατανομές $\chi^2(k_1)$ και $\chi^2(k_2)$ αντίστοιχα. Είναι παράδοση στη Στατιστική τα αποτελέσματα να εκφράζονται συναρτήσει των κανονικοποιημένων μεταβλητών Y_1/k_1 και Y_2/k_2. Η κατανομή της

$$\frac{Y_1/k_1}{Y_2/k_2}$$

είναι γνωστή σαν κατανομή F με k_1 και k_2 βαθμούς ελευθερίας, και συμβολίζεται με $F(k_1, k_2)$.

Θεώρημα 8 *Έστω Y_1 και Y_2 ανεξάρτητες τυχαίες μεταβλητές με κατανομές τις $\chi^2(k_1)$ και $\chi^2(k_2)$. Τότε, η τυχαία μεταβλητή*

$$\frac{Y_1/k_1}{Y_2/k_2}$$

που ακολουθεί την κατανομή $F(k_1, k_2)$, έχει πυκνότητα τη συνάρτηση f που ορίζεται από τις $f(x) = 0$ αν $x \leq 0$ και

$$f(x) = \frac{(k_1/k_2)\Gamma[(k_1+k_2)/2](k_1x/k_2)^{(k_1/2)-1}}{\Gamma(k_1/2)\Gamma(k_2/2)[1+(k_1x/k_2)]^{(k_1+k_2)/2}}, \quad x > 0. \tag{39}$$

Απόδειξη: Από το Θεώρημα 3, η τυχαία μεταβλητή Y_1/Y_2 έχει πυκνότητα τη συνάρτηση g που δίνεται από την (24) με $\alpha_1 = k_1/2$ και $\alpha_2 = k_2/2$. Άρα η πυκνότητα f της k_2Y_1/k_1Y_2 δίνεται από την

$$f(x) = \frac{k_1}{k_2} g\left(\frac{k_1 x}{k_2}\right),$$

και η (39) προκύπτει από την (24). □

Μπορούμε να εφαρμόσουμε το προηγούμενο αποτέλεσμα στις τυχαίες μεταβλητές $X_1, \ldots, X_n$ που ορίστηκαν στην αρχή αυτής της παραγράφου. Έστω $1 \leq m < n$. Από το Θεώρημα 4, οι τυχαίες μεταβλητές

$$\frac{X_1^2 + \ldots + X_m^2}{\sigma^2} \quad \text{και} \quad \frac{X_{m+1}^2 + \ldots + X_n^2}{\sigma^2}$$

είναι ανεξάρτητες. Αφού ακολουθούν τις κατανομές $\chi^2(m)$ και $\chi^2(n-m)$ αντίστοιχα, βλέπουμε ότι η τυχαία μεταβλητή

$$\frac{(X_1^2+\ldots+X_m^2)/m}{(X_{m+1}^2+\ldots+X_n^2)/(n-m)}$$

έχει την κατανομή $F(m,n-m)$ και την πυκνότητα που δίνεται από την (39), με $k_1=m$ και $k_2=n-m$. Στον Τόμο II δίνουμε πίνακες των κατανομών F.

Η περίπτωση $m=1$ είναι ιδιαίτερα σημαντική. Η τυχαία μεταβλητή

$$\frac{X_1^2}{(X_2^2+\ldots+X_n^2)/(n-1)}$$

ακολουθεί την κατανομή $F(1,n-1)$. Χρησιμοποιώντας αυτό το γεγονός μπορούμε να βρούμε την κατανομή της

$$Y=\frac{X_1}{\sqrt{(X_2^2+\ldots+X_n^2)/(n-1)}}.$$

Αφού η X_1 έχει συμμετρική πυκνότητα και είναι ανεξάρτητη από την τυχαία μεταβλητή $\sqrt{(X_2^2+\ldots+X_n^2)/(n-1)}$, από το Θεώρημα 2 του Κεφαλαίου 5 συμπεραίνουμε ότι η Y έχει μία συμμετρική συνάρτηση πυκνότητας f_Y. Από το Παράδειγμα 5 του Κεφαλαίου 5, η πυκνότητα f_{Y^2} συνδέεται με την f_Y μέσω της

$$f_{Y^2}(z)=\frac{1}{2\sqrt{z}}\left(f_Y(-\sqrt{z})+f_Y(\sqrt{z})\right), \qquad Z>0.$$

Χρησιμοποιώντας τη συμμετρία της f_Y και θέτοντας $z=y^2$ βλέπουμε ότι

$$f_Y(y)=|y|f_{Y^2}(y^2).$$

Αφού η Y^2 έχει την πυκνότητα $F(1,n-1)$ που δίνεται από την (39) με $k_1=1$ και $k_2=k=n-1$, βλέπουμε τώρα ότι

$$f_Y(y)=\frac{|y|(1/k)\Gamma[(k+1)/2](y^2/k)^{-1/2}}{\Gamma(1/2)\Gamma(k/2)[1+(y^2/k)]^{(k+1)/2}}.$$

Αφού $\Gamma(1/2)=\sqrt{\pi}$, η τελευταία παράσταση γράφεται απλούστερα ως εξής:

$$(40) \qquad f_Y(y)=\frac{\Gamma[(k+1)/2][1+(y^2/k)]^{-(k+1)/2}}{\sqrt{k\pi}\,\Gamma(k/2)}, \qquad -\infty<y<\infty.$$

Αν μία τυχαία μεταβλητή έχει πυκνότητα που δίνεται από την (40), τότε λέμε ότι ακολουθεί κατανομή t με k βαθμούς ελευθερίας. Παρατηρούμε ότι η κατανομή t με ένα βαθμό ελευθερίας είναι ακριβώς η κατανομή Cauchy που συζητήσαμε στο Κεφάλαιο 4. Στον Τόμο II δίνουμε πίνακες των κατανομών t.

Η κατανομή της τυχαίας μεταβλητής

$$Y=\frac{X_1}{\sqrt{(X_2^2+\ldots+X_n^2)/(n-1)}},$$

η οποία είναι κατανομή t με $n-1$ βαθμούς ελευθερίας, εξαρτάται μόνο από το γεγονός ότι οι

$$\frac{X_1}{\sigma} \quad \kappa\alpha\iota \quad \frac{X_2^2+\ldots+X_n^2}{\sigma^2}$$

είναι ανεξάρτητες και ακολουθούν τις κατανομές $n(0,1)$ και $\chi^2(n-1)$ αντίστοιχα. Επομένως, καταλήγουμε στο εξής:

Θεώρημα 9 *Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν τις κατανομές $n(0,1)$ και $\chi^2(n-1)$ αντίστοιχα. Τότε, η*

$$\frac{X}{\sqrt{Y/k}}$$

ακολουθεί κατανομή t με k βαθμούς ελευθερίας.

6.7 Πολυδιάστατες αλλαγές μεταβλητών*

Έστω $X_1,\ldots,X_n$ συνεχείς τυχαίες μεταβλητές με από κοινού πυκνότητα την f. Έστω $Y_1,\ldots,Y_n$ τυχαίες μεταβλητές που ορίζονται μέσω των X_j. Σε αυτήν την παράγραφο θα συζητήσουμε μία μέθοδο για να βρίσκουμε την από κοινού πυκνότητα των Y_i συναρτήσει της f. Θα περιοριστούμε στην περίπτωση που οι Y_i ορίζονται σαν γραμμικές συναρτήσεις των X_j.

Υποθέτουμε λοιπόν ότι

$$Y_i = \sum_{j=1}^{n} a_{ij}X_j, \quad i=1,\ldots,n.$$

Οι σταθεροί συντελεστές a_{ij} προσδιορίζουν έναν $n\times n$ πίνακα

$$A=[a_{ij}]=\begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ a_{n1} & \ldots & a_{nn} \end{bmatrix}.$$

Για κάθε τέτοιον πίνακα θεωρούμε την ορίζουσά του

$$\det A=\begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ a_{n1} & \ldots & a_{nn} \end{vmatrix}.$$

Αν $\det A\neq 0$ τότε ορίζεται ο αντίστροφος πίνακας $B=[b_{ij}]$ για τον οποίο $BA=I$, δηλαδή

$$\sum_{k=1}^{n} b_{ik}a_{kj} = \begin{cases} 1 & ,\, i=j \\ 0 & ,\, i\neq j. \end{cases} \tag{41}$$

Οι σταθερές b_{ij} υπολογίζονται αν για κάθε i επιλύσουμε το σύστημα (41) που αποτελείται από n εξισώσεις με n αγνώστους, τους $b_{i1}, \ldots, b_{in}$. Αλλιώς, οι σταθερές b_{ij} ορίζονται μονοσήμαντα αν ζητήσουμε οι εξισώσεις

$$y_i = \sum_{j=1}^{n} a_{ij} x_j, \quad i = 1, \ldots, n,$$

να έχουν λύση τους

$$x_i = \sum_{j=1}^{n} b_{ij} y_j, \quad i = 1, \ldots, n. \tag{42}$$

Θεώρημα 10 *Έστω $X_1, \ldots, X_n$ συνεχείς τυχαίες μεταβλητές με από κοινού πυκνότητα την f και ας υποθέσουμε ότι οι τυχαίες μεταβλητές $Y_1, \ldots, Y_n$ ορίζονται από τις*

$$Y_i = \sum_{j=1}^{n} a_{ij} X_j, \quad i = 1, \ldots, n,$$

και ότι ο πίνακας $A = [a_{ij}]$ έχει μη μηδενική ορίζουσα $\det A$. Τότε, οι $Y_1, \ldots, Y_n$ έχουν από κοινού πυκνότητα την $f_{Y_1,\ldots,Y_n}$ που δίνεται από την

$$f_{Y_1,\ldots,Y_n}(y_1, \ldots, y_n) = \frac{1}{|\det A|} f(x_1, \ldots, x_n),$$

όπου τα x_i ορίζονται μέσω των y_i από την (42) ή σαν η μοναδική λύση των εξισώσεων $y_i = \sum_{j=1}^{n} a_{ij} x_j$.

Το θεώρημα αυτό, το οποίο δεν θα αποδείξουμε εδώ, είναι ισοδύναμο με ένα θεώρημα το οποίο αποδεικνύεται σε προχωρημένα μαθήματα απειροστικού λογισμού και αφορά «Ιακωβιανές» σε ένα γενικότερο πλαίσιο. Χρησιμοποιώντας αυτό το γενικό αποτέλεσμα του απειροστικού λογισμού, μπορούμε να επεκτείνουμε το παραπάνω θεώρημα και σε μη γραμμικές αλλαγές μεταβλητών. Θα περιγράψουμε αυτήν την επέκταση εν συντομία, αν και δεν θα τη χρειαστούμε στη συνέχεια.

Έστω ότι οι Y_i ορίζονται συναρτήσει των X_j από τις

$$Y_i = g_i(X_1, \ldots, X_n), \quad i = 1, \ldots, n.$$

Θεωρούμε τις αντίστοιχες εξισώσεις

$$y_i = g_i(x_1, \ldots, x_n), \quad i = 1, \ldots, n. \tag{44}$$

Ας υποθέσουμε ότι αυτές οι εξισώσεις ορίζουν τα x_j μονοσήμαντα συναρτήσει των y_i, ότι οι μερικές παράγωγοι $\partial y_i / \partial x_j$ υπάρχουν και είναι συνεχείς, και ότι η Ιακωβιανή

$$J(x_1, \ldots, x_n) = \begin{vmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \frac{\partial y_n}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_n} \end{vmatrix}$$

δεν μηδενίζεται πουθενά. Τότε, οι τυχαίες μεταβλητές $Y_1, \ldots, Y_n$ είναι συνεχείς και έχουν από κοινού πυκνότητα που δίνεται από την

$$f_{Y_1,\ldots,Y_n}(y_1, \ldots, y_n) = \frac{1}{|J(x_1, \ldots, x_n)|} f(x_1, \ldots, x_n), \tag{45}$$

όπου τα x_j ορίζονται πεπλεγμένα από τα y_i μέσω της (44). Αυτός ο τύπος αλλαγής μεταβλητής επεκτείνεται ακόμα περισσότερο αν ζητήσουμε οι συναρτήσεις g_i να ορίζονται σε κάποιο ανοικτό υποσύνολο S του $\mathbb{R}^n$ με την ιδιότητα

$$P((X_1, \ldots, X_n) \in S) = 1.$$

Στην ειδική περίπτωση που $y_i = \sum_{j=1}^n a_{ij} x_j$, βλέπουμε ότι $\partial y_i / \partial x_j = a_{ij}$ και η $J(x_1, \ldots, x_n)$ είναι σταθερή και ίση με $\det A = \det[a_{ij}]$. Είναι λοιπόν φανερό ότι η (45) ανάγεται στην (43) στη γραμμική περίπτωση.

Παράδειγμα 13. Έστω $X_1, \ldots, X_n$ ανεξάρτητες τυχαίες μεταβλητές που έχουν εκθετική πυκνότητα με παράμετρο λ. Ορίζουμε $Y_1, \ldots, Y_n$ θέτοντας $Y_i = X_1 + \ldots + X_i$, $1 \leq i \leq n$. Βρείτε την από κοινού πυκνότητα των $Y_1, \ldots, Y_n$.

Ο πίνακας $[a_{ij}]$ είναι ο

$$\begin{vmatrix} 1 & 0 & & \ldots & 0 \\ 1 & 1 & 0 & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & 0 \\ 1 & 1 & & \ldots & 1 \end{vmatrix}$$

Η ορίζουσά του είναι προφανώς ίση με 1. Οι εξισώσεις

$$y_i = x_1 + \ldots + x_i, \qquad i = 1, \ldots, n,$$

έχουν λύση την $x_1 = y_1$ και

$$x_i = y_i - y_{i-1}, \qquad i = 2, \ldots, n.$$

Η από κοινού πυκνότητα των $X_1, \ldots, X_n$ δίνεται από την

$$f(x_1, \ldots, x_n) = \begin{cases} \lambda^n e^{-\lambda(x_1 + \ldots + x_n)} & , x_1, \ldots, x_n > 0 \\ 0 & , \text{αλλιώς.} \end{cases} \tag{46}$$

Επομένως η από κοινού πυκνότητα $f_{Y_1,\ldots,Y_n}$ δίνεται από την

$$f_{Y_1,\ldots,Y_n}(y_1, \ldots, y_n) = \begin{cases} \lambda^n e^{-\lambda y_n} & , 0 < y_1 < \ldots < y_n \\ 0 & , \text{αλλιώς.} \end{cases} \tag{47}$$

Μπορούμε φυσικά να εφαρμόσουμε το θεώρημα και στην αντίστροφη κατεύθυνση. Αν οι $Y_1, \ldots, Y_n$ έχουν την από κοινού πυκνότητα στην (47) και οι τυχαίες μεταβλητές $X_1, \ldots, X_n$ ορίζονται από τις $X_1 = Y_1$ και $X_i = Y_i - Y_{i-1}$, $2 \leq i \leq n$, τότε οι X_j έχουν την από κοινού πυκνότητα f στην (46). Με άλλα λόγια, οι $X_1, \ldots, X_n$ είναι ανεξάρτητες και καθεμία ακολουθεί εκθετική κατανομή με παράμετρο λ. Θα χρησιμοποιήσουμε αυτό το αποτέλεσμα στο Κεφάλαιο 9 κατά τη μελέτη των ανελίξεων Poisson.

6.8 Ασκήσεις

1 Έστω X και Y συνεχείς τυχαίες μεταβλητές που έχουν από κοινού πυκνότητα την f. Βρείτε την από κοινού συνάρτηση κατανομής και την από κοινού πυκνότητα των τυχαίων μεταβλητών $W = a + bX$ και $Z = c + dY$, όπου $b > 0$ και $d > 0$. Δείξτε ότι αν οι X και Y είναι ανεξάρτητες, τότε οι W και Z είναι ανεξάρτητες.

2 Έστω X και Y συνεχείς τυχαίες μεταβλητές που έχουν από κοινού συνάρτηση κατανομής την F και από κοινού πυκνότητα την f. Βρείτε την από κοινού συνάρτηση κατανομής και την από κοινού πυκνότητα των τυχαίων μεταβλητών $W = X^2$ και $Z = Y^2$. Δείξτε ότι αν οι X και Y είναι ανεξάρτητες, τότε οι W και Z είναι ανεξάρτητες.

3 Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές ομοιόμορφα κατανεμημένες στο $(0, 1)$. Βρείτε τις
(α) $P(|X - Y| \le 0.5)$,
(β) $P\left(\left|\frac{X}{Y} - 1\right| \le 0.5\right)$,
(γ) $P(Y \ge X \mid Y \ge 1/2)$.

4 Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές που έχουν την κανονική πυκνότητα $n(0, \sigma^2)$. Βρείτε την $P(X^2 + Y^2 \le 1)$. *Υπόδειξη:* Χρησιμοποιήστε πολικές συντεταγμένες.

5 Υποθέτουμε ότι οι X και Y έχουν από κοινού πυκνότητα f που είναι ομοιόμορφη στο εσωτερικό του τριγώνου με κορυφές τα $(0, 0)$, $(2, 0)$, και $(1, 2)$. Βρείτε την $P(X \le 1, Y \le 1)$.

6 Υποθέτουμε ότι οι χρόνοι που χρειάζονται δύο φοιτητές για να λύσουν ένα πρόβλημα είναι ανεξάρτητοι και ακολουθούν εκθετική κατανομή με παράμετρο λ. Βρείτε την πιθανότητα ο πρώτος φοιτητής να χρειαστεί τουλάχιστον διπλάσιο χρόνο από τον δεύτερο για να λύσει το πρόβλημα.

7 Έστω X και Y συνεχείς τυχαίες μεταβλητές με από κοινού πυκνότητα την f που δίνεται από την $f(x, y) = \lambda^2 e^{-\lambda y}$, $0 \le x \le y$, και $f(x, y) = 0$ αλλιώς. Βρείτε τις περιθώριες πυκνότητες των X και Y. Βρείτε την από κοινού συνάρτηση κατανομής των X και Y.

8 Ορίζουμε $f(x, y) = c(y - x)^\alpha$, $0 \le x < y \le 1$, και $f(x, y) = 0$ αλλιώς.
(α) Για ποιές τιμές του α μπορούμε να επιλέξουμε c ώστε να κάνουμε την f πυκνότητα;
(β) Πως πρέπει να επιλέξουμε το c (όταν αυτό γίνεται) ώστε η f να γίνει πυκνότητα;
(γ) Βρείτε τις περιθώριες πυκνότητες της f.

9 Ορίζουμε $f(x, y) = ce^{-(x^2 - xy + 4y^2)/2}$, $-\infty < x, y < \infty$. Πώς πρέπει να επιλέξουμε το c ώστε η f να γίνει πυκνότητα; Βρείτε τις περιθώριες πυκνότητες της f.

10 Έστω X και Y ανεξάρτητες συνεχείς τυχαίες μεταβλητές με από κοινού πυκνότητα την f. Δώστε έναν τύπο για την πυκνότητα της $Z = Y - X$.

11 Έστω X και Y ανεξάρτητες συνεχείς τυχαίες μεταβλητές με περιθώριες πυκνότητες που δίνονται πιό κάτω. Βρείτε την πυκνότητα της $Z = X + Y$.

(α) Οι X και Y ακολουθούν εκθετική κατανομή με παραμέτρους λ_1 και λ_2, όπου $\lambda_1 \neq \lambda_2$.

(β) Η X είναι ομοιόμορφα κατανεμημένη στο $(0,1)$ και η Y ακολουθεί εκθετική κατανομή με παράμετρο λ.

12 Υποθέτουμε ότι οι X και Y έχουν από κοινού πυκνότητα την f της Άσκησης 8. Βρείτε την πυκνότητα της $Z = X + Y$.

13 Υποθέτουμε ότι οι X και Y είναι ανεξάρτητες και ομοιόμορφα κατανεμημένες στο (a,b). Βρείτε την πυκνότητα της $Z = |X - Y|$.

14 Έστω X και Y συνεχείς τυχαίες μεταβλητές με από κοινού πυκνότητα την f. Δώστε έναν τύπο για την πυκνότητα της $Z = aX + bY$, όπου $b \neq 0$.

15 Έστω f μία πυκνότητα Βήτα με παραμέτρους $\alpha_1 > 1$ και $\alpha_2 > 1$. Ποιά είναι η μέγιστη τιμή της f;

16 Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές με κανονικές πυκνότητες, τις $n(\mu_1, \sigma_1^2)$ και $n(\mu_2, \sigma_2^2)$. Βρείτε την πυκνότητα της $Z = X - Y$.

17 Επιλέγουμε τυχαία ένα σημείο στο επίπεδο με τέτοιον τρόπο ώστε οι x και y συντεταγμένες του να είναι ανεξάρτητα κατανεμημένες σύμφωνα με την κανονική πυκνότητα $n(0, \sigma^2)$. Βρείτε τη συνάρτηση πυκνότητας της τυχαίας μεταβλητής R που συμβολίζει την απόσταση του σημείου από την αρχή των αξόνων. (Η πυκνότητα αυτή μελετάται από τους ηλεκτρολόγους μηχανικούς, και είναι γνωστή σαν πυκνότητα Rayleigh.)

18 Έστω X και Y συνεχείς τυχαίες μεταβλητές με από κοινού πυκνότητα την f. Βρείτε τύπο για την πυκνότητα της $Z = XY$.

19 Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές που έχουν την κανονική πυκνότητα $n(0, \sigma^2)$. Δείξτε ότι η Y/X και η $Y/|X|$ έχουν πυκνότητα Cauchy.

20 Έστω X και Y όπως στην Άσκηση 19. Βρείτε την πυκνότητα της $Z = |Y|/|X|$.

21 Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν εκθετική κατανομή με παράμετρο λ. Βρείτε την πυκνότητα της $Z = Y/X$.

22 Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές με πυκνότητες γάμμα, τις $\Gamma(\alpha_1, \lambda)$ και $\Gamma(\alpha_2, \lambda)$ αντίστοιχα. Βρείτε την πυκνότητα της $Z = X/(X+Y)$. *Υπόδειξη:* Εκφράστε την Z συναρτήσει της Y/X.

23 Υποθέτουμε ότι οι X και Y έχουν από κοινού πυκνότητα την f που δίνεται παρακάτω. Βρείτε την δεσμευμένη πυκνότητα $f_{Y|X}$ σε κάθε περίπτωση:

(α) για την f της Άσκησης 7.
(β) για την f της Άσκησης 8.
(γ) για την f της Άσκησης 9.

24 Υποθέτουμε ότι οι X και Y κατανέμονται όπως στο Παράδειγμα 7. Βρείτε την $P(Y \leq 2 \mid X = 1)$.

25 Δείξτε ότι η περιθώρια πυκνότητα f_Y στο Παράδειγμα 9 είναι αρνητική διωνυμική με παραμέτρους α και $p = \beta/(\beta+1)$. *Υπόδειξη:* Χρησιμοποιήστε τον τύπο (36) του Κεφαλαίου 5.

26 Έστω Y μία διακριτή τυχαία μεταβλητή που ακολουθεί διωνυμική κατανομή με παραμέτρους n και p. Υποθέτουμε τώρα ότι το p μεταβάλλεται σαν μία τυχαία μεταβλητή Π που έχει πυκνότητα Βήτα με παραμέτρους α_1 και α_2. Βρείτε τη δεσμευμένη πυκνότητα της Π δεδομένου ότι $Y = y$.

27 Έστω ότι η Y ακολουθεί εκθετική κατανομή με παράμετρο λ. Υποθέτουμε ότι το λ μεταβάλλεται σαν μια τυχαία μεταβλητή Λ που έχει την πυκνότητα γάμμα $\Gamma(\alpha, \beta)$. Βρείτε την περιθώρια πυκνότητα της Y και τη δεσμευμένη πυκνότητα της Λ δεδομένου ότι $Y = y$.

28 Έστω X_1, X_2, X_3 οι τρείς συνιστώσες της ταχύτητας ενός μορίου κάποιου αερίου. Υποθέτουμε ότι οι X_1, X_2, X_3 είναι ανεξάρτητες, και καθεμία έχει την κανονική πυκνότητα $n(0, \sigma^2)$. Στη Φυσική λέμε ότι το μέτρο της ταχύτητας $Y = (X_1^2 + X_2^2 + X_3^2)^{1/2}$ ακολουθεί την κατανομή Maxwell. Βρείτε την f_Y.

29 Έστω $X_1, \ldots, X_n$ ανεξάρτητες τυχαίες μεταβλητές που έχουν την ίδια κανονική πυκνότητα. Δείξτε ότι υπάρχουν σταθερές A_n και B_n ώστε η

$$\frac{X_1 + \ldots + X_n - A_n}{B_n}$$

να έχει την ίδια πυκνότητα με την X_1.

30 Έστω X_1, X_2, X_3 ανεξάρτητες τυχαίες μεταβλητές ομοιόμορφα κατανεμημένες στο $(0, 1)$. Βρείτε την πυκνότητα της τυχαίας μεταβλητής $Y = X_1 + X_2 + X_3$. Βρείτε την $P(X_1 + X_2 + X_3 \leq 2)$.

31 Έστω ότι η X_1 επιλέγεται ομοιόμορφα στο $(0, 1)$, η X_2 επιλέγεται ομοιόμορφα στο $(0, X_1)$, και η X_3 επιλέγεται ομοιόμορφα στο $(0, X_2)$. Βρείτε την από κοινού πυκνότητα των X_1, X_2, X_3 και την περιθώρια πυκνότητα της X_3.

32 Έστω $U_1, \ldots, U_n$ ανεξάρτητες τυχαίες μεταβλητές, ομοιόμορφα κατανεμημένες στο $(0, 1)$. Θεωρούμε τις X_k, $k = 1, \ldots, n$ και R όπως στην Παράγραφο 6.5.
(α) Βρείτε την από κοινού πυκνότητα των X_1 και X_n.
(β) Βρείτε την πυκνότητα της R.
(γ) Βρείτε την πυκνότητα της X_k.

33 Έστω $U_1, \ldots, U_n$ ανεξάρτητες τυχαίες μεταβλητές, καθεμία από τις οποίες έχει εκθετική πυκνότητα με παράμετρο λ. Βρείτε τη συνάρτηση πυκνότητας της $X_1 = \min(U_1, \ldots, U_n)$.

34 Δώστε τύπο για την πυκνότητα $\chi^2(n)$.

35 Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν τις κατανομές $\chi^2(m)$ και $\chi^2(n)$ αντίστοιχα. Βρείτε την πυκνότητα της $Z = X/(X+Y)$. *Υπόδειξη:* Χρησιμοποιήστε το αποτέλεσμα της Άσκησης 22.

36 Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές που έχουν την τυπική κανονική κατανομή. Βρείτε την από κοινού πυκνότητα των $aX + bY$ και $bX - aY$, όπου $a^2 + b^2 > 0$. Χρησιμοποιήστε την απάντησή σας για να δώσετε δεύτερη απόδειξη του Θεωρήματος 2.

37 Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές, καθεμία από τις οποίες έχει πυκνότητα την f. Βρείτε την από κοινού πυκνότητα των X και $Z = X + Y$.

38 Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές, καθεμία από τις οποίες έχει εκθετική πυκνότητα με παράμετρο λ. Βρείτε τη δεσμευμένη πυκνότητα της X δεδομένου ότι $Z = X + Y = z$. *Υπόδειξη:* Χρησιμοποιήστε το αποτέλεσμα της Άσκησης 37.

39 Λύστε την Άσκηση 38 με την υπόθεση ότι οι X και Y είναι ομοιόμορφα κατανεμημένες στο $(0, c)$.

40 Έστω U και V ανεξάρτητες τυχαίες μεταβλητές, καθεμία από τις οποίες έχει την τυπική κανονική πυκνότητα. Θέτουμε $Z = \rho U + \sqrt{1-\rho^2}V$, όπου $-1 < \rho < 1$.
(α) Βρείτε την πυκνότητα της Z.
(β) Βρείτε την από κοινού πυκνότητα των U και Z.
(γ) Βρείτε την από κοινού πυκνότητα των $X = \mu_1 + \sigma_1 U$ και $Y = \mu_2 + \sigma_2 Z$, όπου $\sigma_1 > 0$ και $\sigma_2 > 0$. Αυτή η πυκνότητα είναι γνωστή σαν *διδιάστατη κανονική* πυκνότητα.
(δ) Βρείτε τη δεσμευμένη πυκνότητα της Y δεδομένου ότι $X = x$.

41 Έστω X και Y θετικές συνεχείς τυχαίες μεταβλητές που έχουν από κοινού πυκνότητα την f. Θέτουμε $W = Y/X$ και $Z = X + Y$. Βρείτε την από κοινού πυκνότητα των W και Z συναρτήσει της f. *Υπόδειξη:* Χρησιμοποιήστε την Εξίσωση (45).

42 Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές που έχουν τις πυκνότητες γάμμα $\Gamma(\alpha_1, \lambda)$ και $\Gamma(\alpha_2, \lambda)$ αντίστοιχα. Χρησιμοποιώντας την Άσκηση 41 δείξτε ότι οι Y/X και $X + Y$ είναι ανεξάρτητες τυχαίες μεταβλητές.

43 Έστω R και Θ ανεξάρτητες τυχαίες μεταβλητές εκ των οποίων η R έχει την πυκνότητα Rayleigh

$$f_R(r) = \begin{cases} \sigma^{-2} r e^{-r^2/2\sigma^2} & , r \geq 0 \\ 0 & , r < 0 \end{cases}$$

και η Θ είναι ομοιόμορφα κατανεμημένη στο $(-\pi, \pi)$. Δείξτε ότι οι $X = R\cos\Theta$ και $Y = R\sin\Theta$ είναι ανεξάρτητες τυχαίες μεταβλητές και έχουν την κανονική πυκνότητα $n(0, \sigma^2)$. *Υπόδειξη:* Χρησιμοποιήστε την (45).

Κεφάλαιο 7

Μέση τιμή και το Κεντρικό Οριακό Θεώρημα

Στις πρώτες τέσσερις παραγράφους αυτού του κεφαλαίου επεκτείνουμε τον ορισμό και τις ιδιότητες της μέσης τιμής σε τυχαίες μεταβλητές που δεν είναι αναγκαστικά διακριτές. Στην Παράγραφο 7.5 συζητάμε το Κεντρικό Οριακό Θεώρημα. Το θεώρημα αυτό, ένα από τα πιό σημαντικά στην θεωρία πιθανοτήτων, δικαιολογεί το γεγονός ότι πολλές συναρτήσεις κατανομής προσεγγίζονται από κατάλληλες κανονικές συναρτήσεις κατανομής.

