

ΔΗΜΗΤΡΗΣ ΧΕΛΙΩΤΗΣ
Αναπληρωτής καθηγητής
Τμήμα Μαθηματικών
Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών

Ένα δεύτερο μάθημα στις πιθανότητες



**Ελληνικά Ακαδημαϊκά Ηλεκτρονικά
Συγγράμματα και Βοηθήματα**
www.kallipos.gr

Ένα δεύτερο μάθημα στις πιθανότητες

Συγγραφή

Δημήτρης Χελιώτης

Κριτικός αναγνώστης

Μιχάλης Λουλάκης

Συντελεστές έκδοσης

Γλωσσική Επιμέλεια: Θεόφιλος Τραμπούλης



Το παρόν έργο αδειοδοτείται υπό τους όρους της άδειας Creative Commons Αναφορά Δημιουργού - Μη Εμπορική Χρήση - Όχι Παράγωγα Έργα 3.0. Για να δείτε ένα αντίγραφο της άδειας αυτής επισκεφτείτε τον ιστότοπο

<https://creativecommons.org/licenses/by-nc-sa/3.0/gr/>

ΣΥΝΔΕΣΜΟΣ ΕΛΛΗΝΙΚΩΝ ΑΚΑΔΗΜΑΪΚΩΝ ΒΙΒΛΙΟΘΗΚΩΝ

Εθνικό Μετσόβιο Πολυτεχνείο

Ηρώων Πολυτεχνείου 9, 15780 Ζωγράφου

www.kallipos.gr

ISBN:978-960-603-296-7

Περιεχόμενα

Πρόλογος	vi
Σύμβολα	ix
1 σ-άλγεβρες	1
1.1 σ-άλγεβρες	1
1.2 Παραγόμενη σ-άλγεβρα	2
1.3 Τα σύνολα Borel	3
1.4 \liminf και \limsup ακολουθίας συνόλων	5
2 Μέτρα	8
2.1 Μέτρα σε μετρήσιμο χώρο	8
2.2 Το μέτρο Lebesgue	10
2.3 Διακριτά μέτρα πιθανότητας	11
2.4 Στήριγμα μέτρου	13
3 Ισότητα πεπερασμένων μέτρων	16
3.1 Κλάσεις Dynkin	16
3.2 Το θεώρημα π-λ	17
4 Μετρήσιμες συναρτήσεις	19
4.1 Μετρήσιμες συναρτήσεις	19
4.2 Σ-άλγεβρα παραγόμενη από συναρτήσεις	23
5 Το ολοκλήρωμα Lebesgue	27
5.1 Ολοκλήρωμα Lebesgue. Ορισμός	27
5.2 Ειδικές περιπτώσεις	29
5.3 Η οπτική του ολοκληρώματος Lebesgue	30
5.4 Ιδιότητες του ολοκληρώματος	31
5.5 Κατασκευή διακριτών τυχαίων μεταβλητών	35
5.6 Οι χώροι \mathcal{L}^p με $p \in [1, \infty)$	36
5.7 Ο χώρος \mathcal{L}^∞	37
5.8 Τα βασικά οριακά θεωρήματα	38
6 Κατανομή τυχαίας μεταβλητής και ολοκλήρωση	45
6.1 Ο τύπος αλλαγής μεταβλητής	45
6.2 Κατανομές στο \mathbb{R} με πυκνότητα	48
6.3 Διακριτές κατανομές	50

ΠΕΡΙΕΧΟΜΕΝΑ

7 Κατανομές στο \mathbb{R}	53
7.1 Η συνάρτηση κατανομής μέτρου	53
7.2 Η συνάρτηση κατανομής τυχαίας μεταβλητής	56
7.3 Η συνάρτηση ποσοστημορίων*	57
7.4 Είδη κατανομών στο \mathbb{R}	58
8 Τρόποι σύγκλισης ακολουθιών τυχαίων μεταβλητών	61
9 Μέτρα γινόμενο	67
9.1 Γινόμενο χώρων μέτρου. Πεπερασμένο πλήθος	67
9.2 Ολοκλήρωση σε χώρο γινόμενο	68
9.3 Γινόμενο χώρων πιθανότητας. Αυθαίρετο πλήθος	70
10 Ανεξαρτησία	73
10.1 Ανεξαρτησία για οικογένειες συνόλων και τυχαίες μεταβλητές	73
10.2 Ανεξαρτησία και ομαδοποίηση	77
10.3 Ανεξαρτησία=Μέτρο γινόμενο	77
10.4 Κατασκευή ανεξάρτητων τυχαίων μεταβλητών με δεδομένη κατανομή	78
11 Τα λήμματα Borel-Cantelli και ο νόμος 0-1 του Kolmogorov	82
11.1 Τα λήμματα Borel-Cantelli	82
11.2 Ο νόμος 0-1 του Kolmogorov*	86
12 Ο Ισχυρός Νόμος των Μεγάλων Αριθμών	92
12.1 Το θεώρημα	92
12.2 Δύο εφαρμογές*	94
13 Χαρακτηριστικές συναρτήσεις	98
13.1 Ολοκλήρωμα μιγαδικής συνάρτησης	98
13.2 Χαρακτηριστικές συναρτήσεις	98
13.3 Ο τύπος αντιστροφής	102
13.4 Ροπογεννήτριες	104
13.5 Χαρακτηριστικές συναρτήσεις μέσω ροπογεννητριών*	106
13.6 Άθροισμα ανεξάρτητων τυχαίων μεταβλητών	108
14 Σύγκλιση κατά κατανομή	111
14.1 Σύγκλιση κατά κατανομή	111
14.2 Η αναπαράσταση Skorokhod	116
14.3 Σφιχτότητα και υπακολουθιακά όρια	117
15 Σύγκλιση κατά κατανομή και χαρακτηριστικές συναρτήσεις	122
15.1 Το Θεώρημα Συνέχειας του Lévy	122
15.2 Εφαρμογές	124
16 Το Κεντρικό Οριακό Θεώρημα	128
16.1 Προετοιμασία	128
16.2 Το Κεντρικό Οριακό Θεώρημα	129
16.3 Εφαρμογές	130

ΠΕΡΙΕΧΟΜΕΝΑ

16.4	Ισχύς των προσεγγίσεων*	131
16.5	Η σύγκλιση στο κεντρικό οριακό θεώρημα*	132
17	Μεγάλες αποκλίσεις*	134
17.1	Η έννοια της μεγάλης απόκλισης	134
17.2	Γιατί οι μεγάλες αποκλίσεις είναι σημαντικές	134
17.3	Η αρχή μεγάλων αποκλίσεων	135
17.4	Το Θεώρημα Cramer	136
	Παραρτήματα	142
A'	Αναλυτικά αποτελέσματα	142
A'.1	Αθροίσματα	142
A'.2	Όρια γινομένων	142
A'.3	Η συνάρτηση Γάμμα	144
A'.4	Παραγωγή ολοκληρώματος που εξαρτάται από παράμετρο	145
B'	Τεχνικές αποδείξεις	147
	Υποδείξεις για επιλεγμένες ασκήσεις	155
	Βιβλιογραφία	172
	Ευρετήριο ελληνικών όρων	173
	Ευρετήριο ξενόγλωσσων όρων	175
	Μετάφραση ορολογίας	176

Πρόλογος

Οι σημειώσεις αυτές είναι κατάλληλες για ένα προχωρημένο προπτυχιακό¹ μάθημα σε τμήμα Μαθηματικών πάνω στις Πιθανότητες. Προσπαιτούν καλή γνώση των στοιχειωδών Πιθανοτήτων (που καλύπτονται στο πρώτο προπτυχιακό μάθημα Πιθανοτήτων) και του Απειροστικού Λογισμού (κυρίως μιας μεταβλητής) και κάποια τριβή με την Πραγματική Ανάλυση (μετρικοί χώροι και συνεχείς συναρτήσεις σε μετρικούς χώρους).

Δύο είναι οι στόχοι των σημειώσεων. Πρώτα να δούμε την ορολογία της μετροθεωρητικής θεωρίας πιθανοτήτων (Κεφάλαια 1-10, 13, 14) και έπειτα να μελετήσουμε την οριακή συμπεριφορά ακολουθιών τυχαίων μεταβλητών. Ασχολούμαστε με τα δύο σημαντικότερα είδη σύγκλισης της θεωρίας πιθανοτήτων, τη σχεδόν βέβαιη και την κατά κατανομή. Έτσι, καλύπτουμε τη βασική τεχνική για την απόδειξη αποτελεσμάτων για καθεμία από αυτές (Κεφάλαιο 11 για την πρώτη, Κεφάλαιο 15 για τη δεύτερη) και τέλος βλέπουμε τα δύο πιο αντιπροσωπευτικά θεώρηματα που αυτές εμφανίζονται: τον Ισχυρό Νόμο των Μεγάλων Αριθμών και το Κεντρικό Οριακό Θεώρημα.

Αυτή η δουλειά βάζει σε αυστηρό πλαίσιο την ύλη που καλύπτεται σε μαθήματα στοιχειωδών Πιθανοτήτων. Πολλά αποτελέσματα τα οποία είχαμε μάθει να χρησιμοποιούμε σε εκείνα τα μαθήματα χωρίς όμως αποδείξεις, εδώ τα αποδεικνύουμε.

Στο κομμάτι των σημειώσεων που αφορούν τη Θεωρία Μέτρου (ιδιαίτερα στα Κεφάλαια 4 και 5) δεν δίνουμε αποδείξεις για όλα τα αποτελέσματα που διατυπώνουμε γιατί είναι πέρα από τους στόχους μας. Από τα μετέπειτα κεφάλαια, δεν δίνουμε απόδειξη για το θεώρημα Prokhorov (Θεώρημα 14.24). Για κάποια άλλα αποτελέσματα, η απόδειξη παρατίθεται στο Παράρτημα Β' γιατί, μολονότι δεν είναι τόσο σημαντική για την κατανόηση της θεωρίας, είναι, σε δεύτερη ανάγνωση, προσιτή και ωφέλιμη.

Σε ένα εξάμηνο (26 διαλέξεις των 45+45 λεπτών) είναι δυνατόν να καλυφθούν όλες οι σημειώσεις εκτός από τις παραγράφους που σημειώνονται με αστερίσκο. Τα Κεφάλαια 1-5, 8, 9 συνήθως διδάσκονται σε μαθήματα Θεωρίας Μέτρου, και με μεγαλύτερη λεπτομέρεια. Οι αναγνώστες που έχουν ήδη πάρει ανάλογο μάθημα μπορούν απλώς να ρίξουν μια γρήγορη ματιά σε αυτά τα κεφάλαια για να δουν την ορολογία και τον συμβολισμό που θα χρησιμοποιηθούν στη συνέχεια. Ωστόσο, συνιστάται θερμά η επίλυση ασκήσεων από τα κεφάλαια αυτά. Για τη μείωση του χρόνου που αφιερώνεται στη Θεωρία Μέτρου συνιστάται η παράλειψη του Κεφαλαίου 3. Το θεώρημα π-λ, που παρουσιάζεται εκεί, χρησιμοποιείται μόνο στις αποδείξεις των Θεωρημάτων 7.6, 10.6, 10.10, 11.12.

Η πλειονότητα των ασκήσεων είναι ασκήσεις τριβής και εξοικείωσης με τις έννοιες. Κάποιες είναι προεκτάσεις της θεωρίας. Όσες έχουν αστερίσκο είναι λίγο δυσκολότερες.

Βάση για αυτές τις σημειώσεις υπήρξαν οι σημειώσεις παραδόσεων του μαθήματος Πιθανότητες II όπως διδάχθηκε το εαρινό εξάμηνο του ακαδημαϊκού έτους 2012-2013. Ακολουθήσαμε τότε κυρίως το βιβλίο Probability Essentials των Jacod-Protter. Τα υπόλοιπα βιβλία που τις επηρέασαν αναγράφονται στη βιβλιογραφία. Ευχαριστώ τον κ. Παύλο Τσατσούλη, ο οποίος έγραψε το πρώτο προσχέδιο των σημειώσεων σε Latex.

Επίσης ευχαριστώ: Τον συνάδελφο Αντώνη Τσολομύτη για τη βοήθεια σε θέματα Latex και τη

¹Η και μεταπτυχιακό μάθημα σε master Χρηματοοικονομικών Μαθηματικών ή Στατιστικής. Δεν καλύπτουν κάποια θέματα που εξετάζονται σε μεταπτυχιακά μαθήματα τμημάτων Μαθηματικών όπως: Ομοιόμορφη ολοκληρωσιμότητα, ο ασθενής νόμος των μεγάλων αριθμών και το κεντρικό οριακό θεώρημα για τριγωνικούς πίνακες. Σύγκλιση στην κατανομή Poisson. Ευσταθείς κατανομές.

διαμόρφωση της εμφάνισης του κειμένου. Δύο ανώνυμους αξιολογητές που στο πλαίσιο του προγράμματος «Κάλλιπος» έκαναν χρήσιμες παρατηρήσεις σε προηγούμενη έκδοση των σημειώσεων. Τον συνάδελφο Σάμη Τρέβεζα για διορθώσεις και σχόλια. Τον συνάδελφο Μιχάλη Λουλάκη που μελέτησε προσεκτικά τις σημειώσεις και έκανε πολλές διορθώσεις και προτάσεις για προσθήκες/αλλαγές και ασκήσεις που συνέβαλαν στη σημαντική βελτίωση των σημειώσεων.

Τα θεμελιώδη θεωρήματα

Το πρώτο μισό του 20ου αιώνα, το βασικότερο αντικείμενο μελέτης της Θεωρίας Πιθανοτήτων ήταν το άθροισμα $S_n := X_1 + X_2 + \dots + X_n$ ανεξάρτητων και ισόνομων τυχαίων μεταβλητών και το ενδιαφέρον επικεντρώθηκε στη συμπεριφορά του S_n για μεγάλα n για τις διάφορες επιλογές που έχει η κοινή κατανομή των X_i . Ας έχουμε για τα επόμενα στο μυαλό μας την ειδική περίπτωση που αυτή η κατανομή είναι η ομοιόμορφη στο δισύνολο $\{-1, 1\}$. Τότε το S_n είναι το συνολικό κέρδος μας μετά από n ανεξάρτητα παιχνίδια σε καθένα από τα οποία κερδίζουμε 1 με πιθανότητα 1/2 και χάνουμε 1 πάλι με πιθανότητα 1/2. Τι μπορούμε να πούμε για αυτό το κέρδος;

Τα θεμελιώδη θεωρήματα που αποδείχθηκαν για το S_n είναι τα εξής.

(α) Ισχυρός Νόμος των Μεγάλων Αριθμών (Borel, 1909· Kolmogorov, 1933). Υποθέτουμε ότι $\mathbf{E}(X_1) = 0$. Με πιθανότητα 1,

$$\lim_{n \rightarrow \infty} \frac{S_n}{n} = 0.$$

(β) Κεντρικό Οριακό Θεώρημα (Lindeberg, Lévy 1922). Υποθέτουμε ότι $\mathbf{E}(X_1) = 0$, $\text{Var}(X_1) = \sigma^2 \in (0, \infty)$. Για $n \rightarrow \infty$,

$$\frac{S_n}{\sqrt{n}} \Rightarrow Y \sim N(0, \sigma^2),$$

όπου \Rightarrow δηλώνει τη σύγκλιση κατά κατανομή και $Y \sim N(0, \sigma^2)$ σημαίνει ότι η τυχαία μεταβλητή Y ακολουθεί την κατανομή $N(0, \sigma^2)$.

(γ) Νόμος Επαναλαμβανόμενου Λογαρίθμου (Khinchine, 1923· Kolmogorov, 1929· Hartman-Wintner, 1941). Υποθέτουμε ότι $\mathbf{E}(X_1) = 0$ και $\text{Var}(X_1) = \sigma^2 \in (0, \infty)$. Με πιθανότητα 1,

$$\liminf_{n \rightarrow \infty} \frac{S_n}{\sigma \sqrt{2n \log \log n}} = -1, \quad (1)$$

$$\limsup_{n \rightarrow \infty} \frac{S_n}{\sigma \sqrt{2n \log \log n}} = 1. \quad (2)$$

(δ) Μεγάλες αποκλίσεις. Θεώρημα Cramer (1938). Υποθέτουμε ότι $\mathbf{E}(X_1) = 0$. Υπάρχει συνάρτηση $I : \mathcal{B}(\mathbb{R}) \rightarrow [0, \infty]$ ώστε, για κάθε $A \subset \mathbb{R}$ σύνολο Borel (που επιπλέον ικανοποιεί μια τεχνική συνθήκη), να ισχύει

$$\mathbf{P}\left(\frac{S_n}{n} \in A\right) \approx e^{-nI(A)}.$$

Εδώ, $a_n \approx b_n$ σημαίνει ότι $\lim_{n \rightarrow \infty} \log a_n / \log b_n = 1$, δηλαδή οι a_n, b_n είναι ισοδύναμες σε λογαριθμική κλίμακα. Μάλιστα, αν υποθέσουμε ότι η ροπογεννήτρια της X_1 είναι πεπερασμένη σε περιοχή του 0, τότε η συνάρτηση I είναι τέτοια ώστε $I(A) > 0$ όταν $0 \notin \bar{A}$.

Αν αντί της υπόθεσης $\mathbf{E}(X_1) = 0$ πιο πάνω έχουμε απλώς ότι η μέση τιμή $\mu := \mathbf{E}(X_1)$ ορίζεται και είναι πραγματικός αριθμός, τότε τα θεωρήματα ισχύουν πάλι αρκεί να αντικαταστήσουμε την S_n με την $S_n - \mu n$.

Τώρα κάποια σχόλια στα πιο πάνω θεωρήματα.

Ένα γεγονός για το S_n το λέμε τυπικό αν, καθώς το n τείνει στο άπειρο, το γεγονός έχει πιθανότητα φραγμένη μακριά από το 0 (π.χ. είναι μεγαλύτερη του 10^{-6} για όλα τα n), αλλιώς το λέμε μη τυπικό.

Τα θεωρήματα (α), (β) αφορούν τυπικά γεγονότα για το S_n (τυπική συμπεριφορά της $(S_n)_{n \geq 1}$). Το (α) αναφέρεται στο γεγονός $|S_n/n| < \varepsilon$ (για οποιοδήποτε $\varepsilon > 0$) και συνεπάγεται ότι η πιθανότητά του τείνει στο 1. Το (β) λέει ότι το S_n είναι τυπικά της τάξης \sqrt{n} . Μάλιστα, για $a < b$, το γεγονός $a\sigma < S_n/\sqrt{n} < b\sigma$ έχει πιθανότητα που συγκλίνει στον θετικό αριθμό $\Phi(b) - \Phi(a)$, άρα είναι τυπικό.

Τα (γ), (δ) αφορούν μη τυπική συμπεριφορά του S_n .

Σχετικά με το (γ). Ξέρουμε ήδη από το (β) ότι το S_n είναι τυπικά της τάξης του \sqrt{n} , και ένα γεγονός της μορφής

$$A_n := \left\{ \frac{S_n}{\sqrt{n}} \in ((1 - \varepsilon)a_n, (1 + \varepsilon)a_n) \right\}$$

με $a_n \rightarrow \infty$ έχει πιθανότητα που τείνει στο 0. Ειδικά για την επιλογή $a_n := \sigma \sqrt{2 \log \log n}$ έχουμε ότι για οποιοδήποτε δεδομένο n η πιθανότητα του A_n είναι πολύ μικρή (και γίνεται αμελητέα για μεγάλο n), όμως σύμφωνα με το (γ), αν παρατηρήσουμε ολόκληρη την τροχιά της $(S_n)_{n \geq 1}$, θα διαπιστώσουμε ότι το S_n κατορθώνει άπειρες φορές να πραγματοποιήσει το A_n , δηλαδή να γίνει της τάξης $\sqrt{n}a_n$ (όπως και της τάξης $-\sqrt{n}a_n$). Η $(a_n)_{n \geq 1}$ δίνει πόσο μεγάλο μπορεί να γίνει το S_n άπειρες φορές. Έτσι, για παράδειγμα, στην περίπτωση που η X_1 είναι ομοιόμορφη στο $\{-1, 1\}$, αν και το γεγονός $S_n = n/2$ είναι εφικτό (έχει θετική πιθανότητα), με πιθανότητα 1 δεν μπορεί να γίνει άπειρες φορές.

Σχετικά με το (δ). Το (α) συνεπάγεται ότι για οποιοδήποτε σύνολο A με $0 \notin \bar{A}$ η πιθανότητα του γεγονότος $S_n/n \in A$ τείνει στο 0. Άρα μιλάμε για ένα μη τυπικό γεγονός. Το Θεώρημα Cramer προσδιορίζει το ρυθμό μείωσης της πιθανότητας του.

Σε αυτές τις σημειώσεις θα δούμε τις αποδείξεις των (α), (β), (δ).

Τα θεωρήματα (α), (β), (γ), (δ) και οι τεχνικές απόδειξής τους λειτουργούν ως υπόδειγμα για την ασυμπτωτική μελέτη ακολουθιών τυχαίων μεταβλητών που ενδεχομένως είναι πιο σύνθετες από την S_n .

Δημήτρης Χελιώτης
27 Οκτωβρίου 2020

Σύμβολα

\mathbb{N} το σύνολο των μη αρνητικών ακεραίων $\{0, 1, 2, \dots\}$.

\mathbb{N}^+ το σύνολο των θετικών ακεραίων $\{1, 2, \dots\}$.

Για $n \in \mathbb{N}^+$,

$$[n] := \{1, 2, \dots, n\}.$$

Για ακολουθία συνόλων $(A_n)_{n \geq 1}$,

$$\liminf_n A_n := \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k,$$

$$\limsup_n A_n := \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k.$$

Για ακολουθία πραγματικών αριθμών $(x_n)_{n \geq 1}$,

$$\underline{\lim}_n x_n := \sup_{n \geq 1} \inf_{k \geq n} x_k,$$

$$\overline{\lim}_n x_n := \inf_{n \geq 1} \sup_{k \geq n} x_k.$$

Για A, B σύνολα, $A \subset B$: το A είναι υποσύνολο του B (όχι απαραίτητα γνήσιο).

Για A, B σύνολα, $A \setminus B := \{x \in A : x \notin B\}$.

Για X σύνολο,

$$\mathcal{P}(X) := \{A : A \subset X\},$$

το δυναμοσύνολο του X .

Για X μετρικό χώρο,

$\mathcal{B}(X)$: η σ -άλγεβρα των Borel υποσυνόλων του X .

Για A υποσύνολο τοπολογικού χώρου (π.χ., μετρικού χώρου),

A° : το εσωτερικό του A ,

\bar{A} : η κλειστότητα του A ,

$\partial A := \bar{A} \setminus A^\circ$: το σύνορο του A .

Για $x, y \in \mathbb{R}$,

$$x \wedge y := \min\{x, y\},$$

$$x \vee y := \max\{x, y\}.$$

Για $x \in \mathbb{R}$,

$$x^+ = x \vee 0 = \begin{cases} x & \text{αν } x > 0, \\ 0 & \text{αν } x \leq 0, \end{cases} \quad x^- = (-x) \vee 0 = \begin{cases} -x & \text{αν } x < 0, \\ 0 & \text{αν } x \geq 0. \end{cases}$$

Το θετικό και το αρνητικό μέρος του x .

Δείκτριες

Χρησιμοποιούμε δυο ειδών δείκτριες συναρτήσεις. Παραδείγματα χρήσης τους είναι τα εξής.

1) $\mathbf{1}_A$, με $A \subset \mathbb{R}$ και πεδίο ορισμού το \mathbb{R} .

$$\mathbf{1}_A(x) := \begin{cases} 1 & \text{αν } x \in A, \\ 0 & \text{αν } x \in \mathbb{R} \setminus A. \end{cases}$$

2) $\mathbf{1}_{x \in A}$, $\mathbf{1}_{x+y>1}$, $\mathbf{1}_{f \leq 0}$ με $f : S \rightarrow \mathbb{R}$ συνάρτηση.

Η $\mathbf{1}_{x \in A}$ είναι άλλος τρόπος γραφής της $\mathbf{1}_A(x)$.

Η $\mathbf{1}_{x+y>1}$ παίρνει την τιμή 1 αν $x + y > 1$ και την τιμή 0 διαφορετικά.

Η $\mathbf{1}_{f \leq 0}$ είναι συνάρτηση με πεδίο ορισμού το S και παίρνει την τιμή 1 αν $f(x) \leq 0$ και την τιμή 0 διαφορετικά.

Το πρώτο είδος είναι το πιο συνηθισμένο, από τη συνολοθεωρία: δείκτρια συνόλου. Το δεύτερο έχει τη γενική μορφή $\mathbf{1}_P$, όπου P είναι μια λογική πρόταση, και η τιμή της δείκτριας είναι

$$\mathbf{1}_P = \begin{cases} 1 & \text{αν η πρόταση } P \text{ είναι αληθής,} \\ 0 & \text{αν η πρόταση } P \text{ είναι ψευδής.} \end{cases}$$

1

σ-άλγεβρες

1.1 σ-άλγεβρες

Έστω X σύνολο. Συμβολίζουμε με $\mathcal{P}(X)$ το δυναμοσύνολο του X , δηλαδή $\mathcal{P}(X) = \{A : A \subset X\}$.

Ορισμός 1.1. Μια οικογένεια $\mathcal{A} \subset \mathcal{P}(X)$ λέγεται **σ-άλγεβρα** στο X αν έχει τις εξής ιδιότητες:

- (i) $\emptyset \in \mathcal{A}$.
- (ii) Αν $A \in \mathcal{A}$, τότε $X \setminus A \in \mathcal{A}$.
- (iii) Η \mathcal{A} είναι κλειστή στις άπειρες αριθμήσιμες ενώσεις, δηλαδή αν $(A_n)_{n \in \mathbb{N}}$ είναι ακολουθία στοιχείων της \mathcal{A} , τότε $\cup_{n \in \mathbb{N}} A_n \in \mathcal{A}$.

Παρατήρηση 1.2. Για κάθε σ-άλγεβρα \mathcal{A} ισχύει επίσης ότι

- $X \in \mathcal{A}$ λόγω των (i) και (ii) εφόσον το X είναι το συμπλήρωμα του \emptyset .
- Η \mathcal{A} είναι κλειστή στις πεπερασμένες ενώσεις γιατί αν A_1, A_2, \dots, A_k είναι στοιχεία της τότε ορίζοντας $A_j := \emptyset$ για κάθε $j \geq k + 1$ έχουμε ότι $\cup_{j=1}^k A_j = \cup_{j=1}^{\infty} A_j$ και η τελευταία ένωση είναι στοιχείο της \mathcal{A} λόγω του (iii) του ορισμού.
- Η \mathcal{A} είναι κλειστή στις αριθμήσιμες τομές γιατί για J αριθμήσιμο σύνολο και $(A_i)_{i \in J}$ στοιχεία της \mathcal{A} έχουμε $\cap_{n \in J} A_n = X \setminus \{ \cup_{n \in J} (X \setminus A_n) \}$ και λόγω της προηγούμενης παρατήρησης και των (ii), (iii) του ορισμού έπεται το συμπέρασμα.
- Αν $A, B \in \mathcal{A}$, τότε $A \setminus B \in \mathcal{A}$ γιατί $A \setminus B = A \cap (X \setminus B)$.

Παράδειγμα 1.3. Αν X σύνολο, τότε οι οικογένειες

$$\begin{aligned}\mathcal{A}_1 &:= \{\emptyset, X\}, \\ \mathcal{A}_2 &:= \mathcal{P}(X),\end{aligned}$$

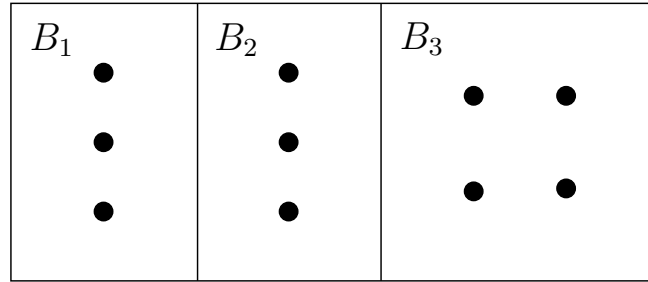
είναι σ-άλγεβρες στο X . Η πρώτη είναι η ελάχιστη δυνατή και η δεύτερη η μέγιστη δυνατή σ-άλγεβρα στο X . Δηλαδή, για οποιαδήποτε σ-άλγεβρα \mathcal{A} στο X ισχύει $\mathcal{A}_1 \subset \mathcal{A} \subset \mathcal{A}_2$. Η \mathcal{A}_1 ονομάζεται η τετριμμένη σ-άλγεβρα στο X .

Παράδειγμα 1.4. Έστω $X = \{1, 2, \dots, 10\}$. Θέτουμε $B_1 := \{1, 2, 3\}$, $B_2 := \{4, 5, 6\}$, $B_3 := \{7, 8, 9, 10\}$ (δες Σχήμα 1.1). Η οικογένεια

$$\mathcal{A} = \{\emptyset, X, B_1, B_2, B_3, B_1 \cup B_2, B_1 \cup B_3, B_2 \cup B_3\}$$

είναι σ-άλγεβρα στο X . Για παράδειγμα, το συμπλήρωμα του B_2 είναι το $B_1 \cup B_3$ το οποίο βρίσκεται και αυτό στην \mathcal{A} .

σ-άλγεβρες



Σχήμα 1.1: Μία διαμέριση του $\{1, 2, \dots, 10\}$.

Αντίθετα, η

$$\mathcal{B} = \{\emptyset, X, B_1, B_2, B_1 \cup B_2\}$$

δεν είναι σ-άλγεβρα γιατί, ενώ είναι κλειστή στις ενώσεις, δεν είναι κλειστή στα συμπληρώματα. Τα συμπληρώματα των $B_1, B_2, B_1 \cup B_2$ δεν περιέχονται στη \mathcal{B} .

Παράδειγμα 1.5. Έστω $X = \mathbb{R}$. Η οικογένεια

$$\mathcal{A} := \{A \subset \mathbb{R} : A \text{ αριθμησιμο ή } \mathbb{R} \setminus A \text{ αριθμησιμο}\}$$

είναι σ-άλγεβρα (εύκολη άσκηση).

Πρόταση 1.6. Έστω X σύνολο και $(\mathcal{A}_i)_{i \in I}$ οικογένεια σ-αλγεβρών στο X . Τότε, η

$$\mathcal{H} = \bigcap_{i \in I} \mathcal{A}_i := \{A \subset X : A \in \mathcal{A}_i \quad \forall i \in I\}$$

είναι σ-άλγεβρα στο X .

Απόδειξη. Προφανώς το \emptyset ανήκει στην \mathcal{H} γιατί είναι στοιχείο κάθε σ-άλγεβρας \mathcal{A}_i στο X . Αν $A \in \mathcal{A}_i$ για κάθε $i \in I$, τότε, επειδή κάθε \mathcal{A}_i είναι σ-άλγεβρα, έπεται ότι

$$X \setminus A \in \mathcal{A}_i \quad \forall i \in I,$$

δηλαδή $X \setminus A \in \bigcap_{i \in I} \mathcal{A}_i$. Όμοια αποδεικνύουμε ότι η $\bigcap_{i \in I} \mathcal{A}_i$ είναι κλειστή στις αριθμησιμες ενώσεις. Έπεται λοιπόν ότι η \mathcal{H} είναι σ-άλγεβρα. ■

1.2 Παραγόμενη σ-άλγεβρα

Ορισμός 1.7. Έστω X σύνολο και $C \subset \mathcal{P}(X)$. Το C δεν είναι απαραίτητα σ-άλγεβρα. Ορίζουμε

$$\mathcal{J}(C) := \{\mathcal{A} \subset \mathcal{P}(X) : \mathcal{A} \supset C \text{ και η } \mathcal{A} \text{ είναι σ-άλγεβρα}\},$$

δηλαδή το σύνολο των σ-αλγεβρών στο X που καθεμία τους περιέχει την οικογένεια C . Η σ-άλγεβρα που παράγεται από την C ορίζεται ως η τομή όλων των σ-άλγεβρων που περιέχουν την C και συμβολίζεται με $\sigma(C)$, δηλαδή

$$\sigma(C) := \bigcap_{\mathcal{A} \in \mathcal{J}(C)} \mathcal{A}.$$

Η $\sigma(C)$ περιέχει ακριβώς όλα τα $B \subset X$ με την ιδιότητα $B \in \mathcal{A}$ για κάθε σ-άλγεβρα \mathcal{A} στο X με $\mathcal{A} \supset C$.

Από την Πρόταση 1.6, έπεται ότι η $\sigma(C)$ είναι πράγματι σ-άλγεβρα που περιέχει την οικογένεια C και από την κατασκευή της είναι η μικρότερη με την ιδιότητα αυτή. Δηλαδή περιέχεται σε οποιαδήποτε σ-άλγεβρα περιέχει την C . Προφανώς, αν η C είναι σ-άλγεβρα, τότε $\sigma(C) = C$.

Μπορούμε να έχουμε στο μυαλό μας ότι η $\sigma(C)$ προκύπτει με την εξής αναδρομική διαδικασία. Ξεκινάμε με την C και, αν αυτή δεν είναι σ-άλγεβρα, π.χ., γιατί το συμπλήρωμα ενός στοιχείου της ή κάποια αριθμήσιμη ένωση στοιχείων της δεν είναι στοιχείο της, προσθέτουμε σε αυτή το σύνολο που ανακαλύψαμε ότι της λείπει. Αυτό πρέπει να το κάνουμε πολλές φορές με τη νέα οικογένεια που προκύπτει μετά την προσθήκη κάθε συνόλου. Κάποια στιγμή φτάνουμε σε μια οικογένεια που είναι σ-άλγεβρα και τότε σταματάμε, βρήκαμε τη $\sigma(C)$.

Στα πιο κάτω παραδείγματα, αυτό είναι το σκεπτικό που μας οδηγεί. Βέβαια για την απόδειξη ακολουθούμε τον τυπικό ορισμό της $\sigma(C)$.

Παράδειγμα 1.8. Έστω X μη κενό σύνολο με τουλάχιστον δύο στοιχεία και $\emptyset \subsetneq A \subsetneq X$. Η σ-άλγεβρα που παράγεται από την οικογένεια $C := \{A\}$ είναι η $\mathcal{B} := \{\emptyset, X, A, A^c\}$. Πράγματι, η \mathcal{B} είναι σ-άλγεβρα, και οποιαδήποτε σ-άλγεβρα \mathcal{A} περιέχει το A θα πρέπει να περιέχει και το A^c (και τα \emptyset, X βέβαια), άρα $\mathcal{A} \supset \mathcal{B}$.

Παράδειγμα 1.9. Επιστρέφουμε στο Παράδειγμα 1.4. Η οικογένεια \mathcal{B} δεν είναι σ-άλγεβρα, και με σκεπτικό όπως στο προηγούμενο παράδειγμα διαπιστώνουμε ότι $\sigma(\mathcal{B}) = \mathcal{A}$.

Παράδειγμα 1.10 (Σ-άλγεβρα παραγόμενη από αριθμήσιμη διαμέριση). Έστω X σύνολο και $C := \{A_i : i \in I\}$ διαμέριση του X (δηλαδή τα A_i είναι μη κενά σύνολα, ξένα ανά δύο, με ένωση το X), με I αριθμήσιμο (πεπερασμένο ή άπειρο). Για τη σ-άλγεβρα που παράγει η C έχουμε την εξής απλή περιγραφή:

$$\sigma(C) = \{\cup_{i \in J} A_i : J \subset I\}. \quad (1.1)$$

Δηλαδή ένα σύνολο της $\sigma(C)$ είναι ένωση κάποιων στοιχείων της διαμέρισης C .

Για να το δούμε αυτό, ας ονομάσουμε \mathcal{A} το σύνολο στο δεξί μέλος της παραπάνω σχέσης. Τότε έχουμε τα εξής:

- (i) Οποιαδήποτε σ-άλγεβρα \mathcal{A}_1 περιέχει την C πρέπει να περιέχει την \mathcal{A} . Γιατί οποιαδήποτε ένωση $\cup_{i \in J} A_i$ είναι αριθμήσιμη αφού το I είναι αριθμήσιμο. Και εφόσον η \mathcal{A}_1 είναι σ-άλγεβρα και περιέχει τα A_i με $i \in J$, θα περιέχει και την ένωσή τους.
- (ii) Η οικογένεια \mathcal{A} περιέχει τη C . Πράγματι, οποιοδήποτε σύνολο της C είναι της μορφής A_{i_0} για κάποιο $i_0 \in I$. Η επιλογή $J = \{i_0\} \subset I$ στην περιγραφή στοιχείων της \mathcal{A} δίνει $\cup_{i \in J} A_i = A_{i_0} \in \mathcal{A}$.
- (iii) Η \mathcal{A} είναι σ-άλγεβρα. Πράγματι, η επιλογή $J = \emptyset$ δίνει $\cup_{i \in J} A_i = \emptyset$. Επίσης, αν πάρουμε A της μορφής $A = \cup_{i \in J} A_i$ για κάποιο $J \subset I$, τότε $X \setminus A = \cup_{i \in I \setminus J} A_i$ που είναι στοιχείο της \mathcal{A} . Τέλος, αν έχουμε ακολουθία $(B_n)_{n \geq 1}$ στοιχείων της \mathcal{A} με $B_n = \cup_{i \in J_n} A_i$ όπου $J_n \subset I$ για κάθε $n \geq 1$, τότε για $J := \cup_{n=1}^{\infty} J_n$ έχουμε $\cup_{n=1}^{\infty} B_n = \cup_{i \in J} A_i$, που πάλι είναι στοιχείο της \mathcal{A} .

Η πρώτη παρατήρηση συνεπάγεται ότι $\mathcal{A} \subset \sigma(C)$, ενώ οι υπόλοιπες δύο δίνουν ότι $\sigma(C) \subset \mathcal{A}$. Έτσι η (1.1) αποδείχθηκε.

1.3 Τα σύνολα Borel

Αν (X, d) είναι ένας μετρικός χώρος, τότε ένα $A \subset X$ το λέμε ανοιχτό σύνολο αν για κάθε $x \in A$ υπάρχει $\delta > 0$ ώστε $B(x, \delta) := \{y \in X : d(y, x) < \delta\} \subset A$, δηλαδή η σφαίρα ακτίνας δ γύρω από το x είναι υποσύνολο του A . Γενικά, το δ εξαρτάται από το x .

σ-άλγεβρες

Τα σύνολα $\mathbb{R}^n, \mathbb{C}^n$ θα τα θεωρούμε μετρικούς χώρους με μετρική την Ευκλείδεια. Δηλαδή

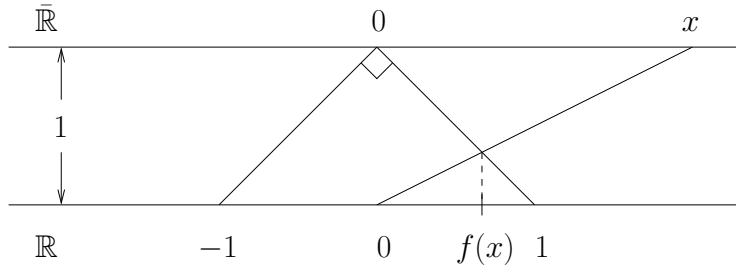
$$d(x, y) = \sqrt{\sum_{k=1}^n |x_k - y_k|^2}$$

για $x, y \in \mathbb{R}^n$ (ή $x, y \in \mathbb{C}^n$) με $|\cdot|$ να συμβολίζει την απόλυτη τιμή πραγματικού αριθμού (το μέτρο μιγαδικού, αντίστοιχα).

Θα χρειαστεί πιο κάτω να δουλέψουμε και με το σύνολο $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$. Αυτό θα το θεωρούμε μετρικό χώρο με μετρική την

$$d(x, y) = |f(x) - f(y)|, \tag{1.2}$$

όπου $f(x) = x/(1 + |x|)$ για $x \in \mathbb{R}$ και $f(-\infty) = -1, f(\infty) = 1$. Απεικονίζουμε τον $\bar{\mathbb{R}}$ στο τμήμα $[-1, 1]$ μέσω της f και μετά ορίζουμε την απόσταση δύο σημείων του ως τη συνηθισμένη απόσταση των εικόνων τους στο $[-1, 1]$ (δες Σχήμα 1.2). Ο περιορισμός της μετρικής αυτής στο \mathbb{R} είναι μια



Σχήμα 1.2: Γεωμετρική ερμηνεία της απεικόνισης $f(x) = x/(1 + |x|)$. Το $\mathbb{R} \cup \{-\infty, \infty\}$ απεικονίζεται στο $[-1, 1]$.

μετρική ισοδύναμη με τη συνηθισμένη μετρική του \mathbb{R} . Σχετικά με το ∞ , εύκολα βλέπει κανείς ότι για $0 < \varepsilon < 1$ η σφαίρα ακτίνας ε γύρω από το ∞ είναι η ημιευθεία $(1 - \varepsilon^{-1}, \infty]$ (αντίστοιχη παρατήρηση ισχύει και για το $-\infty$).

Γενικά, η οικογένεια \mathcal{T} των ανοικτών συνόλων σε έναν μετρικό χώρο δεν είναι σ-άλγεβρα. Για παράδειγμα, στο \mathbb{R} το $A = (0, 1)$ είναι ανοικτό, ενώ το συμπλήρωμά του δεν είναι. Θα μας φανεί χρήσιμο όμως να θεωρήσουμε τη σ-άλγεβρα που παράγουν τα ανοιχτά σύνολα.

Ορισμός 1.11. Έστω (X, d) μετρικός χώρος. Η σ-άλγεβρα $\sigma(\mathcal{T})$ που παράγεται από την οικογένεια \mathcal{T} των ανοικτών συνόλων του X ονομάζεται **σ-άλγεβρα Borel** και τα στοιχεία της **σύνολα Borel**. Συνήθως συμβολίζουμε τη $\sigma(\mathcal{T})$ με $\mathcal{B}(X)$.

Η $\mathcal{B}(X)$ είναι η μικρότερη σ-άλγεβρα που περιέχει τα ανοιχτά σύνολα.

Πρόταση 1.12. Κάθε ανοικτό ή κλειστό υποσύνολο ενός μετρικού χώρου (X, \mathcal{T}) είναι σύνολο Borel.

Απόδειξη. Απο τον ορισμό των συνόλων Borel έχουμε $\mathcal{T} \subset \sigma(\mathcal{T}) =: \mathcal{B}(X)$. Αν F είναι κλειστό, τότε $X \setminus F \in \mathcal{B}(X)$ ως ανοικτό. Αλλά η $\mathcal{B}(X)$ είναι σ-άλγεβρα, οπότε πρέπει και το συμπλήρωμα του $X \setminus F$ να περιέχεται στην $\mathcal{B}(X)$. Άρα $F \in \mathcal{B}(X)$. ■

Πρόταση 1.13. Κάθε υποδιάστημα του \mathbb{R} είναι σύνολο Borel.

Απόδειξη. Τα διάφορα σενάρια για ένα υποδιάστημα είναι

$$(-\infty, a], [a, \infty), (-\infty, a), (a, \infty), (a, b), [a, b], (a, b], [a, b).$$

Τα πρώτα δύο είναι κλειστά σύνολα, τα επόμενα τρία είναι ανοιχτά και το $[a, b]$ είναι κλειστό. Για το $(a, b]$ γράφουμε

$$(a, b] = \mathbb{R} \setminus ((-\infty, a] \cup (b, \infty)).$$

Επειδή η $\mathcal{B}(\mathbb{R})$ είναι σ-άλγεβρα και $(-\infty, a], (b, \infty) \in \mathcal{B}(\mathbb{R})$, έπεται ότι $(-\infty, a] \cup (b, \infty) \in \mathcal{B}(\mathbb{R})$ και $\mathbb{R} \setminus ((-\infty, a] \cup (b, \infty)) \in \mathcal{B}(\mathbb{R})$. Όμοια δείχνουμε ότι $[a, b) \in \mathcal{B}(\mathbb{R})$. ■

Επειδή η $\mathcal{B}(\mathbb{R})$ είναι σ-άλγεβρα και περιέχει όλα τα διαστήματα, έπεται ότι όλα τα σύνολα που φτιάχνουμε ξεκινώντας από διαστήματα και εφαρμόζοντας αριθμήσιμο πλήθος επαναλήψεων τις πράξεις της ένωσης, της τομής, και του συμπληρώματος θα είναι επίσης στοιχεία της $\mathcal{B}(\mathbb{R})$.

Θεώρημα 1.14. Έστω \mathcal{F} η οικογένεια των κλειστών συνόλων του \mathbb{R} και

$$\begin{aligned} \mathcal{A}_1 &:= \{(-\infty, b] : b \in \mathbb{R}\}, \\ \mathcal{A}_2 &:= \{(a, b] : a < b, \quad a, b \in \mathbb{R}\}, \\ \mathcal{A}_3 &:= \{(a, b) : a < b, \quad a, b \in \mathbb{R}\}. \end{aligned}$$

Τότε $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{F}) = \sigma(\mathcal{A}_1) = \sigma(\mathcal{A}_2) = \sigma(\mathcal{A}_3)$.

Απόδειξη. Θα δείξουμε ότι

$$\mathcal{B}(\mathbb{R}) \supset \sigma(\mathcal{F}) \supset \sigma(\mathcal{A}_1) \supset \sigma(\mathcal{A}_2) \supset \sigma(\mathcal{A}_3) \supset \mathcal{B}(\mathbb{R}).$$

Η $\mathcal{B}(\mathbb{R})$ περιέχει τα ανοικτά σύνολα άρα και τα συμπληρώματά τους, δηλαδή τα κλειστά σύνολα, συνεπώς και τη $\sigma(\mathcal{F})$. Τα διαστήματα της \mathcal{A}_1 είναι κλειστά, άρα $\sigma(\mathcal{F}) \supset \sigma(\mathcal{A}_1)$. Για τη $\sigma(\mathcal{A}_1) \supset \sigma(\mathcal{A}_2)$ παρατηρούμε ότι $(a, b] = (-\infty, b] \setminus (-\infty, a]$ και για τη $\sigma(\mathcal{A}_2) \supset \sigma(\mathcal{A}_3)$ ότι $(a, b) = \bigcup_{n=1}^{\infty} (a, b - \frac{1}{n}]$. Τέλος, για τη $\sigma(\mathcal{A}_3) \supset \mathcal{B}(\mathbb{R})$ γνωρίζουμε ότι κάθε μη κενό ανοικτό σύνολο στο \mathbb{R} γράφεται ως αριθμήσιμη ένωση ανοικτών φραγμένων διαστημάτων, άρα $\mathcal{T} \subset \sigma(\mathcal{A}_3)$, που δίνει το συμπέρασμα. ■

Επίσης, η $\mathcal{B}(\mathbb{R})$ παράγεται από την οικογένεια

$$\mathcal{A} = \{(-\infty, q] : q \in \mathbb{Q}\},$$

η οποία είναι αριθμήσιμη. Αυτό ισχύει γιατί $\mathcal{A} \subset \mathcal{A}_1 = \mathcal{B}(\mathbb{R})$ (η \mathcal{A}_1 ορίστηκε στο Θεώρημα 1.14) και $\mathcal{A}_1 \subset \sigma(\mathcal{A})$ γιατί για κάθε $b \in \mathbb{R}$ έχουμε

$$(-\infty, b] = \bigcap_{q > b, q \in \mathbb{Q}} (-\infty, q]$$

και το σύνολο στο δεξί μέλος είναι στοιχείο της $\sigma(\mathcal{A})$. Άρα $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{A}_1) \subset \sigma(\mathcal{A})$.

1.4 Liminf και limsup ακολουθίας συνόλων

Έστω X σύνολο και $(A_n)_{n \geq 1}$ ακολουθία υποσυνόλων του. Ορίζουμε τα σύνολα

$$\liminf_{n \geq 1} A_n := \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k, \tag{1.3}$$

$$\limsup_{n \geq 1} A_n := \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k, \tag{1.4}$$

τα οποία λέμε liminf και limsup αντίστοιχα της ακολουθίας $(A_n)_{n \geq 1}$. Είναι και τα δύο υποσύνολα της ένωσης $\bigcup_{n=1}^{\infty} A_n$.

σ-άλγεβρες

Με λόγια, ένα $x \in X$ ανήκει στο $\liminf_{n \geq 1} A_n$ αν από έναν δείκτη και μετά ανήκει σε όλα τα στοιχεία της ακολουθίας $(A_n)_{n \geq 1}$, ενώ ανήκει στο $\limsup_{n \geq 1} A_n$ αν ανήκει σε άπειρα από τα $(A_n)_{n \geq 1}$. Από τον τυπικό ορισμό ή από αυτή την περιγραφή είναι σαφές ότι $\liminf_{n \geq 1} A_n \subset \limsup_{n \geq 1} A_n$. Το να είναι ένα x μέλος του $\liminf_{n \geq 1} A_n$ είναι ισχυρότερη απαίτηση και έτσι λιγότερα x την ικανοποιούν.

Παράδειγμα 1.15. Αν $X = \mathbb{R}$ και $A_n = [-n, n]$ για κάθε $n \geq 1$, τότε $\liminf_{n \geq 1} A_n := \mathbb{R}$. Για κάθε πραγματικό αριθμό x υπάρχει φυσικός $n(x)$ ώστε ο x να ανήκει σε όλα τα A_n με $n \geq n(x)$ (άρα αυτά τα οποία εξαιρούνται έχουν πεπερασμένο πλήθος). Κάθε x έχει το δικό του $n(x)$. Μάλιστα, μπορούμε να πάρουμε $n(x) := \lceil |x| \rceil + 1$.

Παράδειγμα 1.16. Αν $X = \mathbb{R}$ και $A_n = \{x : nx \text{ είναι ακέραιος}\} = \mathbb{Z}/n$ για κάθε $n \in \mathbb{N}^+$, τότε

$$\liminf_{n \geq 1} A_n := \mathbb{Z},$$

$$\limsup_{n \geq 1} A_n := \mathbb{Q}.$$

Σχετικά με τον δεύτερο ισχυρισμό, για κάθε ρητό $x = p/q$ με q θετικό ακέραιο έχουμε ότι $x \in A_n$ για κάθε n που είναι πολλαπλάσιο του q . Άρα έχουμε την \supset . Και επειδή όλα τα A_n είναι υποσύνολα του \mathbb{Q} , έχουμε ότι και $\limsup_{n \geq 1} A_n \subset \mathbb{Q}$. Η απόδειξη του πρώτου ισχυρισμού αφήνεται στον αναγνώστη.

Αν οι όροι της ακολουθίας $(A_n)_{n \geq 1}$ είναι στοιχεία μιας σ-άλγεβρας \mathcal{A} , τότε και τα $\liminf_{n \geq 1} A_n$, $\limsup_{n \geq 1} A_n$ είναι επίσης στοιχεία της \mathcal{A} . Ας το δούμε για το \limsup . Επειδή η \mathcal{A} είναι κλειστή στις αριθμήσιμες ενώσεις, για κάθε $n \geq 1$ έχουμε $B_n := \cup_{k=n} A_k \in \mathcal{A}$, και επειδή είναι κλειστή στις αριθμήσιμες τομές έχουμε $\limsup_{n \geq 1} A_n = \cap_{n \geq 1} B_n \in \mathcal{A}$.

Ασκήσεις

1.1 Έστω $X := \{\alpha, \beta, \gamma, \delta\}$ και

$$\mathcal{A}_1 := \{\emptyset, X, \{\beta, \gamma\}\},$$

$$\mathcal{A}_2 := \{\emptyset, X, \{\beta, \gamma\}, \{\alpha, \delta\}\}.$$

(α) Είναι οι $\mathcal{A}_1, \mathcal{A}_2$ σ-άλγεβρες;

(β) Να δείξετε ότι $\sigma(\mathcal{A}_1) = \mathcal{A}_2$.

1.2 Σε αυτή την άσκηση παίρνουμε $X := \mathbb{R}$.

(α) Να δείξετε ότι η $\mathcal{A} := \{A \subset \mathbb{R} : A \text{ αριθμήσιμο ή } \mathbb{R} \setminus A \text{ αριθμήσιμο}\}$ είναι σ-άλγεβρα και $\mathcal{A} \subset \mathcal{B}(\mathbb{R})$.

(β) Για την οικογένεια $\mathcal{A}_0 := \{\{x\} : x \in \mathbb{R}\}$ να δείξετε ότι $\sigma(\mathcal{A}_0) = \mathcal{A}$.

(γ) Να δείξετε ότι η οικογένεια $\mathcal{A}_1 := \{A \subset \mathbb{R} : A \text{ πεπερασμένο ή } \mathbb{R} \setminus A \text{ πεπερασμένο}\}$ δεν είναι σ-άλγεβρα.

1.3 Αν στο Παράδειγμα 1.10 η διαμέριση \mathcal{C} του X έχει υπεραριθμήσιμο πλήθος στοιχείων (δηλαδή το I είναι υπεραριθμήσιμο, άρα και το X υπεραριθμήσιμο), να δοθεί περιγραφή της παραγόμενης σ-άλγεβρας $\sigma(\mathcal{C})$.

1.4 Έστω $\mathcal{A} := \{[-\infty, x] : x \in \mathbb{R}\}$. Να δειχθεί ότι $\sigma(\mathcal{A}) = \mathcal{B}(\overline{\mathbb{R}})$.

1.5 Να δείξετε ότι η $\mathcal{B}(\mathbb{R})$ δεν παράγεται από διαμέριση.

1.6 Έστω $(A_n)_{n \geq 1}$ ακολουθία στοιχείων μιας σ-άλγεβρας \mathcal{A} . Να δείξετε ότι υπάρχει ακολουθία $(B_n)_{n \geq 1}$ στοιχείων της \mathcal{A} τα οποία είναι ξένα ανά δύο ώστε $B_n \subset A_n$ για κάθε $n \geq 1$ και $\cup_{n=1}^{\infty} A_n = \cup_{n=1}^{\infty} B_n$.

1.7 Έστω $f : X \rightarrow Y$ συνάρτηση.

σ-άλγεβρες

(α) Αν \mathcal{A} είναι σ-άλγεβρα στο X , θέτουμε

$$\mathcal{B} := \{B \subset Y : f^{-1}(B) \in \mathcal{A}\}.$$

Να δείξετε ότι η \mathcal{B} είναι σ-άλγεβρα στο Y .

(β) Αν \mathcal{B} είναι σ-άλγεβρα στο Y , θέτουμε

$$\mathcal{A} := \{f^{-1}(B) : B \in \mathcal{B}\}.$$

Να δείξετε ότι η \mathcal{A} είναι σ-άλγεβρα στο X .

Υπενθυμίζουμε ότι για $B \subset Y$, συμβολίζουμε με $f^{-1}(B)$ το σύνολο $\{x \in X : f(x) \in B\}$. Το f^{-1} εδώ δεν σημαίνει «αντίστροφη της f ». Δεν είναι απαραίτητο να είναι αντιστρέψιμη η f .

1.8 Έστω X σύνολο και $(A_n)_{n \geq 1}$ ακολουθία με στοιχεία υποσύνολα του X . Να δειχθεί ότι

$$\begin{aligned} \overline{\lim}_{n \rightarrow \infty} \mathbf{1}_{A_n}(x) &= \mathbf{1}_{\limsup_{n \geq 1} A_n}(x), \\ \underline{\lim}_{n \rightarrow \infty} \mathbf{1}_{A_n}(x) &= \mathbf{1}_{\liminf_{n \geq 1} A_n}(x) \end{aligned}$$

για κάθε $x \in X$. Για $A \subset X$, $\mathbf{1}_A : X \rightarrow \{0, 1\}$ δηλώνει τη δείκτρια συνάρτηση του συνόλου A .

1.9 Έστω X σύνολο, $(f_n)_{n \geq 1}$ ακολουθία συναρτήσεων με $f_n : X \rightarrow \mathbb{R}$ για κάθε $n \geq 1$, και $c \in \mathbb{R}$ δεδομένο. Για καθένα από τα ακόλουθα ζεύγη συνόλων να εξεταστεί αν μεταξύ τους ισχύει κάποια από τις σχέσεις $\subset, =, \supset$.

(α) $\{x \in X : \overline{\lim}_{n \rightarrow \infty} f_n(x) < c\}$, $\liminf_{n \geq 1} \{x \in X : f_n(x) < c\}$.

(β) $\{x \in X : \overline{\lim}_{n \rightarrow \infty} f_n(x) \leq c\}$, $\liminf_{n \geq 1} \{x \in X : f_n(x) \leq c\}$.

(γ) $\{x \in X : \overline{\lim}_{n \rightarrow \infty} f_n(x) \leq c\}$, $\limsup_{n \geq 1} \{x \in X : f_n(x) \leq c\}$.

(δ) $\{x \in X : \overline{\lim}_{n \rightarrow \infty} f_n(x) \geq c\}$, $\limsup_{n \geq 1} \{x \in X : f_n(x) \geq c\}$.

(ε) $\{x \in X : \overline{\lim}_{n \rightarrow \infty} f_n(x) > c\}$, $\limsup_{n \geq 1} \{x \in X : f_n(x) > c\}$.

1.10 Έστω $(\Omega, \mathcal{A}, \mathbf{P})$ χώρος πιθανότητας.

(α) Θέτουμε $\mathcal{F} := \{A \in \mathcal{A} : \mathbf{P}(A) = 0 \text{ ή } \mathbf{P}(A) = 1\}$. Να δειχθεί ότι η \mathcal{F} είναι σ-άλγεβρα.

(β) Θέτουμε $C := \{A \in \mathcal{A} : \mathbf{P}(A) = 0\}$. Δείξτε ότι η C δεν είναι σ-άλγεβρα. Ποια είναι η σ-άλγεβρα, $\sigma(C)$, που παράγεται από την C ;

2

Μέτρα

2.1 Μέτρα σε μετρήσιμο χώρο

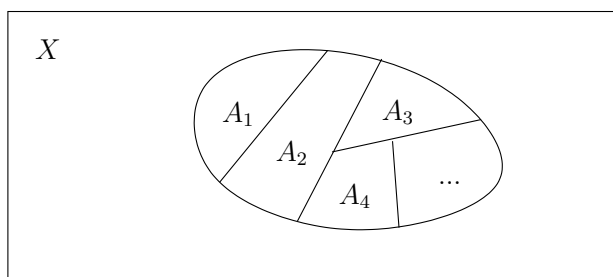
Έστω X σύνολο και \mathcal{A} μια σ -άλγεβρα στο X . Ονομάζουμε το ζεύγος (X, \mathcal{A}) μετρήσιμο χώρο.

Ορισμός 2.1. Μέτρο στον (X, \mathcal{A}) λέμε κάθε συνάρτηση $\mu : \mathcal{A} \rightarrow [0, \infty]$ που ικανοποιεί τις ιδιότητες:

(i) $\mu(\emptyset) = 0$.

(ii) $\mu(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$ για κάθε ακολουθία $(A_n)_{n \geq 1}$ ξένων ανά δύο στοιχείων της \mathcal{A} .

Η τριάδα (X, \mathcal{A}, μ) λέγεται **χώρος μέτρου** και τα στοιχεία της \mathcal{A} **μετρήσιμα σύνολα**. Η ιδιότητα (ii) του ορισμού λέγεται αριθμήσιμη προσθετικότητα.



Σχήμα 2.1: Για την ιδιότητα (ii) του ορισμού του μέτρου.

Παράδειγμα 2.2. (Αριθμητικό μέτρο) Έστω X ένα σύνολο, $\mathcal{A} = \mathcal{P}(X)$, και

$$\mu(A) := \begin{cases} n & \text{αν το } A \text{ είναι πεπερασμένο και έχει ακριβώς } n \text{ στοιχεία,} \\ \infty & \text{αν το } A \text{ είναι άπειροσύνολο} \end{cases}$$

για κάθε $A \in \mathcal{A}$. Το μ είναι το αριθμητικό μέτρο στο X .

Παράδειγμα 2.3. (Μέτρο Dirac) Έστω X ένα σύνολο, $\mathcal{A} = \mathcal{P}(X)$, και $x_0 \in X$ ένα δεδομένο σημείο του X . Ορίζουμε

$$\delta_{x_0}(A) := \begin{cases} 1 & \text{αν } x_0 \in A, \\ 0 & \text{αν } x_0 \in X \setminus A \end{cases}$$

για κάθε $A \in \mathcal{A}$. Η συνάρτηση δ_{x_0} είναι μέτρο και ονομάζεται μέτρο Dirac στο x_0 .

Παράδειγμα 2.4. (Μέτρο περιορισμός) Αν μ είναι ένα μέτρο στον (X, \mathcal{A}) και $C \in \mathcal{A}$, τότε το σύνολο

$$\mathcal{A}_C := \{A \in \mathcal{A} : A \subset C\}$$

είναι σ -άλγεβρα στο C . Η συνάρτηση $\mu_C : \mathcal{A}_C \rightarrow [0, \infty]$ που ορίζεται ως $\mu_C(A) = \mu(A)$ για κάθε $A \in \mathcal{A}_C$ είναι μέτρο στον μετρήσιμο χώρο (C, \mathcal{A}_C) [Άσκηση] και ονομάζεται περιορισμός του μ στο C .

Τα αξιώματα στον ορισμό του μέτρου συνεπάγονται αρκετές ιδιότητες για μια τέτοια συνάρτηση. Καταγράφουμε στην παρακάτω πρόταση κάποιες που στη συνέχεια θα χρησιμοποιήσουμε επανειλημμένα.

Πρόταση 2.5. Έστω μ ένα μέτρο στον (X, \mathcal{A}) . Τότε:

(i) $\mu(\cup_{k=1}^n A_k) = \sum_{k=1}^n \mu(A_k)$ για κάθε $n \geq 1$ και $\{A_n : 1 \leq k \leq n\}$ ξένα ανά δύο στοιχεία της \mathcal{A} .

(ii) Αν $A, B \in \mathcal{A}$, με $A \subset B$, τότε $\mu(A) \leq \mu(B)$, και αν $\mu(A) < \infty$, τότε $\mu(B \setminus A) = \mu(B) - \mu(A)$.

(iii) $\mu(\cup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mu(A_n)$ για κάθε ακολουθία $(A_n)_{n \geq 1}$ στοιχείων της \mathcal{A} .

(iv) Αν $(A_n)_{n \geq 1}$ είναι αύξουσα ακολουθία στοιχείων της \mathcal{A} , τότε $\mu(\cup_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} \mu(A_n)$.

(v) Αν $(A_n)_{n \geq 1}$ είναι φθίνουσα ακολουθία στοιχείων της \mathcal{A} με $\mu(A_1) < \infty$, τότε $\mu(\cap_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} \mu(A_n)$.

Απόδειξη. (i) Τα σύνολα της ακολουθίας $(B_k)_{k \geq 1}$ με

$$B_k := \begin{cases} A_k & \text{αν } k \in \{1, 2, \dots, n\}, \\ \emptyset & \text{αν } k \in \mathbb{N}, k \geq n+1. \end{cases}$$

είναι στοιχεία της \mathcal{A} ξένα ανά δύο. Οπότε η ιδιότητα (ii) του ορισμού του μέτρου δίνει

$$\mu(\cup_{k=1}^n A_k) = \mu(\cup_{k=1}^{\infty} B_k) = \sum_{k=1}^{\infty} \mu(B_k) = \sum_{k=1}^n \mu(A_k)$$

αφού $\mu(\emptyset) = 0$.

(ii) Το B είναι η ένωση των ξένων συνόλων $A, B \setminus A$, οπότε με βάση το (i) της πρότασης,

$$\mu(B) = \mu(A \cup (B \setminus A)) = \mu(A) + \mu(B \setminus A).$$

Επειδή $\mu(B \setminus A) \geq 0$, έπεται ότι $\mu(B) \geq \mu(A)$. Τώρα, όταν $\mu(A) < \infty$, το αφαιρούμε από την πιο πάνω ισότητα, και παίρνουμε ότι $\mu(B \setminus A) = \mu(B) - \mu(A)$.