7.1 Μέση τιμή συνεχών τυχαίων μεταβλητών

Στο Κεφάλαιο 4 δώσαμε τον ορισμό της μέσης τιμής μιάς διακριτής τυχαίας μεταβλητής X που έχει πυκνότητα f. Λέμε ότι η X έχει πεπερασμένη μέση τιμή αν $\sum_x |x|f(x) < \infty$, και σ' αυτήν την περίπτωση ορίζουμε την μέση τιμή $\mathbb{E}X$ της X θέτοντας

$$\mathbb{E}X = \sum_x x f(x).$$

Ο ευκολότερος τρόπος για να ορίσουμε την μέση τιμή μιάς συνεχούς τυχαίας μεταβλητής είναι κατ' αναλογία προς την διακριτή περίπτωση.

Ορισμός 1 *Έστω X συνεχής τυχαία μεταβλητή με πυκνότητα f. Λέμε ότι η X έχει πεπερασμένη μέση τιμή αν*

$$\int_{-\infty}^{\infty} |x|f(x)dx < \infty,$$

και στην περίπτωση αυτή ορίζουμε την μέση τιμή της X θέτοντας

$$\mathbb{E}X = \int_{-\infty}^{\infty} x f(x) dx.$$

Χρησιμοποιώντας αυτόν τον ορισμό μπορούμε εύκολα να υπολογίσουμε τις μέσες τιμές των συνεχών τυχαίων μεταβλητών που οι πυκνότητές τους περιγράφτηκαν στα Κεφάλαια 5 και 6.

Παράδειγμα 1 Υποθέτουμε ότι η X είναι ομοιόμορφα κατανεμημένη στο (a,b). Τότε,

$$\mathbb{E}X = \int_a^b x \left(\frac{1}{b-a}\right) dx = \left(\frac{1}{b-a}\right) \frac{x^2}{2}\Big|_a^b = \frac{a+b}{2}.$$

Παράδειγμα 2 Υποθέτουμε ότι η X έχει την πυκνότητα γάμμα $\Gamma(\alpha, \lambda)$. Τότε,

$$\begin{aligned}
\mathbb{E}X &= \int_0^{\infty} x \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} dx \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^{\infty} x^\alpha e^{-\lambda x} dx \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+1)}{\lambda^{\alpha+1}} \\
&= \frac{\alpha}{\lambda},
\end{aligned}$$

όπου χρησιμοποιήσαμε τους τύπους (34) και (36) του Κεφαλαίου 5. Θέτοντας $\alpha = 1$ βλέπουμε ότι αν η X έχει εκθετική πυκνότητα με παράμετρο λ, τότε $\mathbb{E}X = \lambda^{-1}$.

Παράδειγμα 3 Έστω ότι η X έχει την πυκνότητα Cauchy f που δίνεται από την

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad -\infty < x < \infty.$$

Τότε η X δεν έχει πεπερασμένη μέση τιμή. Γιατί

$$\begin{aligned}
\int_{-\infty}^{\infty} |x| \frac{1}{\pi(1+x^2)} dx &= \frac{2}{\pi} \int_0^{\infty} \frac{x}{1+x^2} dx \\
&= \frac{2}{\pi} \lim_{c\to\infty} \int_0^c \frac{x}{1+x^2} dx \\
&= \frac{1}{\pi} \lim_{c\to\infty} \log(1+x^2)\Big|_0^c \\
&= \infty.
\end{aligned}$$

7.2 Ένας γενικός ορισμός της μέσης τιμής

Ο ορισμός της μέσης τιμής που δώσαμε στην Παράγραφο 7.1 είναι ασφαλώς κατάλληλος από την υπολογιστική άποψη στην περίπτωση των συνεχών τυχαίων μεταβλητών που έχουν πυκνότητες. Για να δώσουμε όμως έναν γενικό ορισμό της μέσης τιμής, είναι προτιμότερο να επεκτείνουμε την έννοια της μέσης τιμής απευθείας από την διακριτή στην γενική περίπτωση. Θα παρουσιάσουμε μόνο τις βασικές ιδέες που οδηγούν στον γενικό ορισμό της μέσης τιμής. Οι ακριβείς λεπτομέρειες απαιτούν γνώσεις από την θεωρία μέτρου και ολοκλήρωσης. Στην συζήτηση που ακολουθεί θα υποθέσουμε ότι όλες οι τυχαίες μεταβλητές που εμφανίζονται είναι ορισμένες σε έναν σταθερό χώρο πιθανότητας $(\Omega, \mathcal{A}, P)$.

Έστω X και Y διακριτές τυχαίες μεταβλητές τέτοιες ώστε, για κάποιο $\varepsilon > 0$, $P(|X - Y| \leq \varepsilon) = 1$. Από τα Θεωρήματα 2(iii) και 3 του Κεφαλαίου 4 έπεται ότι αν η Y έχει πεπερασμένη μέση τιμή, τότε η X έχει πεπερασμένη μέση τιμή και $|\mathbb{E}X - \mathbb{E}Y| \leq \varepsilon$. Έπεται ακόμα ότι αν η Y δεν έχει πεπερασμένη μέση τιμή, τότε το ίδιο ισχύει και για την X. Με οποιονδήποτε γενικό ορισμό της μέσης τιμής, οι ιδιότητες αυτές θα πρέπει να συνεχίσουν να ισχύουν.

Ας υποθέσουμε τα παραπάνω, και ας θεωρήσουμε τυχούσα τυχαία μεταβλητή X. Έστω ότι θέλουμε να υπολογίσουμε την $\mathbb{E}X$ με σφάλμα το πολύ ίσο με ε, για κάποιο $\varepsilon > 0$. Το μόνο που χρειάζεται να κάνουμε είναι να βρούμε μία διακριτή τυχαία μεταβλητή Y που να ικανοποιεί την $P(|X-Y| \leq \varepsilon) = 1$ και να υπολογίσουμε την $\mathbb{E}Y$ με τις μεθόδους του Κεφαλαίου 4.

Είναι εύκολο να βρούμε τέτοιες προσεγγίσεις για την X. Έστω X_ε η διακριτή τυχαία μεταβλητή που ορίζεται από την

$$(1) \qquad X_\varepsilon = \varepsilon k, \quad \varepsilon k \leq X < \varepsilon(k+1), \quad k \in \mathbb{Z}.$$

Μπορούμε να ορίσουμε αυτήν την τυχαία μεταβλητή με την βοήθεια της συνάρτησης του ακεραίου μέρους, θέτοντας $X_\varepsilon = \varepsilon[X/\varepsilon]$. Αν $\varepsilon = 10^{-n}$ για κάποιον μη αρνητικό ακέραιο n, τότε ο $X_\varepsilon(\omega)$ προκύπτει από τον $X(\omega)$ αν γράψουμε τον $X(\omega)$ σε δεκαδική μορφή και διαγράψουμε όλα τα ψηφία που βρίσκονται n ή περισότερες θέσεις μακριά από την υποδιαστολή. Έπεται άμεσα από την (1) ότι

$$X(\omega) - \varepsilon < X_\varepsilon(\omega) \leq X(\omega), \quad \omega \in \Omega,$$

άρα $P(|X - X_\varepsilon| \leq \varepsilon) = 1$. Η πυκνότητα της X_ε ισούται με $P(\varepsilon k \leq X < \varepsilon(k+1))$ αν $x = \varepsilon k$ για τον ακέραιο k, και ισούται με 0 αλλιώς.

Η τυχαία μεταβλητή X_ε έχει πεπερασμένη μέση τιμή αν και μόνο αν

$$\sum_x |x| f_{X_\varepsilon}(x) = \sum_k |\varepsilon k| P(\varepsilon k \leq X < \varepsilon(k+1)) < \infty,$$

και τότε

$$\mathbb{E}(X_\varepsilon) = \sum_k \varepsilon k P(\varepsilon k \leq X < \varepsilon(k+1)) < \infty.$$

Όλες αυτές οι παραστάσεις εκφράζονται μέσω της F_X. Γιατί

$$P(\varepsilon k \leq X < \varepsilon(k+1)) = P(X < \varepsilon(k+1)) - P(X < \varepsilon k)$$

και από την ισότητα (5) του Κεφαλαίου 5, η $P(X < x) = F(x-)$ ισχύει για κάθε x. Το θεώρημα που ακολουθεί, το οποίο διατυπώνουμε χωρίς απόδειξη, θα χρησιμοποιηθεί για να δώσουμε έναν γενικό ορισμό της μέσης τιμής.

Θεώρημα 1 *Έστω X τυχαία μεταβλητή, και έστω $X_\varepsilon, \varepsilon > 0$, η τυχαία μεταβλητή που ορίζεται από την (1). Αν η X_ε έχει πεπερασμένη μέση τιμή για κάποιο $\varepsilon > 0$, τότε η X_ε έχει πεπερασμένη μέση τιμή για κάθε $\varepsilon > 0$ και το*

$$\lim_{\varepsilon \to 0} \mathbb{E} X_\varepsilon$$

υπάρχει και είναι πεπερασμένο.

Το θεώρημα αυτό και η συζήτηση που προηγήθηκε προτείνουν τον ακόλουθο γενικό ορισμό της μέσης τιμής.

Ορισμός 2 *Έστω X τυχαία μεταβλητή, και έστω $X_\varepsilon, \varepsilon > 0$, η τυχαία μεταβλητή που ορίζεται από την (1). Αν η X_ε έχει πεπερασμένη μέση τιμή για κάποιο $\varepsilon > 0$, λέμε ότι η X έχει πεπερασμένη μέση τιμή, και ορίζουμε την μέση τιμή της $\mathbb{E}X$ μέσω της*

$$\mathbb{E}X = \lim_{\varepsilon \to 0} \mathbb{E} X_\varepsilon .$$

Σε αντίθετη περίπτωση, λέμε ότι η X δεν έχει πεπερασμένη μέση τιμή.

Από την συζήτηση που προηγήθηκε του Θεωρήματος 1 έπεται ότι ο ορισμός της $\mathbb{E}X$ μπορεί να δοθεί μέσω της συνάρτησης κατανομής της X και ότι αν δύο τυχαίες μεταβλητές έχουν την ίδια συνάρτηση κατανομής, τότε οι μέσες τιμές τους συμπίπτουν (ή, δεν ορίζονται και οι δύο). Χρησιμοποιώντας τεχνικές από την θεωρία μέτρου και ολοκλήρωσης, μπορούμε να δείξουμε ότι ο Ορισμός 2 δίνει το ίδιο αποτέλεσμα με αυτό που δίνουν οι προηγούμενοι ορισμοί μας στις ειδικές περιπτώσεις που η X είναι διακριτή ή η X είναι συνεχής τυχαία μεταβλητή με πυκνότητα. Υπάρχει ένα ανάλογο του Θεωρήματος 1 του Κεφαλαίου 4, το οποίο διατυπώνουμε χωρίς απόδειξη. Σ' αυτό το θεώρημα, η ϕ μπορεί να είναι οποιαδήποτε συνάρτηση από αυτές που συναντάμε στον απειροστικό λογισμό.

Θεώρημα 2 *Έστω $X_1, \ldots, X_n$ συνεχείς τυχαίες μεταβλητές με από κοινού πυκνότητα την f, και έστω Z τυχαία μεταβλητή που ορίζεται με την βοήθεια των $X_1, \ldots, X_n$ από την $Z = \phi(X_1, \ldots, X_n)$. Τότε η Z έχει πεπερασμένη μέση τιμή αν και μόνο αν*

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} |\phi(x_1, \ldots, x_n)| f(x_1, \ldots, x_n) dx_1 \ldots dx_n < \infty,$$

και τότε,

$$\mathbb{E}Z = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \phi(x_1, \ldots, x_n) f(x_1, \ldots, x_n) dx_1 \ldots dx_n .$$

Μπορούμε να δείξουμε ότι οι βασικές ιδιότητες της μέσης τιμής που αποδείξαμε στο Κεφάλαιο 4 για διακριτές τυχαίες μεταβλητές ισχύουν γενικά. Ειδικότερα, τα Θεωρήματα 2,3, και 4 του Κεφαλαίου 4 ισχύουν και θα τα χρησιμοποιούμε ελέυθερα.

Όπως και στην διακριτή περίπτωση, μερικές φορές λέμε την $\mathbb{E}X$ *μέσο* της X. Οι ορισμοί των ροπών, των κεντρικών ροπών, της διασποράς, της τυπικής απόκλισης, της συνδιακύμανσης, και του συντελεστή συσχέτισης, που δόθηκαν στο Κεφάλαιο 4 για διακριτές τυχαίες μεταβλητές, βασίζονται μόνο στην έννοια της μέσης τιμής και επεκτείνονται άμεσα στην γενική περίπτωση.

Γενικά, όπως στην διακριτή περίπτωση, αν η X έχει ροπή τάξης r, τότε η X έχει ροπή τάξης k για κάθε $k \le r$. Τα Θεωρήματα 6 και 7 του Κεφαλαίου 4 ισχύουν επίσης γενικά. Ο αναγνώστης πρέπει να θυμηθεί τους ορισμούς και τα θεωρήματα του Κεφαλαίου 4 πριν προχωρήσει στην επόμενη παράγραφο.

7.3 Ροπές συνεχών τυχαίων μεταβλητών

Έστω X συνεχής τυχαία μεταβλητή με πυκνότητα την f και μέσο μ. Αν η X έχει πεπερασμένη ροπή τάξης m, τότε από το Θεώρημα 2

$$\mathbb{E}X^m = \int_{-\infty}^{\infty} x^m f(x)dx$$

και

$$\mathbb{E}(X-\mu)^m = \int_{-\infty}^{\infty} (x-\mu)^m f(x)dx.$$

Ειδικότερα, αν η X έχει πεπερασμένη ροπή δεύτερης τάξης, η διασπορά της σ^2 δίνεται από την

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx.$$

Παρατηρήστε ότι $\sigma^2 > 0$. Γιατί αν $\sigma^2 = 0$, τότε το επιχείρημα της Παραγράφου 4.3 δείχνει ότι $P(X = \mu) = 1$, το οποίο αντιφάσκει προς την υπόθεση ότι η X είναι συνεχής τυχαία μεταβλητή.

Παράδειγμα 4 Υποθέτουμε ότι η X έχει την πυκνότητα γάμμα $\Gamma(\alpha, \lambda)$. Βρείτε τις ροπές και την διασπορά της X.

Η ροπή τάξης m της X ισούται με

$$\begin{aligned}\mathbb{E}X^m &= \int_0^{\infty} x^m \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} dx \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^{\infty} x^{m+\alpha-1} e^{-\lambda x} dx,\end{aligned}$$

οπότε από τις (34) και (36) του Κεφαλαίου 5,

$$\begin{aligned}(2) \qquad \mathbb{E}X^m &= \frac{\lambda^\alpha \Gamma(m+\alpha)}{\lambda^{m+\alpha}\Gamma(\alpha)} \\ &= \frac{\alpha(\alpha+1)\dots(\alpha+m-1)}{\lambda^m}.\end{aligned}$$

Η διασπορά της X ισούται με

$$\sigma^2 = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \frac{\alpha(\alpha+1)}{\lambda^2} - \left(\frac{\alpha}{\lambda}\right)^2 = \frac{\alpha}{\lambda^2}.$$

Θέτοντας $\alpha = 1$, βλέπουμε ότι αν η X έχει εκθετική πυκνότητα με παράμετρο λ, τότε $\mathbb{E}X^m = m!\lambda^{-m}$ και η X έχει διασπορά λ^{-2}. Για μία δεύτερη ειδική περίπτωση, ας υποθέσουμε ότι η X έχει την κατανομή $\chi^2(n)$, δηλαδή την κατανομή $\Gamma(n/2, 1/2)$. Τότε,

$$\mathbb{E}X = \frac{n/2}{1/2} = n \quad \textit{και} \quad \mathrm{Var}X = \frac{n/2}{(1/2)^2} = 2n.$$

Συχνά επωφελούμαστε από την συμμετρία για να υπολογίσουμε ροπές. Για παράδειγμα, αν η X έχει συμμετρική πυκνότητα, αν η $\mathbb{E}X^m$ υπάρχει, και αν m είναι ένας περιττός θετικός ακέραιος, τότε $\mathbb{E}X^m = 0$. Για να το δείτε, παρατηρήστε ότι από το Θεώρημα 2 του Κεφαλαίου 5, οι X και $-X$ έχουν την ίδια συνάρτηση κατανομής. Άρα, οι X^m και $(-X)^m = -X^m$ έχουν την ίδια συνάρτηση κατανομής, άρα και την ίδια μέση τιμή. Μ' άλλα λόγια, $\mathbb{E}X^m = \mathbb{E}(-X^m) = -\mathbb{E}X^m$, το οποίο σημαίνει ότι $\mathbb{E}X^m = 0$.

Παράδειγμα 5 Έστω ότι η X έχει την κανονική πυκνότητα $n(\mu, \sigma^2)$. Βρείτε τον μέσο και τις κεντρικές ροπές της X.

Η τυχαία μεταβλητή $X - \mu$ έχει την κανονική πυκνότητα $n(0, \sigma^2)$, η οποία είναι συμμετρική πυκνότητα. Άρα $\mathbb{E}(X-\mu)^2 = 0$ αν ο m είναι περιττός θετικός ακέραιος. Ειδικότερα $\mathbb{E}(X - \mu) = 0$, δηλαδή βλέπουμε ότι η παράμετρος μ της κανονικής πυκνότητας $n(\mu, \sigma^2)$ είναι απλώς ο μέσος της πυκνότητας. Έπεται τώρα ότι όλες οι περιττές κεντρικές ροπές της X είναι ίσες με μηδέν. Για να υπολογίσουμε τις άρτιες κεντρικές ροπές χρησιμοποιούμε το γεγονός (βλέπε Παράγραφο 5.3.3) ότι η $Y = (X-\mu)^2$ έχει την πυκνότητα γάμμα $\Gamma(1/2, 1/2\sigma^2)$. Αφού $\mathbb{E}(X-\mu)^m = \mathbb{E}Y^{m/2}$ αν ο m είναι άρτιος, από το Παράδειγμα 4 βλέπουμε ότι

$$\begin{aligned}
\mathbb{E}(X-\mu)^m &= \frac{\Gamma\left(\frac{m+1}{2}\right)}{\left(\frac{1}{2\sigma^2}\right)^{m/2}\Gamma\left(\frac{1}{2}\right)} \\
&= \frac{\frac{1}{2}\frac{3}{2}\cdots\left(\frac{m-1}{2}\right)}{\left(\frac{1}{2\sigma^2}\right)^{m/2}} \\
&= \sigma^m 1.3\ldots(m-1).
\end{aligned}$$

Χρησιμοποιώντας τους τύπους (35) και (38) του Κεφαλαίου 5, καταλήγουμε στον ισοδύναμο τύπο

$$\mathbb{E}(X-\mu)^m = \frac{m!}{2^{m/2}\left(\frac{m}{2}\right)!}\sigma^m. \tag{3}$$

Ειδικότερα, σ^2 είναι η διασπορά της X, και $\mathbb{E}(X-\mu)^4 = 3\sigma^4$.

Έστω X και Y συνεχείς τυχαίες μεταβλητές με από κοινού πυκνότητα την f, μέσους μ_X και μ_Y, και πεπερασμένες δεύτερες ροπές. Τότε, η συνδιακύμανσή τους δίνεται από την

$$\mathbb{E}(X-\mu_X)(Y-\mu_Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_X)(y-\mu_Y)f(x,y)dxdy. \tag{4}$$

Παράδειγμα 6 Έστω ότι οι X και Y έχουν από κοινού πυκνότητα την f του Παραδείγματος 2 του Κεφαλαίου 6. Βρείτε τον συντελεστή συσχέτισης των X και Y.

Σύμφωνα με το Παράδειγμα 2 του Κεφαλαίου 6

$$\begin{aligned} f(x,y) &= \frac{\sqrt{3}}{4\pi}e^{-[(x^2-xy+y^2)/2]} \\ &= \frac{\sqrt{3}}{4\pi}e^{-3x^2/8}e^{-[(y-x/2)^2/2]}. \end{aligned}$$

Στο παράδειγμα εκείνο είδαμε ότι καθεμία από τις X και Y έχει την κανονική πυκνότητα $n(0,4/3)$. Άρα $\mu_X=\mu_Y=0$ και $\mathrm{Var}X=\mathrm{Var}Y=4/3$. Από την ισότητα (4) και τον δεύτερο τρόπο γραφής της f, έχουμε

$$\mathbb{E}XY = \frac{\sqrt{3}}{2\sqrt{2\pi}}\int_{-\infty}^{\infty}xe^{-(3x^2/8)}dx\int_{-\infty}^{\infty}y\frac{1}{\sqrt{2\pi}}e^{-[(y-x/2)^2/2]}dy.$$

Όμως

$$\int_{-\infty}^{\infty}y\frac{1}{\sqrt{2\pi}}e^{-[(y-x/2)^2/2]}dy = \int_{-\infty}^{\infty}\left(u+\frac{x}{2}\right)\frac{1}{\sqrt{2\pi}}e^{-(u^2/2)}du = \frac{x}{2},$$

επομένως

$$\begin{aligned} \mathbb{E}XY &= \frac{1}{2\left(\frac{2}{\sqrt{3}}\right)\sqrt{2\pi}}\int_{-\infty}^{\infty}x^2e^{-(3x^2/8)}dx \\ &= \frac{1}{2}\int_{-\infty}^{\infty}x^2n(x;0,4/3)dx \\ &= \frac{1}{2}\frac{4}{3}=\frac{2}{3}. \end{aligned}$$

Ο συντελεστής συσχέτισης ρ των X και Y δίνεται από την

$$\rho = \frac{\mathbb{E}XY}{\sqrt{\mathrm{Var}X}\sqrt{\mathrm{Var}Y}} = \frac{2/3}{\sqrt{4/3}\sqrt{4/3}} = \frac{1}{2}.$$

Παράδειγμα 7 Έστω $U_1,\ldots,U_n$ ανεξάρτητες τυχαίες μεταβλητές ομοιόμορφα κατανεμημένες στο $(0,1)$. Θεωρούμε τις

$$X = \min(U_1,\ldots,U_n)$$

και

$$Y = \max(U_1, \ldots, U_n).$$

Βρείτε τις ροπές των X και Y και τον συντελεστή συσχέτισής τους.

Οι τυχαίες αυτές μεταβλητές μελετήθηκαν στην Παράγραφο 6.5 (όπου τις συμβολίζαμε με X_1 και X_n). Εξειδικεύοντας τα αποτελέσματα εκείνης της παραγράφου στις U_i που είναι ομοιόμορφα κατανεμημένες, βλέπουμε ότι οι X και Y έχουν από κοινού πυκνότητα f που ορίζεται από την

$$(5) \qquad f(x,y) = \begin{cases} n(n-1)(y-x)^{n-2} & , 0 \le x \le y \le 1, \\ 0 & , \text{αλλιώς.} \end{cases}$$

Εκείνοι οι αναγνώστες που παρέλειψαν την Παράγραφο 6.5 μπορούν να σκέφτονται αυτό το πρόβλημα σαν ένα πρόβλημα εύρεσης των ροπών και του συντελεστή συσχέτισης δύο τυχαίων μεταβλητών X και Y που η από κοινού πυκνότητά τους δίνεται από την (5).

Η ροπή τάξης m της X δίνεται από την

$$\begin{aligned} \mathbb{E}X^m &= n(n-1)\int_0^1 x^m dx \int_x^1 (y-x)^{n-2} dy \\ &= n(n-1)\int_0^1 x^m dx \frac{(y-x)^{n-1}}{n-1}\Big|_{y=x}^{y=1} \\ &= n\int_0^1 x^m (1-x)^{n-1} dx. \end{aligned}$$

Το ορισμένο ολοκλήρωμα που εμφανίζεται σ' αυτήν την παράσταση είναι ένα ολοκλήρωμα Βήτα που δίνεται από τον τύπο (19) του Κεφαλαίου 6. Από τον τύπο αυτό βλέπουμε ότι

$$\mathbb{E}X^m = \frac{n\Gamma(m+1)\Gamma(n)}{\Gamma(m+n+1)} = \frac{m!n!}{(m+n)!}.$$

Ειδικότερα, $\mathbb{E}X = 1/(n+1)$ και $\mathbb{E}X^2 = 2/(n+1)(n+2)$. Έπεται ότι

$$\mathrm{Var}X = (\mathbb{E}X^2) - (\mathbb{E}X)^2 = \frac{n}{(n+1)^2(n+2)}.$$

Η ροπή τάξης m της Y δίνεται από την

$$\begin{aligned} \mathbb{E}Y^m &= n(n-1)\int_0^1 y^m dy \int_0^y (y-x)^{n-2} dx \\ &= n(n-1)\int_0^1 y^m dy \frac{(y-x)^{n-1}(-1)}{n-1}\Big|_{x=0}^{x=y} \\ &= n\int_0^1 y^{m+n-1} dy \\ &= \frac{n}{m+n}. \end{aligned}$$

Άρα $\mathbb{E}Y = n/(n+1)$ και

$$\mathrm{Var}Y = \frac{n}{n+2} - \left(\frac{n}{n+1}\right)^2 = \frac{n}{(n+1)^2(n+2)}.$$

Θα μπορούσαμε να υπολογίσουμε αυτές τις ποσότητες και από τις περιθώριες πυκνότητες των X και Y.

Για να βρούμε την συνδιακύμανση των X και Y ξεκινάμε με την

$$\mathbb{E}XY = n(n-1)\int_0^1 y dy \int_0^y x(y-x)^{n-2}dx.$$

Αφού

$$x(y-x)^{n-2} = y(y-x)^{n-2} - (y-x)^{n-1},$$

βλέπουμε ότι

$$\begin{aligned}
\mathbb{E}XY &= n(n-1)\int_0^1 y^2 dy \int_0^y (y-x)^{n-2}dx \\
&- n(n-1)\int_0^1 y dy \int_0^y (y-x)^{n-1}dx \\
&= n(n-1)\int_0^1 y^2 dy \frac{(y-x)^{n-1}(-1)}{n-1}\Big|_{x=0}^{x=y} \\
&- n(n-1)\int_0^1 y dy \frac{(y-x)^{n}(-1)}{n}\Big|_{x=0}^{x=y} \\
&= n\int_0^1 y^{n+1}dy - (n-1)\int_0^1 y^{n+1}dy \\
&= \frac{1}{n+2}.
\end{aligned}$$

Επομένως

$$\begin{aligned}
\mathrm{Cov}(X,Y) &= \mathbb{E}XY - \mathbb{E}X\mathbb{E}Y \\
&= \frac{1}{n+2} - \frac{n}{(n+1)^2} \\
&= \frac{1}{(n+1)^2(n+2)}.
\end{aligned}$$

Τέλος βρίσκουμε τον συντελεστή συσχέτισης των X και Y,

$$\begin{aligned}
\rho &= \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}X}\sqrt{\mathrm{Var}Y}} \\
&= \frac{1}{(n+1)^2(n+2)}\sqrt{\frac{n}{(n+1)^2(n+2)}} \\
&= \frac{1}{n}.
\end{aligned}$$

7.4 Δεσμευμένη μέση τιμή

Έστω X και Y συνεχείς τυχαίες μεταβλητές με από κοινού πυκνότητα την f, και ας υποθέσουμε ότι η Y έχει πεπερασμένη μέση τιμή. Στην παράγραφο 6.3 ορίσαμε την δεσμευμένη πυκνότητα της Y δεδομένου ότι $X = x$ μέσω της

$$f_{Y|X}(y|x) = \begin{cases} \frac{f(x,y)}{f_X(x)} & , 0 < f_X(x) < \infty, \\ 0 & , \text{αλλιώς.} \end{cases}$$

Για κάθε x που ικανοποιεί την $0 < f_X(x) < \infty$ η συνάρτηση $f_{Y|X}(y|x)$, $-\infty < y < \infty$, είναι συνάρτηση πυκνότητας σύμφωνα με τον ορισμό 5 του Κεφαλαίου 5. Επομένως μπορούμε να μιλάμε για τις διάφορες ροπές αυτής της πυκνότητας. Ο μέσος της λέγεται *δεσμευμένη μέση τιμή* της Y δεδομένου ότι $X = x$ και συμβολίζεται με $\mathbb{E}[Y|X = x]$ ή $\mathbb{E}[Y|x]$. Άρα

$$\begin{aligned} (6) \qquad \mathbb{E}[Y|X = x] &= \int_{-\infty}^{\infty} yf(y|x)dy \\ &= \int_{-\infty}^{\infty} \frac{yf(x,y)}{f_X(x)}dy \end{aligned}$$

όταν $0 < f_X(x) < \infty$. Ορίζουμε $\mathbb{E}[Y|X = x] = 0$ αλλιώς. Στην Στατιστική, η συνάρτηση m που ορίζεται από την $m(x) = \mathbb{E}[Y|X = x]$ λέγεται *συνάρτηση παλινδρόμησης* της Y επί της X.

Οι δεσμευμένες μέσες τιμές εμφανίζονται σε πολλά προβλήματα στατιστικής, όπως θα δούμε στον Τόμο ΙΙ. Είναι επίσης σημαντικές, από μία πιό γενική οπτική γωνία, στην προχωρημένη θεωρία πιθανοτήτων. Θα περιοριστούμε σε μερικά στοιχειώδη παραδείγματα δεσμευμένων μέσων τιμών. Η γενική θεωρία είναι αρκετά πολύπλοκη και δεν θα την χρειαστούμε σ' αυτό το βιβλίο.

Παράδειγμα 8 Έστω ότι οι X και Y έχουν από κοινού πυκνότητα την f του Παραδείγματος 2 του Κεφαλαίου 6. Βρείτε την δεσμευμένη μέση τιμή της Y δεδομένου ότι $X = x$.

Στο Παράδειγμα 7 του Κεφαλαίου 6 είδαμε ότι η δεσμευμένη πυκνότητα της Y δεδομένου ότι $X = x$ είναι η κανονική πυκνότητα $n(x/2, 1)$, η οποία έχει μέσο $x/2$. Άρα

$$\mathbb{E}[Y|X = x] = \frac{x}{2}.$$

Παράδειγμα 9 Έστω X και Y τυχαίες μεταβλητές με από κοινού πυκνότητα την f που δίνεται από την (5). Στην προηγούμενη παράγραφο υπολογίσαμε διάφορες ροπές των X και Y. Εδώ θα υπολογίσουμε την δεσμευμένη πυκνότητα και την δεσμευμένη μέση τιμή της Y δεδομένου ότι $X = x$.

Η περιθώρια πυκνότητα της X δίνεται από την

$$\begin{aligned} f_X(x) &= n(n-1)\int_x^1 (y-x)^{n-2}dy \\ &= n(1-x)^{n-1}, \quad 0 \le x \le 1, \end{aligned}$$

και $f_X(x) = 0$ αλλιώς. Άρα, αν $0 \le x \le 1$,

$$f(y|x) = \begin{cases} \frac{(n-1)(y-x)^{n-2}}{(1-x)^{n-1}} & , x \le y < 1, \\ 0 & , \text{αλλιώς}. \end{cases}$$

Επομένως, αν $0 \le x < 1$,

$$\begin{aligned} \mathbb{E}[Y|X=x] &= \int_{-\infty}^{\infty} y f(y|x) dy \\ &= (n-1)(1-x)^{1-n} \int_x^1 y(y-x)^{n-2} dy \\ &= (n-1)(1-x)^{1-n} \int_x^1 [(y-x)^{n-1} - x(y-x)^{n-2}] dy \\ &= (n-1)(1-x)^{1-n} \left[\frac{(1-x)^n}{n} + \frac{x(1-x)^{n-1}}{n-1} \right] \\ &= \frac{(n-1)(1-x)}{n} + x \\ &= \frac{n-1+x}{n}. \end{aligned}$$

Συχνά βολεύει να υπολογίζουμε την μέση τιμή της Y χρησιμοποιώντας τον τύπο

$$\mathbb{E}Y = \int_{-\infty}^{\infty} \mathbb{E}[Y|X=x] f_X(x) dx. \tag{7}$$

Ο τύπος αυτός έπεται άμεσα από την (6). Γιατί

$$\begin{aligned} \int_{-\infty}^{\infty} \mathbb{E}[Y|X=x] f_X(x) dx &= \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} y f(x,y) dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x,y) dx dy \\ &= \mathbb{E}Y. \end{aligned}$$

Εφαρμόζοντας αυτόν τον τύπο στο Παράδειγμα 9, παίρνουμε

$$\begin{aligned} \mathbb{E}Y &= \int_0^1 \frac{n-1+x}{n} n(1-x)^{n-1} dx \\ &= n \int_0^1 (1-x)^{n-1} dx - \int_0^1 (1-x)^n dx \\ &= 1 - \frac{1}{n+1} = \frac{n}{n+1}, \end{aligned}$$

κάτι που συμφωνεί με την απάντηση που δώσαμε στο Παράδειγμα 7.

Μπορούμε φυσικά να ορίσουμε με τον ίδιο τρόπο την δεσμευμένη μέση τιμή για διακριτές τυχαίες μεταβλητές. Στο Κεφάλαιο 4 υπήρχαν κάποιες σχετικές ασκήσεις.

7.5 Το Κεντρικό Οριακό Θεώρημα

Στην παράγραφο αυτή, με $X_1, X_2, \ldots$ συμβολίζουμε ανεξάρτητες, ισοκατανεμημένες τυχαίες μεταβλητές με μέσο μ και πεπερασμένη μη μηδενική διασπορά σ^2. Ενδιαφερόμαστε για την κατανομή της $S_n = X_1 + \ldots + X_n$. Παρατηρήστε ότι η S_n έχει μέσο $n\mu$ και διασπορά $n\sigma^2$.

Ας υποθέσουμε ότι η X_1 έχει πυκνότητα f. Τότε, για κάθε $n \geq 1$, η S_n θα έχει μία πυκνότητα f_{S_n}. Έχουμε $f_{S_1} = f$, και οι άλλες πυκνότητες υπολογίζονται διαδοχικά αν χρησιμοποιήσουμε τους τύπους των Κεφαλαίων 3 και 6 για την πυκνότητα του αθροίσματος δύο ανεξάρτητων τυχαίων μεταβλητών. Έχουμε

$$f_{S_n}(x) = \sum_y f_{S_{n-1}}(y) f(x-y)$$

ή

$$f_{S_n}(x) = \int_{-\infty}^{\infty} f_{S_{n-1}}(y) f(x-y) dy$$

ανάλογα με το αν η X_1 είναι διακριτή ή συνεχής τυχαία μεταβλητή. Για κάποιες ειδικές f (π.χ. την διωνυμική, την αρνητική διωνυμική, την Poisson, την κανονική, και την γάμμα), μπορούμε να βρούμε απλούς τύπους για την f_{S_n}. Γενικά, όμως, πρέπει να καταφύγουμε σε αριθμητικές μεθόδους.

Ένα από τα πιό σημαντικά και αξιοσημείωτα αποτελέσματα της θεωρίας πιθανοτήτων είναι ότι για μεγάλες τιμές του n η κατανομή της S_n εξαρτάται από την κατανομή της X_1 ουσιαστικά μόνο μέσω των μ και σ^2. Αποτελέσματα αυτού του είδους περιγράφονται ευκολότερα αν θεωρήσουμε την κανονικοποιημένη τυχαία μεταβλητή

$$S_n^* = \frac{S_n - \mathbb{E}S_n}{\sqrt{\mathrm{Var} S_n}} = \frac{S_n - n\mu}{\sigma\sqrt{n}},$$

η οποία έχει μέσο 0 και διασπορά 1.

Για να πάρετε μια ιδέα για την συμπεριφορά της συνάρτησης κατανομής της S_n^* καθώς το $n \to \infty$, ας θεωρήσουμε πρώτα μια ειδική περίπτωση στην οποία αυτή η συνάρτηση κατανομής υπολογίζεται εύκολα και ακριβώς. Υποθέτουμε λοιπόν ότι η X_1 είναι κανονικά κατανεμημένη με μέσο μ και διασπορά σ^2. Τότε, από τα αποτελέσματα του Κεφαλαίου 6, η S_n^* είναι κανονικά κατανεμημένη με μέσο 0 και διασπορά 1 ή, με άλλα λόγια, η S_n^* έχει την τυπική κανονική συνάρτηση κατανομής Φ.

Υποθέτουμε στη συνέχεια ότι η X_1 παίρνει τις τιμές 0 και 1 με πιθανότητες p και $1-p$ αντίστοιχα. Τότε, όπως είδαμε στο Κεφάλαιο 3, η S_n ακολουθεί διωνυμική κατανομή με παραμέτρους n και p. Δηλαδή,

$$P(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Οι DeMoivre (1667-1754) και Laplace (1749-1827) ανακάλυψαν ότι σ' αυτήν την περίπτωση η συνάρτηση κατανομής της S_n^* πλησιάζει την Φ, την τυπική κανονική συνάρτηση κατανομής, καθώς το $n \to \infty$.

Πιό πρόσφατα, βρέθηκαν πολλές επεκτάσεις του οριακού θεωρήματος των DeMoivre και Laplace, οι οποίες είναι όλες γνωστές σαν «κεντρικά οριακά θεωρήματα». Το απλούστερο και γνωστότερο από αυτά τα αποτελέσματα αποδείχτηκε από τον Lindeberg το 1922:

Θεώρημα 3 (Κεντρικό Οριακό Θεώρημα) *Έστω $X_1, X_2, \ldots$ ανεξάρτητες, ισοκατανεμημένες τυχαίες μεταβλητές με μέσο μ και πεπερασμένη μη μηδενική διασπορά σ^2. Θέτουμε $S_n = X_1 + \ldots + X_n$. Τότε*

$$\lim_{n\to\infty} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \le x\right) = \Phi(x), \qquad -\infty < x < \infty. \tag{8}$$

Η γενικότητα αυτού του θεωρήματος είναι αξιοσημείωτη. Η τυχαία μεταβλητή X_1 μπορεί να είναι διακριτή, συνεχής, ή τίποτα από τα δύο. Επιπλέον, το συμπέρασμα ισχύει ακόμα κι αν η X_1 δεν έχει ροπές υψηλότερης τάξης από την δεύτερη. Ένα άλλο εντυπωσιακό σημείο του θεωρήματος είναι ότι η οριακή συνάρτηση κατανομής της S_n^* είναι ανεξάρτητη από την συγκεκριμένη κατανομή της X_1 (αρκεί βέβαια να ικανοποιούνται οι υποθέσεις του θεωρήματος). Δεν πρέπει, όμως, να προκαλεί έκπληξη το γεγονός ότι η Φ είναι αυτή η οριακή κατανομή. Γιατί είδαμε ότι αυτό ισχύει στην περίπτωση που η X_1 είναι είτε η κανονική ή η διωνυμική κατανομή.

Η απόδειξη του Κεντρικού Οριακού Θεωρήματος αναβάλλεται για το Κεφάλαιο 8, γιατί απαιτεί προχωρημένες τεχνικές σχετικές με τις χαρακτηριστικές συναρτήσεις, τις οποίες θα μελετήσουμε εκεί. Μπορούμε να δώσουμε μία στοιχειώδη αλλά κοπιαστική απόδειξη του οριακού θεωρήματος των DeMoivre και Laplace, την ειδική περίπτωση του Κεντρικού Οριακού Θεωρήματος κατά την οποία η X_1 είναι διωνυμικά κατανεμημένη. Υπάρχουν στοιχειώδεις τρόποι για να εξηγηθεί το Κεντρικό Οριακό Θεώρημα, δεν είναι όμως αποδείξεις. Ένας τέτοιος τρόπος είναι να δείξουμε ότι, για οποιονδήποτε θετικό ακέραιο m, αν η X_1 έχει πεπερασμένη ροπή τάξης m, τότε το

$$\lim_{n\to\infty} \mathbb{E}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \le x\right)^m$$

υπάρχει και ισούται με την ροπή τάξης m της τυπικής κανονικής κατανομής.

Στο στάδιο που βρισκόμαστε, είναι προτιμότερο να καταλάβουμε τι σημαίνει το Κεντρικό Οριακό Θεώρημα και πώς εφαρμόζεται σε τυπικές εφαρμογές.

Παράδειγμα 10 Έστω $X_1, X_2, \ldots$ ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν κατανομή Poisson με παράμετρο λ. Στο Κεφάλαιο 4 είδαμε ότι $\mu = \sigma^2 = \lambda$ και ότι η S_n ακολουθεί κατανομή Poisson με παράμετρο $n\lambda$. Από το Κεντρικό Οριακό Θεώρημα έπεται ότι

$$\lim_{n\to\infty} P\left(\frac{S_n - n\lambda}{\sqrt{n\lambda}} \le x\right) = \Phi(x), \qquad -\infty < x < \infty.$$

Μπορούμε να επεκτείνουμε το αποτέλεσμα αυτού του παραδείγματος και να δείξουμε ότι αν η X_t είναι τυχαία μεταβλητή που ακολουθεί κατανομή Poisson με παράμετρο $\lambda = t$, τότε

$$\lim_{t\to\infty} P\left(\frac{X_t - \mathbb{E}X_t}{\sqrt{\mathrm{Var}X_t}} \le x\right) = \Phi(x), \qquad -\infty < x < \infty. \tag{9}$$

Η ισότητα (9) ισχύει επίσης αν η X_t είναι τυχαία μεταβλητή που ακολουθεί κατανομή γάμμα $\Gamma(t,\lambda)$ για σταθερό λ, ή την αρνητική διωνυμική κατανομή με παραμέτρους $\alpha = t$ και σταθερό p.

7.5.1 Κανονικές προσεγγίσεις

Το Κεντρικό Οριακό Θεώρημα υποδεικνύει ισχυρά ότι για μεγάλα n θα μπορούσαμε προσεγγιστικά να γράψουμε

$$P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right) \approx \Phi(x), \qquad -\infty < x < \infty,$$

ή ισοδύναμα

$$\begin{aligned} (10) \qquad P(S_n \leq x) &\approx \Phi\left(\frac{x - n\mu}{\sigma\sqrt{n}}\right) \\ &= \Phi\left(\frac{x - \mathbb{E}S_n}{\sqrt{\mathrm{Var} S_n}}\right), \qquad -\infty < x < \infty. \end{aligned}$$

Λέμε ότι η (10) είναι ένας τύπος *κανονικής προσέγγισης*. Σύμφωνα με αυτόν τον τύπο, προσεγγίζουμε την συνάρτηση κατανομής της S_n από την τυπική κανονική συνάρτηση κατανομής που έχει τον ίδιο μέσο και την ίδια διασπορά. Μια δυσκολία που αντιμετωπίζουμε κατά την εφαρμογή του τύπου κανονικής προσέγγισης είναι η επιλογή του n ώστε η (10) να ισχύει με τον επιθυμητό βαθμό ακρίβειας. Διάφορες αριθμητικές μελέτες έχουν δείξει ότι στις τυπικές πρακτικές εφαρμογές η επιλογή $n = 25$ είναι αρκετή για να πάρουμε χρήσιμα συμπεράσματα από την (10).

Σαν ένα παράδειγμα στο οποίο εφαρμόζεται η κανονική προσέγγιση, θεωρούμε ανεξάρτητες τυχαίες μεταβλητές $X_1, X_2, \ldots$, καθεμία από τις οποίες έχει εκθετική πυκνότητα με παράμετρο $\lambda = 1$. Τότε η (10) παίρνει την μορφή

$$(11) \qquad P(S_n \leq x) \approx \Phi\left(\frac{x - n}{\sqrt{n}}\right), \qquad -\infty < x < \infty.$$

Στο Σχήμα 1 δίνουμε γραφικές παραστάσεις (για $n = 10$) που δείχνουν την ακρίβεια της κανονικής προσέγγισης.

Παράδειγμα 11 Υποθέτουμε ότι ο χρόνος ζωής κάποιου τύπου λυχνίας, μετά την εγκατάστασή της, είναι εκθετικά κατανεμημένος με μέσο χρόνο ζωής τις 10 μέρες. Όταν καίγεται μία τέτοια λυχνία, την αντικαθιστούμε με μία πανομοιότυπη. Βρείτε την πιθανότητα να χρειαστούμε περισσότερες από 50 λυχνίες στην διάρκεια ενός έτους.

Για να λύσουμε αυτό το πρόβλημα, συμβολίζουμε με X_n τον χρόνο ζωής της n-στής λυχνίας. Υποθέτουμε ότι οι $X_1, X_2, \ldots$ είναι ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν εκθετική κατανομή με μέσο 10, δηλαδή παράμετρο $\lambda = 1/10$. Τότε, η $S_n = X_1 + \ldots + X_n$ δίνει τον χρόνο μέχρι την στιγμή που θα καεί η n-στή λυχνία. Θέλουμε να βρούμε την $P(S_{50} < 365)$. Η S_{50} έχει μέσο $50\lambda^{-1} = 500$ και διασπορά $50\lambda^{-2} = 5000$. Επομένως, από τον τύπο κανονικής προσέγγισης (10)

$$\begin{aligned} P(S_{50} < 365) &\approx \Phi\left(\frac{365 - 500}{\sqrt{5000}}\right) \\ &= \Phi(-1.91) = 0.028. \end{aligned}$$

Είναι λοιπόν πολύ απίθανο να χρειαστούμε περισσότερες από 50 λυχνίες.

Ας υποθέσουμε ότι η S_n είναι συνεχής τυχαία μεταβλητή με πυκνότητα f_{S_n}. Αν παραγωγίσουμε τα δύο μέλη της (10) παίρνουμε

$$f_{S_n}(x) \approx \frac{1}{\sigma\sqrt{n}}\phi\left(\frac{x - n\mu}{\sigma\sqrt{n}}\right), \quad -\infty < x < \infty. \tag{12}$$

Αν και ο τρόπος με τον οποίο πήραμε την (12) απέχει πολύ από το να είναι απόδειξη, η (12) δίνει πραγματικά καλή προσέγγιση για μεγάλα n (με την πρόσθετη ήπια συνθήκη ότι, για κάποιο n, η f_{S_n} είναι φραγμένη συνάρτηση).

Σαν παράδειγμα αυτής της προσέγγισης, παίρνουμε την X_1 να ακολουθεί εκθετική κατανομή με παράμετρο $\lambda = 1$, οπότε η (11) εφαρμόζεται. Τότε, η (12) μας δίνει

$$f_{S_n}(x) \approx \frac{1}{\sqrt{n}}\phi\left(\frac{x - n}{\sqrt{n}}\right), \quad -\infty < n < \infty. \tag{13}$$

Στο Σχήμα 2 δίνουμε γραφικές παραστάσεις (για $n = 10$) που δείχνουν την ακρίβεια αυτής της προσέγγισης.

Οι μορφές του Κεντρικού Οριακού Θεωρήματος που περιέχουν πυκνότητες αντί για συναρτήσεις κατανομής λέγονται «τοπικά» κεντρικά οριακά θεωρήματα. Είναι εξίσου σημαντικές, ειδικά στην προχωρημένη θεωρία πιθανοτήτων.

Υπάρχει μία προσέγγιση όμοια με την (12) που αφορά διακριτές τυχαίες μεταβλητές. Φυσικά, η ακριβής διατύπωση ενός τέτοιου τύπου προσέγγισης εξαρτάται από την φύση των πιθανών τιμών της S_n, δηλαδή εκείνων των τιμών του x για τις οποίες $f_{S_n}(x) = P(S_n = x) > 0$. Για απλότητα κάνουμε τις εξής δύο υποθέσεις:

(i) Αν x είναι μία πιθανή τιμή της X_1, τότε ο x είναι ακέραιος.

(ii) Αν a είναι μία πιθανή τιμή της X_1, τότε ο μέγιστος κοινός διαιρέτης του συνόλου $\{x - a | x \text{ πιθανή τιμή της } X_1\}$ είναι ίσος με ένα.

Αποκλείουμε, για παράδειγμα, μία τυχαία μεταβλητή X_1 για την οποία $P(X_1 = 1) = P(X_1 = 3) = 1/2$, γιατί τότε ο μέγιστος κοινός διαιρέτης του παραπάνω συνόλου είναι 2. Με τις υποθέσεις (i) και (ii), η προσέγγιση

$$f_{S_n}(x) \approx \frac{1}{\sigma\sqrt{n}} \phi\left(\frac{x - n\mu}{\sigma\sqrt{n}}\right), \tag{14}$$

για x ακέραιο, ισχύει για μεγάλα n.

Παράδειγμα 12 Έστω ότι η X_1 είναι διωνυμική τυχαία μεταβλητή που παίρνει τις τιμές 1 και 0 με πιθανότητες p και $1 - p$ αντίστοιχα. Τότε οι (i) και (ii) ισχύουν και η (14) εφαρμόζεται με $\mu = p$ και $\sigma^2 = p(1-p)$. Αφού η S_n ακολουθεί διωνυμική κατανομή με παραμέτρους n και p, παίρνουμε την προσέγγιση

$$\begin{aligned} f_{S_n}(x) &= \binom{n}{x} p^x (1-p)^{n-x} \\ &\approx \frac{1}{\sqrt{np(1-p)}} \phi\left(\frac{x - np}{\sqrt{np(1-p)}}\right), \end{aligned} \tag{15}$$

για x ακέραιο. Η προσέγγιση αυτή φαίνεται στο Σχήμα 3 (για $n = 10$ και $p = 0.3$).

Από το Σχήμα 3 οδηγούμαστε σε μία άλλη μέθοδο προσέγγισης της $f_{S_n}(x)$ στην διακριτή περίπτωση, δηλαδή, το ολοκλήρωμα του δεξιού μέλους της (14) στο $[x - 1/2, x + 1/2]$. Εκφράζοντας αυτό το ολοκλήρωμα συναρτήσει της Φ παίρνουμε σαν εναλλακτική προς την (14) την

$$
\begin{aligned}
\text{(16)} \qquad f_{S_n}(x) \quad &\approx \quad \Phi\left(\frac{x + (1/2) - n\mu}{\sigma\sqrt{n}}\right) \\
&- \quad \Phi\left(\frac{x - (1/2) - n\mu}{\sigma\sqrt{n}}\right),
\end{aligned}
$$

για x ακέραιο. Το εμβαδόν του γραμμοσκιασμένου χωρίου στο Σχήμα 3 είναι μία προσέγγιση της $P(S_n = 5)$.

Τέλος, αν αθροίσουμε την (16) στο σύνολο $\{\ldots, x - 2, x - 1, x\}$ οδηγούμαστε στην προσέγγιση

$$
\text{(17)} \qquad P(S_n \leq x) \approx \Phi\left(\frac{x + (1/2) - n\mu}{\sigma\sqrt{n}}\right),
$$

για x ακέραιο. Όταν η S_n είναι διακριτή και οι συνθήκες (i) και (ii) ικανοποιούνται, τότε η (17) συνήθως δίνει πιό ακριβή αποτελέσματα από την αρχική κανονική προσέγγιση του τύπου (10). Στο Σχήμα 4 συγκρίνουμε τις προσεγγίσεις που δίνουν οι τύποι (10) και (17) στην περίπτωση που η S_n ακολουθεί διωνυμική κατανομή με παραμέτρους $n = 10$ και $p = 0.3$.

Παράδειγμα 13 Ένας μπασκετμπωλίστας γνωρίζει ότι, κατά μέσο όρο, επιτυγχάνει στο 60% των ελεύθερων βολών που επιχειρεί. Ποιά είναι η πιθανότητα σε 25 προσπάθειες να ευστοχήσει σε περισσότερες από τις μισές;

Με βάση τις υποθέσεις του προβλήματος, δεχόμαστε ότι η S_n που μετράει το πλήθος των επιτυχιών σε n βολές ακολουθεί διωνυμική κατανομή με παραμέτρους n και $p = 0.6$. Αφού $P(S_n \geq x) = 1 - P(S_n \leq x - 1)$, η (17) μας δίνει την προσέγγιση

$$P(S_n \geq x) \approx 1 - \Phi\left(\frac{x - (1/2) - n\mu}{\sigma\sqrt{n}}\right), \tag{18}$$

για x ακέραιο. Στην περίπτωσή μας, $n\mu = 25(0.6) = 15$ και $\sigma\sqrt{n} = \sqrt{25(0.6)(0.4)} = 5\sqrt{0.24}$. Άρα,

$$\begin{aligned} P(S_{25} \geq 13 &\approx 1 - \Phi\left(\frac{13 - (1/2) - 15}{5\sqrt{0.24}}\right) \\ &= 1 - \Phi(-1.02) \\ &= \Phi(1.02) = 0.846. \end{aligned}$$

7.5.2 Εφαρμογές στη δειγματοληψία

Το Κεντρικό Οριακό Θεώρημα και οι αντίστοιχοι τύποι κανονικής προσέγγισης μπορούν να θεωρηθούν εκλεπτύνσεις του Ασθενούς Νόμου των Μεγάλων Αριθμών που συζητήσαμε στο Κεφάλαιο 4. Υπενθυμίζουμε ότι αυτός ο νόμος μάς λέει ότι για μεγάλα n, το S_n/n πρέπει να βρίσκεται κοντά στον μ με πιθανότητα κοντά στο 1. Όμως, ο ασθενής νόμος από μόνος του δεν μάς δίνει καμμία πληροφορία για την ακρίβεια μιάς τέτοιας εκτίμησης. Όπως είδαμε στο Κεφάλαιο 4, η ανισότητα του Chebyshev ρίχνει κάποιο φώς σ' αυτό το ερώτημα.

Ο τύπος (10) της κανονικής προσέγγισης είναι χρήσιμος και σ' αυτό το πλαίσιο. Για κάθε $c > 0$

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq c\right) = P(S_n \leq n\mu - nc) + P(S_n \geq n\mu + nc)$$

$$\begin{aligned} &\approx \quad \Phi\left(\frac{-nc}{\sigma\sqrt{n}}\right) + 1 - \Phi\left(\frac{nc}{\sigma\sqrt{n}}\right) \\ &= \quad 2\left[1 - \Phi\left(\frac{c\sqrt{n}}{\sigma}\right)\right]. \end{aligned}$$

Μ' άλλα λόγια

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq c\right) \approx 2(1 - \Phi(\delta)), \tag{19}$$

όπου

$$\delta = \frac{c\sqrt{n}}{\sigma}. \tag{20}$$

Παράδειγμα 14 Ένα δείγμα μεγέθους n πρόκειται να χρησιμοποιηθεί για να προσδιοριστεί το ποσοστό του πληθυσμού που σκοπεύει να ασκήσει το εκλογικό του δικαίωμα σε μία εκλογή. Θέτουμε $X_k = 1$ αν το k-στό άτομο του δείγματος σκοπεύει να ψηφίσει, και $X_k = 0$ αλλιώς. Υποθέτουμε ότι οι $X_1, \ldots, X_n$ είναι ανεξάρτητες, ισοκατανεμημένες τυχαίες μεταβλητές με $P(X_1 = 1) = p$ και $P(X_1 = 0) = 1 - p$. Τότε, $\mu = p$ και $\sigma^2 = p(1-p)$. Θα υποθέσουμε ακόμα ότι το p είναι αρκετά κοντά στο 0.5, οπότε η $\sigma = \sqrt{p(1-p)}$ προσεγγίζεται ικανοποιητικά από την $\sigma \approx 1/2$ (παρατηρήστε ότι η σ παίρνει μέγιστη τιμή $1/2$ όταν $p = 0.5$, και ότι αν το p κινείται μεταξύ 0.3 και 0.7, τότε η σ μένει πάνω από το 0.458 που είναι κοντά στο $1/2$). Η τυχαία μεταβλητή S_n/n συμβολίζει το ποσοστό των ατόμων στο δείγμα που πρόκειται να ψηφίσουν, και μπορούμε να την χρησιμοποιήσουμε για να εκτιμήσουμε την αληθή αλλά άγνωστη πιθανότητα p. Θα χρησιμοποιήσουμε κανονικές προσεγγίσεις για να λύσουμε τα εξής τρία προβλήματα:

(i) Έστω ότι $n = 900$. Βρείτε την πιθανότητα να ισχύει

$$\left|\frac{S_n}{n} - p\right| \geq 0.025.$$

(ii) Έστω ότι $n = 900$. Βρείτε c τέτοιο ώστε

$$P\left(\left|\frac{S_n}{n} - p\right| \geq c\right) = 0.01.$$

(iii) Βρείτε n τέτοιο ώστε

$$P\left(\left|\frac{S_n}{n} - p\right| \geq 0.025\right) = 0.01.$$

Λύση του (i). Από την (20)

$$\delta = \frac{(0.025)\sqrt{900}}{0.5} = 1.5,$$

οπότε από την (19)

$$\begin{aligned} P\left(\left|\frac{S_n}{n} - p\right| \geq 0.025\right) &\approx 2(1-\Phi(1.5)) \\ &= 2(0.067) = 0.134. \end{aligned}$$

Λύση του (ii). Επιλέγουμε πρώτα δ τέτοιο ώστε $2(1-\Phi(\delta)) = 0.01$, δηλαδή $\Phi(\delta) = 0.995$. Από τον Πίνακα I βλέπουμε ότι $\delta = 2.58$. Λύνοντας την (20) ως προς c παίρνουμε

$$c = \frac{\delta\sigma}{\sqrt{n}} = \frac{(2.58)(0.5)}{\sqrt{900}} = 0.043.$$

Λύση του (iii). Όπως στο (ii) έχουμε $\delta = 2.58$. Λύνοντας την (20) ως προς n βρίσκουμε

$$n = \frac{\delta^2\sigma^2}{c^2} = \frac{(2.58)^2(0.25)}{(0.025)^2} = 2663.$$

Αξίζει τον κόπο να συγκρίνουμε τα αποτελέσματα των (ii) και (iii). Και στις δύο περιπτώσεις,

$$P\left(\left|\frac{S_n}{n} - p\right| \geq c\right) \approx 0.01.$$

Στο (ii), $c = 0.043$ και $n = 900$, ενώ στο (iii), $c = 0.025$ και $n = 2663$. Πηγαίνοντας από το (ii) στο (iii), για να μικρύνουμε το c κατά $43/25$ είμαστε υποχρεωμένοι να αυξήσουμε το n κατά το τετράγωνο αυτού του παράγοντα. Αυτό ισχύει γενικά αν θέλουμε να κρατήσουμε την

$$P\left(\left|\frac{S_n}{n} - p\right| \geq c\right)$$

σταθερή. Γιατί τότε το δ προσδιορίζεται από την (19) και, από την (20), το n συνδέεται με το c μέσω της $n = \delta^2\sigma^2/c^2$. Για τον ίδιο λόγο, αν αυξήσουμε το n κατά έναν παράγοντα, το c θα μειωθεί κατά την τετραγωνική ρίζα αυτού του παράγοντα.

7.6 Ασκήσεις

1 Έστω ότι η X έχει πυκνότητα βήτα με παραμέτρους α_1 και α_2. Βρείτε την $\mathbb{E}X$.

2 Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές με πυκνότητες γάμμα, τις $\Gamma(\alpha_1, \lambda)$ και $\Gamma(\alpha_2, \lambda)$ αντίστοιχα. Θέτουμε $Z = Y/X$. Για ποιές τιμές των α_1 και α_2 η Z έχει πεπερασμένη μέση τιμή; Βρείτε την $\mathbb{E}Z$ όταν υπάρχει. *Υπόδειξη:* Δείτε το Θεώρημα 3 του Κεφαλαίου 6 και την σχετική συζήτηση.

3 Υποθέτουμε ότι η X έχει την κανονική πυκνότητα $n(0, \sigma^2)$. Βρείτε την $\mathbb{E}|X|$. *Υπόδειξη:* Χρησιμοποιήστε το αποτέλεσμα της Άσκησης 31 του Κεφαλαίου 5.