(iii) Έστω $B_1 := A_1$ και $B_n := A_n \setminus (A_1 \cup A_2 \cup \dots \cup A_{n-1})$ για κάθε $n \geq 2$. Τα $\{B_n : n \geq 1\}$ είναι ξένα ανά δύο στοιχεία της \mathcal{A} , $B_n \subset A_n$ για κάθε $n \geq 1$, και $\cup_{n=1}^{\infty} B_n = \cup_{n=1}^{\infty} A_n$. Άρα

$$\mu(\cup_{n=1}^{\infty} A_n) = \mu(\cup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mu(B_n) \leq \sum_{n=1}^{\infty} \mu(A_n).$$

Η δεύτερη ισότητα ισχύει γιατί τα $(B_n)_{n \geq 1}$ είναι ξένα ανά δύο, και η ανισότητα λόγω της $B_n \subset A_n$ και του μέρους (ii) της πρότασης.

Μέτρα

(iv) Έστω $B_1 := A_1$ και $B_n := A_n \setminus A_{n-1}$ για κάθε $n \geq 2$. Τα $\{B_n : n \geq 1\}$ είναι ξένα ανά δύο στοιχεία της \mathcal{A} , $\cup_{k=1}^n B_k = A_n$, και $\cup_{n=1}^{\infty} B_n = \cup_{n=1}^{\infty} A_n$. Έτσι,

$$\mu(\cup_{n=1}^{\infty} A_n) = \mu(\cup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mu(B_n) = \lim_{n \rightarrow \infty} \sum_{k=1}^n \mu(B_k) = \lim_{n \rightarrow \infty} \mu(\cup_{k=1}^n B_k) = \lim_{n \rightarrow \infty} \mu(A_n).$$

(v) Θέτουμε $B_n = A_1 \setminus A_n$ για κάθε $n \geq 1$. Για την ακολουθία $(B_n)_{n \geq 1}$ εφαρμόζεται το προηγούμενο μέρος της πρότασης και δίνει $\lim_{n \rightarrow \infty} \mu(A_1 \setminus \cap_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} \mu(A_1 \setminus A_n)$. Έπειτα, το ότι $\mu(A_1) < \infty$ δίνει $\mu(\cap_{n=1}^{\infty} A_n) < \infty$ και $\mu(A_n) < \infty$ για κάθε $n \geq 1$, οπότε χρησιμοποιώντας και το μέρος (ii) της πρότασης παίρνουμε το ζητούμενο. ■

Ορισμός 2.6. Ένα μέτρο μ σε έναν μετρήσιμο χώρο (X, \mathcal{A}) λέγεται **πεπερασμένο** αν $\mu(X) < \infty$, και **μέτρο πιθανότητας** αν $\mu(X) = 1$.

Αντίστοιχα, ο χώρος μέτρου (X, \mathcal{A}, μ) λέγεται **χώρος πεπερασμένου μέτρου** ή **χώρος πιθανότητας**. Για έναν χώρο πιθανότητας συνήθως χρησιμοποιείται ο συμβολισμός $(\Omega, \mathcal{F}, \mathbf{P})$.

Ορολογία. Έστω χώρος μέτρου (X, \mathcal{A}, μ) και $x \in X$. Λέμε ότι το x είναι **άτομο** του μ αν $\{x\} \in \mathcal{A}$ και $\mu(\{x\}) > 0$.

2.2 Το μέτρο Lebesgue

Στον χώρο $X = \mathbb{R}$ θεωρούμε τη σ -άλγεβρα $\mathcal{A} = \mathcal{B}(\mathbb{R})$. Είναι δυνατόν να οριστεί ένα μέτρο λ στον χώρο $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ ώστε

$$\lambda(I) = \text{μήκος}(I) \tag{*}$$

για κάθε διάστημα $I \subset \mathbb{R}$. Για παράδειγμα, για $a < b$ πραγματικούς, έχουμε $\lambda((a, b)) = \lambda([a, b]) = b - a$, $\lambda((a, \infty)) = \infty$ και $\lambda([0.1, 2) \cup (3, 4) \cup (5, 5.3)) = 1.9 + 1 + 0.3 = 3.2$.

Πώς μπορούμε να ορίσουμε μια τέτοια συνάρτηση; Ξέρουμε τις τιμές της στα διαστήματα, τα οποία είναι στοιχεία του $\mathcal{B}(\mathbb{R})$, και οι ιδιότητες του μέτρου καθορίζουν μοναδικά τις τιμές της σε ενώσεις διαστημάτων. Αυτό όμως δεν αρκεί. Χρειάζεται να την επεκτείνουμε σε όλο το $\mathcal{B}(\mathbb{R})$. Αποδεικνύεται ότι μια τέτοια επέκταση είναι δυνατή και γίνεται μοναδικά. Δηλαδή υπάρχει μοναδικό μέτρο στο $\mathcal{B}(\mathbb{R})$ που ικανοποιεί την (*). Την κατασκευή αυτού του μέτρου μπορεί να βρει ο αναγνώστης σε βιβλία θεωρίας μέτρου [για παράδειγμα, Κεφάλαιο 3 του Κουμουλλής Γ. και Νεγρεπόντης Σ. (1991)].

Τα Lebesgue-μετρήσιμα σύνολα

Μπορεί άραγε το λ να επεκταθεί σε όλο το $\mathcal{P}(\mathbb{R})$; Δηλαδή να αποδώσουμε σε κάθε υποσύνολο του \mathbb{R} έναν αριθμό που θα είναι το «μήκος» του. Αποδεικνύεται ότι το λ μπορεί να επεκταθεί μέχρι ένα σύνολο \mathcal{M} , που είναι μάλιστα σ -άλγεβρα, με $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{M} \subseteq \mathcal{P}(\mathbb{R})$, αλλά όχι παραπάνω. Γιατί όχι παραπάνω; Είναι τόσο δύσκολο να επεκτείνουμε μια συνάρτηση; Το δύσκολο δεν είναι να την επεκτείνουμε, αλλά να την επεκτείνουμε με τέτοιο τρόπο ώστε να ικανοποιεί την (ii) του Ορισμού 2.1. Αυτή η συνθήκη βάζει τόσες πολλές απαιτήσεις στη λ ώστε να μην υπάρχει καμία $\lambda : \mathcal{P}(\mathbb{R}) \rightarrow [0, \infty]$ που να μπορεί να τις ικανοποιήσει όλες.

Το σύνολο \mathcal{M} έχει την εξής ακριβή περιγραφή:

$$\mathcal{M} := \sigma(\mathcal{B}(\mathbb{R}) \cup \mathcal{N}) = \{A \cup N : A \in \mathcal{B}(\mathbb{R}), N \in \mathcal{N}\},$$

όπου $\mathcal{N} := \{A \subset \mathbb{R} : \text{υπάρχει } B \in \mathcal{B}(\mathbb{R}) \text{ με } B \supset A \text{ και } \lambda(B) = 0\}$. Η πρώτη ισότητα είναι ορισμός ενώ η απόδειξη της δεύτερης ισότητας αφήνεται ως άσκηση. Τα στοιχεία του \mathcal{N} δεν είναι όλα σύνολα

Μέτρα

Borel, ενώ φαίνεται λογικό ότι καθένα πρέπει να έχει μέτρο 0. Τα στοιχεία του \mathcal{M} ονομάζονται **Lebesgue-μετρήσιμα** υποσύνολα του \mathbb{R} . Έχοντας στη διάθεσή μας το μέτρο Lebesgue ορισμένο στο $\mathcal{B}(\mathbb{R})$ το επεκτείνουμε στο \mathcal{M} θέτοντας $\lambda(A \cup N) = \lambda(A)$ για κάθε σύνολο $A \cup N \in \mathcal{M}$. Αποδεικνύεται ότι αυτός ο ορισμός είναι καλός (δηλαδή δεν εξαρτάται από την αναπαράσταση $A \cup N$ του τυχόντος στοιχείου της \mathcal{M}) και η επεκταση του λ στο \mathcal{M} παραμένει μέτρο.

Παρατήρηση 2.7. (Σύνολα με μέτρο Lebesgue 0) Κάθε μονοσύνολο $\{x\} \subset \mathbb{R}$ έχει μέτρο Lebesgue 0 αφού $\{x\} = [x, x]$ είναι ένα διάστημα με μήκος 0. Έπεται από την (ii) του Ορισμού 2.1 του μέτρου ότι κάθε αριθμήσιμο σύνολο έχει επίσης μέτρο Lebesgue 0. Έτσι, το \mathbb{Q} , ενώ είναι ένα πυκνό υποσύνολο του \mathbb{R} και κατά μια έννοια «μεγάλο» σύνολο, έχει μέτρο 0. Υπάρχουν όμως και υπεραριθμήσιμα σύνολα με μέτρο 0, με πιο γνωστό παράδειγμα το σύνολο Cantor. Αυτό γράφεται ως $C = \bigcap_{n=1}^{\infty} C_n$ όπου το C_n είναι ένωση 2^n ξένων διαστημάτων, το καθένα με μήκος 3^{-n} (από τη συνήθη κατασκευή του C). Άρα $\lambda(C) \leq \lambda(C_n) = 2^n 3^{-n} \rightarrow 0$ καθώς $n \rightarrow \infty$.

2.3 Διακριτά μέτρα πιθανότητας

Έστω Ω αριθμήσιμο σύνολο και $\mathcal{F} := \mathcal{P}(\Omega)$.

Κάθε μέτρο πιθανότητας \mathbf{P} στον μετρήσιμο χώρο (Ω, \mathcal{F}) καθορίζεται από τις ποσότητες $f(\omega) := \mathbf{P}(\{\omega\})$, δηλαδή από τις μάζες των μονοσυνόλων. Πράγματι, αν $A \subset \Omega$, τότε το A , ως αριθμήσιμο, γράφεται ως αριθμήσιμη ένωση μονοσυνόλων, $A = \bigcup_{\omega \in A} \{\omega\}$, και επομένως

$$\mathbf{P}(A) = \sum_{\omega \in A} \mathbf{P}(\{\omega\}) = \sum_{\omega \in A} f(\omega). \quad (2.1)$$

Έπειτα, οι τιμές $f(\omega) := \mathbf{P}(\{\omega\})$ ικανοποιούν τις

$$f(\omega) \geq 0 \text{ για κάθε } \omega \in \Omega \text{ και} \quad (2.2)$$

$$\sum_{\omega \in \Omega} f(\omega) = 1. \quad (2.3)$$

Ο πρώτος ισχυρισμός είναι προφανής ενώ για τον δεύτερο υπολογίζουμε

$$\sum_{\omega \in \Omega} f(\omega) = \sum_{\omega \in \Omega} \mathbf{P}(\{\omega\}) = \mathbf{P}(\bigcup_{\omega \in \Omega} \{\omega\}) = \mathbf{P}(\Omega) = 1.$$

Αν θέλουμε να κατασκευάσουμε ένα μέτρο στον χώρο (Ω, \mathcal{F}) , είναι χρήσιμη η επόμενη πρόταση.

Πρόταση 2.8. Αν η $f : \Omega \rightarrow \mathbb{R}$ ικανοποιεί τις (2.2), (2.3), τότε υπάρχει μέτρο πιθανότητας \mathbf{P} στον (Ω, \mathcal{F}) με $\mathbf{P}(\{\omega\}) = f(\omega)$ για κάθε $\omega \in \Omega$.

Απόδειξη. Ορίζουμε τη συνάρτηση $\mathbf{P} : \mathcal{F} \rightarrow [0, 1]$ μέσω της (2.1) για κάθε $A \subset \Omega$, και θα δείξουμε ότι είναι μέτρο πιθανότητας. Προφανώς $\mathbf{P}(\emptyset) = 0$. Έστω $(A_i)_{i \geq 1}$ ακολουθία ξένων ανά δύο υποσυνόλων του Ω . Θεωρούμε μια αρίθμηση $\{\omega_j : j \in I\}$ του Ω όπου $I = \mathbb{N}^+$ ή $I = \{1, 2, \dots, n\}$ για κάποιο $n \in \mathbb{N}^+$. Όταν $I = \mathbb{N}^+$, θέτουμε $n = \infty$. Τότε

$$\mathbf{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{j=1}^n \mathbf{1}_{\omega_j \in \bigcup_{i=1}^{\infty} A_i} f(\omega_j) = \sum_{j=1}^n \sum_{i=1}^{\infty} \mathbf{1}_{\omega_j \in A_i} f(\omega_j) = \sum_{i=1}^{\infty} \sum_{j=1}^n \mathbf{1}_{\omega_j \in A_i} f(\omega_j) = \sum_{i=1}^{\infty} \mathbf{P}(A_i).$$

Τέλος, $\mathbf{P}(\Omega) = \sum_{\omega \in \Omega} f(\omega) = 1$ λόγω της (2.3). ■

Παράδειγμα 2.9. (Ρίψη νομίσματος) Για το πείραμα ρίψης ενός νομίσματος που έχει πιθανότητα $p \in [0, 1]$ να φέρει κορώνα και $1 - p$ να φέρει γράμματα, ένας φυσιολογικός χώρος πιθανότητας προκύπτει με εφαρμογή της προηγούμενης πρότασης. Παίρνουμε $\Omega := \{K, \Gamma\}$, $f(K) = p$, $f(\Gamma) = 1 - p$. Προκύπτει έτσι ένα μέτρο πιθανότητας, έστω $\mathbf{P}^{(p)}$, και τελικά ο χώρος πιθανότητας είναι ο $(\{K, \Gamma\}, \mathcal{P}(\{K, \Gamma\}), \mathbf{P}^{(p)})$.

Παράδειγμα 2.10. (Κατανομή Poisson) Έστω $\Omega = \mathbb{N}$ και $\lambda > 0$. Για κάθε $k \in \mathbb{N}$ έστω $p_k = e^{-\lambda} \frac{\lambda^k}{k!}$. Η $(p_k)_{k \in \mathbb{N}}$ ικανοποιεί τις απαιτήσεις της Πρότασης 2.8. Πράγματι, $p_k > 0$ για κάθε $k \in \mathbb{N}$ και $\sum_{k \in \mathbb{N}} e^{-\lambda} \frac{\lambda^k}{k!} = 1$. Συνεπώς ορίζεται μέτρο πιθανότητας \mathbf{P} στον $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ έτσι ώστε $\mathbf{P}(\{k\}) = p_k$ για κάθε $k \in \mathbb{N}$. Το μέτρο αυτό λέγεται κατανομή **Poisson** με παράμετρο λ .

Παράδειγμα 2.11. * (Το μοντέλο Ising) Έστω $N \geq 1$ φυσικός, $V := [-N, N]^2 \cap \mathbb{Z}^2$,

$$\Omega := \{-1, 1\}^V = \{s \mid s : V \rightarrow \{-1, 1\} \text{ συνάρτηση} \},$$

και $\mathcal{F} := \mathcal{P}(\Omega)$. Το V είναι ένα πεπερασμένο πλέγμα μέσα στο \mathbb{Z}^2 σε σχήμα τετραγώνου. Η εικόνα που έχουμε στο μυαλό μας είναι ότι σε κάθε σημείο του V υπάρχει ένα ηλεκτρόνιο του οποίου η μαγνητική ροπή παίρνει μόνο μία από τις τιμές $-1, 1$. Τα σημεία του Ω τα λέμε σχηματισμούς και κάθε σχηματισμός αναθέτει σε κάθε σημείο του πλέγματος V , δηλαδή σε κάθε ηλεκτρόνιο, μια από τις τιμές $-1, 1$. Γειτονικά σημεία ενός σημείου $x := (x_1, x_2) \in \mathbb{Z}^2$ λέμε τα $(x_1 - 1, x_2)$, $(x_1 + 1, x_2)$, $(x_1, x_2 - 1)$, $(x_1, x_2 + 1)$. Αν το y είναι γειτονικό του x , τότε και το x είναι γειτονικό του y , και γράφουμε $\langle x, y \rangle$.

Έστω $J, h \in \mathbb{R}$ σταθερές. Για κάθε σχηματισμό $s \in \Omega$, ορίζουμε $A(s) := J \sum_{x, y \in V: \langle x, y \rangle} s(x)s(y) + h \sum_{x \in V} s(x)$ και

$$f(s) = \frac{1}{Z} e^{A(s)},$$

όπου $Z := \sum_{r \in \Omega} e^{A(r)}$ είναι ο κατάλληλος αριθμός ώστε αθροίζοντας την f πάνω σε όλα τα $s \in \Omega$ να παίρνουμε 1. Η συνάρτηση f ορίζει, με εφαρμογή της Πρότασης 2.8, ένα μέτρο πιθανότητας στον Ω . Ας υποθέσουμε ότι $J > 0$ και $h = 0$. Τότε μεγαλύτερη πιθανότητα έχουν σχηματισμοί που δίνουν το ίδιο πρόσημο σε πολλά γειτονικά σημεία.

Αυτό το μέτρο πιθανότητας καθώς και γενικεύσεις του έχουν χρησιμοποιηθεί για την κατανόηση των μαγνητικών ιδιοτήτων της ύλης.

Ορισμός 2.12. Έστω Ω πεπερασμένο σύνολο. Ένα μέτρο πιθανότητας \mathbf{P} στον $(\Omega, \mathcal{P}(\Omega))$ λέγεται **ομοιόμορφο** αν υπάρχει $c > 0$ έτσι ώστε $\mathbf{P}(\{\omega\}) = c$ για κάθε $\omega \in \Omega$, δηλαδή το \mathbf{P} δίνει την ίδια μάζα σε κάθε $\omega \in \Omega$.

Από τον Ορισμό 2.12 συμπεραίνουμε ότι

$$\mathbf{P}(A) = \frac{|A|}{|\Omega|} \quad \text{για κάθε } A \subset \Omega.$$

Πράγματι, εφόσον το \mathbf{P} είναι μέτρο πιθανότητας,

$$1 = \sum_{\omega \in \Omega} \mathbf{P}(\{\omega\}) = \sum_{\omega \in \Omega} c = c|\Omega|.$$

Άρα $c = 1/|\Omega|$. Όμως $\mathbf{P}(A) = \sum_{\omega \in A} \mathbf{P}(\{\omega\}) = \sum_{\omega \in A} c = c|A|$, από το οποίο προκύπτει το ζητούμενο.

Τα ομοιόμορφα μέτρα μοντελοποιούν πειράματα που έχουν «ισοπίθανα» αποτελέσματα. Δυο ενδεχόμενα A_1, A_2 σε έναν πεπερασμένο δειγματικό χώρο τα λέμε «ισοπίθανα» αν οι πληροφορίες που έχουμε για αυτά δεν δίνουν προβάδισμα στην πραγματοποίηση του ενός έναντι του άλλου (αρχή της αδιαφορίας). Για παράδειγμα, τα ενδεχόμενα $\{2\}, \{5\}$ σε ένα συμμετρικό και ομοιογενές ζάρι. Αντίθετα, στα ενδεχόμενα «κερδίζω το Λόττο», «δεν κερδίζω το Λόττο», σαφώς έχει προβάδισμα το δεύτερο. Συνήθως το «ισοπίθανα» είναι συνέπεια συμμετρίας.

Παράδειγμα 2.13. Θεωρούμε το πείραμα ρίψης ενός αμερόληπτου ζαριού. Τότε $\Omega = \{1, 2, 3, 4, 5, 6\}$ και $\mathcal{A} = \mathcal{P}(\Omega)$. Το κατάλληλο μέτρο που μοντελοποιεί το πείραμα είναι το \mathbf{P} με $\mathbf{P}(\{\omega\}) = 1/6$ για κάθε $\omega \in \Omega$. Δηλαδή το ομοιόμορφο μέτρο πιθανότητας στον $(\Omega, \mathcal{P}(\Omega))$.

Παράδειγμα 2.14. (Λόττο) Μια κάλπη περιέχει 49 σφαιρίδια αριθμημένα $1, 2, 3, \dots, 49$. Επιλέγουμε 6 από αυτά στην τύχη. Ο χώρος πιθανότητας που μοντελοποιεί το πείραμα έχει δειγματικό χώρο

$$\Omega := \{A \subset \{1, 2, \dots, 49\} : |A| = 6\}.$$

Κάθε στοιχείο του Ω είναι μια δυνατή εξαγωγή. Ο πληθάριθος του Ω είναι

$$|\Omega| = \binom{49}{6} = 13,983,816.$$

Για λόγους συμμετρίας, όλες οι εξαγωγές είναι ισοπίθανες και άρα το σωστό μέτρο πιθανότητας για το πείραμα είναι το ομοιόμορφο μέτρο πιθανότητας \mathbf{P} στον $(\Omega, \mathcal{P}(\Omega))$. Η πιθανότητα του ενδεχομένου $D := \{A \in \Omega : \{8, 9\} \subset A\}$, που σημαίνει ότι στην εξάδα περιέχονται οι αριθμοί 8, 9, είναι

$$\mathbf{P}(D) = \frac{|D|}{|\Omega|} = \frac{\binom{47}{4}}{\binom{49}{6}} = \frac{5}{392} \approx 0.012$$

2.4 Στήριγμα μέτρου

Στο \mathbb{R} θεωρούμε το μέτρο $\nu : \mathcal{P}(\mathbb{R}) \rightarrow [0, \infty]$ με $\nu := \sum_{k \in \mathbb{Z}} \delta_k$, δηλαδή με $\nu(A) = |A \cap \mathbb{Z}|$ για κάθε $A \subset \mathbb{R}$. Το ν έχει την ιδιότητα $\nu(\mathbb{R} \setminus \mathbb{Z}) = 0$, δηλαδή δίνει όλη τη μάζα στο \mathbb{Z} , που είναι ένα γνήσιο υποσύνολο του χώρου \mathbb{R} . Και μάλιστα το \mathbb{Z} είναι το μικρότερο σύνολο με αυτή την ιδιότητα. Αυτό το ιδιαίτερο σύνολο θέλει να πιάσει¹ στη γενική περίπτωση η έννοια του στήριγματος.

Ορισμός 2.15. Έστω (X, \mathcal{A}) μετρήσιμος χώρος ώστε ο X να είναι μετρικός χώρος και η σ -άλγεβρα \mathcal{A} να περιέχει τα σύνολα Borel του X , δηλαδή $\mathcal{B}(X) \subset \mathcal{A}$. Αν μ είναι ένα μέτρο στον (X, \mathcal{A}) , **στήριγμα** (ή και **φορέα**) του μ λέμε το σύνολο

$$\text{supp}(\mu) := \{x \in X : \mu(U) > 0 \text{ για κάθε ανοιχτό υποσύνολο του } X \text{ με } x \in U\} \quad (2.4)$$

$$= X \setminus \{V : V \subset X \text{ ανοιχτό με } \mu(V) = 0\}. \quad (2.5)$$

Το ότι ισχύει η δεύτερη ισότητα αφήνεται ως άσκηση. Από αυτήν γίνεται σαφές ότι το στήριγμα είναι κλειστό σύνολο.

Έστω W η ένωση των ανοιχτών συνόλων στη (2.5). Σε πολλές περιπτώσεις, για παράδειγμα όταν ο X είναι διαχωρίσιμος μετρικός χώρος [Άσκηση 2.10], το W έχει $\mu(W) = 0$ (και τέτοιες είναι όλες οι περιπτώσεις που θα μας απασχολήσουν). Τότε το στήριγμα είναι το μικρότερο κλειστό

¹Περίπου, γιατί δεξ τι συμβαίνει στο Παράδειγμα 2.16(ii) πιο κάτω.

Μέτρα

υποσύνολο του X στο οποίο το μ συγκεντρώνει τη μάζα του (δηλαδή δίνει μέτρο 0 στο συμπλήρωμά αυτού του συνόλου). Πράγματι, το $X \setminus W$ είναι κλειστό με συμπλήρωμα που έχει μέτρο 0, αλλά και αν F είναι ένα άλλο οποιοδήποτε κλειστό σύνολο με $\mu(X \setminus F) = 0$, τότε πρέπει $X \setminus F \subset W$ από τον ορισμό του W , άρα $X \setminus W \subset F$.

Είναι όμως επίσης δυνατόν να ισχύει $\mu(W) > 0$. Μάλιστα υπάρχει μέτρο πιθανότητας σε κατάλληλο χώρο για το οποίο ισχύει $W = X$, οπότε το στήριγμα του είναι το κενό σύνολο [είναι το μέτρο Dieudonné, δες Άσκηση 12.3 στο Κουμουλλής Γ. και Νεγρεπόντης Σ. (1991)].

Παράδειγμα 2.16. (i) Έστω $\lambda > 0$ και μ_1 το διακριτό μέτρο πιθανότητας στο \mathbb{R} που δίνει σε κάθε $k \in \mathbb{N}$ μάζα $e^{-\lambda} \lambda^k / k!$. Το στήριγμά του είναι το \mathbb{N} .

(ii) Έστω μ_2 το μέτρο στο \mathbb{R} με $\mu_2(A) = |A \cap \mathbb{Q}|$ για κάθε $A \subset \mathbb{R}$. Αυτό είναι ένα διακριτό μέτρο (δίνει μάζα 1 σε κάθε ρητό). Παρόλο που $\mu_2(\mathbb{R} \setminus \mathbb{Q}) = 0$ (δηλαδή η μάζα του μ_2 είναι συγκεντρωμένη στο \mathbb{Q}), το στήριγμα του μ_2 είναι το \mathbb{R} γιατί οποιοδήποτε ανοιχτό σύνολο γύρω από οποιονδήποτε πραγματικό αριθμό έχει θετικό μέτρο αφού περιέχει κάποιον ρητό.

Ασκήσεις

2.1 Έστω $\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_n, n \in \mathbb{N}^+$, μέτρα πιθανότητας στον (Ω, \mathcal{F}) και $\lambda_1, \lambda_2, \dots, \lambda_n \in [0, 1]$ με $\sum_{i=1}^n \lambda_i = 1$. Να δειχθεί ότι ο κυρτός συνδυασμός

$$\mathbf{Q} = \sum_{i=1}^n \lambda_i \mathbf{P}_i$$

των μέτρων $\mathbf{P}_i, i = 1, 2, \dots, n$, είναι μέτρο πιθανότητας στον (Ω, \mathcal{F}) .

2.2 Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας και $(A_n)_{n \geq 1}$ ακολουθία ξένων ανά δύο στοιχείων της \mathcal{F} . Να δείξετε ότι $\lim_{n \rightarrow \infty} \mathbf{P}(A_n) = 0$.

2.3 Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας και $(A_n)_{n \geq 1}$ ακολουθία στοιχείων της \mathcal{F} .

(α) Αν $\mathbf{P}(A_n) = 0$ για κάθε $n \geq 1$, τότε $\mathbf{P}(\cup_{n=1}^{\infty} A_n) = 0$.

(β) Αν $\mathbf{P}(A_n) = 1$ για κάθε $n \geq 1$, τότε $\mathbf{P}(\cap_{n=1}^{\infty} A_n) = 1$.

2.4 Να βρεθεί χώρος πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$ και $(A_i)_{i \in I}, (B_i)_{i \in I'}$ οικογένειες στοιχείων της \mathcal{F} ώστε

(α) $\mathbf{P}(A_i) = 0$ για κάθε $i \in I, \cup_{i \in I} A_i \in \mathcal{F}$, αλλά $\mathbf{P}(\cup_{i \in I} A_i) \neq 0$.

(β) $\mathbf{P}(B_i) = 1$ για κάθε $i \in I',$ αλλά $\cap_{i \in I'} B_i = \emptyset$.

2.5 Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας και $(A_\beta)_{\beta \in B}$ οικογένεια ξένων ανά δύο στοιχείων της \mathcal{F} . Αν $\mathbf{P}(A_\beta) > 0$ για κάθε $\beta \in B$, να δείξετε ότι το B είναι αριθμήσιμο.

2.6 Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας και $(A_n)_{n \geq 1}$ ακολουθία στοιχείων της \mathcal{F} . Να δειχθεί ότι

$$\mathbf{P}(\liminf_{n \geq 1} A_n) \leq \underline{\lim}_{n \rightarrow \infty} \mathbf{P}(A_n) \leq \overline{\lim}_{n \rightarrow \infty} \mathbf{P}(A_n) \leq \mathbf{P}(\limsup_{n \geq 1} A_n). \quad (2.6)$$

2.7 Έστω μ_1 το μέτρο στο Παράδειγμα 2.16(i) περιορισμένο στη σ -άλγεβρα $\mathcal{B}(\mathbb{R})$, λ το μέτρο Lebesgue στο \mathbb{R} , και $\nu : \mathcal{B}(\mathbb{R}) \rightarrow [0, \infty]$ το μέτρο που ορίζεται ως $\nu(A) = \lambda(A \cap [2, 5])$ για κάθε $A \in \mathcal{B}(\mathbb{R})$. Θέτουμε $\mu = \mu_1 + \nu$. Ποιο είναι το $\text{supp}(\mu)$;

2.8* Έστω $F \subset \mathbb{R}$ κλειστό. Να δειχθεί ότι υπάρχει μέτρο στο \mathbb{R} με στήριγμα το F .

2.9 Έστω $X := \mathbb{R}$ και \mathcal{A} η σ -άλγεβρα που ορίστηκε στο Παράδειγμα 1.5. Ορίζουμε $\mu : \mathcal{A} \rightarrow [0, \infty]$ με

$$\mu(A) = \begin{cases} 0 & \text{αν το } A \text{ είναι αριθμήσιμο,} \\ 1 & \text{αν το } \mathbb{R} \setminus A \text{ είναι αριθμήσιμο.} \end{cases}$$

Μέτρα

Να δειχθεί ότι το μ είναι μέτρο πιθανότητας στον (X, \mathcal{A}) .

2.10 Έστω X διαχωρίσιμος μετρικός χώρος.

(α) Να δειχθεί ότι υπάρχει αριθμήσιμη ακολουθία $(U_n)_{n \in \mathbb{N}}$ ανοιχτών υποσυνόλων του X ώστε για κάθε ανοιχτό $A \subset X$ να υπάρχει $I \subset \mathbb{N}$ με $A = \cup_{i \in I} U_i$.

[Υπόδειξη: Έστω $(x_i)_{i \in \mathbb{N}}$ αριθμήσιμο πυκνό υποσύνολο του X . Θεωρούμε το σύνολο των σφαιρών $\{B(x_i, 1/j) : i \in \mathbb{N}, j \in \mathbb{N}^+\}$.]

(β) Έστω μ ένα μέτρο στον $(X, \mathcal{B}(X))$ και $W := \cup\{V : V \subset X \text{ ανοιχτό με } \mu(V) = 0\}$. Να δειχθεί ότι $\mu(W) = 0$.

3

Ισότητα πεπερασμένων μέτρων

Στο κεφάλαιο αυτό θα δούμε ένα τεχνικό αποτέλεσμα, το λεγόμενο θεώρημα π-λ, που στόχο έχει να διευκολύνει την απόδειξη ιδιοτήτων για τα στοιχεία μιας σ-άλγεβρας.

3.1 Κλάσεις Dynkin

Ορισμός 3.1. Έστω X σύνολο. Μια οικογένεια $\mathcal{D} \subset \mathcal{P}(X)$ λέγεται κλάση **Dynkin** στο X αν έχει τις εξής ιδιότητες:

- (i) $X \in \mathcal{D}$.
- (ii) Αν $A, B \in \mathcal{D}$ και $A \subset B$, τότε $B \setminus A \in \mathcal{D}$.
- (iii) Αν η $(A_n)_{n \in \mathbb{N}}$ είναι αύξουσα ακολουθία στην \mathcal{D} , τότε $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{D}$.

Παρατήρηση 3.2. Κάθε σ-άλγεβρα είναι κλάση Dynkin. Το αντίστροφο όμως δεν ισχύει (Άσκηση 3.2). Δηλαδή είναι ευκολότερο ένα σύνολο να είναι κλάση Dynkin.

Όπως και στην περίπτωση των σ-αλγεβρών, για κάθε οικογένεια $C \subset \mathcal{P}(X)$ ενός συνόλου X , υπάρχει η ελάχιστη κλάση Dynkin που την περιέχει. Αυτή περιγράφεται ως η τομή όλων των κλάσεων Dynkin που περιέχουν τη C . Συνήθως τη συμβολίζουμε με $\delta(C)$. Συνοψίζοντας παίρνουμε την ακόλουθη πρόταση η απόδειξη της οποίας είναι παρόμοια με αυτή στην περίπτωση των σ-αλγεβρών.

Πρόταση 3.3. Έστω X σύνολο και $C \subset \mathcal{P}(X)$. Θέτουμε

$$\mathcal{J} := \{\mathcal{D} \subset \mathcal{P}(X) : \mathcal{D} \text{ είναι κλάση Dynkin και } C \subset \mathcal{D}\}.$$

Τότε η οικογένεια

$$\delta(C) := \bigcap_{\mathcal{D} \in \mathcal{J}} \mathcal{D}$$

- είναι κλάση Dynkin στο X και
- είναι η μικρότερη κλάση Dynkin που περιέχει τη C . Δηλαδή περιέχεται σε κάθε κλάση Dynkin που περιέχει τη C .

Ονομάζουμε τη $\delta(C)$ κλάση Dynkin που παράγεται από την C .

Παρατήρηση 3.4. Εύκολα παρατηρεί κανείς ότι $\delta(C) \subset \sigma(C)$ εφόσον η $\sigma(C)$ είναι κλάση Dynkin και περιέχει τη C .

Λέμε ότι μια οικογένεια C συνόλων είναι *κλειστή στις πεπερασμένες τομές* αν για κάθε $n \in \mathbb{N}^+$ και $A_1, A_2, \dots, A_n \in C$ ισχύει ότι $A_1 \cap A_2 \cap \dots \cap A_n \in C$. Προφανώς αρκεί να ισχύει η συνθήκη αυτή για $n = 2$ και έπειτα οι υπόλοιπες περιπτώσεις αποδεικνύονται επαγωγικά.

Η επόμενη πρόταση δίνει μια απλή συνθήκη ώστε μια κλάση Dynkin να είναι σ-άλγεβρα.

Πρόταση 3.5. Έστω X σύνολο και \mathcal{D} κλάση Dynkin στο X . Αν η \mathcal{D} είναι κλειστή στις πεπερασμένες τομές, τότε η \mathcal{D} είναι σ -άλγεβρα στο X .

Απόδειξη. Από τον Ορισμό 3.1 της κλάσης Dynkin έχουμε ότι $X \in \mathcal{D}$ λόγω του (i) και, αν $A \in \mathcal{D}$, τότε $X \setminus A \in \mathcal{D}$ λόγω των (i) και (ii). Επίσης, η \mathcal{D} είναι κλειστή στις πεπερασμένες ενώσεις εφόσον είναι κλειστή στις πεπερασμένες τομές και στα συμπληρώματα. Μένει να δείξουμε ότι είναι κλειστή τις αριθμήσιμες ενώσεις. Αν $(A_n)_{n \geq 1}$ είναι στοιχεία της \mathcal{D} τότε για κάθε $n \in \mathbb{N}^+$ το $B_n := \cup_{k=1}^n A_k \in \mathcal{D}$ και η $(B_n)_{n \geq 1}$ είναι αύξουσα ακολουθία στην \mathcal{D} . Άρα $\cup_{n=1}^{\infty} B_n \in \mathcal{D}$. Όμως $\cup_{n=1}^{\infty} B_n = \cup_{n=1}^{\infty} A_n$. ■

3.2 Το θεώρημα π-λ

Θεώρημα 3.6 (Θεώρημα π-λ). Έστω X σύνολο και $C \subset \mathcal{P}(X)$ οικογένεια κλειστή στις πεπερασμένες τομές. Τότε $\delta(C) = \sigma(C)$.

Απόδειξη. Ισχύει ότι $\delta(C) \subset \sigma(C)$. Άρα, αν δείξουμε ότι η $\delta(C)$ είναι σ -άλγεβρα, τότε $\sigma(C) \subset \delta(C)$, εφόσον η $\sigma(C)$ είναι η ελάχιστη σ -άλγεβρα που περιέχει τη C . Με βάση την Πρόταση 3.5, αρκεί να δείξουμε ότι η $\delta(C)$ είναι κλειστή στις πεπερασμένες τομές, δηλαδή

$$A, B \in \delta(C) \Rightarrow A \cap B \in \delta(C). \quad (3.1)$$

Γνωρίζουμε το συμπέρασμα της συνεπαγωγής για $A, B \in C$, οπότε το σχέδιο είναι να την ενισχύσουμε σε δύο βήματα. Δηλαδή να δείξουμε ότι ισχύει πρώτα για $A \in C, B \in \delta(C)$ και έπειτα για $A \in \delta(C), B \in \delta(C)$.

Για κάθε $A \subset X$ θέτουμε

$$\mathcal{D}(A) = \{U \in \delta(C) : A \cap U \in \delta(C)\}.$$

Αυτό το σύνολο περιέχει τα σύνολα που «τέμνονται ωραία» με το A .

Βήμα 1. Για $A \in C$, έχουμε:

- $C \subset \mathcal{D}(A)$ εφόσον η C είναι κλειστή στις πεπερασμένες τομές.
- Η $\mathcal{D}(A)$ είναι κλάση Dynkin (η απόδειξη αφήνεται ως άσκηση).

Άρα $\mathcal{D}(A) \supset \delta(C)$, εφόσον η $\delta(C)$ είναι η ελάχιστη κλάση Dynkin που περιέχει τη C . Όμως $\mathcal{D}(A) \subset \delta(C)$ από τον ορισμό της $\mathcal{D}(A)$. Τελικά $\mathcal{D}(A) = \delta(C)$, που σημαίνει ότι

$$A \in C, B \in \delta(C) \Rightarrow A \cap B \in \delta(C). \quad (3.2)$$

Βήμα 2. Για $B \in \delta(C)$, έχουμε:

- $C \subset \mathcal{D}(B)$ από την (3.2).
- Η $\mathcal{D}(B)$ είναι κλάση Dynkin (η απόδειξη αφήνεται ως άσκηση).

Άρα, όπως στο Βήμα 1, έχουμε ότι $\mathcal{D}(B) = \delta(C)$, δηλαδή ισχύει η (3.1), και το θεώρημα αποδείχθηκε. ■

Μια οικογένεια $C \subset \mathcal{P}(X)$ λέγεται **π-σύστημα** αν είναι κλειστή στις πεπερασμένες τομές, ενώ λέγεται **λ-σύστημα** αν είναι κλάση Dynkin. Σε αυτή την ορολογία οφείλεται το όνομα του προηγούμενου θεωρήματος. Μια ισοδύναμη διατύπωσή του δίνεται στην Άσκηση 3.3.

Χαρακτηριστική εφαρμογή του θεωρήματος είναι η εξής: Για $X = \mathbb{R}$, η οικογένεια $C = \{(-\infty, x] : x \in \mathbb{R}\}$ είναι κλειστή στις πεπερασμένες τομές και άρα $\delta(C) = \sigma(C)$. Ξέρουμε όμως ότι $\sigma(C) = \mathcal{B}(\mathbb{R})$. Άρα παίρνουμε ακόμα μία περιγραφή των συνόλων Borel ως $\delta(C)$.

Μια σημαντική συνέπεια του θεωρήματος π-λ είναι το ακόλουθο αποτέλεσμα.

Πόρισμα 3.7. Έστω X σύνολο, $\mathcal{A} \subset \mathcal{P}(X)$ σ -άλγεβρα, $\mathcal{C} \subset \mathcal{A}$ οικογένεια κλειστή στις πεπερασμένες τομές ώστε $\sigma(\mathcal{C}) = \mathcal{A}$, και μ, ν πεπερασμένα μέτρα στον (X, \mathcal{A}) με $\mu(X) = \nu(X)$. Αν $\mu(A) = \nu(A)$ για κάθε $A \in \mathcal{C}$, τότε $\mu = \nu$ στην \mathcal{A} .

Απόδειξη. Έστω $\mathcal{B} = \{A \in \mathcal{A} : \mu(A) = \nu(A)\}$. Τότε,

- $\mathcal{C} \subset \mathcal{B} \subset \sigma(\mathcal{C})$.
- Η \mathcal{B} είναι κλάση Dynkin.

Πράγματι, ο πρώτος ισχυρισμός είναι προφανής και για τον δεύτερο έχουμε

- (i) $X \in \mathcal{B}$ από υπόθεση.
- (ii) Αν $A, B \in \mathcal{B}$, με $A \subset B$, τότε $\mu(B \setminus A) = \mu(B) - \mu(A) = \nu(B) - \nu(A) = \nu(B \setminus A)$.
- (iii) Αν $(A_n)_{n \in \mathbb{N}}$ είναι αύξουσα ακολουθία στη \mathcal{B} , τότε

$$\mu(\cup_{n \in \mathbb{N}} A_n) = \lim_{n \rightarrow \infty} \mu(A_n) = \lim_{n \rightarrow \infty} \nu(A_n) = \nu(\cup_{n \in \mathbb{N}} A_n).$$

Άρα $\cup_{n \in \mathbb{N}} A_n \in \mathcal{B}$.

Εφόσον η \mathcal{B} είναι κλάση Dynkin, έχουμε ότι $\delta(\mathcal{C}) \subset \mathcal{B} \subset \mathcal{A}$. Όμως από το θεώρημα π-λ, $\delta(\mathcal{C}) = \sigma(\mathcal{C}) (= \mathcal{A}$ από υπόθεση), και τελικά $\mathcal{B} = \mathcal{A}$, από το οποίο προκύπτει το ζητούμενο. ■

Το προηγούμενο πόρισμα δίνει ότι αν έχουμε δύο μέτρα πιθανότητας μ, ν στον $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ για τα οποία ισχύει $\mu((-\infty, x]) = \nu((-\infty, x])$ για κάθε $x \in \mathbb{R}$, τότε $\mu = \nu$.

Ασκήσεις

3.1 Έστω $(\Omega, \mathcal{A}, \mathbf{P})$ χώρος πιθανότητας και $U \in \mathcal{A}$ δεδομένο. Θέτουμε

$$\mathcal{C} := \{A \in \mathcal{A} : \mathbf{P}(A \cap U) = \mathbf{P}(A)\mathbf{P}(U)\}.$$

Να δειχθεί ότι η \mathcal{C} είναι κλάση Dynkin.

3.2 Έστω $\Omega := \{1, 2, 3, 4\}$ και

$$\mathcal{A} := \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{1, 4\}, \{1, 2, 3, 4\}, \emptyset\}.$$

Να δειχθεί ότι η \mathcal{A} είναι κλάση Dynkin αλλά δεν είναι σ -άλγεβρα στο Ω .

3.3 Έστω X σύνολο, $\mathcal{C}_1 \subset \mathcal{C}_2 \subset \mathcal{P}(X)$ ώστε η \mathcal{C}_1 να είναι κλειστή στις πεπερασμένες τομές και η \mathcal{C}_2 να είναι κλάση Dynkin. Να δειχθεί ότι $\sigma(\mathcal{C}_1) \subset \mathcal{C}_2$.

3.4 Έστω X ένα μη κενό σύνολο. Να αποδειχθεί ότι ένα $\mathcal{A} \subset \mathcal{P}(X)$ είναι κλάση Dynkin αν και μόνο αν ικανοποιούνται οι εξής τρεις συνθήκες.

- (i) $X \in \mathcal{A}$.
- (ii) Αν $A \in \mathcal{A}$ τότε $X \setminus A \in \mathcal{A}$.
- (iii) Αν $(A_n)_{n \in \mathbb{N}^+}$ είναι ακολουθία ξένων ανα δύο στοιχείων της \mathcal{A} , τότε $\cup_{n=1}^{\infty} A_n \in \mathcal{A}$.

4

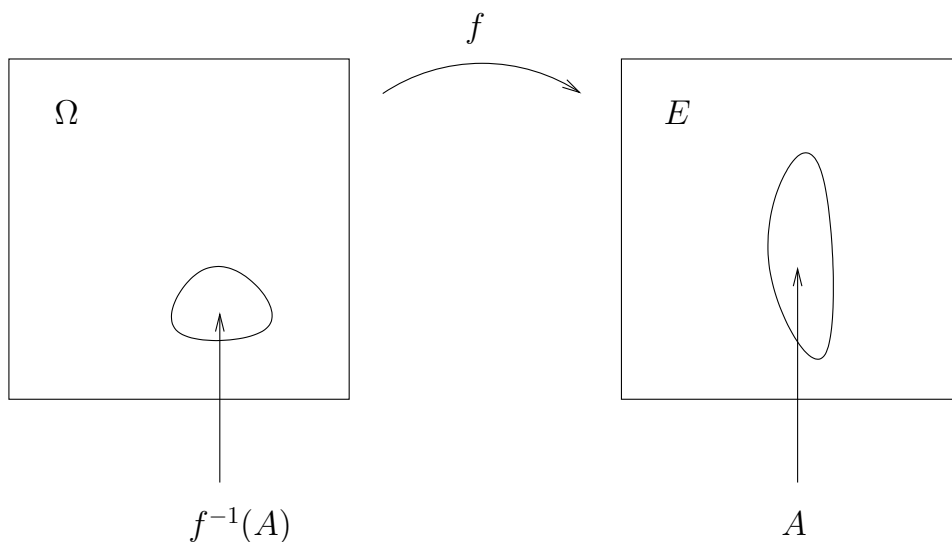
Μετρήσιμες συναρτήσεις

4.1 Μετρήσιμες συναρτήσεις

Ορισμός 4.1. Έστω (Ω, \mathcal{F}) , (E, \mathcal{E}) μετρήσιμοι χώροι. Μια συνάρτηση $f : \Omega \rightarrow E$ λέγεται \mathcal{F}/\mathcal{E} -μετρήσιμη αν

$$f^{-1}(A) \in \mathcal{F} \quad \text{για κάθε } A \in \mathcal{E}. \quad (4.1)$$

Συμβολίζουμε το σύνολο $\{f^{-1}(A) : A \in \mathcal{E}\}$ με $f^{-1}(\mathcal{E})$. Οπότε η απαίτηση του ορισμού της μετρησιμότητας γράφεται $f^{-1}(\mathcal{E}) \subset \mathcal{F}$.



Σχήμα 4.1: Μια f όπως στον Ορισμό 4.1

Ορολογία: 1. Μια \mathcal{F}/\mathcal{E} -μετρήσιμη συνάρτηση τη λέμε \mathcal{F} -μετρήσιμη ή \mathcal{E} -μετρήσιμη ή απλώς μετρήσιμη αν είναι σαφές ποια είναι η σ -άλγεβρα που δεν αναφέρουμε.

2. Όταν ο χώρος Ω ή/και ο E είναι μετρικός χώρος (π.χ. υποσύνολο ενός από τους χώρους \mathbb{R}^d , $[-\infty, \infty]$, $[0, \infty]$, \mathbb{C}), εκτός αν αναφέρεται κάτι διαφορετικό, θα θεωρείται ότι η σ -άλγεβρα στον Ω και στον E είναι η σ -άλγεβρα των υποσυνόλων Borel του Ω και του E . Και τότε, π.χ., \mathcal{F} -μετρήσιμη σημαίνει $\mathcal{F}/\mathcal{B}(E)$ μετρήσιμη. Στην περίπτωση που ο Ω είναι μετρικός χώρος και η \mathcal{E} εννοείται, ονομάζουμε **Borel-μετρήσιμη** κάθε f η οποία είναι $\mathcal{B}(\Omega)/\mathcal{E}$ -μετρήσιμη.

3. Όταν $(\Omega, \mathcal{F}) = (\mathbb{R}, \mathcal{M})$ και η \mathcal{E} εννοείται, τότε κάθε \mathcal{M}/\mathcal{E} -μετρήσιμη συνάρτηση λέγεται **Lebesgue-μετρήσιμη** συνάρτηση. Προφανώς κάθε Borel-μετρήσιμη συνάρτηση είναι Lebesgue-μετρήσιμη αφού $\mathcal{B}(\mathbb{R}) \subset \mathcal{M}$. [Δες Παράγραφο 2.2 για τον ορισμό του \mathcal{M} .]

4. Σε έναν χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$, κάθε μετρήσιμη συνάρτηση λέγεται **τυχαία μεταβλητή**. Συμβολίζουμε τις τυχαίες μεταβλητές με κεφαλαία γράμματα X, Y, \dots , σε αντίθεση με τη σύμβαση που υιοθετούμε στον απειροστικό λογισμό και την πραγματική ανάλυση.

Για το σύνολο $f^{-1}(A) := \{\omega \in \Omega : f(\omega) \in A\}$ συνήθως χρησιμοποιούμε τον συμβολισμό $\{f \in A\}$. Όμοια, αν $E = \mathbb{R}$, το $\{f < a\}$ συμβολίζει το σύνολο $\{\omega \in \Omega : f(\omega) < a\}$ και $\{f^2 < f + 1\}$ το $\{\omega \in \Omega : f^2(\omega) < f(\omega) + 1\}$.

Παρατήρηση 4.2. Γιατί απαιτούμε από μια συνάρτηση $X : \Omega \rightarrow E$ να έχει την ιδιότητα (4.1); Γιατί, όταν ορίσουμε ένα μέτρο πιθανότητας \mathbf{P} στην \mathcal{F} , θέλουμε να μπορούμε να εξετάζουμε πιθανότητες της μορφής $\mathbf{P}(X \in A)$, όπου $A \in \mathcal{E}$, δηλαδή $\mathbf{P}(X^{-1}(A))$. Πρέπει επομένως το $X^{-1}(A)$ να ανήκει στο πεδίο ορισμού της \mathbf{P} , το οποίο είναι η \mathcal{F} .

Κάθε σταθερή συνάρτηση $f : \Omega \rightarrow E$ μεταξύ δύο μετρήσιμων χώρων (Ω, \mathcal{F}) , (E, \mathcal{E}) είναι μετρήσιμη. Γιατί αν $f(x) = c$ για κάθε $x \in \Omega$, όπου $c \in E$ δεδομένο, τότε για οποιοδήποτε $A \in \mathcal{E}$ έχουμε

$$f^{-1}(A) = \begin{cases} \Omega & \text{αν } c \in A, \\ \emptyset & \text{αν } c \in E \setminus A, \end{cases}$$

το οποίο είναι πάντοτε στοιχείο της \mathcal{F} .

Πρόταση 4.3. Έστω (Ω, \mathcal{F}) , (E, \mathcal{E}) μετρήσιμοι χώροι, $f : \Omega \rightarrow E$ συνάρτηση, και $C \subset \mathcal{E}$ οικογένεια ώστε $\sigma(C) = \mathcal{E}$. Τότε $f^{-1}(C) \subset \mathcal{F}$ αν και μόνο αν $f^{-1}(\mathcal{E}) \subset \mathcal{F}$.

Βέβαια ο έλεγχος $f^{-1}(C) \subset \mathcal{F}$ είναι ευκολότερος από τον $f^{-1}(\mathcal{E}) \subset \mathcal{F}$. Έτσι η πρόταση κάνει ευκολότερο τον έλεγχο της μετρησιμότητας μιας συνάρτησης.

Απόδειξη. \Rightarrow Έστω $\mathcal{B} := \{A \in \mathcal{E} : f^{-1}(A) \in \mathcal{F}\}$. Τότε η $C \subset \mathcal{B}$ και εύκολα βλέπουμε ότι η \mathcal{B} είναι σ-άλγεβρα (Άσκηση 1.7). Συνεπώς, $\sigma(C) \subset \mathcal{B}$. Όμως $\sigma(C) = \mathcal{E}$ και $\mathcal{B} \subset \mathcal{E}$. Άρα $\mathcal{B} = \mathcal{E}$.

\Leftarrow Προφανές αφού $C \subset \mathcal{E}$. ■

Ακολουθούν δύο συνέπειες της πρότασης.

Πόρισμα 4.4. Έστω (Ω, \mathcal{F}) μετρήσιμος χώρος και $f : \Omega \rightarrow \mathbb{R}$ συνάρτηση. Τότε η f είναι μετρήσιμη αν και μόνο αν $f^{-1}((-\infty, a]) \in \mathcal{F}$ για κάθε $a \in \mathbb{R}$.

Απόδειξη. Αν $C = \{(-\infty, a] : a \in \mathbb{R}\}$, γνωρίζουμε ότι $\sigma(C) = \mathcal{B}(\mathbb{R})$. Άρα, από την Πρόταση 4.3 προκύπτει το ζητούμενο. ■

Το ίδιο αποτέλεσμα ισχύει αν αντικαταστήσουμε τα διαστήματα $(-\infty, a]$ με $(-\infty, a)$ ή γενικά με οποιαδήποτε οικογένεια διαστημάτων που παράγουν την $\mathcal{B}(\mathbb{R})$. Επίσης, αντίστοιχο συμπέρασμα προκύπτει αν έχουμε μετρήσιμη συνάρτηση με τιμές στο $[-\infty, \infty]$.

Υπάρχουν μετρήσιμες συναρτήσεις; Είναι πολλές; Καταρχάς, θα δούμε αμέσως ότι όλες οι συνεχείς συναρτήσεις είναι μετρήσιμες. Υπενθυμίζουμε ότι, αν (Ω, d_1) , (E, d_2) μετρικοί χώροι, μια συνάρτηση $f : X \rightarrow Y$ είναι συνεχής αν για κάθε $V \subset E$ ανοιχτό έχουμε ότι $f^{-1}(V)$ είναι ανοιχτό. Δηλαδή αν η αντίστροφη εικόνα κάθε ανοιχτού συνόλου είναι ανοιχτό σύνολο.

Πόρισμα 4.5. Έστω (Ω, d_1) , (E, d_2) μετρικοί χώροι και $f : \Omega \rightarrow E$ συνεχής συνάρτηση. Αν $\mathcal{F} = \mathcal{B}(\Omega)$ και $\mathcal{E} = \mathcal{B}(E)$, τότε η f είναι \mathcal{F}/\mathcal{E} μετρήσιμη.

Απόδειξη. Για την οικογένεια \mathcal{S} των ανοιχτών συνόλων του E έχουμε ότι $\sigma(\mathcal{S}) = \mathcal{E}$ και όλα τα στοιχεία του $f^{-1}(\mathcal{S})$ είναι ανοιχτά σύνολα (αφού η f είναι συνεχής) και άρα $f^{-1}(\mathcal{S}) \subset \mathcal{F}$. Το συμπέρασμα έπεται από την Πρόταση 4.3. ■

Παραθέτουμε χωρίς απόδειξη τις βασικές ιδιότητες κλειστότητας του συνόλου των μετρήσιμων συναρτήσεων. Εν ολίγοις, αν ξεκινήσει κανείς με μετρήσιμες συναρτήσεις και τις συνδυάσει με κάποιον «φυσιολογικό» τρόπο, προκύπτουν πάλι μετρήσιμες συναρτήσεις.

Πρόταση 4.6. Έστω $f, g : \Omega \rightarrow [-\infty, \infty]$ μετρήσιμες συναρτήσεις στον μετρήσιμο χώρο (Ω, \mathcal{F}) και $a \in \mathbb{R}$. Τότε μετρήσιμες είναι επίσης οι συναρτήσεις

$$af, |f|, f + g, fg, f/g, \min\{f, g\}, \max\{f, g\}, f^+, f^-,$$

όπου καθεμία ορίζεται έτσι ώστε να είναι σταθερή και ίση με μία αυθαίρετη πεπερασμένη σταθερά στο σύνολο των σημείων που ο τύπος της δεν μπορεί να την ορίσει (δηλαδή σε σημεία που εμφανίζονται εκφράσεις της μορφής $1/0, 0/0, \infty/0, \infty/\infty, \infty - \infty, 0 \cdot \infty$).

Πρόταση 4.7. Έστω $(f_n)_{n \geq 1}$ ακολουθία μετρήσιμων συναρτήσεων στον μετρήσιμο χώρο (Ω, \mathcal{F}) και με τιμές στο $[-\infty, \infty]$. Τότε:

(i) Οι συναρτήσεις

$$\inf_{n \geq 1} f_n, \sup_{n \geq 1} f_n, \liminf_{n \rightarrow \infty} f_n, \overline{\lim}_{n \rightarrow \infty} f_n,$$

είναι επίσης μετρήσιμες.

(ii) Αν η $(f_n)_{n \geq 1}$ συγκλίνει σημειακά σε μια συνάρτηση f , τότε η $f = \lim_{n \rightarrow \infty} f_n$ είναι μετρήσιμη συνάρτηση.

Αντιπαραβάλετε την προηγούμενη πρόταση με το γεγονός ότι γενικά το σημειακό όριο συνεχών συναρτήσεων δεν είναι συνεχής συνάρτηση. Η μετρησιμότητα είναι πιο ανθεκτική σε μετασχηματισμούς.

Πρόταση 4.8. Έστω $(\Omega, \mathcal{F}), (E, \mathcal{E}), (G, \mathcal{G})$ μετρήσιμοι χώροι και $f : \Omega \rightarrow E, g : E \rightarrow G$ μετρήσιμες συναρτήσεις. Τότε η $g \circ f : \Omega \rightarrow G$ είναι \mathcal{F}/\mathcal{G} μετρήσιμη.

Απόδειξη. Έστω $A \in \mathcal{G}$. Τότε, $(g \circ f)^{-1}(A) = f^{-1}(g^{-1}(A))$. Όμως $g^{-1}(A) \in \mathcal{E}$, άρα $f^{-1}(g^{-1}(A)) \in \mathcal{F}$, από το οποίο προκύπτει το ζητούμενο. ■

Πρόταση 4.9. Έστω (Ω, \mathcal{F}) μετρήσιμος χώρος. Τότε:

(i) Για $A \subset \Omega$, η $\mathbf{1}_A$ είναι μετρήσιμη αν και μόνο αν $A \in \mathcal{F}$.

(ii) Αν $f_1, f_2, \dots, f_n : \Omega \rightarrow \mathbb{R}, n \geq 1$ είναι $\mathcal{F}/\mathcal{B}(\mathbb{R})$ μετρήσιμες συναρτήσεις και $g : \mathbb{R}^n \rightarrow \mathbb{R}$ είναι $\mathcal{B}(\mathbb{R}^n)/\mathcal{B}(\mathbb{R})$ μετρήσιμη, τότε η $g(f_1, f_2, \dots, f_n) : \Omega \rightarrow \mathbb{R}$ είναι μετρήσιμη.

Απόδειξη. Θα δείξουμε μόνο το (i). Η απόδειξη του (ii) περιγράφεται στην Άσκηση 9.5. Αν $B \in \mathcal{B}(\mathbb{R})$, έχουμε

$$(\mathbf{1}_A)^{-1}(B) = \begin{cases} \emptyset & \text{αν } 0, 1 \notin B, \\ \Omega \setminus A & \text{αν } 0 \in B, 1 \notin B, \\ A & \text{αν } 0 \notin B, 1 \in B, \\ \Omega & \text{αν } 0, 1 \in B. \end{cases} \quad (4.2)$$

Αν η $\mathbf{1}_A$ είναι μετρήσιμη, τότε για $B = \{1\}$, έχουμε $(\mathbf{1}_A)^{-1}(B) \in \mathcal{F}$, δηλαδή $A \in \mathcal{F}$. Αντίστροφα, αν $A \in \mathcal{F}$, τότε από την (4.2) έχουμε $(\mathbf{1}_A)^{-1}(B) \in \mathcal{F}$ για κάθε $B \in \mathcal{B}(\mathbb{R})$. ■

Παράδειγμα 4.10. Έστω $(f_n)_{n \geq 1}$ ακολουθία μετρήσιμων συναρτήσεων σε μετρήσιμο χώρο (Ω, \mathcal{F}) με τιμές στο \mathbb{R} . Θέτουμε $T := \min\{k \in \mathbb{N}^+ : f_k > 0\}$ με τη σύμβαση $\min \emptyset = \infty$. Τότε η T είναι μετρήσιμη γιατί για $k \in \mathbb{N}^+$ ισχύει

$$\{T \leq k\} = \{f_1 > 0\} \cup \{f_2 > 0\} \cup \dots \cup \{f_k > 0\} \in \mathcal{F}.$$

Για k μη θετικό ακέραιο έχουμε $\{T \leq k\} = \emptyset$, ενώ για κάθε πραγματικό x έχουμε $\{T \leq x\} = \{T \leq [x]\}$.

Επίσης, για οποιοδήποτε $n \geq 1$, η $\cos(f_1 + f_2 + \dots + f_n)$ είναι μετρήσιμη λόγω του (ii) της προηγούμενης πρότασης και του ότι η $(x_1, x_2, \dots, x_n) \rightarrow \cos(x_1 + x_2 + \dots + x_n)$ είναι συνεχής.

Ορισμός 4.11. Μια συνάρτηση $f : \Omega \rightarrow \mathbb{R}$ λέγεται **απλή** αν η εικόνα της είναι πεπερασμένο σύνολο.

Αν οι διαφορετικές τιμές που παίρνει μια απλή συνάρτηση f είναι a_1, a_2, \dots, a_n και θέσουμε $A_i := f^{-1}(\{a_i\})$, τότε η $\{A_1, A_2, \dots, A_n\}$ είναι διαμέριση του Ω , και η f γράφεται

$$f = \sum_{i=1}^n a_i \mathbf{1}_{A_i}. \quad (4.3)$$

Προφανώς μια απλή f είναι μετρήσιμη αν και μόνο αν τα σύνολα A_1, A_2, \dots, A_n είναι μετρήσιμα.

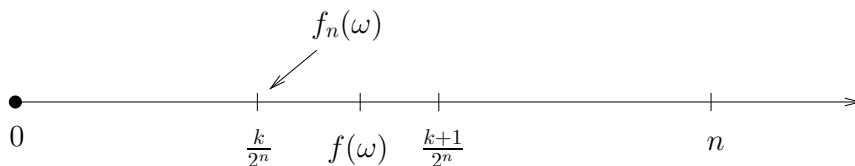
Μια απλή συνάρτηση δεν γράφεται μοναδικά ως γραμμικός συνδυασμός από δείκτριες συναρτήσεις. Αν τα A_1, A_2, \dots, A_n δεν είναι απαραίτητα ξένα, τότε η σχέση (4.3) ορίζει πάλι μια απλή συνάρτηση. Αν όμως ζητήσουμε τα A_1, A_2, \dots, A_n να είναι διαμέριση του Ω (δηλαδή μη κενά, ξένα ανά δύο, με ένωση το Ω) και οι αριθμοί a_1, \dots, a_n διαφορετικοί μεταξύ τους, τότε η γραφή (4.3) είναι μοναδική (με μόνη ελευθερία στη σειρά με την οποία αριθμούμε τα σύνολα και τους αριθμούς) και ονομάζεται **κανονική μορφή** της f .

Πρόταση 4.12. Έστω $f : \Omega \rightarrow [0, \infty]$ μετρήσιμη συνάρτηση στον μετρήσιμο χώρο (Ω, \mathcal{F}) . Τότε υπάρχει μια αύξουσα ακολουθία $(f_n)_{n \geq 1}$ μη αρνητικών, απλών, μετρήσιμων συναρτήσεων, $f_n : \Omega \rightarrow [0, \infty)$ για κάθε $n \geq 1$, ώστε $f = \lim_{n \rightarrow \infty} f_n$ κατά σημείο.

Το ότι η ακολουθία $(f_n)_{n \geq 1}$ είναι αύξουσα σημαίνει ότι $f_n(\omega) \leq f_{n+1}(\omega)$ για κάθε $\omega \in \Omega$ και $n \geq 1$.

Απόδειξη. Για $n \geq 1$, θέτουμε

$$f_n(\omega) := \begin{cases} \frac{k}{2^n} & \text{αν } f(\omega) \in \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right) \text{ με } k \in \mathbb{N}, 0 \leq k \leq n2^n - 1, \\ n & \text{αν } f(\omega) \geq n. \end{cases}$$



Σχήμα 4.2: Ο ορισμός της προσέγγισης f_n . Όλες οι τιμές πάνω από n απεικονίζονται στο n . Στο διάστημα $[0, n]$ η προσέγγιση γίνεται με λάθος το πολύ $1/2^n$.

Κάθε f_n είναι μη αρνητική, μετρήσιμη, και απλή αφού το σύνολο τιμών της είναι πεπερασμένο, και παίρνει την τιμή $k/2^n$, όπου $0 \leq k \leq n2^n - 1$, στο μετρήσιμο σύνολο $f^{-1}([k2^{-n}, (k+1)2^{-n}))$ και την τιμή n στο $f^{-1}([n, \infty))$.