4 Έστω ότι η X έχει εκθετική πυκνότητα με παράμετρο λ, και ότι η X_ε ορίζεται μέσω των X και $\varepsilon > 0$ από την (1). Ποιά είναι η κατανομή της $X_\varepsilon/\varepsilon$; Βρείτε την $\mathbb{E}X_\varepsilon$ και υπολογίστε το όριό της καθώς $\varepsilon \to 0$.

5 Έστω ότι η X έχει πυκνότητα βήτα με παραμέτρους α_1 και α_2. Βρείτε τις ροπές και την διασπορά της X.

6 Έστω ότι η X ακολουθεί κατανομή χ^2 με n βαθμούς ελευθερίας. Βρείτε τον μέσο της $Y = \sqrt{X}$.

7 Έστω X η τυχαία μεταβλητή του Παραδείγματος 7. Βρείτε την $\mathbb{E}X^m$ χρησιμοποιώντας την περιθώρια πυκνότητα f_X.

8 Έστω Z η τυχαία μεταβλητή της Άσκησης 2. Βρείτε την διασπορά της Z.

9 Έστω U_1 και U_2 ανεξάρτητες τυχαίες μεταβλητές που έχουν εκθετική πυκνότητα με παράμετρο λ. Θέτουμε $Y = \max(U_1, U_2)$. Βρείτε τον μέσο και την διασπορά της Y (συμβουλευτείτε την Παράγραφο 6.5).

10 Έστω X η τυχαία μεταβλητή του Παραδείγματος 1 του Κεφαλαίου 5. Βρείτε τον μέσο και την διασπορά της X.

11 Έστω X η τυχαία μεταβλητή του Παραδείγματος 1 του Κεφαλαίου 6. Βρείτε τον μέσο και την διασπορά της X. *Υπόδειξη:* Η $\mathbb{E}X^2$ ανάγεται σε ολοκλήρωμα Βήτα.

12 Βρείτε τον μέσο και την διασπορά της τυχαίας μεταβλητής Z από την Άσκηση 17 του Κεφαλαίου 6.

13 Βρείτε τον μέσο και την διασπορά της τυχαίας μεταβλητής Y από την Άσκηση 28 του Κεφαλαίου 6.

14 Έστω X το ημίτονο γωνίας (σε ακτίνια) που επιλέγεται ομοιόμορφα από το $(-\pi/2, \pi/2)$. Βρείτε τον μέσο και την διασπορά της X.

15 Έστω ότι η X έχει την κανονική πυκνότητα $n(0, \sigma^2)$. Βρείτε τον μέσο και την διασπορά καθεμίας από τις τυχαίες μεταβλητές:
(α) $|X|$,
(β) X^2,
(γ) e^{tX}.

16 Η X έχει την πυκνότητα γάμμα $\Gamma(\alpha, \lambda)$. Για ποιούς πραγματικούς t η e^{tX} έχει πεπερασμένη μέση τιμή; Βρείτε την $\mathbb{E}e^{tX}$ γι' αυτές τις τιμές του t.

17 Η X έχει την πυκνότητα γάμμα $\Gamma(\alpha, \lambda)$. Για ποιούς πραγματικούς r η X^r έχει πεπερασμένη μέση τιμή; Βρείτε την $\mathbb{E}X^r$ γι' αυτές τις τιμές του r.

18 Έστω X μία μη αρνητική συνεχής τυχαία μεταβλητή με πυκνότητα f και συνάρτηση κατανομής F. Δείξτε ότι η X έχει πεπερασμένη μέση τιμή αν και μόνο αν

$$\int_0^\infty (1 - F(x))dx < \infty$$

και τότε

$$\mathbb{E}X = \int_0^{\infty} (1 - F(x))dx.$$

Υπόδειξη: Δείτε την απόδειξη του Θεωρήματος 5 του Κεφαλαίου 4.

19 Έστω X_k η k-στή διατεταγμένη στατιστική συνάρτηση ενός δείγματος από τις τυχαίες μεταβλητές $U_1, \ldots, U_n$ που είναι ανεξάρτητες και ομοιόμορφα κατανεμημένες στο $(0,1)$. Βρείτε τον μέσο και την διασπορά της X_k.

20 Έστω X και Y όπως στο Παράδειγμα 7, και έστω $R = Y - X$. Βρείτε τον μέσο και την διασπορά της R. *Υπόδειξη:* Χρησιμοποιήστε την ισότητα (16) του Κεφαλαίου 4.

21 Δίνονται X και Y με πυκνότητα f όπως στην Άσκηση 9 του Κεφαλαίου 6. Βρείτε τον συντελεστή συσχέτισης των X και Y.

22 Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές, όπου η X έχει την κανονική πυκνότητα $n(\mu, \sigma^2)$ και η Y έχει την πυκνότητα γάμμα $\Gamma(\alpha, \lambda)$. Βρείτε τον μέσο και την διασπορά της τυχαίας μεταβλητής $Z = XY$.

23 Έστω X και Y τυχαίες μεταβλητές με μέσο 0, διασπορά 1, και συντελεστή συσχέτισης ρ. Δείξτε ότι οι $X - \rho Y$ και Y είναι ασυσχέτιστες, και ότι η $X - \rho Y$ έχει μέσο 0 και διασπορά $1 - \rho^2$.

24 Έστω X, Y, και Z τυχαίες μεταβλητές με μέσο 0 και μοναδιαία διασπορά. Έστω ρ_1 ο συντελεστής συσχέτισης των X και Y, ρ_2 ο συντελεστής συσχέτισης των Y και Z, και ρ_3 ο συντελεστής συσχέτισης των X και Z. Δείξτε ότι

$$\rho_3 \geq \rho_1\rho_2 - \sqrt{1 - \rho_1^2}\sqrt{1 - \rho_2^2}.$$

Υπόδειξη: Γράψτε

$$XZ = [\rho_1 Y + (X - \rho_1 Y)][\rho_2 Y + (Z - \rho_2 Y)],$$

και χρησιμοποιήστε το προηγούμενο πρόβλημα και την ανισότητα του Schwarz.

25 Έστω X, Y, και Z όπως στο προηγούμενο πρόβλημα. Υποθέτουμε ότι $\rho_1 \geq 0.9$ και $\rho_2 \geq 0.8$. Τι μπορούμε να πούμε για τον ρ_3;

26 Έστω X και Y με από κοινού πυκνότητα f που είναι ομοιόμορφη στο εσωτερικό του τριγώνου με κορυφές τα $(0,0)$, $(2,0)$, και $(1,2)$. Βρείτε την δεσμευμένη μέση τιμή της Y με δεδομένη την X.

27 Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές με πυκνότητες τις $\Gamma(\alpha_1, \lambda)$ και $\Gamma(\alpha_2, \lambda)$ αντίστοιχα. Θέτουμε $Z = X + Y$. Βρείτε την δεσμευμένη μέση τιμή της X με δεδομένη την Z.

28 Έστω Π και Y τυχαίες μεταβλητές όπως στην Άσκηση 26 του Κεφαλαίου 6. Βρείτε την δεσμευμένη μέση τιμή της Π με δεδομένη την Y.

29 Έστω X και Y συνεχείς τυχαίες μεταβλητές που έχουν από κοινού πυκνότητα. Υποθέτουμε ότι οι Y και $\phi(X)Y$ έχουν πεπερασμένη μέση τιμή. Δείξτε ότι

$$\mathbb{E}\phi(X)Y = \int_{-\infty}^{\infty} \phi(x)\mathbb{E}[Y|X = x]f_X(x)dx.$$

30 Έστω X και Y συνεχείς τυχαίες μεταβλητές που έχουν από κοινού πυκνότητα, και έστω $\mathrm{Var}[Y|X = x]$ η διασπορά της δεσμευμένης πυκνότητας της Y δεδομένου ότι $X = x$. Δείξτε ότι αν $\mathbb{E}[Y|X = x] = \mu$ ανεξάρτητα από την X, τότε $\mathbb{E}Y = \mu$ και

$$\mathrm{Var}Y = \int_{-\infty}^{\infty} \mathrm{Var}[Y|X = x]f_X(x)dx.$$

31 Έστω $X_1, X_2, \ldots$ ανεξάρτητες, ισοκατανεμημένες τυχαίες μεταβλητές με μέσο 0 και πεπερασμένη μη μηδενική διασπορά σ^2. Θέτουμε $S_n = X_1 + \ldots + X_n$. Δείξτε ότι αν η X_1 έχει πεπερασμένη τρίτη ροπή, τότε $\mathbb{E}S_n^3 = n\mathbb{E}X_1^3$ και

$$\lim_{n\to\infty} \mathbb{E}\left(\frac{S_n}{\sigma\sqrt{n}}\right)^3 = 0,$$

που είναι η τρίτη ροπή της τυπικής κανονικής κατανομής.
Υπόδειξη: Ο όρος $3n(n-1)$ προκύπτει από την παράσταση

$$\binom{n}{2}\frac{4!}{2!2!}.$$

33 Έστω ότι η X έχει την πυκνότητα γάμμα $\Gamma(\alpha, \lambda)$. Βρείτε την κανονική προσέγγιση της $P(X \leq x)$.

34 Έστω $X_1, X_2, \ldots$ ανεξάρτητες, κανονικά κατανεμημένες τυχαίες μεταβλητές με μέσο 0 και διασπορά σ^2.

(α) Ποιός είναι ο μέσος και ποιά η διασπορά της τυχαίας μεταβλητής X_1^2;

(β) Πώς πρέπει να προσεγγίσουμε την $P(X_1^2 + \ldots + X_n^2 \leq x)$ με την βοήθεια της Φ;

35 Έστω $X_1, X_2, \ldots$ ανεξάρτητες, κανονικά κατανεμημένες τυχαίες μεταβλητές με μέσο 0 και διασπορά σ^2 (δείτε την προηγούμενη άσκηση).

(α) Βρείτε την $P(X_1^2 + \ldots + X_{100}^2 \leq 120)$.

(β) Βρείτε την $P(80 \leq X_1^2 + \ldots + X_{100}^2 \leq 120)$.

(α) Βρείτε c τέτοιο ώστε $P(X_1^2 + \ldots + X_{100}^2 \leq 100 + c) = 0.95$.

(α) Βρείτε c τέτοιο ώστε $P(100 - c \leq X_1^2 + \ldots + X_{100}^2 \leq 100 + c) = 0.95$.

36 Ένας δρομέας προσπαθεί να τρέξει 100 μέτρα σε 100 βήματα, σε μία άτυπη κούρσα. Τα βήματά του είναι ανεξάρτητα κατανεμημένα με μέσο $\mu = 0.97$ μέτρα και τυπική απόκλιση $\sigma = 0.1$ μέτρα. Βρείτε την πιθανότητα τα 100 βήματά του να διαφέρουν από 100 μέτρα το πολύ 5 μέτρα.

37 Στρογγυλοποιούμε είκοσι αριθμούς στον πλησιέστερο ακέραιο και τους προσθέτουμε. Υποθέτουμε ότι τα σφάλματα στρογγύλευσης είναι ανεξάρτητα και ομοιόμορφα κατανεμημένα στο $(-1/2, 1/2)$. Βρείτε την πιθανότητα το άθροισμά μας να διαφέρει από το άθροισμα των αρχικών είκοσι αριθμών περισσότερο από 3.

38 Ρίχνουμε ένα αμερόληπτο νόμισμα μέχρι να εμφανιστεί 100 φορές κορώνα. Βρείτε την πιθανότητα να χρειαστούν τουλάχιστον 226 ρίψεις.

39 Στο προηγούμενο πρόβλημα βρείτε την πιθανότητα να χρειαστούν ακριβώς 226 ρίψεις.

40 Υποθέτουμε ότι η X ακολουθεί κατανομή Poisson με παράμετρο λ.

(α) Πώς θα προσεγγίζατε την $f_X(x)$ με βάση την τυπική κανονική πυκνότητα ϕ;

(β) Πώς θα προσεγγίζατε την $f_X(x)$ με βάση την τυπική κανονική συνάρτηση κατανομής Φ;

41 Έστω ότι η S_n ακολουθεί διωνυμική κατανομή με παραμέτρους n και $p = 1/2$. Πώς συμπεριφέρεται η $P(S_{2n} = n)$ όταν το n είναι μεγάλο; *Υπόδειξη:* Χρησιμοποιήστε την προσέγγιση (15).

42 Οι παίκτες Α και Β κάνουν μία σειρά στοιχημάτων της 1 δραχμής, για καθένα από τα οποία η πιθανότητα επιτυχίας κάθε παίκτη είναι 1/2. Έστω S_n το καθαρό ποσό που έχει κερδίσει ο παίκτης Α μετά από n στοιχήματα. Πώς συμπεριφέρεται η $P(S_{2n} = 0)$ όταν το n είναι μεγάλο; *Υπόδειξη:* Δείτε το προηγούμενο πρόβλημα. Γιατί δεν μπορούμε να εφαρμόσουμε απευθείας την προσέγγιση (15) σ' αυτήν την περίπτωση;

43 Οι υποψήφιοι Α και Β κατεβαίνουν για την δημαρχία, και το 55% των εκλεκτόρων είναι υπέρ του υποψηφίου Β. Ποιά είναι η πιθανότητα σε ένα δείγμα μεγέθους 100 οι μισοί τουλάχιστον από το δείγμα να είναι υπέρ του υποψηφίου Α;

44 Μία εταιρεία δημοσκοπήσεων παίρνει δείγμα 1200 ψηφοφόρων για να εκτιμήσει το ποσοστό εκείνων που σκοπεύουν να ψηφίσουν τον υποψήφιο Α σε κάποια εκλογή. Πόσο μεγάλο πρέπει να είναι το πραγματικό ποσοστό p του υποψηφίου Α ώστε να είναι 95% βέβαιο ότι η πλειοψηφία των ατόμων στο δείγμα θα τον ψηφίσει;

45 Ας υποθέσουμε ότι ο υποψήφιος Α του προηγούμενου προβλήματος επιμένει ότι το μέγεθος του δείγματος πρέπει να αυξηθεί σε έναν αριθμό n ώστε αν το 51% των ψηφοφόρων τον υποστηρίζουν να είναι 95% βέβαιος ότι θα αποσπάσει πλειοψηφία στο δείγμα. Πόσο μεγάλο πρέπει να είναι το n;

46 Λύστε την Άσκηση 27 του Κεφαλαίου 4 χρησιμοποιώντας κανονική προσέγγιση.

Κεφάλαιο 8

Ροπογεννήτριες και χαρακτηριστικές συναρτήσεις

8.1 Ροπογεννήτριες

Η *ροπογεννήτρια* $M_X(t)$ μιάς τυχαίας μεταβλητής X ορίζεται από την

$$M_X(t) = \mathbb{E}e^{tX}.$$

Το πεδίο ορισμού της M_X αποτελείται από όλους τους πραγματικούς αριθμούς t για τους οποίους η e^{tX} έχει πεπερασμένη μέση τιμή.

Παράδειγμα 1 Έστω X κανονικά κατανεμημένη με μέσο μ και διασπορά σ^2. Τότε,

$$\begin{aligned} M_X(t) = \mathbb{E}e^{tX} &= \int_{-\infty}^{\infty} e^{tx}\frac{1}{\sigma\sqrt{2\pi}}e^{-[(x-\mu)^2/2\sigma^2]}dx \\ &= \int_{-\infty}^{\infty} e^{t(y+\mu)}\frac{1}{\sigma\sqrt{2\pi}}e^{-y^2/2\sigma^2}dy \\ &= e^{\mu t}\int_{-\infty}^{\infty}\frac{1}{\sigma\sqrt{2\pi}}e^{ty-(y^2/2\sigma^2)}dy. \end{aligned}$$

Όμως

$$ty - \frac{y^2}{2\sigma^2} = -\frac{(y-\sigma^2 t)^2}{2\sigma^2} + \frac{\sigma^2 t^2}{2}.$$

Επομένως,

$$M_X(t) = e^{\mu t}e^{\sigma^2 t^2/2}\int_{-\infty}^{\infty}\frac{1}{\sigma\sqrt{2\pi}}e^{-[(y-\sigma^2 t)^2/2\sigma^2]}dt.$$

Αφού το τελευταίο ολοκλήρωμα ισούται με το ολοκλήρωμα της κανονικής πυκνότητας $n(\sigma^2 t, \sigma^2)$, η τιμή του είναι ένα. Επομένως,

$$(1) \qquad M_X(t) = e^{\mu t} e^{\sigma^2 t^2/2}, \qquad -\infty < t < \infty.$$

Παράδειγμα 2 Έστω ότι η X έχει την πυκνότητα γάμμα με παραμέτρους α και λ. Τότε,

$$\begin{aligned} M_X(t) &= \int_0^\infty e^{tx} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} dx \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha-1} e^{-(\lambda-t)x} dx \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha)}{(\lambda-t)^\alpha} \end{aligned}$$

αν $-\infty < t < \lambda$. Το ολοκλήρωμα αποκλίνει αν $\lambda \le t < \infty$. Άρα,

$$(2) \qquad M_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^\alpha, \qquad -\infty < t < \lambda.$$

Ας υποθέσουμε τώρα ότι η X είναι διακριτή τυχαία μεταβλητή, και ότι όλες οι πιθανές τιμές της είναι μη αρνητικοί ακέραιοι. Τότε

$$M_X(t) = \sum_{n=0}^\infty e^{nt} P(X = n).$$

Στο Κεφάλαιο 3 ορίσαμε την γεννήτρια συνάρτηση πιθανότητας για τυχαίες μεταβλητές αυτού του είδους, θέτοντας

$$\Phi_X(t) = \sum_{n=0}^\infty t^n P(X = n).$$

Από αυτούς τους δύο τύπους είναι φανερό ότι

$$(3) \qquad M_X(t) = \Phi_X(e^t).$$

Ο τύπος (3) μας επιτρέπει να προσδιορίσουμε την ροπογεννήτρια απευθείας από την γεννήτρια συνάρτηση πιθανότητας. Για παράδειγμα, αν η X ακολουθεί διωνυμική κατανομή με παραμέτρους n και p, τότε όπως είδαμε στο Παράδειγμα 16 του Κεφαλαίου 3,

$$\Phi_X(t) = (pt + 1 - p)^n.$$

Έπεται άμεσα ότι

$$M_X(t) = (pe^t + 1 - p)^n.$$

Όμοια, αν η X ακολουθεί κατανομή Poisson με παράμετρο λ, τότε σύμφωνα με το Παράδειγμα 18 του Κεφαλαίου 3,

$$\Phi_X(t) = e^{\lambda(t-1)}.$$

Επομένως,

$$M_X(t) = e^{\lambda(e^t-1)}.$$

Θα μπορούσαμε φυσικά, στα παραπάνω δύο παραδείγματα, να υπολογίσουμε την $M_X(t)$ απευθείας από τον ορισμό της ροπογεννήτριας.

Αν X και Y είναι ανεξάρτητες τυχαίες μεταβλητές, τότε οι e^{tX} και e^{tY} είναι επίσης ανεξάρτητες. Επομένως,

$$\begin{aligned} M_{X+Y}(t) = \mathbb{E}e^{t(X+Y)} &= \mathbb{E}e^{tX}e^{tY} = \mathbb{E}e^{tX}\mathbb{E}e^{tY} \\ &= M_X(t)M_Y(t). \end{aligned}$$

Βλέπουμε λοιπόν εύκολα ότι αν οι $X_1, \ldots, X_n$ είναι ανεξάρτητες και ισόνομες, τότε

$$M_{X_1+\ldots+X_n}(t) = (M_{X_1}(t))^n. \tag{4}$$

Για να δούμε γιατί η $M_X(t)$ λέγεται ροπογεννήτρια, γράφουμε

$$M_X(t) = \mathbb{E}e^{tX} = \mathbb{E}\sum_{n=0}^{\infty}\frac{t^nX^n}{n!}.$$

Υποθέτουμε ότι η $M_X(t)$ ορίζεται καλά αν $-t_0 < t < t_0$, όπου t_0 θετικός πραγματικός αριθμός. Σ' αυτήν την περίπτωση μπορούμε να δείξουμε ότι στην τελευταία παράσταση για την $M_X(t)$ επιτρέπεται να εναλλάξουμε την μέση τιμή με την άθροιση. Δηλαδή,

$$M_X(t) = \sum_{n=0}^{\infty}\frac{\mathbb{E}X^n}{n!}t^n \tag{5}$$

αν $-t_0 < t < t_0$. Ειδικότερα, αν η $M_X(t)$ ορίζεται για κάθε t, τότε η (5) ισχύει για κάθε t. Η σειρά Taylor της $M_X(t)$ είναι η

$$M_X(t) = \sum_{n=0}^{\infty}\frac{t^n}{n!}\frac{d^n}{dt^n}M_X(t)\bigg|_{t=0}. \tag{6}$$

Συγκρίνοντας τους συντελεστές του t^n στις (5) και (6), βλέπουμε ότι

$$\mathbb{E}X^n = \frac{d^n}{dt^n}M_X(t)\bigg|_{t=0}. \tag{7}$$

Παράδειγμα 3 Έστω ότι η X ακολουθεί κανονική κατανομή με μέσο 0 και διασπορά σ^2. Βρείτε τις ροπές της X χρησιμοποιώντας ροπογεννήτριες.

Παρατηρούμε πρώτα από την (1) ότι

$$M_X(t) = e^{\sigma^2 t^2/2} = \sum_{n=0}^{\infty} \left(\frac{\sigma^2 t^2}{2}\right)^n \frac{1}{n!} = \sum_{n=0}^{\infty} \frac{\sigma^{2n}}{2^n n!} t^{2n}.$$

Άρα οι περιττές ροπές της X είναι όλες μηδέν, και οι άρτιες ροπές δίνονται από την

$$\frac{\mathbb{E}X^{2n}}{(2n)!} = \frac{\sigma^{2n}}{2^n n!},$$

ή, ισοδύναμα,

$$\mathbb{E}X^{2n} = \frac{\sigma^{2n}(2n)!}{2^n n!}.$$

Αυτό συμφωνεί με το αποτέλεσμα που βρήκαμε στο Κεφάλαιο 7.

Το ίδιο παράδειγμα θα μας βοηθήσει να δούμε τον τρόπο με τον οποίο εφαρμόζεται η (7). Αφού

$$\frac{d}{dt} e^{\sigma^2 t^2/2} = \sigma^2 t e^{\sigma^2 t^2/2}$$

και

$$\frac{d^2}{dt^2} e^{\sigma^2 t^2/2} = (\sigma^2 + \sigma^4 t^2) e^{\sigma^2 t^2/2},$$

έπεται ότι

$$\left.\frac{d}{dt} e^{\sigma^2 t^2/2}\right|_{t=0} = 0$$

και

$$\left.\frac{d^2}{dt^2} e^{\sigma^2 t^2/2}\right|_{t=0} = \sigma^2,$$

παίρνουμε δηλαδή τις δύο πρώτες ροπές της X.

8.2 Χαρακτηριστικές συναρτήσεις

Η *χαρακτηριστική συνάρτηση* μιάς τυχαίας μεταβλητής X ορίζεται από την

$$\varphi_X(t) = \mathbb{E}e^{itX}, \qquad -\infty < t < \infty,$$

όπου $i = \sqrt{-1}$. Οι χαρακτηριστικές συναρτήσεις είναι λίγο πιό πολύπλοκες από τις ροπογεννήτριες, αφού στον ορισμό τους εμπλέκονται μιγαδικοί αριθμοί. Έχουν όμως δύο σημαντικά πλεονεκτήματα σε σύγκριση με τις ροπογεννήτριες. Πρώτον, η $\varphi_X(t)$ είναι πεπερασμένη για όλες τις τυχαίες μεταβλητές X και όλους τους πραγματικούς αριθμούς t. Δεύτερον, η συνάρτηση κατανομής της X, αλλά και η συνάρτηση πυκνότητας της X αν υπάρχει, προκύπτουν από την χαρακτηριστική συνάρτηση μέσω ενός «τύπου αντιστροφής». Χρησιμοποιώντας ιδιότητες των χαρακτηριστικών συναρτήσεων θα μπορέσουμε να αποδείξουμε τον Ασθενή Νόμο των Μεγάλων Αριθμών και το Κεντρικό Οριακό Θεώρημα, κάτι που δεν είμαστε σε θέση να πετύχουμε με τις ροπογεννήτριες.

Πριν μελετήσουμε τις χαρακτηριστικές συναρτήσεις θα ανακεφαλαιώσουμε εν συντομία τα απαιτούμενα στοιχεία από τη θεωρία των μιγαδικών συναρτήσεων.

Μπορούμε να γράφουμε κάθε μιγαδικό αριθμό z στη μορφή $z = x + iy$, όπου x και y είναι πραγματικοί αριθμοί. Το μέτρο $|z|$ ενός μιγαδικού αριθμού ορίζεται από την $|z| = (x^2 + y^2)^{1/2}$. Η απόσταση μεταξύ δύο μιγαδικών αριθμών z_1 και z_2 δίνεται από το $|z_1 - z_2|$.

Αν μία συνάρτηση μιάς πραγματικής μεταβλητής έχει ανάπτυγμα σε δυναμοσειρά με θετική ακτίνα σύγκλισης, μπορούμε να χρησιμοποιήσουμε αυτή τη δυναμοσειρά για να ορίσουμε μία συνάρτηση μιάς μιγαδικής μεταβλητής. Για παράδειγμα, ορίζουμε

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

για κάθε μιγαδικό αριθμό z. Η σχέση

$$e^{z_1+z_2} = e^{z_1} e^{z_2}$$

εξακολουθεί να ισχύει για όλα τα ζεύγη μιγαδικών αριθμών z_1 και z_2. Θέτοντας $z = it$, όπου t είναι ένας πραγματικός αριθμός, βλέπουμε ότι

$$\begin{aligned} e^{it} &= \sum_{n=0}^{\infty} \frac{(it)^n}{n!} \\ &= \left(1 + it - \frac{t^2}{2} - \frac{it^3}{3!} + \frac{t^4}{4!} + \frac{it^5}{5!} - \dots\right) \\ &= \left(1 - \frac{t^2}{2!} + \frac{t^4}{4!} - \dots\right) + i\left(t - \frac{t^3}{3!} + \frac{t^5}{5!} - \dots\right). \end{aligned}$$

Παρατηρούμε ότι οι δύο δυναμοσειρές στην τελευταία παράσταση είναι τα αναπτύγματα των $\cos t$ και $\sin t$, επομένως

$$e^{it} = \cos t + i \sin t. \tag{8}$$

Χρησιμοποιώντας το γεγονός ότι $\cos(-t) = \cos t$ και $\sin(-t) = -\sin t$, βλέπουμε ότι

$$e^{-it} = \cos t - i \sin t.$$

Από αυτούς τους τύπους, λύνοντας ως προς $\cos t$ και $\sin t$, παίρνουμε

$$\cos t = \frac{e^{it} + e^{-it}}{2} \quad , \quad \sin t = \frac{e^{it} - e^{-it}}{2i}. \tag{9}$$

Από την (8) βλέπουμε επίσης ότι

$$|e^{it}| = \left(\cos^2 t + \sin^2 t\right)^{1/2} = 1.$$

Αν $f(t)$ και $g(t)$ είναι πραγματικές συναρτήσεις του t, τότε η $h(t) = f(t) + ig(t)$ ορίζει μία μιγαδική συνάρτηση του t. Μπορούμε να παραγωγίσουμε την $h(t)$ παραγωγίζοντας τις $f(t)$ και $g(t)$ χωριστά. Δηλαδή,

$$h'(t) = f'(t) + ig'(t),$$

με την προϋπόθεση ότι οι $f'(t)$ και $g'(t)$ υπάρχουν. Ομοίως, ορίζουμε

$$\int_a^b h(t)dt = \int_a^b f(t)dt + i\int_a^b g(t)dt,$$

με την προϋπόθεση ότι τα ολοκληρώματα των f και g υπάρχουν. Ο τύπος

$$\frac{d}{dt}e^{ct} = ce^{ct}$$

ισχύει για κάθε μιγαδική σταθερά c. Το θεμελιώδες θεώρημα του απειροστικού λογισμού εξακολουθεί να ισχύει. Ειδικότερα, αν c είναι μία μη μηδενική μιγαδική σταθερά, τότε

$$\int_a^b e^{ct}dt = \frac{e^{cb} - e^{ca}}{c}.$$

Μία μιγαδική τυχαία μεταβλητή Z είναι μία συνάρτηση της μορφής $Z = X + iY$, όπου X και Y πραγματικές τυχαίες μεταβλητές. Η μέση τιμή $\mathbb{E}Z$ της Z ορίζεται από την

$$\mathbb{E}Z = \mathbb{E}(X + iY) = \mathbb{E}X + i\mathbb{E}Y$$

αν οι $\mathbb{E}X$ και $\mathbb{E}Y$ είναι καλά ορισμένες. Ακριβώς όπως και στην περίπτωση των πραγματικών τυχαίων μεταβλητών, η Z έχει πεπερασμένη μέση τιμή αν και μόνο αν $\mathbb{E}|Z| < \infty$, εξακολουθεί δε να ισχύει η

$$|\mathbb{E}Z| \le \mathbb{E}|Z|.$$

Η σχέση

$$\mathbb{E}(a_1 Z_1 + a_2 Z_2) = a_1\mathbb{E}Z_1 + a_2\mathbb{E}Z_2$$

ισχύει για οποιεσδήποτε μιγαδικές σταθερές a_1 και a_2 και οποιεσδήποτε μιγαδικές τυχαίες μεταβλητές Z_1 και Z_2 που έχουν πεπερασμένη μέση τιμή.

Συμφωνούμε όταν γράφουμε X και Y, με ή χωρίς δείκτες, να εννοούμε πραγματικές τυχαίες μεταβλητές. Για παράδειγμα, από την φράση «έστω X μία τυχαία μεταβλητή» καταλαβαίνουμε ότι η X παίρνει πραγματικές τιμές.

Υποθέτουμε τώρα ότι X είναι μία τυχαία μεταβλητή και t είναι μία σταθερά (το σύμβολο t χρησιμοποιείται αποκλειστικά για πραγματικές σταθερές). Τότε $|e^{itX}| =$

1, άρα η e^{itX} έχει πεπερασμένη μέση τιμή και η χαρακτηριστική συνάρτηση $\varphi_X(t)$, $-\infty < t < \infty$, που ορίζεται από την

$$\varphi_X(t) = \mathbb{E}e^{itX}, \qquad -\infty < t < \infty,$$

είναι καλά ορισμένη. Βλέπουμε ότι $\varphi_X(0) = \mathbb{E}e^0 = \mathbb{E}1 = 1$ και, για κάθε $-\infty < t < \infty$,

$$|\varphi_X(t)| = |\mathbb{E}e^{itX}| \le \mathbb{E}|e^{itX}| = \mathbb{E}1 = 1.$$

Ο λόγος για τον οποίο οι χαρακτηριστικές συναρτήσεις παίρνουν πεπερασμένη τιμή για κάθε t ενώ οι ροπογεννήτριες δεν είναι εν γένει πεπερασμένες, είναι ότι η $e^{it}, -\infty < t < \infty$ είναι φραγμένη ενώ η $e^t, -\infty < t < \infty$, δεν είναι φραγμένη.

Παράδειγμα 4. Έστω X μία τυχαία μεταβλητή που παίρνει την τιμή a με πιθανότητα ένα. Τότε,

$$\varphi_X(t) = \mathbb{E}e^{itX} = e^{ita}, \qquad -\infty < t < \infty.$$

Ειδικότερα, αν η X παίρνει την τιμή μηδέν με πιθανότητα ένα, τότε η χαρακτηριστική της συνάρτηση είναι ταυτοτικά ίση με 1.

Αν X είναι μία τυχαία μεταβλητή και a και b είναι πραγματικές σταθερές, τότε

$$\begin{aligned} \varphi_{a+bX}(t) &= \mathbb{E}e^{it(a+bX)} \\ &= \mathbb{E}e^{ita}e^{itbX} \\ &= e^{ita}\mathbb{E}e^{itbX}, \end{aligned}$$

επομένως

$$\varphi_{a+bX}(t) = e^{ita}\varphi_X(bt), \qquad -\infty < t < \infty. \tag{10}$$

Παράδειγμα 5. Υποθέτουμε ότι η U είναι ομοιόμορφα κατανεμημένη στο $(-1, 1)$. Τότε, αν $t \neq 0$,

$$\begin{aligned} \varphi_U(t) &= \int_{-1}^{1} e^{itu}\frac{1}{2}du \\ &= \frac{1}{2}\frac{e^{itu}}{it}\bigg|_{-1}^{1} \\ &= \frac{1}{2}\left(\frac{e^{it} - e^{-it}}{it}\right) \\ &= \frac{\sin t}{t}. \end{aligned}$$

Αν $a < b$, θέτουμε

$$X = \frac{a+b}{2} + \left(\frac{b-a}{2}\right)U.$$

Τότε η X είναι ομοιόμορφα κατανεμημένη στο (a,b) και από την (10), αν $t \neq 0$,

$$\varphi_X(t) = e^{it(a+b)/2}\frac{\sin((b-a)t/2)}{(b-a)t/2}.$$

Εναλλακτικά, μπορούμε να γράψουμε

$$\begin{aligned}\varphi_X(t) &= \int_a^b e^{itx}\frac{1}{b-a}dx \\ &= \frac{1}{b-a}\left.\frac{e^{itx}}{it}\right|_a^b \\ &= \frac{e^{ibt}-e^{iat}}{it(b-a)}.\end{aligned}$$

Χρησιμοποιώντας την (9) μπορείτε εύκολα να ελέγξετε ότι οι δύο απαντήσεις συμφωνούν.

Παράδειγμα 6. Υποθέτουμε ότι η X ακολουθεί εκθετική κατανομή με παράμετρο λ. Τότε,

$$\begin{aligned}\varphi_X(t) &= \int_0^\infty e^{itx}\lambda e^{-\lambda x}dx \\ &= \lambda\int_0^\infty e^{-(\lambda-it)x}dx \\ &= \frac{\lambda}{\lambda-it}e^{-(\lambda-it)x}\Big|_\infty^0.\end{aligned}$$

Όμως $\lim_{x\to\infty} e^{-(\lambda-it)x} = 0$ και η e^{itx} είναι φραγμένη ως προς x, επομένως

$$\lim_{x\to\infty} e^{-(\lambda-it)x} = \lim_{x\to\infty} e^{-\lambda x}e^{itx} = 0.$$

Έπεται ότι

$$\varphi_X(t) = \frac{\lambda}{\lambda-it}.$$

Έστω ότι X και Y είναι ανεξάρτητες τυχαίες μεταβλητές. Τότε, οι e^{itX} και e^{itY} είναι επίσης ανεξάρτητες τυχαίες μεταβλητές. Επομένως,

$$\varphi_{X+Y}(t) = \mathbb{E}e^{it(X+Y)} = \mathbb{E}e^{itX}e^{itY} = \mathbb{E}e^{itX}\cdot\mathbb{E}e^{itY},$$

άρα

$$\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t), \qquad -\infty < t < \infty. \tag{11}$$

Ο τύπος (11) επεκτείνεται άμεσα με συνέπεια το γεγονός ότι η χαρακτηριστική συνάρτηση του αθροίσματος πεπερασμένου πλήθους ανεξάρτητων τυχαίων μεταβλητών είναι το γινόμενο των χαρακτηριστικών τους συναρτήσεων.

Αποδεικνύεται ότι η $\varphi_X(t)$ είναι συνεχής συνάρτηση του t. Επιπλέον, αν η X έχει πεπερασμένη ροπή τάξης n, τότε η $\varphi_X^{(n)}(t)$ υπάρχει και υπολογίζεται από την

$$\varphi_X^{(n)}(t) = \frac{d^n}{dt^n}\mathbb{E}e^{itX} = \mathbb{E}\frac{d^n}{dt^n}e^{itX} = \mathbb{E}(iX)^n e^{itX}.$$

Ειδικότερα,

$$\varphi_X^{(n)}(0) = i^n \mathbb{E}X^n. \tag{12}$$

Μπορούμε να επιχειρήσουμε να αναπτύξουμε την $\varphi_X(t)$ σε δυναμοσειρά σύμφωνα με τον τύπο

$$\varphi_X(t) = \mathbb{E}e^{itX} = \mathbb{E}\sum_{n=0}^{\infty}\frac{(itX)^n}{n!} = \sum_{n=0}^{\infty}\frac{i^n\mathbb{E}X^n}{n!}t^n. \tag{13}$$

Ας υποθέσουμε ότι η

$$M_X(t) = \sum_{n=0}^{\infty}\frac{\mathbb{E}X^n}{n!}t^n$$

είναι πεπερασμένη στο $-t_0 < t < t_0$ για κάποιον θετικό αριθμό t_0. Τότε η (13) ισχύει στο $-t_0 < t < t_0$.

Παράδειγμα 7. Υποθέτουμε ότι η X ακολουθεί κανονική κατανομή με μέσο 0 και διασπορά σ^2. Βρείτε την $\varphi_X(t)$.

Από το Κεφάλαιο 7 γνωρίζουμε ότι $\mathbb{E}X^n = 0$ για κάθε περιττό φυσικό αριθμό n. Επιπλέον, αν $n = 2k$ είναι ένας άρτιος φυσικός, τότε

$$\mathbb{E}X^n = \mathbb{E}X^{2k} = \frac{\sigma^{2k}(2k)!}{2^k k!}.$$

Επομένως,

$$\varphi_X(t) = \sum_{k=0}^{\infty}\frac{i^{2k}\mathbb{E}X^{2k}}{(2k)!}t^{2k} = \sum_{k=0}^{\infty}\frac{(-\sigma^2t^2/2)^k}{k!} = e^{-\sigma^2t^2/2}.$$

Πιό γενικά, έστω ότι η X ακολουθεί κανονική κατανομή με μέσο μ και διασπορά σ^2. Τότε η $Y = X - \mu$ ακολουθεί κανονική κατανομή με μέσο 0 και διασπορά σ^2. Αφού $X = \mu + Y$, από τον τύπο (10) και το Παράδειγμα 7 βλέπουμε ότι

$$\varphi_X(t) = e^{it\mu}e^{-\sigma^2t^2/2}, \qquad -\infty < t < \infty. \tag{14}$$

Έστω X τυχαία μεταβλητή της οποίας η ροπογεννήτρια $M_X(t)$ είναι πεπερασμένη στο $-t_0 < t < t_0$ για κάποιον θετικό αριθμό t_0. Αφού

$$M_X(t) = \mathbb{E}e^{tX}$$

και

$$\varphi_X(t) = \mathbb{E}e^{itX},$$

θα περιμέναμε να ισχύει η

$$\varphi_X(t) = M_X(it). \tag{15}$$

Με άλλα λόγια, θα περιμέναμε ότι αντικαθιστώντας το t με it στον τύπο για την ροπογεννήτρια, παίρνουμε τον τύπο για την χαρακτηριστική συνάρτηση. Αυτό πράγματι ισχύει, για να κατανοήσει όμως κανείς σε βάθος τη σχέση μεταξύ των δύο εννοιών απαιτείται να γνωρίζει την έννοια της αναλυτικής συνέχισης από τη θεωρία των μιγαδικών συναρτήσεων.

Για να δώσουμε ένα παράδειγμα της (15), θεωρούμε μία τυχαία μεταβλητή X που ακολουθεί κανονική κατανομή με μέσο μ και διασπορά σ^2. Τότε, όπως έχουμε ήδη δεί,

$$M_X(t) = e^{\mu t}e^{\sigma^2 t^2/2},$$

επομένως

$$\begin{aligned} M_X(it) &= e^{\mu(it)}e^{\sigma^2(it)^2/2} \\ &= e^{i\mu t}e^{-\sigma^2 t^2/2}, \end{aligned}$$

το οποίο συμπίπτει με την $\varphi_X(t)$ από την (14).

8.3 Τύποι αντιστροφής και το Θεώρημα Συνεχείας

Έστω X μία τυχαία μεταβλητή με ακέραιες τιμές. Η χαρακτηριστική της συνάρτηση δίνεται από την

$$\varphi_X(t) = \sum_{-\infty}^{\infty} e^{ijt} f_X(j).$$

Μία από τις πιό χρήσιμες ιδιότητες της $\varphi_X(t)$ είναι ότι με τη βοήθειά της μπορούμε να υπολογίσουμε την $f_X(k)$. Πιό συγκεκριμένα, έχουμε τον «τύπο αντιστροφής»

$$f_X(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ikt}\varphi_X(t)dt. \tag{16}$$

Για να επαληθεύσουμε την (16) γράφουμε το δεξιό μέλος αυτής της σχέσης στη μορφή

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ikt}\left[\sum_{-\infty}^{\infty} e^{ijt} f_X(j)\right] dt.$$

Σύμφωνα με ένα θεώρημα από τη θεωρία ολοκλήρωσης μπορούμε να εναλλάξουμε το ολοκλήρωμα με το άθροισμα, οπότε παίρνουμε την παράσταση

$$\sum_{-\infty}^{\infty} f_X(j)\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(j-k)t}dt.$$

Για να ολοκληρώσουμε την απόδειξη της (16) πρέπει να δείξουμε ότι η τελευταία παράσταση ισούται με $f_X(k)$. Για το σκοπό αυτό αρκεί να δείξουμε ότι

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(j-k)t}dt = \begin{cases} 1 & , \text{αν } j = k, \\ 0 & , \text{αν } j \neq k. \end{cases} \tag{17}$$

Η σχέση (17) είναι προφανής όταν $j = k$, γιατί σε αυτήν την περίπτωση $e^{i(j-k)t} = 1$ για κάθε t. Αν $j \neq k$, τότε

$$\begin{aligned} \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(j-k)t}dt &= \frac{e^{i(j-k)t}\big|_{-\pi}^{\pi}}{2\pi i(j-k)} \\ &= \frac{e^{i(j-k)t} - e^{-i(j-k)t}}{2\pi i(j-k)} \\ &= \frac{\sin(j-k)\pi}{\pi(j-k)} = 0, \end{aligned}$$

αφού $\sin m\pi = 0$ για όλους τους ακεραίους m. Αυτό ολοκληρώνει την απόδειξη της (17), άρα και της (16).

Παράδειγμα 8. Έστω $X_1, X_2, \ldots, X_n$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με ακέραιες τιμές. Θέτουμε $S_n = X_1 + \ldots + X_n$. Τότε $\varphi_{S_n}(t) = [\varphi_{X_1}(t)]^n$, επομένως από την (16),

$$f_{S_n}(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ikt}[\varphi_{X_1}(t)]^n dt. \tag{18}$$

Η σχέση (18) είναι η βάση για όλες τις μεθόδους ανάλυσης της συμπεριφοράς της $f_{S_n}(k)$ για μεγάλες τιμές του n και, ειδικότερα, η βάση για την απόδειξη του «τοπικού» Κεντρικού Οριακού Θεωρήματος που συζητήσαμε στο Κεφάλαιο 7.

Υπάρχει μία παραλλαγή της (16) για συνεχείς τυχαίες μεταβλητές. Έστω X μία τυχαία μεταβλητή της οποίας η χαρακτηριστική συνάρτηση $\varphi_X(t)$ είναι ολοκληρώσιμη, δηλαδή,

$$\int_{-\infty}^{\infty} |\varphi_X(t)|dt < \infty.$$

Αποδεικνύεται ότι σε αυτήν την περίπτωση η X είναι μία συνεχής τυχαία μεταβλητή με πυκνότητα f_X που δίνεται από την

$$f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ixt}\varphi_X(t)dt. \tag{19}$$

Παράδειγμα 9. Υποθέτουμε ότι η X ακολουθεί κανονική κατανομή με μέσο 0 και διασπορά σ^2. Θα αποδείξουμε απευθείας ότι η (19) ισχύει για μία τέτοια τυχαία μεταβλητή. Από το Παράδειγμα 7 γνωρίζουμε ότι η X έχει χαρακτηριστική συνάρτηση την $\varphi_X(t) = e^{-\sigma^2 t^2/2}$. Από τον ορισμό της χαρακτηριστικής συνάρτησης αυτό σημαίνει ότι

$$e^{-\sigma^2 t^2/2} = \int_{-\infty}^{\infty} e^{itx} \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2} dx.$$

Αν αντικαταστήσουμε το t με $-t$ και το σ με $1/\sigma$ σε αυτήν την ισότητα, παίρνουμε

$$e^{-t^2/2\sigma^2} = \int_{-\infty}^{\infty} e^{-itx} \frac{\sigma}{\sqrt{2\pi}} e^{-\sigma^2 x^2/2} dx,$$

ή ισοδύναμα,

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-t^2/2\sigma^2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} e^{-\sigma^2 x^2/2} dx.$$

Τέλος, αν εναλλάξουμε τους ρόλους των συμβόλων x και t στην τελευταία ισότητα, παίρνουμε

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} e^{-\sigma^2 t^2/2} dx,$$

που είναι ακριβώς η (19) σε αυτήν την ειδική περίπτωση.

Έστω X τυχούσα τυχαία μεταβλητή. Έστω Y μία τυχαία μεταβλητή που είναι ανεξάρτητη από την X και ακολουθεί την τυπική κανονική κατανομή, και έστω c μία θετική σταθερά. Τότε η $X + cY$ έχει χαρακτηριστική συνάρτηση την

$$\varphi_X(t) e^{-c^2 t^2/2}.$$

Αφού η $\varphi_X(t)$ είναι φραγμένη από 1 κατ' απόλυτη τιμή και η $e^{-c^2t^2/2}$ είναι ολοκληρώσιμη, έπεται ότι η $X+cY$ έχει ολοκληρώσιμη χαρακτηριστική συνάρτηση. Επομένως εφαρμόζεται η (19) και η $X + cY$ είναι συνεχής τυχαία μεταβλητή με πυκνότητα την

$$f_{X+cY}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi_X(t) e^{-c^2 t^2/2} dt.$$

Αν ολοκληρώσουμε τα δύο μέλη αυτής της ισότητας στο $a \le x \le b$ και εναλλάξουμε τη σειρά της ολοκλήρωσης, συμπεραίνουμε ότι

$$\begin{aligned} P(a \le X + cY \le b) &= \frac{1}{2\pi} \int_a^b \left(\int_{-\infty}^{\infty} e^{-itx} \varphi_X(t) e^{-c^2t^2/2} dt \right) dx \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_a^b e^{-itx} dx \right) \varphi_X(t) e^{-c^2t^2/2} dt \end{aligned}$$

δηλαδή

$$(20) \qquad P(a \le X + cY \le b) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\frac{e^{-ibt} - e^{-iat}}{-it} \right) \varphi_X(t) e^{-c^2t^2/2} dt.$$

Η σημασία της (20) έγκειται στο ότι ισχύει για οποιαδήποτε τυχαία μεταβλητή X.

Το δεξιό μέλος της (20) εξαρτάται από την X μόνο μέσω της $\varphi_X(t)$. Χρησιμοποιώντας αυτό το γεγονός και αφήνοντας το $c \to 0$ στην (20), μπορούμε να δείξουμε ότι η συνάρτηση κατανομής της X προσδιορίζεται από την χαρακτηριστική της συνάρτηση. Το αποτέλεσμα αυτό είναι γνωστό σαν «θεώρημα μοναδικότητας» και διατυπώνεται ως εξής:

Θεώρημα 1 *Αν δύο τυχαίες μεταβλητές έχουν την ίδια χαρακτηριστική συνάρτηση, τότε έχουν την ίδια συνάρτηση κατανομής.*

Παράδειγμα 10. Χρησιμοποιώντας το θεώρημα μοναδικότητας δείξτε ότι το άθροισμα δύο ανεξάρτητων κανονικά κατανεμημένων τυχαίων μεταβλητών ακολουθεί κανονική κατανομή.

Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν τις κατανομές $n(\mu_1, \sigma_1^2)$ και $n(\mu_2, \sigma_2^2)$ αντιστοίχως. Τότε

$$\varphi_X(t) = e^{i\mu_1 t} e^{-\sigma_1^2 t^2/2}$$

και

$$\varphi_Y(t) = e^{i\mu_2 t} e^{-\sigma_2^2 t^2/2}.$$

Επομένως,

$$\varphi_{X+Y}(t) = e^{i(\mu_1+\mu_2)t} e^{-(\sigma_1^2+\sigma_2^2)t^2/2}.$$

Άρα, η χαρακτηριστική συνάρτηση της $X+Y$ συμπίπτει με αυτήν μιάς τυχαίας μεταβλητής που ακολουθεί κανονική κατανομή με μέσο $\mu_1+\mu_2$ και διασπορά $\sigma_1^2+\sigma_2^2$. Από το θεώρημα μοναδικότητας, η $X+Y$ πρέπει να ακολουθεί αυτήν την κανονική κατανομή.

Η πιό σημαντική εφαρμογή του τύπου αντιστροφής (20) είναι ότι χρησιμοποιώντας τον μπορούμε να αποδείξουμε το επόμενο αποτέλεσμα, το οποίο είναι βασικό για την απόδειξη του Ασθενούς Νόμου των Μεγάλων Αριθμών και του Κεντρικού Οριακού Θεωρήματος.

Θεώρημα 2 *Έστω X_n, $n \geq 1$, και X τυχαίες μεταβλητές τέτοιες ώστε*

$$\lim_{n\to\infty} \varphi_{X_n}(t) = \varphi_X(t), \qquad -\infty < t < \infty. \tag{21}$$

Τότε,

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x) \tag{22}$$

σε όλα τα σημεία x στα οποία η F_X είναι συνεχής.

Το θεώρημα αυτό ισχυρίζεται ότι η σύγκλιση των χαρακτηριστικών συναρτήσεων έχει σαν συνέπεια την σύγκλιση των αντιστοίχων συναρτήσεων κατανομής ή, με άλλα λόγια, ότι οι συναρτήσεις κατανομής «εξαρτώνται συνεχώς» από τις χαρακτηριστικές τους συναρτήσεις. Για το λόγο αυτό, το Θεώρημα 2 είναι γνωστό και ως «Θεώρημα Συνεχείας».

Η απόδειξη αυτού του θεωρήματος είναι αρκετά πολύπλοκη. Δεν θα παρουσιάσουμε τις λεπτομέρειες της απόδειξης, θα δώσουμε όμως εν συντομία μερικές από τις κύριες ιδέες μιάς από τις μεθόδους απόδειξης.

Επιλέγουμε πρώτα μία τυχαία μεταβλητή Y που ακολουθεί την τυπική κανονική κατανομή και είναι ανεξάρτητη από καθεμία εκ των τυχαίων μεταβλητών X_n, $n \geq 1$.

Έστω $a < b$ και έστω c μία θετική σταθερά. Από τον τύπο αντιστροφής (20)

$$P(a \leq X_n + cY \leq b) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\frac{e^{-ibt} - e^{-iat}}{-it} \right) \varphi_{X_n}(t) e^{-c^2t^2/2} dt \tag{23}$$

και

$$P(a \leq X + cY \leq b) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\frac{e^{-ibt} - e^{-iat}}{-it} \right) \varphi_{X}(t) e^{-c^2t^2/2} dt. \tag{24}$$

Από την υπόθεση, $\varphi_{X_n}(t) \to \varphi_X(t)$ καθώς $n \to \infty$. Από ένα θεώρημα της θεωρίας ολοκλήρωσης έπεται ότι το δεξιό μέλος της (23) συγκλίνει στο δεξιό μέλος της (24) καθώς $n \to \infty$. Επομένως,

$$\lim_{n \to \infty} P(a \leq X_n + cY \leq b) = P(a \leq X + cY \leq b). \tag{25}$$

Για την απόδειξη του θεωρήματος απαιτούνται δύο ακόμα βήματα. Πρέπει πρώτα να δείξουμε (αφήνοντας το $a \to -\infty$ στην (25)) ότι

$$\lim_{n \to \infty} P(X_n + cY \leq b) = P(X + cY \leq b). \tag{26}$$

Τέλος, πρέπει να δείξουμε (αφήνοντας το $c \to 0$ στην (26)) ότι

$$\lim_{n \to \infty} P(X_n \leq b) = P(X \leq b)$$

με την προϋπόθεση ότι $P(X = b) = 0$. Το τελευταίο συμπέρασμα είναι ισοδύναμο με το συμπέρασμα του θεωρήματος.

8.4 Ο Ασθενής Νόμος των Μεγάλων Αριθμών και το Κεντρικό Οριακό Θεώρημα

Σε αυτήν την παράγραφο θα χρησιμοποιήσουμε το Θεώρημα Συνεχείας για να αποδείξουμε τα δύο σημαντικά θεωρήματα της θεωρίας πιθανοτήτων που αναφέρονται στον τίτλο αυτής της παραγράφου. Σε προηγούμενα κεφάλαια συζητήσαμε και τα δύο θεωρήματα χωρίς απόδειξη. Για την απόδειξή τους θα χρειαστεί πρώτα να μελετήσουμε την ασυμπτωτική συμπεριφορά της $\log \varphi_X(t)$ κοντά στο $t = 0$.

Έστω z ένας μιγαδικός αριθμός τέτοιος ώστε $|z-1| < 1$. Μπορούμε να ορίσουμε τον $\log z$ μέσω της δυναμοσειράς

$$\log z = (z-1) - \frac{(z-1)^2}{2} + \frac{(z-1)^3}{3} - \ldots$$

(αν $|z-1| \geq 1$, χρειαζόμαστε άλλον τρόπο ορισμού του $\log z$). Από αυτόν τον ορισμό προκύπτουν οι συνήθεις ιδιότητες $\log 1 = 0$,

$$e^{\log z} = z, \qquad |z-1| < 1,$$

και αν $h(t)$, $a < t < b$, είναι μία παραγωγίσιμη μιγαδική συνάρτηση τέτοια ώστε $|h(t) - 1| < 1$, τότε

$$\frac{d}{dt} \log h(t) = \frac{h'(t)}{h(t)}.$$

Έστω X μία τυχαία μεταβλητή με χαρακτηριστική συνάρτηση την $\varphi_X(t)$. Τότε, η $\varphi_X(t)$ είναι συνεχής και $\varphi_X(0) = 1$. Άρα, η $\log \varphi_X(t)$ είναι καλά ορισμένη για t κοντά στο 0 και $\log \varphi_X(0) = 0$.

Ας υποθέσουμε τώρα ότι η X έχει πεπερασμένο μέσο μ. Τότε η $\varphi_X(t)$ είναι παραγωγίσιμη και, από την (12), $\varphi_X'(0) = i\mu$. Άρα,

$$\begin{aligned} \lim_{t\to 0} \frac{\log \varphi_X(t)}{t} &= \lim_{t\to 0} \frac{\log \varphi_x(t) - \log \varphi_X(0)}{t-0} \\ &= \frac{d}{dt} \log \varphi_x(t)\bigg|_{t=0} \\ &= \frac{\varphi_X'(0)}{\varphi_X(0)} \\ &= i\mu. \end{aligned}$$

Επομένως,

$$\lim_{t\to 0} \frac{\log \varphi_X(t) - i\mu t}{t} = 0. \tag{27}$$

Ας υποθέσουμε τώρα ότι η X έχει και πεπερασμένη διασπορά σ^2. Τότε η $\varphi_X(t)$ είναι δύο φορές παραγωγίσιμη και από την (12)

$$\varphi_X''(0) = -\mathbb{E}X^2 = -(\mu^2 + \sigma^2).$$

Εφαρμόζοντας τον κανόνα του l' Hôspital παίρνουμε

$$\begin{aligned} \lim_{t\to 0} \frac{\log \varphi_X(t) - i\mu t}{t^2} &= \lim_{t\to 0} \frac{\frac{\varphi_X'(t)}{\varphi_X(t)} - i\mu}{2t} \\ &= \lim_{t\to 0} \frac{\varphi_X'(t) - i\mu\varphi_X(t)}{2t\varphi_X(t)} \\ &= \lim_{t\to 0} \frac{\varphi_X'(t) - i\mu\varphi_X(t)}{2t}. \end{aligned}$$

Με δεύτερη εφαρμογή του κανόνα του l' Hôspital βλέπουμε ότι

$$\begin{aligned} \lim_{t\to 0} \frac{\log \varphi_X(t) - i\mu t}{t^2} &= \frac{\varphi_X''(0) - i\mu\varphi_X'(0)}{2} \\ &= \frac{-(\mu^2 + \sigma^2) - (i\mu)^2}{2} \\ &= \frac{-\mu^2 - \sigma^2 + \mu^2}{2}. \end{aligned}$$

Με άλλα λόγια,

$$\lim_{t\to 0} \frac{\log \varphi_X(t) - i\mu t}{t^2} = -\frac{\sigma^2}{2}. \tag{28}$$

Θεώρημα 3 (Ασθενής Νόμος των Μεγάλων Αριθμών) *Έστω* $X_1, X_2, \ldots$ *ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με πεπερασμένο μέσο* μ. *Θέτουμε* $S_n = X_1 + \ldots + X_n$. *Τότε, για κάθε* $\varepsilon > 0$

$$\lim_{n\to\infty} P\left(|\frac{S_n}{n} - \mu| > \varepsilon\right) = 0. \tag{29}$$

Απόδειξη. Η χαρακτηριστική συνάρτηση της

$$\frac{S_n}{n} - \mu = \frac{X_1 + \ldots + X_n}{n} - \mu$$

είναι η

$$e^{-i\mu t}\left(\varphi_{X_1}(t/n)\right)^n.$$

Έστω t σταθερό. Τότε για n αρκετά μεγάλο, ο t/n είναι αρκετά κοντά στο μηδέν, οπότε ο $\log \varphi_{X_1}(t/n)$ ορίζεται καλά και

$$e^{-i\mu t}\left(\varphi_{X_1}(t/n)\right)^n = \exp\left[n(\log \varphi_{X_1}(t/n) - i\mu(t/n))\right]. \tag{30}$$

Στη συνέχεια ισχυριζόμαστε ότι

$$\lim_{n\to\infty} n(\log \varphi_{X_1}(t/n) - i\mu(t/n)) = 0. \tag{31}$$

Η ισότητα (31) ισχύει προφανώς όταν $t = 0$, αφού $\log \varphi_{X_1}(0) = \log 1 = 0$. Αν $t \neq 0$ μπορούμε να γράψουμε το αριστερό μέλος της (31) στη μορφή

$$t \lim_{n\to\infty} \frac{\log \varphi_{X_1}(t/n) - i\mu(t/n)}{t/n}.$$

Όμως $t/n \to 0$ καθώς $n \to \infty$, επομένως το τελευταίο όριο είναι 0 από την (27). Αυτό ολοκληρώνει την απόδειξη της (31). Από τις (30) και (31) έπεται ότι η χαρακτηριστική συνάρτηση της

$$\frac{S_n}{n} - \mu$$

πλησιάζει το 1 καθώς $n \to \infty$. Όμως η 1 είναι η χαρακτηριστική συνάρτηση μιάς τυχαίας μεταβλητής X που ικανοποιεί την $P(X = 0) = 1$. Η συνάρτηση κατανομής της X δίνεται από την

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

Αυτή η συνάρτηση κατανομής είναι συνεχής παντού, εκτός από το $x = 0$. Έστω $\varepsilon > 0$. Από το Θεώρημα Συνεχείας,

$$\lim_{n\to\infty} P\left(\frac{S_n}{n} - \mu \le -\varepsilon\right) = F_X(-\varepsilon) = 0 \tag{32}$$

και

$$\lim_{n\to\infty} P\left(\frac{S_n}{n} - \mu \le \varepsilon\right) = F_X(\varepsilon) = 1.$$

Η τελευταία σχέση μάς δείχνει ότι

$$\lim_{n\to\infty} P\left(\frac{S_n}{n} - \mu > \varepsilon\right) = 0,$$

το οποίο μαζί με την (32) μάς δίνει την (29). □

Για το επόμενο θεώρημα θα χρειαστεί να θυμηθείτε ότι $\Phi(x)$ είναι η τυπική κανονική συνάρτηση κατανομής που δίνεται από την

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy, \qquad -\infty < x < \infty.$$

Θα χρειαστούμε επίσης την παρατήρηση ότι αυτή η συνάρτηση κατανομής είναι συνεχής για όλες τις τιμές του x.