Για το $f = \lim_{n \rightarrow \infty} f_n$. Αν $f(\omega) < \infty$, παίρνουμε φυσικό $n_0 > f(\omega)$. Για $n \geq n_0$ έχουμε $f_n(\omega) \leq f(\omega) \leq f_n(\omega) + 2^{-n}$, άρα $|f_n(\omega) - f(\omega)| < 2^{-n}$ και $\lim_{n \rightarrow \infty} f_n(\omega) = f(\omega)$. Αν $f(\omega) = \infty$, τότε $f_n(\omega) = n \rightarrow \infty$ για $n \rightarrow \infty$.

Για το ότι η ακολουθία είναι αύξουσα, παρατηρούμε τα εξής:

- Αν $f(\omega) = \infty$, τότε $f_n(\omega) = n$, που είναι αύξουσα ακολουθία.
- Αν $f(\omega) < \infty$, έστω $n \geq 1$, θα δείξουμε ότι $f_n(\omega) \leq f_{n+1}(\omega)$. Έχουμε τις εξής περιπτώσεις:

(α) $f(\omega) < n$.

(β) $f(\omega) \in [n, n+1)$.

(γ) $f(\omega) \geq n+1$.

Για το (α) παρατηρούμε ότι το $f_n(\omega)$ θα ισούται με το αριστερό άκρο του διαστήματος $[k2^{-n}, (k+1)2^{-n})$ στο οποίο ανήκει το $f(\omega)$. Για τον καθορισμό του $f_{n+1}(\omega)$, χωρίζουμε το $[k2^{-n}, (k+1)2^{-n})$ σε δύο μισά, τα

$$\left[\frac{2k}{2^{n+1}}, \frac{2k+1}{2^{n+1}} \right), \left[\frac{2k+1}{2^{n+1}}, \frac{2k+2}{2^{n+1}} \right)$$

και το $f_{n+1}(\omega)$ ισούται με το αριστερό άκρο του μισού στο οποίο ανήκει το $f(\omega)$. Άρα είναι τουλάχιστον $k2^{-n} = f_n(\omega)$. Οι περιπτώσεις (β) και (γ) αφήνονται ως άσκηση. ■

4.2 Σ-άλγεβρα παραγόμενη από συναρτήσεις

Ορισμός 4.13. Έστω Ω σύνολο. Για μια συνάρτηση $f : \Omega \rightarrow \bar{\mathbb{R}}$, σ-άλγεβρα παραγόμενη από την f ονομάζουμε το σύνολο

$$\sigma(f) := \{f^{-1}(A) : A \in \mathcal{B}(\bar{\mathbb{R}})\} = f^{-1}(\mathcal{B}(\bar{\mathbb{R}})).$$

Το ότι αυτό το σύνολο είναι σ-άλγεβρα το έχουμε δει στην Άσκηση 1.7 (β). Αυτή είναι η ελάχιστη σ-άλγεβρα \mathcal{A} στο Ω η οποία κάνει την f μετρήσιμη στον (Ω, \mathcal{A}) . Βέβαια, αν η f είναι μετρήσιμη στον (Ω, \mathcal{F}) , τότε θα έχουμε $\sigma(f) \subset \mathcal{F}$.

Παράδειγμα 4.14. Η $f : \mathbb{R} \rightarrow \mathbb{R}$ με $f(x) = -\mathbf{1}_{x < 0} + \mathbf{1}_{x \geq 0}$ παράγει τη σ-άλγεβρα

$$\{\mathbb{R}, (-\infty, 0), [0, \infty), \emptyset\}$$

αφού παίρνει μόνο τις τιμές $-1, 1$ και οι αντίστροφες εικόνες αυτών των τιμών είναι τα διαστήματα $(-\infty, 0), [0, \infty)$ αντίστοιχα. Οι λεπτομέρειες της απόδειξης αφήνονται ως άσκηση.

Παράδειγμα 4.15. Η συνάρτηση ακέραιο μέρος $f(x) = [x]$ για κάθε $x \in \mathbb{R}$ παράγει τη σ-άλγεβρα $\sigma(f) = \sigma(C)$ όπου $C := \{[k, k+1) : k \in \mathbb{Z}\}$ [Ο ισχυρισμός αυτός αφήνεται ως άσκηση. Παρατηρούμε ότι η f παίρνει τιμές στο \mathbb{Z} και $f^{-1}(\{k\}) = [k, k+1)$ για κάθε $k \in \mathbb{Z}$]. Η C είναι μια διαμέριση του \mathbb{R} .

Παράδειγμα 4.16. Παίρνουμε $\Omega = \{-1, 1\}^{\mathbb{N}^+}$. Μπορούμε να δούμε αυτό το σύνολο ως τον δειγματικό χώρο για μια ακολουθία ρίψεων ενός νομίσματος. Το -1 παριστά το αποτέλεσμα «Κορώνα» και το 1 το αποτέλεσμα «Γράμματα». Για $n \in \mathbb{N}^+$, ορίζουμε τη συνάρτηση $X_n : \Omega \rightarrow \mathbb{R}$ με $X_n(\omega) = \omega_n$,

όπου $\omega = (\omega_n)_{n \geq 1} \in \Omega$. Δηλαδή η X_n είναι η προβολή στη n -οστή συντεταγμένη. Η X_n παίρνει μόνο δύο τιμές. Οπότε η $\sigma(X_n)$ είναι ακριβώς το σύνολο $\{\emptyset, \Omega, A_{n,-1}, A_{n,1}\}$, με

$$A_{n,-1} := X_n^{-1}(\{-1\}) = \{\omega \in \Omega : \omega_n = -1\} = \{-1, 1\}^{n-1} \times \{-1\} \times \{-1, 1\}^{\mathbb{N}^+ \setminus [n]},$$

$$A_{n,1} := X_n^{-1}(\{1\}) = \{\omega \in \Omega : \omega_n = 1\} = \{-1, 1\}^{n-1} \times \{1\} \times \{-1, 1\}^{\mathbb{N}^+ \setminus [n]},$$

όπου $[n] := \{1, 2, \dots, n\}$.

Η επόμενη πρόταση εξετάζει τη δομή των $\sigma(f)$ -μετρήσιμων συναρτήσεων. Δεν θα τη χρειαστούμε στα επόμενα.

Πρόταση 4.17. Έστω $\Omega, f, \sigma(f)$ όπως στον πιο πάνω ορισμό και $g : \Omega \rightarrow \bar{\mathbb{R}}$ μια συνάρτηση. Τα εξής είναι ισοδύναμα.

(i) Η g είναι $s(f)$ -μετρήσιμη.

(ii) Υπάρχει μία $h : \bar{\mathbb{R}} \rightarrow \bar{\mathbb{R}}$ Borel-μετρήσιμη ώστε $g(x) = h(f(x))$ για κάθε $x \in \Omega$.

Απόδειξη. (ii) \Rightarrow (i) Αν $B \in \mathcal{B}(\bar{\mathbb{R}})$, τότε $g^{-1}(B) = f^{-1}(h^{-1}(B)) \in \sigma(f)$ από τον ορισμό της $s(f)$ αφού $h^{-1}(B) \in \mathcal{B}(\bar{\mathbb{R}})$.

(i) \Rightarrow (ii)

• Αν η g είναι δείκτρια, δηλαδή $g = \mathbf{1}_A$ για κάποιο $A \in \sigma(f)$. Τότε υπάρχει $B \in \mathcal{B}(\bar{\mathbb{R}})$ ώστε $A = f^{-1}(B)$. Έτσι, για $x \in \Omega$ έχουμε

$$g(x) = \mathbf{1}_A(x) = \mathbf{1}_B(f(x)),$$

επομένως ο ισχυρισμός είναι σωστός με $h = \mathbf{1}_B$.

• Αν η g είναι απλή και $\sigma(f)$ -μετρήσιμη με τιμές στο $[0, \infty)$, έστω $g = \sum_{i=1}^n a_i \mathbf{1}_{A_i}$ με τα $A_i \in \sigma(f)$, τότε από την προηγούμενη περίπτωση, υπάρχουν $h_i : \bar{\mathbb{R}} \rightarrow \{0, 1\}$ Borel-μετρήσιμες ώστε $\mathbf{1}_{A_i} = h_i(f)$. Έτσι, θέτοντας $h = \sum_{i=1}^n a_i h_i$, έχουμε ότι η h παίρνει τιμές στο $[0, \infty)$ και $g = h(f)$.

• Αν η g παίρνει τιμές στο $[0, \infty]$ και είναι $\sigma(f)$ μετρήσιμη, τότε υπάρχει αύξουσα ακολουθία απλών μετρήσιμων συναρτήσεων $(g_n)_{n \geq 1}$ που συγκλίνει σημειακά στην g . Κάθε g_n γράφεται ως $g_n = h_n(f)$ για κάποια $h_n : \bar{\mathbb{R}} \rightarrow [0, \infty)$ Borel-μετρήσιμη. Θέτουμε $h = \sup_{n \geq 1} h_n$. Η h είναι Borel-μετρήσιμη [Πρόταση 4.7] με τιμές στο $[0, \infty]$ και $h(f(x)) = \sup_{n \geq 1} h_n(f(x)) = \sup_{n \geq 1} g_n(x) = g(x)$.

• Αν η g παίρνει τιμές στο $\bar{\mathbb{R}}$ και είναι $\sigma(f)$ μετρήσιμη, τότε είναι επίσης $\sigma(f)$ -μετρήσιμες και οι g^-, g^+ . Από την προηγούμενη περίπτωση, υπάρχουν $h_1, h_2 : \bar{\mathbb{R}} \rightarrow [0, \infty]$ Borel-μετρήσιμες ώστε $g^- = h_1(f), g^+ = h_2(f)$. Θέτουμε $D = \{y \in \bar{\mathbb{R}} : h_1(y) = h_2(y) = \infty\}$ και $h(y) = (h_2(y) - h_1(y)) \mathbf{1}_{D^c}(y)$. Τότε

$$g(x) = g^+(x) - g^-(x) = h_2(f(x)) - h_1(f(x)) = h(f(x)). \quad \blacksquare$$

Ορισμός 4.18. Έστω Ω σύνολο. Αν $\{f_i : i \in I\}$ είναι οικογένεια συναρτήσεων στο Ω με τιμές στο $[-\infty, \infty]$, σ-άλγεβρα παραγόμενη από τις συναρτήσεις $\{f_i : i \in I\}$ ονομάζουμε το σύνολο

$$\sigma(\{f_i : i \in I\}) := \sigma\left(\cup_{i \in I} \sigma(f_i)\right). \quad (4.4)$$

Το σύνολο στο δεξί μέλος έχει οριστεί στην Παράγραφο 1.2. Αυτή είναι η ελάχιστη σ-άλγεβρα που κάνει όλες τις $\{f_i : i \in I\}$ μετρήσιμες. Αν $I = \{1, 2, \dots, n\}$, τη συμβολίζουμε με $\sigma(f_1, f_2, \dots, f_n)$.

Παράδειγμα 4.19. Έστω Ω σύνολο, $n \geq 2$, και $f_1, f_2, \dots, f_n : \Omega \rightarrow \mathbb{R}$. Τότε

$$\sigma(f_1 + f_2 + \dots + f_n) \subset \sigma(f_1, f_2, \dots, f_n).$$

Πράγματι, οι συναρτήσεις f_1, f_2, \dots, f_n είναι $\sigma(f_1, f_2, \dots, f_n)$ -μετρήσιμες και από την Πρόταση 4.6, είναι $\sigma(f_1, f_2, \dots, f_n)$ -μετρήσιμη και η συνάρτηση $f_1 + f_2 + \dots + f_n$. Όμως η $\sigma(f_1 + f_2 + \dots + f_n)$ είναι η ελάχιστη σ-άλγεβρα που κάνει την $f_1 + f_2 + \dots + f_n$ μετρήσιμη. Ο ισχυρισμός έπεται.

Παράδειγμα 4.20. Συνεχίζουμε από το Παράδειγμα 4.16. Θα περιγράψουμε τη σ-άλγεβρα $\mathcal{F}_n := \sigma(\{X_1, X_2, \dots, X_n\})$. Για δεδομένη ακολουθία $s = (s_1, s_2, \dots, s_n) \in \{-1, 1\}^n$ θεωρούμε το σύνολο

$$\begin{aligned} A_s &:= \{(s_1, s_2, \dots, s_n, x_{n+1}, x_{n+2}, \dots) : x_i \in \{-1, 1\} \text{ για κάθε } i \geq n+1\} \\ &= X_1^{-1}(\{s_1\}) \cap X_2^{-1}(\{s_2\}) \cap \dots \cap X_n^{-1}(\{s_n\}). \end{aligned}$$

Δηλαδή το A_s περιέχει όλες τις άπειρες ακολουθίες από -1 και 1 που το αρχικό τους τμήμα είναι το s και μετά είναι ελεύθερες να έχουν ότι θέλουν. Για μια ακολουθία που ανήκει στο A_s , η συμπεριφορά της ως τον χρόνο n είναι γνωστή.

Ισχυρισμός: Η \mathcal{F}_n είναι η σ-άλγεβρα που παράγεται από τη διαμέριση $C := \{A_s : s \in \{-1, 1\}^n\}$ του Ω .

Από τον ορισμό της, η \mathcal{F}_n πρέπει να περιέχει τα $X_i^{-1}(\{s_i\})$ για $i = 1, 2, \dots, n$. Άρα, ως σ-άλγεβρα, περιέχει και το A_s , που είναι πεπερασμένη τομή των $X_i^{-1}(\{s_i\})$. Επομένως, $\sigma(C) \subset \mathcal{F}_n$. Από την άλλη, κάθε X_i με $1 \leq i \leq n$ είναι μετρήσιμη ως προς τη $\sigma(C)$. Για παράδειγμα,

$$X_i^{-1}(\{1\}) = \cup_{s \in \{-1, 1\}^n : s_i = 1} A_s$$

είναι πεπερασμένη ένωση στοιχείων της $\sigma(C)$, άρα στοιχείο της. Από την ελαχιστότητα της \mathcal{F}_n , έπεται ότι $\mathcal{F}_n \subset \sigma(C)$ και ο ισχυρισμός αποδείχθηκε.

Ασκήσεις

4.1 Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας. Να δείξετε ότι για μια $X : \Omega \rightarrow \mathbb{R}$, τα ακόλουθα είναι ισοδύναμα:

- (α) $X^{-1}(A) \in \mathcal{F}$ για κάθε $A \in \mathcal{B}(\mathbb{R})$.
- (β) $X^{-1}(A) \in \mathcal{F}$ για κάθε $A \subset \mathbb{R}$ ανοιχτό σύνολο.
- (γ) $X^{-1}([a, b]) \in \mathcal{F}$ για κάθε $a < b$ πραγματικούς αριθμούς.

4.2 Έστω $X : \Omega \rightarrow [-\infty, \infty]$ τυχαία μεταβλητή στον μετρήσιμο χώρο (Ω, \mathcal{F}) . Να δείξετε ότι $\{X = -\infty\}, \{X = \infty\} \in \mathcal{F}$.

4.3 (Μετρήσιμες συναρτήσεις σε σ-άλγεβρα παραγόμενη από αριθμήσιμη διαμέριση) Έστω $C := \{A_i : i \in I\}$ μια αριθμήσιμη διαμέριση ενός συνόλου Ω , και $\mathcal{F} := \sigma(C)$ (Παράδειγμα 1.10). Να δειχθεί ότι μια συνάρτηση $f : \Omega \rightarrow \mathbb{R}$ είναι $\mathcal{F}/\mathcal{B}(\mathbb{R})$ μετρήσιμη αν και μόνο αν είναι σταθερή σε κάθε σύνολο της διαμέρισης.

4.4 Έστω $(X_n)_{n \geq 1}$ ακολουθία τυχαίων μεταβλητών σε έναν χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$ με $X_n : \Omega \rightarrow \mathbb{R}$ για κάθε $n \in \mathbb{N}^+$. Να δείξετε ότι τα παρακάτω σύνολα είναι στοιχεία της \mathcal{F} .

- (α) $\{\lim_{n \rightarrow \infty} X_n = -\infty\}, \{\lim_{n \rightarrow \infty} X_n = \infty\}$.
- (β) $\{\lim_{n \rightarrow \infty} X_n \text{ υπάρχει και είναι πραγματικός αριθμός}\}$.

4.5 Έστω $f : \mathbb{R} \rightarrow \mathbb{R}$ μονότονη συνάρτηση. Να δειχθεί ότι είναι Borel-μετρήσιμη.

4.6 Έστω (Ω, \mathcal{F}) μετρήσιμος χώρος. Αν $f, g : \Omega \rightarrow \mathbb{R}$ μετρήσιμες, να δείξετε ότι το $\{f = g\}$ είναι μετρήσιμο.

4.7 Έστω $(X_n)_{n \geq 1}$ ακολουθία τυχαίων μεταβλητών στον χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$ με τιμές στο \mathbb{R} . Θέτουμε $T := \min\{n \geq 1 : X_n > 2\}$ με τη σύμβαση $\min \emptyset = \infty$.

- (α) Να δειχθεί ότι $\{T = \infty\} \in \mathcal{F}$.
- (β) Να δειχθεί ότι η T είναι τυχαία μεταβλητή.

4.8 Έστω $X = (X_1, X_2, \dots, X_n)$ τυχαία μεταβλητή με τιμές στον \mathbb{R}^n . Θεωρούμε δείκτες $1 \leq i_1 < i_2 < \dots < i_k \leq n$ όπου $1 \leq k < n$. Να δειχθεί ότι η συνάρτηση $Y := (X_{i_1}, X_{i_2}, \dots, X_{i_k})$ είναι τυχαία μεταβλητή.

[Υπόδειξη: Ισχύει $Y = p(X)$, όπου p η προβολή $p : \mathbb{R}^n \rightarrow \mathbb{R}^k$ που απεικονίζει το (x_1, x_2, \dots, x_n) στο $(x_{i_1}, x_{i_2}, \dots, x_{i_k})$.]

Μετρήσιμες συναρτήσεις

4.9 Να δειχθεί ότι πράγματι το δεξί μέλος της (4.4) είναι η μικρότερη σ -άλγεβρα \mathcal{A} που κάνει όλες τις $\{f_i : i \in I\}$ \mathcal{A} -μετρήσιμες.

4.10 Σε αυτή την άσκηση θεωρούμε το πεδίο τιμών της $f : \mathbb{R} \rightarrow \mathbb{R}$, δηλαδή το \mathbb{R} , εφοδιασμένο με τη σ -άλγεβρα των συνόλων Borel. Περιγράψτε τη $\sigma(f)$ στην περίπτωση που

(α) $f(x) = x^3$,

(β) $f(x) = x^2$.

4.11 Έστω X τυχαία μεταβλητή με τιμές στο \mathbb{R} . Αν $\mathbf{P}(X > 1) > 0$, τότε υπάρχει $\varepsilon > 0$ ώστε $\mathbf{P}(X > 1 + \varepsilon) > 0$.

5

Το ολοκλήρωμα Lebesgue

5.1 Ολοκλήρωμα Lebesgue. Ορισμός

Έστω (X, \mathcal{A}, μ) χώρος μέτρου. Στην παράγραφο αυτή θα ορίσουμε το ολοκλήρωμα οποιασδήποτε $\mathcal{A}/\mathcal{B}([-\infty, \infty])$ μετρήσιμης συνάρτησης $f : X \rightarrow [-\infty, \infty]$. Αυτό θα το κάνουμε σε τρία βήματα. Πρώτα για $f \geq 0$ απλή μετρήσιμη, έπειτα για $f \geq 0$ μετρήσιμη, και τέλος για f μετρήσιμη με τιμές στο $[-\infty, \infty]$.

Βήμα 1: $f \geq 0$ απλή μετρήσιμη.

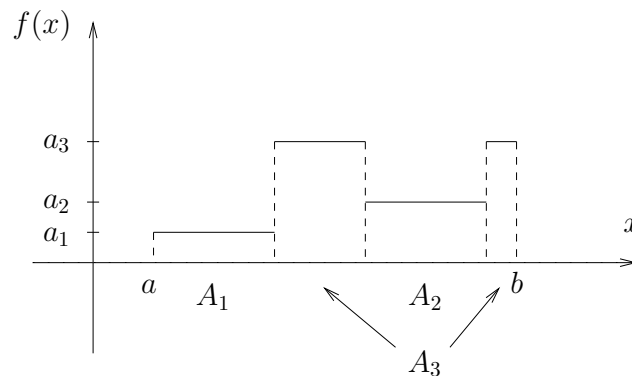
Ορισμός 5.1. Έστω $f : X \rightarrow [0, \infty]$ απλή μετρήσιμη συνάρτηση με κανονική μορφή $f = \sum_{i=1}^n a_i \mathbf{1}_{A_i}$. Ορίζουμε το ολοκλήρωμα Lebesgue της f ως προς το μέτρο μ ως εξής:

$$\int f \, d\mu := \sum_{i=1}^n a_i \mu(A_i), \quad (5.1)$$

με τη σύμβαση $0 \cdot \infty = 0$.

Το ολοκλήρωμα είναι στοιχείο του $[0, \infty]$.

Είναι φυσιολογικός αυτός ο ορισμός; Ας τον ελέγξουμε στην περίπτωση που το X είναι ένα κλειστό και φραγμένο διάστημα $[a, b]$ του \mathbb{R} και μ είναι το μέτρο Lebesgue λ στο $[a, b]$.



Σχήμα 5.1: Ολοκλήρωμα απλής συνάρτησης.

Η συνάρτηση f στο Σχήμα 5.1 είναι απλή και μάλιστα τα σύνολα στα οποία παίρνει διαφορετικές τιμές είναι το καθένα διάστημα ή ένωση διαστημάτων. Με βάση τον προηγούμενο ορισμό, το ολοκλήρωμα της είναι

$$a_1 \lambda(A_1) + a_2 \lambda(A_2) + a_3 \lambda(A_3).$$

Αυτό είναι το εμβαδόν κάτω από το γράφημα της f . Όπως στο ολοκλήρωμα Riemann, έτσι και εδώ, τα γινόμενα $a_i \lambda(A_i)$ δίνουν εμβαδόν ορθογωνίου: ύψος επί βάση. Μόνο που εδώ η βάση ενδέχεται να μην είναι διάστημα. Σε κάθε περίπτωση όμως, το μήκος της βάσης μετριέται σωστά από το μέτρο Lebesgue.

Βήμα 2: $f \geq 0$ μετρήσιμη.

Ορισμός 5.2. Έστω $f : X \rightarrow [0, \infty]$ μετρήσιμη συνάρτηση. Το ολοκλήρωμα Lebesgue της f ως προς το μέτρο μ ορίζεται ως εξής:

$$\int f d\mu := \sup \left\{ \int s d\mu : s \text{ απλή, μετρήσιμη με } 0 \leq s \leq f \right\}. \quad (5.2)$$

Παρατήρηση 5.3. Ο Ορισμός 5.2, στην περίπτωση που η f είναι απλή, συμφωνεί με τον Ορισμό 5.1.

Βήμα 3: $f : X \rightarrow [-\infty, \infty]$ μετρήσιμη.

Ορισμός 5.4. Έστω $f : X \rightarrow [-\infty, \infty]$ μετρήσιμη συνάρτηση. Το ολοκλήρωμα Lebesgue της f ως προς το μέτρο μ ορίζεται ως εξής:

$$\int f d\mu := \int f^+ d\mu - \int f^- d\mu \quad (5.3)$$

εφόσον στο δεξί μέλος της ισότητας δεν εμφανίζεται απροσδιοριστία της μορφής $\infty - \infty$. Στην περίπτωση που το ολοκλήρωμα είναι πραγματικός αριθμός, λέμε ότι η συνάρτηση f είναι **(Lebesgue) ολοκληρώσιμη**.

Για το $\int f d\mu$ χρησιμοποιούμε επίσης τον συμβολισμό $\int f(x) d\mu(x)$.

Παρατήρηση 5.5. (i) Τα $\int f^+ d\mu$ και $\int f^- d\mu$ που εμφανίζονται στον Ορισμό 5.4 ορίζονται από τον Ορισμό 5.2.

(ii) Το ολοκλήρωμα μιας μετρήσιμης συνάρτησης, όταν αυτό ορίζεται, είναι στοιχείο του $[-\infty, \infty]$.

(iii) Μια μετρήσιμη συνάρτηση f είναι ολοκληρώσιμη αν και μόνο αν και τα δύο ολοκληρώματα $\int f^- d\mu, \int f^+ d\mu$ είναι πεπερασμένα.

(iv) Για μια $f \geq 0$ μετρήσιμη συνάρτηση, θεωρούμε την ακολουθία $(f_n)_{n \geq 1}$ των απλών συναρτήσεων που ορίστηκαν στην Πρόταση 4.12. Αποδεικνύεται (με χρήση του Θεωρήματος 5.25 παρακάτω) ότι

$$\int f d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu.$$

Δηλαδή το ολοκλήρωμα είναι το όριο των ολοκληρωμάτων αυτών των απλών συναρτήσεων. Αυτό είναι αντίστοιχο της προσέγγισης του ολοκληρώματος Riemann μιας Riemann-ολοκληρώσιμης συνάρτησης από τα ολοκληρώματα κλιμακωτών συναρτήσεων.

Παρατήρηση 5.6. Όποτε μια f γράφεται ως $f = \sum_{i=1}^n a_i \mathbf{1}_{A_i}$ με τα A_i μετρήσιμα και τα $a_i \geq 0$, τότε πάλι ισχύει ο τύπος (5.1). Δηλαδή δεν είναι απαραίτητο η γραφή $f = \sum_{i=1}^n a_i \mathbf{1}_{A_i}$ να αντιστοιχεί στην κανονική μορφή της f . Ενδεχομένως κάποια από τα $\{A_i : 1 \leq i \leq n\}$ να τέμνονται και κάποια από τα $\{a_i : 1 \leq i \leq n\}$ να είναι ίσα. Το ίδιο ισχύει και όταν $n = \infty$, δηλαδή $f = \sum_{i=1}^{\infty} a_i \mathbf{1}_{A_i}$ με τα a_i, A_i όπως πριν. Και οι δύο ισχυρισμοί έπονται από το Πόρισμα 5.31(i) πιο κάτω.

5.2 Ειδικές περιπτώσεις

Θα δούμε εδώ τις περιπτώσεις που το μέτρο μ της προηγούμενης παραγράφου είναι το αριθμητικό μέτρο στο \mathbb{N} ή το μέτρο Lebesgue σε ένα διάστημα στο \mathbb{R} .

ΑΡΙΘΜΗΤΙΚΟ ΜΕΤΡΟ. Αν πάρουμε μ το αριθμητικό μέτρο στον $X = \mathbb{N}$ (Παράδειγμα 2.2) και $f : \mathbb{N} \rightarrow [0, \infty)$ συνάρτηση [όλες οι συναρτήσεις σε αυτόν τον χώρο είναι μετρήσιμες], τότε

$$\int f d\mu = \sum_{n=0}^{\infty} f(n). \quad (5.4)$$

[Αυτό έπεται από την Παρατήρηση 5.6 αφού η f γράφεται ως $f = \sum_{n=0}^{\infty} f(n)\mathbf{1}_{\{n\}}$ και κάθε μονοσύνολο έχει αριθμητικό μέτρο 1.

Μια απόδειξη με βάση μόνο τον ορισμό του ολοκληρώματος είναι η ακόλουθη.

Αν η f είναι απλή, μη αρνητική, που γράφεται σε κανονική μορφή ως $f = \sum_{i=1}^n a_i \mathbf{1}_{A_i}$, τότε

$$\int f d\mu = \sum_{i=1}^n a_i |A_i| = \sum_{i=1}^n a_i \sum_{j \in A_i} 1 = \sum_{i=1}^n \sum_{j \in A_i} f(j) = \sum_{j \in \mathbb{N}} f(j)$$

Χρησιμοποιήσαμε το ότι $f(j) = a_i$ για κάθε $j \in A_i$ και ότι η $\{A_1, A_2, \dots, A_n\}$ είναι διαμέριση του \mathbb{N} . Έτσι η (5.4) αποδείχθηκε για τέτοιες f .

Έστω τώρα $f : \mathbb{N} \rightarrow [0, \infty)$. Με βάση το πιο πάνω, η απλή συνάρτηση $s_n := \sum_{k=0}^n f(k)\mathbf{1}_{\{k\}}$ (όπου $n \in \mathbb{N}^+$) έχει ολοκλήρωμα $\int s_n d\mu = \sum_{k=0}^n s_n(k) = \sum_{k=0}^n f(k)$. Έτσι, εφόσον $0 \leq s_n \leq f$, ο Ορισμός (5.2) δίνει ότι $\int f d\mu \geq \sum_{k=0}^n f(k)$. Αυτό συνεπάγεται την (5.4) στην περίπτωση που το δεξί μέλος της (5.4) ισούται με ∞ . Στην περίπτωση που το δεξί μέλος είναι πεπερασμένο, πρέπει να δείξουμε και την αντίστροφη ανισότητα, δηλαδή ότι $\int f d\mu \leq \sum_{k=0}^{\infty} f(k)$. Αν s είναι μια απλή συνάρτηση με $0 \leq s \leq f$ που έχει κανονική μορφή $\sum_{i=1}^m b_i \mathbf{1}_{B_i}$, τότε κάθε i με $b_i > 0$ έχει B_i πεπερασμένο (διαφορετικά, $\sum_{k=1}^{\infty} f(k) = \infty$), άρα υπάρχει $n_0 \in \mathbb{N}^+$ ώστε $s \leq s_{n_0}$. Χρησιμοποιώντας αυτή την ανισότητα μαζί με το ότι η (5.4) ισχύει για απλές μη αρνητικές συναρτήσεις, παίρνουμε ότι $\int s d\mu \leq \int s_n d\mu \leq \sum_{k=0}^{\infty} f(k)$, και το ζητούμενο δείχθηκε.]

Έπειτα είναι απλό να δούμε ότι η (5.4) ισχύει για κάθε $f : \mathbb{N} \rightarrow [-\infty, \infty]$ με $\sum_{n=0}^{\infty} |f(n)| < \infty$. Άρα το άθροισμα θετικής ή απολύτως συγκλίνουσας σειράς είναι ειδική περίπτωση του ολοκληρώματος Lebesgue. Όμως αθροίσματα σειρών που συγκλίνουν υπό συνθήκη, όπως η $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$, δεν καλύπτονται (το ολοκλήρωμα Lebesgue της $f(n) := (-1)^n/n$ ως προς το αριθμητικό μέτρο δεν ορίζεται αφού $\int f^- d\mu = \int f^+ d\mu = \infty$). Το ολοκλήρωμα Lebesgue δεν προσθέτει ποσότητες με κάποια «σειρά».

ΜΕΤΡΟ Lebesgue ΣΕ ΦΡΑΓΜΕΝΟ ΔΙΑΣΤΗΜΑ. Έστω $a < b$ πραγματικοί αριθμοί. Στο σύνολο $[a, b]$ θεωρούμε τη σ -άλγεβρα $\mathcal{M}_{[a,b]}$ των Lebesgue-μετρήσιμων υποσυνόλων του $[a, b]$ και τον περιορισμό του μέτρου Lebesgue στο $[a, b]$, τον οποίο συμβολίζουμε επίσης με λ [δες Παράδειγμα 2.4]. Το ολοκλήρωμα Lebesgue στον χώρο μέτρου $([a, b], \mathcal{M}_{[a,b]}, \lambda)$ επεκτείνει γνήσια το ολοκλήρωμα Riemann. Δηλαδή αν η $f : [a, b] \rightarrow \mathbb{R}$ είναι Riemann-ολοκληρώσιμη τότε είναι Lebesgue-μετρήσιμη και Lebesgue-ολοκληρώσιμη ως προς το λ και

$$\int_{[a,b]} f(x) d\lambda(x) = \int_a^b f(x) dx.$$

Το δεξί μέλος είναι το ολοκλήρωμα Riemann της f [δες Stein and Shakarchi (2005), Κεφάλαιο 2, Θεώρημα 1.5 για την απόδειξη αυτού του ισχυρισμού]. Για αυτόν τον λόγο, συνήθως συμβολίζουμε

το ολοκλήρωμα Lebesgue οποιασδήποτε Lebesgue-ολοκληρώσιμης f με $\int_a^b f(x) dx$ ακόμα και αν η f δεν είναι Riemann-ολοκληρώσιμη. Η επέκταση είναι γνήσια γιατί η συνάρτηση $\mathbf{1}_{\mathbb{Q} \cap [a,b]}$ δεν είναι Riemann-ολοκληρώσιμη (εύκολη άσκηση) ενώ είναι Lebesgue-ολοκληρώσιμη με ολοκλήρωμα 0 [αποδεικνύεται εύκολα στοιχειωδώς αλλά και έπεται άμεσα από την Πρόταση 5.12(i) πιο κάτω].

ΜΕΤΡΟ Lebesgue ΣΤΟ \mathbb{R} . Αν για μια $f : \mathbb{R} \rightarrow \mathbb{R}$ υπάρχει¹ το γενικευμένο ολοκλήρωμα Riemann $\int_{-\infty}^{\infty} f(x) dx$ και ισχύει μια από τις $\int_{-\infty}^{\infty} |f(x)| dx < \infty, f \geq 0$, τότε η f είναι Lebesgue-μετρήσιμη και Lebesgue-ολοκληρώσιμη ως προς το λ και

$$\int_{\mathbb{R}} f(x) d\lambda(x) = \int_{-\infty}^{\infty} f(x) dx.$$

Το δεξί μέλος είναι το γενικευμένο ολοκλήρωμα Riemann της f . Τώρα όμως ενδέχεται για μια συνάρτηση να υπάρχει το γενικευμένο ολοκλήρωμα Riemann αλλά να μην είναι Lebesgue-ολοκληρώσιμη ως προς το λ . Το κλασικό παράδειγμα είναι η $f(x) = \mathbf{1}_{x>0}(\sin x)/x$. Για αυτήν υπάρχει το γενικευμένο ολοκλήρωμα Riemann

$$\int_0^{\infty} \frac{\sin x}{x} dx = \frac{\pi}{2},$$

αλλά

$$\int_{\mathbb{R}} f^-(x) d\lambda(x) = \int_{\mathbb{R}} f^+(x) d\lambda(x) = \infty,$$

οπότε το ολοκλήρωμα Lebesgue της f δεν μπορεί να οριστεί.

5.3 Η οπτική του ολοκληρώματος Lebesgue

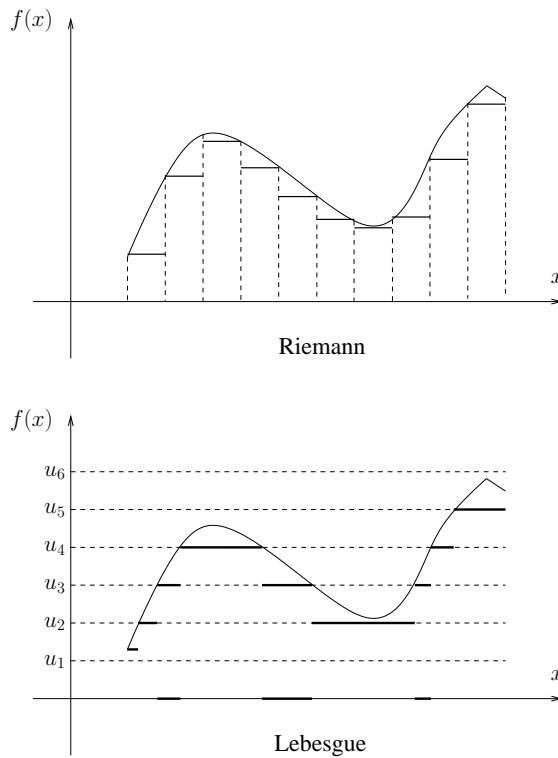
Είναι ενδιαφέρον να συγκρίνουμε τις οριακές διαδικασίες που δίνουν τα ολοκληρώματα Riemann και Lebesgue σε μια περίπτωση συνάρτησης/χώρου που και τα δύο ολοκληρώματα έχουν νόημα [Ως οριακή διαδικασία για το Lebesgue παίρνουμε αυτήν που περιγράφεται στην Παρατήρηση 5.5 (iv)]. Πιο συγκεκριμένα, παίρνουμε $a < b$ πραγματικούς αριθμούς και μια $f : [a, b] \rightarrow [0, \infty)$ φραγμένη και Riemann-ολοκληρώσιμη. Τότε η f είναι επίσης Lebesgue-μετρήσιμη και μη αρνητική, οπότε και το ολοκλήρωμα Lebesgue ορίζεται.

Για το Riemann, διαμερίζουμε το πεδίο ορισμού της f σε τμήματα ίσου μήκους [δες σχήμα 5.2]. Σε καθένα από αυτά, η f έχει μια δεδομένη ελάχιστη τιμή. Πολλαπλασιάζουμε αυτή την ελάχιστη τιμή με το μήκος του τμήματος για να βρούμε τη συνεισφορά του τμήματος στην προσέγγιση του ολοκληρώματος. Έπειτα προσθέτουμε τις συνεισφορές όλων των τμημάτων. Καθώς το μήκος των τμημάτων τείνει στο μηδέν, παίρνουμε την τιμή του ολοκληρώματος Riemann της f .

Για το Lebesgue, διαμερίζουμε το σύνολο τιμών της f σε τμήματα ίσου μήκους. Αυτή η διαμέριση δίνει μια απλή συνάρτηση (το γράφημά της είναι τα χοντρά ευθύγραμμα τμήματα στην κάτω γραφική παράσταση στο Σχήμα 5.2 εκτός από εκείνα που είναι στον άξονα), που είναι μία από τις f_n της Πρότασης 4.12. Ας πάρουμε ένα τμήμα, π.χ. το $[u_3, u_4)$. Η συνεισφορά του στην προσέγγιση $\int f_n d\lambda$ του ολοκληρώματος είναι το εμβαδόν $u_3 \lambda(f^{-1}([u_3, u_4)))$ του «παραλληλογράμμου» με ύψος u_3 και βάση $f^{-1}([u_3, u_4))$ [σχέση (5.1)]. Στη συγκεκριμένη περίπτωση, το $f^{-1}([u_3, u_4))$ είναι ένωση τριών διαστημάτων, σημειωμένων με χοντρή γραμμή στον x -άξονα. Το μήκος αυτής της βάσης είναι το μέτρο Lebesgue του συνόλου $f^{-1}([u_3, u_4))$. Πάλι προσθέτουμε τις συνεισφορές

¹ Δηλαδή η f είναι Riemann-ολοκληρώσιμη στο $[a, b]$ για κάθε πραγματικούς $a < b$ και το $\lim_{a \rightarrow -\infty, b \rightarrow \infty} \int_a^b f(x) dx$ υπάρχει στο $[-\infty, \infty]$.

Το ολοκλήρωμα Lebesgue



Σχήμα 5.2: Η διαφορά οπτικής των ολοκληρωμάτων Riemann και Lebesgue.

όλων των τμημάτων και, καθώς το μήκος τους τείνει στο μηδέν ($n \rightarrow \infty$), παίρνουμε την τιμή του ολοκληρώματος Lebesgue της f .

Το μη τετριμμένο της διαδικασίας για το ολοκλήρωμα Lebesgue είναι ότι πρέπει να είμαστε σε θέση να υπολογίσουμε το μήκος του συνόλου $f^{-1}([u_{k-1}, u_k])$. Στο πιο πάνω παράδειγμα, έτυχε αυτό να είναι ένωση τριών διαστημάτων και είναι προφανές ποιο πρέπει να ονομάσουμε μήκος του. Θα μπορούσε όμως να είναι ένα πολύ περίεργο σύνολο, ειδικά όταν η f δεν είναι συνεχής. Η έννοια μήκους για τα Lebesgue-μετρήσιμα υποσύνολα του \mathbb{R} δίνεται ακριβώς από το μέτρο Lebesgue τους, του οποίου η κατασκευή δεν είναι απλή και γι' αυτό ακριβώς την παραλείψαμε σε αυτές τις σημειώσεις.

Κλείνοντας αυτή τη σύγκριση, να παρατηρήσουμε το εξής πολύ σημαντικό. Για τον ορισμό του ολοκληρώματος Lebesgue, το πεδίο ορισμού, X , της συνάρτησης $f : X \rightarrow \mathbb{R}$ που θέλουμε να ολοκληρώσουμε αρκεί να είναι εφοδιασμένο με μία σ -άλγεβρα και ένα μέτρο. Δεν είναι αναγκαίο να έχει κάποια άλλη δομή (μετρικού ή διανυσματικού χώρου) όπως είναι οι \mathbb{R}^d στους οποίους έχουμε ορίσει το ολοκλήρωμα Riemann. Για το ολοκλήρωμα Riemann χρησιμοποιούμε την επιπλέον δομή με κρίσιμο τρόπο.

5.4 Ιδιότητες του ολοκληρώματος

Πρόταση 5.7. Έστω $f, g : X \rightarrow [-\infty, +\infty]$ μετρήσιμες συναρτήσεις των οποίων το ολοκλήρωμα ορίζεται. Τότε:

$$(i) \int af \, d\mu = a \int f \, d\mu, \text{ για } a \in \mathbb{R}.$$

$$(ii) \int (f + g) \, d\mu = \int f \, d\mu + \int g \, d\mu.$$

$$(iii) \text{ Αν } f \leq g, \text{ τότε } \int f \, d\mu \leq \int g \, d\mu.$$

$$(iv) \left| \int f \, d\mu \right| \leq \int |f| \, d\mu.$$

Η (ii) ισχύει με την προϋπόθεση ότι $\mu(\{f = -g = \infty\} \cup \{f = -g = -\infty\}) = 0$ και στο δεξί της μέλος δεν εμφανίζεται η μορφή $\infty - \infty$.

Παρατήρηση 5.8. Για f όπως στην προηγούμενη πρόταση, η σχέση $|f| = f^- + f^+$ και η ιδιότητα (ii) δίνουν ότι

$$\int |f| \, d\mu = \int f^- \, d\mu + \int f^+ \, d\mu. \quad (5.5)$$

Επομένως, η f είναι ολοκληρώσιμη, δηλαδή έχει ολοκλήρωμα πραγματικό αριθμό, αν και μόνο αν $\int |f| \, d\mu < \infty$.

Αν (X, \mathcal{A}, μ) είναι χώρος μέτρου, $f : X \rightarrow [-\infty, \infty]$ είναι μετρήσιμη, και $A \in \mathcal{A}$, ορίζουμε το ολοκλήρωμα της f ως προς το μέτρο μ στο A ως

$$\int_A f \, d\mu := \int f \mathbf{1}_A \, d\mu$$

εφόσον ορίζεται το δεξί μέλος της ισότητας. Όταν $A = X$, το $\int_X f \, d\mu$ είναι απλώς το $\int f \, d\mu$. Επίσης, εύκολα βλέπουμε ότι αν $f \geq 0$ και $A \subset B$ στοιχεία της \mathcal{A} , ισχύει ότι $\int_A f \, d\mu \leq \int_B f \, d\mu$.

Στην περίπτωση ενός χώρου πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$, το ολοκλήρωμα μιας τυχαίας μεταβλητής $X : \Omega \rightarrow \mathbb{R}$ λέγεται μέση τιμή της X και, αντί του $\int X \, d\mathbf{P}$, χρησιμοποιούμε τον συμβολισμό $\mathbf{E}(X)$. Συνοψίζουμε στον επόμενο ορισμό.

Ορισμός 5.9. Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας και $X : \Omega \rightarrow [-\infty, \infty]$ τυχαία μεταβλητή. Η μέση τιμή της X ορίζεται ως

$$\mathbf{E}(X) := \int X \, d\mathbf{P}$$

εφόσον το δεξί μέλος της ισότητας ορίζεται. Πολλές φορές γράφουμε την $\mathbf{E}(X)$ και ως $\mathbf{E}X$. Επίσης γράφουμε $\mathbf{E}_{\mathbf{P}}$ αντί του \mathbf{E} αν θέλουμε να κάνουμε ξεκάθαρο ότι η ολοκλήρωση γίνεται ως προς το μέτρο πιθανότητας \mathbf{P} .

Στο επόμενο κεφάλαιο θα δούμε τη σχέση αυτού του ορισμού με τον ορισμό της μέσης τιμής που δίνεται στις στοιχειώδεις πιθανότητες [σχέσεις (6.6), (6.10)].

Παρατήρηση 5.10. Αντίστοιχα, αν $A \in \mathcal{F}$, ορίζουμε τη μέση τιμή της X πάνω στο A ως $\mathbf{E}(X\mathbf{1}_A)$ και τη συμβολίζουμε με $\mathbf{E}(X; A)$ εφόσον η $\mathbf{E}(X\mathbf{1}_A)$ μπορεί να οριστεί.

Δύο ειδικές περιπτώσεις μέσης τιμής είναι οι εξής:

(i) Αν η X ισούται με μια σταθερά $c \in \mathbb{R}$, τότε $\mathbf{E}(X) = c$ γιατί η X είναι απλή.

(ii) Αν $X = \mathbf{1}_A$, $A \in \mathcal{F}$, τότε $\mathbf{E}(X) = \mathbf{P}(A)$.

Το (ii) σε συνδυασμό με τις ιδιότητες της μέσης τιμής (Προτάσεις 5.7, 5.14) είναι πολύ χρήσιμο (Ασκήσεις 5.1, 5.2, 5.18).

Παρατήρηση 5.11. Σε έναν χώρο μέτρου (X, \mathcal{A}, μ) , λέμε ότι μια ιδιότητα $\Psi(x)$ που αφορά σημεία x του X (παραδειγμα τέτοιας $\Psi(x)$ είναι η «το όριο $\lim_{n \rightarrow \infty} f_n(x)$ υπάρχει», όπου $(f_n)_{n \geq 1}$ είναι ακολουθία συναρτήσεων με $f_n : X \rightarrow \mathbb{R}$ για κάθε n) ισχύει **μ -σχεδόν παντού**, ή **μ -σχεδόν για κάθε** $x \in X$, αν υπάρχει $A \in \mathcal{A}$ με $A \supset \{x \in X : \eta \Psi(x) \text{ δεν ισχύει}\}$ και $\mu(A) = 0$. Θα θέλαμε να δώσουμε ως ορισμό το ότι το σύνολο στο οποίο η ιδιότητα δεν ισχύει, δηλαδή το $\{x \in X : \eta \Psi(x) \text{ δεν ισχύει}\}$, έχει μέτρο 0. Όμως επειδή αυτό το σύνολο δεν είναι απαραίτητα μετρήσιμο, δίνουμε τον πιο πάνω ορισμό. Αν το μ είναι μέτρο πιθανότητας, λέμε ότι η $\Psi(x)$ ισχύει **με μ -πιθανότητα 1**, ή **μ -σχεδόν βέβαια**. Αν είναι σαφές ποιο είναι το μέτρο μ , το παραλείπουμε στις παραπάνω εκφράσεις.

Πρόταση 5.12. Έστω (X, \mathcal{A}, μ) χώρος μέτρου και $f, g : X \rightarrow [0, \infty]$ μετρήσιμες συναρτήσεις. Τότε:

(i) $\int f \, d\mu = 0$ αν και μόνο αν $\mu(\{f \neq 0\}) = 0$.

(ii) Αν $\mu(\{f \neq g\}) = 0$, τότε $\int f \, d\mu = \int g \, d\mu$.

(iii) Αν $\int f \, d\mu < \infty$, τότε $\mu(f = \infty) = 0$.

Απόδειξη. (i) « \Rightarrow » Έστω ότι $\int f \, d\mu = 0$. Θέτουμε $A_n := \{f \geq \frac{1}{n}\}$, $n \in \mathbb{N}^+$. Τότε,

$$0 = \int f \, d\mu \geq \int_{A_n} f \, d\mu = \int f \mathbf{1}_{A_n} \, d\mu \geq \int \frac{1}{n} \mathbf{1}_{A_n} \, d\mu = \frac{1}{n} \mu(A_n) \quad \text{για κάθε } n \in \mathbb{N}^+.$$

Άρα $\mu(A_n) = 0$ για κάθε $n \in \mathbb{N}^+$. Όμως, $\{f > 0\} = \cup_{n \geq 1} A_n$ και $\mu(\{f > 0\}) \leq \sum_{n \geq 1} \mu(A_n) = 0$. Συνεπώς, $\mu(\{f \neq 0\}) = 0$.

« \Leftarrow » Αν η f είναι απλή όπως στον Ορισμό 5.1, τότε από την υπόθεση πρέπει $\mu(A_i) = 0$ για κάθε i με $a_i > 0$, και άρα $\int f \, d\mu = 0$. Στη γενική περίπτωση, αν πάρουμε απλή, μετρήσιμη s με $0 \leq s \leq f$, τότε $\mu(\{s \neq 0\}) \leq \mu(\{f \neq 0\}) = 0$, και όπως δείξαμε πριν, πρέπει να ισχύει $\int s \, d\mu = 0$. Το συμπέρασμα έπεται.

(ii) $f \leq g + (f - g)\mathbf{1}_{f > g}$. Άρα

$$\int f \, d\mu \leq \int g \, d\mu + \int (f - g)\mathbf{1}_{f > g} \, d\mu = \int g \, d\mu.$$

Η ισότητα έπεται από το μέρος (i) της πρότασης γιατί η $(f - g)\mathbf{1}_{f > g}$ είναι μη αρνητική και

$$\mu(\{(f - g)\mathbf{1}_{f > g} \neq 0\}) \leq \mu(\{f \neq g\}) = 0.$$

Αλλάζοντας τους ρόλους των f, g , παίρνουμε την αντίστροφη ανισότητα και ο ισχυρισμός αποδείχθηκε.

(iii) Για δεδομένο $n \in \mathbb{N}$, η συνάρτηση $s := n\mathbf{1}_{f = \infty}$ είναι απλή, μετρήσιμη και $0 \leq s \leq f$. Άρα

$$\int f \, d\mu \geq \int s \, d\mu = n \times \mu(f = \infty).$$

Αν $\mu(f = \infty) > 0$, τότε για $n \rightarrow \infty$ παίρνουμε $\int f \, d\mu = \infty$. Άτοπο. ■

Παρατήρηση 5.13. Εύκολα βλέπουμε ότι η ιδιότητα (ii) της Πρότασης 5.12 ισχύει και για μετρήσιμες συναρτήσεις με τιμές στο $[-\infty, \infty]$ των οποίων το ολοκλήρωμα ορίζεται.

Στην περίπτωση ενός χώρου πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$, η Πρόταση 5.12 παίρνει την εξής μορφή.

Πρόταση 5.14. Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας και $X, Y : \Omega \rightarrow [0, \infty]$ τυχαίες μεταβλητές. Τότε:

- (i) $\mathbf{E}(X) = 0$ αν και μόνο αν $\mathbf{P}(X = 0) = 1$.
- (ii) Αν $\mathbf{P}(X = Y) = 1$, τότε $\mathbf{E}(X) = \mathbf{E}(Y)$.
- (iii) Αν $\mathbf{E}(X) < \infty$, τότε $\mathbf{P}(X = \infty) = 0$.

Ορισμός 5.15. Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας και $X : \Omega \rightarrow \mathbb{R}$ τυχαία μεταβλητή με $\mathbf{E}|X| < \infty$. Ορίζουμε τη διασπορά $\text{Var}(X)$ της X ως εξής:

$$\text{Var}(X) := \mathbf{E}(\{X - \mathbf{E}(X)\}^2).$$

Η μέση τιμή $\mathbf{E}(X)$, όπως έχουμε ήδη σημειώσει (Παρατήρηση 5.8), είναι πραγματικός αριθμός λόγω της $\mathbf{E}|X| < \infty$. Η διασπορά, όμως, ενδέχεται να παίρνει την τιμή ∞ . Ένας χρήσιμος τύπος για τη διασπορά, που προκύπτει εύκολα από τον ορισμό της, είναι ο $\text{Var}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2$. Έτσι βλέπουμε ότι αν $\mathbf{E}(X^2) < \infty$, τότε $\text{Var}(X) < \infty$.

Η διασπορά είναι ένα μέτρο της μεταβλητότητας της τυχαίας μεταβλητής γύρω από τη μέση της τιμή. Έτσι, όταν $\text{Var}(X) = 0$, αναμένουμε η X να είναι συγκεντρωμένη στη μέση τιμή. Ισχύει το εξής

$$\text{Var}(X) = 0 \Leftrightarrow \mathbf{P}(X = c) = 1 \tag{5.6}$$

με $c = \mathbf{E}(X)$. Απόδειξη χρειάζεται μόνο η κατεύθυνση \Rightarrow . Η $\text{Var}(X) = 0$ σημαίνει ότι η μη αρνητική τυχαία μεταβλητή $\{X - \mathbf{E}(X)\}^2$ έχει μέση τιμή μηδέν. Με βάση την Πρόταση 5.14(i), η $X - \mathbf{E}(X) = 0$ με πιθανότητα 1, που είναι το ζητούμενο.

Δίνουμε τώρα δύο σημαντικές ανισότητες διατυπωμένες στη γλώσσα των πιθανοτήτων. Αντίστοιχες ισχύουν και στην περίπτωση μετρήσιμων συναρτήσεων σε τυχόντα χώρο μέτρου. Εκφράζουν το γεγονός ότι η πιθανότητα μια τυχαία μεταβλητή να βρεθεί μακριά από τη μέση της τιμή είναι μικρή.

Πρόταση 5.16. Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας. Τότε ισχύουν τα εξής:

- (i) (Ανισότητα Markov) Αν $X : \Omega \rightarrow [0, \infty]$ τυχαία μεταβλητή και $a > 0$, τότε

$$\mathbf{P}(X \geq a) \leq \frac{\mathbf{E}(X)}{a}.$$

- (ii) (Ανισότητα Chebyshev) Αν $X : \Omega \rightarrow [-\infty, \infty]$ τυχαία μεταβλητή με $\mathbf{E}|X| < \infty$ και $a > 0$, τότε

$$\mathbf{P}(|X - \mathbf{E}(X)| \geq a) \leq \frac{\text{Var}(X)}{a^2}.$$

Απόδειξη. (i) Χρησιμοποιούμε τη μονοτονία της μέσης τιμής. Έχουμε $X \geq a\mathbf{1}_{X \geq a}$. Άρα

$$\mathbf{E}(X) \geq \mathbf{E}(a\mathbf{1}_{X \geq a}) = a\mathbf{P}(X \geq a).$$

(ii) Εφαρμόζουμε το (i) στην τυχαία μεταβλητή $|X - \mathbf{E}(X)|^2$. Δηλαδή

$$\mathbf{P}(|X - \mathbf{E}(X)| \geq a) = \mathbf{P}(|X - \mathbf{E}(X)|^2 \geq a^2) \leq \frac{\mathbf{E}\{|X - \mathbf{E}(X)|^2\}}{a^2} = \frac{\text{Var}(X)}{a^2}.$$

■

Τέλος, καταγράφουμε χωρίς απόδειξη μια χρήσιμη ιδιότητα της μέσης τιμής (δες Άσκηση 5.3).

Πρόταση 5.17 (Ανισότητα Jensen). Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας, $X : \Omega \rightarrow \mathbb{R}$ τυχαία μεταβλητή με $\mathbf{E}|X| < \infty$, και $\Phi : I \rightarrow \mathbb{R}$ κυρτή συνάρτηση σε ένα διάστημα $I \subset \mathbb{R}$ με $\mathbf{P}(\{X \in I\}) = 1$ και $\mathbf{E}|\Phi(X)| < \infty$. Τότε

$$\Phi(\mathbf{E}\{X\}) \leq \mathbf{E}\{\Phi(X)\}. \quad (5.7)$$

Αν η $\Phi : I \rightarrow \mathbb{R}$ είναι συνάρτηση, με το $I \subset \mathbb{R}$ διάστημα, η οποία ικανοποιεί την (5.7) για όλες τις επιλογές χώρων πιθανότητας και τυχαίων μεταβλητών X , τότε η Φ πρέπει να είναι κυρτή. Πράγματι, αν $x, y \in I$ και $\lambda \in (0, 1)$, θεωρούμε τον χώρο πιθανότητας με $\Omega = \{x, y\}$, $\mathcal{F} = \mathcal{P}(\Omega)$, $\mathbf{P}(\{x\}) = \lambda$, $\mathbf{P}(\{y\}) = 1 - \lambda$ και την τυχαία μεταβλητή $X(\omega) = \omega$ για κάθε $\omega \in \Omega$. Τότε η (5.7) γράφεται

$$\Phi(\lambda x + (1 - \lambda)y) \leq \lambda\Phi(x) + (1 - \lambda)\Phi(y),$$

που είναι ο ορισμός της κυρτότητας.

5.5 Κατασκευή διακριτών τυχαίων μεταβλητών

Η ανισότητα Markov δεν ισχύει χωρίς την υπόθεση $X \geq 0$. Αν γνωρίζουμε ότι $\mathbf{E}(X) = 1$, δεν έπεται ότι $\mathbf{P}(X \geq 4) \leq 1/4$. Ένα αντιπαράδειγμα είναι το εξής. Έστω τυχαία μεταβλητή X που παίρνει τις τιμές 10, -80 με πιθανότητες 9/10 και 1/10 αντίστοιχα. Τότε $\mathbf{E}(X) = 1$ αλλά $\mathbf{P}(X \geq 4) = 9/10$.

Πώς όμως ξέρουμε ότι υπάρχει τέτοια τυχαία μεταβλητή X ; Για την κατασκευή παραδειγμάτων, συχνά θα χρειαστεί να πούμε:

$$\text{θεωρούμε τυχαία μεταβλητή } X \text{ που παίρνει την τιμή } a_i \text{ με πιθανότητα } p_i \text{ για κάθε } i \in I, \quad (5.8)$$

όπου $\{a_i : i \in I\} \subset \mathbb{R}$, $I = \mathbb{N}^+$ ή $I = \{1, 2, \dots, n\}$ για κάποιο $n \in \mathbb{N}^+$, και $(p_i)_{i \in I}$ είναι στοιχεία του $[0, 1]$ με $\sum_{i \in I} p_i = 1$. Τα a_i είναι διαφορετικά ανά δύο μεταξύ τους.

Μια κατασκευή της X είναι ως εξής: Θέτουμε $s_0 := 0$, $s_i := p_1 + p_2 + \dots + p_i$ για κάθε $i \in I$. Ως χώρο πιθανότητας παίρνουμε τον $([0, 1], \mathcal{B}([0, 1]), \lambda)$ και θέτουμε $X : [0, 1] \rightarrow \mathbb{R}$ με $X(\omega) = a_i$ αν $\omega \in [s_{i-1}, s_i)$. Τότε $\lambda(\{X = a_i\}) = \lambda([s_{i-1}, s_i)) = p_i$, δηλαδή η X έχει την ιδιότητα που θέλουμε.

Στο εξής θα περιγράψουμε τυχαίες μεταβλητές όπως στην (5.8) χωρίς να δικαιολογούμε την ύπαρξή τους. Περισσότερα για την κατασκευή τυχαίων μεταβλητών με τιμές στο \mathbb{R} θα δούμε στην Παράγραφο 7.3.

5.6 Οι χώροι \mathcal{L}^p με $p \in [1, \infty)$

Ορισμός 5.18. Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας, $X : \Omega \rightarrow [-\infty, \infty]$ τυχαία μεταβλητή και $p \in [1, \infty)$. Ορίζουμε

$$\|X\|_p := \{\mathbf{E}(|X|^p)\}^{1/p}$$

και

$$\mathcal{L}^p(\Omega, \mathcal{F}, \mathbf{P}) := \{X \mid X : \Omega \rightarrow [-\infty, \infty] \text{ τυχαία μεταβλητή και } \|X\|_p < \infty\}.$$

Όταν είναι σαφές ποιος είναι ο χώρος Ω και ποια η σ -άλγεβρα \mathcal{F} , θα γράφουμε $\mathcal{L}^p(\mathbf{P})$ αντί $\mathcal{L}^p(\Omega, \mathcal{F}, \mathbf{P})$, ενώ όταν είναι σαφές και ποιο είναι το μέτρο \mathbf{P} , τότε γράφουμε απλώς \mathcal{L}^p .

Παρατήρηση 5.19. Αποδεικνύεται ότι η συνάρτηση $\|\cdot\|_p : \mathcal{L}^p \rightarrow [0, \infty)$ ικανοποιεί τις ιδιότητες:

- (i) $\|\lambda X\|_p = |\lambda| \|X\|_p$ για κάθε $\lambda \in \mathbb{R}$ και $X \in \mathcal{L}^p$.
- (ii) $\|X + Y\|_p \leq \|X\|_p + \|Y\|_p$ για κάθε $X, Y \in \mathcal{L}^p$.

Έπεται ότι το σύνολο $\mathcal{L}^p(\mathbf{P})$, εφοδιασμένο με τις πράξεις της πρόσθεσης συναρτήσεων και πολλαπλασιασμού αριθμού με συνάρτηση, είναι διανυσματικός χώρος.

Πρόταση 5.20. (Ανισότητα Cauchy-Schwarz) Έστω $X, Y : \Omega \rightarrow [-\infty, \infty]$ τυχαίες μεταβλητές στοιχεία του \mathcal{L}^2 . Τότε $XY \in \mathcal{L}^1$ και

$$|\mathbf{E}(XY)| \leq \|X\|_2 \|Y\|_2.$$

Απόδειξη. Επειδή $2|XY| \leq X^2 + Y^2$, έπεται ότι $XY \in \mathcal{L}^1$. Έπειτα, για κάθε $\lambda \in \mathbb{R}$ έχουμε

$$0 \leq \mathbf{E}((\lambda X + Y)^2) = \lambda^2 \mathbf{E}(X^2) + 2\lambda \mathbf{E}(XY) + \mathbf{E}(Y^2). \quad (5.9)$$

Η διακρίνουσα της τετραγωνικής μορφής ως προς λ στην (5.9) είναι

$$4 \mathbf{E}(XY)^2 - 4 \mathbf{E}(X^2) \mathbf{E}(Y^2),$$

και πρέπει να είναι ≤ 0 γιατί η μορφή είναι μη αρνητική για κάθε $\lambda \in \mathbb{R}$. Άρα

$$|\mathbf{E}(XY)| \leq \mathbf{E}(X^2)^{1/2} \mathbf{E}(Y^2)^{1/2},$$

το οποίο είναι το ζητούμενο. ■

Γενικότερη της ανισότητας Cauchy-Schwarz είναι η ανισότητα Hölder. Τη διατυπώνουμε, χωρίς απόδειξη, στην επόμενη πρόταση.

Πρόταση 5.21. (Ανισότητα Hölder) Έστω $p, q \in (1, \infty)$ με $p^{-1} + q^{-1} = 1$ και $X, Y : \Omega \rightarrow [-\infty, \infty]$ τυχαίες μεταβλητές με $\|X\|_p < \infty, \|Y\|_q < \infty$. Τότε $XY \in \mathcal{L}^1$ και

$$|\mathbf{E}(XY)| \leq \|X\|_p \|Y\|_q.$$

Πρόταση 5.22. Έστω X τυχαία μεταβλητή με τιμές στο $[-\infty, \infty]$. Τότε, για $1 \leq r < s$, ισχύει ότι

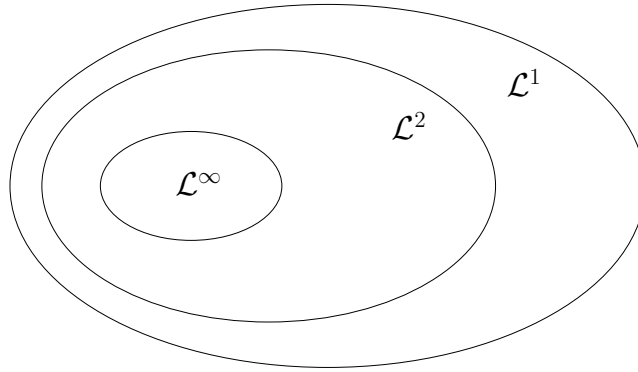
$$\|X\|_r \leq \|X\|_s.$$

Απόδειξη. Αν $X \notin \mathcal{L}^s$, τότε το δεξί μέλος της ανισότητας είναι ∞ και έτσι η ανισότητα ισχύει προφανώς. Αν $X \in \mathcal{L}^s$ τότε εφαρμόζουμε την ανισότητα Hölder πιο πάνω όπου τη θέση της X έχει η $|X|^r$, τη θέση της Y έχει η σταθερή συνάρτηση 1, και $p = s/r, q = s/(s-r)$. Τότε,

$$\mathbf{E}|X|^r = \mathbf{E}(|X|^r \cdot 1) \leq \{\mathbf{E}(|X|^s)\}^{r/s} \{\mathbf{E}(1^q)\}^{1/q} = \{\mathbf{E}(|X|^s)\}^{r/s},$$

και έτσι προκύπτει το ζητούμενο. ■

Η Πρόταση 5.22 μας λέει ότι αν $1 \leq r < s$, τότε $\mathcal{L}^s(\mathbf{P}) \subset \mathcal{L}^r(\mathbf{P})$ (Σχήμα 5.3). Ο εγκλεισμός αυτός όμως έπεται και πιο εύκολα αν παρατηρήσουμε ότι $|X|^r \leq |X|^s + 1$ (το 1 καλύπτει την περίπτωση που $|X(\omega)| < 1$).



Σχήμα 5.3: $\mathcal{L}^s \subset \mathcal{L}^r$ για $1 \leq r < s$.

Ορισμός 5.23. Έστω $X, Y : \Omega \rightarrow \mathbb{R}$ τυχαίες μεταβλητές ώστε $\mathbf{E}|X|, \mathbf{E}|Y| < \infty$ και η $\mathbf{E}(XY)$ ορίζεται (στο $[-\infty, \infty]$). **Συνδιακύμανση** των X, Y ονομάζουμε την ποσότητα

$$\text{Cov}(X, Y) = \mathbf{E}\{(X - \mathbf{E}X)(Y - \mathbf{E}Y)\},$$

η οποία είναι στοιχείο του $[-\infty, \infty]$.

Πρόταση 5.24. Έστω $X, Y \in \mathcal{L}^2$. Τότε

$$|\text{Cov}(X, Y)| \leq \sqrt{\text{Var}(X)} \sqrt{\text{Var}(Y)}.$$

Απόδειξη. Η ανισότητα Cauchy-Schwarz δίνει

$$\begin{aligned} |\text{Cov}(X, Y)| &= |\mathbf{E}\{(X - \mathbf{E}X)(Y - \mathbf{E}Y)\}| \\ &\leq \sqrt{\mathbf{E}\{(X - \mathbf{E}X)^2\}} \sqrt{\mathbf{E}\{(Y - \mathbf{E}Y)^2\}} = \sqrt{\text{Var}(X)} \sqrt{\text{Var}(Y)}. \end{aligned}$$

■

5.7 Ο χώρος \mathcal{L}^∞

Μια τυχαία μεταβλητή $X : \Omega \rightarrow [-\infty, \infty]$ στον $(\Omega, \mathcal{F}, \mathbf{P})$ λέμε ότι είναι **φραγμένη με πιθανότητα 1** αν υπάρχει $M \in \mathbb{R}$ ώστε $\mathbf{P}(|X| \leq M) = 1$. Έπειτα, για κάθε τυχαία μεταβλητή $X : \Omega \rightarrow [-\infty, \infty]$ θέτουμε

$$\begin{aligned} \text{essinf } X &:= \sup\{M \in \mathbb{R} : \mathbf{P}(X \geq M) = 1\}, \\ \text{esssup } X &:= \inf\{M \in \mathbb{R} : \mathbf{P}(X \leq M) = 1\}, \\ \|X\|_\infty &:= \inf\{M > 0 : \mathbf{P}(|X| \leq M) = 1\}. \end{aligned}$$

Οι ποσότητες αυτές ονομάζονται **ουσιώδες infimum**, **ουσιώδες supremum**, και **άπειρο νόρμα** της X αντίστοιχα. Υπενθυμίζουμε ότι $\inf \emptyset = \infty$ και $\sup \emptyset = -\infty$. Τέλος, θέτουμε

$$\mathcal{L}^\infty := \{X \mid X : \Omega \rightarrow [-\infty, \infty] \text{ είναι τυχαία μεταβλητή και } \|X\|_\infty < \infty\}.$$

Επειδή μια σταθερή συνάρτηση έχει πεπερασμένη μέση τιμή (το \mathbf{P} είναι πεπερασμένο μέτρο), έχουμε ότι

$$\mathcal{L}^\infty \subset \mathcal{L}^p$$

για κάθε $1 \leq p < \infty$.

5.8 Τα βασικά οριακά θεωρήματα

Έστω (X, \mathcal{A}, μ) χώρος μέτρου και ακολουθία $(f_n)_{n \in \mathbb{N}}$ μετρήσιμων συναρτήσεων με τιμές στο $[-\infty, \infty]$ που συγκλίνουν σημειακά σε μια συνάρτηση f . Πολλές φορές μας ενδιαφέρει ο υπολογισμός του ορίου $\lim_{n \rightarrow \infty} \int f_n \, d\mu$, και μπαίνουμε στον πειρασμό να μαντέψουμε ότι

$$\lim_{n \rightarrow \infty} \int f_n \, d\mu = \int \lim_{n \rightarrow \infty} f_n \, d\mu, \quad (5.10)$$

δηλαδή το όριο μπαίνει μέσα στο ολοκλήρωμα. Αυτό όμως δεν γίνεται πάντοτε. Το πρόβλημα αυτό είναι το αντικείμενο των βασικών θεωρημάτων σύγκλισης για το ολοκλήρωμα Lebesgue. Τα διατυπώνουμε αλλά παραλείπουμε τις αποδείξεις τους.

Θεώρημα 5.25 (Θεώρημα μονότονης σύγκλισης). Έστω $(f_n)_{n \in \mathbb{N}}$, με $f_n : X \rightarrow [0, +\infty]$ για κάθε $n \in \mathbb{N}$, αύξουσα ακολουθία μετρήσιμων συναρτήσεων. Θέτουμε $f = \lim_{n \rightarrow \infty} f_n$. Τότε

$$\lim_{n \rightarrow \infty} \int f_n \, d\mu = \int f \, d\mu.$$

Το όριο $\lim_{n \rightarrow \infty} f_n$ υπάρχει γιατί η $(f_n)_{n \in \mathbb{N}}$ είναι αύξουσα. Και όμοια, το όριο στο αριστερό μέλος της τελευταίας ισότητας υπάρχει γιατί η ακολουθία των ολοκληρωμάτων είναι αύξουσα.

Θεώρημα 5.26 (Λήμμα Fatou). Έστω $(f_n)_{n \in \mathbb{N}}$, με $f_n : X \rightarrow [0, \infty]$, ακολουθία μετρήσιμων συναρτήσεων. Τότε

$$\int \underline{\lim}_{n \rightarrow \infty} f_n \, d\mu \leq \underline{\lim}_{n \rightarrow \infty} \int f_n \, d\mu.$$

Θεώρημα 5.27 (Θεώρημα κυριαρχημένης σύγκλισης). Έστω $f_n : X \rightarrow \mathbb{R}$ μετρήσιμη για κάθε $n \in \mathbb{N}$ έτσι ώστε $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ και $|f_n(x)| \leq g(x)$ σχεδόν παντού, όπου $g : X \rightarrow [0, \infty]$ είναι μετρήσιμη με $\int g \, d\mu < \infty$. Τότε $\int |f| \, d\mu < \infty$ και

$$\lim_{n \rightarrow \infty} \int f_n \, d\mu = \int f \, d\mu. \quad (5.11)$$

Όταν $|f_n| \leq g$ για κάθε $n \in \mathbb{N}$, λέμε ότι η ακολουθία $(f_n)_{n \in \mathbb{N}}$ κυριαρχείται από τη συνάρτηση g . Η κρίσιμη συνθήκη του θεωρήματος κυριαρχημένης σύγκλισης είναι ότι η ακολουθία $(f_n)_{n \in \mathbb{N}}$ κυριαρχείται από ολοκληρώσιμη συνάρτηση.

Σε έναν χώρο πεπερασμένου μέτρου, οι σταθερές συναρτήσεις είναι ολοκληρώσιμες. Γιατί μια $g = M$ (όπου $M \in \mathbb{R}$ σταθερά) έχει ολοκλήρωμα $M\mu(X)$, το οποίο είναι πραγματικός αριθμός. Έτσι το θεώρημα κυριαρχημένης σύγκλισης έχει την εξής χρήσιμη συνέπεια.

Θεώρημα 5.28 (Θεώρημα φραγμένης σύγκλισης). Έστω (X, \mathcal{A}, μ) χώρος πεπερασμένου μέτρου και $f_n : X \rightarrow \mathbb{R}$ μετρήσιμη για κάθε $n \in \mathbb{N}$, με $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ και $|f_n(x)| \leq M$ σχεδόν παντού, όπου $M < \infty$ σταθερά. Τότε $\int |f| d\mu < \infty$ και

$$\lim_{n \rightarrow \infty} \int f_n d\mu = \int f d\mu.$$

Με άμεση εφαρμογή του θεωρήματος φραγμένης σύγκλισης βλέπουμε ότι

$$\lim_{n \rightarrow \infty} \int_0^A |\sin x|^n dx = 0$$

για κάθε $A > 0$. Αυτό γιατί το μέτρο Lebesgue στο $[0, A]$ είναι πεπερασμένο μέτρο, η ακολουθία $|\sin x|^n$ είναι φραγμένη από το 1, και $\lim_{n \rightarrow \infty} |\sin x|^n = 0$ για όλα τα $x \in [0, A]$ εκτός από ένα πεπερασμένο σύνολο (αυτό των αριθμών του $[0, A]$ που είναι της μορφής $(2k + 1)\pi/2$ με $k \in \mathbb{Z}$) το οποίο όμως έχει μέτρο Lebesgue 0.

Αντιπαράδειγμα [Αποτυχία ισχύος της (5.10)]: Θεωρούμε τον χώρο πιθανότητας $([0, 1], \mathcal{B}([0, 1]), \mathbf{P})$, όπου \mathbf{P} είναι το μέτρο Lebesgue περιορισμένο στο $[0, 1]$ (Παράδειγμα 2.4). Θέτουμε

$$X_n(x) = n \mathbf{1}_{(0, 1/n)}(x) \text{ για κάθε } x \in [0, 1].$$

Η X_n είναι απλή τυχαία μεταβλητή, και $\lim_{n \rightarrow \infty} X_n = 0$. Έχουμε ότι

$$\mathbf{E}(\lim_{n \rightarrow \infty} X_n) = \mathbf{E}(0) = 0$$

και

$$\mathbf{E}(X_n) = n \mathbf{P}((0, 1/n]) = n \frac{1}{n} = 1.$$

Άρα $\mathbf{E}(\lim_{n \rightarrow \infty} X_n) < \lim_{n \rightarrow \infty} \mathbf{E}(X_n)$. Δηλαδή έχουμε γνήσια ανισότητα στο λήμμα Fatou, και το θεώρημα κυριαρχημένης σύγκλισης δεν εφαρμόζεται. Αυτό δεν μας κάνει εντύπωση γιατί η ακολουθία X_n δεν κυριαρχείται από κάποια ολοκληρώσιμη συνάρτηση. Πράγματι, η μικρότερη g που ικανοποιεί την $|X_n(x)| \leq g(x)$ για κάθε $x \in [0, 1]$ και $n \geq 1$ είναι η $\sup_{n \geq 1} X_n(x) = [1/x] \mathbf{1}_{x \in (0, 1]}$ (Άσκηση), της οποίας το ολοκλήρωμα ως προς το μέτρο \mathbf{P} είναι ∞ .

Παράδειγμα 5.29 (Το θεώρημα κυριαρχημένης σύγκλισης και ένα ερώτημα απειροστικού λογισμού). Θα υπολογίσουμε το όριο $\lim_{n \rightarrow \infty} I_n$ όπου

$$I_n := n^{k+1} \int_0^1 \left(\frac{1-x}{1+x} \right)^n x^k dx$$

με $k \in (-1, \infty)$. Η αντικατάσταση $y = nx$ δίνει

$$I_n = \int_0^n \left(\frac{n-y}{n+y} \right)^n y^k dy = \int_0^\infty \left(\frac{n-y}{n+y} \right)^n y^k \mathbf{1}_{y \in [0, n]} dy.$$

Για σταθερό $y > 0$, ο ολοκληρωτέος συγκλίνει στο $e^{-2y} y^k$ αφού

$$\left(\frac{n-y}{n+y} \right)^n \mathbf{1}_{y \in [0, n]} = \left(1 - \frac{2y}{n+y} \right)^n \mathbf{1}_{y \in [0, n]} \rightarrow e^{-2y}$$

Επίσης, φράσσουμε τον ολοκληρωτέο ως εξής

$$0 \leq \left(\frac{n-y}{n+y}\right)^n y^k \mathbf{1}_{y \in [0, n]} \leq e^{-\frac{2yn}{n+y}} y^k \mathbf{1}_{y \in [0, n]} \leq e^{-2y} y^k =: g(y).$$

Στην πρώτη ανισότητα χρησιμοποιήσαμε την $1+x \leq e^x$, ενώ στη δεύτερη το ότι $y \in [0, n]$ για τα y που το αριστερό μέλος είναι θετικό. Η g έχει $\int_0^\infty g(x) dx < \infty$, οπότε το θεώρημα κυριαρχημένης σύγκλισης εφαρμόζεται και δίνει ότι

$$\lim_{n \rightarrow \infty} I_n = \int_0^\infty y^k e^{-2y} dy = 2^{-k-1} \Gamma(k+1).$$

Παρατήρηση 5.30. (Περικοπή συνάρτησης) Για $f : X \rightarrow \mathbb{R}$ μετρήσιμη, ορίζουμε την ακολουθία μετρήσιμων συναρτήσεων $(f_n)_{n \geq 1}$ με

$$f_n(x) = f(x) \cdot \mathbf{1}_{|f(x)| \leq n} = \begin{cases} f(x) & \text{αν } |f(x)| \leq n, \\ 0 & \text{αν } |f(x)| > n, \end{cases}$$

για κάθε $n \in \mathbb{N}^+$ και $x \in X$. Λέμε ότι η f_n προκύπτει από την f με περικοπή. Η ακολουθία αυτή έχει τις εξής ιδιότητες:

- (i) Κάθε f_n είναι φραγμένη.
- (ii) $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ για κάθε $x \in X$.

Δηλαδή είναι ακολουθία φραγμένων συναρτήσεων που προσεγγίζουν σημειακά την f .

Δεδομένου ότι μια σειρά είναι το όριο των μερικών αθροισμάτων της και ότι το ολοκλήρωμα είναι γραμμικό, τα παραπάνω θεωρήματα δίνουν το εξής πόρισμα.

Πόρισμα 5.31. Έστω $(f_n)_{n \geq 1}$, με $f_n : X \rightarrow [-\infty, \infty]$, ακολουθία μετρήσιμων συναρτήσεων.

- (i) (Θεώρημα Βερρο-Λεβί) Αν f_n μη αρνητική για κάθε $n \in \mathbb{N}^+$, τότε

$$\int \left(\sum_{n=1}^{\infty} f_n \right) d\mu = \sum_{n=1}^{\infty} \int f_n d\mu. \quad (5.12)$$

- (ii) Αν οι f_n παίρνουν τιμές στο $[-\infty, \infty]$ και $\sum_{n=1}^{\infty} \int |f_n| d\mu < \infty$, τότε η σειρά $\sum_{n=1}^{\infty} f_n$ συγκλίνει σχεδόν παντού σε μια μετρήσιμη συνάρτηση με τιμές στο \mathbb{R} , ισχύει η (5.12), και τα δύο μέλη της είναι πραγματικοί αριθμοί.

Απόδειξη. Θέτουμε $g_n = \sum_{k=1}^n f_k$ για κάθε $n \geq 1$.

- (i) Τότε η $(g_n)_{n \geq 1}$ είναι αύξουσα ακολουθία μη αρνητικών συναρτήσεων και αν $g = \sum_{k=1}^{\infty} f_k$, ισχύει ότι $\lim_{n \rightarrow \infty} g_n = g$. Επίσης, $\int g_n d\mu = \sum_{k=1}^n \int f_k d\mu$ λόγω γραμμικότητας. Το ζητούμενο προκύπτει από το θεώρημα μοτότονης σύγκλισης (Θεώρημα 5.25).

- (ii) Εφαρμόζουμε το (i) για την ακολουθία $(|f_n|)_{n \geq 1}$. Τότε

$$\int \left(\sum_{n=1}^{\infty} |f_n| \right) d\mu = \sum_{n=1}^{\infty} \int |f_n| d\mu. \quad (5.13)$$

Όμως

$$\left| \sum_{k=1}^n f_k \right| \leq \sum_{k=1}^{\infty} |f_k|$$

για κάθε $n \geq 1$ και, από υπόθεση, η $g = \sum_{k=1}^{\infty} |f_k|$ έχει πεπερασμένο ολοκλήρωμα. Συνεπώς, εφαρμόζοντας το θεώρημα κυριαρχημένης σύγκλισης (Θεώρημα 5.27) για την $(g_n)_{n \geq 1}$, έχουμε ότι

$$\int \left(\sum_{n=1}^{\infty} f_n \right) d\mu = \sum_{n=1}^{\infty} \int f_n d\mu,$$

και από την (5.13) και την Πρόταση 5.12(iii), ισχύει ότι η $\sum_{n=1}^{\infty} f_n$ παίρνει πραγματικές τιμές σχεδόν παντού. ■

Παρατήρηση 5.32. Οι σειρές $\sum_{n=1}^{\infty} f_n$, $\sum_{n=1}^{\infty} \int f_n d\mu$ στο (i) και η $\sum_{n=1}^{\infty} |f_n|$ στο (ii) συγκλίνουν, με ενδεχόμενη τιμή το ∞ , γιατί είναι σειρές μη αρνητικών όρων.

Παράδειγμα 5.33. (Ορισμός μέτρου μέσω πυκνότητας) Έστω $f : X \rightarrow [0, \infty]$ μετρήσιμη συνάρτηση σε έναν χώρο μέτρου (X, \mathcal{A}, μ) . Τότε η συνάρτηση $\nu : \mathcal{A} \rightarrow [0, \infty]$ με

$$\nu(A) := \int_A f d\mu = \int f \mathbf{1}_A d\mu$$

για κάθε $A \in \mathcal{A}$ είναι μέτρο. Επιπλέον, για $A \in \mathcal{A}$, ισχύει ότι αν $\mu(A) = 0$, τότε $\nu(A) = 0$.

Πράγματι, η ν είναι μη αρνητική και $\nu(\emptyset) = \int f \mathbf{1}_{\emptyset} d\mu = 0$. Έπειτα, για $(A_n)_{n \geq 1}$ ακολουθία ξένων ανα δύο στοιχείων της \mathcal{A} , έχουμε $\mathbf{1}_{\cup_{n=1}^{\infty} A_n} = \sum_{n=1}^{\infty} \mathbf{1}_{A_n}$. Συνεπώς

$$\nu(\cup_{n=1}^{\infty} A_n) = \int \left(f \sum_{n=1}^{\infty} \mathbf{1}_{A_n} \right) d\mu = \sum_{n=1}^{\infty} \int f \mathbf{1}_{A_n} d\mu = \sum_{n=1}^{\infty} \nu(A_n).$$

Στη δεύτερη ισότητα χρησιμοποιούμε το θεώρημα Beppo-Levi [Πόρισμα 5.31 (i)]. Τέλος, αν $\mu(A) = 0$, τότε $\mu(f \mathbf{1}_A = 0) = 1$, και από την Πρόταση 5.12 (i) έχουμε ότι $\nu(A) = \int f \mathbf{1}_A d\mu = \int 0 d\mu = 0$.

Παρατήρηση 5.34. Η συνάρτηση f στο Παράδειγμα 5.33 λέγεται **πυκνότητα** του ν ως προς το μέτρο μ καθώς και **παράγωγος Radon-Nikodym** του ν ως προς μ . Γράφουμε

$$f = \frac{d\nu}{d\mu}. \quad (5.14)$$

Αν επιπλέον $\int f d\mu = 1$, τότε το ν είναι μέτρο πιθανότητας στον (X, \mathcal{A}) .

Ασκήσεις

5.1 (Η αρχή εγκλεισμού-αποκλεισμού για πιθανότητες) Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας. Να δειχθεί ότι για κάθε $A_1, A_2, \dots, A_n \in \mathcal{F}$ ισχύει

$$\mathbf{P}(\cup_{i=1}^n A_i) = \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \mathbf{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}).$$

5.2* Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας. Αν $n \in \mathbb{N}^+$ και τα $A_1, A_2, \dots, A_n \in \mathcal{F}$ ικανοποιούν $\mathbf{P}(A_1) + \dots + \mathbf{P}(A_n) > k-1$ για κάποιον $k \in \mathbb{N}^+$, τότε υπάρχουν $1 \leq i_1 < \dots < i_k \leq n$ με $\mathbf{P}(A_{i_1} \cap \dots \cap A_{i_k}) > 0$.

Το ολοκλήρωμα Lebesgue

5.3 (Η ανισότητα Jensen) Έστω X τυχαία μεταβλητή με τιμές σε ένα διάστημα $I \subset \mathbb{R}$ και $\phi : I \rightarrow \mathbb{R}$ κυρτή συνάρτηση. Αν οι $\mathbf{E}X, \mathbf{E}\{\phi(X)\}$ ορίζονται και είναι πραγματικοί αριθμοί, τότε

$$\phi(\mathbf{E}X) \leq \mathbf{E}\{\phi(X)\}.$$

[Υπόδειξη: Έστω $a := \mathbf{E}X$. Αν το a είναι εσωτερικό σημείο του I , τότε υπάρχει $\lambda \in \mathbb{R}$ ώστε $\phi(x) \geq \phi(a) + \lambda(x - a)$ για κάθε $x \in I$ (απειροστικός λογισμός). Θέτουμε όπου x την τυχαία μεταβλητή X . Αν το a είναι άκρο του I , τότε με πιθανότητα 1 η X είναι σταθερή.]

5.4 Έστω X τυχαία μεταβλητή με τιμές στο $[0, \infty)$. Τότε

$$(\mathbf{E}X)^p \begin{cases} \leq \mathbf{E}(X^p) & \text{αν } p \geq 1, \\ \geq \mathbf{E}(X^p) & \text{αν } 0 \leq p < 1. \end{cases}$$

Αν οι $\mathbf{E}X, \mathbf{E}(\log X)$ ορίζονται και είναι πραγματικοί αριθμοί, τότε

$$\log \mathbf{E}X \geq \mathbf{E}(\log X).$$

5.5 Έστω X τυχαία μεταβλητή και έστω ότι για κάποιο $a > 0$ ισχύει $\mathbf{E}(e^{aX}) < \infty$. Να δειχθεί ότι υπάρχει $C > 0$ σταθερά ώστε για κάθε $t \in \mathbb{R}$ να ισχύει $\mathbf{P}(X > t) \leq Ce^{-at}$. Δηλαδή η «ουρά» της X προς τα δεξιά φθίνει γρήγορα, τουλάχιστον με ταχύτητα e^{-at} .

5.6 Έστω X μη αρνητική τυχαία μεταβλητή με $0 < \mathbf{E}X < \infty$ και $a \in (0, 1)$. Τότε

(α)

$$\mathbf{P}(X \leq a \mathbf{E}X) \leq \frac{\text{Var}(X)}{(1-a)^2(\mathbf{E}X)^2}.$$

(β)* (Ανισότητα Paley-Zygmund)

$$\mathbf{P}(X > a \mathbf{E}X) \geq (1-a)^2 \frac{(\mathbf{E}X)^2}{\mathbf{E}(X^2)}.$$

5.7 Έστω X, Y τυχαίες μεταβλητές με τιμές στο $(0, \infty)$ ώστε $XY \geq 1$ παντού. Να δειχθεί ότι

$$\mathbf{E}(X) \mathbf{E}(Y) \geq 1.$$

Ειδικότερα

$$\mathbf{E}\left(\frac{1}{X}\right) \geq \frac{1}{\mathbf{E}(X)}.$$

5.8 Θεωρούμε στον $\Omega = [2, 3]$ το μέτρο Lebesgue λ , που είναι μέτρο πιθανότητας. Θέτουμε

$$f(x) = \begin{cases} x^2 & \text{αν } x \in [2, 3] \setminus \{2 + n^{-1} : n \in \mathbb{N}^+\}, \\ (-1)^n n & \text{αν } x = 2 + n^{-1} \text{ με } n \in \mathbb{N}^+. \end{cases}$$

Να δειχθεί ότι $\inf f = -\infty, \sup f = \infty, \text{essinf } f = 4, \text{esssup } f = 9$.

5.9 Έστω X τυχαία μεταβλητή με τιμές στο \mathbb{R} . Να δειχθεί ότι

$$\lim_{n \rightarrow \infty} \{\mathbf{E}(e^{nX})\}^{1/n} = e^{\text{esssup } X}.$$

5.10 Έστω X τυχαία μεταβλητή με τιμές στο \mathbb{R} . Να δείξετε ότι $\lim_{n \rightarrow \infty} \mathbf{P}(|X| > n) = 0$.

5.11 Να υπολογιστεί το όριο $\lim_{n \rightarrow \infty} \int_1^\infty f_n(x) dx$ στην περίπτωση που

(α) $f_n(x) = \{\sin(nx)\}/(1 + nx^3)$.

(β) $f_n(x) = \frac{1}{1+nx} \mathbf{1}_{[1, e^n]}(x)$.

Το ολοκλήρωμα Lebesgue

5.12 Έστω $f : [0, 1] \rightarrow \mathbb{R}$ συνεχής. Να δειχθεί ότι

$$\lim_{n \rightarrow \infty} n \int_0^1 x^n f(x) dx = f(1).$$

5.13 (Κυριαρχημένη σύγκλιση με υπεραριθμισμό σύνολο δεικτών) Έστω (X, \mathcal{A}, μ) χώρος μέτρου και για κάθε $t > 0$ μετρήσιμη συνάρτηση $f_t : X \rightarrow \mathbb{R}$. Υποθέτουμε ότι υπάρχει το όριο $\lim_{t \rightarrow \infty} f_t(x) =: f(x)$ για κάθε $x \in X$ και υπάρχει $g : X \rightarrow [0, \infty]$ μετρήσιμη με $\int g(x) d\mu(x) < \infty$ και $|f_t(x)| \leq g(x)$ για κάθε $x \in X$ και $t > 0$. Να δειχθεί ότι η f είναι μετρήσιμη, ολοκληρώσιμη και

$$\lim_{t \rightarrow \infty} \int f_t(x) d\mu(x) = \int f(x) d\mu(x).$$

5.14 Έστω $X \in \mathcal{L}^1(\mathbf{P})$ και $E_n := \{|X| \geq n\}$ για κάθε $n \in \mathbb{N}$. Να δείξετε ότι $n\mathbf{P}(E_n) \rightarrow 0$ για $n \rightarrow \infty$.

5.15 Έστω $1 \leq r < s$ και $X \in \mathcal{L}^r$. Θέτουμε $X_n := X\mathbf{1}_{|X| \leq n}$ για κάθε $n \in \mathbb{N}^+$. Να δειχθεί ότι $X_n \in \mathcal{L}^s$ και $\lim_{n \rightarrow \infty} \mathbf{E}(|X_n - X|^r) = 0$.

5.16 Έστω X τυχαία μεταβλητή με τιμές στο $[0, \infty]$. Να δείξετε ότι

(α)

$$\lim_{\varepsilon \rightarrow 0^+} \frac{1}{\varepsilon} \mathbf{E}(X; X < \varepsilon) = 0,$$

(β)

$$\lim_{M \rightarrow \infty} \frac{1}{M} \mathbf{E}(X; X < M) = 0.$$

5.17* Έστω X τυχαία μεταβλητή με τιμές στο \mathbb{R} ώστε $\mathbf{E}(X^2) = \infty$. Να δειχθεί ότι

$$\lim_{M \rightarrow \infty} \frac{\{\mathbf{E}(X\mathbf{1}_{|X| \leq M})\}^2}{\mathbf{E}(X^2\mathbf{1}_{|X| \leq M})} = 0.$$

5.18 Έστω X τυχαία μεταβλητή με τιμές στο $\mathbb{N} \cup \{\infty\}$. Να δείξετε ότι

$$\mathbf{E}(X) = \sum_{k \geq 1} \mathbf{P}(X \geq k).$$

5.19 Έστω X τυχαία μεταβλητή με τιμές στο $[0, \infty]$. Να δείξετε ότι

$$\sum_{k=1}^{\infty} \mathbf{P}(X \geq k) \leq \mathbf{E}X \leq 1 + \sum_{k=1}^{\infty} \mathbf{P}(X \geq k).$$

5.20* Έστω X τυχαία μεταβλητή με τιμές στο $[0, \infty)$ και $\mathbf{E}(X) < \infty$. Έστω και $c > 1$. Να δείξετε ότι

$$\sum_{k=1}^{\infty} c^k \mathbf{P}(X \geq c^k) < \infty.$$

5.21 Έστω X τυχαία μεταβλητή στον χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$ με $\mathbf{E}|X| < \infty$ και $\mathbf{E}(X\mathbf{1}_A) = 0$ για κάθε $A \in \mathcal{F}$. Να δειχθεί ότι $X = 0$ με πιθανότητα 1.

5.22 (Λήμμα Scheffe) Έστω (X, \mathcal{A}, μ) χώρος μέτρου, και $(f_n)_{n \in \mathbb{N}^+}$ ακολουθία συναρτήσεων ώστε $f_n : X \rightarrow [0, \infty)$ και $\int f_n d\mu < \infty$ για κάθε n . Υποθέτουμε ότι $\lim_{n \rightarrow \infty} f_n = f$ σχεδόν παντού και $\int f d\mu < \infty$. Να δειχθεί ότι

$$\lim_{n \rightarrow \infty} \int |f_n - f| d\mu = 0 \Leftrightarrow \lim_{n \rightarrow \infty} \int f_n d\mu = \int f d\mu.$$

5.23 Έστω X τυχαία μεταβλητή με τιμές στο \mathbb{R} .

(α) Να δειχθεί ότι $\mathbf{E}|X| < \infty$ αν και μόνο αν υπάρχει $n \in \mathbb{N}$ με $\mathbf{E}(|X|\mathbf{1}_{|X| > n}) < \infty$.

Το ολοκλήρωμα Lebesgue

(β) Να δειχθεί ότι το $\lim_{n \rightarrow \infty} \mathbf{E}(|X| \mathbf{1}_{|X| > n})$ υπάρχει και είναι 0 ή ∞ .

5.24 Έστω X τυχαία μεταβλητή με τιμές στο \mathbb{R} . Θέτουμε $M_X(t) = \mathbf{E}(e^{tX})$ για κάθε $t \in \mathbb{R}$ [η ροπογεννήτρια της X]. Να δειχθεί ότι η $\log M_X$ είναι κυρτή στο \mathbb{R} .

Μια συνέπεια αυτού είναι ότι το $\{t \in \mathbb{R} : M_X(t) < \infty\}$ είναι διάστημα.

5.25 Έστω X τυχαία μεταβλητή με τιμές στο $[0, \infty)$ ώστε $\mathbf{P}(X > 0) > 0$ και $\mathbf{E}(X^n) < \infty$ για κάθε $n \in \mathbb{N}^+$. Να προσδιοριστεί το όριο

$$\lim_{n \rightarrow \infty} \frac{\mathbf{E}(X^{n+1})}{\mathbf{E}(X^n)}.$$

6

Κατανομή τυχαίας μεταβλητής και ολοκλήρωση

Με όσα έχουμε δει ως τώρα, η μέση τιμή προσδιορίζεται μόνο μέσω της διαδικασίας της Παραγράφου 5.1, η οποία δεν είναι εύχρηστη γενικά. Από την άλλη, στις στοιχειώδεις πιθανότητες η μέση τιμή μιας τυχαίας μεταβλητής με τιμές στο \mathbb{R} , ανάλογα με το είδος της (διακριτή/συνεχής), ορίζεται μέσω ενός αθροίσματος ή ολοκληρώματος. Σε αυτό το κεφάλαιο, θα ανακτήσουμε, ως θεωρήματα, αυτές τις εκφράσεις για τη μέση τιμή.

Κεντρική έννοια σε αυτή τη διαδικασία είναι η κατανομή τυχαίας μεταβλητής.

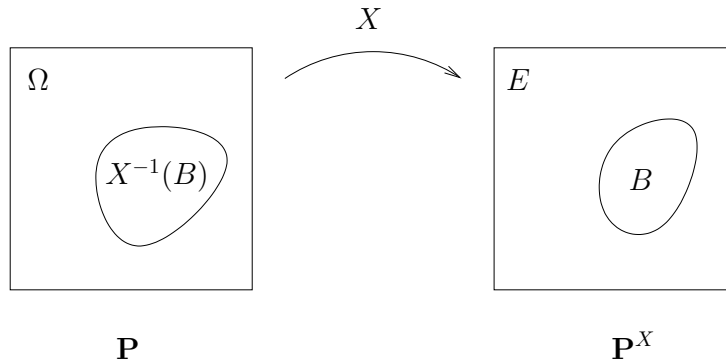
6.1 Ο τύπος αλλαγής μεταβλητής

Ορισμός 6.1. Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας, (E, \mathcal{E}) μετρήσιμος χώρος, και $X : \Omega \rightarrow E$ τυχαία μεταβλητή. Η απεικόνιση $\mathbf{P}^X : \mathcal{E} \rightarrow [0, 1]$ με

$$\mathbf{P}^X(B) := \mathbf{P}(X^{-1}(B)) = \mathbf{P}(X \in B)$$

για κάθε $B \in \mathcal{E}$ είναι μέτρο πιθανότητας και λέγεται **κατανομή** της X .

Το ότι το \mathbf{P}^X είναι μέτρο πιθανότητας αφήνεται ως άσκηση.



Σχήμα 6.1: Η τυχαία μεταβλητή X «μεταφέρει» το μέτρο \mathbf{P} στον χώρο E δίνοντας το μέτρο \mathbf{P}^X .

Η επόμενη πρόταση μεταφέρει τον υπολογισμό ενός ολοκληρώματος από τον Ω στον E .

Πρόταση 6.2. Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας, (E, \mathcal{E}) μετρήσιμος χώρος, και $X : \Omega \rightarrow E$ τυχαία μεταβλητή με κατανομή \mathbf{P}^X . Για κάθε $h : E \rightarrow [0, \infty]$ μετρήσιμη συνάρτηση ισχύει

$$\mathbf{E}_{\mathbf{P}}\{h(X)\} = \mathbf{E}_{\mathbf{P}^X}(h). \tag{6.1}$$

Επίσης, αν η $h : E \rightarrow [-\infty, \infty]$ είναι μετρήσιμη, τότε ή και τα δύο μέλη της (6.1) ορίζονται και είναι ίσα ή και τα δύο δεν ορίζονται.

Το αριστερό μέλος της (6.1) είναι η μέση τιμή της $h \circ X$ στο Ω ως προς το μέτρο \mathbf{P} και το δεξί μέλος της (6.1) είναι η μέση τιμή της h στο E ως προς το μέτρο \mathbf{P}^X . Αυτό γίνεται ακόμη πιο καθαρό αν τη γράψουμε ως

$$\int h(X(\omega)) d\mathbf{P}(\omega) = \int h(x) d\mathbf{P}^X(x).$$

Απόδειξη. Βήμα 1. Αν $h = \mathbf{1}_A$ με $A \in \mathcal{E}$, τότε $h(X(\omega)) = \mathbf{1}_{\{\omega: X(\omega) \in A\}}$. Δηλαδή, $h(X) = \mathbf{1}_{X^{-1}(A)}$ με $X^{-1}(A) \in \mathcal{F}$. Έχουμε λοιπόν

$$\mathbf{E}_{\mathbf{P}}(\mathbf{1}_A(X)) = \mathbf{P}(X^{-1}(A)) = \mathbf{P}^X(A) = \mathbf{E}_{\mathbf{P}^X}(\mathbf{1}_A),$$

άρα η (6.1) ισχύει.

Βήμα 2. Αν η h είναι μη αρνητική απλή, τότε $h = \sum_{i=1}^n a_i \mathbf{1}_{A_i}$, με $a_i \in [0, \infty)$ και $A_i \in \mathcal{E}$ για κάθε $i \in \{1, 2, \dots, n\}$. Τότε

$$\mathbf{E}_{\mathbf{P}}(h(X)) = \mathbf{E}_{\mathbf{P}}\left(\sum_{i=1}^n a_i \mathbf{1}_{A_i}(X)\right) = \sum_{i=1}^n a_i \mathbf{E}_{\mathbf{P}}(\mathbf{1}_{A_i}(X)) = \sum_{i=1}^n a_i \mathbf{E}_{\mathbf{P}^X}(\mathbf{1}_{A_i}) = \mathbf{E}_{\mathbf{P}^X}\left(\sum_{i=1}^n a_i \mathbf{1}_{A_i}\right) = \mathbf{E}_{\mathbf{P}^X}(h).$$

Στην τρίτη ισότητα χρησιμοποιήσαμε το προηγούμενο βήμα.

Βήμα 3. Αν $h \geq 0$ μετρήσιμη, τότε από την Πρόταση 4.12 υπάρχει αύξουσα ακολουθία $(h_n)_{n \in \mathbb{N}}$ μη αρνητικών, απλών, μετρήσιμων συναρτήσεων με $\lim_{n \rightarrow \infty} h_n(x) = h(x)$ για κάθε $x \in \mathcal{E}$. Τότε $\lim_{n \rightarrow \infty} h_n(X(\omega)) = h(X(\omega))$ για κάθε $\omega \in \Omega$ και από το προηγούμενο βήμα

$$\mathbf{E}(h_n(X)) = \mathbf{E}_{\mathbf{P}^X}(h_n) \quad \text{για κάθε } n \in \mathbb{N}.$$

Για $n \rightarrow \infty$, από το θεώρημα μονότονης σύγκλισης (Θεώρημα 5.25) έχουμε

$$\mathbf{E}_{\mathbf{P}}(h(X)) = \lim_{n \rightarrow \infty} \mathbf{E}_{\mathbf{P}}(h_n(X)) = \lim_{n \rightarrow \infty} \mathbf{E}_{\mathbf{P}^X}(h_n) = \mathbf{E}_{\mathbf{P}^X}(h).$$

Βήμα 4. Αν h μετρήσιμη συνάρτηση με τιμές στο $[-\infty, \infty]$, τότε τη γράφουμε ως $h = h^+ - h^-$. Από το προηγούμενο βήμα,

$$\mathbf{E}_{\mathbf{P}}(h^+(X)) = \mathbf{E}_{\mathbf{P}^X}(h^+), \quad (6.2)$$

$$\mathbf{E}_{\mathbf{P}}(h^-(X)) = \mathbf{E}_{\mathbf{P}^X}(h^-). \quad (6.3)$$

Το αριστερό μέλος της (6.1) δεν ορίζεται αν και μόνο αν το αριστερό μέλος των (6.2), (6.3) ισούται με ∞ , ενώ το δεξί μέλος της (6.1) δεν ορίζεται αν και μόνο αν το δεξί μέλος των (6.2), (6.3) ισούται με ∞ . Άρα ή και τα δύο μέλη της (6.1) ορίζονται ή και τα δύο δεν ορίζονται.

Τώρα, στην περίπτωση που και τα δύο μέλη της (6.1) ορίζονται, οι (6.2), (6.3) δίνουν

$$\mathbf{E}_{\mathbf{P}}(h(X)) = \mathbf{E}_{\mathbf{P}}(h^+(X)) - \mathbf{E}_{\mathbf{P}}(h^-(X)) = \mathbf{E}_{\mathbf{P}^X}(h^+) - \mathbf{E}_{\mathbf{P}^X}(h^-) = \mathbf{E}_{\mathbf{P}^X}(h). \quad \blacksquare$$

Η τεχνική απόδειξης της προηγούμενης πρότασης είναι πολύ συνηθισμένη στη Θεωρία Μέτρου. Θα την ονομάζουμε στο εξής **Τυπική Μηχανή**.

Παρατήρηση 6.3. Η τυπική μηχανή. Έστω (X, \mathcal{A}, μ) χώρος μέτρου και ότι θέλουμε να αποδείξουμε ότι μια πρόταση $Q(f)$ ισχύει για όλες τις μετρήσιμες συναρτήσεις $f : X \rightarrow [-\infty, \infty]$. Ακολουθούμε τα εξής βήματα.

- (i) Δείχνουμε την Q για κάθε $f = \mathbf{1}_A$ όπου $A \in \mathcal{A}$.

- (ii) Δείχνουμε την Q για κάθε $f \geq 0$ μετρήσιμη και απλή.
- (iii) Δείχνουμε την Q για κάθε $f \geq 0$ μετρήσιμη.
- (iv) Δείχνουμε την Q για κάθε $f : X \rightarrow [-\infty, \infty]$ μετρήσιμη.

Συνήθως συμβαίνει το εξής: Το (i) είναι συνέπεια ορισμού. Το (ii) έπεται από το (i) και τη γραμμικότητα της ιδιότητας Q . Το (iii) έπεται από το (ii) εφόσον γράψουμε $f = \lim s_n$ για κατάλληλη αύξουσα ακολουθία απλών μη αρνητικών μετρήσιμων συναρτήσεων. Αν η Q εμπλέκει το ολοκλήρωμα της f , χρησιμοποιούμε το θεώρημα μονότονης σύγκλισης. Το (iv) έπεται από το (iii) εφόσον γράψουμε $f = f^+ - f^-$.

Έχουμε δει ήδη μια εφαρμογή της Τυπικής Μηχανής στη απόδειξη της Πρότασης 4.17 και θα δούμε ακόμα τρεις στις αποδείξεις της Πρότασης 6.9, του Θεωρήματος 10.8 καθώς και στην Άσκηση 6.4.

Παρατήρηση 6.4. Η Πρόταση 6.2 ισχύει και όταν το \mathbf{P} είναι μέτρο, όχι απαραίτητα πιθανότητας. Τότε το μέτρο \mathbf{P}^X ονομάζεται μέτρο εικόνα του \mathbf{P} μέσω της απεικόνισης X και δεν είναι απαραίτητα μέτρο πιθανότητας. Γράφουμε τότε την (6.1) ως

$$\int h(X) d\mathbf{P} = \int h d\mathbf{P}^X. \quad (6.4)$$

Το ολοκλήρωμα στο αριστερό μέλος είναι πάνω στο σύνολο Ω ενώ στο δεξί πάνω στο E .

Παρατήρηση 6.5. Ισοκατανεμημένες τυχαίες μεταβλητές. Αν (E, \mathcal{E}) είναι μετρήσιμος χώρος, $(\Omega_1, \mathcal{F}_1, \mathbf{P}_1), (\Omega_2, \mathcal{F}_2, \mathbf{P}_2)$ χώροι πιθανότητας, και

$$X : \Omega_1 \rightarrow E, \quad Y : \Omega_2 \rightarrow E$$

είναι δύο τυχαίες μεταβλητές με την ίδια κατανομή (τα πεδία ορισμού τους, Ω_1, Ω_2 , ενδέχεται να είναι διαφορετικά), τότε για οποιοδήποτε σύνολο $A \in \mathcal{E}$ έχουμε $\mathbf{P}_1(X \in A) = \mathbf{P}_2(Y \in A)$. Αυτό γιατί η πρώτη πιθανότητα ισούται με $\mathbf{P}_1^X(A)$, ενώ η δεύτερη με $\mathbf{P}_2^Y(A)$ και $\mathbf{P}_1^X = \mathbf{P}_2^Y$. Και όμοια, για οποιαδήποτε μετρήσιμη συνάρτηση $h : E \rightarrow \mathbb{R}$ ώστε η $\mathbf{E}_{\mathbf{P}_1}(h(X))$ να ορίζεται, ισχύει $\mathbf{E}_{\mathbf{P}_1}(h(X)) = \mathbf{E}_{\mathbf{P}_2}(h(Y))$. Γιατί και οι δύο μέσες τιμές μπορούν να εκφραστούν (λόγω της παραπάνω πρότασης) μέσω των κατανομών $\mathbf{P}_1^X, \mathbf{P}_2^Y$ οι οποίες ταυτίζονται. Για παράδειγμα, αν $E = \mathbb{R}$, θα ισχύει $\mathbf{E}_{\mathbf{P}_1} X^k = \mathbf{E}_{\mathbf{P}_2} Y^k$ για κάθε $k \in \mathbb{N}$ εφόσον οι ποσότητες αυτές ορίζονται.

Χοντρικά, σε οποιονδήποτε υπολογισμό εμπλέκεται η X μπορούμε να την αντικαταστήσουμε με την Y . Η δικαιολόγηση γίνεται με χρήση της πιο πάνω πρότασης.

Έτσι, σχεδόν για όλα τα προβλήματα πιθανοτήτων, αυτό που μας ενδιαφέρει σε μια τυχαία μεταβλητή είναι μόνο η κατανομή της ενώ ο χώρος πιθανότητας στον οποίο ορίζεται είναι εντελώς αδιάφορος.

Στην ειδική περίπτωση δύο τυχαίων μεταβλητών $X, Y : \Omega \rightarrow E$ (δηλαδή ορισμένων στον ίδιο χώρο πιθανότητας) με $\mathbf{P}(X = Y) = 1$, οι X, Y είναι ισοκατανεμημένες γιατί, για οποιοδήποτε μετρησιμο υποσύνολο A του E , οι τυχαίες μεταβλητές $\mathbf{1}_A(X), \mathbf{1}_A(Y)$ είναι ίσες με πιθανότητα 1. Άρα με βάση την Πρόταση 5.14(ii),

$$\mathbf{P}(X \in A) = \mathbf{E}\{\mathbf{1}_A(X)\} = \mathbf{E}\{\mathbf{1}_A(Y)\} = \mathbf{P}(Y \in A).$$

Ορολογία: Δύο τυχαίες μεταβλητές X, Y που παίρνουν τιμές σε κοινό μετρήσιμο χώρο (E, \mathcal{E}) λέγονται **ισόνομες**, ή και **ισοκατανεμημένες**, αν έχουν την ίδια κατανομή. Δηλαδή τα μέτρα $\mathbf{P}_1^X, \mathbf{P}_2^Y$ στον (E, \mathcal{E}) ταυτίζονται. Γράφουμε $X \stackrel{d}{=} Y$ [το d από το distribution (κατανομή)].

6.2 Κατανομές στο \mathbb{R} με πυκνότητα

Η Πρόταση 6.2 μας ενδιαφέρει κυρίως στην περίπτωση που $E = \mathbb{R}$ και η κατανομή της X προκύπτει από πυκνότητα.

Ορισμός 6.6. Έστω \mathbf{P} μέτρο πιθανότητας στον $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, λ το μέτρο Lebesgue (Παράγραφος 2.2), και $f : \mathbb{R} \rightarrow [0, \infty]$ Lebesgue-μετρήσιμη συνάρτηση. Η f λέγεται **πυκνότητα** του \mathbf{P} αν

$$\mathbf{P}(A) = \int_A f(x) d\lambda(x) \quad \text{για κάθε } A \in \mathcal{B}(\mathbb{R}). \quad (6.5)$$

Βέβαια $\lambda(\{f = \infty\}) = 0$ λόγω της Πρότασης 5.12(iii) αφού $\int_{\mathbb{R}} f d\lambda = 1 < \infty$.

Η πυκνότητα ενός μέτρου (αν αυτό έχει) δεν είναι μοναδική. Γιατί αν ένα μέτρο \mathbf{P} έχει πυκνότητα f , τότε αλλάζοντας την f σε ένα σύνολο που έχει μέτρο Lebesgue μηδέν, παίρνουμε μια νέα συνάρτηση \tilde{f} , η οποία είναι και αυτή πυκνότητα του \mathbf{P} . Αυτό έπεται από τον ορισμό της πυκνότητας και την Πρόταση 5.12(ii).

Αν ένα μέτρο όπως στο ορισμό έχει πυκνότητα f , τότε μπορούμε να υποθέσουμε ότι αυτή είναι Borel-μετρήσιμη. Συγκεκριμένα, υπάρχει $g : \mathbb{R} \rightarrow [0, \infty]$ Borel-μετρήσιμη ώστε $\lambda(\{f \neq g\}) = 0$ [δες Πρόταση 6.14 (ii) στο Κουμουλλής Γ. και Νεγρεπόντης Σ. (1991)], και άρα η g είναι επίσης πυκνότητα του \mathbf{P} .

Ένα μέτρο με πυκνότητα δεν έχει μοναδική πυκνότητα με την αυστηρή έννοια του όρου αλλά ουσιαστικά, με την οπτική της θεωρίας μέτρου, η πυκνότητα είναι μοναδική. Δηλαδή ισχύει το εξής.

Πρόταση 6.7. Αν δύο συναρτήσεις f_1, f_2 είναι πυκνότητες για το ίδιο μέτρο πιθανότητας \mathbf{P} στο \mathbb{R} , τότε $\lambda(\{f_1 \neq f_2\}) = 0$.

Απόδειξη. Έστω $g_1, g_2 : \mathbb{R} \rightarrow [0, \infty]$ Borel-μετρήσιμες συναρτήσεις ώστε $\lambda(\{f_1 \neq g_1\}) = \lambda(\{f_2 \neq g_2\}) = 0$. Θεωρούμε το σύνολο $A := \{g_1 > g_2\}$ το οποίο είναι Borel. Επειδή $\mathbf{P}(A) = \int_A g_1 d\lambda = \int_A g_2 d\lambda$, έχουμε

$$0 = \int_A (g_1 - g_2) d\lambda = \int (g_1 - g_2) \mathbf{1}_A d\lambda.$$

Όμως $(g_1 - g_2) \mathbf{1}_A \geq 0$ και έτσι η Πρόταση 5.12(i) δίνει ότι $\lambda(\{(g_1 - g_2) \mathbf{1}_A \neq 0\}) = 0$. Όμως $\{(g_1 - g_2) \mathbf{1}_A \neq 0\} = A$. Επομένως $\lambda(\{g_1 > g_2\}) = 0$. Αντιστρέφοντας τους ρόλους των g_1, g_2 , παίρνουμε $\lambda(\{g_1 < g_2\}) = 0$ και έτσι $\lambda(g_1 \neq g_2) = 0$. Τέλος, επειδή $\{f_1 \neq f_2\} \subset \{f_1 \neq g_1\} \cup \{f_2 \neq g_2\} \cup \{g_1 \neq g_2\}$, έπεται το συμπέρασμα. ■

Συνδυάζοντας αυτή την πρόταση με την Πρόταση 5.12(ii), συμπεραίνουμε ότι για υπολογισμούς ολοκληρωμάτων ως προς το μέτρο λ που εμπλέκουν μια πυκνότητα του \mathbf{P} , οποιαδήποτε άλλη πυκνότητα του \mathbf{P} δίνει το ίδιο αποτέλεσμα και επομένως θεωρούμε την πυκνότητα ουσιαστικά μοναδική.

Όπως παρατηρήσαμε πιο πάνω, η σχέση (6.5) για $A = \mathbb{R}$ δίνει ότι $\int_{\mathbb{R}} f d\lambda = 1$. Τώρα, αν έχουμε μια $f : \mathbb{R} \rightarrow [0, \infty]$ που είναι Lebesgue-μετρήσιμη με $\int_{\mathbb{R}} f d\lambda = 1$, τότε είναι ευκολο να δούμε ότι η (6.5) ορίζει ένα μέτρο πιθανότητας στο \mathbb{R} (δες Παράδειγμα 5.33). Άρα πυκνότητες κατανομών στο \mathbb{R} είναι ακριβώς οι μη αρνητικές Lebesgue-μετρήσιμες συναρτήσεις στο \mathbb{R} με ολοκλήρωμα 1 ως προς το μέτρο Lebesgue.

Ορισμός 6.8. Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας, $X : \Omega \rightarrow \mathbb{R}$ τυχαία μεταβλητή, και $f : \mathbb{R} \rightarrow [0, \infty]$ Lebesgue-μετρήσιμη συνάρτηση. Λέμε ότι η f είναι μια **πυκνότητα** της τυχαίας μεταβλητής X αν είναι πυκνότητα της κατανομής, \mathbf{P}^X , της X .

Επιστρέφουμε στην ειδική περίπτωση της Πρότασης 6.2 όπου $E = \mathbb{R}$ και η X έχει πυκνότητα.

Πρόταση 6.9. Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας και $X : \Omega \rightarrow \mathbb{R}$ τυχαία μεταβλητή με πυκνότητα $f : \mathbb{R} \rightarrow [0, \infty]$. Αν η $h : \mathbb{R} \rightarrow [-\infty, \infty]$ είναι Borel-μετρήσιμη, τότε

$$\mathbf{E}_{\mathbf{P}}\{h(X)\} = \int h(x)f(x) \, d\lambda(x) \quad (6.6)$$

όποτε ένα από τα δύο μέλη της ισότητας ορίζεται (δηλαδή τότε ορίζεται και το άλλο και είναι ίσα).

Το αριστερό μέλος της (6.6), από την Πρόταση 6.2, ισούται με $\mathbf{E}_{\mathbf{P}^X}(h)$ και αυτό θα δείξουμε στην απόδειξη ότι ισούται με το δεξί μέλος. Έτσι ο υπολογισμός της μέσης τιμής $\mathbf{E}_{\mathbf{P}}\{h(X)\}$ στο χώρο $(\Omega, \mathcal{F}, \mathbf{P})$ ανάγεται αρχικά σε υπολογισμό στο χώρο $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbf{P}^X)$ και τελικά σε ένα συνηθισμένο ολοκλήρωμα μιας μεταβλητής. Το δεξί μέλος της (6.6) είναι ο τρόπος με τον οποίο υπολογίζαμε τη μέση τιμή $\mathbf{E}\{h(X)\}$ για τυχαίες μεταβλητές X με πυκνότητα στις στοιχειώδεις πιθανότητες.

Απόδειξη. Θα δείξουμε ότι το δεξί μέλος της (6.6) ισούται επίσης με $\mathbf{E}_{\mathbf{P}^X}(h)$. Χρησιμοποιούμε την τυπική μηχανή (Παρατήρηση 6.3).

Αν $h = \mathbf{1}_A$ με $A \in \mathcal{B}(\mathbb{R})$, τότε

$$\begin{aligned} \int h \, d\mathbf{P}^X &= \int \mathbf{1}_A \, d\mathbf{P}^X = \mathbf{P}^X(A) = \mathbf{P}(X \in A) \\ &= \int_A f(x) \, d\lambda(x) = \int \mathbf{1}_A(x)f(x) \, d\lambda(x) = \int h(x)f(x) \, d\lambda(x). \end{aligned}$$

Η τρίτη ισότητα είναι ο ορισμός της κατανομής \mathbf{P}^X . Η τέταρτη είναι ο ορισμός της πυκνότητας.

Αν $h \geq 0$ απλή μετρήσιμη, τότε λόγω γραμμικότητας, από τα προηγούμενα προκύπτει το ζητούμενο.

Αν $h \geq 0$ μετρήσιμη, τότε από την Πρόταση 4.12 υπάρχει αύξουσα ακολουθία $(h_n)_{n \in \mathbb{N}}$ μη αρνητικών, απλών, μετρήσιμων συναρτήσεων με $\lim_{n \rightarrow \infty} h_n = h$. Άρα, σε συνδυασμό με το θεώρημα μονότονης σύγκλισης, έχουμε

$$\int h \, d\mathbf{P}^X = \lim_{n \rightarrow \infty} \int h_n \, d\mathbf{P}^X = \lim_{n \rightarrow \infty} \int h_n(x)f(x) \, d\lambda(x) = \int h(x)f(x) \, d\lambda(x).$$

Τέλος, αν η h είναι μετρήσιμη ώστε ένα από τα δύο μέλη της (6.6) να ορίζεται, από τα προηγούμενα έχουμε

$$\begin{aligned} \int h^+ \, d\mathbf{P}^X &= \int h^+(x)f(x) \, d\lambda(x), \\ \int h^- \, d\mathbf{P}^X &= \int h^-(x)f(x) \, d\lambda(x), \end{aligned}$$

και επομένως, αφού $h = h^+ - h^-$, έχουμε

$$\int h \, d\mathbf{P}^X = \int h^+ \, d\mathbf{P}^X - \int h^- \, d\mathbf{P}^X = \int h^+(x)f(x) \, d\lambda(x) - \int h^-(x)f(x) \, d\lambda(x) = \int h(x)f(x) \, d\lambda(x).$$

Από την υπόθεση, δεν υπάρχει κάποιο από τα μέλη των ισοτήτων στην τελευταία γραμμή που να δίνει τη μορφή $\infty - \infty$. ■

Παράδειγμα 6.10. Έστω $X : \Omega \rightarrow \mathbb{R}$ τυχαία μεταβλητή με πυκνότητα $f(x) = \frac{1}{\pi(1+x^2)}$ για κάθε $x \in \mathbb{R}$ (δηλαδή ακολουθεί την κατανομή Cauchy). Η $\mathbf{E}(X)$ δεν ορίζεται. Πράγματι, από την Πρόταση 6.9, για τη συνάρτηση $h : \mathbb{R} \rightarrow \mathbb{R}$ με $h(a) = a^+$ για κάθε $a \in \mathbb{R}$, έχουμε

$$\begin{aligned} \mathbf{E}(X^+) &= \mathbf{E}(h(X)) = \int h(x)f(x) dx = \int x^+ f(x) dx \\ &= \int_0^{\infty} x \frac{1}{\pi(1+x^2)} dx \geq \frac{1}{\pi} \int_1^{\infty} \frac{x}{1+x^2} dx \\ &\geq \frac{1}{\pi} \int_1^{\infty} \frac{1}{2x} dx = \frac{1}{2\pi} \int_1^{\infty} \frac{1}{x} dx = \infty. \end{aligned}$$

Όμοια, $\mathbf{E}(X^-) = \infty$.

Ορισμός 6.11. Έστω X τυχαία μεταβλητή με τιμές στο \mathbb{R} . Λέμε ότι η X έχει συμμετρική κατανομή αν $X \stackrel{d}{=} -X$, δηλαδή, για κάθε $A \in \mathcal{B}(\mathbb{R})$ ισχύει $\mathbf{P}(X \in A) = \mathbf{P}(-X \in A)$.

Παράδειγμα 6.12. Έστω Y τυχαία μεταβλητή με τιμές στο \mathbb{R} και πυκνότητα f άρτια συνάρτηση. Τότε η Y έχει συμμετρική κατανομή. Πράγματι, για κάθε $A \in \mathcal{B}(\mathbb{R})$ έχουμε

$$\mathbf{P}(Y \in A) = \int_A f(x) d\lambda(x) = \int_{-A} f(-x) d\lambda(x) = \int_{-A} f(x) d\lambda(x) = \mathbf{P}(Y \in -A).$$

Η δεύτερη ισότητα είναι συνέπεια της (6.4) με $\Omega = E = \mathbb{R}$, $\mathbf{P} = \lambda$ (το μέτρο Lebesgue στο \mathbb{R}), $X(\omega) = -\omega$ και $h(x) = f(x) \cdot \mathbf{1}_{x \in A}$ (χρησιμοποιούμε και την $\lambda^X = \lambda$).

Παράδειγμα συμμετρικής τυχαίας μεταβλητής είναι μια $Y \sim N(0, \sigma^2)$. Όμως μια $Y \sim \text{Exp}(\lambda)$, $\lambda > 0$, δεν έχει συμμετρική κατανομή.

6.3 Διακριτές κατανομές

Διακριτή κατανομή σε ένα σύνολο E λέμε ένα μέτρο πιθανότητας \mathbf{P} στον μετρήσιμο χώρο $(E, \mathcal{P}(E))$ για το οποίο υπάρχει ένα αριθμήσιμο σύνολο $S \subset E$ ώστε $\mathbf{P}(S) = 1$. Για $A \subset S$, γράφουμε $A = \cup_{x \in A} \{x\}$, και επειδή το A είναι αριθμήσιμο, ισχύει

$$\mathbf{P}(A) = \sum_{x \in A} \mathbf{P}(\{x\}). \quad (6.7)$$

Το \mathbf{P} δίνει μάζα $\mathbf{P}(\{x\})$ σε κάθε σημείο $x \in S$ και μάζα μηδέν στο $E \setminus S$. Έτσι, για κάθε $A \subset E$ έχουμε $\mathbf{P}(A) = \mathbf{P}(A \cap S)$ και

$$\mathbf{P}(A) = \sum_{x \in A} \mathbf{P}(\{x\}). \quad (6.8)$$

Το άθροισμα στο δεξί μέλος έχει αριθμήσιμο πλήθος μη μηδενικών όρων. Αντιστοιχούν στα σημεία του $A \cap S$.

Η ολοκλήρωση ως προς το \mathbf{P} είναι απλή υπόθεση. Έχουμε το εξής.

Πρόταση 6.13. Έστω \mathbf{P} διακριτή κατανομή στο E . Τότε

$$\int h(x) d\mathbf{P}(x) = \sum_{x \in E} h(x) \mathbf{P}(\{x\}) \quad (6.9)$$

για κάθε $h : E \rightarrow [-\infty, \infty]$ για την οποία ένα από τα δύο μέλη ορίζεται. Δηλαδή για κάθε $h : E \rightarrow [-\infty, \infty]$ ή και τα δύο μέλη ορίζονται και ισούνται ή και τα δύο δεν ορίζονται.

Δεν έχουμε κάποια απαίτηση μετρησιμότητας από την h αφού η σ -άλγεβρα στο πεδίο ορισμού της h είναι η $\mathcal{P}(E)$.

Απόδειξη. Αν $h = \mathbf{1}_A$ με $A \subset E$, τότε η (6.9) είναι η (6.8). Αν $h \geq 0$ απλή, όπως στο δεξί μέλος της (4.3), τότε

$$\begin{aligned} \int h(x) d\mathbf{P}(x) &= \sum_{i=1}^n a_i \mathbf{P}(A_i) = \sum_{i=1}^n a_i \sum_{x \in A_i} \mathbf{P}(\{x\}) = \sum_{i=1}^n \sum_{x \in A_i} h(x) \mathbf{P}(\{x\}) \\ &= \sum_{x \in E} h(x) \mathbf{P}(\{x\}). \end{aligned}$$

Έστω τώρα $h \geq 0$ μετρήσιμη. Αν το S είναι πεπερασμένο, η απόδειξη τελείωσε γιατί η h είναι απλή και αναγόμαστε στην προηγούμενη περίπτωση. Αν το S είναι άπειρο αριθμήσιμο, θεωρούμε $(s_n)_{n \geq 1}$ μια αρίθμηση του. Για $h : E \rightarrow [0, \infty]$ και κάθε $n \geq 1$ θέτουμε $h_n = \sum_{k=1}^n h(s_k) \mathbf{1}_{\{s_k\}}$. Η h_n είναι απλή με $0 \leq h_n \leq h$ και η ακολουθία $(h_n)_{n \geq 1}$ είναι αύξουσα και συγκλίνει στην h . Άρα, από το θεώρημα μονότονης σύγκλισης,

$$\int h(x) d\mathbf{P}(x) = \lim_{n \rightarrow \infty} \int h_n(x) d\mathbf{P}(x) = \lim_{n \rightarrow \infty} \sum_{k=1}^n h(s_k) \mathbf{P}(\{s_k\}) = \sum_{x \in S} h(x) \mathbf{P}(\{x\}) = \sum_{x \in E} h(x) \mathbf{P}(\{x\}).$$

Η περίπτωση που η h παίρνει τιμές στο $[-\infty, \infty]$ αντιμετωπίζεται όπως και στην απόδειξη της Πρότασης 6.9. ■

Διακριτή τυχαία μεταβλητή στο E λέμε μια τυχαία μεταβλητή $X : \Omega \rightarrow E$ της οποίας η εικόνα, $S := X(\Omega)$, είναι αριθμήσιμο σύνολο. Η κατανομή της, \mathbf{P}^X , είναι μια διακριτή κατανομή αφού $\mathbf{P}^X(S) = 1$. Ισχύει

$$\mathbf{E}\{h(X)\} = \sum_{x \in E} h(x) \mathbf{P}(X = x) \quad (6.10)$$

για όλες τις $h : \mathbb{R} \rightarrow [-\infty, \infty]$ για τις οποίες κάποιο από τα δύο μέλη της ισότητας έχει νόημα. Αυτό προκύπτει από την (6.1) και την (6.9) εφαρμοσμένη στο μέτρο \mathbf{P}^X , το οποίο έχει $\mathbf{P}^X(\{x\}) = \mathbf{P}(X = x)$ για κάθε $x \in E$. Ονομάζουμε τη συνάρτηση $f : E \rightarrow [0, 1]$ με $f(x) := \mathbf{P}(X = x)$ **συνάρτηση πιθανότητας** της X .