Θεώρημα 4 (Κεντρικό Οριακό Θεώρημα). *Έστω $X_1, X_2, \ldots$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με μέσο μ και πεπερασμένη μη μηδενική διασπορά σ^2. Τότε,*

$$\lim_{n\to\infty} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \le x\right) = \Phi(x), \qquad -\infty < x < \infty.$$

Απόδειξη. Θέτουμε

$$S_n^* = \frac{S_n - n\mu}{\sigma\sqrt{n}}.$$

Τότε, για σταθερό t και n αρκετά μεγάλο,

$$\begin{aligned} \varphi_{S_n^*}(t) &= e^{-in\mu t/\sigma\sqrt{n}} \varphi_{S_n}(t/\sigma\sqrt{n}) \\ &= e^{-in\mu t/\sigma\sqrt{n}} \left(\varphi_{X_1}(t/\sigma\sqrt{n})\right)^n, \end{aligned}$$

δηλαδή,

$$\varphi_{S_n^*}(t) = \exp\left[n(\log \varphi_{X_1}(t/\sigma\sqrt{n}) - i\mu(t/\sigma\sqrt{N}))\right]. \tag{33}$$

Ο επόμενος ισχυρισμός μας είναι ότι

$$\lim_{n\to\infty} n(\log \varphi_{X_1}(t/\sigma\sqrt{n}) - i\mu(t/\sigma\sqrt{N})) = -\frac{t^2}{2}. \tag{34}$$

Αν $t = 0$, τότε τα δύο μέλη της (34) είναι ίσα με μηδέν, οπότε η (34) προφανώς ισχύει. Αν $t \neq 0$, μπορούμε να γράψουμε το αριστερό μέλος της (34) στη μορφή

$$\frac{t^2}{\sigma^2} \lim_{n\to\infty} \frac{\log \varphi_{X_1}(t/\sigma\sqrt{n}) - i\mu(t/\sigma\sqrt{N})}{(t/\sigma\sqrt{n})^2},$$

το οποίο από την (28) ισούται με

$$\frac{t^2}{\sigma^2}\left(-\frac{\sigma^2}{2}\right) = -\frac{t^2}{2}.$$

Άρα η (34) ισχύει για κάθε t. Από τις (33) και (34) βλέπουμε ότι

$$\lim_{n\to\infty} \varphi_{S_n^*}(t) = e^{-t^2/2}, \qquad -\infty < t < \infty.$$

Σύμφωνα με το Παράδειγμα 7, η $e^{-t^2/2}$ είναι η χαρακτηριστική συνάρτηση μιάς τυχαίας μεταβλητής X που έχει την τυπική κανονική συνάρτηση κατανομής $\Phi(x)$. Από το Θεώρημα Συνεχείας συμπεραίνουμε ότι

$$\lim_{n\to\infty} P(S_n^* \leq x) = \Phi(x), \qquad -\infty < x < \infty,$$

το οποίο είναι το ζητούμενο συμπέρασμα. □

8.5 Ασκήσεις

1 Έστω ότι η X είναι ομοιόμορφα κατανεμημένη στο (a, b). Βρείτε την $M_X(t)$.

2 Εκφράστε την ροπογεννήτρια της $Y = a + bX$ συναρτήσει της $M_X(t)$ (όπου a και b είναι σταθερές).

3 Υποθέτουμε ότι η X ακολουθεί κατανομή Poisson με παράμετρο λ. Χρησιμοποιήστε ροπογεννήτριες για να βρείτε τον μέσο και την διασπορά της X.

4 Έστω ότι η X ακολουθεί αρνητική διωνυμική κατανομή με παραμέτρους α και p.
(α) Βρείτε την ροπογεννήτρια της X.
(β) Χρησιμοποιώντας αυτήν την ροπογεννήτρια, βρείτε τον μέσο και την διασπορά της X.

5 Έστω X συνεχής τυχαία μεταβλητή με πυκνότητα $f_X(x) = (1/2)e^{-|x|}$, $-\infty < x < \infty$.
(α) Δείξτε ότι $M_X(t) = 1/(1-t^2)$, $-1 < t < 1$.
(β) Χρησιμοποιώντας αυτήν την ροπογεννήτρια, βρείτε τύπο για την $\mathbb{E}X^{2n}$ (παρατηρήστε ότι οι περιττές ροπές της X είναι όλες μηδέν).

6 Έστω ότι η X ακολουθεί διωνυμική κατανομή με παραμέτρους n και p.
(α) Βρείτε τις $dM_X(t)/dt$ και $d^2M_X(t)/dt^2$.

(β) Χρησιμοποιήστε το (α) και τον τύπο (7) για να υπολογίσετε τον μέσο και την διασπορά της X.

7 Έστω $X_1, \ldots, X_n$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές τέτοιες ώστε η $M_X(t)$ να ορίζεται για κάθε t. Χρησιμοποιώντας ροπογεννήτριες, δείξτε ότι

$$\begin{aligned}\mathbb{E}(X_1 + \ldots + X_n)^3 &= n\mathbb{E}X_1^3 + 3n(n-1)\mathbb{E}X_1^2\mathbb{E}X_1 \\ &+ n(n-1)(n-2)(\mathbb{E}X_1)^3.\end{aligned}$$

Υπόδειξη: Βρείτε την $(d^3/dt^3)(M_{X_1}(t))^n|_{t=0}$.

8 Έστω X τυχαία μεταβλητή τέτοια ώστε η $M_X(t)$ να ορίζεται για κάθε t. Μιμούμενοι το επιχείρημα της απόδειξης της ανισότητας του Chebyshev, δείξτε ότι

$$P(X \geq x) \leq e^{-tx}M_X(t), \quad t \geq 0.$$

Έπεται ότι

$$P(X \geq x) \leq \min_{t \geq 0} e^{-tx}M_X(t),$$

αρκεί η $e^{-tx}M_X(t)$ να παίρνει ελάχιστη τιμή στο $0 \leq t < \infty$.

9 Έστω ότι η X έχει κατανομή γάμμα με παραμέτρους α και λ. Χρησιμοποιώντας το αποτέλεσμα της Άσκησης 8, δείξτε ότι $P(X \geq 2\alpha/\lambda) \leq (2/e)^\alpha$.

10 Έστω ότι η X έχει κατανομή Poisson με παράμετρο λ. Βρείτε την $\varphi_X(t)$.

11 Έστω ότι η X έχει γεωμετρική κατανομή με παράμετρο p. Βρείτε την $\varphi_X(t)$.

12 Έστω $X_1, X_2, \ldots, X_n$ ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν γεωμετρική κατανομή με παράμετρο p. Βρείτε την χαρακτηριστική συνάρτηση της $X = X_1 + \ldots + X_n$.

13 Έστω $X_1, X_2, \ldots, X_n$ ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν εκθετική κατανομή με παράμετρο λ. Βρείτε την χαρακτηριστική συνάρτηση της $X = X_1 + \ldots + X_n$.

14 Έστω X μία διακριτή τυχαία μεταβλητή της οποίας όλες οι πιθανές τιμές είναι μη αρνητικοί ακέραιοι. Ποιά είναι η αναμενόμενη σχέση ανάμεσα στην χαρακτηριστική συνάρτηση της X και τη γεννήτρια συνάρτηση πιθανότητας της X; Θυμηθείτε τις σχέσεις (3) και (15).

15 Έστω X οποιαδήποτε τυχαία μεταβλητή.
(α) Δείξτε ότι $\varphi_X(t) = \mathbb{E}\cos tX + i\mathbb{E}\sin tX$.
(β) Δείξτε ότι $\varphi_{-X}(t) = \mathbb{E}\cos tX - i\mathbb{E}\sin tX$.
(γ) Δείξτε ότι $\varphi_{-X}(t) = \varphi_X(-t)$.

16 Έστω X συμμετρική τυχαία μεταβλητή, δηλαδή, τέτοια ώστε οι X και $-X$ να έχουν την ίδια συνάρτηση κατανομής.
(α) Δείξτε ότι $\mathbb{E}\sin tX = 0$ και ότι η $\varphi_X(t)$ παίρνει πραγματικές τιμές.
(β) Δείξτε ότι $\varphi_X(-t) = \varphi_X(t)$.

17 Έστω X και Y δύο ανεξάρτητες και ισόνομες τυχαίες μεταβλητές. Δείξτε ότι $\varphi_{X-Y}(t) = |\varphi_X(t)|^2$. *Υπόδειξη:* Χρησιμοποιήστε την Άσκηση 15.

18 Έστω X τυχαία μεταβλητή τέτοια ώστε η $\varphi_X(t)$ να παίρνει πραγματικές τιμές.

(α) Δείξτε ότι οι X και $-X$ έχουν την ίδια χαρακτηριστική συνάρτηση (χρησιμοποιήστε την Άσκηση 15).

(β) Γιατί έπεται ότι οι X και $-X$ έχουν την ίδια συνάρτηση κατανομής;

19 Έστω X συνεχής τυχαία μεταβλητή με πυκνότητα την $f_X(x) = (1/2)e^{-|x|}$, $-\infty < x < \infty$.

(α) Δείξτε ότι $\varphi_X(t) = 1/(1+t^2)$.

(β) Χρησιμοποιήστε το (α) και τον τύπο αντιστροφής (19) για να δείξετε ότι

$$e^{-|x|} = \int_{-\infty}^{\infty} e^{-itx} \frac{1}{\pi(1+t^2)} dt.$$

(γ) Χρησιμοποιώντας το (β) δείξτε ότι

$$e^{-|x|} = \int_{-\infty}^{\infty} e^{ixt} \frac{1}{\pi(1+t^2)} dt.$$

20 Έστω X τυχαία μεταβλητή που έχει πυκνότητα Cauchy

$$f_X(x) = \frac{1}{\pi(1+x^2)}, \qquad -\infty < x < \infty.$$

Δείξτε ότι $\varphi_X(t) = e^{-|x|}$, $-\infty < t < \infty$. *Υπόδειξη:* Εναλλάξτε τους ρόλους των x και t στην Άσκηση 19.

21 Έστω X και Y ανεξάρτητες τυχαίες μεταβλητές που έχουν πυκνότητα Cauchy.

(α) Βρείτε τις χαρακτηριστικές συναρτήσεις των $X+Y$ και $(X+Y)/2$.

(β) Γιατί έπεται ότι και η $(X+Y)/2$ έχει πυκνότητα Cauchy;

22 Επεκτείνοντας το αποτέλεσμα της Άσκησης 21, δείξτε ότι αν $X_1, X_2, \ldots X_n$ είναι ανεξάρτητες τυχαίες μεταβλητές καθεμία από τις οποίες έχει πυκνότητα Cauchy, τότε η $(X_1 + \ldots + X_n)/n$ έχει επίσης πυκνότητα Cauchy.

23 Για κάθε $\lambda > 0$ θεωρούμε μία τυχαία μεταβλητή X_λ που ακολουθεί κατανομή Poisson με παράμετρο λ.

(α) Χρησιμοποιώντας επιχειρήματα ανάλογα με αυτά της απόδειξης του Κεντρικού Οριακού Θεωρήματος, δείξτε ότι αν $-\infty < t < \infty$,

$$\lim_{\lambda\to\infty} \mathbb{E}e^{it(X_\lambda - \lambda)/\sqrt{\lambda}} = \lim_{\lambda\to\infty} \exp\left[\lambda(e^{it/\sqrt{\lambda}} - 1 - it/\sqrt{\lambda})\right] = e^{-t^2/2}.$$

(β) Ποιό συμπέρασμα προκύπτει από το (α) με κατάλληλη τροποποίηση του Θεωρήματος Συνεχείας;

Κεφάλαιο 9

Τυχαίοι περίπατοι και Διαδικασίες Poisson

Σε αυτό το κεφάλαιο θα συζητήσουμε δύο στοιχειώδη αλλά σημαντικά παραδείγματα στοχαστικών διαδικασιών. Με τον όρο *στοχαστική διαδικασία* μπορούμε να εννοούμε οποιαδήποτε συλλογή τυχαίων μεταβλητών. Συνήθως, όμως, όταν αναφερόμαστε σε μία στοχαστική διαδικασία έχουμε στο νού μας μία διαδικασία που έχει αρκετή πρόσθετη δομή ώστε να μπορούν να εξαχθούν ενδιαφέροντα και χρήσιμα συμπεράσματα. Αυτό αληθεύει σαφώς για τα δύο παραδείγματα που θα μάς απασχολήσουν στο παρόν κεφάλαιο. Το υλικό που αφορά τις διαδικασίες Poisson, το δεύτερο παράδειγμά μας, είναι ανεξάρτητο από αυτό των δύο πρώτων παραγράφων στις οποίες μελετάμε τυχαίους περιπάτους.

9.1 Τυχαίοι περίπατοι

Θεωρούμε την εξής ακολουθία παιχνιδιών: στη διάρκεια του n-στού παιχνιδιού, παρατηρούμε μία τυχαία μεταβλητή X_n και κάθε παίκτης που συμμετέχει στο n-στό παιχνίδι εισπράττει ποσό X_n από το καζίνο (φυσικά, αν $X_n < 0$ ο παίκτης στην πραγματικότητα πληρώνει ποσό $-X_n$ στο καζίνο).

Ας παρακολουθήσουμε την πρόοδο ενός παίκτη που ξεκινάει με αρχικό κεφάλαιο x. Με S_n, $n \geq 0$, συμβολίζουμε το κεφάλαιό του μετά από n παιχνίδια. Τότε, $S_0 = x$ και

$$S_n = x + X_1 + \ldots + X_n, \quad n \geq 1.$$

Η ακολουθία των τυχαίων μεταβλητών $S_0, S_1, S_2, \ldots$ είναι ένα παράδειγμα στοχαστικής διαδικασίας. Για να αντλήσουμε ενδιαφέροντα συμπεράσματα, θα υποθέσουμε ότι οι τυχαίες μεταβλητές $X_1, X_2, \ldots$ είναι ανεξάρτητες και ισόνομες. Με αυτήν την υπόθεση, λέμε ότι η διαδικασία $S_0, S_1, \ldots$ λέγεται *τυχαίος περίπατος*.

Θα υποθέσουμε επιπλέον ότι οι X_k έχουν πεπερασμένο μέσο μ. Αν ένας παίκτης συμμετάσχει στα πρώτα n παιχνίδια, το αναμενόμενο κεφάλαιό του με τη συμπλήρωση του n-στού παιχνιδιού είναι

$$\mathbb{E}S_n = x + n\mu. \tag{1}$$

Έστω όμως ότι ο παίκτης επιλέγει δύο αριθμούς $a \leq x$ και $b \geq x$ και συμφωνεί με τον εαυτό του ότι θα σταματήσει να παίζει αν το κεφάλαιό του γίνει μικρότερο ή ίσο από a ή μεγαλύτερο ή ίσο από b. Τότε, το πλήθος T των παιχνιδιών στα οποία θα παίξει είναι μία τυχαία μεταβλητή που ορίζεται από την

$$T = \min(n \geq 0 \mid S_n \leq a \ \ \eta \ \ S_n \geq b).$$

Για να είμαστε σε θέση να εγγυηθούμε ότι $S_n \leq a$ ή $S_n \geq b$ για κάποιο n, υποθέτουμε ότι

$$P(X_k = 0) < 1. \tag{2}$$

Αποδεικνύεται ότι η τυχαία μεταβλητή T παίρνει πεπερασμένες τιμές (με πιθανότητα 1) και, ακόμα ισχυρότερα, η $P(T > n)$ φθίνει εκθετικά καθώς $n \to \infty$. Αυτό σημαίνει ότι για κάποιες θετικές σταθερές M και $c < 1$,

$$P(T > n) < Mc^n, \qquad n = 0, 1, 2, \ldots \tag{3}$$

Η απόδειξη της (3) δεν είναι δύσκολη, την παραλείπουμε όμως για να αφήσουμε χώρο για αποτελέσματα που παρουσιάζουν πολύ μεγαλύτερο ενδιαφέρον. Από την (3) και το Θεώρημα 5 του Κεφαλαίου 4, έπεται ότι η $\mathbb{E}T$ και όλες οι ροπές υψηλότερης τάξης της T είναι πεπερασμένες.

Αν ο παίκτης εγκαταλείψει το παιχνίδι μετά το T-στό παιχνίδι, το κεφάλαιό του θα είναι S_T (δείτε το Σχήμα 1). Μία πολύ γνωστή ταυτότητα του Abraham Wald εκφράζει το αναμενόμενο κεφάλαιο τη στιγμή που ο παίκτης εγκαταλείπει το παιχνίδι συναρτήσει του αναμενόμενου πλήθους των παιχνιδιών στα οποία συμμετέχει. Συγκεκριμένα, η ταυτότητα του Wald μάς λέει ότι

$$\mathbb{E}S_T = x + \mu\mathbb{E}T. \tag{4}$$

Παρατηρήστε την αξιοσημείωτη ομοιότητα της ταυτότητας του Wald με την (1).

Για την απόδειξη της ταυτότητας του Wald θα χρησιμοποιήσουμε νέο συμβολισμό. Έστω A τυχόν ενδεχόμενο. Με 1_A συμβολίζουμε την δείκτρια τυχαία μεταβλητή του A, δηλαδή την τυχαία μεταβλητή που παίρνει την τιμή 1 αν συμβαίνει το A και την τιμή 0 αν δεν συμβαίνει το A. Από τον ορισμό έχουμε $1_A + 1_{A^c} = 1$. Χρησιμοποιώντας αυτόν το συμβολισμό μπορούμε να γράψουμε

$$S_T = x + \sum_{j=1}^{T} X_j = x + \sum_{j=1}^{\infty} X_j 1_{\{T \geq j\}}.$$

Δεδομένου ότι το συμπλήρωμα του ενδεχομένου $\{T \geq j\}$ είναι το ενδεχόμενο $\{T < j\}$, βλέπουμε ότι

$$S_T = x + \sum_{j=1}^{T} X_j = x + \sum_{j=1}^{\infty} X_j (1 - 1_{\{T<j\}}), \tag{5}$$

επομένως

$$\mathbb{E}S_T = x + \mathbb{E}\sum_{j=1}^{T} X_j = x + \sum_{j=1}^{\infty} X_j (1 - 1_{\{T<j\}}). \tag{6}$$

Με χρήση θεωρίας μέτρου αποδεικνύεται ότι μπορούμε να εναλλάξουμε τη μέση τιμή με την άθροιση στην (6). Άρα,

$$\mathbb{E}S_T = x + \sum_{j=1}^{T} X_j = x + \sum_{j=1}^{\infty} \mathbb{E}[X_j (1 - 1_{\{T<j\}})]. \tag{7}$$

Για να αποφασίσουμε αν $T < j$ ή όχι, αρκεί να κοιτάξουμε τις τυχαίες μεταβλητές $X_1, X_2, \ldots, X_{j-1}$. Αυτό σημαίνει ότι οι τυχαίες μεταβλητές X_j και $(1 - 1_{\{T<j\}})$ είναι ανεξάρτητες. Επομένως,

$$\begin{aligned} \mathbb{E}[X_j (1 - 1_{\{T,j\}})] &= \mathbb{E}X_j \mathbb{E}(1 - 1_{\{T<j\}}) \\ &= \mu(1 - P(T < j)) \\ &= \mu P(T \geq j). \end{aligned}$$

Από την (6) και το Θεώρημα 5 του Κεφαλαίου 4 συμπεραίνουμε τώρα ότι

$$\begin{aligned} \mathbb{E}S_T &= x + \mu \sum_{j=1}^{\infty} P(T \geq j) \\ &= x + \mu\mathbb{E}T, \end{aligned}$$

το οποίο ολοκληρώνει την απόδειξη της ταυτότητας του Wald.

Αν οι X_n έχουν μέσο $\mu = 0$ και πεπερασμένη διασπορά σ^2, υπάρχει μία δεύτερη ταυτότητα του Wald, η οποία μάς λέει ότι

$$\text{Var} S_T = \sigma^2 \mathbb{E} T. \tag{8}$$

Αφού $\mathbb{E} S_T = x$ από την (4), η (8) είναι ισοδύναμη με την

$$\mathbb{E}(S_T - x)^2 = \sigma^2 \mathbb{E} T. \tag{9}$$

Θα αποδείξουμε τώρα την (9). Από την (5),

$$S_T - x = \sum_{j=1}^{\infty} X_j (1 - 1_{\{T<j\}}),$$

επομένως

$$\begin{aligned} (S_T - x)^2 &= \sum_{j=1}^{\infty} X_j (1 - 1_{\{T<j\}}) \sum_{k=1}^{\infty} X_k (1 - 1_{\{T<k\}}) \\ &= \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} X_j (1 - 1_{\{T<j\}}) X_k (1 - 1_{\{T<k\}}). \end{aligned}$$

Παίρνοντας μέσες τιμές, μπορούμε κι εδώ να κάνουμε εναλλαγή με την άθροιση. Συμπεραίνουμε λοιπόν ότι

$$\mathbb{E}(S_T - x)^2 = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \mathbb{E}\left[X_j (1 - 1_{\{T<j\}}) X_k (1 - 1_{\{T<k\}}) \right]. \tag{10}$$

Στη συνέχεια υπολογίζουμε τους όρους του διπλού αθροίσματος. Θεωρούμε πρώτα εκείνους τους όρους που αντιστοιχούν σε τιμές των j και k με $j < k$. Η τυχαία μεταβλητή

$$X_j (1 - 1_{\{T<j\}})(1 - 1_{\{T<k\}})$$

εξαρτάται μόνο από τις $X_1, X_2, \ldots, X_{k-1}$, επομένως είναι ανεξάρτητη από την τυχαία μεταβλητή X_k. Αφού $\mathbb{E} X_k = \mu = 0$, βλέπουμε ότι

$$\mathbb{E}\left[X_j (1 - 1_{\{T<j\}}) X_k (1 - 1_{\{T<k\}}) \right] = \mathbb{E}\left[X_j (1 - 1_{\{T<j\}})(1 - 1_{\{T<k\}}) \right] \mathbb{E} X_k = 0.$$

Όμοια, οι όροι στο δεξιό μέλος της (10) για τους οποίους $j > k$ μηδενίζονται. Όταν $j = k$, παίρνουμε

$$\mathbb{E}\left[X_j^2 (1 - 1_{\{T<j\}})^2 \right].$$

Η τυχαία μεταβλητή $(1 - 1_{\{T<j\}})$ εξαρτάται μόνο από τις $X_1, X_2, \ldots, X_{j-1}$, επομένως είναι ανεξάρτητη από την X_j. Αφού αυτή η τυχαία μεταβλητή παίρνει μόνο τις τιμές 0 και 1, βλέπουμε ότι

$$(1 - 1_{\{T<j\}})^2 = 1 - 1_{\{T<j\}}.$$

Άρα,

$$\begin{aligned}\mathbb{E}\left[X_j^2(1-1_{\{T<j\}})^2\right] &= \mathbb{E}\left[X_j^2(1-1_{\{T<j\}})\right]\\ &= \mathbb{E}X_j^2\mathbb{E}(1-1_{\{T<j\}})\\ &= \mathrm{Var}X_j^2(1-P(T<j))\\ &= \sigma^2 P(T\geq j).\end{aligned}$$

Από την (10) και το Θεώρημα 5 του Κεφαλαίου 4 συμπεραίνουμε τώρα ότι

$$\mathbb{E}(S_T-x)^2=\sigma^2\sum_{j=1}^{\infty}P(T\geq j)=\sigma^2\mathbb{E}T,$$

το οποίο αποδεικνύει την (9), άρα και την (8).

9.2 Απλοί τυχαίοι περίπατοι

Σε αυτήν την παράγραφο θα υποθέσουμε ότι $a\leq x\leq b$, $a<b$, και οι a,b και x είναι ακέραιοι. Οι δύο ταυτότητες της προηγουμένης παραγράφου εφαρμόζονται πολύ πιό εύκολα αν είναι γνωστό ότι

(11) $$P(S_T=a\;\eta\;b)=P(S_T=a)+P(S_T=b)=1.$$

Αυτό βέβαια ισχύει σίγουρα αν

(12) $$P(X_k=-1,0\;\eta\;1)=1,$$

οπότε ο τυχαίος περίπατος $S_0,S_1,\ldots$ λέγεται *απλός τυχαίος περίπατος*. Η βασική ιδιότητα που διαχωρίζει τους απλούς τυχαίους περιπάτους από άλλους τυχαίους περιπάτους είναι ότι δεν «υπερπηδούν» τα ακέραια σημεία. Θέτουμε

$$\begin{aligned}p&=P\{X_k=1\},\\ q&=P\{X_k=-1\},\\ r&=P\{X_k=0\}.\end{aligned}$$

Τότε $p\geq 0$, $q\geq 0$, $r\geq 0$, και $p+q+r=1$. Η υπόθεση (2) μάς εξασφαλίζει ότι $r<1$. Από την (11) έπεται ότι

(13) $$\begin{aligned}\mathbb{E}S_T &= aP(S_T=a)+bP(S_T=b)\\ &= a(1-P(S_T=b))+bP(S_T=b).\end{aligned}$$

Στην περίπτωση των απλών τυχαίων περιπάτων μπορούμε να επιλύσουμε ως προς $P(S_T=a)$, $P(S_T=b)$, $\mathbb{E}S_T$ και $\mathbb{E}T$. Ας δούμε πρώτα την περίπτωση $p=q$. Τότε $\mu=0$ και η ταυτότητα του Wald (4) παίρνει τη μορφή $\mathbb{E}S_T=x$. Άρα, από την (13),

$$x=a(1-P(S_T=b))+bP(S_T=b).$$

Άμεση συνέπεια είναι η

$$P(S_T = b) = \frac{x-a}{b-a} \tag{14}$$

και η

$$P(S_T = a) = \frac{b-x}{b-a}. \tag{15}$$

Η ταυτότητα του Wald δεν δίνει καμμία πληροφορία για την $\mathbb{E}T$ όταν $\mu = 0$. Εφαρμόζοντας όμως την ταυτότητα (8) σε αυτήν την περίπτωση, βλέπουμε ότι $\sigma^2 = p + q = 1 - r$ και

$$\mathrm{Var} S_T = \sigma^2 \mathbb{E}T = (1-r)\mathbb{E}T.$$

Τώρα,

$$\begin{aligned}
\mathrm{Var} S_T &= \mathbb{E}S_T^2 - (\mathbb{E}S_T)^2 \\
&= b^2 P(S_T = b) + a^2 P(S_T = a) - x^2 \\
&= \frac{b^2(x-a) + a^2(b-x)}{b-a} - x^2 \\
&= (ax + bx - ab) - x^2 \\
&= (x-a)(b-x).
\end{aligned}$$

Επομένως, αν $p = q$,

$$\mathbb{E}T = \frac{(x-a)(b-x)}{1-r}. \tag{16}$$

Αν $r = 0$ και $p = q = 1/2$,

$$\mathbb{E}T = (x-a)(b-x). \tag{17}$$

Παράδειγμα 1. Δύο παίκτες που έχουν αρχικά κεφάλαια 5 και 10 δολλαρίων αντιστοίχως, συμφωνούν να παίξουν μία σειρά στοιχημάτων του ενός δολλαρίου μέχρι τη στιγμή που κάποιος από τους δύο θα χρεωκοπήσει. Υποθέτουμε ότι οι εκβάσεις των στοιχημάτων είναι ανεξάρτητες και ότι κάθε παίκτης έχει πιθανότητα επιτυχίας 1/2 σε κάθε στοίχημα. Βρείτε την πιθανότητα να χρεωκοπήσει ο παίκτης που είχε αρχικό κεφάλαιο 10 δολλαρίων. Βρείτε τον αναμενόμενο αριθμό των στοιχημάτων που θα απαιτηθούν.

Το πρόβλημα ταιριάζει στο σχήμα μας, αν συμβολίσουμε με S_n το κεφάλαιο του λιγότερο εύπορου παίκτη μετά από n στοιχήματα και αν επιλέξουμε $p = q = 1/2$, $x = 5$, $a = 0$ και $b = 15$. Η απάντηση στο πρώτο ερώτημα είναι

$$P(S_T = b) = \frac{5-0}{15-0} = \frac{1}{3}.$$

Η απάντηση στο δεύτερο ερώτημα είναι

$$\mathbb{E}T = (5-0)(15-5) = 50.$$

Ας υποθέσουμε τώρα ότι $p \neq q$. Για να αποφύγουμε τετριμμένες εξαιρέσεις, θα υποθέσουμε ακόμα ότι $p > 0$ και $q > 0$. Όταν $p \neq q$, δεν μπορούμε να χρησιμοποιήσουμε την ταυτότητα του Wald για να βρούμε την $P(S_T = b)$. Απαιτείται λοιπόν μία άλλη προσέγγιση.