Ο τύπος (6.10) είναι γνωστός από τις στοιχειωδεις πιθανότητες.

Ασκήσεις

6.1 Έστω $(\Omega_1, \mathcal{F}_1, \mathbf{P}_1), (\Omega_2, \mathcal{F}_2, \mathbf{P}_2)$ χώροι πιθανότητας, $(E, \mathcal{E}), (G, \mathcal{G})$ μετρήσιμοι χώροι, $X : \Omega_1 \rightarrow E, Y : \Omega_2 \rightarrow E$ τυχαίες μεταβλητές, και $f : E \rightarrow G$ μετρήσιμη συνάρτηση. Αν $X \stackrel{d}{=} Y$, να δειχθεί ότι $f(X) \stackrel{d}{=} f(Y)$.

6.2 (Τυχαία μεταβλητή με δεδομένη κατανομή) Έστω (E, \mathcal{E}) μετρήσιμος χώρος και ν μέτρο πιθανότητας σε αυτόν. Να κατασκευαστεί χώρος πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$ και τυχαία μεταβλητή $X : \Omega \rightarrow E$ έτσι ώστε η κατανομή της X να είναι ν .

[Υπόδειξη: Παίρνουμε $(\Omega, \mathcal{F}, \mathbf{P}) = (E, \mathcal{E}, \nu)$.]

6.3 Να δειχθεί ότι οι μέσες τιμές στην ισότητα (6.1) ισούνται επίσης με

$$\int_{\mathbb{R}} t d\mathbf{P}^{h(X)}(t).$$

Κατανομή τυχαίας μεταβλητής και ολοκλήρωση

6.4 Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας και X, \mathbf{Q} όπως στο Παράδειγμα 5.33. Αν $Y : \Omega \rightarrow \mathbb{R}$ τυχαία μεταβλητή, να δείξετε ότι

$$\int Y d\mathbf{Q} = \mathbf{E}_{\mathbf{P}}(YX)$$

όποτε κάποια από τις δύο ποσότητες ορίζεται (για παράδειγμα, όταν $Y \geq 0$ ή $\mathbf{E}_{\mathbf{P}}(|Y|X) < \infty$). Η άσκηση αυτή γενικεύει την Πρόταση 6.9.

6.5 Έστω X τυχαία μεταβλητή με κατανομή την κανονική $N(0, 1)$. Για κάθε $x > 0$ ναδειχθεί ότι

$$\frac{x}{x^2 + 1} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \leq \mathbf{P}(X > x) \leq \frac{1}{x} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \quad (6.11)$$

Δηλαδή, για μεγάλο x , έχουμε $\mathbf{P}(X > x) \sim cx^{-1}e^{-x^2/2}$ με $c = 1/\sqrt{2\pi}$.

6.6 Έστω X συνεχής τυχαία μεταβλητή με τιμές στο \mathbb{R} και πυκνότητα άρτια συνάρτηση. Ναδειχθεί ότι:

(α) $\mathbf{E}\{h(X)\} = 2 \mathbf{E}\{h(X)\mathbf{1}_{X>0}\}$ για κάθε $h : \mathbb{R} \rightarrow \mathbb{R}$ μετρήσιμη άρτια συνάρτηση με $\mathbf{E}|h(X)| < \infty$.

(β) $\mathbf{E}\{h(X)\} = 0$ για κάθε $h : \mathbb{R} \rightarrow \mathbb{R}$ μετρήσιμη περιττή συνάρτηση με $\mathbf{E}|h(X)| < \infty$.

6.7 Έστω X συνεχής τυχαία μεταβλητή με τιμές στο \mathbb{R} και πυκνότητα f . Ναδειχθεί ότι $\mathbf{P}(f(X) = 0) = 0$.

6.8 Έστω X τυχαία μεταβλητή με $\mathbf{E}(X^2) < \infty$. Ναδειχθεί ότι η μέση τιμή $\mathbf{E}X$ είναι το μοναδικό σημείο που ελαχιστοποιεί τη συνάρτηση $g(a) = \mathbf{E}\{(X - a)^2\}$.

7

Κατανομές στο \mathbb{R}

7.1 Η συνάρτηση κατανομής μέτρου

Στην παράγραφο αυτή θα μιλήσουμε για μέτρα πιθανότητας στον $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Αυτά τα μέτρα τα λέμε και *κατανομές* στο \mathbb{R} .

Ορισμός 7.1. Έστω \mathbf{P} μέτρο πιθανότητας στον $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. **Συνάρτηση κατανομής** του \mathbf{P} λέγεται η συνάρτηση $F : \mathbb{R} \rightarrow [0, 1]$ με

$$F(x) = \mathbf{P}((-\infty, x]) \quad \text{για κάθε } x \in \mathbb{R}.$$

Δηλαδή, η $F(x)$ μετράει τη μάζα που δίνει το μέτρο \mathbf{P} στην ημιευθεία $(-\infty, x]$.

Παράδειγμα 7.2. Έστω $x_0 \in \mathbb{R}$ και δ_{x_0} το μέτρο Dirac στο x_0 με πεδίο ορισμού τη σ -άλγεβρα $\mathcal{P}(\mathbb{R})$. Η συνάρτηση κατανομής του δ_{x_0} είναι η

$$F(x) = \begin{cases} 0 & \text{αν } x < x_0, \\ 1 & \text{αν } x \geq x_0. \end{cases}$$

Παρακάτω, χρησιμοποιούνται οι εξής συμβολισμοί: Για $h : \mathbb{R} \rightarrow \mathbb{R}$ και $x_0 \in \mathbb{R}$,

$$\begin{aligned} h(x_0-) &:= \lim_{x \rightarrow x_0^-} h(x), & h(x_0+) &:= \lim_{x \rightarrow x_0^+} h(x), \\ h(-\infty) &:= \lim_{x \rightarrow -\infty} h(x), & h(\infty) &:= \lim_{x \rightarrow \infty} h(x), \end{aligned}$$

όποτε τα όρια υπάρχουν.

Παρατήρηση 7.3 (Όρια μονότονων συναρτήσεων). Αν η $h : \mathbb{R} \rightarrow \mathbb{R}$ είναι αύξουσα, τότε υπάρχουν τα όρια $h(-\infty), h(\infty), h(x_0-), h(x_0+)$ για κάθε $x_0 \in \mathbb{R}$. Μάλιστα έχουν τις εξής τιμές

$$\begin{aligned} h(-\infty) &= \inf\{h(x) : x \in \mathbb{R}\}, \\ h(\infty) &= \sup\{h(x) : x \in \mathbb{R}\}, \\ h(x_0-) &= \sup\{h(x) : x < x_0\}, \\ h(x_0+) &= \inf\{h(x) : x > x_0\}. \end{aligned}$$

Ας αποδείξουμε την ύπαρξη του $h(x_0-)$ και την έκφραση για την τιμή του. Έστω $\ell := \sup\{h(x) : x < x_0\}$, το υποψήφιο όριο. Για $M < \ell$, από τον ορισμό του \sup , υπάρχει $x_1 < x_0$ ώστε $h(x_1) > M$. Τότε για κάθε $x \in (x_1, x_0)$ θα έχουμε $M < h(x) \leq \ell$, ετσι η $h(x_0-) = \ell$ αποδείκτηκε.

Παρόμοια είναι η απόδειξη και για τα άλλα όρια. Για τα $h(-\infty), h(\infty)$ ενδεχόμενες τιμές είναι τα $-\infty, \infty$ αντίστοιχα.

Πρόταση 7.4. Έστω \mathbf{P} μέτρο πιθανότητας στον $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ και F η συνάρτηση κατανομής του \mathbf{P} . Τότε:

(i) Η F είναι αύξουσα συνάρτηση.

(ii) Η F είναι δεξιά συνεχής.

(iii) $F(-\infty) = 0$ και $F(\infty) = 1$.

Απόδειξη. (i) Έστω $x \leq y$. Τότε επειδή $(-\infty, x] \subset (-\infty, y]$ το (ii) της Πρότασης 2.5 δίνει ότι

$$F(x) = \mathbf{P}((-\infty, x]) \leq \mathbf{P}((-\infty, y]) = F(y).$$

(ii) Έστω $x_0 \in \mathbb{R}$. Επειδή η F είναι αύξουσα, το $F(x_0+)$ υπάρχει, και έχουμε

$$\begin{aligned} F(x_0+) &= \lim_{n \rightarrow \infty} F\left(x_0 + \frac{1}{n}\right) = \lim_{n \rightarrow \infty} \mathbf{P}\left(\left(-\infty, x_0 + \frac{1}{n}\right]\right) = \mathbf{P}\left(\bigcap_{n \in \mathbb{N}^+} \left(-\infty, x_0 + \frac{1}{n}\right]\right) \\ &= \mathbf{P}((-\infty, x_0]) = F(x_0). \end{aligned}$$

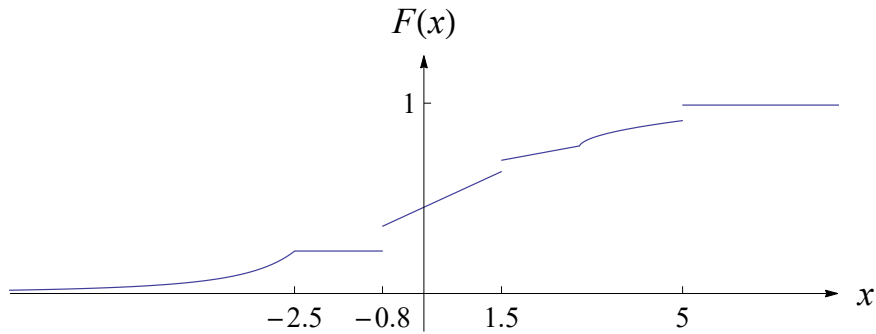
(iii) Επειδή η F είναι αύξουσα, τα όρια υπάρχουν και

$$\lim_{x \rightarrow -\infty} F(x) = \lim_{n \rightarrow \infty} F(-n) = \lim_{n \rightarrow \infty} \mathbf{P}((-\infty, -n]) = \mathbf{P}\left(\bigcap_{n \in \mathbb{N}} (-\infty, -n]\right) = \mathbf{P}(\emptyset) = 0,$$

ενώ

$$\lim_{x \rightarrow \infty} F(x) = \lim_{n \rightarrow \infty} F(n) = \lim_{n \rightarrow \infty} \mathbf{P}((-\infty, n]) = \mathbf{P}\left(\bigcup_{n \in \mathbb{N}} (-\infty, n]\right) = \mathbf{P}(\mathbb{R}) = 1.$$

■



Σχήμα 7.1: Μια συνάρτηση κατανομής

Σημειώνουμε ότι η F ικανοποιεί τη σχέση

$$\mathbf{P}((x, y]) = F(y) - F(x) \tag{7.1}$$

για κάθε x, y πραγματικούς αριθμούς με $x \leq y$ γιατί

$$\mathbf{P}((x, y]) = \mathbf{P}((-\infty, y] \setminus (-\infty, x]) = \mathbf{P}((-\infty, y]) - \mathbf{P}((-\infty, x]) = F(y) - F(x).$$

Το ίδιο ισχύει για κάθε $-\infty \leq x \leq y \leq \infty$ με τις συμβάσεις $(x, \infty] = (x, \infty)$ και $(x, x] = \emptyset$. Χρήσιμη επίσης είναι η σχέση (Άσκηση 7.1)

$$\mathbf{P}(\{x\}) = F(x) - F(x-) \tag{7.2}$$

για κάθε $x \in \mathbb{R}$. Επομένως η F έχει άλμα (είναι ασυνεχής) ακριβώς στα σημεία στα οποία το \mathbf{P} δίνει θετικό μέτρο. Υπενθυμίζουμε ότι μια αύξουσα συνάρτηση σε ένα διάστημα είναι ασυνεχής σε ένα σημείο μόνο αν τα πλευρικά της όρια σε αυτό (τα οποία υπάρχουν) είναι διαφορετικά.

Έστω ότι μας δίνεται στο Σχήμα 7.1 η συνάρτηση κατανομής ενός μέτρου πιθανότητας \mathbf{P} . Παρατηρούμε ότι ικανοποιεί τις ιδιότητες (i)-(iii) της Πρότασης 7.4. Τι συμπεράσματα μπορούμε να βγάλουμε από το σχήμα για το \mathbf{P} ;

(1) Η F έχει άλμα ακριβώς στα σημεία $-0.8, 1.5, 5$. Επομένως $\mathbf{P}(\{x\}) > 0$ όταν το x πάρει μια από αυτές τις τρεις τιμές ενώ διαφορετικά, $\mathbf{P}(\{x\}) = 0$.

(2) Η F είναι σταθερή στο $[-2.5, -0.8)$ και συνεχής στο -2.5 . Άρα

$$\mathbf{P}([-2.5, -0.8)) = F(-0.8-) - F(-2.5-) = 0$$

[δες Άσκηση 7.1]. Βέβαια $\mathbf{P}([-2.5, -0.8]) > 0$ λόγω του άλματος στο -0.8 .

(3) $\mathbf{P}((x, y]) > 0$ για κάθε διάστημα $(x, y]$ με $x < y$ που περιέχεται στο $(-\infty, -2.5]$ ή στο $[-0.8, 5]$ γιατί σε αυτά τα δύο διαστήματα η F είναι γνησίως αύξουσα.

(4) $F(x) > 0$ για κάθε $x \in \mathbb{R}$, οπότε $\mathbf{P}((-\infty, x]) > 0$ για κάθε x , όσο μικρό και να είναι αυτό.

(5) $F(5) = 1$. Άρα $\mathbf{P}((5, \infty)) = 0$.

Παράδειγμα 7.5 (Διακριτό μέτρο πιθανότητας στο \mathbb{R}). Έστω $S \subset \mathbb{R}$ αριθμήσιμο και $(a_t)_{t \in S}$ θετικοί αριθμοί έτσι ώστε $\sum_{t \in S} a_t = 1$ (για παράδειγμα, $S = \mathbb{N}$, $a_k = 1/2^{k+1}$ για κάθε $k \in \mathbb{N}$). Ορίζουμε

$$\mathbf{P}(A) = \sum_{t \in A} a_t.$$

Το \mathbf{P} είναι μέτρο πιθανότητας στην $\mathcal{P}(\mathbb{R})$. Για τη συνάρτηση κατανομής, F , του \mathbf{P} έχουμε

$$F(x) = \sum_{t \leq x} a_t$$

και από την Άσκηση 7.1(α),

$$F(x) - F(x-) = \mathbf{P}(\{x\}) = \begin{cases} a_x & \text{αν } x \in S, \\ 0 & \text{αν } x \in \mathbb{R} \setminus S. \end{cases}$$

Δηλαδή, η F είναι ασυνεχής ακριβώς στα σημεία του S .

Αν το \mathbf{P} έχει πυκνότητα, τότε η συνάρτηση κατανομής του είναι συνεχής γιατί για κάθε $x \in \mathbb{R}$ έχουμε

$$F(x) - F(x-) = \mathbf{P}(\{x\}) = \int_{\{x\}} f(t) d\lambda(t) = 0$$

αφού $\lambda(\{x\}) = 0$. Συμπεραίνουμε από αυτό ότι υπάρχουν μέτρα πιθανότητας στο \mathbb{R} χωρίς πυκνότητα αφού σαφώς υπάρχουν μέτρα με συνάρτηση κατανομής ασυνεχή (δες το προηγούμενο παράδειγμα).

Η σημαντικότητα της συνάρτησης κατανομής πηγάζει από το επόμενο θεώρημα, το οποίο λέει ότι η F κωδικοποιεί πλήρως ένα μέτρο. Κρατώντας την F αντί του \mathbf{P} , καμία πληροφορία δεν έχει χαθεί.

Θεώρημα 7.6. Έστω \mathbf{P}, \mathbf{Q} μέτρα πιθανότητας στον $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ με την ίδια συνάρτηση κατανομής. Τότε $\mathbf{P} = \mathbf{Q}$.

Η απόδειξη του δόθηκε αμέσως μετά το Πρόσμμα 3.7.

Παρατήρηση 7.7. Αν ξέρουμε τη συνάρτηση κατανομής F ενός μέτρου πιθανότητας \mathbf{P} , τότε γνωρίζουμε τις τιμές του σε σύνολα που προκύπτουν από διαστήματα της μορφής $(x, y]$ με συνήθεις συνολοθεωρητικές πράξεις χρησιμοποιώντας την (7.1) και τις σχέσεις της Άσκησης 7.1.

Ποιες συναρτήσεις $F : \mathbb{R} \rightarrow \mathbb{R}$ προκύπτουν ως συναρτήσεις κατανομής μέτρων πιθανότητας στον $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$; Η απάντηση δίνεται στο επόμενο θεώρημα.

Θεώρημα 7.8. Μια συνάρτηση $F : \mathbb{R} \rightarrow \mathbb{R}$ είναι συνάρτηση κατανομής ενός μέτρου πιθανότητας \mathbf{P} στον $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ αν και μόνο αν ισχύουν τα (i)-(iii) της Πρότασης 7.4.

Απόδειξη. Τη συνεπαγωγή \Rightarrow την είδαμε στην Πρόταση 7.4. Τη συνεπαγωγή \Leftarrow θα την αποδείξουμε στην Παράγραφο 7.3. ■

Το θεώρημα αυτό μας επιτρέπει να δείξουμε την ύπαρξη μέτρων ορίζοντας μόνο τη συνάρτηση κατανομής τους. Δεν είναι απαραίτητο να ορίσουμε την τιμή τους σε κάθε Borel υποσύνολο του \mathbb{R} . Παραδείγματα δίνονται στις ασκήσεις.

7.2 Η συνάρτηση κατανομής τυχαίας μεταβλητής

Αν $X : \Omega \rightarrow \mathbb{R}$ είναι τυχαία μεταβλητή, τότε **συνάρτηση κατανομής** της X λέμε τη συνάρτηση $F_X : \mathbb{R} \rightarrow [0, 1]$ με

$$F_X(x) := \mathbf{P}(X \leq x)$$

για κάθε $x \in \mathbb{R}$. Αν \mathbf{P}^X είναι η κατανομή της X , τότε $F_X(x) = \mathbf{P}^X((-\infty, x])$. Επομένως, με βάση τον Ορισμό 7.1, η F_X είναι η συνάρτηση κατανομής του μέτρου \mathbf{P}^X . Έτσι, τα αποτελέσματα της προηγούμενης Παραγράφου συνεπάγονται ότι η F_X έχει τις εξής ιδιότητες.

- (i) Η F_X είναι αύξουσα.
- (ii) Η F_X είναι δεξιά συνεχής.
- (iii) $F_X(-\infty) = 0$ και $F_X(\infty) = 1$.
- (iv) $\mathbf{P}(a < X \leq b) = F_X(b) - F_X(a)$ για κάθε $a < b, a, b \in \mathbb{R}$.
- (v) $\mathbf{P}(X = a) = F_X(a) - F_X(a-)$ για κάθε $a \in \mathbb{R}$.

Στην επόμενη παράγραφο θα δούμε ότι αν μια συνάρτηση F έχει τις ιδιότητες (i)-(iii) πιο πάνω, τότε υπάρχει τυχαία μεταβλητή X που να έχει την F ως συνάρτηση κατανομής.

Σημειώνουμε επίσης ότι το Θεώρημα 7.6 έχει την εξής συνέπεια.

Πρόσμμα 7.9. Αν X, Y είναι τυχαίες μεταβλητές με τιμές στο \mathbb{R} ώστε $F_X = F_Y$, τότε $X \stackrel{d}{=} Y$.

7.3 Η συνάρτηση ποσοστημορίων*

Έστω $F : \mathbb{R} \rightarrow \mathbb{R}$ που ικανοποιεί τις (i)-(iii) της Πρότασης 7.4. Ορίζουμε τη συνάρτηση $Q : (0, 1) \rightarrow \mathbb{R}$ με

$$Q(t) := \inf\{x \in \mathbb{R} : F(x) \geq t\}. \quad (7.3)$$

Η Q είναι καλά ορισμένη, γιατί $F(-\infty) = 0$ και $F(\infty) = 1$, και λέγεται **συνάρτηση ποσοστημορίων** της F . Όταν η F είναι γνησίως αύξουσα και συνεχής, τότε $Q = F^{-1}$ (η αντίστροφη της F). Η Q είναι αύξουσα, και για $t \in (0, 1), z \in \mathbb{R}$ έχουμε ότι

$$Q(t) \leq z \Leftrightarrow t \leq F(z). \quad (7.4)$$

[Η κατεύθυνση \Leftarrow είναι προφανής. Για την κατεύθυνση \Rightarrow , από τις ιδιότητες του \inf , υπάρχει φθίνουσα ακολουθία $(x_n)_{n \geq 1}$ με όριο το $Q(t)$ ώστε $F(x_n) \geq t$ για κάθε $n \geq 1$. Έτσι, επειδή η F είναι αύξουσα και δεξιά συνεχής, έχουμε $F(z) \geq F(Q(t)) = \lim_{n \rightarrow \infty} F(x_n) \geq t$.]

Η Q ως αύξουσα είναι μετρήσιμη (Άσκηση 4.5). Χρησιμοποιώντας την θα δείξουμε το δύσκολο κομμάτι του Θεωρήματος 7.8.

Απόδειξη του Θεωρήματος 7.8: Έστω F που ικανοποιεί τα (i), (ii) και (iii) της Πρότασης 7.4. Θέτουμε $\mathbf{P} : \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ με

$$\mathbf{P}(A) := \lambda(Q^{-1}(A))$$

όπου λ είναι το μέτρο Lebesgue στο \mathbb{R} . Ο ορισμός είναι καλός γιατί η Q είναι μετρήσιμη. Θεωρώντας τον χώρο πιθανότητας $((0, 1), \mathcal{B}((0, 1)), \lambda)$, έχουμε ότι το μέτρο \mathbf{P} είναι η κατανομή της τυχαίας μεταβλητής $Q : (0, 1) \rightarrow \mathbb{R}$. Επομένως, είναι μέτρο πιθανότητας στον $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

ΙΣΧΥΡΙΣΜΟΣ: Το \mathbf{P} έχει συνάρτηση κατανομής F .

Πράγματι, για $x \in \mathbb{R}$, χρησιμοποιώντας την (7.4), έχουμε

$$Q^{-1}((-\infty, x]) = \{t \in (0, 1) : Q(t) \leq x\} = \{t \in (0, 1) : t \leq F(x)\} = (0, F(x)],$$

οπότε $\mathbf{P}((-\infty, x]) = \lambda((0, F(x)]) = F(x)$. ■

Παράφραση της ιδέας της προηγούμενης απόδειξης, με όρους τυχαίων μεταβλητών, είναι η εξής πρόταση.

Πρόταση 7.10. Έστω $F : \mathbb{R} \rightarrow \mathbb{R}$ που ικανοποιεί τις (i)-(iii) της Πρότασης 7.4, Q όπως στην (7.3), και U τυχαία μεταβλητή με κατανομή την ομοιόμορφη στο $(0, 1)$. Τότε η $X := Q(U)$ έχει συνάρτηση κατανομής F .

Απόδειξη. Έχουμε

$$\mathbf{P}(Q(U) \leq x) = \mathbf{P}(U \leq F(x)) = F(x).$$

Στην πρώτη ισότητα χρησιμοποιήσαμε την (7.4). Στη δεύτερη, ότι $F(x) \in [0, 1]$ και ότι η U ακολουθεί την ομοιόμορφη κατανομή στο $(0, 1)$. ■

Η προηγούμενη πρόταση έχει τις εξής δύο εφαρμογές:

(α) Δίνει μια μέθοδο προσομοίωσης τυχαίων μεταβλητών. Αν έχουμε έναν μηχανισμό που παράγει ομοιόμορφες τυχαίες μεταβλητές, τότε μπορούμε να παραγάγουμε και οποιαδήποτε άλλη τυχαία μεταβλητή για την οποία είναι εύκολο να υπολογίσουμε τη συνάρτηση ποσοστημορίων της, Q .

Παράδειγμα 7.11. (i) Η συνάρτηση κατανομής της κατανομής $\exp(2)$ είναι $F(x) = (1 - e^{-2x})\mathbf{1}_{x \geq 0}$ και

$$Q(t) = -\frac{1}{2} \log(1 - t)$$

για κάθε $t \in (0, 1)$. Άρα με βάση την προηγούμενη πρόταση, έχουμε ότι, αν η U έχει την ομοιόμορφη κατανομή στο $(0, 1)$, τότε η $-(1/2) \log(1 - U)$ ακολουθεί την $\exp(2)$.

- (ii) Αν $p \in (0, 1)$, η κατανομή $\text{Bernoulli}(p)$ έχει συνάρτηση κατανομής $F(x) = (1 - p) \cdot \mathbf{1}_{[0,1)}(x) + \mathbf{1}_{[1,\infty)}(x)$ και συνάρτηση ποσοστημορίων $Q(t) = \mathbf{1}_{(1-p,1)}(t) = \mathbf{1}_{1-t < p}$ για κάθε $t \in (0, 1)$. Έτσι, αν $U \sim U(0, 1)$, τότε η $\mathbf{1}_{1-U < p} \sim \text{Bernoulli}(p)$. Επειδή και η $1 - U \sim U(0, 1)$, έχουμε επίσης $\mathbf{1}_{U < p} \sim \text{Bernoulli}(p)$, πράγμα το οποίο βλέπουμε και άμεσα.

(β) Κατασκευάζει σε κοινό χώρο πιθανότητας τυχαίες μεταβλητές με δεδομένη κατανομή. Δηλαδή αν έχουμε $(X_i)_{i \in I}$ τυχαίες μεταβλητές (όχι απαραίτητα σε κοινό χώρο πιθανότητας) με τιμές στο \mathbb{R} που έχουν αντίστοιχα συναρτήσεις ποσοστημορίων $(Q_i)_{i \in I}$, τότε θεωρούμε μια τυχαία μεταβλητή $U \sim U(0, 1)$ σε κάποιον χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$ και θέτουμε $Y_i := Q_i(U)$ για κάθε $i \in I$. Τότε $Y_i \stackrel{d}{=} X_i$ για κάθε i και όλες οι $(Y_i)_{i \in I}$ είναι ορισμένες στον $(\Omega, \mathcal{F}, \mathbf{P})$.

Παράδειγμα 7.12. Έστω $U \sim U(0, 1)$. Για κάθε $p \in (0, 1)$, ορίζουμε την τυχαία μεταβλητή $X_p := \mathbf{1}_{U < p}$. Οι $(X_p)_{p \in (0,1)}$ είναι ορισμένες στον ίδιο χώρο πιθανότητας και $X_p \sim \text{Bernoulli}(p)$ για κάθε p . Όλες οι X_p παράγονται από μια τυχαία μεταβλητή, την U .

7.4 Είδη κατανομών στο \mathbb{R}

Ανάμεσα σε όλες τις κατανομές (δηλαδή μέτρα πιθανότητας) στο \mathbb{R} ξεχωρίζουμε τα εξής δύο είδη:

- (i) Διακριτές.
- (ii) Συνεχείς.

Ορίσαμε τις διακριτές σε γενικότερο πλαίσιο στην προηγούμενη παράγραφο. Έπειτα, λέμε μια κατανομή ν **συνεχή** αν η συνάρτηση κατανομής της, $F(x) := \nu((-\infty, x])$, είναι συνεχής συνάρτηση. Αυτό ισοδυναμεί με $\nu(\{x\}) = 0$ για κάθε $x \in \mathbb{R}$ [Άσκηση 7.1(α)], δηλαδή η κατανομή ν δεν έχει άτομα.

Ανάμεσα στις συνεχείς κατανομές ξεχωρίζουμε τα εξής δύο είδη:

- (i) Απολύτως συνεχείς.
- (ii) Ιδιάζουσες.

Αυτός ο διαχωρισμός προκύπτει από τη σχέση που έχει μια κατανομή με το μέτρο Lebesgue λ του \mathbb{R} . Έτσι, μια κατανομή ν στο \mathbb{R} λέγεται **απολύτως συνεχής** αν έχει πυκνότητα, ενώ λέγεται **ιδιάζουσα** αν υπάρχει $A \in \mathcal{B}(\mathbb{R})$ με $\lambda(A) = 0$ και $\nu(A^c) = 0$. Δηλαδή το ν κατανέμει όλη του την μάζα σε ένα σύνολο (το A) που έχει μέτρο Lebesgue μηδέν. Επειδή το ν είναι μέτρο πιθανότητας μπορούμε να γράψουμε ισοδύναμα ότι $\lambda(A) = 0$ και $\nu(A) = 1$.

Μια ισοδύναμη περιγραφή για τις ιδιάζουσες είναι αυτές των οποίων η συνάρτηση κατανομής, F , είναι συνεχής με παράγωγο $F'(x) = 0$, λ -σχεδόν παντού στο \mathbb{R} . Κατασκευή τέτοιας κατανομής γίνεται στην Άσκηση 7.8.

Αν ν_1, ν_2 είναι κατανομές που η πρώτη είναι διακριτή και η δεύτερη συνεχής, τότε η $(\nu_1 + \nu_2)/2$ είναι κατανομή που δεν είναι ούτε διακριτή ούτε συνεχής. Κάθε κατανομή έχει μια τέτοια ανάλυση σε κυρτό συνδυασμό κατανομών από τα «καθαρά» είδη.

Θεώρημα 7.13. Αν μ είναι κατανομή στο \mathbb{R} , τότε γράφεται ως κυρτός συνδυασμός τριών κατανομών μ_d, μ_{ac}, μ_s με τη μ_d διακριτή, τη μ_{ac} απολύτως συνεχή, και τη μ_s ιδιάζουσα.

Κατανομές στο \mathbb{R}

Δηλαδή υπάρχουν $\lambda_1, \lambda_2, \lambda_3 \in [0, 1]$ με $\lambda_1 + \lambda_2 + \lambda_3 = 1$ και μ_d, μ_{ac}, μ_s όπως στην εκφώνηση ώστε

$$\mu = \lambda_1 \mu_d + \lambda_2 \mu_{ac} + \lambda_3 \mu_s.$$

Τα $\lambda_1, \lambda_2, \lambda_3$ καθορίζονται μοναδικά και κάθε κατανομή στο δεξί μέλος με μη μηδενικό συντελεστή καθορίζεται μοναδικά. Για την απόδειξη του θεωρήματος απαιτούνται προχωρημένες γνώσεις της θεωρίας παραγωγίσιων συναρτήσεων [δες Κεφάλαιο 14 στο Κουμουλλής Γ. και Νεγρεπόντης Σ. (1991)]. Μπορεί να τη δει κανείς, για παράδειγμα, στο Παπαδάτος Ν. (2006), Παράγραφος 4.3. Κατανομές που στην ανάλυσή τους έχουν τουλάχιστον δύο από τα $\lambda_1, \lambda_2, \lambda_3$ διαφορετικά από το 0 τις λέμε **μεικτές**.

Αντίστοιχα, σε μια τυχαία μεταβλητή $X : \Omega \rightarrow \mathbb{R}$ αποδίδουμε τον χαρακτηρισμό διακριτή, συνεχής, απολύτως συνεχής, ιδιάζουσα, ή μεικτή, ανάλογα με το τι είναι η κατανομή της. Ωστόσο, εδώ η χρήση του όρου «συνεχής» είναι καταχρηστική γιατί αποκαλώντας τη X συνεχή (ως τ.μ.) δεν σημαίνει ότι είναι συνεχής συνάρτηση. Μάλιστα ενδέχεται να μην έχει νόημα να εξετάσουμε αν η X είναι συνεχής ως συνάρτηση γιατί ο Ω δεν έχει απαραίτητα δομή μετρικού χώρου.

Ασκήσεις

7.1 Έστω F συνάρτηση κατανομής ενός μέτρου \mathbf{P} στον $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Για $x, y \in \mathbb{R}$ με $x < y$ να δειχθεί ότι

- (α) $\mathbf{P}(\{x\}) = F(x) - F(x-)$.
- (β) $\mathbf{P}([x, y]) = F(y) - F(x-)$.
- (γ) $\mathbf{P}([x, y)) = F(y-) - F(x-)$.
- (δ) $\mathbf{P}((x, y)) = F(y-) - F(x)$.

7.2 Έστω \mathbf{P} μέτρο πιθανότητας στο \mathbb{R} και F η συνάρτηση κατανομής του \mathbf{P} . Να δείξετε ότι η F μπορεί να έχει το πολύ αριθμήσιμο πλήθος αλμάτων.

7.3 Έστω \mathbf{P}_1 κατανομή στο \mathbb{R} , με πυκνότητα $f(x) = e^{-x} \mathbf{1}_{x>0}$, και \mathbf{P}_2 κατανομή στο \mathbb{R} που δίνει μάζα $\frac{1}{2}$ σε καθένα από τα $\{-2\}, \{3\}$. Για $\lambda \in (0, 1)$ και, θεωρώντας τον κυρτό συνδυασμό $\mathbf{P} = \lambda \mathbf{P}_1 + (1 - \lambda) \mathbf{P}_2$ των \mathbf{P}_1 και \mathbf{P}_2 , να υπολογιστούν

- (α) η $\mathbf{P}((0, 4))$,
- (β) η συνάρτηση κατανομής του \mathbf{P} .

7.4 Έστω X τυχαία μεταβλητή με συνάρτηση κατανομής F για την οποία γνωρίζουμε ότι

$$F(x) = \begin{cases} \frac{1}{3} \sqrt{x} & \text{αν } x \in (0, 1), \\ \frac{1}{2} x & \text{αν } x \in (1, 2). \end{cases}$$

- (α) Να προσδιοριστούν οι τιμές $F(-1), F(1), F(3)$.
- (β) Να προσδιοριστούν οι πιθανότητες $\mathbf{P}(X < 1), \mathbf{P}(X = 1), \mathbf{P}(0.25 < X < 1.5)$.

7.5 Έστω $F : \mathbb{R} \rightarrow [0, 1]$ συνάρτηση που ικανοποιεί τις ιδιότητες (i)-(iii) της Πρότασης 7.4 και η οποία επιπλέον είναι συνεχής. Να δειχθεί ότι η F είναι ομοιόμορφα συνεχής.

7.6 Έστω X_1, X_2 τυχαίες μεταβλητές σε χώρους πιθανότητας $(\Omega_1, \mathcal{F}_1, \mathbf{P}_1), (\Omega_2, \mathcal{F}_2, \mathbf{P}_2)$ αντίστοιχα. Λέμε ότι η X_1 είναι **στοχαστικά μικρότερη** από την X_2 , και γράφουμε $X_1 \leq_{st} X_2$, αν $\mathbf{P}_1(X_1 > x) \leq \mathbf{P}_2(X_2 > x)$ για κάθε $x \in \mathbb{R}$ (ισοδύναμα, $F_1(x) \geq F_2(x)$ για κάθε $x \in \mathbb{R}$). Δηλαδή το να είναι η X_1 μεγάλη είναι λιγότερο πιθανό από το να είναι η X_2 . Να δειχθεί ότι τα εξής είναι ισοδύναμα:

- (α) $X_1 \leq_{st} X_2$.
- (β) Υπάρχει χώρος πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$ και τυχαίες μεταβλητές Y_1, Y_2 ορισμένες σε αυτόν ώστε $Y_1 \stackrel{d}{=} X_1, Y_2 \stackrel{d}{=} X_2$ και $\mathbf{P}(Y_1 \leq Y_2) = 1$.

[Υπόδειξη: Χρησιμοποιούμε τη συνάρτηση ποσοστημορίων.]

Κατανομές στο \mathbb{R}

(γ) $\mathbf{E}_1\{h(X_1)\} \leq \mathbf{E}_2\{h(X_2)\}$ για κάθε $h : \mathbb{R} \rightarrow \mathbb{R}$ αύξουσα ώστε οι μέσες τιμές να ορίζονται.

7.7 [Μεροληψία με βάση το μέγεθος] Έστω X τυχαία μεταβλητή με τιμές στο $[0, \infty)$ και με $\mu := \mathbf{E}(X) \in (0, \infty)$. Να δειχθεί ότι:

(α) Υπάρχει τυχαία μεταβλητή X^* ώστε

$$\mathbf{P}(X^* \leq x) = \frac{\mathbf{E}(X \mathbf{1}_{X \leq x})}{\mu} \quad (7.5)$$

για κάθε $x \in \mathbb{R}$. Η X^* λέγεται η **μεροληπτική με βάση το μέγεθος εκδοχή της X** .

(β) Για κάθε $h : \mathbb{R} \rightarrow \mathbb{R}$ Borel-μετρήσιμη με τιμές στο $[0, \infty]$ ή με $\mathbf{E}|h(X^*)| < \infty$ ισχύει

$$\mathbf{E}\{h(X^*)\} = \frac{1}{\mu} \mathbf{E}\{Xh(X)\}. \quad (7.6)$$

(γ) Αν η X παίρνει τιμές σε ένα αριθμήσιμο $S \subset \mathbb{R}$ και έχει συνάρτηση πιθανότητας f_X , τότε η X^* έχει συνάρτηση πιθανότητας

$$f_{X^*}(x) = \frac{1}{\mu} x f_X(x) \quad (7.7)$$

για κάθε $x \in \mathbb{R}$.

(δ) Αν η X έχει πυκνότητα f_X , τότε η X^* έχει πυκνότητα

$$f_{X^*}(x) = \frac{1}{\mu} x f_X(x) \quad (7.8)$$

για κάθε $x \in \mathbb{R}$.

7.8 (Η κατανομή Cantor) Η συνάρτηση Cantor $\phi : [0, 1] \rightarrow [0, 1]$ ορίζεται ως εξής. Αν έχουν οριστεί οι τιμές της ϕ στα άκρα ενός διαστήματος $I = [a, b] \subset [0, 1]$, τότε το να εφαρμόσουμε τη διαδικασία T στη ϕ στο I σημαίνει να χωρίσουμε το I σε τρία διαδοχικά διαστήματα μήκους $|I|/3$ το καθένα και στην κλειστότητα του μεσαίου διαστήματος να ορίσουμε τη ϕ να παίρνει την τιμή $(\phi(a) + \phi(b))/2$. Ορίζουμε λοιπόν $\phi(0) = 0, \phi(1) = 1$ και εφαρμόζουμε στη ϕ τη διαδικασία T στο $[0, 1]$. Έτσι, οι τιμές της ϕ έχουν οριστεί στο $[1/3, 2/3]$. Έπειτα εφαρμόζουμε τη διαδικασία T στα διαστήματα $[0, 1/3], [2/3, 1]$. Στα άκρα τους οι τιμές της ϕ είναι καθορισμένες από τα προηγούμενα βήματα. Συνεχίζουμε με τον ίδιο τρόπο επ' άπειρον. Με αυτό τον τρόπο καθορίζονται οι τιμές της ϕ στο συμπλήρωμα του συνόλου Cantor C (και σε ένα αριθμήσιμο πλήθος σημείων ακόμα). Για τα υπόλοιπα σημεία του $[0, 1]$ ορίζουμε $\phi(x) := \sup\{\phi(t) : t < x, t \in [0, 1] \setminus C\}$.

(α) Να δειχθεί ότι η ϕ είναι αύξουσα και το σύνολο τιμών της είναι πυκνό στο $[0, 1]$.

(β) Να δειχθεί ότι η ϕ είναι συνεχής και $\phi'(x) = 0$ για κάθε $x \in [0, 1] \setminus C$.

[Υπόδειξη: Η συνέχεια έπεται από το (α).]

(γ) Ορίζουμε $F(x) = \phi(x)$ για $x \in [0, 1]$, $F(x) = 0$ για $x < 0$, και $F(x) = 1$ για $x > 1$. Η F έχει τις ιδιότητες συνάρτησης κατανομής, οπότε υπάρχει μέτρο πιθανότητας μ στο \mathbb{R} που την έχει ως συνάρτηση κατανομής. Να δειχθεί ότι $\mu([0, 1] \setminus C) = 0$.

(δ) Να δειχθεί ότι το μ δεν έχει άτομα (δηλαδή σημεία x με $\mu(\{x\}) > 0$, άρα είναι συνεχής κατανομή) αλλά δεν προκύπτει από πυκνότητα.

8

Τρόποι σύγκλισης ακολουθιών τυχαίων μεταβλητών

Για μια ακολουθία $(X_n)_{n \in \mathbb{N}}$ τυχαίων μεταβλητών σε χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$ και με τιμές στο \mathbb{R} γνωρίζουμε από τον απειροστικό λογισμό δύο βασικούς τρόπους σύγκλισης, την κατά σημείο και την ομοιόμορφη. Σε αυτό το κεφάλαιο θα ορίσουμε κάποιες νέες, φυσιολογικές έννοιες σύγκλισης ως προς τις οποίες μια ακολουθία είναι ευκολότερο να συγκλίνει και επιπλέον είναι χρήσιμες στις εφαρμογές των Πιθανοτήτων στη Στατιστική και αλλού.

Ορισμός 8.1. Έστω $(X_n)_{n \in \mathbb{N}}$ ακολουθία τυχαίων μεταβλητών όπως πιο πάνω.

- (i) Λέμε ότι η $(X_n)_{n \in \mathbb{N}}$ συγκλίνει σε μια τυχαία μεταβλητή X **με πιθανότητα 1** ή **σχεδόν βέβαια**, και γράφουμε $X_n \xrightarrow{\sigma.β.} X$, αν

$$\mathbf{P}\left(\lim_{n \rightarrow \infty} X_n = X\right) = 1,$$

δηλαδή

$$\mathbf{P}\left(\{\omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\}\right) = 1. \quad (8.1)$$

- (ii) Για $p \geq 1$ και $X_n, X \in \mathcal{L}^p$ για κάθε $n \in \mathbb{N}$, λέμε ότι η $(X_n)_{n \in \mathbb{N}}$ συγκλίνει στη X **στον \mathcal{L}^p** , και γράφουμε $X_n \xrightarrow{\mathcal{L}^p} X$, αν

$$\lim_{n \rightarrow \infty} \mathbf{E}(|X_n - X|^p) = 0. \quad (8.2)$$

- (iii) Λέμε ότι η $(X_n)_{n \in \mathbb{N}}$ συγκλίνει στη X **κατά πιθανότητα**, και γράφουμε $X_n \xrightarrow{\mathbf{P}} X$, αν

$$\lim_{n \rightarrow \infty} \mathbf{P}(|X_n - X| > \epsilon) = 0 \quad (8.3)$$

για κάθε $\epsilon > 0$.

Και για τα τρία είδη σύγκλισης του προηγούμενου ορισμού είναι απαραίτητο οι τυχαίες μεταβλητές της ακολουθίας $(X_n)_{n \in \mathbb{N}}$ και η οριακή τυχαία μεταβλητή X να ορίζονται στον ίδιο χώρο πιθανότητας. Παντού σε αυτό το κεφάλαιο κάνουμε αυτή την υπόθεση χωρίς να το αναφέρουμε. Επίσης θα υποθέτουμε ότι παίρνουν τιμές στο \mathbb{R} .

Αν θέλουμε¹ ο πιο πάνω ορισμός να καλύπτει και τυχαίες μεταβλητές με τιμές στο $[-\infty, \infty]$, τότε αρκεί να αντικαταστήσουμε την (8.3) με την

$$\lim_{n \rightarrow \infty} \mathbf{P}(d(X_n, X) > \epsilon) = 0 \quad (8.4)$$

όπου d είναι η μετρική στον $[-\infty, \infty]$ που ορίσαμε στη σχέση (1.2). Για τη σύγκλιση στον \mathcal{L}^p , η υπόθεση ότι οι $X, X_n \in \mathcal{L}^p$ δίνει ότι με πιθανότητα 1 παίρνουν τιμές στο \mathbb{R} , οπότε η ποσότητα $X_n - X$ ορίζεται και είναι πραγματικός αριθμός με πιθανότητα 1.

¹Είναι φυσιολογικό να χρειαζόμαστε αυτή την επέκταση του ορισμού γιατί το σημειακό όριο μιας ακολουθίας τυχαίων μεταβλητών με τιμές στο \mathbb{R} μπορεί σε κάποια σημεία να παίρνει κάποια από τις τιμές $-\infty, \infty$.

Στην έκφραση $X_n \xrightarrow{\sigma.\beta.} X$ πρέπει κανείς να συμπληρώσει «καθώς $n \rightarrow \infty$ ». Θα παραλείψουμε αυτή την επιπλέον φράση. Όμως είναι απαραίτητη όταν υπάρχει πρόβλημα σύγκλισης, π.χ., όταν έχουμε μια ακολουθία με δύο παραμέτρους $(X_{m,n})_{m,n \in \mathbb{N}}$. Το ίδιο ισχύει και για τα υπόλοιπα είδη σύγκλισης.

Αν $X_n \rightarrow X$ κατά σημείο, τότε βέβαια $X_n \xrightarrow{\sigma.\beta.} X$ αφού $\{\lim_{n \rightarrow \infty} X_n = X\} = \Omega$. Στο επόμενο θεώρημα, βλέπουμε επίσης ότι η σχεδόν βέβαιη σύγκλιση και η σύγκλιση στον \mathcal{L}^p είναι ισχυρότερες από τη σύγκλιση κατά πιθανότητα.

Θεώρημα 8.2. Έστω $X, (X_n)_{n \in \mathbb{N}}$ τυχαίες μεταβλητές και $p \geq 1$.

(i) Αν $X_n \xrightarrow{\mathcal{L}^p} X$, τότε $X_n \xrightarrow{\mathbf{P}} X$.

(ii) Αν $X_n \xrightarrow{\sigma.\beta.} X$, τότε $X_n \xrightarrow{\mathbf{P}} X$.

Απόδειξη. Έστω $\epsilon > 0$.

(i) Για κάθε $n \in \mathbb{N}$, χρησιμοποιώντας την ανισότητα Μαρκον, έχουμε

$$\mathbf{P}(|X_n - X| > \epsilon) = \mathbf{P}(|X_n - X|^p > \epsilon^p) \leq \frac{1}{\epsilon^p} \mathbf{E}(|X_n - X|^p).$$

Για $n \rightarrow \infty$ προκύπτει το ζητούμενο.

(ii) Για κάθε $n \in \mathbb{N}$,

$$\mathbf{P}(|X_n - X| > \epsilon) = \mathbf{E}(\mathbf{1}_{|X_n - X| > \epsilon}) = \mathbf{E}(g_n),$$

όπου $g_n = \mathbf{1}_{|X_n - X| > \epsilon}$. Η $X_n \xrightarrow{\sigma.\beta.} X$ δίνει $g_n \xrightarrow{\sigma.\beta.} 0$. Επίσης $|g_n| \leq 1$, άρα από το θεώρημα φραγμένης σύγκλισης,

$$\lim_{n \rightarrow \infty} \mathbf{E}(g_n) = \mathbf{E}(\lim_{n \rightarrow \infty} g_n) = 0. \quad \blacksquare$$

Παράδειγμα 8.3. (i) Έστω $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}([0, 1])$ και \mathbf{P} το μέτρο Lebesgue στο $[0, 1]$. Θεωρούμε την τυχαία μεταβλητή $X = 0$ και για $n \in \mathbb{N}^+$ την τυχαία μεταβλητή

$$X_n(\omega) = \begin{cases} n & \text{αν } \omega \in (0, \frac{1}{n}), \\ 0 & \text{αν } \omega \in [0, 1] \setminus (0, \frac{1}{n}). \end{cases}$$

Τότε $X_n(\omega) \rightarrow X(\omega)$ για κάθε $\omega \in \Omega$, άρα $X_n \xrightarrow{\sigma.\beta.} X$. Επιπλέον, $X_n \xrightarrow{\mathbf{P}} X$. Όμως $X_n \not\xrightarrow{\mathcal{L}^p} X$ για $p \geq 1$. Πράγματι,

$$\mathbf{E}(|X_n - X|^p) = \mathbf{E}(|X_n|^p) = n^p \mathbf{P}(X_n = n) + 0^p \mathbf{P}(X_n = 0) = n^p \frac{1}{n} = n^{p-1} \rightarrow 0.$$

καθώς $n \rightarrow \infty$ γιατί $p \geq 1$.

(ii) Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ όπως στο (i). Για κάθε $k \in \mathbb{N}^+$ χωρίζουμε το $[0, 1]$ σε 2^k διαδοχικά κλειστά διαστήματα ίδιου μήκους, $J_1^k, J_2^k, \dots, J_{2^k}^k$. Αριθμούμε τα $\{J_r^k : k \geq 1, 1 \leq r \leq 2^k\}$ σε μια ακολουθία $(I_n)_{n \in \mathbb{N}}$ ώστε το $J_{r_1}^{\mu}$ εμφανίζεται νωρίτερα από το $J_{r_2}^{\nu}$ αν $\mu < \nu$ ή αν $\mu = \nu$ και $r_1 < r_2$. Θεωρούμε την τυχαία μεταβλητή $X_n = \mathbf{1}_{I_n}$, για κάθε $n \in \mathbb{N}$. Τότε, για $p \geq 1$,

$$\mathbf{E}(|X_n - 0|^p) = \mathbf{E}(|X_n|^p) = \mathbf{P}(I_n) \rightarrow 0 \text{ για } n \rightarrow \infty.$$

Τρόποι σύγκλισης ακολουθιών τυχαίων μεταβλητών

[Μάλιστα η τελευταία πιθανότητα ανήκει στο διάστημα $[1/(n+2), 2/(n+2)]$]. Άρα $X_n \xrightarrow{\mathcal{L}^p} 0$. Συνεπώς, $X_n \xrightarrow{\mathbf{P}} 0$. Όμως η X_n δεν συγκλίνει σε κάποια τυχαία μεταβλητή σχεδόν βέβαια. Πράγματι,

$$\underline{\lim}_{n \rightarrow \infty} X_n(\omega) = 0 < 1 = \overline{\lim}_{n \rightarrow \infty} X_n(\omega)$$

για κάθε $\omega \in \Omega$ αφού κάθε $\omega \in \Omega$ ανήκει σε άπειρα από τα I_n αλλά και σε άπειρα από τα $[0, 1] \setminus I_n$. Άρα $\mathbf{P}(\lim_{n \rightarrow \infty} X_n \text{ υπάρχει}) = 0$.

Θεώρημα 8.4. Έστω $(X_n)_{n \in \mathbb{N}}$ ακολουθία τυχαίων μεταβλητών και X τυχαία μεταβλητή έτσι ώστε $X_n \xrightarrow{\mathbf{P}} X$. Τότε υπάρχει υπακολουθία $(X_{n_k})_{k \in \mathbb{N}^+}$ έτσι ώστε $X_{n_k} \xrightarrow{\sigma.\beta.} X$.

Προσοχή. Η υπακολουθία δεν εξαρτάται από το $\omega \in \Omega$. Δηλαδή υπάρχει μία $(n_k)_{k \geq 1}$ γνησίως αύξουσα ακολουθία φυσικών (π.χ. η $n_k = k!$) ώστε σχεδόν για όλα τα $\omega \in \Omega$ να ισχύει $X_{n_k}(\omega) \rightarrow X(\omega)$.

Απόδειξη. Επιλέγουμε αναδρομικά μια γνησίως αύξουσα ακολουθία φυσικών, $(n_k)_{k \in \mathbb{N}^+}$, έτσι ώστε

$$\mathbf{P}\left(|X_{n_k} - X| > \frac{1}{k}\right) < \frac{1}{2^k}$$

για κάθε $k \in \mathbb{N}^+$. Αυτό είναι δυνατόν γιατί $X_n \xrightarrow{\mathbf{P}} X$. Έπειτα θεωρούμε τα σύνολα $A_k = \{|X_{n_k} - X| > 1/k\}$, $k \in \mathbb{N}^+$, τα οποία ικανοποιούν την

$$\sum_{k=1}^{\infty} \mathbf{P}(A_k) \leq \sum_{k=1}^{\infty} \frac{1}{2^k} < \infty.$$

Αυτό συνεπάγεται ότι $\mathbf{P}(\limsup_{k \geq 1} A_k) = 0$ γιατί

$$\mathbf{P}(\limsup_{k \geq 1} A_k) = \mathbf{P}(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k) = \lim_{n \rightarrow \infty} \mathbf{P}(\bigcup_{k=n}^{\infty} A_k) \leq \overline{\lim}_{n \rightarrow \infty} \sum_{k=n}^{\infty} \mathbf{P}(A_k) = 0.$$

Τότε το σύνολο $\Omega \setminus \limsup_{k \geq 1} A_k$ έχει πιθανότητα 1 και για κάθε $\omega \in \Omega \setminus \limsup_{k \geq 1} A_k$ υπάρχει $k(\omega) \in \mathbb{N}$ έτσι ώστε για κάθε $k \geq k(\omega)$ να ισχύει $\omega \notin A_k$, δηλαδή $|X_{n_k}(\omega) - X(\omega)| \leq 1/k$, συνεπώς $X_{n_k}(\omega) \rightarrow X(\omega)$. Άρα $X_{n_k} \xrightarrow{\sigma.\beta.} X$. ■

Θεώρημα 8.5. Έστω $p \geq 1$ και $(X_n)_{n \in \mathbb{N}}$, X όπως στο Θεώρημα 8.4 με την επιπλέον υπόθεση ότι υπάρχει $Y \in \mathcal{L}^p$ ώστε $|X_n| \leq Y$ για κάθε $n \in \mathbb{N}$. Τότε $X \in \mathcal{L}^p$ και $X_n \xrightarrow{\mathcal{L}^p} X$.

Απόδειξη. Από το Θεώρημα 8.4 υπάρχει υπακολουθία $(X_{n_k})_{k \in \mathbb{N}}$ τέτοια ώστε $X_{n_k} \xrightarrow{\sigma.\beta.} X$. Συνεπώς,

$$\mathbf{E}(|X|^p) = \mathbf{E}(\lim_{k \rightarrow \infty} |X_{n_k}|^p) \leq \underline{\lim}_{k \rightarrow \infty} \mathbf{E}(|X_{n_k}|^p) \leq \mathbf{E}(|Y|^p),$$

σύμφωνα με το λήμμα Fatou. Άρα $X \in \mathcal{L}^p$.

Έστω ότι $X_n \xrightarrow{\mathcal{L}^p} X$. Τότε υπάρχει $\epsilon > 0$ και υπακολουθία $(X_{\lambda_n})_{n \in \mathbb{N}}$ έτσι ώστε

$$\mathbf{E}(|X_{\lambda_n} - X|^p) \geq \epsilon \quad \text{για κάθε } n \in \mathbb{N}. \quad (8.5)$$

Εφόσον $X_{\lambda_n} \xrightarrow{\mathbf{P}} X$, από το Θεώρημα 8.4, υπάρχει υπακολουθία $(X_{\lambda_{n_k}})_{k \in \mathbb{N}}$ της $(X_{\lambda_n})_{n \in \mathbb{N}}$ έτσι ώστε $X_{\lambda_{n_k}} \xrightarrow{\sigma.\beta.} X$. Βεβαίως $|X| \leq Y$. Για κάθε $k \in \mathbb{N}$, έστω $Z_k = |X_{\lambda_{n_k}} - X|^p$. Η ακολουθία $(Z_k)_{k \geq 1}$ συγκλίνει στο 0 σχεδόν βεβαίως και επειδή ²

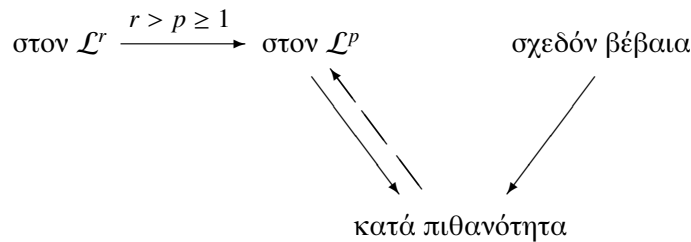
$$|Z_k| \leq 2^p(|X_{\lambda_{n_k}}|^p + |X|^p) \leq 2^p \cdot 2 \cdot Y^p,$$

κυριαρχείται από την $2^{p+1}Y^p$, που έχει πεπερασμένη μέση τιμή αφού $Y \in \mathcal{L}^p$. Από το θεώρημα κυριαρχημένης σύγκλισης,

$$\lim_{n \rightarrow \infty} \mathbf{E}(Z_n) = \mathbf{E}(\lim_{n \rightarrow \infty} Z_n) = 0,$$

το οποίο συγκρούεται με την (8.5). Συνεπώς, $X_n \xrightarrow{\mathcal{L}^p} X$. ■

Παρατηρήστε ότι λόγω του Παραδείγματος 8.3(i) η σχεδόν βέβαιη σύγκλιση δεν συνεπάγεται σύγκλιση στον \mathcal{L}^p (για $p \geq 1$). Χρειάζεται να υποθέσουμε κάτι επιπλέον για να πάρουμε αυτή τη σύγκλιση. Επειδή η σχεδόν βέβαιη σύγκλιση συνεπάγεται την κατά πιθανότητα, το προηγούμενο θεώρημα δίνει ότι, όταν $X_n \rightarrow X$ σχεδόν βέβαια και υπάρχει $Y \in \mathcal{L}^p$ με $|X_n| \leq Y$ για όλα τα $n \in \mathbb{N}$, τότε $X \in \mathcal{L}^p$ και $X_n \xrightarrow{\mathcal{L}^p} X$.



Σχήμα 8.1: Σχέσεις μεταξύ των διαφόρων τρόπων σύγκλισης. Το διακεκομμένο βέλος δηλώνει ότι η συνεπαγωγή ισχύει αν έχουμε κάποια επιπλέον υπόθεση, όπως η κυριαρχιση της ακολουθίας από στοιχείο του \mathcal{L}^p . Παραλείψαμε το βέλος $\mathcal{L}^r \rightarrow$ κατά πιθανότητα. Για όλες τις υπόλοιπες κατευθύνσεις που δεν υπάρχει συμπαγές βέλος υπάρχουν αντιπαραδείγματα.

Οι συνεχείς συναρτήσεις διατηρούν τη σχεδόν βέβαιη σύγκλιση και τη σύγκλιση κατά πιθανότητα. Πιο συγκεκριμένα, έχουμε το ακόλουθο αποτέλεσμα.

Πρόταση 8.6. Έστω $f : \mathbb{R} \rightarrow \mathbb{R}$ συνεχής συνάρτηση και $(X_n)_{n \in \mathbb{N}}$, X τυχαίες μεταβλητές.

(i) Αν $X_n \xrightarrow{\sigma.\beta.} X$, τότε $f(X_n) \xrightarrow{\sigma.\beta.} f(X)$.

(ii) Αν $X_n \xrightarrow{\mathbf{P}} X$, τότε $f(X_n) \xrightarrow{\mathbf{P}} f(X)$.

Απόδειξη. (i) Έστω $A = \{\omega \in \Omega : X_n(\omega) \rightarrow X(\omega)\}$. Τότε $\mathbf{P}(A) = 1$ και για $\omega \in A$ ισχύει ότι $f(X_n(\omega)) \rightarrow f(X(\omega))$ εφόσον η f είναι συνεχής. Άρα, αν $B = \{\omega \in \Omega : f(X_n(\omega)) \rightarrow f(X(\omega))\}$, έχουμε ότι $A \subset B$, συνεπώς $\mathbf{P}(B) = 1$, από το οποίο προκύπτει το ζητούμενο. Σαφώς $A, B \in \mathcal{F}$ (Άσκηση 4.4).

(ii) Έστω ότι $f(X_n) \not\xrightarrow{\mathbf{P}} f(X)$. Τότε υπάρχουν $\epsilon > 0$, $\delta > 0$, και γνήσια αύξουσα ακολουθία $(n_k)_{k \in \mathbb{N}}$ έτσι ώστε $\mathbf{P}(|f(X_{n_k}) - f(X)| > \epsilon) \geq \delta$ για κάθε $k \in \mathbb{N}$. Μπορούμε να υποθέσουμε ότι

²Χρησιμοποιούμε το ότι $|a + b|^p \leq (2 \max(|a|, |b|))^p \leq 2^p(|a|^p + |b|^p)$.

Τρόποι σύγκλισης ακολουθιών τυχαίων μεταβλητών

$\mathbf{P}(|f(X_n) - f(X)| > \epsilon) \geq \delta$ για κάθε $n \in \mathbb{N}$, διαφορετικά δουλεύουμε όμοια με την ακολουθία τυχαίων μεταβλητών $(Y_k)_{k \in \mathbb{N}}$, όπου $Y_k = X_{n_k}$.

Εφόσον $X_n \xrightarrow{\mathbf{P}} X$, από το Θεώρημα 8.4 υπάρχει υπακολουθία $(X_{\lambda_n})_{n \in \mathbb{N}}$ της $(X_n)_{n \in \mathbb{N}}$ έτσι ώστε $X_{\lambda_n} \xrightarrow{\sigma.\beta.} X$. Από το (i), $f(X_{\lambda_n}) \xrightarrow{\sigma.\beta.} f(X)$, άρα $f(X_{\lambda_n}) \xrightarrow{\mathbf{P}} f(X)$ το οποίο είναι άτοπο εφόσον $\mathbf{P}(|f(X_{\lambda_n}) - f(X)| > \epsilon) \geq \delta$ για κάθε $n \in \mathbb{N}$. Συνεπώς, $f(X_n) \xrightarrow{\mathbf{P}} f(X)$. ■

Για τη σύγκλιση στον \mathcal{L}^p το αντίστοιχο αποτέλεσμα δεν ισχύει αφού δεν είναι καν απαραίτητο η $f(Y)$ να ανήκει στον \mathcal{L}^p αν $Y \in \mathcal{L}^p$ και f συνεχής.

Ασκήσεις

Στις ασκήσεις αυτού του κεφαλαίου, όπου εμφανίζεται ακολουθία τυχαίων μεταβλητών, θεωρούμε ότι όλες τους ορίζονται στον ίδιο χώρο πιθανότητας.

8.1 Έστω $(X_n)_{n \geq 1}$ ακολουθία τυχαίων μεταβλητών με τιμές στο \mathbb{R} . Για $\epsilon > 0$ και $n \geq 1$ θέτουμε $A_n^\epsilon := \{|X_n| \geq \epsilon\}$. Να δείξετε ότι τα εξής είναι ισοδύναμα:

(α) $\mathbf{P}(\lim_{n \rightarrow \infty} X_n = 0) = 1$.

(β) $\mathbf{P}(\limsup_{n \geq 1} A_n^\epsilon) = 0$ για κάθε $\epsilon > 0$.

8.2 Για $X, Y : \Omega \rightarrow \mathbb{R}$ τυχαίες μεταβλητές ορίζουμε

$$d(X, Y) := \mathbf{E} \left\{ \frac{|X - Y|}{1 + |X - Y|} \right\} \in [0, 1].$$

(α) Για $X, Y, Z : \Omega \rightarrow \mathbb{R}$ τυχαίες μεταβλητές να δειχθεί ότι:

(i) $d(X, Y) = 0$ αν και μόνο αν $\mathbf{P}(X = Y) = 1$.

(ii) $d(X, Y) = d(Y, X)$.

(iii) $d(X, Z) \leq d(X, Y) + d(Y, Z)$.

(β) Έστω $X, (X_n)_{n \geq 1}$ τυχαίες μεταβλητές με τιμές στο \mathbb{R} . Να δειχθεί ότι $X_n \rightarrow X$ κατά πιθανότητα αν και μόνο αν $d(X_n, X) \rightarrow 0$.

8.3 Έστω $(X_n)_{n \geq 1}$ ακολουθία τυχαίων μεταβλητών με τιμές στο \mathbb{R} . Να δειχθεί ότι $X_n \rightarrow 0$ κατά πιθανότητα αν και μόνο αν $\lim_{n \rightarrow \infty} \mathbf{E}(|X_n| \wedge 1) = 0$.

8.4 Έστω $(X_n)_{n \geq 1}$ ακολουθία ισόνομων τυχαίων μεταβλητών με τιμές στο \mathbb{R} . Να δειχθεί ότι

$$\lim_{n \rightarrow \infty} \frac{X_n}{n} = 0$$

κατά πιθανότητα.

8.5 (Μοναδικότητα του ορίου) Αν μια ακολουθία τυχαίων μεταβλητών $(X_n)_{n \geq 1}$ συγκλίνει κατά πιθανότητα στην τυχαία μεταβλητή X αλλά και στην τυχαία μεταβλητή Y , τότε $X = Y$ με πιθανότητα 1.

8.6 Έστω $X, (X_n)_{n \geq 1}$ τυχαίες μεταβλητές και $s \geq 1$. Αν $X_n \rightarrow X$ στον \mathcal{L}^s , να δειχθεί ότι:

(α) $\mathbf{E}(|X_n|^s) \rightarrow \mathbf{E}(|X|^s)$ για $n \rightarrow \infty$.

(β) $X_n \rightarrow X$ στον \mathcal{L}^r για κάθε $r \in [1, s]$.

8.7 Για κάθε $n \in \mathbb{N}^+$, θέτουμε $\epsilon_n := 1/(2n^2)$. Έστω $(X_n)_{n \in \mathbb{N}^+}$ ακολουθία τυχαίων μεταβλητών σε κοινό χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$ ώστε

$$X_n = \begin{cases} 0 & \text{με πιθανότητα } \epsilon_n, \\ 1 & \text{με πιθανότητα } 1 - 2\epsilon_n, \\ n^2 & \text{με πιθανότητα } \epsilon_n \end{cases}$$

Τρόποι σύγκλισης ακολουθιών τυχαίων μεταβλητών

για κάθε $n \in \mathbb{N}^+$. Εξετάστε κατά πόσον η $(X_n)_{n \in \mathbb{N}^+}$ συγκλίνει σε κάποια τυχαία μεταβλητή X

(α) κατά πιθανότητα,

(β) στον \mathcal{L}^1 .

9

Μέτρα γινόμενο

9.1 Γινόμενο χώρων μέτρου. Πεπερασμένο πλήθος

Ένα μέτρο μ σε έναν μετρήσιμο χώρο (X, \mathcal{A}) λέγεται **σ -πεπερασμένο** αν υπάρχει ακολουθία $(C_n)_{n \geq 1}$ στοιχείων της \mathcal{A} ώστε $\cup_{n=1}^{\infty} C_n = X$ και $\mu(C_n) < \infty$ για κάθε $n \geq 1$. Και τότε ο χώρος (X, \mathcal{A}, μ) λέγεται χώρος σ -πεπερασμένου μέτρου.

Το μέτρο Lebesgue στον $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ όπως και το αριθμητικό μέτρο στον $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ είναι σ -πεπερασμένα, ενώ το αριθμητικό μέτρο στον $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$ δεν είναι.

Έστω τώρα $(X, \mathcal{A}, \mu), (Y, \mathcal{B}, \nu)$ δύο χώροι σ -πεπερασμένου μέτρου. Θα ορίσουμε έναν νέο χώρο μέτρου,

$$(X \times Y, \mathcal{A} \otimes \mathcal{B}, \mu \otimes \nu),$$

τον οποίο θα ονομάσουμε το γινόμενό τους.

Μετρήσιμο ορθογώνιο στον $X \times Y$ λέμε κάθε σύνολο της μορφής $A \times B$ με $A \in \mathcal{A}, B \in \mathcal{B}$. Σ -άλγεβρα γινόμενο των \mathcal{A}, \mathcal{B} ονομάζουμε τη σ -άλγεβρα που παράγεται από τα μετρήσιμα ορθογώνια, δηλαδή

$$\mathcal{A} \otimes \mathcal{B} = \sigma(\{A \times B : A \in \mathcal{A}, B \in \mathcal{B}\}).$$

Εκτός από τετριμμένες περιπτώσεις, η σ -άλγεβρα γινόμενο περιέχει πολύ περισσότερα σύνολα από τα μετρήσιμα ορθογώνια. Για παράδειγμα,

$$\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^2). \quad (9.1)$$

[Για τον εγκλεισμό \supset παρατηρούμε ότι κάθε ανοιχτό υποσύνολο του \mathbb{R}^2 γράφεται ως αριθμήσιμη ένωση $\cup_{n=1}^{\infty} (\alpha_n, \beta_n) \times (\gamma_n, \delta_n)$. Η απόδειξη της (9.1) αφήνεται ως άσκηση.] Το $\mathcal{B}(\mathbb{R}^2)$ περιέχει όλα τα «ενδιαφέροντα» υποσύνολα του \mathbb{R}^2 .

Ορίζουμε μια συνάρτηση m στα μετρήσιμα ορθογώνια ως εξής:

$$m(A \times B) = \mu(A)\nu(B) \quad (9.2)$$

για κάθε $A \in \mathcal{A}, B \in \mathcal{B}$. Αυτή η απαίτηση είναι αρκετή ώστε να ορίσει μοναδικά ένα μέτρο στον χώρο γινόμενο. Δηλαδή, αποδεικνύεται ότι υπάρχει μοναδικό μέτρο m στον μετρήσιμο χώρο $(X \times Y, \mathcal{A} \otimes \mathcal{B})$ που να ικανοποιεί την (9.2). Το m ονομάζεται **μέτρο γινόμενο** των μ, ν και θα το συμβολίζουμε με $\mu \otimes \nu$. Ο χώρος $(X \times Y, \mathcal{A} \otimes \mathcal{B}, \mu \otimes \nu)$ ονομάζεται **χώρος γινόμενο** των $(X, \mathcal{A}, \mu), (Y, \mathcal{B}, \nu)$.

Παράδειγμα 9.1. (i) Έστω μ_1 το αριθμητικό μέτρο στον $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. Ο χώρος γινόμενο των $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu_1), (\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu_1)$, δηλαδή ο $(\mathbb{N}^2, \mathcal{P}(\mathbb{N}) \otimes \mathcal{P}(\mathbb{N}), \mu_1 \otimes \mu_1)$, είναι κάτι απλό. Πρώτα $\mathcal{P}(\mathbb{N}) \otimes \mathcal{P}(\mathbb{N}) = \mathcal{P}(\mathbb{N}^2)$ αφού η σ -άλγεβρα γινόμενο περιέχει τα μονοσύνολα $\{(m, n)\} = \{m\} \times \{n\}$ και έπειτα $\mu_1 \otimes \mu_1$ είναι το αριθμητικό μέτρο στον \mathbb{N}^2 αφού κάθε μονοσύνολο $\{(m, n)\}$, ως μετρήσιμο ορθογώνιο, έχει μέτρο

$$(\mu_1 \otimes \mu_1)(\{m\} \times \{n\}) = \mu_1(\{m\})\mu_1(\{n\}) = 1.$$

(ii) Έστω λ_1 το μέτρο Lebesgue στον $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Για τον χώρο γινόμενο των $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda_1)$, $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda_1)$ ισχύει ότι $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^2)$, ενώ το μέτρο γινόμενο $\lambda_2 := \lambda_1 \otimes \lambda_1$ είναι το μέτρο που σε κάθε Borel υποσύνολο του \mathbb{R}^2 δίνει το *εμβαδόν* του. Ονομάζεται μέτρο Lebesgue στον \mathbb{R}^2 .

Ανάλογα ορίζεται ο χώρος γινόμενο για πεπερασμένο πλήθος σ -πεπερασμένων χώρων μέτρου $(X_i, \mathcal{A}_i, \mu_i), i = 1, 2, \dots, n$. Το μέτρο γινόμενο $m := \mu_1 \otimes \mu_2 \otimes \dots \otimes \mu_n$ είναι το μοναδικό μέτρο στη σ -άλγεβρα γινόμενο με την ιδιότητα

$$m(A_1 \times A_2 \times \dots \times A_n) = \mu_1(A_1)\mu_2(A_2)\dots\mu_n(A_n)$$

για κάθε $A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2, \dots, A_n \in \mathcal{A}_n$.

Το γινόμενο του $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda_1)$ n φορές με τον εαυτό του δίνει τον $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \lambda_n)$. Το μέτρο $\lambda_n := \lambda_1 \otimes \lambda_1 \otimes \dots \otimes \lambda_1$ ονομάζεται μέτρο Lebesgue στον \mathbb{R}^n . Εκτός από τα λ_1, λ_2 , γνώριμο είναι και το λ_3 , το οποίο δίνει τον όγκο κάθε Borel υποσυνόλου του \mathbb{R}^3 .

9.2 Ολοκλήρωση σε χώρο γινόμενο

Ένα ολοκλήρωμα ως προς το μέτρο γινόμενο σε έναν χώρο γινόμενο ανάγεται στον υπολογισμό ολοκληρωμάτων στους χώρους που είναι παράγοντες του γινομένου. Όπως ακριβώς στον απειροστικό λογισμό, όπου υπολογίζουμε ένα διπλό ολοκλήρωμα (στον \mathbb{R}^2) με δύο διαδοχικές ολοκληρώσεις στον \mathbb{R} .

Στην παράγραφο αυτή θα δούμε το ανάλογο αποτέλεσμα στην περίπτωση του γινομένου δύο χώρων σ -πεπερασμένου μέτρου (X, \mathcal{A}, μ) και (Y, \mathcal{B}, ν) . Δεν θα δώσουμε όμως τις αποδείξεις για τα θεωρήματα που θα διατυπώσουμε. Οι αποδείξεις ακολουθούν τη διαδικασία της τυπικής μηχανής (Παρατήρηση 6.3).

Το επόμενο αποτέλεσμα αφορά τη σ -άλγεβρα γινόμενο. Λέει ότι αν σε μια μετρήσιμη συνάρτηση δύο μεταβλητών σταθεροποιήσουμε τη μία, παίρνουμε μια συνάρτηση μιας μεταβλητής η οποία είναι πάλι μετρήσιμη.

Θεώρημα 9.2. Έστω $(X, \mathcal{A}), (Y, \mathcal{B})$ μετρήσιμοι χώροι και $f : X \times Y \rightarrow [-\infty, \infty], \mathcal{A} \otimes \mathcal{B} / \mathcal{B}([-\infty, \infty])$ μετρήσιμη συνάρτηση. Τότε:

(i) Για κάθε $x \in X$, η συνάρτηση $y \mapsto f(x, y)$ είναι $\mathcal{B} / \mathcal{B}([-\infty, \infty])$ μετρήσιμη.

(ii) Για κάθε $y \in Y$, η συνάρτηση $x \mapsto f(x, y)$ είναι $\mathcal{A} / \mathcal{B}([-\infty, \infty])$ μετρήσιμη.

Το πρώτο αποτέλεσμα για το ολοκλήρωμα ως προς το μέτρο γινόμενο αφορά μη αρνητικές συναρτήσεις.

Θεώρημα 9.3 (Tonelli). Έστω $(X, \mathcal{A}, \mu), (Y, \mathcal{B}, \nu)$ χώροι σ -πεπερασμένου μέτρου, $(X \times Y, \mathcal{A} \otimes \mathcal{B}, \mu \otimes \nu)$ ο χώρος γινόμενο και $f : X \times Y \rightarrow [0, \infty]$ μετρήσιμη συνάρτηση. Τότε οι συναρτήσεις

$$x \mapsto \int f(x, y) d\nu(y), \quad y \mapsto \int f(x, y) d\mu(x) \tag{9.3}$$

είναι $\mathcal{A} / \mathcal{B}([-\infty, \infty]), \mathcal{B} / \mathcal{B}([-\infty, \infty])$ μετρήσιμες, αντίστοιχα, και

$$\int f(x, y) d(\mu \otimes \nu)(x, y) = \int \left(\int f(x, y) d\nu(y) \right) d\mu(x) = \int \left(\int f(x, y) d\mu(x) \right) d\nu(y). \tag{9.4}$$

Τα ολοκληρώματα στην (9.3) ορίζονται γιατί από το Θεώρημα 9.2 οι συναρτήσεις τις οποίες ολοκληρώνουμε είναι μετρήσιμες και επιπλέον είναι μη αρνητικές.

Όταν η συνάρτηση την οποία ολοκληρώνουμε δεν διατηρεί απαραίτητα πρόσημο, έχουμε το εξής αποτέλεσμα.

Θεώρημα 9.4 (Fubini). Έστω $(X, \mathcal{A}, \mu), (Y, \mathcal{B}, \nu)$ χώροι σ -πεπερασμένου μέτρου, $(X \times Y, \mathcal{A} \otimes \mathcal{B}, \mu \otimes \nu)$ ο χώρος γινόμενο και $f : X \times Y \rightarrow [-\infty, \infty]$ μετρήσιμη συνάρτηση. Αν $\int |f(x, y)| d(\mu \otimes \nu)(x, y) < \infty$, τότε ισχύουν οι ισχυρισμοί του Θεωρήματος 9.3.

Από το θεώρημα Tonelli,

$$\int |f(x, y)| d(\mu \otimes \nu)(x, y) = \int \left(\int |f(x, y)| d\nu(y) \right) d\mu(x) = \int \left(\int |f(x, y)| d\mu(x) \right) d\nu(y). \quad (9.5)$$

Έτσι, όταν εφαρμόζουμε το θεώρημα Fubini και θέλουμε να ελεγχουμε αν το ολοκλήρωμα $\int |f(x, y)| d(\mu \otimes \nu)(x, y)$ είναι πεπερασμένο, ελέγχουμε αν είναι πεπερασμένο κάποιο από τα δύο διαδοχικά ολοκληρώματα στην (9.5).

Παρατήρηση 9.5. Συνήθως από το θεώρημα Fubini χρησιμοποιούμε τη δεύτερη ισότητα στην (9.4), δηλαδή την ισότητα των διαδοχικών ολοκληρωμάτων, για να αλλάξουμε σειρά ολοκλήρωσης. Θα γράψουμε τώρα αυτή την ισότητα όταν τα δύο μέτρα είναι κάποια από τα εξής τρία: το μέτρο Lebesgue στο \mathbb{R} , το αριθμητικό μέτρο στο \mathbb{N} , ή ένα μέτρο πιθανότητας \mathbf{P} . Υπό τις προϋποθέσεις του θεωρήματος κάθε φορά, έχουμε τις εξής ισότητες:

(i) Με $a_{n,k} = f(n, k)$,

$$\sum_{n=1}^{\infty} \sum_{k=1}^{\infty} a_{n,k} = \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} a_{n,k}.$$

(ii) Με $g_n(x) = f(n, x)$,

$$\int_{\mathbb{R}} \sum_{n=1}^{\infty} g_n(x) dx = \sum_{n=1}^{\infty} \int_{\mathbb{R}} g_n(x) dx.$$

(iii) Με $X_n(\omega) = f(n, \omega)$,

$$\mathbf{E} \left(\sum_{n=1}^{\infty} X_n \right) = \sum_{n=1}^{\infty} \mathbf{E}(X_n).$$

(iv)

$$\mathbf{E} \left(\int_{\mathbb{R}} f(x, \omega) dx \right) = \int_{\mathbb{R}} \mathbf{E}(f(x, \omega)) dx.$$

Εφαρμογές αυτών των ισοτήτων θα δούμε στο υπόλοιπο αυτών των σημειώσεων, στη θεωρία ή στις ασκήσεις. Για τώρα, θα δούμε μια εφαρμογή της πρώτης και μια της τέταρτης.

Παράδειγμα 9.6. Για κάθε $n \geq 1$ θέτουμε $s_n = \sum_{k=1}^n 1/k$. Θα δείξουμε ότι

$$\sum_{n=1}^{\infty} \frac{s_n}{n(n+1)} = \frac{\pi^2}{6}.$$

Μέτρα γινόμενο

Πράγματι, γράφουμε το άθροισμα ως

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{1}{n(n+1)} \sum_{k=1}^n \frac{1}{k} &= \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \frac{1}{n(n+1)} \frac{1}{k} \mathbf{1}_{k \leq n} = \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} \frac{1}{n(n+1)} \frac{1}{k} \mathbf{1}_{k \leq n} \\ &= \sum_{k=1}^{\infty} \frac{1}{k} \sum_{n=k}^{\infty} \frac{1}{n(n+1)} = \sum_{k=1}^{\infty} \frac{1}{k} \sum_{n=k}^{\infty} \left(\frac{1}{n} - \frac{1}{n+1} \right) = \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}. \end{aligned}$$

Στη δεύτερη ισότητα αλλάξαμε σειρά άθροισης, κάτι το οποίο επιτρέπεται καθώς όλοι οι προσθεταίοι είναι μη αρνητικοί (θεώρημα Tonelli). Στην προτελευταία ισότητα απλώς αθροίσαμε την τηλεσκοπική σειρά.

Παράδειγμα 9.7. Αν X είναι τυχαία μεταβλητή με τιμές στο $[0, \infty]$, τότε ισχύει

$$\mathbf{E} X = \int_0^{\infty} \mathbf{P}(X > t) dt. \quad (9.6)$$

Πράγματι το δεξί μέλος γράφεται

$$\int_0^{\infty} \mathbf{E}(\mathbf{1}_{X>t}) dt = \mathbf{E} \left(\int_0^{\infty} \mathbf{1}_{X>t} dt \right) = \mathbf{E} \left(\int_0^X 1 dt \right) = \mathbf{E} X.$$

Εδώ η συνάρτηση δύο μεταβλητών την οποία ολοκληρώνουμε είναι η $f(t, \omega) = \mathbf{1}_{X(\omega)>t}$. Η εναλλαγή σειράς ολοκλήρωσης επιτρέπεται γιατί η συνάρτηση f είναι μη αρνητική, οπότε εφαρμόζεται το θεώρημα Tonelli.

Το ολοκλήρωμα στο δεξί μέλος της (9.6) ισούται επίσης με $\int_0^{\infty} \mathbf{P}(X \geq t) dt$. Αυτό γιατί οι συναρτήσεις

$$\begin{aligned} t &\mapsto \mathbf{P}(X \geq t), \\ t &\mapsto \mathbf{P}(X > t) \end{aligned}$$

διαφέρουν σε ένα αριθμησιμο σύνολο (στα σημεία t με $\mathbf{P}(X = t) > 0$), άρα είναι λ -σχεδόν παντού ίσες.

9.3 Γινόμενο χώρων πιθανότητας. Αυθαίρετο πλήθος

Θα ορίσουμε σε αυτή την παράγραφο τον χώρο γινόμενο αυθαίρετου (ενδεχομένως και άπειρου) πλήθους χώρων μέτρου. Μας ενδιαφέρει μόνο η περίπτωση που είναι όλοι χώροι πιθανότητας.

Έστω λοιπόν σύνολο δεικτών $I \neq \emptyset$ και $(\Omega_i, \mathcal{F}_i, \mathbf{P}_i)$ χώρος πιθανότητας για κάθε $i \in I$. Θεωρούμε τον χώρο γινόμενο

$$\Omega = \prod_{i \in I} \Omega_i = \{(\omega_i)_{i \in I} : \omega_i \in \Omega_i \text{ για κάθε } i \in I\} = \left\{ \omega : I \rightarrow \cup_{i \in I} \Omega_i \mid \omega_i \in \Omega_i \text{ για κάθε } i \in I \right\}.$$

Μετρήσιμο ορθογώνιο στο Ω λέμε κάθε $A \subset \Omega$ της μορφής

$$A = \prod_{i \in I} A_i$$

ώστε $A_i \in \mathcal{F}_i$ για κάθε $i \in I$, και με το σύνολο $J = \{i \in I : A_i \neq \Omega_i\}$ πεπερασμένο.

Δηλαδή ένα μετρήσιμο ορθογώνιο είναι καρτεσιανό γινόμενο μετρήσιμων συνόλων, αλλά μόνο πεπερασμένα από αυτά διαφέρουν από τον δειγματικό χώρο του οποίου είναι υποσύνολα. Συμβολίζουμε με $\otimes_{i \in I} \mathcal{F}_i$ τη σ -άλγεβρα που παράγει το σύνολο των μετρήσιμων ορθογωνίων. Δηλαδή θέτουμε

$$\otimes_{i \in I} \mathcal{F}_i := \sigma(\{A \subset \Omega : A \text{ μετρήσιμο ορθογώνιο}\}).$$

Μέτρα γινόμενο

Για ένα μετρήσιμο ορθογώνιο A όπως πριν, ορίζουμε

$$\mathbf{P}(A) := \prod_{i \in I} \mathbf{P}_i(A_i) = \prod_{i \in J} \mathbf{P}_i(A_i).$$

Το πρώτο γινόμενο δεν πρέπει να μας ανησυχεί γιατί, ακόμα και το I να είναι άπειρο, μόνο πεπερασμένοι όροι του γινομένου είναι διαφορετικοί του 1. Η δεύτερη ισότητα ισχύει γιατί ακριβώς παραλείπουμε όρους του γινομένου που είναι σίγουρα 1, δηλαδή αυτούς με $i \in I \setminus J$.

Αποδεικνύεται ότι η \mathbf{P} επεκτείνεται μοναδικά σε μέτρο πιθανότητας στη σ -άλγεβρα $\otimes_{i \in I} \mathcal{F}_i$. Ονομάζουμε αυτή την επέκταση **μέτρο γινόμενο** των $(\mathbf{P}_i)_{i \in I}$ και το συμβολίζουμε με $\otimes_{i \in I} \mathbf{P}_i$. Αν το I είναι πεπερασμένο, έστω $I = \{1, 2, \dots, n\}$, το συμβολίζουμε με $\mathbf{P}_1 \otimes \mathbf{P}_2 \otimes \dots \otimes \mathbf{P}_n$.

Έχουμε λοιπόν ορίσει έναν νέο χώρο πιθανότητας, τον **χώρο γινόμενο**

$$\left(\prod_{i \in I} \Omega_i, \otimes_{i \in I} \mathcal{F}_i, \otimes_{i \in I} \mathbf{P}_i \right)$$

των $\{(\Omega_i, \mathcal{F}_i, \mathbf{P}_i) : i \in I\}$. Η χρησιμότητά του θα φανεί στην Παράγραφο 10.4.

Παράδειγμα 9.8 (Ένας υπολογισμός σε χώρο γινόμενο). Θεωρούμε το πείραμα ρίψης ενός νομίσματος άπειρες (αριθμησίμες) φορές που φέρνει K με πιθανότητα $p \in (0, 1)$. Ας δούμε τον χώρο πιθανότητας του πειράματος.

Για $i = 1, 2, 3, \dots$ η i -οστή ρίψη έχει χώρο πιθανότητας $(\Omega_i, \mathcal{F}_i, \mathbf{P}_i) = (\{K, \Gamma\}, \mathcal{P}(\{K, \Gamma\}), \mathbf{P}^p)$, όπου \mathbf{P}^p το μέτρο με $\mathbf{P}^p(\{K\}) = p$ και $\mathbf{P}^p(\{\Gamma\}) = 1 - p$. Ο χώρος πιθανότητας για όλο το πείραμα είναι ο χώρος γινόμενο των $\{(\Omega_i, \mathcal{F}_i, \mathbf{P}_i) : i \in \mathbb{N}^+\}$.

Ας υπολογίσουμε τώρα την πιθανότητα στις ρίψεις 2, 3 και 5 το αποτέλεσμα να είναι K, K, Γ αντίστοιχα. Το ενδεχόμενο είναι το μετρήσιμο ορθογώνιο

$$A = \{K, \Gamma\} \times \{K\} \times \{K\} \times \{K, \Gamma\} \times \{\Gamma\} \times \prod_{i \geq 6} \Omega_i.$$

Άρα

$$\mathbf{P}(A) = \mathbf{P}_2(\{K\}) \mathbf{P}_3(\{K\}) \mathbf{P}_5(\{\Gamma\}) = p^2(1 - p).$$

Ασκήσεις

9.1 Έστω $(q_k)_{k \geq 1}$ μια αρίθμηση των ρητών του $(0, 1)$. Ορίζουμε $f : \mathbb{R} \rightarrow [0, \infty]$ ως

$$f(x) = \sum_{n=1}^{\infty} \frac{1}{n^2} \frac{1}{\sqrt{|x - q_n|}}.$$

για κάθε $x \in \mathbb{R}$. Να δείξετε ότι το μέτρο Lebesgue των σημείων του $(0, 1)$ στα οποία η f απειρίζεται είναι 0, δηλαδή ότι η f είναι σχεδόν παντού πεπερασμένη.

[Υπόδειξη: Υπολογίστε το ολοκλήρωμα της f στο $(0, 1)$ ως προς το μέτρο Lebesgue.]

9.2 Έστω X τυχαία μεταβλητή με τιμές στο $[0, \infty)$.

(α) Για $g : [0, \infty) \rightarrow \mathbb{R}$ παραγωγίσιμη με συνεχή παράγωγο να δειχθεί ότι

$$\mathbf{E}(g(X)) = g(0) + \int_0^{\infty} g'(t) \mathbf{P}(X > t) dt$$

υποθέτοντας ότι $g' \geq 0$ ή ότι το ολοκλήρωμα συγκλίνει απολυτα.

Μέτρα γινόμενο

(β) Για $p > 0$ ισχύει

$$\mathbf{E}(X^p) = p \int_0^\infty t^{p-1} \mathbf{P}(X > t) dt,$$

και επιπλέον, αν η X παίρνει ακέραιες μη αρνητικές τιμές, τότε

$$\mathbf{E}(X^p) = \sum_{k=1}^{\infty} \mathbf{P}(X \geq k)(k^p - (k-1)^p).$$

9.3 Έστω X τυχαία μεταβλητή με τιμές στο \mathbb{R} για την οποία ορίζεται η $\mathbf{E}(X)$. Ναδειχθεί ότι

$$\mathbf{E}(X) = \int_0^\infty \mathbf{P}(X \geq t) dt - \int_{-\infty}^0 \mathbf{P}(X \leq t) dt = \int_{-\infty}^\infty \{\mathbf{P}(X \geq t) - \mathbf{1}_{t < 0}\} dt.$$

9.4 Διάμεσο ενός μέτρου πιθανότητας ν στο \mathbb{R} λέμε οποιονδήποτε αριθμό m ικανοποιεί τις

$$\nu((-\infty, m]) \geq \frac{1}{2}, \nu([m, \infty)) \geq \frac{1}{2}.$$

Διάμεσο μιας τυχαίας μεταβλητής με τιμές στο \mathbb{R} λέμε οποιονδήποτε διάμεσο της κατανομής της.

(α) Ναδειχθεί ότι κάθε μέτρο πιθανότητας στο \mathbb{R} έχει τουλάχιστον έναν διάμεσο και να δοθεί παράδειγμα όπου ο διάμεσος δεν είναι μοναδικός.

(β) Έστω X τυχαία μεταβλητή με $\mathbf{E}|X| < \infty$ και της οποίας η συνάρτηση κατανομής είναι συνεχής. Ναδειχθεί ότι οι διάμεσοι της X είναι ακριβώς τα σημεία στα οποία παίρνει το ολικό της ελάχιστο η συνάρτηση $f(a) := \mathbf{E}|X - a|$.

(γ) Έστω X τυχαία μεταβλητή με μέση τιμή μ , διάμεσο m , και διασπορά σ^2 . Ναδειχθεί ότι

$$|m - \mu| \leq \sigma$$

και η ισότητα ισχύει μόνο όταν η X είναι σταθερή.

9.5 Αποδείξτε την Πρόταση 4.9(ii) δείχνοντας πρώτα ότι ισχύει για κάθε g της μορφής $\mathbf{1}_A$ όπου

(α) A είναι μετρήσιμο ορθογώνιο στον \mathbb{R}^n , δηλαδή της μορφής $A_1 \times A_2 \times \cdots \times A_n$ με $A_i \in \mathcal{B}(\mathbb{R})$ για κάθε i .

(β) $A \in \mathcal{B}(\mathbb{R}^n)$. [Υπόδειξη: Χρησιμοποιήστε το θεώρημα π-λ.]

10

Ανεξαρτησία

10.1 Ανεξαρτησία για οικογένειες συνόλων και τυχαίες μεταβλητές

Στην παράγραφο αυτή δουλεύουμε σε χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$.

Δίνουμε καταρχάς τον ορισμό της ανεξαρτησίας για ενδεχόμενα, σύνολα ενδεχομένων, και τυχαίες μεταβλητές.

Ορισμός 10.1. Έστω $(A_i)_{i \in I}$ στοιχεία της \mathcal{F} . Τα $(A_i)_{i \in I}$ λέγονται **ανεξάρτητα** αν για κάθε $J \subset I$ πεπερασμένο ισχύει ότι

$$\mathbf{P}(\cap_{i \in J} A_i) = \prod_{i \in J} \mathbf{P}(A_i). \quad (10.1)$$

Η τομή και το γινόμενο στην τελευταία ισότητα έχουν πεπερασμένο πλήθος όρων.

Ορισμός 10.2. Έστω $(\mathcal{F}_i)_{i \in I}$ οικογένεια υποσυνόλων της \mathcal{F} (δηλαδή $\mathcal{F}_i \subset \mathcal{F}$ για κάθε $i \in I$). Η οικογένεια $(\mathcal{F}_i)_{i \in I}$ λέγεται **ανεξάρτητη** αν για κάθε $J \subset I$ πεπερασμένο και $A_i \in \mathcal{F}_i$ για κάθε $i \in J$ ισχύει η (10.1). Λέμε επίσης ότι τα $(\mathcal{F}_i)_{i \in I}$ είναι ανεξάρτητα.

Ορισμός 10.3. Έστω $\{(E_i, \mathcal{E}_i) : i \in I\}$ μετρήσιμοι χώροι και $(X_i)_{i \in I}$ οικογένεια τυχαίων μεταβλητών με $X_i : \Omega \rightarrow E_i$ για κάθε $i \in I$. Οι $(X_i)_{i \in I}$ λέγονται **ανεξάρτητες** αν η οικογένεια των σ -άλγεβρων $(\sigma(X_i))_{i \in I}$, που είναι υποσύνολα της \mathcal{F} , είναι ανεξάρτητη.

Παρατήρηση 10.4. Ο Ορισμός 10.3, σύμφωνα με τον Ορισμό 10.2, απαιτεί

$$\mathbf{P}(X_{i_1} \in A_{i_1}, X_{i_2} \in A_{i_2}, \dots, X_{i_n} \in A_{i_n}) = \mathbf{P}(X_{i_1} \in A_{i_1}) \mathbf{P}(X_{i_2} \in A_{i_2}) \cdots \mathbf{P}(X_{i_n} \in A_{i_n}) \quad (10.2)$$

για κάθε $n \geq 2$, κάθε επιλογή δεικτών $i_1, i_2, \dots, i_n \in I$ διαφορετικών ανά δύο, και κάθε $A_{i_1} \in \mathcal{E}_{i_1}, \dots, A_{i_n} \in \mathcal{E}_{i_n}$, αφού κάθε στοιχείο μιας $\sigma(X_i)$ είναι της μορφής $X_i^{-1}(A_i) = \{X_i \in A_i\}$ με $A_i \in \mathcal{E}_i$. Το ενδεχόμενο στο αριστερό μέλος της (10.2) είναι συντομογραφία του ενδεχομένου $X_{i_1}^{-1}(A_{i_1}) \cap X_{i_2}^{-1}(A_{i_2}) \cap \cdots \cap X_{i_n}^{-1}(A_{i_n})$.

Σύμβαση: Στο εξής, όποτε λέμε ότι κάποιες τυχαίες μεταβλητές είναι ανεξάρτητες, θα εννοείται ότι ορίζονται στον ίδιο χώρο πιθανότητας, δηλαδή έχουν το ίδιο πεδίο ορισμού. Αυτό εξασφαλίζει ότι το ενδεχόμενο $X_{i_1}^{-1}(A_{i_1}) \cap X_{i_2}^{-1}(A_{i_2}) \cap \cdots \cap X_{i_n}^{-1}(A_{i_n})$ στο αριστερό μέλος της (10.2) είναι στοιχείο της \mathcal{F} , που είναι το πεδίο ορισμού της \mathbf{P} .

Παράδειγμα 10.5 (Δύο ανεξάρτητες τυχαίες μεταβλητές). Θεωρούμε το πείραμα δύο ρίψεων ενός νομίσματος που φέρνει κορώνα με πιθανότητα p . Το σύνηθες μέτρο που μοντελοποιεί το πείραμα είναι τέτοιο ώστε τα αποτελέσματα των δύο ρίψεων να είναι ανεξάρτητες τυχαίες μεταβλητές.

Πιο συγκεκριμένα, ο δειγματικός μας χώρος είναι ο $\Omega = \{K, \Gamma\} \times \{K, \Gamma\}$ και σ -άλγεβρα η $\mathcal{F} = \mathcal{P}(\Omega)$. Για $\omega \in \Omega$ θέτουμε

$$p_\omega = \begin{cases} p(1-p) & \text{αν } \omega = (K, \Gamma) \text{ ή } (\Gamma, K), \\ p^2 & \text{αν } \omega = (K, K), \\ (1-p)^2 & \text{αν } \omega = (\Gamma, \Gamma). \end{cases}$$

Ανεξαρτησία

Έστω \mathbf{P} το μοναδικό μέτρο πιθανότητας στην \mathcal{F} με $\mathbf{P}(\omega) = p_\omega$. Έστω $E = \{K, \Gamma\}$ και $\mathcal{E} = \mathcal{P}(E)$. Θεωρούμε τις τυχαίες μεταβλητές $X, Y : \Omega \rightarrow E$, με $X((x, y)) = x$ και $Y((x, y)) = y$. Δηλαδή, η X είναι η ένδειξη της πρώτης ρίψης και η Y η ένδειξη της δεύτερης.

ΙΣΧΥΡΙΣΜΟΣ: Οι X, Y είναι ανεξάρτητες.

Θα δείξουμε ότι για κάθε $A, B \in \mathcal{E}$ ισχύει

$$\mathbf{P}(X \in A, Y \in B) = \mathbf{P}(X \in A) \mathbf{P}(Y \in B). \quad (10.3)$$

- Αν $A = \emptyset$ ή $B = \emptyset$, η (10.3) ισχύει.
- Αν $A = \{K, \Gamma\}$, τότε $\{X \in A, Y \in B\} = \{Y \in B\}$, και $\mathbf{P}(X \in A) = 1$. Άρα η (10.3) πάλι ισχύει.
- Αν $B = \{K, \Gamma\}$, η (10.3) αποδεικνύεται όμοια.

Τέλος, μένουν οι περιπτώσεις που τα A, B είναι μονοσύνολα. Για παράδειγμα, αν $A = \{K\}$ και $B = \{\Gamma\}$, έχουμε

$$\begin{aligned} \mathbf{P}(X = K, Y = \Gamma) &= \mathbf{P}(X^{-1}(\{K\}) \cap Y^{-1}(\{\Gamma\})) = \mathbf{P}(\{(K, \Gamma), (K, K)\} \cap \{(K, \Gamma), (\Gamma, \Gamma)\}) \\ &= \mathbf{P}(\{K, \Gamma\}) = p(1 - p). \end{aligned}$$

Όμως

$$\begin{aligned} \mathbf{P}(X = K) &= \mathbf{P}(\{(K, \Gamma), (K, K)\}) = p(1 - p) + pp = p \\ \mathbf{P}(Y = \Gamma) &= \mathbf{P}(\{(K, \Gamma), (\Gamma, \Gamma)\}) = p(1 - p) + (1 - p)^2 = 1 - p, \end{aligned}$$

που το γινόμενο τους είναι $p(1 - p) = \mathbf{P}(X = K, Y = \Gamma)$, και έτσι η (10.3) ισχύει πάλι.

Όμοια αποδεικνύονται και οι υπόλοιπες περιπτώσεις όπου τα A, B είναι μονοσύνολα.

Το επόμενο θεώρημα διευκολύνει τον έλεγχο ανεξαρτησίας δύο τυχαίων μεταβλητών.

Θεώρημα 10.6. Έστω $(E, \mathcal{E}), (G, \mathcal{G})$ μετρήσιμοι χώροι και $X : \Omega \rightarrow E, Y : \Omega \rightarrow G$ τυχαίες μεταβλητές. Θεωρούμε τη σχέση

$$\mathbf{P}(X \in A, Y \in B) = \mathbf{P}(X \in A) \mathbf{P}(Y \in B). \quad (*)$$

Τα εξής είναι ισοδύναμα:

- Οι X, Y είναι ανεξάρτητες.
- $H (*)$ ισχύει για κάθε $A \in \mathcal{E}$ και $B \in \mathcal{G}$.
- $H (*)$ ισχύει για κάθε $A \in \mathcal{C}, B \in \mathcal{D}$, όπου \mathcal{C}, \mathcal{D} οικογένειες κλειστές στις πεπερασμένες τομές με $\sigma(\mathcal{C}) = \mathcal{E}, \sigma(\mathcal{D}) = \mathcal{G}$.

Σημαντική είναι η ισοδυναμία των (i) και (iii). Δηλαδή αρκεί να ελέγξουμε την (*) για λιγότερα σύνολα από ότι απαιτεί η (ii) ώστε να διαπιστώσουμε την ανεξαρτησία των X, Y .

Απόδειξη. Η (ii) είναι αναδιατύπωση του ορισμού της ανεξαρτησίας και προφανώς συνεπάγεται την (iii). Μένει να δείξουμε ότι η (iii) συνεπάγεται την (ii).

Έστω $A \in \mathcal{C}$ και

$$\mathcal{D}_1(A) := \{B \in \mathcal{G} : \text{η } (*) \text{ ισχύει για τα } A, B\}.$$

Έχουμε ότι $\mathcal{D} \subset \mathcal{D}_1(A)$ από υπόθεση και η $\mathcal{D}_1(A)$ είναι κλάση Dynkin (αποδεικνύεται όμοια όπως στην Άσκηση 3.1). Άρα $\mathcal{D} \subset \mathcal{D}_1(A)$. Επειδή η \mathcal{D} είναι κλειστή στις πεπερασμένες τομές, το

Ανεξαρτησία

θεώρημα π-λ δίνει ότι $\sigma(\mathcal{D}) = \delta(\mathcal{D})$. Άρα $\sigma(\mathcal{D}) \subset \mathcal{D}_1(A)$, δηλαδή η (*) ισχύει για κάθε $A \in \mathcal{C}$ και $B \in \mathcal{G}$. Τώρα για $B \in \mathcal{G}$ θέτουμε

$$\mathcal{D}_2(B) := \{A \in \mathcal{E} : \eta (*) \text{ ισχύει για τα } A, B\}.$$

Όμοια, όπως με το $\mathcal{D}_1(A)$, δείχνουμε ότι $\mathcal{D}_2(B) = \mathcal{E}$, και έτσι αποδείχθηκε η (ii). ■

Στις στοιχειώδεις πιθανότητες μαθαίνουμε (χωρίς απόδειξη) ότι δύο τυχαίες μεταβλητές X, Y είναι ανεξάρτητες αν και μόνο αν η από κοινού συνάρτηση κατανομή τους, $F_{X,Y}$, γράφεται ως $F_{X,Y}(x, y) = F_X(y)F_Y(y)$ για κάθε $x, y \in \mathbb{R}$. Τώρα είμαστε σε θέση να το αποδείξουμε.

Πόρισμα 10.7. Έστω $X, Y : \Omega \rightarrow \mathbb{R}$ τυχαίες μεταβλητές. Τότε οι X, Y είναι ανεξάρτητες αν και μόνο αν

$$\mathbf{P}(X \leq x, Y \leq y) = \mathbf{P}(X \leq x)\mathbf{P}(Y \leq y)$$

για κάθε $x, y \in \mathbb{R}$.

Απόδειξη. Προκύπτει από το Θεώρημα 10.6 αν πάρουμε $C = \mathcal{D} = \{(-\infty, a] : a \in \mathbb{R}\}$. ■

Θεώρημα 10.8. Έστω $X, Y : \Omega \rightarrow [-\infty, \infty]$ ανεξάρτητες τυχαίες μεταβλητές μη αρνητικές ή με $\mathbf{E}|X|, \mathbf{E}|Y| < \infty$. Τότε

$$\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y).$$

Απόδειξη. Το σχέδιο της απόδειξης είναι να δείξουμε κάτι φαινομενικά ισχυρότερο. Δηλαδή το εξής.

ΙΣΧΥΡΙΣΜΟΣ:

$$\mathbf{E}\{f(X)g(Y)\} = \mathbf{E}\{f(X)\}\mathbf{E}\{g(Y)\} \quad (10.4)$$

για κάθε $f : [-\infty, \infty] \rightarrow [-\infty, \infty]$, $g : [-\infty, \infty] \rightarrow [-\infty, \infty]$ μετρήσιμες που είναι μη αρνητικές ή ικανοποιούν τις $\mathbf{E}|f(X)|, \mathbf{E}|g(Y)| < \infty$.

Θα δείξουμε την (10.4) σταδιακά με τον γνωστό τρόπο (Τυπική Μηχανή).

Βήμα 1. Αν $f = \mathbf{1}_A$ και $g = \mathbf{1}_B$, όπου $A, B \in \mathcal{B}([-\infty, \infty])$, έχουμε ότι

$$\begin{aligned} \mathbf{E}(f(X)g(Y)) &= \mathbf{E}(\mathbf{1}_{X \in A}\mathbf{1}_{Y \in B}) = \mathbf{E}(\mathbf{1}_{\{X \in A\} \cap \{Y \in B\}}) = \mathbf{P}(X \in A, Y \in B) \\ &= \mathbf{P}(X \in A)\mathbf{P}(Y \in B) = \mathbf{E}(\mathbf{1}_{X \in A})\mathbf{E}(\mathbf{1}_{Y \in B}). \end{aligned}$$

Η τέταρτη ισότητα ισχύει γιατί οι X, Y είναι ανεξάρτητες. Και έτσι προκύπτει η (10.4) για τις συγκεκριμένες f, g .

Βήμα 2. Αν $f, g \geq 0$ απλές μετρήσιμες, έστω

$$f = \sum_{i=1}^m a_i \mathbf{1}_{A_i}, \quad g = \sum_{j=1}^n b_j \mathbf{1}_{B_j},$$

σε κανονική μορφή, όπου $A_i, B_j \in \mathcal{B}([-\infty, \infty])$ και $a_i, b_j \in [0, \infty)$, τότε

$$\mathbf{E}(f(X)g(Y)) = \mathbf{E}\left(\sum_{i=1}^m \sum_{j=1}^n a_i b_j \mathbf{1}_{A_i}(X) \mathbf{1}_{B_j}(Y)\right).$$

Και λόγω γραμμικότητας και του προηγούμενου βήματος, η τελευταία μέση τιμή ισούται με

$$\sum_{i=1}^m \sum_{j=1}^n a_i b_j \mathbf{E}(\mathbf{1}_{A_i}(X)) \mathbf{E}(\mathbf{1}_{B_j}(Y)) = \mathbf{E}(f(X)) \mathbf{E}(g(Y)).$$

Βήμα 3. Αν $f, g \geq 0$ μετρήσιμες, τότε υπάρχουν αύξουσες ακολουθίες $(r_n)_{n \in \mathbb{N}}$ και $(s_n)_{n \in \mathbb{N}}$ μη αρνητικών απλών συναρτήσεων έτσι ώστε

$$\lim_{n \rightarrow \infty} r_n = f, \quad \lim_{n \rightarrow \infty} s_n = g.$$

Συνεπώς, από τα προηγούμενα έχουμε ότι

$$\mathbf{E}(r_n(X)s_n(Y)) = \mathbf{E}(r_n(X))\mathbf{E}(s_n(Y)) \quad \text{για κάθε } n \in \mathbb{N}.$$

Επειδή οι ακολουθίες $(r_n)_{n \in \mathbb{N}}$ και $(s_n)_{n \in \mathbb{N}}$ είναι αύξουσες, για $n \rightarrow \infty$, από το θεώρημα μονότονης σύγκλισης έχουμε ότι

$$\mathbf{E}(f(X)g(Y)) = \mathbf{E}(f(X))\mathbf{E}(g(Y)).$$

Δηλαδή η (10.4) ισχύει για τις f, g .

Βήμα 4. Αν οι f, g είναι μετρήσιμες με $\mathbf{E}|f(X)|, \mathbf{E}|g(Y)| < \infty$, έχουμε

$$\begin{aligned} \mathbf{E}\{f(X)g(Y)\} &= \mathbf{E}\{(f^+(X) - f^-(X))(g^+(Y) - g^-(Y))\} \\ &= \mathbf{E}\{f^+(X)g^+(Y)\} - \mathbf{E}\{f^+(X)g^-(Y)\} - \mathbf{E}\{f^-(X)g^+(Y)\} + \mathbf{E}\{f^-(X)g^-(Y)\} \\ &= \mathbf{E}(f^+(X))\mathbf{E}(g^+(Y)) - \mathbf{E}(f^+(X))\mathbf{E}(g^-(Y)) - \mathbf{E}(f^-(X))\mathbf{E}(g^+(Y)) + \mathbf{E}(f^-(X))\mathbf{E}(g^-(Y)) \\ &= \mathbf{E}(f(X))\mathbf{E}(g(Y)). \end{aligned}$$

Στην τρίτη ισότητα χρησιμοποιήσαμε το ότι η (10.4) ισχύει για μη αρνητικές μετρήσιμες. Επίσης στον τελευταίο υπολογισμό δεν εμφανίζεται πουθενά κάποια απροσδιόριστη μορφή $\infty - \infty$ γιατί οι f, g ικανοποιούν $\mathbf{E}|f(X)|, \mathbf{E}|g(Y)| < \infty$. Έτσι ο ισχυρισμός αποδείχθηκε.

Επιστρέφουμε τώρα στην απόδειξη του θεωρήματος. Αν $X, Y \geq 0$, εφαρμόζουμε τον ισχυρισμό για

$$f(x) = g(x) = x^+ = \begin{cases} x & \text{αν } x \in [0, \infty], \\ 0 & \text{αν } x \in [-\infty, 0). \end{cases}$$

Στην περίπτωση που $\mathbf{E}|X|, \mathbf{E}|Y| < \infty$, εφαρμόζουμε τον ισχυρισμό για $f(x) = g(x) = x$ για κάθε $x \in [-\infty, \infty]$. ■

Παρατήρηση 10.9. Ανάλογα των Θεωρημάτων 10.6, 10.8 ισχύουν αν αντί δύο έχουμε περισσότερες ανεξάρτητες τυχαίες μεταβλητές, έστω X_1, X_2, \dots, X_n . Για παράδειγμα, η αντίστοιχη της (10.4) είναι η

$$\mathbf{E}\{f_1(X_1)f_2(X_2)\cdots f_n(X_n)\} = \mathbf{E}\{f_1(X_1)\}\mathbf{E}\{f_2(X_2)\}\cdots \mathbf{E}\{f_n(X_n)\}$$

με τις f_1, f_2, \dots, f_n μετρήσιμες και μη αρνητικές ή με $\mathbf{E}|f_k(X_k)| < \infty$ για κάθε k . Η απόδειξη των αντίστοιχων αυτών ισχυρισμών γίνεται με επαγωγή.

Υπενθυμίζουμε εδώ ότι για $(X_k)_{1 \leq k \leq n}$ τυχαίες μεταβλητές στον ίδιο χώρο πιθανότητας, με πραγματικές τιμές και με $\mathbf{E}(X_k^2) < \infty$ για κάθε $k \in \{1, 2, \dots, n\}$, ισχύει

$$\text{Var}(X_1 + X_2 + \cdots + X_n) = \text{Var}(X_1) + \cdots + \text{Var}(X_n) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j). \quad (10.5)$$

Η απόδειξη γίνεται όπως ακριβώς την έχουμε δει στις στοιχειώδεις πιθανότητες. Όταν οι $(X_k)_{1 \leq k \leq n}$ είναι ανεξάρτητες, το προηγούμενο θεώρημα δίνει ότι όλες οι συνδιακυμάνσεις είναι 0, οπότε

$$\text{Var}(X_1 + X_2 + \cdots + X_n) = \text{Var}(X_1) + \cdots + \text{Var}(X_n). \quad (10.6)$$

10.2 Ανεξαρτησία και ομαδοποίηση

Αν οι τυχαίες μεταβλητές X, Y, Z, V, W είναι ανεξάρτητες μεταξύ τους, περιμένουμε και οι $X + Y, Z^2W, |V|$ να είναι ανεξάρτητες εφόσον χρησιμοποιούν διαφορετικά ανεξάρτητα συστατικά. Θα διατυπώσουμε ένα θεώρημα που δίνει αποτελέσματα αυτής της μορφής. Προηγουμένως, διατυπώνουμε το αντίστοιχο αποτέλεσμα για οικογένειες συνόλων. Η απόδειξή του δίνεται στο Παράρτημα Β'.

Θεώρημα 10.10. Έστω $(\mathcal{F}_i)_{i \in I}$ ανεξάρτητη οικογένεια υποσυνόλων της \mathcal{F} και $\{I_j : j \in J\}$ διαμέριση ¹ του συνόλου δεικτών I . Για κάθε $j \in J$ θεωρούμε τη σ-άλγεβρα

$$\mathcal{G}_j := \sigma\left(\cup_{i \in I_j} \mathcal{F}_i\right).$$

Οι $(\mathcal{G}_j)_{j \in J}$ είναι ανεξάρτητες.

Και το αποτέλεσμα που αφορά τυχαίες μεταβλητές είναι το εξής.

Θεώρημα 10.11. Έστω $(X_i)_{i \in I}$ ανεξάρτητες τυχαίες μεταβλητές, $\{I_j : j \in J\}$ διαμέριση του συνόλου δεικτών I , και για κάθε $j \in J$, μετρήσιμη συνάρτηση $f_j : \mathbb{R}^{I_j} \rightarrow \mathbb{R}$. Για κάθε $j \in J$ θεωρούμε τη συνάρτηση $Y_j := f_j((X_i)_{i \in I_j}) : \Omega \rightarrow \mathbb{R}$. Οι $(Y_j)_{j \in J}$ είναι ανεξάρτητες τυχαίες μεταβλητές.

Για αυτό το θεώρημα, σε πρώτη ανάγνωση καλό είναι να υποθέσει κανείς ότι τα σύνολα I, I_j είναι πεπερασμένα και επομένως οι συναρτήσεις f_j ορίζονται σε χώρους της μορφής \mathbb{R}^d . Σε όλες τις περιπτώσεις θεωρούμε ότι ο \mathbb{R}^{I_j} είναι εφοδιασμένος με τη σ-άλγεβρα γινόμενο $\otimes_{i \in I_j} \mathcal{B}(\mathbb{R})$ (όλοι οι όροι του γινομένου είναι ίδιοι). Η απόδειξη και αυτού του θεωρήματος δίνεται στο Παράρτημα Β'.

Εφαρμόζοντάς το έχουμε, για παράδειγμα, ότι αν οι τυχαίες μεταβλητές $(X_n)_{n \geq 1}$ είναι ανεξάρτητες, τότε και οι $(Y_n)_{n \geq 1}$ με $Y_n := \sum_{i=2^{n+1}}^{2^{n+2}} X_i$ για κάθε $n \geq 1$ είναι ανεξάρτητες. Όμοια, είναι ανεξάρτητα και τα σύνολα $\{X_{2n} + X_{2n+1} > 0\}, n \geq 1$.

10.3 Ανεξαρτησία=Μέτρο γινόμενο

Πρόταση 10.12. Έστω $X := (X_1, X_2, \dots, X_n)$ τυχαία μεταβλητή με τιμές στον \mathbb{R}^n . Τότε οι X_1, X_2, \dots, X_n είναι ανεξάρτητες αν και μόνο αν

$$\mathbf{P}^X = \mathbf{P}^{X_1} \otimes \mathbf{P}^{X_2} \otimes \dots \otimes \mathbf{P}^{X_n}.$$

Απόδειξη. \Rightarrow Η τιμή του μέτρου \mathbf{P}^X σε ένα μετρήσιμο ορθογώνιο $A := A_1 \times A_2 \times \dots \times A_n$ του \mathbb{R}^n είναι

$$\begin{aligned} \mathbf{P}^X(A) &= \mathbf{P}(X_1 \in A_1, \dots, X_n \in A_n) = \mathbf{P}(X_1 \in A_1) \mathbf{P}(X_2 \in A_2) \dots \mathbf{P}(X_n \in A_n) \\ &= \mathbf{P}^{X_1}(A_1) \mathbf{P}^{X_2}(A_2) \dots \mathbf{P}^{X_n}(A_n). \end{aligned}$$

Στη δεύτερη ισότητα χρησιμοποιήσαμε την ανεξαρτησία των X_1, \dots, X_n . Όμως το $\mathbf{P}^{X_1} \otimes \mathbf{P}^{X_2} \otimes \dots \otimes \mathbf{P}^{X_n}$ είναι το μοναδικό μέτρο που παίρνει την τιμή $\mathbf{P}^{X_1}(A_1) \mathbf{P}^{X_2}(A_2) \dots \mathbf{P}^{X_n}(A_n)$ στο A . Η ζητούμενη ισότητα έπεται.

\Leftarrow Ελέγχουμε τη σχέση (10.2). Έστω $A_1, A_2, \dots, A_n \in \mathcal{B}(\mathbb{R})$. Τότε

$$\begin{aligned} \mathbf{P}(X_1 \in A_1, \dots, X_n \in A_n) &= \mathbf{P}^{X_1} \otimes \mathbf{P}^{X_2} \otimes \dots \otimes \mathbf{P}^{X_n}(A_1 \times A_2 \times \dots \times A_n) \\ &= \mathbf{P}^{X_1}(A_1) \mathbf{P}^{X_2}(A_2) \dots \mathbf{P}^{X_n}(A_n) \\ &= \mathbf{P}(X_1 \in A_1) \mathbf{P}(X_2 \in A_2) \dots \mathbf{P}(X_n \in A_n). \end{aligned}$$

¹ Δηλαδή τα $I_j (j \in J)$, είναι μη κενά, ξένα ανα δύο, και έχουν ένωση το I .

Στην πρώτη ισότητα χρησιμοποιήσαμε την υπόθεση και στη δεύτερη τον ορισμό του μέτρου γινομένου. ■

Η προηγούμενη πρόταση σε συνδυασμό με τα θεωρήματα Tonelli και Fubini μας δίνει τον πιο συνηθισμένο τρόπο υπολογισμού της μέσης τιμής ποσοτήτων που είναι συναρτήσεις ανεξάρτητων τυχαίων μεταβλητών. Ο τρόπος αυτός φαίνεται στο επόμενο παράδειγμα.

Παράδειγμα 10.13. Έστω X, Y ανεξάρτητες τυχαίες μεταβλητές ώστε $X \sim \text{Exp}(2)$ και $Y \sim U(0, 1)$. Θα υπολογίσουμε τη μέση τιμή $\mathbf{E}(e^{XY})$. Από την Πρόταση 10.12 και το Θεώρημα Tonelli έχουμε ότι

$$\mathbf{E}(e^{XY}) = \int \int e^{xy} d\mathbf{P}_X(x) d\mathbf{P}_Y(y) = \int g(y) d\mathbf{P}_Y(y) = \mathbf{E}(g(Y))$$

όπου $g(y) = \int e^{xy} d\mathbf{P}_X(x) = \mathbf{E}(e^{Xy})$. Αυτό σημαίνει ότι βρίσκουμε τη μέση τιμή σε δύο βήματα. Πρώτα σταθεροποιούμε την τυχαία μεταβλητή Y και παίρνουμε μέση τιμή μόνο ως προς την X . Το αποτέλεσμα είναι η τυχαία μεταβλητή $g(Y)$. Έπειτα παίρνουμε τη μέση τιμή της $g(Y)$. Οι υπολογισμοί έχουν ως εξής:

$$g(y) = \mathbf{E}(e^{Xy}) = \frac{2}{2-y}$$

και έπειτα

$$\mathbf{E}\{g(Y)\} = \int_0^1 g(y) dy = 2 \log 2.$$

Βέβαια θα μπορούσαμε να κάνουμε την ολοκλήρωση με άλλη σειρά. Δηλαδή να πάρουμε πρώτα μέση τιμή ως προς Y (με το X σταθερό) και μετά ως προς X .

10.4 Κατασκευή ανεξάρτητων τυχαίων μεταβλητών με δεδομένη κατανομή

Υπάρχουν ανεξάρτητες τυχαίες μεταβλητές; Πώς μπορούμε να κατασκευάσουμε τέτοιες και βέβαια τον χώρο πιθανότητας στον οποίο αυτές ορίζονται; Τη λύση σε αυτά τα προβλήματα δίνουν οι χώροι γινόμενο, τους οποίους είδαμε στην Παράγραφο 9.3. Σε αυτή την παράγραφο θα τους χρησιμοποιήσουμε για να κατασκευάσουμε ανεξάρτητες τυχαίες μεταβλητές με επιθυμητές ιδιότητες.

Υπενθυμίζουμε ότι αν (E, \mathcal{E}) είναι μετρήσιμος χώρος και $X : \Omega \rightarrow E$ τυχαία μεταβλητή, κατανομή της X λέμε το μέτρο πιθανότητας \mathbf{P}^X που ορίζεται στον (E, \mathcal{E}) με $\mathbf{P}^X(A) = \mathbf{P}(X^{-1}(A))$ για κάθε $A \in \mathcal{E}$.

Έστω I σύνολο δεικτών και $(E_i, \mathcal{E}_i, \mathbf{P}_i)_{i \in I}$ οικογένεια χώρων πιθανότητας. Θέλουμε να κατασκευάσουμε έναν χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$ και τυχαίες μεταβλητές $X_i : \Omega \rightarrow E_i$ έτσι ώστε

- (α) Η κατανομή της X_i να είναι η \mathbf{P}_i για κάθε $i \in I$.
- (β) Η $(X_i)_{i \in I}$ να είναι οικογένεια ανεξάρτητων τυχαίων μεταβλητών.

Το να κάνουμε ένα από τα (α) ή (β) είναι εύκολο. Αυτό που είναι μη τετριμμένο είναι να κάνουμε και τα δύο μαζί. Και το κατορθώνουμε με χρήση του χώρου γινομένου ως εξής.

Έστω $\Omega = \prod_{i \in I} E_i, \mathcal{F} = \otimes_{i \in I} \mathcal{E}_i$, και \mathbf{P} το μέτρο γινόμενο των $\mathbf{P}_i, i \in I$. Για κάθε $r \in I$ ορίζουμε $X_r : \Omega \rightarrow E_r$ ως

$$X_r((\omega_i)_{i \in I}) = \omega_r,$$

δηλαδή η X_r είναι η προβολή στην r -συντεταγμένη.

Ανεξαρτησία

Εύκολα βλέπουμε ότι η X_r είναι τυχαία μεταβλητή γιατί για $A \in \mathcal{E}_r$ έχουμε

$$X_r^{-1}(A) = \prod_{i \in I} A_i, \text{ με } A_i = \begin{cases} \Omega & \text{αν } i \neq r, \\ A & \text{αν } i = r. \end{cases}$$

Άρα $X_r^{-1}(A) \in \mathcal{F}$ ως μετρήσιμο ορθογώνιο.

Θα δείξουμε τώρα ότι πράγματι αυτή η κατασκευή ικανοποιεί τις απαιτήσεις του προβλήματος.

Πρόταση 10.14. Έστω $(\Omega, \mathcal{F}, \mathbf{P})$, $(E_i, \mathcal{E}_i, \mathbf{P}_i)_{i \in I}$ όπως πριν και $X_r : \Omega \rightarrow E_r$, $r \in I$, η r -προβολή. Τότε:

- (i) Η X_r έχει κατανομή \mathbf{P}_r .
- (ii) Οι $(X_r)_{r \in I}$ είναι ανεξάρτητες.

Απόδειξη. (i) Έστω $A \in \mathcal{E}_r$. Έχουμε,

$$\mathbf{P}^{X_r}(A) = \mathbf{P}(X_r^{-1}(A)) = \mathbf{P}\left(\prod_{i \in I} A_i\right) = \prod_{i \in I} \mathbf{P}_i(A_i) = \mathbf{P}_r(A),$$

γράφοντας το $X_r^{-1}(A)$ ως μετρήσιμο ορθογώνιο όπως προηγουμένως. Άρα, $\mathbf{P}^{X_r} = \mathbf{P}_r$.

(ii) Έστω $J = \{j_1, j_2, \dots, j_n\} \subset I$, και $B_{j_1} \in \mathcal{E}_{j_1}, B_{j_2} \in \mathcal{E}_{j_2}, \dots, B_{j_n} \in \mathcal{E}_{j_n}$. Τότε

$$\mathbf{P}(\{X_{j_1} \in B_{j_1}\} \cap \{X_{j_2} \in B_{j_2}\} \cap \dots \cap \{X_{j_n} \in B_{j_n}\}) = \mathbf{P}\left(\prod_{i \in I} A_i\right),$$

όπου

$$A_i = \begin{cases} \Omega_i & \text{αν } i \in I \setminus J, \\ B_i, & \text{αν } i \in J. \end{cases}$$

Από τον ορισμό του \mathbf{P} έχουμε ότι

$$\begin{aligned} \mathbf{P}\left(\prod_{i \in I} A_i\right) &= \prod_{i \in I} \mathbf{P}(A_i) = \prod_{i \in J} \mathbf{P}(A_i) = \mathbf{P}_{j_1}(B_{j_1}) \mathbf{P}_{j_2}(B_{j_2}) \dots \mathbf{P}_{j_n}(B_{j_n}) \\ &= \mathbf{P}(X_{j_1} \in B_{j_1}) \mathbf{P}(X_{j_2} \in B_{j_2}) \dots \mathbf{P}(X_{j_n} \in B_{j_n}). \end{aligned}$$

Συνεπώς οι $\{X_r : r \in I\}$ είναι ανεξάρτητες. ■

Μια συνέπεια της Πρότασης 10.14 είναι ότι για δεδομένη κατανομή \mathbf{Q} σε έναν μετρήσιμο χώρο (E, \mathcal{E}) και σύνολο I υπάρχει σύνολο ανεξάρτητων και ισόνομων τυχαίων μεταβλητών $(X_i)_{i \in I}$ που καθεμία έχει κατανομή \mathbf{Q} .

Ασκήσεις

Στις ασκήσεις πιο κάτω, εκτός αν δηλώνεται διαφορετικά, οι τυχαίες μεταβλητές που εμφανίζονται παίρνουν τιμές στο \mathbb{R} .

10.1 Έστω $\mathcal{F}_1, \mathcal{F}_2 \subset \mathcal{F}$ σ -άλγεβρες στον $(\Omega, \mathcal{F}, \mathbf{P})$. Αν για κάθε $A \in \mathcal{F}_1$ ισχύει ότι $\mathbf{P}(A) = 0$ ή $\mathbf{P}(A) = 1$, να δείξετε ότι οι $\mathcal{F}_1, \mathcal{F}_2$ είναι ανεξάρτητες.

10.2 Έστω X, Y όπως στη διατύπωση του Θεωρήματος 10.6. Να δειχθεί ότι τα εξής είναι ισοδύναμα:

- (α) Οι X, Y είναι ανεξάρτητες.
- (β) Οι $f(X), g(Y)$ είναι ανεξάρτητες για κάθε $f : E \rightarrow \mathbb{R}, g : G \rightarrow \mathbb{R}$ μετρήσιμες.

Ανεξαρτησία

10.3 Έστω X, Y ανεξάρτητες τυχαίες μεταβλητές ώστε $\mathbf{P}(X = Y) = 1$. Να δειχθεί ότι υπάρχει $c \in \mathbb{R}$ ώστε $\mathbf{P}(X = c) = 1$, δηλαδή με πιθανότητα 1 η X είναι σταθερή τυχαία μεταβλητή.

10.4 Έστω X, Y ανεξάρτητες ισόνομες τυχαίες μεταβλητές με πεπερασμένη δεύτερη ροπή (δηλαδή $\mathbf{E}(X^2), \mathbf{E}(Y^2) < \infty$) και διασπορά σ^2 . Να δειχθεί ότι $\mathbf{E}\{(X - Y)^2/2\} = \sigma^2$.

10.5 Έστω X, Y τυχαίες μεταβλητές με πεπερασμένη δεύτερη ροπή ώστε οι $X, X - Y$ να είναι ανεξάρτητες αλλά και οι $Y, X - Y$ να είναι ανεξάρτητες. Να δειχθεί ότι υπάρχει $c \in \mathbb{R}$ ώστε $\mathbf{P}(X - Y = c) = 1$.
[Υπόδειξη: $Y = X - (X - Y)$.]