Ορίζουμε $f(x)$ για ακεραίους x στο $[a, b]$, θέτοντας $f(x)$ την πιθανότητα να ισχύει $S_T = b$ δεδομένου ότι $S_0 = x$. Παρατηρούμε πρώτα ότι η f ικανοποιεί την εξίσωση διαφορών

$$f(x) = pf(x+1) + qf(x-1) + rf(x), \qquad a < x < b. \tag{18}$$

Αυτό είναι σωστό γιατί

$$\begin{aligned} f(x) &= P(S_T = b) \\ &= p \cdot P(S_T = b \mid X_1 = 1) + q \cdot P(S_T = b \mid X_1 = -1) \\ &\quad + r \cdot P(S_T = b \mid X_1 = 0) \end{aligned}$$

και

$$P(S_T = b \mid X_1 = i) = f(x+i), \qquad i = 1, -1, 0.$$

Εκτός από την (18) έχουμε τις προφανείς συνοριακές συνθήκες

$$f(a) = 0 \qquad \textit{και} \qquad f(b) = 1. \tag{19}$$

Από την (18) και την $1 - r = p + q$, βλέπουμε ότι

$$f(x+1) - f(x) = \frac{q}{p}[f(x) - f(x-1)], \qquad a < x < b. \tag{20}$$

Θέτουμε $c = f(a+1) = f(a+1) - f(a)$. Τότε η (20) μάς δίνει

$$f(x+1) - f(x) = c\left(\frac{p}{q}\right)^{x-a}, \qquad a \le x < b.$$

Χρησιμοποιώντας τον τύπο για το άθροισμα γεωμετρικής προόδου παίρνουμε

$$\begin{aligned} f(x) = f(x) - f(a) &= \sum_{y=a}^{x-1} (f(y+1) - f(y)) \\ &= \sum_{y=a}^{x-1} c\left(\frac{q}{p}\right)^{y-a} \\ &= c\frac{1 - (q/p)^{x-a}}{1 - (q/p)}, \qquad a \le x \le b. \end{aligned}$$

Από την ειδική περίπτωση $f(b) = 1$, έχουμε

$$c = \frac{1 - (q/p)}{1 - (q/p)^{b-a}}.$$

Αντικαθιστώντας στην παράσταση που δίνει την $f(x)$ συμπεραίνουμε ότι

$$f(x) = \frac{1-(q/p)^{x-a}}{1-(q/p)^{b-a}}.$$

Με άλλα λόγια, έχουμε δείξει ότι αν $p \neq q$ και $p > 0$, τότε

$$P(S_T = b) = \frac{1-(q/p)^{x-a}}{1-(q/p)^{b-a}}, \quad a \leq x \leq b. \tag{21}$$

Από την (21) έπεται άμεσα ότι, με τις ίδιες υποθέσεις,

$$P(S_T = a) = \frac{(q/p)^{x-a}-(q/p)^{b-a}}{1-(q/p)^{b-a}}, \quad a \leq x \leq b. \tag{22}$$

Από τις (13) και (21),

$$\mathbb{E}S_T = (b-a)\frac{1-(q/p)^{x-a}}{1-(q/p)^{b-a}} + a. \tag{23}$$

Αφού $\mu = p - q$, από την ταυτότητα του Wald βλέπουμε ότι

$$\mathbb{E}T = \left(\frac{b-a}{p-q}\right)\frac{1-(q/p)^{x-a}}{1-(q/p)^{b-a}} - \frac{x-a}{p-q}, \quad a \leq x \leq b. \tag{24}$$

Παράδειγμα 2. Τροποποιούμε το προηγούμενο παράδειγμα υποθέτοντας ότι ο πιό εύπορος παίκτης έχει μεγαλύτερες ικανότητες, οπότε η πιθανότητά του να κερδίσει κάθε στοίχημα είναι 0.6, ενώ η πιθανότητα του άλλου παίκτη είναι 0.4. Βρείτε την πιθανότητα να χρεωκοπήσει ο εύπορος παίκτης, το αναμενόμενο κέρδος αυτού του παίκτη, καθώς και το αναμενόμενο πλήθος στοιχημάτων.

Θα πάρουμε $p = 0.4$ και $q = 0.6$. Η πιθανότητα να χρεωκοπήσει ο εύπορος παίκτης είναι

$$P(S_T = 15) = \frac{1-(0.6/0.4)^5}{1-(0.6/0.4)^{15}} = 0.0151.$$

Για να βρούμε το αναμενόμενο κέρδος του ιδίου παίκτη, παρατηρούμε πρώτα ότι το αναμενόμενο κεφάλαιο του άλλου παίκτη όταν θα σταματήσουν το παιχνίδι είναι

$$\mathbb{E}S_T = 15P(S_T = 15) = 15(0.0151) = 0.23.$$

Άρα το αναμενόμενο κέρδος του εύπορου παίκτη ή η αναμενόμενη ζημία του άλλου παίκτη είναι 5.00-0.23=4.77, ένα καλό ποσοστό του αρχικού κεφαλαίου του φτωχότερου παίκτη. Ο αναμενόμενος αριθμός στοιχημάτων είναι

$$\mathbb{E}T = \frac{\mathbb{E}S_T - x}{\mu} = \frac{-4.77}{-0.2} \equiv 24.$$

Αφήνουμε το $b \to \infty$ στην (21). Αποδεικνύεται ότι το αριστερό μέλος της (21) συγκλίνει στην $P(S_n > a, \forall n \geq 0)$. Αν $q < p$, το δεξιό μέλος της (21) συγκλίνει στον $1-(q/p)^{x-a}$. Αν $q>p$, το δεξιό μέλος της (21) συγκλίνει στο 0. Αν $q=p$, το δεξιό μέλος της (14) συγκλίνει στο 0. Άρα, αν $a \le x = S_0$,

$$(25) \qquad P(S_n > a, \forall n \ge 0) = \begin{cases} 1 - \left(\frac{q}{p}\right)^{x-a} & , \text{αν } q < p, \\ 0 & , \text{αν } q \ge p. \end{cases}$$

Όμοια, αν $b \ge x = S_0$,

$$P(S_n < b, \forall n \ge 0) = \begin{cases} 1 - \left(\frac{p}{q}\right)^{b-x} & , \text{αν } p < q, \\ 0 & , \text{αν } p \ge q. \end{cases}$$

Παράδειγμα 3. Μία χαρτοπαικτική λέσχη διαθέτει κεφάλαιο εκατό χιλιάδων δολλαρίων. Ένας απείρως πλούσιος παίκτης προσπαθεί να καταστρέψει τη λέσχη. Του επιτρέπουν να προσπαθήσει, και αποφασίζει να ποντάρει 1000 δολλάρια κάθε φορά. Αν ο παίκτης έχει πιθανότητα 0.49 να κερδίσει και 0.51 να χάσει σε κάθε παιχνίδι, ποιά είναι η πιθανότητα να επιτύχει το στόχο του κάποια στιγμή;

Έστω S_n το κεφάλαιο της λέσχης (σε πολλαπλάσια των 1000 δολλαρίων) μετά από n παιχνίδια. Τότε, $p=0.51$, $q=0.49$, $x=100$ και $a=0$. Από την (25), η πιθανότητα να χρεωκοπήσει η λέσχη είναι

$$1 - P(S_n > 0, \forall n \ge 0) = \left(\frac{q}{p}\right)^{x-a} = \left(\frac{0.49}{0.51}\right)^{100} = 0.018.$$

Έστω A ένα υποσύνολο των ακεραίων (στις εφαρμογές, το A θα έχει 0,1 ή 2 στοιχεία). Αν $x \notin A$ και $y \notin A$, συμβολίζουμε με $P_A(x,y)$ την πιθανότητα ένας απλός τυχαίος περίπατος που ξεκινάει από το x να περάσει από το y κάποια χρονική στιγμή προτού να περάσει από το A. Αν $x \in A$ ή $y \in A$ θέτουμε $P_A(x,y)=0$. Οι πιθανότητες αυτές υπολογίζονται με τη βοήθεια των τύπων που συζητήσαμε σε αυτήν την παράγραφο.

Παράδειγμα 4. Έστω ότι $p=q$. Βρείτε την $P_{\{a,b\}}(y,y)$, όπου $a<y<b$.

Μετά από ένα βήμα, ο τυχαίος περίπατος βρίσκεται στο $y-1$, στο y ή το $y+1$ με αντίστοιχες πιθανότητες p, $1-2p$ και p. Από το $y-1$, η πιθανότητα επιστροφής στο y πριν το a είναι $(y-a-1)/(y-a)$. Από το $y+1$, η πιθανότητα επιστροφής στο y πριν το b είναι $(b-y-1)/(b-y)$. Άρα η πιθανότητα επιστροφής στο y πριν να περάσει από το a ή το b δίνεται από την

$$P_{\{a,b\}}(y,y) = p\frac{y-a-1}{y-a} + 1 - 2p + p\frac{b-y-1}{b-y}$$

δηλαδή

$$(26) \qquad P_{\{a,b\}}(y,y) = 1 - \frac{p(b-a)}{(y-a)(b-y)}.$$

Αν $x \notin A$ και $y \notin A$, συμβολίζουμε με $G_A(x,y)$ το αναμενόμενο πλήθος επισκέψεων στο y ενός τυχαίου περιπάτου που ξεκινάει από το x (για θετικές τιμές του n), πριν περάσει από το A. Θέτουμε $G_A(x,y) = 0$ αν $x \in A$ ή $y \in A$. Δεν είναι δύσκολο να δείξουμε ότι το πλήθος των επιστροφών από το y στο y πριν το πέρασμα από το A ακολουθεί γεωμετρική κατανομή με παράμετρο $p = 1 - P_A(y,y)$. Άρα, από το Παράδειγμα 3 του Κεφαλαίου 4,

$$G_A(y,y) = \frac{P_A(y,y)}{1 - P_A(y,y)}. \tag{27}$$

Αν $x \neq y$, τότε

$$G_A(x,y) = P_A(x,y)(1 + G_A(y,y)). \tag{28}$$

Για να έχουμε θετικό αριθμό επισκέψεων στο y πριν περάσουμε από το A, πρέπει βέβαια να φτάσουμε στο y πριν από το A. Το ενδεχόμενο αυτό έχει πιθανότητα $P_A(x,y)$. Αν πράγματι φτάσουμε στο y πριν από το A, τότε ο συνολικός αριθμός επισκέψεων στο y πριν να περάσουμε από το A είναι 1 συν το πλήθος των επιστροφών από το y στο y πριν να συναντήσουμε το A. Αυτό εξηγεί την (28). Από τις (27) και (28) συμπεραίνουμε ότι

$$G_A(x,y) = \frac{P_A(x,y)}{1 - P_A(y,y)} \tag{29}$$

για κάθε x και y.

Παράδειγμα 5. Ας επιστρέψουμε στο πρώτο παράδειγμα αυτής της παραγράφου. Βρείτε το αναμενόμενο πλήθος των στιγμών $n > 1$ κατά τις οποίες οι δύο παίκτες θα επανέλθουν στα αρχικά τους κεφάλαια πριν κάποιος από τους δύο να χρεωκοπήσει.

Υπενθυμίζουμε ότι σε αυτό το παράδειγμα $p = q = 1/2$, $a = 0$, $x = 5$ και $b = 15$. Η πιθανότητα επιστροφής στα αρχικά κεφάλαια πριν να χρεωκοπήσει κάποιος από τους παίκτες είναι, από την (26),

$$P_{\{0,15\}}(5,5) = 1 - \frac{(1/2)(15)}{5 \cdot 10} = 0.85.$$

Από την (27) έπεται ότι το αναμενόμενο πλήθος των στιγμών κατά τις οποίες οι δύο παίκτες θα επανέλθουν στα αρχικά τους κεφάλαια πριν κάποιος από τους δύο να χρεωκοπήσει είναι

$$\begin{aligned} G_{\{0,15\}}(5,5) &= \frac{P_{\{0,15\}}(5,5)}{1 - P_{\{0,15\}}(5,5)} \\ &= \frac{0.85}{0.15} = 5.67. \end{aligned}$$

9.3 Κατασκευή μιάς διαδικασίας Poisson

Στις υπόλοιπες παραγράφους αυτού του κεφαλαίου θα θεωρήσουμε ένα πιθανοθεωρητικό μοντέλο για την τυχαία κατανομή σωματιδίων στο χώρο ή γεγονότων στο χρόνο. Τέτοια μοντέλα βρίσκουν εφαρμογές σε πολλά και διαφορετικά πεδία. Για να δώσουμε ένα παράδειγμα, ας υποθέσουμε ότι έχουμε ένα κομμάτι ραδιενεργού υλικού και ότι παρατηρούμε τις χρονικές στιγμές στις οποίες συμβαίνουν διασπάσεις του. Υποθέτουμε ότι το πείραμα ξεκινάει τη χρονική στιγμή 0 και συμβολίζουμε με D_m τη χρονική στιγμή της m-στής διάσπασης. Όπως είδαμε στο Παράδειγμα 2 του Κεφαλαίου 1, οι νόμοι της φυσικής μάς λένε ότι οι χρονικές στιγμές D_m πρέπει να θεωρηθούν ως τυχαίες μεταβλητές. Η ακολουθία των σημείων $\{D_1, D_2, \ldots\}$ είναι ένα τυχαίο αριθμήσιμο υποσύνολο του $[0, \infty)$.

Για ένα άλλο παράδειγμα που περιγράφει ουσιαστικά τα ίδια φαινόμενα, ας θεωρήσουμε τις κλήσεις που δέχεται ένα τηλεφωνικό κέντρο. Συμβολίζουμε με D_m τη χρονική στιγμή κατά την οποία φτάνει στο κέντρο η m-στή κλήση. Δεν υπάρχει γνωστός τρόπος ακριβούς υπολογισμού των D_m, μπορούμε όμως να τις μελετήσουμε σαν τυχαίες μεταβλητές και να εξαγάγουμε χρήσιμα συμπεράσματα.

Θεωρούμε ένα πείραμα του εξής τύπου. Εμβαπτίζουμε ένα κομμάτι βαμβάκι σε έναν δοκιμαστικό σωλήνα στον οποίο αιωρούνται βακτηρίδια και κατόπιν αλείφουμε με αυτό την επιφάνεια ενός δίσκου ο οποίος παρέχει κατάλληλο υπόστρωμα για τον πολλαπλασιασμό των βακτηριδίων. Μετά από μερικές μέρες, γύρω από κάθε σημείο στο οποίο έπεσε ένα βακτηρίδιο εμφανίζεται μία ορατή αποικία βακτηριδίων. Οι θέσεις αυτών των περιοχών καθώς και το συνολικό τους πλήθος είναι τυχαία. Μία εικόνα αυτής της κατάστασης μάς δίνει το Σχήμα 2.

Η θέση των κηλίδων μπορεί να ιδωθεί σαν ένα τυχαίο υποσύνολο του συνόλου των σημείων του δίσκου.

Τα παραδείγματα των ραδιενεργών σωματιδίων και των βακτηριδίων μάς οδήγησαν στη μελέτη μιάς τυχαίας συλλογής σημείων ενός συγκεκριμένου υποσυνόλου S του Ευκλειδείου χώρου. Σε αυτά τα παραδείγματα, τόσο η θέση των «σωματιδίων» όσο και το συνολικό τους πλήθος είναι τυχαία. Σε κάθε τέτοια τυχαία συλλογή αντιστοιχούν διάφορες τυχαίες μεταβλητές όπως το συνολικό πλήθος N σωματιδίων στο S, το πλήθος N_B των σωματιδίων σε ένα προκαθορισμένο υποσύνολο $B \subset S$, και η απόσταση D_m ενός προκαθορισμένου σημείου του S από το m-στό πιό κοντινό σωματίδιο. Στο Σχήμα 3, τα σωματίδια συμβολίζονται με κουκίδες. Στο συγκεκριμένο σχήμα $N = N_S = 4$ και $N_B = 2$. Στο υπόλοιπο αυτού του κεφαλαίου, σκοπός μας είναι να μελετήσουμε τις κατανομές και τις από κοινού κατανομές τυχαίων μεταβλητών αυτού ακριβώς του τύπου.

Φυσικά, υπάρχουν πολλά διαφορετικά μοντέλα σχετικά με την τυχαία κατανομή σωματιδίων. Θα μελετήσουμε ένα από τα πιό στοιχειώδη και πιό σημαντικά μοντέλα του είδους, την διαδικασία Poisson. Η διαδικασία αυτή συνδέεται στενά με την ομοιόμορφη κατανομή σωματιδίων την οποία θα συζητήσουμε πρώτα.

Θεωρούμε λοιπόν ένα σύστημα στο οποίο το συνολικό πλήθος n των σωματιδίων είναι σταθερό. αλλά οι θέσεις των σωματιδίων στο S είναι τυχαίες. Το μοντέλο που μάς εξυπηρετεί είναι ένα μοντέλο στο οποίο αυτά τα n σωματίδια κατανέμονται ανεξάρτητα και ομοιόμορφα σε ένα σύνολο S με πεπερασμένο όγκο. Συμβολίζουμε τον όγκο ενός υποσυνόλου B του S με $|B|$. Τότε, κάθε σωματίδιο βρίσκεται στο B με πιθανότητα $p = |B|/|S|$. Επομένως, το πλήθος N_B των σωματιδίων που βρίσκονται στο B ακολουθεί διωνυμική κατανομή με παραμέτρους n και p. Πιό γενικά, έστω $B_1, B_2, \dots, B_k$ μία διαμέριση του S σε k ξένα υποσύνολα. Θέτουμε $p_j = |B_j|/|S|$. Τότε οι τυχαίες μεταβλητές $N_{B_1}, \dots, N_{B_k}$ ακολουθούν πολυωνυμική κατανομή με παραμέτρους n και $p_1, \dots, p_k$. Επομένως, αν $n_1, \dots, n_k$ είναι μη αρνητικοί ακέραιοι με άθροισμα n,

$$\begin{aligned} P(N_{B_1} = n_1, \dots, N_{B_k} = n_k) &= \frac{n!}{(n_1!)\dots(n_k!)} p_1^{n_1} \dots p_k^{n_k} \\ &= \frac{n!}{|S|^n} \prod_{j=1}^{k} \frac{|B_j|^{n_j}}{n_j!}. \end{aligned}$$

Η διαδικασία Poisson στο S είναι μία τροποποίηση της παραπάνω. Υποθέτουμε τώρα ότι το συνολικό πλήθος $N = N_S$ των σωματιδίων στο S είναι μία τυχαία μεταβλητή που ακολουθεί κατανομή Poisson με παράμετρο $\lambda|S|$. Επιπλέον, υποθέτουμε ότι, με δεδομένο ότι $N = n$, τα n σωματίδια κατανέμονται ανεξάρτητα και ομοιόμορφα στο S. Έστω $B_1, \dots, B_k$ όπως παραπάνω. Η υπόθεσή μας είναι ότι αν $n_1, \dots, n_k$ είναι k μη αρνητικοί ακέραιοι με άθροισμα n, τότε

$$P(N_{B_1} = n_1, \dots, N_{B_k} = n_k \mid N = n) = \frac{n!}{|S|^n} \prod_{j=1}^{k} \frac{|B_j|^{n_j}}{n_j!}.$$

Άρα,

$$P(N = n, N_{B_1} = n_1, \dots, N_{B_k} = n_k) = P(N = n) \frac{n!}{|S|^n} \prod_{j=1}^{k} \frac{|B_j|^{n_j}}{n_j!} \tag{30}$$

$$\begin{aligned} &= \frac{\lambda^n |S|^n}{n!} e^{-\lambda|S|} \frac{n!}{|S|^n} \prod_{j=1}^{k} \frac{|B_j|^{n_j}}{n_j!} \\ &= \lambda^n e^{-\lambda|S|} \prod_{j=1}^{k} \frac{|B_j|^{n_j}}{n_j!}. \end{aligned}$$

Αφού τα σύνολα B_j αποτελούν διαμέριση του S,

$$|S| = |B_1| + \ldots + |B_k|,$$

επομένως μπορούμε να γράψουμε το δεξιό μέλος της (30) στη μορφή

$$\prod_{j=1}^{k} \frac{(\lambda|B_j|)^{n_j}}{n_j!} e^{-\lambda|B_j|}.$$

Τώρα, το ενδεχόμενο $\{N = n, N_{B_1} = n_1, \ldots, N_{B_k} = n_k\}$ συμπίπτει με το ενδεχόμενο $\{N_{B_1} = n_1, \ldots, N_{B_k} = n_k\}$, γιατί $n = n_1 + \ldots + n_k$ και $N = N_{B_1} + \ldots + N_{B_k}$. Άρα,

$$P(N = n, N_{B_1} = n_1, \ldots, N_{B_k} = n_k) = P(N_{B_1} = n_1, \ldots, N_{B_k} = n_k).$$

Έχουμε λοιπόν αποδείξει τον εξής σημαντικό τύπο:

$$P(N_{B_1} = n_1, \ldots, N_{B_k} = n_k) = \prod_{j=1}^{k} \frac{(\lambda|B_j|)^{n_j}}{n_j!} e^{-\lambda|B_j|}. \tag{31}$$

Με άλλα λόγια, *οι τυχαίες μεταβλητές $N_{B_1}, \ldots, N_{B_k}$ είναι ανεξάρτητες και ακολουθούν κατανομές Poisson με παραμέτρους $\lambda|B_j|$ αντίστοιχα.*

Δεν πρέπει να αποτελεί έκπληξη το γεγονός ότι οι τυχαίες μεταβλητές N_{B_j}, $1 \leq j \leq k$ ακολουθούν κατανομή Poisson. Αυτό όμως που προκαλεί εντύπωση είναι το γεγονός ότι είναι ανεξάρτητες γιατί στην περίπτωση που το πλήθος των σωματιδίων είναι σταθερό οι αντίστοιχες ποσότητες είναι εξαρτημένες. Αυτή η ιδιότητα της ανεξαρτησίας καθιστά την διαδικασία Poisson πολύ εύχρηστη στις εφαρμογές.

Στα προηγούμενα παραδείγματα το συνολικό πλήθος των σωματιδίων, είτε σταθερό ή τυχαίο, ήταν πάντα πεπερασμένο, και τα σωματίδια ήταν κατανεμημένα σε σύνολα με πεπερασμένο συνολικό όγκο. Για κάποιους όμως λόγους, είναι θεωρητικά απλούστερο να θεωρήσουμε άπειρα το πλήθος σωματίδια κατανεμημένα σε ένα σύνολο με άπειρο όγκο. Θα μπορούσαμε για παράδειγμα να θέλουμε να κατανείμουμε σωματίδια στον $\mathbb{R}^r$ ή στο $[0, \infty)$, κλπ. Για να καλύψουμε αυτές τις περιπτώσεις, χρειαζόμαστε μία μικρή επέκταση του προηγούμενου μοντέλου.

Η βασική υπόθεση για μία διαδικασία Poisson σε ένα σύνολο S που έχει πεπερασμένο ή άπειρο όγκο είναι: για κάθε επιλογή ξένων υποσυνόλων $B_1, B_2, \ldots, B_k$ του S που έχουν πεπερασμένο όγκο, οι τυχαίες μεταβλητές $N_{B_1}, \ldots, N_{B_k}$ να είναι ανεξάρτητες και να ακολουθούν κατανομή Poisson με παραμέτρους $\lambda|B_1|, \ldots, \lambda|B_k|$ αντίστοιχα. Η σταθερά λ είναι η παράμετρος της διαδικασίας.

Έστω B ένα υποσύνολο του S που έχει πεπερασμένο όγκο. Από τον ορισμό της διαδικασίας Poisson προκύπτει ότι, αν $B_1, B_2, \ldots, B_k$ είναι ξένα υποσύνολα του B που η ένωσή τους είναι το B, και αν $n_1, n_2, \ldots, n_k$ είναι μη αρνητικοί ακέραιοι με άθροισμα n, τότε

$$P(N_{B_1} = n_1, \ldots, N_{B_k} = n_k \mid N = n) = \frac{n!}{|B|^n} \prod_{j=1}^{k} \frac{|B_j|^{n_j}}{n_j!}. \tag{32}$$

Για να επαληθεύσουμε την (32) παρατηρούμε ότι η ζητούμενη πιθανότητα είναι ίση με

$$\frac{P(N_{B_1} = n_1, \ldots, N_{B_k} = n_k)}{P(N_B = n)} = \frac{\prod_{j=1}^{k} e^{-\lambda|B_j|}(\lambda|B_j|)^{n_j}/n_j!}{e^{-\lambda|B|}(\lambda|B|)^n/n!},$$

το οποίο ισούται με το δεξιό μέλος της (32).

Ένας άλλος τρόπος για να βλέπουμε την (32) είναι ο εξής: Δεδομένου ότι υπάρχουν n σωματίδια στο B, η από κοινού κατανομή των τυχαίων μεταβλητών $N_{B_1}, \ldots, N_{B_k}$ συμπίπτει με αυτήν που παίρνουμε όταν κατανέμουμε n σωματίδια ανεξάρτητα και ομοιόμορφα στο B. Η παρατήρηση αυτή είναι πολύ χρήσιμη για την επίλυση κάποιων προβλημάτων στα οποία η διαδικασία Poisson εμφανίζεται σαν συνιστώσα ενός πιό πολύπλοκου συστήματος. Δεν θα εμβαθύνουμε περισσότερο σε αυτήν την πλευρά της διαδικασίας Poisson. (Δείτε όμως τις Ασκήσεις 21 και 31 για απλά παραδείγματα του τρόπου με τον οποίο χρησιμοποιείται).

9.4 Απόσταση από σωματίδια

Ας υποθέσουμε ότι έχουμε μία διαδικασία Poisson σε ένα υποσύνολο S του Ευκλειδείου χώρου. Αν το S έχει πεπερασμένο όγκο, τότε το πλήθος N των σωματιδίων στο S είναι πεπερασμένο. Συμβολίζουμε με $D_1 \leq D_2 \leq \ldots \leq D_N$ τις αποστάσεις αυτών των σωματιδίων από την αρχή των αξόνων, διατεταγμένες σε αύξουσα διάταξη. Αν το S έχει άπειρο όγκο, τότε το πλήθος των σωματιδίων στο S είναι άπειρο, και γράφουμε $D_1 \leq D_2 \leq \ldots \leq D_m \ldots$ για τις αποστάσεις των σωματιδίων από την αρχή των αξόνων, διατεταγμένες πάλι σε αύξουσα διάταξη. Μία τέτοια διάταξη είναι δυνατή γιατί για κάθε θετικό αριθμό r υπάρχουν πεπερασμένα το πλήθος σωματίδια που απέχουν απόσταση μικρότερη του r από την αρχή των αξόνων. Σε αυτήν την παράγραφο θα υπολογίσουμε την κατανομή της D_m για διάφορες επιλογές του συνόλου S.

Δίνουμε πρώτα ένα παράδειγμα στο οποίο αυτές οι αποστάσεις εμφανίζονται με φυσιολογικό τρόπο. Ας υποθέσουμε ότι τα αστέρια που βρίσκονται σε κάποιο σύνολο S του 3-διάστατου Ευκλειδείου χώρου κατανέμονται σύμφωνα με μία διαδικασία Poisson με παράμετρο λ στο S. Υποθέτουμε ακόμα ότι αυτά τα αστέρια είναι εξίσου φωτεινά. Το ποσόν του φωτός που φτάνει στην αρχή των αξόνων από κάποιο αστέρι είναι αντιστρόφως ανάλογο του τετραγώνου της απόστασης του αστεριού από την αρχή των αξόνων. Δηλαδή το ποσόν του φωτός που λαμβάνεται από ένα αστέρι σε απόσταση r από την αρχή των αξόνων είναι K/r^2 για κάποια θετική σταθερά K. Το ποσόν του φωτός που λαμβάνεται από το πλησιέστερο (άρα φαινομενικά λαμπρότερο) αστέρι είναι K/D_1^2. Η συνολική ποσότητα φωτός είναι

$$\sum_m \frac{K}{D_m^2}.$$

Χρησιμοποιώντας προχωρημένες τεχνικές της θεωρίας πιθανοτήτων μπορούμε να αποδείξουμε ότι αν S είναι ολόκληρος ο τριδιάστατος χώρος, τότε το άθροισμα αυτής της σειράς απειρίζεται με πιθανότητα ένα. Αυτό το γεγονός έχει ενδιαφέρουσες συνέπειες στην κοσμολογία.

Θα υπολογίσουμε τώρα την κατανομή της D_m, υποθέτοντας για απλότητα ότι το S έχει άπειρο όγκο. Θέτουμε $S_r = S \cap \{x : |x| \le r\}$ (δηλαδή, S_r είναι το σύνολο των σημείων του S που απέχουν απόσταση το πολύ r από την αρχή των αξόνων) και γράφουμε $\varphi(r)$ για τον όγκο του S_r. Το πλήθος N_{S_r} των σωματιδίων στο S_r ακολουθεί κατανομή Poisson με παράμετρο $\lambda\varphi(r)$. Το ενδεχόμενο $\{D_m \le r\}$ συμπίπτει με το ενδεχόμενο $\{N_{S_r} \ge m\}$. Από τις σχέσεις (39) και (40) του Κεφαλαίου 5,

$$\begin{aligned} (33) \qquad P(D_m \le r) &= P(N_{S_r} \ge m) \\ &= \int_0^{\lambda\varphi(r)} \frac{t^{m-1}e^{-t}}{(m-1)!}dt. \end{aligned}$$

Από την (33) έπεται ότι αν η $\varphi(r)$ είναι παραγωγίσιμη, τότε η D_m έχει συνάρτηση πυκνότητας την f_m που ορίζεται από την

$$(34) \qquad f_m(r) = \frac{\lambda^m \varphi(r)^{m-1}\varphi'(r)e^{-\lambda\varphi(r)}}{(m-1)!}, \qquad r > 0.$$

Αν η $\varphi(r)$ είναι γνησίως αύξουσα, τότε έχει μία συνεχή αντίστροφη συνάρτηση $\varphi^{-1}(r)$. Από την (33) βλέπουμε ότι η τυχαία μεταβλητή $\varphi(D_m)$ ακολουθεί κατανομή Γάμμα $\Gamma(m, \lambda)$.

Σε πολλές σημαντικές περιπτώσεις η $\varphi(r)$ είναι της μορφής $\varphi(r) = cr^d$, όπου c είναι μία θετική αριθμητική σταθερά (αυτό ισχύει, για παράδειγμα, αν $S = \mathbb{R}^d$ ή αν $S = [0, \infty)$). Σε αυτήν την περίπτωση η (34) γίνεται

$$(35) \qquad f_m(r) = \frac{d(c\lambda)^m}{(m-1)!}e^{-c\lambda r^d}, \qquad r > 0.$$

Στις ασκήσεις θα εξετάσουμε διάφορες ειδικές περιπτώσεις αυτής της σχέσης.

Θα χρησιμοποιήσουμε την σχέση (35) για να υπολογίσουμε τις ροπές $\mathbb{E}D_m^j$ σε αυτές τις περιπτώσεις. Έχουμε

$$\begin{aligned} \mathbb{E}D_m^j &= \int_0^\infty r^j f_m(r)dr \\ &= \int_0^\infty \frac{d(c\lambda)^m}{(m-1)!} r^{md+j-1} e^{-c\lambda r^d} dr. \end{aligned}$$

Για να συνεχίσουμε τον υπολογισμό, κάνουμε την αλλαγή μεταβλητής $s = \lambda c r^d$ και ολοκληρώνοντας παίρνουμε

$$\mathbb{E}D_m^j = \frac{(\lambda c)^{-j/d}\Gamma(m + (j/d))}{(m-1)!}.$$

9.5 Χρόνοι αναμονής

Ως τώρα βλέπαμε την διαδικασία Poisson σαν ένα μοντέλο για την κατανομή σωματιδίων στο χώρο. Αν σκεφτόμαστε το σύνολο $[0, \infty)$ σαν τον άξονα του χρόνου, τότε μπορούμε να βλέπουμε μία διαδικασία Poisson στο $[0, \infty)$ σαν την κατανομή των χρονικών στιγμών στις οποίες συμβαίνουν κάποια γεγονότα. Στην αρχή της Παραγράφου 3 αναφέραμε κάποια παραδείγματα στα οποία μπορεί να χρησιμοποιηθεί κατ' αυτόν τον τρόπο μία διαδικασία Poisson στο $[0, \infty)$.