10.6 Έστω X, Y ανεξάρτητες τυχαίες μεταβλητές με $\mathbf{E}|X + Y| < \infty$. Να δειχθεί ότι $\mathbf{E}|X| < \infty$ και $\mathbf{E}|Y| < \infty$.

10.7* Έστω X τυχαία μεταβλητή και $f, g : \mathbb{R} \rightarrow \mathbb{R}$ αύξουσες συναρτήσεις. Υποθέτουμε ότι οι μέσες τιμές $\mathbf{E}\{f(X)g(X)\}, \mathbf{E}\{f(X)\}, \mathbf{E}\{g(X)\}$ ορίζονται και είναι πραγματικοί αριθμοί. Να δειχθεί ότι

$$\mathbf{E}\{f(X)g(X)\} \geq \mathbf{E}\{f(X)\} \mathbf{E}\{g(X)\}.$$

Αυτή είναι ειδική περίπτωση της ανισότητας FKG (Fortuin-Kasteleyn-Ginibre. 1971).

10.8 Έστω X_1, X_2, \dots, X_n ανεξάρτητες τυχαίες μεταβλητές με $X_i \sim \text{Exp}(a_i)$ για κάθε $i = 1, 2, \dots, n$ όπου $a_1, a_2, \dots, a_n \in (0, \infty)$ είναι σταθερές. Ποια η κατανομή της $m := \min\{X_1, X_2, \dots, X_n\}$;

10.9 Έστω $(X_n)_{n \in \mathbb{N}}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών έτσι ώστε $\mathbf{E}(X_1) = 0, \mathbf{E}(X_1^2) = 1, \mathbf{E}(X_1^4) = c < \infty$. Θέτουμε $S_n = \sum_{k=1}^n X_k$ για κάθε $n \geq 1$. Να δείξετε ότι

(α) $\mathbf{E}(S_n^2) = n$.

(β) $\mathbf{E}(S_n^4) = nc + 3n(n - 1)$.

(γ) $\frac{S_n}{n} \xrightarrow{\mathbf{P}} 0$ καθώς $n \rightarrow \infty$.

[Υποδειξη για το (γ): Χρησιμοποιήστε την ανισότητα Markov και το (α).]

10.10 Έστω $(X_n)_{n \geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών καθεμία με κατανομή την ομοιόμορφη στο $(0, 1)$. Για κάθε $n \geq 1$ θεωρούμε τις τυχαίες μεταβλητές

$$m_n = \min\{X_1, X_2, \dots, X_n\},$$

$$M_n = \max\{X_1, X_2, \dots, X_n\}.$$

Να δειχθεί ότι $m_n \rightarrow 0$ και $M_n \rightarrow 1$ κατά πιθανότητα καθώς $n \rightarrow \infty$.

10.11 Για κάθε πίνακα $A \in \mathbb{C}^{n \times n}$, η ορίζουσα του και η permanent του ορίζονται από τους τύπους

$$\det(A) := \sum_{\pi \in S_n} \text{sgn}(\pi) a_{1,\pi(1)} a_{2,\pi(2)} \cdots a_{n,\pi(n)}, \tag{10.7}$$

$$\text{per}(A) := \sum_{\pi \in S_n} a_{1,\pi(1)} a_{2,\pi(2)} \cdots a_{n,\pi(n)}. \tag{10.8}$$

S_n είναι το σύνολο των μεταθέσεων στο $\{1, 2, \dots, n\}$ και $\text{sgn}(\pi)$ είναι το πρόσημο της μετάθεσης π (1 για άρτια μετάθεση και -1 για περιττή).

Έστω $A = (a_{i,j})_{1 \leq i,j \leq n} \in \mathbb{C}^n$ δεδομένος πίνακας. Θεωρούμε ανεξάρτητες τυχαίες μεταβλητές $\{u_{i,j} : 1 \leq i, j \leq n\}$ με τιμές στο \mathbb{R} ώστε καθεμία έχει μέση τιμή 0 και διασπορά 1. Ορίζουμε τον πίνακα $B = (u_{i,j} \sqrt{a_{i,j}})_{1 \leq i,j \leq n}$ όπου $\sqrt{a_{i,j}}$ είναι μία τετραγωνική ρίζα του $a_{i,j}$. Να δειχθεί ότι

$$\mathbf{E}(\det(B^2)) = \text{per}(A).$$

10.12 (α) Έστω X_1, X_2 τυχαίες μεταβλητές ώστε $X_1 \sim \text{Bin}(m, p), X_2 \sim \text{Bin}(n, p)$ όπου $m, n \in \mathbb{N}^+, m < n$ και $p \in [0, 1]$. Να δειχθεί ότι $X_1 \leq_{\text{st}} X_2$. [Δες Άσκηση 7.6 για τον ορισμό της \leq_{st} .]

(β) Έστω X_1, X_2 τυχαίες μεταβλητές ώστε $X_1 \sim \text{Bin}(n, p_1), X_2 \sim \text{Bin}(n, p_2)$ όπου $n \in \mathbb{N}^+, p_1, p_2 \in [0, 1]$ με $p_1 \leq p_2$. Να δειχθεί ότι $X_1 \leq_{\text{st}} X_2$.

Ανεξαρτησία

(γ) Έστω X_1, X_2 τυχαίες μεταβλητές ώστε $X_1 \sim \text{Poisson}(\lambda), X_2 \sim \text{Poisson}(\mu)$ όπου $0 < \lambda < \mu$. Να δειχθεί ότι $X_1 \leq_{\text{st}} X_2$.

[Θεωρούμε δεδομένο ότι το άθροισμα n ανεξάρτητων τυχαίων μεταβλητών $\text{Bernoulli}(p)$ έχει κατανομή $\text{Bin}(n, p)$ και ότι αν οι X, Y είναι ανεξάρτητες και $X \sim \text{Poisson}(a), Y \sim \text{Poisson}(b)$, τότε $X + Y \sim \text{Poisson}(a + b)$].

10.13 Έστω X τυχαία μεταβλητή με τιμές στο $[0, \infty)$ και με $\mu := \mathbf{E}(X) \in (0, \infty)$. Έστω επίσης X^* η μεροληπτική μεγέθους εκδοχή της X (δες Άσκηση 7.7 για τον ορισμό της X^*). Να δειχθεί ότι $X \leq_{\text{st}} X^*$.

10.14 Έστω X, Y ανεξάρτητες τυχαίες μεταβλητές με τιμές στο \mathbb{R} . Να δειχθεί ότι

$$\mathbf{P}(X = Y) = \sum_{x \in \mathbb{R}} \mathbf{P}(X = x) \mathbf{P}(Y = x).$$

Στο άθροισμα στο δεξί μέλος, το πλήθος των μη μηδενικών όρων είναι αριθμησιμο.

11

Τα λήμματα Borel-Cantelli και ο νόμος 0-1 του Kolmogorov

Τα βασικότερα εργαλεία για την απόδειξη θεωρημάτων που αφορούν τη σχεδόν βέβαιη σύγκλιση είναι δύο απλά αποτελέσματα, τα δύο λήμματα Borel-Cantelli, τα οποία θα δούμε σε αυτό το κεφάλαιο. Επίσης θα δούμε τον νόμο 0-1 του Kolmogorov, ο οποίος είναι πολύ χρήσιμος όταν θέλουμε να δείξουμε ότι μια ιδιότητα ισχύει με πιθανότητα 1. Σύμφωνα με αυτόν τον νόμο, αν η ιδιότητα έχει μια συγκεκριμένη μορφή, τότε αναγκαστικά έχει πιθανότητα 0 ή 1. Επομένως για να δείξουμε ότι ισχύει με πιθανότητα 1, αρκεί να δείξουμε ότι ισχύει με θετική πιθανότητα. Και το τελευταίο πολλές φορές είναι σημαντικά ευκολότερο.

11.1 Τα λήμματα Borel-Cantelli

Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας και $(A_n)_{n \geq 1}$ ακολουθία στοιχείων της \mathcal{F} . Υπενθυμίζουμε ότι

$$\limsup_{n \geq 1} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = \{\omega \in \Omega : \omega \text{ ανήκει σε άπειρα } A_n\}.$$

Πρόταση 11.1 (Πρώτο λήμμα Borel-Cantelli). Έστω $(A_n)_{n \geq 1}$ ακολουθία ενδεχομένων στον $(\Omega, \mathcal{F}, \mathbf{P})$. Αν $\sum_{n=1}^{\infty} \mathbf{P}(A_n) < \infty$, τότε

$$\mathbf{P}(\limsup_{n \geq 1} A_n) = 0.$$

Απόδειξη. Θεωρούμε την τυχαία μεταβλητή $X = \sum_{n=1}^{\infty} \mathbf{1}_{A_n}$. Τότε

$$\limsup_{n \geq 1} A_n = \{\omega \in \Omega : \omega \text{ ανήκει σε άπειρα } A_n\} = \{\omega \in \Omega : X(\omega) = \infty\} = \{X = \infty\},$$

και

$$\mathbf{E}(X) = \sum_{n=1}^{\infty} \mathbf{E}(\mathbf{1}_{A_n}) = \sum_{n=1}^{\infty} \mathbf{P}(A_n) < \infty.$$

Επειδή $\mathbf{E}(X) < \infty$ και $X \geq 0$, η Πρόταση 5.14(iii) δίνει ότι $\mathbf{P}(X = \infty) = 0$, δηλαδή $\mathbf{P}(\limsup_{n \geq 1} A_n) = 0$. ■

Παράδειγμα 11.2. Θεωρούμε κάλπη που τη χρονική στιγμή 0 είναι άδεια. Τη χρονική στιγμή n ($n \in \mathbb{N}^+$) προσθέτουμε στην κάλπη ένα σφαιρίδιο που φέρει τον αριθμό n και επιλέγουμε τυχαία ένα από τα n σφαιρίδια της κάλπης (η οποία τότε περιέχει τα σφαιρίδια $1, 2, \dots, n$). Καταγράφουμε τον αριθμό του σφαιριδίου και το επιστρέφουμε στην κάλπη. Για παράδειγμα, μια ακολουθία των αριθμών που εξάγουμε μπορεί να ξεκινάει ως εξής $1, 2, 2, 1, 4, 3, 6, 2, 7, 9, 8, \dots$

Τριάδα λέμε κάθε σύνολο τριών διαδοχικών θέσεων στην ακολουθία που καταλαμβάνονται από τον ίδιο αριθμό. Θα δείξουμε ότι, με πιθανότητα 1, σε αυτή την ακολουθία το πλήθος των

τριάδων είναι πεπερασμένο. Θέτουμε A_n το ενδεχόμενο τις χρονικές στιγμές $n, n+1, n+2$ να επιλεγεί ο ίδιος αριθμός (δηλαδή το σύνολο $\{n, n+1, n+2\}$ να είναι μια τριάδα). Τότε

$$\mathbf{P}(A_n) = \frac{n}{n(n+1)(n+2)} = \frac{1}{(n+1)(n+2)}$$

και $\sum_{n=1}^{\infty} \mathbf{P}(A_n) < \infty$. Το πρώτο λήμμα Borel-Cantelli συνεπάγεται ότι με πιθανότητα 1 συμβαίνουν πεπερασμένο πλήθος από τα A_n .

Η υπόθεση του λήμματος δεν είναι απαραίτητη ώστε να ισχύει το συμπέρασμα. Με άλλα λόγια, το αντίστροφο του λήμματος δεν ισχύει. Ένα αντιπαράδειγμα είναι το εξής. Παίρνουμε λ το μέτρο Lebesgue στον $\Omega := (0, 1)$ (το λ είναι μέτρο πιθανότητας) και $A_n = (0, 1/n)$ για κάθε $n \geq 1$. Τότε $\lambda(A_n) = 1/n$, και η σειρά $\sum_{n=1}^{\infty} \lambda(A_n) = \infty$, ενώ $\limsup_{n \geq 1} A_n = \emptyset$. Κανένας αριθμός δεν ανήκει σε άπειρα από τα A_n .

Λήμμα 11.3. Αν $A_1, A_2, \dots, A_n, n \geq 2$, ανεξάρτητα ενδεχόμενα στον $(\Omega, \mathcal{F}, \mathbf{P})$, τότε και τα $A_1^c, A_2^c, \dots, A_n^c$ είναι ανεξάρτητα.

Απόδειξη. Αυτό είναι συνέπεια του Θεωρήματος 10.10. Επιλέγουμε $I = J = \{1, 2, \dots, n\}$, και $I_j = \{j\}, \mathcal{F}_j = \{A_j\}$ για κάθε $j \in J$. Μπορούμε όμως να το δείξουμε και στοιχειωδώς ξεκινώντας από τον εξής ισχυρισμό.

ΙΣΧΥΡΙΣΜΟΣ: Τα $A_1, A_2, \dots, A_{n-1}, A_n^c$ είναι ανεξάρτητα.

Πράγματι, για $k \leq n$, έστω δείκτες $1 \leq i_1 < i_2 < \dots < i_k \leq n$.

- Αν $i_k < n$, τότε τα $A_{i_1}, A_{i_2}, \dots, A_{i_k}$ είναι ανεξάρτητα από υπόθεση. Επομένως

$$\mathbf{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_{k-1}} \cap A_{i_k}) = \mathbf{P}(A_{i_1}) \mathbf{P}(A_{i_2}) \cdots \mathbf{P}(A_{i_k}) \quad (11.1)$$

- Αν $i_k = n$, τότε

$$\begin{aligned} \mathbf{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_{k-1}} \cap A_n^c) &= \mathbf{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_{k-1}}) - \mathbf{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_{k-1}} \cap A_n) \\ &= \mathbf{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_{k-1}}) - \mathbf{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_{k-1}}) \mathbf{P}(A_n) \\ &= \mathbf{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_{k-1}}) (1 - \mathbf{P}(A_n)) \\ &= \mathbf{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_{k-1}}) \mathbf{P}(A_n^c) \\ &= \mathbf{P}(A_{i_1}) \mathbf{P}(A_{i_2}) \cdots \mathbf{P}(A_n^c). \end{aligned}$$

Στη δεύτερη και στην τελευταία ισότητα χρησιμοποιήσαμε την ανεξαρτησία των A_1, A_2, \dots, A_n .

Άρα και στις δύο περιπτώσεις η πιθανότητα της τομής ισούται με το γινόμενο των πιθανοτήτων, και ο ισχυρισμός αποδείχθηκε.

Χρησιμοποιώντας τον ισχυρισμό, δείχνουμε επαγωγικά ότι τα $A_1^c, A_2^c, \dots, A_n^c$ είναι ανεξάρτητα. ■

Το δεύτερο λήμμα Borel-Cantelli αφορά την περίπτωση που η σειρά $\sum_{n=1}^{\infty} \mathbf{P}(A_n)$ απειρίζεται. Όμως τώρα υποθέτουμε επιπλέον ότι τα $(A_n)_{n \geq 1}$ είναι ανεξάρτητα.

Πρόταση 11.4 (Δεύτερο λήμμα Borel-Cantelli). Έστω $(A_n)_{n \geq 1}$ ακολουθία ανεξάρτητων ενδεχομένων στον $(\Omega, \mathcal{F}, \mathbf{P})$. Αν $\sum_{n=1}^{\infty} \mathbf{P}(A_n) = \infty$, τότε

$$\mathbf{P}(\limsup_{n \geq 1} A_n) = 1.$$

Τα λήμματα Borel-Cantelli και ο νόμος 0-1 του Kolmogorov

Απόδειξη. Θα δείξουμε ότι $\mathbf{P}(\{\limsup_{n \geq 1} A_n\}^c) = 0$. Έχουμε ότι

$$\mathbf{P}(\{\limsup_{n \geq 1} A_n\}^c) = \mathbf{P}(\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k^c) = \lim_{n \rightarrow \infty} \mathbf{P}(\bigcap_{k=n}^{\infty} A_k^c)$$

εφόσον η ακολουθία $(B_n)_{n \geq 1}$ με $B_n = \bigcap_{k=n}^{\infty} A_k^c$ για κάθε $n \geq 1$ είναι αύξουσα. Για δεδομένο $n \geq 1$ έχουμε

$$\begin{aligned} \mathbf{P}(\bigcap_{k=n}^{\infty} A_k^c) &= \lim_{m \rightarrow \infty} \mathbf{P}(\bigcap_{k=n}^m A_k^c) = \lim_{m \rightarrow \infty} \prod_{k=n}^m \mathbf{P}(A_k^c) \\ &= \lim_{m \rightarrow \infty} \prod_{k=n}^m (1 - \mathbf{P}(A_k)) \leq \lim_{m \rightarrow \infty} \prod_{k=n}^m e^{-\mathbf{P}(A_k)} \\ &= \lim_{m \rightarrow \infty} e^{-\sum_{k=n}^m \mathbf{P}(A_k)} = e^{-\sum_{k=n}^{\infty} \mathbf{P}(A_k)} = e^{-\infty} = 0. \end{aligned}$$

Στη δεύτερη ισότητα χρησιμοποιήσαμε το Λήμμα 11.3, ενώ η ανισότητα προκύπτει από την $1 + x \leq e^x$ για κάθε $x \in \mathbb{R}$. Άρα $\mathbf{P}(\{\limsup_{n \geq 1} A_n\}^c) = 0$ και το ζητούμενο αποδείχθηκε. ■

Αν παραλείψουμε την υπόθεση της ανεξαρτησίας από τη διατύπωση του δεύτερου λήμματος Borel-Cantelli, τότε το συμπέρασμά του ενδέχεται να μην ισχύει. Ένα αντιπαράδειγμα είδαμε ακριβώς πριν το Λήμμα 11.3. Σαφώς όμως τα σύνολα $(A_n)_{n \geq 1}$ εκεί δεν είναι ανεξάρτητα.

Παράδειγμα 11.5. (α) Θεωρούμε το πείραμα ρίψης ενός νομίσματος άπειρες (αριθμήσιμες) φορές που φέρνει «κορώνα» (K) με πιθανότητα $p \in (0, 1)$. Θα υπολογίσουμε την πιθανότητα $\mathbf{P}(K \text{ έρχεται άπειρες φορές})$.

Ο χώρος πιθανότητας του πειράματος είναι ο χώρος γινόμενο $(\Omega, \mathcal{F}, \mathbf{P})$ των $(\Omega_n, \mathcal{F}_n, \mathbf{P}_n)_{n \geq 1}$, όπου, για κάθε $n \geq 1$, $\Omega_n = \{K, \Gamma\}$ ($\Gamma =$ το ενδεχόμενο «γράμματα»), $\mathcal{F}_n = \mathcal{P}(\Omega_n)$ και $\mathbf{P}_n = \mathbf{P}^{(p)}$ ($\mathbf{P}^{(p)}$ το μέτρο πιθανότητας με $\mathbf{P}^{(p)}(\{K\}) = p$). Για $n \geq 1$, θεωρούμε το ενδεχόμενο $A_n = \{\text{έρχεται } K \text{ στη } n \text{ ρίψη}\}$ και την τυχαία μεταβλητή $X_n : \Omega \rightarrow \{K, \Gamma\}$ με $X_n(\omega) = \omega_n =$ το αποτέλεσμα της n ρίψης.

Γνωρίζουμε ότι οι $(X_n)_{n \in \mathbb{N}}$ είναι ανεξάρτητες (Πρόταση 10.14) και, εφόσον $A_n = X_n^{-1}(\{K\})$, έχουμε ότι τα $(A_n)_{n \geq 1}$ είναι ανεξάρτητα. Επιπλέον, $\mathbf{P}(A_n) = p$ γιατί το A_n είναι μετρήσιμο ορθογώνιο $\prod_{i \in \mathbb{N}^+} C_i$, με $C_i = \Omega_i$ για κάθε $i \in \mathbb{N}^+ \setminus \{n\}$, $C_n = \{K\}$, και $\mathbf{P}_n(C_n) = p$. Άρα

$$\sum_{n=1}^{\infty} \mathbf{P}(A_n) = \infty.$$

Από το 2ο λήμμα Borel-Cantelli έχουμε $\mathbf{P}(\limsup_{n \geq 1} A_n) = 1$, δηλαδή $\mathbf{P}(K \text{ έρχεται άπειρες φορές}) = 1$.

(β) Στο μοντέλο του Παραδείγματος 11.2, ονομάζουμε ζευγάρι κάθε σύνολο δύο διαδοχικών θέσεων στην ακολουθία που καταλαμβάνονται από τον ίδιο αριθμό. Με πιθανότητα 1, στην ακολουθία που δημιουργούμε το πλήθος των ζευγαριών είναι άπειρο. Για την απόδειξη αυτού, θέτουμε A_n το ενδεχόμενο οι θέσεις $2n - 1, 2n$ να καταλαμβάνονται από τον ίδιο αριθμό. Τότε τα $(A_n)_{n \in \mathbb{N}^+}$ είναι ανεξάρτητα (γιατί αφορούν διαφορετικούς χρόνους στην ακολουθία των εξαγωγών),

$$\mathbf{P}(A_n) = \frac{2n - 1}{(2n - 1) \times 2n} = \frac{1}{2n},$$

και $\sum_{n=1}^{\infty} \mathbf{P}(A_n) = \infty$. Το συμπέρασμα έπεται με εφαρμογή του δεύτερου λήμματος Borel-Cantelli.

Παρατήρηση 11.6. Συνήθως για το $\limsup_{n \geq 1} A_n$ χρησιμοποιούμε τον συμβολισμό

$$\{A_n \text{ συμβαίνει άπειρες φορές}\},$$

και το σκεπτικό του είναι το εξής. Ο χώρος πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$ μοντελοποιεί ένα πείραμα, και μια πραγματοποίηση του πειράματος είναι ένα $\omega \in \Omega$ (στο προηγούμενο παράδειγμα η πραγματοποίηση είναι ένα $\omega \in \{K, \Gamma\}^{\mathbb{N}^+}$). Πολλές φορές παρ' όλ' αυτά, αντιλαμβανόμαστε ότι το πείραμα γίνεται σε πολλά στάδια (π.χ., ρίχνουμε ένα νόμισμα άπειρες φορές, τη μία μετά την άλλη και όχι μονομιάς). Ορίζουμε A_n να είναι ένα σύνολο που αφορά το στάδιο n , και αυτό πραγματοποιείται αν $\omega \in A_n$. Τότε το να συμβεί το A_n άπειρες φορές (δηλαδή για άπειρα n) σημαίνει ακριβώς ότι η πραγματοποίηση ω ανήκει σε άπειρα από τα $\{A_n : n \geq 1\}$. Στο προηγούμενο παράδειγμα πήραμε

$$A_n = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_{n-1} \times \{K\} \times \Omega_{n+1} \times \cdots .$$

Θα χρησιμοποιήσουμε τα λήμματα Borel-Cantelli για να αποδείξουμε σχεδόν βέβαιη σύγκλιση ακολουθιών τυχαίων μεταβλητών. Χαρακτηριστικά είναι τα επόμενα δύο παραδείγματα.

Παράδειγμα 11.7. Έστω $(X_n)_{n \geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με καθεμία να έχει πυκνότητα $f(x) = x^{-2} \mathbf{1}_{x > 1}$. Θα δείξουμε ότι:

(α) Για κάθε $M > 0$ ισχύει

$$\mathbf{P}\left(\overline{\lim}_{n \rightarrow \infty} \frac{X_n}{n} \geq M\right) = 1.$$

(β)

$$\mathbf{P}\left(\overline{\lim}_{n \rightarrow \infty} \frac{X_n}{n} = \infty\right) = 1.$$

Προφανώς το (α) είναι συνέπεια του (β), αλλά θα το χρησιμοποιήσουμε ως πρώτο βήμα για να δείξουμε το (β).

Υπολογίζουμε ότι $\mathbf{P}(X_1 \geq t) = \int_t^\infty f(x) dx = 1/t$ για κάθε $t \geq 1$.

Για το (α), θέτουμε $A_n := \{X_n \geq nM\}$ για κάθε $n \in \mathbb{N}^+$. Για $n \geq 1/M$ ισχύει $\mathbf{P}(A_n) = 1/(nM)$. Τα σύνολα $(A_n)_{n \geq 1}$ είναι ανεξάρτητα μεταξύ τους, αφού οι $(X_n)_{n \geq 1}$ είναι ανεξάρτητες και $A_n = X_n^{-1}([nM, \infty)) \in \sigma(X_n)$, ενώ επίσης ισχύει $\sum_{n=1}^\infty \mathbf{P}(A_n) = \infty$. Το δεύτερο λήμμα Borel-Cantelli συνεπάγεται ότι $\mathbf{P}(\limsup_{n \geq 1} A_n) = 1$. Όμως $\limsup_{n \geq 1} A_n \subset \{\overline{\lim}_{n \rightarrow \infty} X_n/n \geq M\}$, έτσι το (α) δείχτηκε.

Για το (β), θέτουμε $C_M := \{\overline{\lim}_{n \rightarrow \infty} X_n/n \geq M\}$ για κάθε $M > 0$. Τότε $\{\overline{\lim}_{n \rightarrow \infty} X_n/n = \infty\} = \bigcap_{k \in \mathbb{N}^+} C_k$. Επειδή, από το (α), κάθε C_k έχει πιθανότητα 1, έπεται ότι $\mathbf{P}(\{\overline{\lim}_{n \rightarrow \infty} X_n/n = \infty\}) = 1$.

Μια άμεση απόδειξη του (β) προκύπτει αν θεωρήσει κανείς τα σύνολα $B_n := \{X_n \geq n \log n\}$, $n \geq 1$. Αυτά είναι ανεξάρτητα με $\sum_{n=1}^\infty \mathbf{P}(B_n) = \infty$ και στο $\limsup_{n \geq 1} B_n$ ισχύει $X_n/n \geq \log n$ για άπειρα n .

Παράδειγμα 11.8. Έστω $(X_n)_{n \geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με $X_1 \sim \text{Exp}(1)$, δηλαδή με πυκνότητα $f(x) = e^{-x} \mathbf{1}_{x > 0}$. Θα δείξουμε ότι

$$\overline{\lim}_{n \rightarrow \infty} \frac{X_n}{\log n} = 1 \text{ με πιθανότητα } 1.$$

[Μεταφράζοντας τη σημασία του $\overline{\lim}$, αυτό σημαίνει ότι για κάθε $\varepsilon > 0$ ισχύει $X_n/\log n > 1 + \varepsilon$ για πεπερασμένα το πλήθος n , ενώ ισχύει $X_n/\log n > 1 - \varepsilon$ για άπειρα το πλήθος n . Άμεσα βλέπουμε τη συνάφεια των λημμάτων Borel-Cantelli με το ερώτημα.]

Για κάθε $n \geq 1$ και $r > 0$, θέτουμε $A_n^{(r)} = \{X_n \geq r \log n\}$ και

$$C_r := \left\{ \overline{\lim}_{n \rightarrow \infty} \frac{X_n}{\log n} > r \right\}.$$

Τα λήμματα Borel-Cantelli και ο νόμος 0-1 του Kolmogorov

Προφανώς $C_r \subset \limsup_{n \geq 1} A_n^{(r)}$. Επίσης,

$$\mathbf{P}(A_n^{(r)}) = \mathbf{P}(X_n \geq r \log n) = e^{-r \log n} = \frac{1}{n^r}.$$

Έστω $r > 1$. Τότε

$$\sum_{n=1}^{\infty} \mathbf{P}(A_n^{(r)}) < \infty,$$

και από το πρώτο λήμμα Borel-Cantelli έχουμε ότι $\mathbf{P}(\limsup_{n \geq 1} A_n^{(r)}) = 0$, άρα $\mathbf{P}(C_r) = 0$. Και επειδή

$$\left\{ \overline{\lim}_{n \rightarrow \infty} \frac{X_n}{\log n} > 1 \right\} = \cup_{k=1}^{\infty} C_{1+\frac{1}{k}},$$

το σύνολο στο αριστερό μέλος της τελευταίας ισότητας έχει πιθανότητα 0. Επομένως,

$$\mathbf{P}\left(\overline{\lim}_{n \rightarrow \infty} \frac{X_n}{\log n} \leq 1\right) = 1. \quad (11.2)$$

Έστω $r = 1$. Τότε, εφόσον τα $(A_n^{(1)})_{n \geq 1}$ είναι ανεξάρτητα (οι $(X_n)_{n \geq 1}$ είναι ανεξάρτητες) και

$$\sum_{n=1}^{\infty} \mathbf{P}(A_n^{(1)}) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty,$$

από το 2ο λήμμα Borel-Cantelli, έχουμε ότι $\mathbf{P}(\limsup_{n \geq 1} A_n^{(1)}) = 1$. Όμως για $\omega \in \limsup_{n \geq 1} A_n^{(1)}$ ισχύει ότι $X_n(\omega) \geq \log n$ για άπειρα $n \geq 1$, άρα

$$\overline{\lim}_{n \rightarrow \infty} \frac{X_n(\omega)}{\log n} \geq 1,$$

από το οποίο προκύπτει ότι $\mathbf{P}\left(\overline{\lim}_{n \rightarrow \infty} \frac{X_n}{\log n} \geq 1\right) = 1$. Η τελευταία ισότητα μαζί με την (11.2) δίνουν το ζητούμενο. Δηλαδή, στο σύνολο $\left\{ \overline{\lim}_{n \rightarrow \infty} \frac{X_n}{\log n} \leq 1 \right\} \cap \left\{ \overline{\lim}_{n \rightarrow \infty} \frac{X_n}{\log n} \geq 1 \right\}$, που έχει πιθανότητα 1, ισχύει $\overline{\lim}_{n \rightarrow \infty} \frac{X_n}{\log n} = 1$.

Παρατήρηση 11.9. Στα δύο προηγούμενα παραδείγματα, ορίσαμε κάποια σύνολα (τα $A_n, C_M, A_n^{(r)}$, και C_r) και μιλήσαμε για τις πιθανότητες τους. Τυπικά θα έπρεπε προηγουμένως να δείξουμε ότι είναι στοιχεία της \mathcal{F} , δηλαδή είναι μετρήσιμα σύνολα (για παράδειγμα το $C_M \in \mathcal{F}$ γιατί με βάση την Πρόταση 4.7 η $\overline{\lim}_{n \rightarrow \infty} X_n/n$ είναι μετρήσιμη). Δεν το κάναμε, ούτε θα το κάνουμε στο εξής για τα σύνολα που θα ορίζουμε. Όλα θα είναι μετρήσιμα, και η τυπική δικαιολόγηση αφήνεται στον αναγνώστη.

11.2 Ο νόμος 0-1 του Kolmogorov*

Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας, $((E_n, \mathcal{E}_n))_{n \geq 1}$ μετρήσιμοι χώροι και $X_n : \Omega \rightarrow E_n, n \geq 1$, τυχαίες μεταβλητές. Για $n \geq 1$ θέτουμε

$$\mathcal{C}_n := \sigma(\{X_k : k \geq n+1\}),$$

τη σ -άλγεβρα που παράγεται από τις X_{n+1}, X_{n+2}, \dots

Ορισμός 11.10. Η τελική σ-άλγεβρα που παράγεται από τις $(X_n)_{n \geq 1}$ ορίζεται ως

$$\mathcal{C}_\infty = \bigcap_{n=1}^{\infty} \mathcal{C}_n.$$

Τι σημαίνει πρακτικά για ένα ενδεχόμενο A να ανήκει στην \mathcal{C}_∞ ; Σημαίνει ότι για κάθε $n \geq 1$ η πραγματοποίηση ή όχι του A δεν εξαρτάται από την τιμή που παίρνουν οι πρώτες n από τις X_i . Δηλαδή οποιοδήποτε δεδομένο πεπερασμένο πλήθος από τις X_i δεν επηρεάζει την πραγματοποίηση του A . Αυτή η ασαφής περιγραφή θα γίνει ξεκάθαρη στο επόμενο παράδειγμα.

Παράδειγμα 11.11. Έστω $(X_n)_{n \geq 1}$ όπως προηγουμένως με τιμές στο \mathbb{R} . Θεωρούμε τα σύνολα

$$A = \{\omega : \overline{\lim}_{k \rightarrow \infty} X_k(\omega) \geq 1\}, \quad B = \left\{ \omega : \overline{\lim}_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n X_k(\omega) \geq 0 \right\},$$

$$\Gamma = \{\omega : \inf_{n \geq 1} X_n(\omega) \leq 0\}, \quad \Delta = \left\{ \omega : \sum_{n=1}^{\infty} X_n(\omega) \leq 10 \right\}.$$

Τα A, B ανήκουν στην \mathcal{C}_∞ , ενώ τα Γ, Δ δεν ανήκουν σε αυτήν αναγκαστικά.

Πράγματι, όσον αφορά τα A, B , έχουμε ότι για κάθε $m \geq 1$ ισχύει

$$A = \{\omega : \overline{\lim}_{k \rightarrow \infty} X_k(\omega) \geq 1\} = \{\omega : \overline{\lim}_{k \rightarrow \infty} X_{m+1+k}(\omega) \geq 1\} \in \mathcal{C}_m$$

και

$$B = \left\{ \omega : \overline{\lim}_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n X_k(\omega) \geq 0 \right\} = \left\{ \omega : \overline{\lim}_{n \rightarrow \infty} \left(\frac{1}{n} \sum_{k=1}^m X_k(\omega) + \frac{1}{n} \sum_{k=m+1}^n X_k(\omega) \right) \geq 0 \right\}$$

$$= \left\{ \omega : \overline{\lim}_{n \rightarrow \infty} \frac{1}{n} \sum_{k=m+1}^n X_k(\omega) \geq 0 \right\} \in \mathcal{C}_m$$

εφόσον (προσοχή, το άθροισμα δεν εξαρτάται από το n)

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^m X_k(\omega) = 0.$$

Άρα $A, B \in \mathcal{C}_\infty$.

Για τα Γ και Δ αρκεί να παρατηρήσουμε ότι για συγκεκριμένες επιλογές των τυχαίων μεταβλητών $(X_n)_{n \geq 1}$, τα σύνολα Γ και Δ εξαρτώνται από την τιμή του X_1 . Για παράδειγμα, στο Γ , αν οι $(X_n)_{n \geq 1}$ είναι ανεξάρτητες και $X_1 \sim N(0, 1)$, $X_k \sim U(1, 3)$ (ομοιόμορφη στο $(1, 3)$) για κάθε $k \geq 2$, τότε $\Gamma = \{X_1 \leq 0\}$, το οποίο έχει πιθανότητα $1/2$. Αν υποθέσουμε ότι $\Gamma \in \mathcal{C}_\infty$, τότε $\Gamma \in \sigma(\{X_k : k \geq 2\})$, και αυτή η σ-άλγεβρα είναι ανεξάρτητη από την $\sigma(X_1)$, η οποία περιέχει το Γ . Άρα το Γ είναι ανεξάρτητο από το εαυτό του, δηλαδή $\mathbf{P}(\Gamma \cap \Gamma) = \mathbf{P}(\Gamma) \mathbf{P}(\Gamma)$, το οποίο δίνει ότι $\mathbf{P}(\Gamma) \in \{0, 1\}$. Άτοπο.

Πριν κάνουμε τις παραπάνω αποδείξεις για τα A, B , βλέπουμε ότι για δεδομένο ω (δηλαδή για μία πραγματοποίηση του πειράματος), το αν ισχύει $\omega \in A$, δηλαδή το αν το A πραγματοποιήθηκε, δεν εξαρτάται από τις πρώτες τιμές της $(X_n(\omega))_{n \geq 1}$. Για το A , η τιμή του $\overline{\lim}$ μένει η ίδια αν αλλάξουμε, π.χ., τους πρώτους 1000 όρους της ακολουθίας. Το ίδιο συμβαίνει και με το B . Αυτή η παρατήρηση μας πείθει ότι $A, B \in \mathcal{C}_\infty$ και τη χρησιμοποιούμε στην τυπική απόδειξη.

Το βασικό αποτέλεσμα αυτής της παραγράφου αφορά την τελική σ-άλγεβρα ακολουθίας ανεξάρτητων τυχαίων μεταβλητών.

Θεώρημα 11.12 (Νόμος 0-1 του Kolmogorov). Έστω $(X_n)_{n \geq 1}$ ανεξάρτητες τυχαίες μεταβλητές στον $(\Omega, \mathcal{F}, \mathbf{P})$ και \mathcal{C}_∞ η τελική σ-άλγεβρά τους. Αν $C \in \mathcal{C}_\infty$, τότε $\mathbf{P}(C) = 0$ ή 1 .

Η απόδειξη του θεωρήματος δίνεται στο Παράρτημα Β'.

Το θεώρημα χρησιμοποιείται συνήθως για να δείξουμε ότι ένα γεγονός έχει πιθανότητα 1. Δείχνουμε ότι ανήκει στην τελική σ -άλγεβρα και έχει θετική πιθανότητα. Μια τέτοια χρήση γίνεται στην Άσκηση 11.16. Το να δείξει κανείς ότι το γεγονός του ερωτήματος (β) της άσκησης έχει θετική πιθανότητα είναι πολύ απλό, ενώ το να δείξει ότι έχει πιθανότητα 1 είναι αρκετά περίπλοκο αν δεν χρησιμοποιήσουμε τον νόμο 0-1 του Kolmogorov.

Άμεση συνέπεια του νόμου 0-1 είναι το ακόλουθο πόρισμα.

Πόρισμα 11.13. Έστω $(X_n)_{n \in \mathbb{N}}$ ανεξάρτητες τυχαίες μεταβλητές στον $(\Omega, \mathcal{F}, \mathbf{P})$ και $X : \Omega \rightarrow [-\infty, \infty]$ τυχαία μεταβλητή \mathcal{C}_∞ -μετρήσιμη. Τότε η X είναι σταθερή με πιθανότητα 1.

Απόδειξη. Εφόσον η X είναι \mathcal{C}_∞ -μετρήσιμη, τα σύνολα $\{X = -\infty\}, \{X = \infty\}$ είναι στοιχεία της \mathcal{C}_∞ , και επομένως έχουν πιθανότητα 0 ή 1. Αν κάποιο από αυτά έχει πιθανότητα 1, δείχθηκε το ζητούμενο. Διαφορετικά, έχουμε ότι η X παίρνει τιμές στο \mathbb{R} . Σε αυτή την περίπτωση, για τη συνάρτηση κατανομής της, F , το Θεώρημα 11.12 δίνει ότι για κάθε $x \in \mathbb{R}$ ισχύει

$$F(x) = \mathbf{P}(X \leq x) = \mathbf{P}(X^{-1}((-\infty, x])) \in \{0, 1\}. \quad (11.3)$$

Ξέρουμε όμως ότι η F είναι αύξουσα, δεξιά συνεχής και $F(-\infty) = 0, F(\infty) = 1$. Αυτές οι ιδιότητες μαζί με την (11.3) συνεπάγονται ότι υπάρχει $c \in \mathbb{R}$ τέτοιο ώστε

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

Συνεπώς $\mathbf{P}(X = c) = F(c) - F(c-) = 1$, δηλαδή η X ισούται με τη σταθερά c με πιθανότητα 1. ■

Παρατήρηση 11.14. Στην απόδειξη του Πορίσματος 11.13, από το ότι η X είναι \mathcal{C}_∞ -μετρήσιμη χρησιμοποιήσαμε μόνο ότι όλα τα σύνολα της \mathcal{C}_∞ έχουν πιθανότητα 0 ή 1. Έτσι, το ίδιο επιχείρημα δίνει ότι αν η X είναι τυχαία μεταβλητή στον $(\Omega, \mathcal{F}, \mathbf{P})$ με τιμές στο $[-\infty, \infty]$ και $\mathcal{A} \subset \mathcal{F}$ σ -άλγεβρα τέτοια ώστε η X να είναι $\mathcal{A}/\mathcal{B}([-\infty, \infty])$ -μετρήσιμη και $\mathbf{P}(A) \in \{0, 1\}$ για κάθε $A \in \mathcal{A}$, τότε η X είναι σταθερή με πιθανότητα 1.

Επιστρέφουμε στο Παράδειγμα 11.8. Εκεί η $Z := \overline{\lim}_{n \rightarrow \infty} (X_n / \log n)$ είναι μια \mathcal{C}_∞ -μετρήσιμη τυχαία μεταβλητή με τιμές στο $[0, \infty]$ και οι $(X_n)_{n \geq 1}$ είναι ανεξάρτητες. Το Πόρισμα 11.13 εφαρμόζεται. Άρα εκ των προτέρων ξέρουμε ότι η Z είναι σταθερή με πιθανότητα 1.

Άσκήσεις

11.1* Έστω $(A_n)_{n \geq 1}$ ακολουθία ανεξάρτητων ενδεχομένων με $\mathbf{P}(A_n) < 1$ για κάθε $n \geq 1$ και $\mathbf{P}(\cup_{n=1}^{\infty} A_n) = 1$. Να δείξετε ότι $\sum_{n=1}^{\infty} \mathbf{P}(A_n) = \infty$.

11.2 Έστω $(X_n)_{n \geq 1}$ ακολουθία τυχαίων μεταβλητών με $\mathbf{P}(X_n \neq 0) = \frac{1}{n^2}$ για κάθε $n \geq 1$. Να δείξετε ότι με πιθανότητα 1 για κάθε $\omega \in \Omega$, υπάρχει $n(\omega) \in \mathbb{N}$ ώστε $X_n(\omega) = 0$ για κάθε $n \geq n(\omega)$. (Συνεπώς $X_n \rightarrow 0$ με πιθανότητα 1.)

11.3 Στην Άσκηση 10.10, να δειχθεί ότι επίσης $m_n \rightarrow 0$ και $M_n \rightarrow 1$ σχεδόν βέβαια καθώς $n \rightarrow \infty$.

11.4 Έστω $(X_n)_{n \geq 1}$ ακολουθία ανεξάρτητων τυχαίων μεταβλητών ώστε

$$\mathbf{P}(X_n = 1) = \frac{1}{n}, \mathbf{P}(X_n = 0) = 1 - \frac{1}{n}.$$

Τα λήμματα Borel-Cantelli και ο νόμος 0-1 του Kolmogorov

Να δειχθεί ότι

(α) $X_n \rightarrow 0$ κατά πιθανότητα καθώς $n \rightarrow \infty$,

(β) αλλά δεν ισχύει $X_n \rightarrow 0$ σχεδόν βέβαια. Μάλιστα $\mathbf{P}(\lim_{n \rightarrow \infty} X_n = 0) = 0$.

11.5 Έστω $(X_n)_{n \geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με τιμές στο $[0, \infty)$ ώστε $\mathbf{P}(X_1 > 0) > 0$. Να δειχθεί ότι με πιθανότητα 1 ισχύει

$$\sum_{n=1}^{\infty} X_n = \infty.$$

11.6 Έστω $(X_n)_{n \geq 1}$ ακολουθία τυχαίων μεταβλητών με τιμές στο \mathbb{R} . Να δείξετε ότι υπάρχει ακολουθία $(a_n)_{n \geq 1}$ πραγματικών θετικών αριθμών τέτοια ώστε $\mathbf{P}\left(\lim_{n \rightarrow \infty} \frac{X_n}{a_n} = 0\right) = 1$.

11.7 Έστω $(X_n)_{n \geq 1}$ ακολουθία ανεξάρτητων τυχαίων μεταβλητών με τιμές στο \mathbb{R} . Να δείξετε ότι για την τυχαία μεταβλητή $X^* = \sup_{n \geq 1} X_n$ ισχύει ότι $\mathbf{P}(X^* < \infty) = 1$ αν και μόνο αν υπάρχει $M \in \mathbb{R}$ έτσι ώστε $\sum_{n=1}^{\infty} \mathbf{P}(X_n > M) < \infty$.

11.8* Έστω $(X_n)_{n \geq 1}$ ακολουθία ανεξάρτητων τυχαίων μεταβλητών με $X_n \sim N(0, 1)$ για κάθε $n \geq 1$. Να δείξετε ότι με πιθανότητα 1 ισχύει

$$\overline{\lim}_{n \rightarrow \infty} \frac{X_n}{\sqrt{2 \log n}} = 1. \quad (11.4)$$

11.9 Έστω $(X_n)_{n \geq 1}$ ακολουθία ανεξάρτητων τυχαίων μεταβλητών με $X_n \sim \text{Exp}(1)$ για κάθε $n \geq 1$. Αν $M_n := \max\{X_1, X_2, \dots, X_n\}$ για κάθε $n \geq 1$, να δείξετε ότι με πιθανότητα 1 ισχύει

$$\lim_{n \rightarrow \infty} \frac{M_n}{\log n} = 1. \quad (11.5)$$

11.10 Έστω $(X_n)_{n \geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με $\mathbf{P}(X_1 = K) = \mathbf{P}(X_1 = \Gamma) = 1/2$. Οι $(X_n)_{n \geq 1}$ καταγράφουν τα αποτελέσματα ακολουθίας ρίψεων ενός αμερόληπτου νομίσματος. Για κάθε $n \geq 1$ θέτουμε

$$L_n := \max\{m \geq 1 : X_n = X_{n+1} = \dots = X_{n+m-1}\}.$$

Για παράδειγμα, αν οι ρίψεις δώσουν το αποτέλεσμα $(K, K, \Gamma, \Gamma, K, \Gamma, \Gamma, \Gamma, K, \dots)$, τότε $L_1(\omega) = 2$, $L_5(\omega) = 1$, και $L_6(\omega) = 3$. Δείξτε ότι

(α) $\overline{\lim}_{n \rightarrow \infty} \frac{L_n}{\log_2 n} \leq 1$ με πιθανότητα 1.

(β) $\underline{\lim}_{n \rightarrow \infty} L_n = 1$ με πιθανότητα 1.

11.11* Έστω $(X_n)_{n \geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών. Να δείξετε ότι τα ακόλουθα είναι ισοδύναμα:

(α) $\lim_{n \rightarrow \infty} \frac{X_n}{n} = 0$ με πιθανότητα 1.

(β) $\mathbf{E}|X_1| < \infty$.

11.12 Έστω $(X_n)_{n \geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με τιμές στο $[0, \infty)$, και $c > 1$. Να δείξετε ότι τα ακόλουθα είναι ισοδύναμα:

(α) $\mathbf{E}\{(\log X_1)^+\} < \infty$.

(β) $\sum_{n=1}^{\infty} X_n/c^n < \infty$ με πιθανότητα 1.

11.13* Έστω $(X_n)_{n \geq 1}$ ακολουθία ισόνομων και ανεξάρτητων τυχαίων μεταβλητών με τιμές στο \mathbb{R} ώστε η κατανομή της X_1 να μην είναι συγκεντρωμένη σε ένα σημείο (δηλαδή δεν υπάρχει $c \in \mathbb{R}$ με $\mathbf{P}(X_1 = c) = 1$). Να δειχθεί ότι

$$\mathbf{P}(\lim_{n \rightarrow \infty} X_n \text{ υπάρχει}) = 0.$$

11.14 Έστω $(X_n)_{n \geq 1}$ θετικές τυχαίες μεταβλητές. Να δειχθεί ότι με πιθανότητα 1 ισχύει

$$\overline{\lim} \frac{1}{n} \log X_n \leq \overline{\lim} \frac{1}{n} \log \mathbf{E} X_n. \quad (11.6)$$

11.15 Έστω $(X_i)_{i \geq 1}$ ακολουθία τυχαίων μεταβλητών με τιμές στο \mathbb{R} και $C_\infty := \bigcap_{n=1}^{\infty} \sigma(X_n, X_{n+1}, \dots)$ η τελική σ-άλγεβρα.

(α) Υποθέτουμε ότι οι X_i παίρνουν μη αρνητικές τιμές. Ποιες από τις παρακάτω τυχαίες μεταβλητές είναι C_∞ μετρήσιμες;

$$\begin{aligned} \text{(i)} \quad \overline{\lim}_{n \rightarrow \infty} \frac{X_n}{n}, & \quad \text{(ii)} \quad \overline{\lim}_{n \rightarrow \infty} \frac{X_1 + X_2 + \dots + X_n}{n}, & \quad \text{(iii)} \quad \overline{\lim}_{n \rightarrow \infty} n(X_1 + X_2 + \dots + X_n), \\ \text{(iv)} \quad \sum_{k=1}^{\infty} \frac{X_k}{2^k}, & \quad \text{(v)} \quad \overline{\lim}_{n \rightarrow \infty} (X_n + X_{n+1}). \end{aligned}$$

(β) Ποια από τα παρακάτω σύνολα είναι στοιχεία της C_∞ ;

$$\begin{aligned} \text{(i)} \quad \left\{ \sum_{n=1}^{\infty} |X_n| < \infty \right\}, & \quad \text{(ii)} \quad \{X_1 + X_2 + \dots + X_n = 0 \text{ για άπειρα } n\}, \\ \text{(iii)} \quad \left\{ \overline{\lim}_{n \rightarrow \infty} n|X_1 + X_2 + \dots + X_n| \leq 1 \right\}, & \quad \text{(iv)} \quad \left\{ \sum_{n=1}^{\infty} X_n \text{ συγκλίνει σε πραγματικό άριθμό} \right\}, \\ \text{(v)} \quad \left\{ \sum_{k=n}^{2n} X_k > 0 \text{ για άπειρα } n \right\}. \end{aligned}$$

[Σχόλιο: Καταρχάς, τα ερωτήματα να απαντηθούν διαισθητικά. Έπειτα, σχετικά με το (β), για τα σύνολα τα οποία είναι στοιχεία της C_∞ να αποδειχθεί αυτό τυπικά. Για τα υπόλοιπα, να μην αποδειχθεί τίποτε. Για εκείνα δεν ισχυριζόμαστε ότι πάντοτε δεν είναι στοιχεία της C_∞ . Εξαρτάται από τη συγκεκριμένη επιλογή της ακολουθίας $(X_i)_{i \geq 1}$. Παρόμοιο σχόλιο ισχύει για το μέρος (α) της άσκησης.]

11.16 Έστω $(X_n)_{n \geq 1}$ ακολουθία ανεξάρτητων τυχαίων μεταβλητών, καθεμία με κατανομή την τυπική κανονική $N(0, 1)$. Για κάθε $n \geq 1$ θέτουμε $S_n := X_1 + X_2 + \dots + X_n$.

(α) Για κάθε $A > 0$ και $n \geq 1$ να δειχθεί ότι

$$\mathbf{P} \left(\frac{S_n}{\sqrt{n}} \geq A \right) = 1 - \Phi(A) > 0,$$

όπου Φ είναι η συνάρτηση κατανομής της $N(0, 1)$. [Εδώ μπορείτε να χρησιμοποιήσετε πράγματα για κανονικές τυχαίες μεταβλητές και αθροίσματα τους από τις στοιχειώδεις πιθανότητες.]

(β) Για κάθε $A > 0$, με πιθανότητα 1 ισχύει

$$\overline{\lim}_{n \rightarrow \infty} \frac{S_n}{\sqrt{n}} \geq A.$$

(γ) Με πιθανότητα 1 ισχύει

$$\overline{\lim}_{n \rightarrow \infty} \frac{S_n}{\sqrt{n}} = \infty.$$

11.17 Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας και $\{A_i : i \in I\}$ στοιχεία της \mathcal{F} .

(α) Θεωρούμε τις τυχαίες μεταβλητές $X_i := \mathbf{1}_{A_i}, i \in I$. Να δειχθεί ότι οι $\{X_i : i \in I\}$ είναι ανεξάρτητες αν και μόνο αν τα $\{A_i : i \in I\}$ είναι ανεξάρτητα.

Τα λήμματα Borel-Cantelli και ο νόμος 0-1 του Kolmogorov

(β) Υποθέτουμε ότι $I = \mathbb{N}^+$. Αν τα $\{A_n : n \in \mathbb{N}^+\}$ είναι ανεξάρτητα, τότε τα σύνολα $\liminf_{n \geq 1} A_n$, $\limsup_{n \geq 1} A_n$ έχουν πιθανότητα 0 ή 1.

11.18 Έστω $(X_n)_{n \geq 1}$ ακολουθία ανεξάρτητων τυχαίων μεταβλητών με τιμές στο \mathbb{R} . Θεωρούμε τη δυναμοσειρά

$$f(z) = \sum_{n=0}^{\infty} X_n z^n.$$

(α) Ναδειχθεί ότι η ακτίνα σύγκλισης R της f είναι μετρήσιμη ως προς την τελική σ-άλγεβρα των $(X_n)_{n \geq 1}$ και άρα είναι σταθερή με πιθανότητα 1.

(β) Αν υποθέσουμε επιπλέον ότι καθεμία από τις $(X_n)_{n \geq 1}$ έχει κατανομή $N(0, 1)$, τότε με πιθανότητα 1 ισχύει ότι $R = 1$.

11.19 Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας, και $\{A_n : n \geq 1\}$ στοιχεία της \mathcal{F} τα οποία είναι ανα δύο ανεξάρτητα (δηλαδή $\mathbf{P}(A_i \cap A_j) = \mathbf{P}(A_i) \mathbf{P}(A_j)$ για κάθε $i, j \geq 1, i \neq j$). Θέτουμε

$$S_n := \sum_{i=1}^n \mathbf{1}_{A_i}$$

και

$$s_n := \mathbf{E}(S_n) = \sum_{i=1}^n \mathbf{P}(A_i).$$

(α) Ναδειχθεί ότι

$$\mathbf{E}(S_n^2) = s_n + s_n^2 - \sum_{i=1}^n \mathbf{P}(A_i)^2 \leq s_n + s_n^2.$$

(β) Για κάθε $\varepsilon > 0$, ναδειχθεί ότι

$$\mathbf{P}(S_n \geq \varepsilon s_n) \geq (1 - \varepsilon)^2 \frac{s_n^2}{s_n^2 + s_n}.$$

[Υπόδειξη: Άσκηση 5.6.]

(γ)* Αν επιπλέον ισχύει ότι $\sum_{i=1}^{\infty} \mathbf{P}(A_i) = \infty$, ναδειχθεί ότι $\mathbf{P}(\limsup_{n \geq 1} A_n) = 1$.

[Η άσκηση αυτή γενικεύει το 2ο λήμμα Borel-Cantelli κατά το ότι υποθέτουμε τα $\{A_n : n \geq 1\}$ ανά δύο ανεξάρτητα και όχι απαραίτητα πλήρως ανεξάρτητα.]

11.20 Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας και $\{A_n : n \geq 1\}$ στοιχεία της \mathcal{F} για τα οποία ισχύει $\sum_{i=1}^{\infty} \mathbf{P}(A_i) = \infty$ και υπάρχει $C \in (0, \infty)$ ώστε

$$\mathbf{P}(A_i \cap A_j) \leq C \mathbf{P}(A_i) \mathbf{P}(A_j)$$

για κάθε $i, j \geq 1$. Ναδειχθεί ότι

$$\mathbf{P}(\limsup_{n \geq 1} A_n) \geq 1/C > 0.$$

12

Ο Ισχυρός Νόμος των Μεγάλων Αριθμών

Στο κεφάλαιο αυτό παρουσιάζουμε ένα από τα σημαντικότερα αποτελέσματα της Θεωρίας Πιθανοτήτων, τον ισχυρό νόμο των μεγάλων αριθμών. Η διατύπωση που θα αποδείξουμε δεν είναι η ισχυρότερη μορφή του νόμου, όμως η απόδειξή της είναι ευκολότερη τεχνικά και διατηρεί αρκετά από τα στοιχεία της απόδειξης της ισχυρής μορφής.

12.1 Το θεώρημα

Θεώρημα 12.1 (Ισχυρός Νόμος των Μεγάλων Αριθμών). Έστω $(X_n)_{n \geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές στον $(\Omega, \mathcal{F}, \mathbf{P})$ με τιμές στο \mathbb{R} έτσι ώστε $\mathbf{E}(X_1^2) < \infty$. Θέτουμε $\mu = \mathbf{E}(X_1)$ και $S_n = \sum_{k=1}^n X_k$ για κάθε $n \geq 1$. Τότε

$$\lim_{n \rightarrow \infty} \frac{S_n}{n} = \mu \text{ με πιθανότητα } 1.$$

Απόδειξη. Έστω $\sigma^2 = \text{Var}(X_1)$.

Πρώτα θα αποδείξουμε το ζητούμενο για $(X_n)_{n \in \mathbb{N}}$ με τιμές στο $[0, \infty]$. Για $n \geq 1$, θέτουμε $Y_n = \frac{S_n}{n} - \mu$. Τότε

$$\mathbf{E}(Y_n) = \frac{1}{n} \mathbf{E}(S_n) - \mu = \frac{1}{n} n \mathbf{E}(X_1) - \mu = 0,$$

και

$$\mathbf{E}(Y_n^2) = \text{Var}(Y_n^2) = \text{Var}\left(\frac{S_n}{n}\right) = \frac{1}{n^2} \text{Var}(S_n).$$

Όμως $\text{Var}(S_n) = n \text{Var}(X_1)$ εφόσον οι $(X_n)_{n \geq 1}$ είναι ανεξάρτητες. Επομένως,

$$\mathbf{E}(Y_n^2) = \frac{\sigma^2}{n}.$$

Έτσι, για την υπακολουθία $(Y_n^2)_{n \geq 1}$, έχουμε ότι

$$\mathbf{E}\left\{\sum_{n=1}^{\infty} Y_n^2\right\} = \sum_{n=1}^{\infty} \frac{\sigma^2}{n^2} < \infty.$$

Άρα, με πιθανότητα 1, $\sum_{n=1}^{\infty} Y_n^2 < \infty$, συνεπώς $\lim_{n \rightarrow \infty} Y_n^2 = 0$.

Τώρα, από το $\lim_{n \rightarrow \infty} Y_n^2 = 0$, θέλουμε να περάσουμε στο $\lim_{n \rightarrow \infty} Y_n = 0$.

Έστω $k \geq 1$. Θέτουμε $n(k) = [\sqrt{k}]$. Τότε

$$\frac{S_{n(k)^2}}{(n(k)+1)^2} \leq \frac{S_k}{k} \leq \frac{S_{(n(k)+1)^2}}{n(k)^2}$$

εφόσον $n(k) \leq \sqrt{k} \leq n(k) + 1$ και η S_n είναι άθροισμα θετικών όρων. Όμως, για $k \rightarrow \infty$, με πιθανότητα 1,

$$\frac{S_{n(k)^2}}{(n(k)+1)^2} = \frac{S_{n(k)^2}}{n(k)^2} \left(\frac{n(k)}{n(k)+1}\right)^2 \rightarrow \mu$$

και

$$\frac{S_{(n(k)+1)^2}}{n(k)^2} = \frac{S_{(n(k)+1)^2}}{(n(k)+1)^2} \left(\frac{n(k)+1}{n(k)} \right)^2 \rightarrow \mu.$$

Άρα, με πιθανότητα 1, έχουμε ότι $\lim_{k \rightarrow \infty} \frac{S_k}{k} = \mu$.

Στην περίπτωση που οι $(X_n)_{n \geq 1}$ παίρνουν τιμές στο $[-\infty, \infty]$, έχουμε

$$\frac{S_n}{n} = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{X_1^+ + X_2^+ + \dots + X_n^+}{n} - \frac{X_1^- + X_2^- + \dots + X_n^-}{n}.$$

Οι $(X_n^+)_{n \in \mathbb{N}}$, όπως και οι $(X_n^-)_{n \in \mathbb{N}}$, είναι ανεξάρτητες και ισόνομες και $\mathbf{E}((X_1^+)^2) < \infty$, $\mathbf{E}((X_1^-)^2) < \infty$ εφόσον $(X_1^+)^2 \leq X_1^2$, $(X_1^-)^2 \leq X_1^2$. Από τα προηγούμενα, έχουμε ότι τα σύνολα

$$A := \left\{ \omega \in \Omega : \lim_{n \rightarrow \infty} \frac{X_1^+(\omega) + X_2^+(\omega) + \dots + X_n^+(\omega)}{n} = \mathbf{E}(X_1^+) \right\}$$

$$B := \left\{ \omega \in \Omega : \lim_{n \rightarrow \infty} \frac{X_1^-(\omega) + X_2^-(\omega) + \dots + X_n^-(\omega)}{n} = \mathbf{E}(X_1^-) \right\}$$

έχουν πιθανότητα 1. Άρα στο σύνολο $A \cap B$, που έχει πιθανότητα 1, ισχύει

$$\lim_{n \rightarrow \infty} \frac{S_n}{n} = \mathbf{E}(X_1^+) - \mathbf{E}(X_1^-) = \mu. \quad \blacksquare$$

Το συμπέρασμα του θεωρήματος ισχύει και αν αντί της $\mathbf{E}(X_1^2) < \infty$ υποθέσουμε ότι $\mathbf{E}|X_1| < \infty$, δηλαδή κάτι λιγότερο. Αυτή είναι η γενική μορφή του νόμου των μεγάλων αριθμών και στο εξής θα τον θεωρούμε δεδομένο με αυτή, την ισχυρότερη μορφή.

Στην Άσκηση 12.3 διατυπώνεται ένα είδος αντίστροφου του νόμου των μεγάλων αριθμών. Δηλαδή, αν ο μέσος όρος ανεξάρτητων και ισόνομων τυχαίων μεταβλητών συγκλίνει με πιθανότητα 1 σε πεπερασμένο αριθμό, τότε αναγκαστικά $\mathbf{E}|X_1| < \infty$. Άρα η υπόθεση $\mathbf{E}|X_1| < \infty$ στο θεώρημα είναι απαραίτητη για την ισχύ του συμπεράσματος.

Επίσης, στην Άσκηση 12.2 δείχνουμε ότι το συμπέρασμα του νόμου των μεγάλων αριθμών ισχύει ακόμα και όταν η μέση τιμή $\mu := \mathbf{E}(X_1)$ είναι $-\infty$ ή ∞ . Άρα το συμπέρασμα ισχύει πάντοτε όταν η $\mathbf{E}(X_1)$ μπορεί να οριστεί.

Όταν η $\mathbf{E}(X_1)$ δεν μπορεί να οριστεί, δηλαδή όταν $\mathbf{E}(X_1^+) = \mathbf{E}(X_1^-) = \infty$, τότε ισχύει ακριβώς ένα από τα εξής

- (i) $\lim_{n \rightarrow \infty} \frac{S_n}{n} = -\infty$ με πιθανότητα 1.
- (ii) $\lim_{n \rightarrow \infty} \frac{S_n}{n} = \infty$ με πιθανότητα 1.
- (iii) $\underline{\lim}_{n \rightarrow \infty} \frac{S_n}{n} = -\infty$, $\overline{\lim}_{n \rightarrow \infty} \frac{S_n}{n} = \infty$ με πιθανότητα 1.

Ποιο από τα σενάρια συμβαίνει εξαρτάται από την κατανομή της X_1 και υπάρχει μάλιστα κριτήριο που το καθορίζει αλλά δεν θα το διατυπώσουμε [δες Erickson (1973)].

Σύμβαση: Αν $(X_n)_{n \geq 1}$ είναι ακολουθία τυχαίων μεταβλητών με τιμές στο \mathbb{R} , στο εξής θα συμβολίζουμε με S_n το n -οστό μερικό άθροισμά τους.

Πόρισμα 12.2. (Ο Ασθενής Νόμος των Μεγάλων Αριθμών) Έστω $(X_n)_{n \geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με τιμές στο \mathbb{R} έτσι ώστε $\mu := \mathbf{E}(X_1) \in \mathbb{R}$. Τότε

$$\lim_{n \rightarrow \infty} \frac{S_n}{n} = \mu \text{ κατά πιθανότητα.}$$

Απόδειξη. Έπεται από τον ισχυρό νόμο των μεγάλων αριθμών (ισχυρή μορφή) και το ότι η σχεδόν βέβαιη σύγκλιση συνεπάγεται την κατά πιθανότητα [Θεώρημα 8.2(ii)]. ■

Παρατήρηση 12.3 (Η μέθοδος Monte Carlo). Ο νόμος των μεγάλων αριθμών δικαιολογεί την εξής μέθοδο για τον προσεγγιστικό υπολογισμό μέσης τιμής μιας δεδομένης τυχαίας μεταβλητής X . Παράγουμε με κάποιο τρόπο έναν μεγάλο αριθμό ανεξάρτητων πραγματοποιήσεων της τυχαίας μεταβλητής X , έστω x_1, x_2, \dots, x_n , και υπολογίζουμε τον μέσο τους όρο $(x_1 + x_2 + \dots + x_n)/n$. Αυτός ο μέσος όρος είναι η ζητούμενη προσέγγιση.