Προκειμένου να αρχίσουμε να σκεφτόμαστε την διαδικασία Poisson στο $[0, \infty)$ σαν την κατανομή γεγονότων στο χρόνο και όχι σαν την κατανομή σωματιδίων στο $[0, \infty)$, πρέπει να μιλήσουμε με νέους όρους. Αντί να μιλάμε για «σωματίδια» θα μιλάμε για «γεγονότα», η δε απόσταση D_m από το m-στό σωματίδιο γίνεται πλέον η χρονική στιγμή κατά την οποία συμβαίνει το m-στό γεγονός.

Έστω $N(t) = N_{[0,t]}$ το πλήθος των γεγονότων που συμβαίνουν στο χρονικό διάστημα $[0, t]$. Τότε, η $N(t)$ είναι μία τυχαία μεταβλητή που ακολουθεί κατανομή Poisson με παράμετρο λt. Αν $0 \leq s \leq t$, τότε η $N(t) - N(s)$ παριστάνει το πλήθος των γεγονότων που συμβαίνουν στο χρονικό διάστημα $(s, t]$, και ακολουθεί κατανομή Poisson με παράμετρο $\lambda(t - s)$. Πιό γενικά, αν $0 \leq t_1 \leq t_2 \leq \ldots \leq t_n$, τότε οι $N(t_1)$, $N(t_2) - N(t_1), \ldots, N(t_n) - N(t_{n-1})$ είναι ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν κατανομές Poisson με παραμέτρους

$$\lambda t_1, \lambda(t_2 - t_1), \ldots, \lambda(t_n - t_{n-1})$$

αντίστοιχα. Όλα τα παραπάνω προκύπτουν άμεσα από τον ορισμό της διαδικασίας Poisson αν τον μεταφράσουμε στη γλώσσα του χρόνου.

Όπως ήδη αναφέραμε, D_m είναι η στιγμή κατά την οποία συμβαίνει το m-στό γεγονός. Από τα αποτελέσματα της Παραγράφου 9.4 γνωρίζουμε ότι η D_m ακολουθεί την κατανομή γάμμα $\Gamma(m, \lambda)$. Ειδικότερα, η D_1 ακολουθεί εκθετική κατανομή με παράμετρο λ. Θυμηθείτε από το Κεφάλαιο 6 ότι το άθροισμα m ανεξάρτητων και ισόνομων εκθετικών τυχαίων μεταβλητών έχει την κατανομή γάμμα $\Gamma(m, \lambda)$. Ορίζουμε τυχαίες μεταβλητές $W_1, W_2, \ldots, W_n, \ldots$ ως εξής: $W_1 = D_1$, $W_n = D_n - D_{n-1}$, $n \geq 2$. Προφανώς, $D_m = W_1 + \ldots + W_m$. Η συζήτηση που κάναμε μάς οδηγεί στο συμπέρασμα ότι οι $W_1, W_2, \ldots, W_m$ είναι ανεξάρτητες εκθετικά κατανεμημένες τυχαίες μεταβλητές με την ίδια παράμετρο λ. Αυτό ισχύει πραγματικά και δίνει μία πολύ ενδιαφέρουσα και χρήσιμη ιδιότητα της διαδικασίας Poisson στο $[0, \infty)$. Η τυχαία μεταβλητή W_m δεν είναι, φυσικά, τίποτε άλλο από τον χρόνο που μεσολαβεί ανάμεσα στο $(m-1)$-στό και το m-στό γεγονός. Για το λόγο αυτό, οι χρόνοι

$W_1, W_2, \ldots$ ονομάζονται *χρόνοι αναμονής* μεταξύ διαδοχικών γεγονότων της διαδικασίας Poisson.

Θεώρημα 1 *Έστω $W_1, W_2, \ldots, W_n, \ldots$ οι χρόνοι αναμονής μεταξύ διαδοχικών γεγονότων σε μία διαδικασία Poisson στο $[0, \infty)$ με παράμετρο λ. Τότε οι W_n, $n \geq 1$, είναι ανεξάρτητες εκθετικά κατανεμημένες τυχαίες μεταβλητές με τον ίδιο μέσο λ^{-1}.*

Απόδειξη: Έστω f_n η n-διάστατη πυκνότητα που δίνεται από την

$$f_n(t_1, \ldots, t_n) = \begin{cases} \lambda^n e^{-\lambda t_n} & , \text{αν } 0 \leq t_1 \leq t_2 \leq \ldots \leq t_n, \\ \\ 0 & , \text{αλλιώς.} \end{cases}$$

Από το Παράδειγμα 13 του Κεφαλαίου 6, βλέπουμε ότι το θεώρημα ισχύει αν και μόνο αν οι τυχαίες μεταβλητές $D_1, \ldots, D_n$ έχουν από κοινού πυκνότητα την f_n. Αυτό ισχύει αν $n = 1$, γιατί η D_1 ακολουθεί εκθετική κατανομή με παράμετρο λ.

Η αυστηρή απόδειξη στη γενική περίπτωση είναι πιό πολύπλοκη. Πριν δώσουμε αυτήν την απόδειξη θα εξηγήσουμε ευριστικά πώς μπορεί κανείς να βλέπει αυτό το αποτέλεσμα.

Θεωρούμε $0 = t_0 < t_1 < \ldots < t_n$ και επιλέγουμε $h > 0$ τόσο μικρό ώστε $t_{i-1} + h \leq t_i$, $i = 1, 2, \ldots, n$. Τότε (δείτε το Σχήμα 4)

$$\begin{aligned} (36) \qquad & P(t_i < D_i \leq t_i + h, 1 \leq i \leq n) \\ = \; & P(N(t_1) = 0, N(t_1 + h) - N(t_1) = 1, \ldots, \\ & N(t_n) - N(t_{n-1} + h) = 0, N(t_n + h) - N(t_n) \geq 1) \\ = \; & e^{-\lambda t_1} (\lambda h) e^{-\lambda h} \ldots e^{-\lambda (t_n - t_{n-1} - h)} \left[1 - e^{-\lambda h}\right] \\ = \; & \lambda^{n-1} h^{n-1} e^{-\lambda t_n} (1 - e^{-\lambda h}). \end{aligned}$$

Αν γνωρίζαμε ότι οι τυχαίες μεταβλητές $D_1, D_2, \ldots, D_n$ είχαν από κοινού πυκνότητα g_n συνεχή στο σημείο $(t_1, t_2, \ldots, t_n)$, θα μπορούσαμε να συμπεράνουμε ότι

$$P(t_i < D_i \leq t_i + h, 1 \leq i \leq n) = g_n(t_1, \ldots, t_n) h^n + e(h),$$

όπου $e(h)$ είναι μία συνάρτηση του h με την ιδιότητα

$$\lim_{h \to 0^+} \frac{e(h)}{h^n} = 0.$$

Τότε, από την (36) θα παίρναμε

$$\begin{aligned} g(t_1,\ldots,t_n) &= \lim_{h\to 0^+} h^{-n}P(t_i < D_i \le t_i + h) \\ &= \lim_{h\to 0^+} \lambda^{n-1}e^{-\lambda t_n}\frac{1-e^{-\lambda h}}{h} \\ &= \lambda^n e^{-\lambda t_n} \end{aligned}$$

δηλαδή το ζητούμενο.

Θα δώσουμε τώρα μία στοιχειώδη αλλά αυστηρή απόδειξη του Θεωρήματος 1. Η απόδειξη, χωρίς να είναι δύσκολη, είναι αρκετά μακροσκελής, οπότε ο αναγνώστης μπορεί αν θέλει να την παραλείψει.

Έστω F_n η συνάρτηση κατανομής που έχει πυκνότητα την f_n. Από τον ορισμό της f_n έπεται άμεσα ότι για $n \ge 2$

$$f_n(s_1,\ldots,s_n) = f_1(s_1)f_{n-1}(s_2-s_1,\ldots,s_n-s_1).$$

Ολοκληρώνοντας και τα δύο μέλη στο σύνολο $s_1 \le t_1,\ldots,s_n \le t_n$, βλέπουμε ότι

$$F_n(t_1,\ldots,t_n) = \int_0^{t_1} f_1(s_1)F_{n-1}(t_2-s_1,\ldots,t_n-s_1)ds_1. \tag{37}$$

Έστω G_n η από κοινού κατανομή των $D_1, D_2, \ldots, D_n$. Από το Παράδειγμα 10 του Κεφαλαίου 6 βλέπουμε ότι το θεώρημα ισχύει αν και μόνο αν οι τυχαίες μεταβλητές $D_1,\ldots,D_n$ έχουν από κοινού πυκνότητα την f_n, άρα από κοινού κατανομή την F_n. Επομένως, για να αποδείξουμε το θεώρημά μας πρέπει να δείξουμε ότι $F_n = G_n$.

Τώρα, η F_1 είναι απλώς η εκθετική κατανομή με παράμετρο λ. Όπως παρατηρήσαμε, η G_1, η κατανομή της D_1, είναι επίσης εκθετική με παράμετρο λ. Άρα, $F_1 = G_1$.

Ας υποθέσουμε πως μπορούμε να δείξουμε ότι για $n \ge 2$,

$$G_n(t_1,\ldots,t_n) = \int_0^{t_1} f_1(s_1)G_{n-1}(t_2-s_1,\ldots,t_n-s_1)ds_1. \tag{38}$$

Τότε, αφού $G_1 = F_1$, από τις (37) και (38) θα παίρναμε την $G_2 = F_2$. Χρησιμοποιώντας την $G_2 = F_2$ και εφαρμόζοντας ξανά τις (37) και (38) θα δείχναμε ότι $G_3 = F_3$ και ούτω καθεξής. Με άλλα λόγια, αν γνωρίζαμε ότι ισχύει η (38), το θεώρημά μας θα ήταν συνέπεια των (37) και (38) με επαγωγή. Για να αποδείξουμε την (38) μπορούμε να υποθέσουμε ότι $0 \le t_1 \le \ldots \le t_n$, γιατί αλλιώς και τα δύο μέλη της (38) είναι ίσα με μηδέν.

Για να ξεκινήσουμε την απόδειξη, παρατηρούμε ότι το ενδεχόμενο $\{D_i \le t_i\}$ συμπίπτει με το ενδεχόμενο $\{N(t_i) \ge i\}$, επομένως

$$\begin{aligned} \{D_i \le t_i, 1 \le i \le n\} &= \cap_{i=1}^n \{D_i \le t_i\} \\ &= \cap_{i=1}^n \{N(t_i) \ge i\} \\ &= \{N(t_i) \ge i, 1 \le i \le n\}. \end{aligned}$$

Μπορούμε λοιπόν να γράψουμε

$$G_n(t_1, t_2, \ldots, t_n) = P(N(t_i) \geq i, 1 \leq i \leq n).$$

Επομένως, η (38) είναι ισοδύναμη με την

$$P(N(t_i) \geq i, 1 \leq i \leq n) = \int_0^{t_i} f_1(s_1) P(N(t_i - s_i) \geq i - 1, 2 \leq i \leq n) ds_1. \tag{39}$$

Για να αποδείξουμε την (39), παρατηρούμε πρώτα ότι για κάθε $k \geq 1$

$$\frac{(\lambda t)^k}{k!} e^{-\lambda t} = \int_0^t \lambda e^{-\lambda s} \frac{[\lambda(t-s)]^{k-1}}{(k-1)!} e^{-\lambda(t-s)} ds. \tag{40}$$

Πράγματι,

$$\begin{aligned} \int_0^t \lambda e^{-\lambda s} \frac{[\lambda(t-s)]^{k-1}}{(k-1)!} e^{-\lambda(t-s)} ds &= \frac{e^{-\lambda t} \lambda^k}{(k-1)!} \int_0^t (t-s)^{k-1} ds \\ &= \frac{e^{-\lambda t} \lambda^k}{(k-1)!} \int_0^t s^{k-1} ds = \frac{(\lambda t)^k}{k!} e^{-\lambda t}. \end{aligned}$$

Έστω τώρα $0 \leq t_1 \leq t_2 \leq \ldots \leq t_n$ και $1 \leq k_1 \leq k_2 \leq \ldots \leq k_n$. Ισχυριζόμαστε ότι

$$\begin{aligned} & P(N(t_1) = k_1, \ldots, N(t_n) = k_n) \\ = & \int_0^{t_1} \lambda e^{-\lambda s} P(N(t_1 - s) = k_1 - 1, \ldots, N(t_n - s) = k_n - 1) ds. \end{aligned} \tag{41}$$

Για να το δούμε αυτό παρατηρούμε ότι από την (40)

$$\begin{aligned} & P(N(t_1) = k_1, \ldots, N(t_n) = k_n) \\ = & P(N(t_1) = k_1) \prod_{j=2}^n P(N(t_j) - N(t_{j-1}) = k_j - k_{j-1}) \\ = & \int_0^{t_1} \lambda e^{-\lambda s} \frac{[\lambda(t_1 - s)]^{k_j - k_{j-1}} e^{-\lambda(t_1 - s)}}{(k_1 - 1)!} ds \\ & \times \prod_{j=2}^n \frac{[\lambda(t_j - t_{j-1})]^{k_j - k_{j-1}} e^{-\lambda(t_j - t_{j-1})}}{(k_j - k_{j-1})!}. \end{aligned} \tag{42}$$

Από την άλλη πλευρά,

$$\begin{aligned} & \int_0^{t_1} \lambda e^{-\lambda s} P(N(t_1 - s) = k_1 - 1, \ldots, N(t_n - s) = k_n - 1) ds \\ = & \int_0^{t_1} \lambda e^{-\lambda s} P(N(t_1 - s) = k_1 - 1) \\ & \times \prod_{j=2}^n P(N(t_j - s) - N(t_{j-1} - s) = k_j - k_{j-1}) ds \end{aligned} \tag{43}$$

$$= \int_0^{t_1} \lambda e^{-\lambda s} \frac{[\lambda(t_1 - s)]^{k_1 - 1} e^{-\lambda(t_1 - s)}}{(k_1 - 1)!} \times \prod_{j=2}^{n} \frac{[\lambda(t_j - t_{j-1})]^{k_j - k_{j-1}} e^{-\lambda(t_j - t_{j-1})}}{(k_j - k_{j-1})!} ds.$$

Συγκρίνοντας το δεξιό μέλος της (42) με αυτό της (43), βλέπουμε ότι η (41) ισχύει. Η ισότητα (39) προκύπτει τώρα από την (41) αν αθροίσουμε και τα δύο μέλη της (41) πάνω από όλες τις τιμές των $k_1, \ldots, k_n$ για τις οποίες $k_1 \leq k_2 \leq \ldots \leq k_n$ και $k_1 \geq 1, k_2 \geq 2, \ldots, k_n \geq n$. □

9.6 Ασκήσεις

1 Έστω S_n ένας τυχαίος περίπατος με $\mu = 0$ και $S_0 = x$. Υποθέτουμε ότι

$$P(a - c \leq S_T \leq b + d) = 1,$$

όπου $a < x < b$, $c \geq 0$, και $d \geq 0$.
(α) Δείξτε ότι

$$(a - c)P(S_T \leq a) + bP(S_T \geq b) \leq x \leq aP(S_T \leq a) + (b + d)P(S_T \geq b).$$

(β) Δείξτε ότι

$$\frac{x - a}{b - a + d} \leq P(S_T \geq b) \leq \frac{x - a + c}{b - a + c}.$$

2 Ένας χαρτοπαίκτης βάζει μία σειρά στοιχημάτων του 1 δολλαρίου. Αποφασίζει ότι θα εγκαταλείψει το παιχνίδι όταν το κέρδος του φτάσει τα 25 δολλάρια ή η ζημία του φτάσει τα 50 δολλάρια. Ας υποθέσουμε ότι η πιθανότητά του να κερδίσει ή να χάσει κάθε στοίχημα είναι ίση με 1/2.

(α) Βρείτε την πιθανότητα να έχει χάσει 50 δολλάρια όταν θα εγκαταλείψει.

(β) Βρείτε την αναμενόμενη ζημία του.

(γ) Βρείτε το αναμενόμενο πλήθος στοιχημάτων που θα βάλει μέχρι να εγκαταλείψει.

3 Ας υποθέσουμε ότι ο παίκτης της Άσκησης 2 παίζει στη ρουλέτα, με πιθανότητα να κερδίσει ή να χάσει σε κάθε γύρο 9/19 και 10/19 αντίστοιχα. Απαντήστε στα ερωτήματα (α), (β) και (γ) της Άσκησης 2 χρησιμοποιώντας αυτές τις πιθανότητες.

4 Ένας παίκτης βάζει μία σειρά στοιχημάτων με πιθανότητα p να κερδίσει και πιθανότητα $q > p$ να χάσει κάθε στοίχημα. Αποφασίζει να παίξει μέχρι να κερδίσει M_1 δολλάρια ή να χάσει M_2 δολλάρια, όπου M_1 και M_2 είναι θετικοί ακέραιοι. Έχει την επιλογή να ποντάρει ένα ή μισό δολλάριο κάθε φορά. Δείξτε ότι είναι πιό πιθανό να κερδίσει M_1 δολλάρια πριν να χάσει M_2 δολλάρια αν ποντάρει ένα δολλάριο τη φορά παρά αν ποντάρει μισό δολλάριο τη φορά. Πώς γενικεύεται αυτό το συμπέρασμα;

5 Αποδείξτε την (14) λύνοντας κατάλληλη εξίσωση διαφορών.

6 Έστω S_n ένας απλός τυχαίος περίπατος με $p = q = 1/2$ και έστω $a < b$. Βρείτε τις $P_{\{a,b\}}(x, y)$ και $G_{\{a,b\}}(x, y)$ για $a < x < b$ και $a < y < b$.

7 Έστω S_n ένας απλός τυχαίος περίπατος με $p = q = 1/2$. Βρείτε τις $P_{\{0\}}(x, y)$ και $G_{\{0\}}(x, y)$ για $x > 0$ και $y > 0$.

8 Έστω S_n ένας απλός τυχαίος περίπατος με $0 < q < p$. Βρείτε τις $P_\emptyset(x, y)$ και $G_\emptyset(x, y)$.

9 Έστω S_n ένας απλός τυχαίος περίπατος με $0 < q < p$. Βρείτε τις $P_{\{0\}}(-1, y)$ και $G_{\{0\}}(-1, y)$ για $y < 0$.

10 Υποθέτουμε ότι σημεία κατανέμονται στον τριδιάστατο χώρο σύμφωνα με μία διαδικασία Poisson με παράμετρο λ. Κάθε σημείο της διαδικασίας παίρνεται σαν κέντρο μιάς σφαίρας με ακτίνα r. Έστω X το πλήθος των σφαιρών που περιέχουν την αρχή των αξόνων. Δείξτε ότι η X ακολουθεί κατανομή Poisson με μέσο $(4/3)\pi r^3$.

11 Ένα σημείο επιλέγεται τυχαία πάνω σε έναν κύκλο με κέντρο την αρχή των αξόνων και ακτίνα R. Το σημείο αυτό επιλέγεται σαν κέντρο ενός κύκλου ακτίνας X όπου X μία τυχαία μεταβλητή με πυκνότητα f. Βρείτε την πιθανότητα p αυτός ο κύκλος να περιέχει την αρχή των αξόνων.

12 Υποθέτουμε ότι n σημεία επιλέγονται ανεξάρτητα και ομοιόμορφα από έναν κύκλο με κέντρο την αρχή των αξόνων και ακτίνα R. Κάθε σημείο παίρνεται σαν το κέντρο ενός τυχαίου κύκλου που η ακτίνα του έχει πυκνότητα f. Βρείτε, συναρτήσει της p της Άσκησης 11, την πιθανότητα ακριβώς k από τους κύκλους να περιέχουν την αρχή των αξόνων.

13 Βρείτε την απάντηση στην Άσκηση 12 αν τα n σημεία αντικατασταθούν από τυχαίο αριθμό N σημείων που ακολουθεί κατανομή Poisson με μέσο πR^2.

14 Υποθέτουμε ότι N βώλοι κατανέμονται τυχαία σε r δοχεία, όπου το N ακολουθεί κατανομή Poisson με μέσο λ. Έστω Y το πλήθος των άδειων δοχείων. Δείξτε ότι η Y ακολουθεί διωνυμική κατανομή με παραμέτρους r και $p = e^{-\lambda/r}$. *Υπόδειξη:* Αν X_i είναι το πλήθος των βώλων στο δοχείο i, τότε οι $X_1, \ldots, X_r$ είναι ανεξάρτητες τυχαίες μεταβλητές Poisson με μέσο λ/r.

15 Χρησιμοποιώντας το αποτέλεσμα της Άσκησης 14 μπορούμε εύκολα να υπολογίσουμε την πιθανότητα $p_k(r, n)$ να υπάρχουν ακριβώς k άδεια δοχεία όταν μοιράζουμε τυχαία n βώλους σε r δοχεία. Για το σκοπό αυτό, παρατηρήστε πρώτα ότι

$$\begin{aligned} P(Y = k) &= \sum_{n=0}^{\infty} P(N = n) P(Y = k \mid N = n) \\ &= \sum_{n=0}^{\infty} e^{-\lambda} \frac{\lambda^n}{n!} p_k(r, n) \end{aligned}$$

και

$$P(Y = k) = \binom{r}{k} e^{-\lambda k/r} (1 - e^{-\lambda/r})^{r-k}.$$

Άρα,

$$\sum_{n=0}^{\infty} \frac{\lambda^n}{n!} p_k(r,n) = \binom{r}{k} e^{\lambda(r-k)/r}(1-e^{-\lambda/r})^{r-k}.$$

Εξισώνοντας τώρα τους συντελεστές του λ^n θα ξαναπάρετε την (16) του Κεφαλαίου 2.

16 Υποθέτουμε ότι οι χρόνοι των διαδοχικών βλαβών ενός μηχανήματος σχηματίζουν διαδικασία Poisson στο $[0,\infty)$ με παράμετρο λ.

(α) Ποιά είναι η πιθανότητα να παρατηρήσουμε τουλάχιστον μία βλάβη στη διάρκεια της περιόδου $(t, t+h)$, $h>0$;

(β) Ποιά είναι η δεσμευμένη πιθανότητα να παρατηρήσουμε τουλάχιστον μία βλάβη μέχρι την χρονική στιγμή $t+h$, δεδομένου ότι δεν παρουσιάστηκε βλάβη ως τη στιγμή t;

17 Έστω ότι έχουμε μία διαδικασία Poisson στο $[0,\infty)$ με παράμετρο λ. Συμβολίζουμε με Z_t την απόσταση από το t του πλησιέστερου από δεξιά σωματιδίου. Υπολογίστε τη συνάρτηση κατανομής της Z_t.

18 Έστω ότι έχουμε μία διαδικασία Poisson στο $[0,\infty)$ με παράμετρο λ. Συμβολίζουμε με Y_t την απόσταση από το t του πλησιέστερου από αριστερά σωματιδίου. Αν δεν υπάρχουν σωματίδια αριστερά του t, θέτουμε $Y_t = t$. Υπολογίστε τη συνάρτηση κατανομής της Y_t.

19 Για τις Z_t και Y_t των Ασκήσεων 17 και 18,

(α) δείξτε ότι οι Y_t και Z_t είναι ανεξάρτητες.

(β) υπολογίστε την κατανομή της $Z_t + Y_t$.

20 Σωματίδια φτάνουν σε έναν μετρητή σύμφωνα με μία διαδικασία Poisson παραμέτρου λ. Κάθε σωματίδιο προκαλεί μία δόνηση μοναδιαίας διάρκειας. Ο μετρητής καταγράφει το σωματίδιο αν και μόνο αν φτάνει κάποια στιγμή κατά την οποία δεν υπάρχουν δονήσεις. Βρείτε την πιθανότητα να καταγραφεί κάποιο σωματίδιο στο χρονικό διάστημα μεταξύ t και $t+1$. Υποθέστε ότι $t \geq 1$.

21 Θεωρούμε μία διαδικασία Poisson στο $[0,\infty)$ με παράμετρο λ και μία τυχαία μεταβλητή T που είναι ανεξάρτητη από την διαδικασία. Υποθέτουμε ότι η T ακολουθεί εκθετική κατανομή με παράμετρο ν. Έστω N_T το πλήθος των σωματιδίων στο διάστημα $[0,T]$. Υπολογίστε την διακριτή πυκνότητα της N_T.

22 Να επαναλάβετε την Άσκηση 21 αν η T ακολουθεί ομοιόμορφη κατανομή στο $[0,a]$, $a>0$.

23 Θεωρούμε δύο ανεξάρτητες διαδικασίες Poisson στο $[0,\infty)$ με παραμέτρους λ_1 και λ_2 αντίστοιχα. Ποιά είναι η πιθανότητα να έχουμε γεγονός νωρίτερα στην πρώτη διαδικασία από ότι στην δεύτερη;

24 Υποθέτουμε ότι n σωματίδια κατανέμονται ανεξάρτητα και ομοιόμορφα σε ένα δίσκο ακτίνας r. Έστω D_1 η απόσταση του κέντρου του δίσκου από το πλησιέστερο σωματίδιο. Υπολογίστε την πυκνότητα της D_1.

25 Για την D_1 στην Άσκηση 24, υπολογίστε τις ροπές της D_1. *Υπόδειξη:* Με αλλαγή μεταβλητής θα οδηγηθείτε σε ένα ολοκλήρωμα Βήτα.

26 Θεωρούμε μία διαδικασία Poisson στον $\mathbb{R}^r$ με παράμετρο λ. Για κάθε σύνολο A με πεπερασμένο όγκο, γράφουμε N_A για το πλήθος των σωματιδίων στο A.

(α) Υπολογίστε την $\mathbb{E}N_A^2$.

(β) Αν A, B είναι δύο σύνολα με πεπερασμένο όγκο, υπολογίστε την $\mathbb{E}(N_A N_B)$.

27 Έστω $A_1, \ldots, A_n$ ξένα σύνολα με πεπερασμένο όγκο, και όμοια, έστω $B_1, \ldots, B_n$ ξένα σύνολα με πεπερασμένο όγκο. Αν $\alpha_1, \ldots, \alpha_n$ και $\beta_1, \ldots, \beta_n$ είναι πραγματικοί αριθμοί, θέτουμε

$$f(x) = \sum_{i=1}^{n} \alpha_i 1_{A_i}(x)$$

και

$$g(x) = \sum_{i=1}^{n} \beta_i 1_{B_i}(x).$$

Δείξτε ότι για μία διαδικασία Poisson με παράμετρο λ,

$$\begin{aligned}\mathbb{E}\left(\sum_{i=1}^{n} \alpha_i N_{A_i}\right)\left(\sum_{i=1}^{n} \beta_i N_{B_i}\right) &= \lambda^2 \left(\int_{\mathbb{R}^r} f(x)dx\right)\left(\int_{\mathbb{R}^r} g(x)dx\right) \\ &\quad + \lambda \int_{\mathbb{R}^r} f(x)g(x)dx.\end{aligned}$$

28 Στην Άσκηση 27 δείξτε ότι

$$\mathrm{Var}\left(\sum_{i=1}^{n} \alpha_i N_{A_i}\right) = \lambda \int_{\mathbb{R}^r} f^2(x)dx.$$

29 Θεωρούμε μία διαδικασία Poisson με παράμετρο λ στον $\mathbb{R}^3$, και συμβολίζουμε με D_m την απόσταση από την αρχή των αξόνων προς το m-στό πλησιέστερο σωματίδιο.

(α) Βρείτε την πυκνότητα της D_m.

(β) Βρείτε την πυκνότητα της D_m^3.

30 Έστω ότι έχουμε μία διαδικασία Poisson με παράμετρο λ στο άνω ημιεπίπεδο του $\mathbb{R}^2$, δηλαδή, η διαδικασία Poisson είναι στο υποσύνολο $\{(x, y) : y > 0\}$ του $\mathbb{R}^2$.

(α) Ποιά είναι η πυκνότητα της απόστασης D_m από το 0 προς το m-στό πλησιέστερο σωματίδιο;

(β) Βρείτε τις πρώτες δύο ροπές της D_m.

31 Θεωρούμε το εξής σύστημα. Οι χρόνοι κατά τους οποίους σωματίδια φτάνουν στο σύστημα συνιστούν διαδικασία Poisson στο $[0, \infty)$ με παράμετρο λ. Κατόπιν, κάθε σωματίδιο έχει κάποιο χρόνο ζωής ανεξάρτητο από τους χρόνους αφίξεως των σωματιδίων στην διαδικασία και ανεξάρτητο από τους χρόνους ζωής των άλλων σωματιδίων. Υποθέτουμε ότι οι χρόνοι ζωής των σωματιδίων ακολουθούν εκθετική κατανομή με την ίδια παράμετρο μ. Έστω $M(t)$ το πλήθος των σωματιδίων που είναι ζωντανά κατά την χρονική στιγμή t. Υπολογίστε την κατανομή της $M(t)$ ακολουθώντας τα παρακάτω βήματα.

(α) Ας υποθέσουμε ότι ένα σωματίδιο φτάνει σύμφωνα με την ομοιόμορφη κατανομή στο $[0, t]$ και ζεί για ένα τυχαίο χρονικό διάστημα που είναι εκθετικά κατανεμημένο με παράμετρο μ. Βρείτε την πιθανότητα p_t το σωματίδιο να είναι ζωντανό την χρονική στιγμή t.

(β) Χρησιμοποιώντας το γεγονός ότι, δεδομένου ότι $N(t) = n$, τα σωματίδια είναι ανεξάρτητα και ομοιόμορφα κατανεμημένα στο $[0, t]$, δείξτε ότι

$$P(M(t) = k \mid N(t) = n) = \binom{n}{k} p_t^k (1 - p_t)^{n-k}.$$

(γ) Δείξτε τώρα ότι η $M(t)$ ακολουθεί κατανομή Poisson με παράμετρο $\lambda t p_t$.