Με αυτή τη μέθοδο μπορούμε να βρούμε προσεγγίσεις για οποιαδήποτε ποσότητα μπορεί να παρασταθεί ως μέση τιμή κάποιας τυχαίας μεταβλητής. Η ποσότητα ενδέχεται να εμφανίζεται φυσιολογικά ως μέση τιμή σε κάποιο πρόβλημα πιθανοτήτων και να μην μπορεί να υπολογιστεί με κλειστό τύπο. Επίσης μπορεί να μην έχει καμία σχέση με πιθανότητες. Για παράδειγμα, για τον αριθμό π έχουμε την αναπαράσταση $\pi = 4 \mathbf{E}(\mathbf{1}_{U^2+V^2 < 1})$ όπου οι U, V είναι ανεξάρτητες ομοιόμορφες τυχαίες μεταβλητές στο $(-1, 1)$. Η τυχαία μεταβλητή (U, V) είναι ομοιόμορφη στο τετράγωνο $(-1, 1) \times (-1, 1)$ και η μέθοδος Monte Carlo για την προσέγγιση της $\mathbf{E}(\mathbf{1}_{U^2+V^2 < 1})$ παίρνει πολλές πραγματοποιήσεις της (U, V) και υπολογίζει το ποσοστό αυτών που πέφτουν μέσα στον μοναδιαίο δίσκο.

12.2 Δύο εφαρμογές*

Ανανεωτική θεωρία. Έστω $(X_n)_{n \geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με τιμές στο $[0, \infty)$. Θέτουμε

$$\begin{aligned} T_0 &:= 0, \\ T_n &:= X_1 + X_2 + \dots + X_n \end{aligned}$$

για κάθε $n \geq 1$. Η ερμηνεία που σκεφτόμαστε για τις ακολουθίες $(X_n)_{n \geq 1}, (T_n)_{n \geq 1}$ είναι η εξής. Για τον φωτισμό ενός δωματίου, έχουμε άπειρο πλήθος λαμπών, κάθε μία από τις οποίες έχει τυχαίο χρόνο ζωής. X_n είναι ο χρόνος ζωής της n λάμπας. Τη χρονική στιγμή 0 τοποθετούμε τη λάμπα 1. Μόλις αυτή καεί, τοποθετούμε τη λάμπα 2 και συνεχίζουμε όμοια. T_n είναι ο χρόνος που καίγεται η λάμπα n . Τώρα για $t \geq 0$, θέτουμε

$$N_t := \sup\{n \geq 0 : T_n \leq t\}.$$

Η τυχαία μεταβλητή N_t μετράει πόσες λάμπες κήκαν κατά το χρονικό διάστημα $[0, t]$. Έστω ότι $\mu := \mathbf{E}(X_1)$. Αφού κάθε λάμπα ζει περίπου χρόνο μ , για το χρονικό διάστημα $[0, t]$ αναμένουμε να χρειάζονται t/μ λάμπες περίπου. Δηλαδή $N_t \approx t/\mu$.

Θεώρημα 12.4. Με πιθανότητα 1 ισχύει

$$\lim_{t \rightarrow \infty} \frac{N_t}{t} = \frac{1}{\mu}.$$

Απόδειξη. Από τον νόμο των μεγάλων αριθμών, υπάρχει ένα μετρήσιμο σύνολο $A \subset \Omega$ με $\mathbf{P}(A) = 1$ ώστε για κάθε $\omega \in A$ να ισχύει

$$\lim_{n \rightarrow \infty} \frac{T_n}{n} = \mu. \quad (12.1)$$

Από τον ορισμό του N_t έχουμε

$$T_{N_t} \leq t \leq T_{N_t+1},$$

και άρα

$$\frac{T_{N_t}}{N_t} \leq \frac{t}{N_t} \leq \frac{T_{N_t+1}}{N_t} = \frac{T_{N_t+1}}{N_t+1} \frac{N_t+1}{N_t}. \quad (12.2)$$

ΙΣΧΥΡΙΣΜΟΣ: Για κάθε $\omega \in \Omega$ ισχύει $N_t \rightarrow \infty$.

Πράγματι, επειδή η N_t είναι αύξουσα συνάρτηση του t , αν δεν ισχύει ο ισχυρισμός, τότε η N_t θα ήταν φραγμένη. Δηλαδή θα υπήρχε φυσικός $\ell \geq 1$ ώστε $N_t \leq \ell$ για κάθε $t \geq 0$. Άρα $X_1 + X_2 + \dots + X_\ell \geq t$ για κάθε $t \geq 0$, το οποίο δεν μπορεί να ισχύει γιατί οι X_1, X_2, \dots, X_ℓ παίρνουν πραγματικές τιμές (και όχι την τιμή ∞).

Για $\omega \in A$, και χρησιμοποιώντας τον ισχυρισμό και την (12.1), έχουμε

$$\lim_{t \rightarrow \infty} \frac{T_{N_t}}{N_t} = \lim_{t \rightarrow \infty} \frac{T_{N_t+1}}{N_t+1} = \mu.$$

Άρα για $t \rightarrow \infty$, η (12.2) δίνει $\lim_{t \rightarrow \infty} t/N_t = \mu$ που είναι το ζητούμενο. ■

Εντροπία. Έστω X διακριτή τυχαία μεταβλητή, S το (αριθμησιμο) σύνολο τιμών της, και $f(x) = \mathbf{P}(X = x)$ η συνάρτηση πιθανότητας της. **Εντροπία** της X ονομάζουμε τον αριθμό

$$H(X) := - \sum_{x \in S} f(x) \log f(x) \quad (12.3)$$

με τη σύμβαση $0 \log 0 = 0$. Ισχύει $H(X) \geq 0$ γιατί $f(x) \in [0, 1]$. Επίσης, $H(X) = -\mathbf{E}\{\log f(X)\}$.

Η εντροπία της X εκφράζει το μέγεθος της αβεβαιότητας που έχουμε για την τιμή που θα πάρει η X αν επιχειρήσουμε να παραγάγουμε μια πραγματοποίηση της. Ας πούμε ότι η X παίρνει τιμές στο $\{x_1, x_2, \dots, x_k\}$ με αντίστοιχες πιθανότητες $p_j := \mathbf{P}(X = x_j)$. Αν $p_1 = 1$ και $p_2 = \dots = p_k = 0$, τότε $H(X) = 0$, και βέβαια δεν υπάρχει καμία αβεβαιότητα για την πραγματοποίηση της X , θα είναι x_1 . Η εντροπία μεγιστοποιείται όταν $p_1 = p_2 = \dots = p_k = 1/k$ (Άσκηση 12.9). Τότε όλα τα ενδεχόμενα είναι το ίδιο πιθανά και η αβεβαιότητα μας μέγιστη. Σε περίπτωση άνισων πιθανοτήτων, η αβεβαιότητα είναι μικρότερη γιατί περιμένουμε η X να πάρει μία από τις τιμές που έχουν μεγαλύτερη πιθανότητα.

Τον Ορισμό (12.3) υπαγορεύει ένας υπολογισμός που θα δούμε αμέσως τώρα. Ο ίδιος υπολογισμός δίνει και την εμπειρική σημασία της εντροπίας.

Έστω $(X_k)_{k \geq 1}$ ακολουθία ανεξάρτητων και ισόνομων διακριτών τυχαίων μεταβλητών, καθεμία ισόνομη με τη X . Για $n \geq 1$ σταθερό και δεδομένα $x_1, x_2, \dots, x_n \in S$, υπολογίζουμε την πιθανότητα οι n πραγματοποιήσεις (X_1, X_2, \dots, X_n) της X να ταυτιστούν με το διάνυσμα (x_1, x_2, \dots, x_n) . Αυτή ισούται με

$$p_n(x_1, x_2, \dots, x_n) := \mathbf{P}(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = f(x_1)f(x_2) \cdots f(x_n). \quad (12.4)$$

Για παράδειγμα, όταν η X παίρνει τις τιμές 1 και 0 με πιθανότητες $p := 1/3$ και $2/3$ αντίστοιχα, τότε $f(x) = p^x(1-p)^{1-x} \mathbf{1}_{x \in \{0,1\}}$ και η πιθανότητα εμφάνισης της n -άδας $(x_1, x_2, \dots, x_n) \in \{0, 1\}^n$ είναι

$$\left(\frac{1}{3}\right)^{x_1 + \dots + x_n} \left(\frac{2}{3}\right)^{n - x_1 - \dots - x_n}$$

Αυτή η πιθανότητα παίρνει τιμές από $(1/3)^n$ ως και $(2/3)^n$. Δεν είναι σταθερή, η τιμή της εξαρτάται από την επιλογή της n -άδας (x_1, x_2, \dots, x_n) . Όμως η πιθανότητα πραγματοποίησης μιας τυχαίας n -άδας (X_1, X_2, \dots, X_n) είναι περίπου η ίδια σχεδόν για κάθε πραγματοποίηση της (X_1, X_2, \dots, X_n) .

Πιο συγκεκριμένα, στη γενική περίπτωση θέλουμε να εκτιμήσουμε την $p_n(X_1, X_2, \dots, X_n)$. Αυτή η πιθανότητα είναι μια τυχαία μεταβλητή (εξαρτάται από τις X_1, X_2, \dots, X_n), φθίνει εκθετικά με το n και από το νόμο των μεγάλων αριθμών έχουμε

$$\begin{aligned} \frac{1}{n} \log p_n(X_1, X_2, \dots, X_n) &= \frac{1}{n} \log f(X_1)f(X_2) \cdots f(X_n) \\ &= \frac{\log f(X_1) + \log f(X_2) + \cdots + \log f(X_n)}{n} \rightarrow \mathbf{E}(\log f(X_1)) = -H(X) \end{aligned} \quad (12.5)$$

με πιθανότητα 1 καθώς $n \rightarrow \infty$. Δηλαδή, με πιθανότητα 1, για μεγάλο n , ισχύει

$$p_n(X_1, X_2, \dots, X_n) \approx e^{-nH(X)}.$$

Στο παράδειγμα πιο πάνω, $H(X) = -(1/3) \log(1/3) - (2/3) \log(2/3)$ και η $e^{-nH(X)}$ είναι μια πιθανότητα ανάμεσα στις $(1/3)^n, (2/3)^n$.

Ας υποθέσουμε ότι το S είναι πεπερασμένο. Αν ορίσουμε

$$A_n^\varepsilon = \{(x_1, x_2, \dots, x_n) \in S^n : e^{-n(H(X)+\varepsilon)} < p_n(x_1, x_2, \dots, x_n) < e^{-n(H(X)-\varepsilon)}\},$$

τότε $\mathbf{P}(A_n^\varepsilon) \rightarrow 1$ λόγω της (12.5). Ακολουθίες στο A_n^ε τις λέμε ε -τυπικές. Αφού όλες τους έχουν πιθανότητα περίπου $e^{-nH(X)}$ και το άθροισμα των πιθανοτήτων τους είναι σχεδόν 1, το A_n^ε έχει πληθικότητα περίπου $e^{nH(X)}$. Στη γενική περίπτωση ισχύει $H(X) < \log |S|$ (Άσκηση 12.9), οπότε το A_n^ε είναι (από πλευράς πληθικότητας) ένα πολύ μικρό κομμάτι του S^n αφού το S^n έχει πληθικότητα $|S|^n = e^{n \log |S|}$. Παρ' όλα αυτά, το A_n^ε συγκεντρώνει σχεδόν όλη την πιθανότητα.

Η έννοια της εντροπίας είναι κεντρικής σημασίας στη Θεωρία Πληροφορίας και έχει πολλές εφαρμογές [δες Cover and Thomas (2014)].

Άσκησης

12.1 (Ασθενής Νόμος των Μεγάλων Αριθμών. Ασθενής έκδοση.) Έστω $(X_n)_{n \in \mathbb{N}}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με τιμές στο \mathbb{R} έτσι ώστε $\mathbf{E}(X_1^2) < \infty$. Θέτουμε $\mu = \mathbf{E}(X_1)$. Χωρίς χρήση του ισχυρού νόμου των μεγάλων αριθμών να δείχθει ότι

$$\lim_{n \rightarrow \infty} \mathbf{P}\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) = 0 \text{ για κάθε } \varepsilon > 0.$$

Δηλαδή η ακολουθία $\frac{S_n}{n}$ συγκλίνει στο μ κατά πιθανότητα.

12.2* Έστω $(X_n)_{n \in \mathbb{N}}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με τιμές στο \mathbb{R} έτσι ώστε $\mathbf{E}(X_1^+) = \infty$ και $\mathbf{E}(X_1^-) < \infty$. Να δείξετε ότι

$$\lim_{n \rightarrow \infty} \frac{S_n}{n} = \infty$$

με πιθανότητα 1.

12.3* (Αντίστροφο του Νόμου των Μεγάλων Αριθμών) Έστω $(X_n)_{n \geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με τιμές στο \mathbb{R} έτσι ώστε $\lim_{n \rightarrow \infty} \frac{S_n}{n} = \mu$ σχεδόν βέβαια με $\mu \in \mathbb{R}$. Να δείξετε ότι $\mathbf{E}|X_1| < \infty$ και $\mathbf{E}(X_1) = \mu$. [Υπόδειξη: Χρήσιμη είναι η Άσκηση 11.11.]

12.4 Έστω $(X_n)_{n \geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με $X_1 \sim N(1, 3)$. Να δείξετε ότι

$$\lim_{n \rightarrow \infty} \frac{X_1 + X_2 + \dots + X_n}{(X_1)^2 + (X_2)^2 + \dots + (X_n)^2} = \frac{1}{4}$$

με πιθανότητα 1.

Ο Ισχυρός Νόμος των Μεγάλων Αριθμών

12.5 (α)* Έστω $(X_n)_{n \in \mathbb{N}^+}$ ισόνομες τυχαίες μεταβλητές με τιμές στο \mathbb{R} , με $\mathbf{E}(X_1) \in \mathbb{R}$, και οι οποίες είναι m -εξαρτημένες, όπου $m \in \mathbb{N}^+$ δεδομένο. Δηλαδή, για κάθε $r \in \mathbb{N}^+$ και θετικούς ακέραιους $k_1 < k_2 < \dots < k_r$ με $k_{i+1} - k_i \geq m + 1$ για κάθε $i = 1, 2, \dots, r - 1$, οι τυχαίες μεταβλητές $X_{k_1}, X_{k_2}, \dots, X_{k_r}$ είναι ανεξάρτητες. Να δείξετε ότι $\lim_{n \rightarrow \infty} S_n/n = \mathbf{E}(X_1)$ με πιθανότητα 1.

(β) Έστω $(X_n)_{n \geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές ώστε $X_1 \sim U(0, 1)$. Θέτουμε

$$T_n := X_1 X_2 + X_2 X_3 + \dots + X_{n-1} X_n = \sum_{k=2}^n X_{k-1} X_k$$

για κάθε $n \in \mathbb{N}^+$. Να δειχθεί ότι το όριο $\lim_{n \rightarrow \infty} T_n/n$ υπάρχει με πιθανότητα 1 και να υπολογιστεί.

12.6 Έστω $(U_i)_{i \geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών, καθεμία με κατανομή $U(0, 1)$. Να δειχθεί ότι:

(α) $\lim_{n \rightarrow \infty} (U_1 U_2 \dots U_n)^{1/n} = e^{-1}$ με πιθανότητα 1.

(β) $\lim_{n \rightarrow \infty} U_1 U_2 \dots U_n = 0$ με πιθανότητα 1.

(γ) $\lim_{n \rightarrow \infty} \frac{U_1^a + \dots + U_n^a}{n} = \begin{cases} \frac{1}{1+a} & \text{με πιθανότητα 1 αν } a > -1, \\ \infty & \text{με πιθανότητα 1 αν } a \leq -1. \end{cases}$

12.7 Έστω $(X_i)_{i \geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με $\mu = \mathbf{E}(X_1) \in \mathbb{R}$ και $\sigma^2 = \text{Var}(X_1) < \infty$. Να δειχθεί ότι

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n (X_k - \mu)^2 = \sigma^2 \text{ με πιθανότητα 1.}$$

12.8* Έστω $p \in [0, 1]$ και $(X_n)_{n \geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με $\mathbf{P}(X_1 = 0) = 1 - p, \mathbf{P}(X_1 = 1) = p$. Θέτουμε

$$X := \sum_{n=1}^{\infty} \frac{1}{2^n} X_n.$$

Δηλαδή ο X είναι ένας αριθμός που γραμμένος στο δυαδικό σύστημα έχει ψηφία τις τυχαίες μεταβλητές $(X_n)_{n \geq 1}$. Να δειχθεί ότι:

(α) Αν $p = 1/2$, τότε η κατανομή της X είναι ο περιορισμός του μέτρου Lebesgue στο $[0, 1]$.

(β) Αν $p \neq 1/2$, τότε η κατανομή της X είναι ιδιάζουσα.

12.9 Έστω $k \geq 2, S = \{x_1, \dots, x_k\}$ σύνολο με k στοιχεία, και X τυχαία μεταβλητή με τιμές στο S . Να δειχθεί ότι $H(X) \leq \log k$ και η ισότητα ισχύει αν και μόνο αν $\mathbf{P}(X = x_j) = 1/k$ για κάθε $1 \leq j \leq k$.

[Υπόδειξη: Εφαρμόστε κατάλληλα την ανισότητα Jensen.]

12.10 Έστω $(X_k)_{k \geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με τιμές στο $S = \{1, 2, \dots, r\}$ και f η συνάρτηση πιθανότητας της X_1 . Έστω επίσης $(Y_k)_{k \geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με τιμές στο S και g η συνάρτηση πιθανότητας της Y_1 . Υποθέτουμε ότι για $k \in S$ ισχύει $f(k) = 0 \Rightarrow g(k) = 0$. Ορίζουμε την p_n όπως στην (12.4). Δείξτε ότι η πιθανότητα $p_n(Y_1, Y_2, \dots, Y_n)$ φθίνει εκθετικά και υπολογίστε τον ρυθμό μείωσης. Αυτή είναι η πιθανότητα στις πρώτες n συντεταγμένες η ακολουθία X να μοιάζει με δείγμα παρμένο από την Y .

12.11 Έστω $(X_n)_{n \geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με $\mathbf{E}|X_1| < \infty$. Να δειχθεί ότι $S_n/n \rightarrow \mathbf{E}(X_1)$ στον \mathcal{L}^1 .

[Υπόδειξη: Λήμμα Scheffe.]

12.12 Έστω $(X_n)_{n \geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με $\mathbf{E}|X_1| < \infty$ και $(k_n)_{n \geq 1}$ ακολουθία πραγματικών αριθμών με $\lim_{n \rightarrow \infty} k_n = \infty$. Να δειχθεί ότι

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_i \mathbf{1}_{|X_i| > k_n} = 0$$

με πιθανότητα 1.

13

Χαρακτηριστικές συναρτήσεις

13.1 Ολοκλήρωμα μιγαδικής συνάρτησης

Επεκτείνουμε τον ορισμό του ολοκληρώματος Lebesgue σε συναρτήσεις με μιγαδικές τιμές. Συγκεκριμένα, έστω $(\Omega, \mathcal{F}, \mu)$ χώρος μέτρου και $f : \Omega \rightarrow \mathbb{C}$ μετρήσιμη συνάρτηση (δηλαδή $\mathcal{F}/\mathcal{B}(\mathbb{C})$ -μετρήσιμη). Το πραγματικό και φανταστικό μέρος της f , που τα συμβολίζουμε με $\operatorname{Re}(f)$, $\operatorname{Im}(f)$, είναι συναρτήσεις στο Ω με πραγματικές τιμές και είναι εύκολο να δει κανείς ότι είναι επίσης μετρήσιμες. Ορίζουμε το ολοκλήρωμα Lebesgue της f ως προς το μέτρο μ ως εξής:

$$\int f \, d\mu := \int \operatorname{Re}(f) \, d\mu + i \int \operatorname{Im}(f) \, d\mu$$

με την προϋπόθεση ότι τα δύο πραγματικά ολοκληρώματα ορίζονται και είναι πραγματικοί αριθμοί¹. Τότε λέμε ότι η f είναι ολοκληρώσιμη. Ο χώρος των ολοκληρώσιμων συναρτήσεων (εφοδιασμένος με τις πράξεις της πρόσθεσης συναρτήσεων και πολλαπλασιασμού συνάρτησης με μιγαδικό αριθμό) είναι γραμμικός και το ολοκλήρωμα είναι γραμμική συνάρτηση σε αυτόν. Για κάθε ολοκληρώσιμη f ισχύει

$$\left| \int f \, d\mu \right| \leq \int |f| \, d\mu, \quad (13.1)$$

$$\int \bar{f} \, d\mu = \overline{\int f \, d\mu}, \quad (13.2)$$

όπου $|\cdot|$ συμβολίζει το μέτρο μιγαδικού. Η δεύτερη ιδιότητα είναι προφανής, ενώ για την πρώτη γράφουμε το ολοκλήρωμα σε πολική μορφή

$$\int f \, d\mu = e^{i\theta} \left| \int f \, d\mu \right|$$

με $\theta \in [0, 2\pi)$. Τότε

$$\left| \int f \, d\mu \right| = e^{-i\theta} \int f \, d\mu = \int e^{-i\theta} f \, d\mu = \int \operatorname{Re}(e^{-i\theta} f) \, d\mu \leq \int |e^{-i\theta} f| \, d\mu = \int |f| \, d\mu, \quad (13.3)$$

από το οποίο προκύπτει το ζητούμενο. Η τρίτη ισότητα ισχύει γιατί ξέρουμε ότι το αριστερό της μέλος είναι πραγματικός αριθμός.

13.2 Χαρακτηριστικές συναρτήσεις

Ορισμός 13.1. Έστω X τυχαία μεταβλητή σε χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$ με τιμές στο \mathbb{R} . **Χαρακτηριστική συνάρτηση** της X λέμε τη συνάρτηση $\phi_X : \mathbb{R} \rightarrow \mathbb{C}$ με

$$\phi_X(t) = \mathbf{E}(e^{itX}).$$

¹Δηλαδή αν $\int |\operatorname{Re}(f)| \, d\mu, \int |\operatorname{Im}(f)| \, d\mu < \infty$, το οποίο είναι ισοδύναμο με το $\int |f| \, d\mu < \infty$.

Θεώρημα 13.2. Έστω X τυχαία μεταβλητή με τιμές στο \mathbb{R} . Τότε:

(i) $|\phi_X(t)| \leq 1$ για κάθε $t \in \mathbb{R}$.

(ii) $\phi_X(0) = 1$.

(iii) Η ϕ_X είναι ομοιόμορφα συνεχής.

Απόδειξη. (i) Για $t \in \mathbb{R}$, χρησιμοποιώντας την (13.1), έχουμε ότι

$$|\phi_X(t)| = \left| \int e^{itX(\omega)} d\mathbf{P}(\omega) \right| \leq \int |e^{itX(\omega)}| d\mathbf{P}(\omega) = 1.$$

(ii) $\phi_X(0) = \mathbf{E}(e^0) = 1$.

(iii) Αρκεί να δείξουμε ότι για κάθε ακολουθία $(\delta_k)_{k \in \mathbb{N}}$ στον \mathbb{R} με $\delta_k \rightarrow 0$ ισχύει ότι

$$\limsup_{k \rightarrow \infty} \sup_{t \in \mathbb{R}} |\phi_X(t + \delta_k) - \phi_X(t)| = 0.$$

Έστω $t, \zeta \in \mathbb{R}$. Τότε,

$$\begin{aligned} |\phi_X(t + \zeta) - \phi_X(t)| &= |\mathbf{E}(e^{i(t+\zeta)X} - e^{itX})| = |\mathbf{E}\{e^{itX}(e^{i\zeta X} - 1)\}| \\ &\leq \mathbf{E}\{|e^{itX}| |e^{i\zeta X} - 1|\} = \mathbf{E}|e^{i\zeta X} - 1|. \end{aligned}$$

Άρα, αν $(\delta_k)_{k \in \mathbb{N}}$ είναι μηδενική ακολουθία, για $k \in \mathbb{N}$ έχουμε

$$\sup_{t \in \mathbb{R}} |\phi_X(t + \delta_k) - \phi_X(t)| \leq \mathbf{E}|e^{i\delta_k X} - 1|. \quad (13.4)$$

Από το θεώρημα κυριαρχημένη σύγκλισης, το δεξί μέλος της (13.4) τείνει στο 0. Γιατί θέτοντας $f_k(\omega) = |e^{i\delta_k X(\omega)} - 1|$ για κάθε $k \in \mathbb{N}$, έχουμε

(α) $\lim_{k \rightarrow \infty} f_k(\omega) = 0$ για κάθε $\omega \in \Omega$.

(β) $|f_k(\omega)| \leq 2 =: g(\omega)$ για κάθε $\omega \in \Omega$.

(γ) $\mathbf{E}(g) = 2 < \infty$.

Και έτσι προκύπτει το ζητούμενο. ■

Πρόταση 13.3. Έστω X, Y τυχαίες μεταβλητές στον $(\Omega, \mathcal{F}, \mathbf{P})$ με τιμές στον \mathbb{R} και $a, b \in \mathbb{R}$. Τότε για κάθε $t \in \mathbb{R}$ έχουμε

(i) $\phi_X(-t) = \overline{\phi_X(t)}$,

(ii) $\phi_{aX+b}(t) = e^{itb} \phi_X(at)$

(iii) Αν οι X, Y είναι ανεξάρτητες, τότε $\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t)$.

Απόδειξη. (i) $\phi_X(-t) = \mathbf{E}(e^{i(-t)X}) = \mathbf{E}(\overline{e^{itX}}) = \overline{\mathbf{E}(e^{itX})} = \overline{\phi_X(t)}$.

(ii) $\phi_{aX+b}(t) = \mathbf{E}(e^{it(aX+b)}) = e^{itb} \mathbf{E}(e^{iatX}) = e^{itb} \phi_X(at)$.

(iii) $\phi_{X+Y}(t) = \mathbf{E}(e^{it(X+Y)}) = \mathbf{E}(e^{itX} e^{itY}) = \mathbf{E}(e^{itX}) \mathbf{E}(e^{itY}) = \phi_X(t)\phi_Y(t)$, όπου στην τρίτη ισότητα χρησιμοποιήσαμε την ανεξαρτησία των X, Y και το Θεώρημα 10.8. ■

Στο επόμενο παράδειγμα, υπολογίζουμε τη χαρακτηριστική συνάρτηση τυχαίων μεταβλητών που ακολουθούν κάποια από τις γνωστές κατανομές.

Παράδειγμα 13.4. (i) Έστω $X \sim \text{Bin}(n, p)$. Τότε, $\phi_X(t) = (pe^{it} + 1 - p)^n$. Πράγματι,

$$\begin{aligned}\phi_X(t) &= \mathbf{E}(e^{itX}) = \sum_{k=0}^n e^{itk} \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=0}^n \binom{n}{k} (e^{it}p)^k (1-p)^{n-k} = (e^{it}p + 1 - p)^n.\end{aligned}$$

(ii) Έστω $X \sim \text{Poisson}(\lambda)$, $\lambda > 0$. Τότε, $\phi_X(t) = e^{\lambda(e^{it}-1)}$ [αποδεικνύεται όμοια με το (i)].

(iii) Έστω $X \sim U(-a, a)$ με $a > 0$. Τότε

$$\phi_X(t) = \begin{cases} \frac{\sin(at)}{at} & \text{αν } t \in \mathbb{R} \setminus \{0\}, \\ 1 & \text{αν } t = 0. \end{cases}$$

Πράγματι, η X έχει πυκνότητα

$$f_X(x) = \begin{cases} \frac{1}{2a} & \text{αν } x \in (-a, a), \\ 0 & \text{αν } x \in \mathbb{R} \setminus (-a, a). \end{cases}$$

Άρα, για $t \neq 0$, έχουμε

$$\begin{aligned}\phi_X(t) &= \mathbf{E}(e^{itX}) = \int_{-a}^a e^{itx} \frac{1}{2a} dx \\ &= \frac{1}{2a} \int_{-a}^a \cos(tx) dx + i \frac{1}{2a} \int_{-a}^a \sin(tx) dx \\ &= \frac{1}{2a} \frac{\sin(tx)}{t} \Big|_{-a}^a + 0 = \frac{1}{2a} \left(\frac{\sin(ta)}{t} - \frac{\sin(-ta)}{t} \right) = \frac{\sin(ta)}{ta}.\end{aligned}$$

Η τέταρτη ισότητα ισχύει γιατί έχουμε ολοκλήρωμα περιττής συνάρτησης σε διάστημα συμμετρικό γύρω από το 0. Για $t = 0$, προφανώς $\phi_X(0) = 1$.

(iv) Έστω $X \sim N(0, 1)$. Τότε, $\phi_X(t) = e^{-t^2/2}$.

Ο υπολογισμός της χαρακτηριστικής συνάρτησης στην περίπτωση αυτή είναι πιο περίπλοκος. Ένας τρόπος είναι με χρήση επιχειρημάτων από τη Μιγαδική Ανάλυση και θα τον δούμε στο Παράδειγμα 13.12 της Παραγράφου 13.5. Ένας άλλος, όχι και τόσο προφανής τρόπος, είναι ο εξής:

$$\begin{aligned}\phi_X(t) &= \mathbf{E}(e^{itx}) = \int_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \cos(tx) e^{-x^2/2} dx + i \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sin(tx) e^{-x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \cos(tx) e^{-x^2/2} dx.\end{aligned}$$

Η τελευταία ισότητα ισχύει γιατί η συνάρτηση $x \mapsto \sin(tx)e^{-x^2/2}$ είναι περιττή. Πλέον η συνάρτηση $\phi_X(t)$ είναι πραγματική, παραγωγίσιμη, και ισχύει ότι

$$\phi'_X(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (-x) \sin(tx) e^{-x^2/2} dx.$$

Χαρακτηριστικές συναρτήσεις

[Η παραγωγή της κάτω από το ολοκλήρωμα δικαιολογείται με χρήση της Πρότασης Α'5 του Παραρτήματος Α'. Θέτουμε $f(x, t) = e^{-x^2/2} \cos(tx)$. Η κυριαρχούσα συνάρτηση A που απαιτεί η πρόταση είναι η $A(x) = |x|e^{-x^2/2}$, η οποία έχει $\int_{-\infty}^{\infty} A(x) dx < \infty$.]

Ολοκληρώνοντας κατά μέρη, έχουμε

$$\phi'_X(t) = -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t \cos(tx) e^{-x^2/2} dx = -t\phi_X(t).$$

Έτσι καταλήγουμε στη συνήθη διαφορική εξίσωση $\phi'_X(t) = -t\phi_X(t)$, η οποία έχει γενική λύση

$$\phi_X(t) = C e^{-t^2/2}.$$

Και εφόσον $\phi_X(0) = 1$, έχουμε ότι $C = 1$. Άρα $\phi_X(t) = e^{-t^2/2}$.

(v) Έστω $X \sim N(\mu, \sigma^2)$. Τότε

$$\phi_X(t) = e^{it\mu - t^2 \frac{\sigma^2}{2}}.$$

Από τα προηγούμενα, θεωρώντας την τυχαία μεταβλητή $Z = \frac{X-\mu}{\sigma}$, έχουμε ότι $Z \sim N(0, 1)$ και $X = \sigma Z + \mu$. Άρα, $\phi_X(t) = \phi_{\sigma Z + \mu}(t) = e^{it\mu} \phi_Z(t\sigma) = e^{it\mu} e^{-t^2 \frac{\sigma^2}{2}}$.

(vi) Έστω $X \sim \text{Exp}(\lambda)$, $\lambda > 0$. Τότε, $\phi_X(t) = \frac{\lambda}{\lambda - it}$ για κάθε $t \in \mathbb{R}$. Πράγματι,

$$\begin{aligned} \phi_X(t) &= \mathbf{E}(e^{itX}) = \int_0^{\infty} e^{itx} \lambda e^{-\lambda x} dx \\ &= \lambda \int_0^{\infty} e^{x(-\lambda+it)} dx = \lim_{M \rightarrow \infty} \lambda \int_0^M e^{x(-\lambda+it)} dx \\ &= \lim_{M \rightarrow \infty} \lambda \frac{e^{x(-\lambda+it)}}{-\lambda+it} \Big|_0^M = \lim_{M \rightarrow \infty} \lambda \frac{e^{M(-\lambda+it)} - 1}{-\lambda+it} \\ &= \frac{\lambda}{\lambda - it}, \end{aligned}$$

γιατί $\lim_{M \rightarrow \infty} e^{-\lambda M} e^{itM} = 0$ αφού $|e^{itM}| = 1$ και $\lambda > 0$.

(vii) Έστω $X \sim \Gamma(a, \lambda)$, με $a, \lambda > 0$ σταθερές. Δηλαδή η X έχει πυκνότητα

$$f_X(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x} \mathbf{1}_{x \geq 0}.$$

Τότε, $\phi_X(t) = \left(\frac{\lambda}{\lambda - it}\right)^a$ για κάθε $t \in \mathbb{R}$. Αυτό θα προκύψει από τη σχέση ροπογεννητριών και χαρακτηριστικών συναρτήσεων που εκτίθεται στην Παράγραφο 13.5 παρακάτω.

(viii) Έστω ότι X ακολουθεί την κατανομή Cauchy, δηλαδή έχει πυκνότητα

$$f(x) = \frac{1}{\pi} \frac{1}{1+x^2}$$

για κάθε $x \in \mathbb{R}$. Τότε $\phi_X(t) = e^{-|t|}$ για κάθε $t \in \mathbb{R}$.

Πράγματι, αν $t > 0$, τότε θεωρούμε τη μερόμορφη συνάρτηση $G(z) = e^{itz}/(1+z^2)$, η οποία ικανοποιεί $\lim_{z \rightarrow \infty, \text{Im}(z) \geq 0} |z|G(z) = 0$ (γιατί $t > 0$) και της οποίας ο μοναδικός πόλος στο

Χαρακτηριστικές συναρτήσεις

$\{z \in \mathbb{C} : \text{Im}(z) \geq 0\}$ είναι το i . Κατά τα γνωστά από τη μιγαδική ανάλυση (υπολογισμός ολοκληρωμάτων με τη βοήθεια ολοκληρωτικών υπολοίπων), έχουμε

$$\phi_X(t) = \frac{1}{\pi} \int_{\mathbb{R}} e^{itx} \frac{1}{1+x^2} dx = \frac{1}{\pi} 2\pi i \text{Res}(G(z), i) = 2i \frac{e^{i^2 t}}{2i} = e^{-t}.$$

Αν $t < 0$, δουλεύουμε στο ημιεπίπεδο $\{z \in \mathbb{C} : \text{Im}(z) \leq 0\}$ και βρίσκουμε ότι το ολοκλήρωμα ισούται με e^t (εναλλακτικά, κάνουμε αλλαγή μεταβλητής, $y = -x$, στο ολοκλήρωμα και αναγόμενα στην προηγούμενη περίπτωση).

13.3 Ο τύπος αντιστροφής

Η χαρακτηριστική συνάρτηση, ϕ_X , οποιασδήποτε τυχαίας μεταβλητής $X : \Omega \rightarrow \mathbb{R}$ καθορίζεται από την κατανομή, \mathbf{P}^X , της X γιατί

$$\phi_X(t) = \mathbf{E}(e^{itX}) = \int e^{itx} d\mathbf{P}^X(x).$$

Σε αυτή την παράγραφο θα δούμε ότι γνωρίζοντας την ϕ_X μπορούμε να ανακτήσουμε την κατανομή \mathbf{P}^X .

Αν μ είναι μέτρο πιθανότητας στον $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, **μετασχηματισμό Fourier** του μ ονομάζουμε τη συνάρτηση $\phi_\mu : \mathbb{R} \rightarrow \mathbb{C}$ που ορίζεται ως $\phi_\mu(t) := \int e^{itx} d\mu(x)$ για κάθε $t \in \mathbb{R}$. Η ϕ_μ καθορίζει το μέτρο μ μέσω του τύπου αντιστροφής.

Θεώρημα 13.5 (Ο τύπος αντιστροφής). Για κάθε $a < b$ ισχύει

$$\lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-T}^T \frac{e^{-ita} - e^{-itb}}{it} \phi_\mu(t) dt = \mu((a, b)) + \frac{1}{2}\mu(\{a\}) + \frac{1}{2}\mu(\{b\}). \quad (13.5)$$

Η απόδειξη του θεωρήματος δίνεται στο Παράρτημα Β'.

Αν το μ είναι η κατανομή μιας τυχαίας μεταβλητής X , τότε $\phi_\mu = \phi_X$ και το δεξί μέλος της (13.5) ισούται με $\mathbf{P}(a < X < b) + \{\mathbf{P}(X = a) + \mathbf{P}(X = b)\}/2$.

Πόρισμα 13.6 (Θεώρημα Μοναδικότητας). (i) Αν μ, ν είναι μέτρα πιθανότητας στον $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ και $\phi_\mu(t) = \phi_\nu(t)$ για κάθε $t \in \mathbb{R}$, τότε $\mu = \nu$.

(ii) Αν X, Y είναι δυο τυχαίες μεταβλητές με τιμές στο \mathbb{R} και $\phi_X(t) = \phi_Y(t)$ για κάθε $t \in \mathbb{R}$, τότε οι X, Y έχουν την ίδια κατανομή, δηλαδή $\mathbf{P}^X = \mathbf{P}^Y$.

Απόδειξη. (i) Έστω $C := \{x \in \mathbb{R} : \mu(\{x\}) > 0 \text{ ή } \nu(\{x\}) > 0\}$, το οποίο είναι αριθμήσιμο. Εφόσον $\phi_\mu = \phi_\nu$, ο τύπος αντιστροφής δίνει ότι για κάθε $a, b \in \mathbb{R} \setminus C$ ισχύει $\mu([a, b]) = \nu([a, b])$. Εφαρμόζοντας αυτή τη σχέση στα μέλη μιας ακολουθίας $(a_n)_{n \geq 1}$ σημείων του $\mathbb{R} \setminus C$ με $\lim_{n \rightarrow \infty} a_n = -\infty$ και παίρνοντας $n \rightarrow \infty$ βρίσκουμε ότι $\mu((-\infty, b]) = \nu((-\infty, b])$. Τα δύο μέλη της τελευταίας ισότητας είναι δεξιά συνεχείς συναρτήσεις του b . Για οποιοδήποτε $x \in \mathbb{R}$, θεωρούμε φθίνουσα ακολουθία, $(b_n)_{n \geq 1}$, σημείων του $\mathbb{R} \setminus C$ που συγκλίνει στο x . Τότε

$$\mu((-\infty, x]) = \lim_{n \rightarrow \infty} \mu((-\infty, b_n]) = \lim_{n \rightarrow \infty} \nu((-\infty, b_n]) = \nu((-\infty, x]).$$

Το συμπέρασμα έπεται από το Πόρισμα 3.7.

(ii) Έπεται από το (i) γιατί $\phi_Z = \phi_{\mathbf{P}^Z}$ για κάθε τυχαία μεταβλητή Z . ■

Παράδειγμα 13.7. (i) Έστω X_1, X_2, \dots, X_n ανεξάρτητες και ισόνομες τυχαίες μεταβλητές τέτοιες ώστε $X_1 \sim \text{Bernoulli}(p)$. Αν $Y = \sum_{j=1}^n X_j$, τότε $Y \sim \text{Bin}(n, p)$. Πράγματι, η χαρακτηριστική συνάρτηση καθεμίας από τις X_j ισούται με $\phi_{X_1}(t) = e^{it}p + 1 - p$ και από την Πρόταση 13.3(iii) έπεται ότι

$$\phi_Y(t) = \phi_{X_1}(t)^n = (e^{it}p + 1 - p)^n,$$

που είναι η χαρακτηριστική συνάρτηση της $\text{Bin}(n, p)$. Το θεώρημα μοναδικότητας (Πόρισμα 13.6) δίνει ότι $Y \sim \text{Bin}(n, p)$.

(ii) Έστω X, Y ανεξάρτητες τυχαίες μεταβλητές τέτοιες ώστε $X \sim \text{Poisson}(\lambda)$ και $Y \sim \text{Poisson}(\mu)$. Τότε για τη $Z \sim X + Y$ έχουμε ότι $Z \sim \text{Poisson}(\lambda + \mu)$. Πράγματι, καταρχάς παρατηρούμε ότι η χαρακτηριστική συνάρτηση της Z είναι

$$\phi_Z(t) = \phi_X(t)\phi_Y(t) = e^{\lambda(e^{it}-1)}e^{\mu(e^{it}-1)} = e^{(\lambda+\mu)(e^{it}-1)},$$

όπου η πρώτη ισότητα ισχύει γιατί οι X, Y είναι ανεξάρτητες. Συνεπώς, η ϕ_Z είναι η χαρακτηριστική συνάρτηση της κατανομής $\text{Poisson}(\lambda + \mu)$, και από το θεώρημα μοναδικότητας η Z έχει κατανομή $\text{Poisson}(\lambda + \mu)$.

(iii) Έστω $X \sim \text{Bin}(n, p)$ και $Y \sim \text{Bin}(m, p)$ ανεξάρτητες τυχαίες μεταβλητές. Τότε η $Z = X + Y$ έχει κατανομή $\text{Bin}(n + m, p)$. Αυτό προκύπτει εύκολα με τη χρήση χαρακτηριστικών συναρτήσεων ή με χρήση του (i), αναπαριστώντας τις X, Y ως άθροισμα ανεξάρτητων τυχαίων μεταβλητών κατανομής $\text{Bernoulli}(p)$.

(iv) Έστω $X \sim N(\mu, \sigma^2)$ και $Y \sim N(\nu, \tau^2)$ ανεξάρτητες τυχαίες μεταβλητές. Τότε, η $Z = X + Y$ έχει κατανομή $N(\mu + \nu, \sigma^2 + \tau^2)$. Αυτό γιατί η χαρακτηριστική συνάρτηση της Z είναι

$$\phi_Z(t) = \phi_X(t)\phi_Y(t) = e^{it\mu - t^2 \frac{\sigma^2}{2}} e^{it\nu - t^2 \frac{\tau^2}{2}} = e^{it(\mu+\nu) - t^2 \frac{\sigma^2 + \tau^2}{2}}.$$

Η τελευταία συνάρτηση είναι η χαρακτηριστική συνάρτηση της κατανομής $N(\mu + \nu, \sigma^2 + \tau^2)$ και από το θεώρημα μοναδικότητας προκύπτει το ζητούμενο.

(v) Έστω $X \sim \Gamma(a_1, \lambda)$ και $Y \sim \Gamma(a_2, \lambda)$ ανεξάρτητες τυχαίες μεταβλητές. Τότε, δουλεύοντας όμοια με τα προηγούμενα, δείχνουμε ότι η $Z = X + Y$ έχει κατανομή $\Gamma(a_1 + a_2, \lambda)$.

Μια συνέπεια αυτού του αποτελέσματος είναι η εξής αναπαράσταση της κατανομής χ^2 τετράγωνο με p βαθμούς ελευθερίας ($p \in \mathbb{N}^+$). Αν οι Y_1, Y_2, \dots, Y_p είναι ανεξάρτητες τυχαίες μεταβλητές με κατανομή $N(0, 1)$, τότε η

$$X := Y_1^2 + Y_2^2 + \dots + Y_p^2$$

έχει κατανομή χ_p^2 . Υπενθυμίζουμε ότι χ_p^2 είναι η κατανομή $\Gamma(\frac{p}{2}, \frac{1}{2})$.

Για την απόδειξη αυτού του ισχυρισμού, δείχνουμε με τη γνωστή τεχνική από τις στοιχειώδεις πιθανότητες ότι αν $Y \sim N(0, 1)$, τότε η $Y^2 \sim \Gamma(\frac{1}{2}, \frac{1}{2})$ (περίπτωση $p = 1$ του ισχυρισμού). Έπειτα εφαρμόζουμε το αποτέλεσμα που δείξαμε για άθροισμα ανεξάρτητων τυχαίων μεταβλητών που έχουν κατανομή Γάμμα με κοινή παράμετρο κλίμακας λ .

Πόρισμα 13.8. Έστω μ μέτρο πιθανότητας στον $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ ώστε $\int_{\mathbb{R}} |\phi_{\mu}(t)| dt < \infty$. Τότε το μ έχει πυκνότητα

$$f(x) := \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} \phi_{\mu}(t) dt \tag{13.6}$$

για κάθε $x \in \mathbb{R}$, η οποία είναι συνεχής και φραγμένη.

Χαρακτηριστικές συναρτήσεις

Απόδειξη. Η f είναι φραγμένη αφού $|f(x)| \leq (1/2\pi) \int_{\mathbb{R}} |\phi_{\mu}(t)| dt < \infty$ για κάθε $x \in \mathbb{R}$. Επίσης, το θεώρημα κυριαρχημένης σύγκλισης δίνει ότι είναι συνεχής.

Για $a < b$ και $t \neq 0$, γράφουμε $\frac{e^{-ita} - e^{-itb}}{it} = \int_a^b e^{-itx} dx$. Έτσι, για $T > 0$ έχουμε

$$\frac{1}{2\pi} \int_{-T}^T \frac{e^{-ita} - e^{-itb}}{it} \phi_{\mu}(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_a^b e^{-itx} \phi_{\mu}(t) \mathbf{1}_{t \in [-T, T]} dx dt.$$

Επειδή $|e^{-itx} \phi_{\mu}(t) \mathbf{1}_{t \in [-T, T]}| \leq |\phi_{\mu}(t)|$ για κάθε $(x, t) \in [a, b] \times \mathbb{R}$ και $\int_{\mathbb{R}} \int_a^b |\phi_{\mu}(t)| dx dt < \infty$, έπεται από το θεώρημα κυριαρχημένης σύγκλισης (παίρνουμε $T \rightarrow \infty$) ότι το αριστερό μέλος του τύπου αντιστροφής, (13.5), ισούται με

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-ita} - e^{-itb}}{it} \phi_{\mu}(t) dt.$$

Το όριο αυτής της ποσότητας για $b \rightarrow a$ ισούται με 0 (το θεώρημα κυριαρχημένης σύγκλισης εφαρμόζεται). Επομένως, ο τύπος αντιστροφής δίνει ότι $\mu(\{a\}) = 0$ για κάθε $a \in \mathbb{R}$ και με εφαρμογή του θεωρήματος Fubini γράφεται

$$\mu([a, b]) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_a^b e^{-itx} \phi_{\mu}(t) dx dt = \int_a^b f(x) dx. \quad (13.7)$$

Επειδή η f είναι συνεχής, η τελευταία σχέση συνεπάγεται ότι $f \geq 0$. Έπειτα για $a \rightarrow -\infty, b \rightarrow \infty$ δίνει ότι $\int_{-\infty}^{\infty} f(x) dx = 1$. Άρα η συνάρτηση $\nu : \mathcal{B}(\mathbb{R}) \rightarrow [0, \infty]$ με $\nu(A) = \int_A f(x) dx$ για κάθε $A \in \mathcal{B}(\mathbb{R})$ ορίζει μέτρο πιθανότητας (με πυκνότητα την f). Παίρνοντας $a \rightarrow -\infty$ στην (13.7), έχουμε ότι τα μ, ν συμφωνούν στα διαστήματα της μορφής $(-\infty, b]$, άρα $\mu = \nu$ (Πόρισμα 3.7). ■

13.4 Ροπογεννήτριες

Για μια τυχαία μεταβλητή X με τιμές στο $[-\infty, \infty]$ η ροπογεννήτριά της είναι η συνάρτηση $M_X : \mathbb{R} \rightarrow [0, \infty]$ με

$$M_X(t) = \mathbf{E}(e^{tX}).$$

Η M_X ως μέση τιμή θετικής τυχαίας μεταβλητής ορίζεται για κάθε $t \in \mathbb{R}$, απλώς ενδέχεται σε κάποια t να παίρνει την τιμή ∞ . Αν η X παίρνει τιμές στο $[0, \infty]$ (αντίστοιχα, στο $[-\infty, 0]$), τότε η M_X είναι πεπερασμένη για κάθε $t \leq 0$ (αντίστοιχα, για κάθε $t \geq 0$), και μάλιστα $M_X(t) \leq 1$ για εκείνα τα t . Πάντοτε $M_X(0) = 1$, ενώ το δεδομένο $M_X(t) < \infty$ για κάποιο $t \neq 0$ έχει χρήσιμες συνέπειες. Καταγράφουμε μία από αυτές στο επόμενο λήμμα (δες επίσης την Άσκηση 5.5).

Λήμμα 13.9. (i) Αν $\varepsilon > 0$ και $M_X(\varepsilon) < \infty$, τότε $M_X(t) < \infty$ για κάθε $t \in [0, \varepsilon]$ και $\mathbf{E}((X^+)^k) < \infty$ για κάθε $k \in \mathbb{N}$.

(ii) Αν $\varepsilon > 0$ και $M_X(-\varepsilon) < \infty$, τότε $M_X(t) < \infty$ για κάθε $t \in [-\varepsilon, 0]$ και $\mathbf{E}((X^-)^k) < \infty$ για κάθε $k \in \mathbb{N}$.

Απόδειξη. (i) Για $t \in [0, \varepsilon]$ έχουμε $e^{tX} \leq e^{\varepsilon X} + 1$ (παίρνουμε τις περιπτώσεις $X(\omega) \geq 0$ και $X(\omega) < 0$), άρα $M_X(t) < \infty$. Έπειτα, πάλι παίρνοντας περιπτώσεις, έχουμε $0 \leq \varepsilon^k (X^+)^k \leq k! e^{\varepsilon X}$ και το συμπέρασμα έπεται.

(ii) Όμοια όπως στο μέρος (i). ■

Το λήμμα συνεπάγεται ότι το $D_X := \{t \in \mathbb{R} : M_X(t) < \infty\}$ είναι ένα διάστημα που περιέχει το 0. Στη χειρότερη περίπτωση είναι το $\{0\}$. Επίσης, αν η X είναι τυχαία μεταβλητή με $\mathbf{E}(X^-) = \mathbf{E}(X^+) = \infty$, το λήμμα δίνει ότι $D_X = \{0\}$. Παράδειγμα τέτοιας τυχαίας μεταβλητής είναι η Cauchy

(Παράδειγμα 6.10), ενώ και άλλες τυχαίες μεταβλητές με $D_X = \{0\}$ δίνει η Άσκηση 13.5. Όλες αυτές οι τυχαίες μεταβλητές έχουν την ίδια ροπογεννήτρια αλλά διαφορετική κατανομή. Άρα η ροπογεννήτρια δεν χαρακτηρίζει την κατανομή μιας τυχαίας μεταβλητής, δεν την κωδικοποιεί.

Μελετούμε τώρα την περίπτωση που το D_X περιέχει ένα ανοιχτό διάστημα $(-\varepsilon, \varepsilon)$ γύρω από το 0.

Πρόταση 13.10. *Αν υπάρχει $\varepsilon > 0$ ώστε $M_X(-\varepsilon), M_X(\varepsilon) < \infty$, τότε*

- (i) $M_X(t) < \infty$ για κάθε $t \in [-\varepsilon, \varepsilon]$.
- (ii) $\mathbf{E}(|X|^k) < \infty$ για κάθε $k \in \mathbb{N}$.
- (iii) Η M_X αναλύεται σε δυναμοσειρά ως

$$M_X(t) = \sum_{k=0}^{\infty} \frac{\mathbf{E}(X^k)}{k!} t^k \quad (13.8)$$

με ακτίνα σύγκλισης τουλάχιστον ε .

- (iv) $\mathbf{E}(X^k) = M_X^{(k)}(0)$ για κάθε $k \in \mathbb{N}$.

Απόδειξη. (i) Έπεται από το προηγούμενο λήμμα.

(ii) Έπεται από το προηγούμενο λήμμα και το ότι $|X|^k = (X^-)^k + (X^+)^k$.

(iii) Για $t \in (-\varepsilon, \varepsilon)$ έχουμε

$$M_X(t) = \mathbf{E}\left(\sum_{k=0}^{\infty} \frac{t^k X^k}{k!}\right) = \sum_{k=0}^{\infty} \mathbf{E}\left(\frac{t^k X^k}{k!}\right) = \sum_{k=0}^{\infty} \frac{t^k \mathbf{E}(X^k)}{k!}.$$

Η εναλλαγή ολοκληρώματος και αθροίσματος έπεται από το θεώρημα Fubini (εφαρμοσμένο στα μέτρα \mathbf{P} , αριθμητικό μέτρο στο \mathbb{N}) γιατί

$$\mathbf{E}\left(\sum_{k=0}^{\infty} \frac{|t^k X^k|}{k!}\right) = \mathbf{E}(e^{|tX|}) < \mathbf{E}(e^{-tX}) + \mathbf{E}(e^{tX}) < \infty.$$

Επιπλέον, το άθροισμα της δυναμοσειράς (13.8) είναι πεπερασμένο.

(iv) Έπεται από το (iii) και τη θεωρία των δυναμοσειρών. ■

Για να θυμάται κανείς τον τύπο $\mathbf{E}(X^k) = M_X^{(k)}(0)$ χρήσιμη είναι η εξής «απόδειξη» του. Στην $M_X(t) = \mathbf{E}(e^{tX})$ παραγωγίζουμε k φορές και παίρνουμε

$$M_X^{(k)}(t) = \mathbf{E}(X^k e^{tX}). \quad (13.9)$$

Δηλαδή περνάμε την παράγωγο μέσα από τη μέση τιμή. Το ότι αυτό είναι σωστό αποδεικνύεται με χρήση του θεωρήματος κυριαρχημένης σύγκλισης, όπως εξηγείται στην Πρόταση Α'5, αλλά το παραλείπουμε (η απόδειξη όταν $k = 1$ δίνεται στο Λήμμα 17.7). Έπειτα θέτουμε $t = 0$ στην (13.9).

Αντιπαραβάλλουμε τη χαρακτηριστική συνάρτηση με τη ροπογεννήτρια μιας τυχαίας μεταβλητής. Γράφουμε (+) στα προτέρηματα και (-) στα ελαττώματα.

Η χαρακτηριστική συνάρτηση:

- (i) Είναι πάντοτε πεπερασμένος αριθμός. (+)

Χαρακτηριστικές συναρτήσεις

(ii) Χαρακτηρίζει την κατανομή της X . Δύο τυχαίες μεταβλητές με ίδια χαρακτηριστική συνάρτηση έχουν την ίδια κατανομή (Πόρισμα 13.6). (+)

(iii) Ο υπολογισμός της ενδέχεται να εμπλέκει ολοκληρώματα μιγαδικών συναρτήσεων. (-)

Η ροπογεννήτρια:

(i) Ενδέχεται να πάρει την τιμή ∞ . (-)

(ii) Ο υπολογισμός της εμπλέκει ολοκληρώματα ή αθροίσματα στο \mathbb{R} . (+)

(iii) Γενικά, δεν χαρακτηρίζει την κατανομή της X . Δύο τυχαίες μεταβλητές ενδέχεται να έχουν την ίδια ροπογεννήτρια αλλά διαφορετική κατανομή. (-)

(iv) Η υπόθεση $M_X(t) < \infty$ για κάποιο $t \neq 0$ δίνει πληροφορίες για τη X (π.χ., Λήμμα 13.9 και Άσκηση 5.5). (+)

13.5 Χαρακτηριστικές συναρτήσεις μέσω ροπογεννητριών*

Έχοντας υπολογίσει κανείς την ροπογεννήτρια M_X της X είναι δελεαστικό να προσπαθήσει να υπολογίσει τη χαρακτηριστική συνάρτηση ως

$$\phi_X(t) = \mathbf{E}(e^{itX}) \stackrel{!}{=} M_X(it).$$

Ένα πρώτο πρόβλημα είναι ότι το σύμβολο $M_X(it)$ δεν έχει νόημα αφού η M_X έχει πεδίο ορισμού το \mathbb{R} . Ας το παραβλέψουμε. Η ιδέα είναι να βρούμε έναν τύπο για τη M_X στον οποίο να μπορέσουμε να βάλουμε όπου t το it . Και έχουμε παραδείγματα που αυτό δουλεύει. Π.χ. στην περίπτωση που η X ακολουθεί κάποια κανονική ή εκθετική κατανομή.

Ας δούμε τι γίνεται αν η $X \sim N(0, 1)$. Βρίσκουμε ότι $M_X(t) = e^{t^2/2}$. Βάζοντας όπου t το it βρίσκουμε $e^{-t^2/2}$ που είναι ο σωστός τύπος για τη χαρακτηριστική συνάρτηση της X . Είναι δυνατόν όμως να πει κανείς ότι $M_X(t) = e^{t^2/2}$ για κάθε $t \in \mathbb{R}$ και η αντικατάσταση $t \rightarrow it$ δίνει $e^{t^2/2}$, που είναι λάθος. Τι καλύτερο έχει ο τύπος $e^{t^2/2}$ από τον $e^{it^2/2}$;

Πρόταση 13.11. Έστω X πραγματική τυχαία μεταβλητή με ροπογεννήτρια M_X . Υποθέτουμε ότι υπάρχει $\varepsilon > 0$ ώστε

(i) Η M_X είναι πεπερασμένη στο $(-\varepsilon, \varepsilon)$.

(ii) Υπάρχει αναλυτική συνάρτηση $h : \{z \in \mathbb{C} : |\operatorname{Re}(z)| < \varepsilon\} \rightarrow \mathbb{C}$ ώστε το σύνολο $\{t \in \mathbb{R} : M_X(t) = h(t)\}$ να έχει σημείο συσσώρευσης στο $(-\varepsilon, \varepsilon)$.

Τότε $\phi_X(t) = h(it)$ για κάθε $t \in \mathbb{R}$.

Απόδειξη. Έστω $A_\varepsilon := \{z \in \mathbb{C} : |\operatorname{Re}(z)| < \varepsilon\}$. Θέτουμε $g : A_\varepsilon \rightarrow \mathbb{C}$ με $g(z) := \mathbf{E}(e^{zX})$ για κάθε $z \in A_\varepsilon$.

ΙΣΧΥΡΙΣΜΟΣ: Η g είναι καλά ορισμένη² και αναλυτική στο A_ε .

Επειδή $|e^{zX}| = e^{X \operatorname{Re} z}$ και $\mathbf{E}(e^{X \operatorname{Re} z}) < \infty$ από την υπόθεση (i), έπεται ότι η g είναι καλά ορισμένη. Τώρα για $z_0 \in A_\varepsilon$ και $z \in \mathbb{C}$ με $|z| < \varepsilon - |\operatorname{Re}(z_0)|$ ισχύει

$$g(z_0 + z) = \mathbf{E}(e^{z_0 X} e^{zX}) = \mathbf{E} \left\{ e^{z_0 X} \sum_{k=0}^{\infty} \frac{(zX)^k}{k!} \right\} = \sum_{k=0}^{\infty} \frac{\mathbf{E} \{ X^k e^{z_0 X} \}}{k!} z^k. \quad (13.10)$$

²Το καλά ορισμένη σημαίνει ότι η μέση τιμή μπορεί να οριστεί και είναι στοιχείο του \mathbb{C} . Δεν εμφανίζεται κάποια μορφή $\infty - \infty$.

Χρειάζεται δικαιολόγηση μόνο η τελευταία ισότητα. Δηλαδή η αλλαγή σειράς μέσης τιμής και άθροισης. Αυτό έπεται από το θεώρημα Fubini αφού

$$\mathbf{E} \left\{ \sum_{k=0}^{\infty} \left| e^{z_0 X} \frac{(zX)^k}{k!} \right| \right\} = \mathbf{E} \left\{ e^{\operatorname{Re}(z_0)X + |zX|} \right\} \leq \mathbf{E} \left(e^{|X|(|z| + |\operatorname{Re}(z_0)|)} \right) < \infty.$$

Το ότι η τελευταία ποσότητα είναι πεπερασμένη έπεται από το ότι $|z| + |\operatorname{Re}(z_0)| < \varepsilon$ και την υπόθεση (i). Εδώ λοιπόν είναι κρίσιμη η υπόθεση ότι η M_X είναι πεπερασμένη στο $(-\varepsilon, \varepsilon)$. Επίσης συμπεραίνουμε ότι στο δεξί μέλος της (13.10) έχουμε μια δυναμοσειρά του z με πεπερασμένους συντελεστές η οποία συγκλίνει αφού η $g(z_0 + z)$ είναι πεπερασμένη. Έπεται ότι η g αναλύεται σε δυναμοσειρά με κέντρο z_0 και ακτίνα σύγκλισης τουλάχιστον $\varepsilon - |\operatorname{Re}(z_0)| > 0$, πράγμα που αποδεικνύει τον ισχυρισμό.

Από την υπόθεση (ii) το σύνολο των σημείων που ισχύει $g(z) = h(z)$ έχει σημείο συσσώρευσης στο $(-\varepsilon, \varepsilon) \subset A_\varepsilon$. Από την αρχή αναλυτικής συνέχισης οι συναρτήσεις h, g ταυτίζονται στο A_ε . Άρα, για $t \in \mathbb{R}$,

$$\phi_X(t) = g(it) = h(it)$$

αφού $it \in A_\varepsilon$. ■

Επιστρέφοντας στη συζήτηση πριν την πρόταση, το πρόβλημα με την $e^{|z|^2/2}$ είναι ότι δεν είναι αναλυτική συνάρτηση (ούτε καν σε ένα σημείο του \mathbb{C}). Έτσι δεν μπορεί να παίξει τον ρόλο της h που αναφέρει η πρόταση.

Παράδειγμα 13.12. (i) Μια $X \sim N(0, 1)$ έχει ροπογεννήτρια $M_X(t) = e^{t^2/2}$ για κάθε $t \in \mathbb{R}$. Η M_X είναι σαφώς πεπερασμένη σε περιοχή του 0. Η συνάρτηση $h : \mathbb{C} \rightarrow \mathbb{C}$ με $h(z) = e^{z^2/2}$ για κάθε $z \in \mathbb{C}$ είναι αναλυτική σε όλο το \mathbb{C} και συμφωνεί με τη M_X στο \mathbb{R} (είναι η μόνη αναλυτική που το κάνει αυτό). Άρα η Πρόταση 13.11 εφαρμόζεται και δίνει ότι $\phi_X(t) = h(it) = e^{-t^2/2}$ για κάθε $t \in \mathbb{R}$.

(ii) Μια $X \sim \operatorname{Exp}(\lambda)$ ($\lambda > 0$) έχει ροπογεννήτρια

$$M_X(t) = \begin{cases} \frac{\lambda}{\lambda - t} & \text{αν } t < \lambda, \\ \infty & \text{αν } t \geq \lambda. \end{cases}$$

Η M_X είναι πεπερασμένη στην περιοχή $(-\lambda, \lambda)$ του 0. Η συνάρτηση $h : \mathbb{C} \setminus \{\lambda\} \rightarrow \mathbb{C}$ με $h(z) = \lambda/(\lambda - z)$ είναι αναλυτική στο πεδίο ορισμού της [το οποίο περιέχει μια λωρίδα της μορφής $\{z : |\operatorname{Re}(z)| < \varepsilon\}$ με $\varepsilon > 0$, π.χ. με $\varepsilon = \lambda$.] και συμφωνεί με τη M_X στο $(-\infty, \lambda)$. Άρα η Πρόταση 13.11 εφαρμόζεται και δίνει ότι $\phi_X(t) = h(it) = \lambda/(\lambda - it)$ για κάθε $t \in \mathbb{R}$.

Συνέπεια της απόδειξης της Πρότασης 13.11 είναι το εξής θεώρημα μοναδικότητας για ροπογεννήτριες.

Θεώρημα 13.13 (Θεώρημα μοναδικότητας για ροπογεννήτριες). Έστω X, Y τυχαίες μεταβλητές με τιμές στον \mathbb{R} . Αν υπάρχει $\varepsilon > 0$ ώστε

(i) οι M_X, M_Y είναι πεπερασμένες στο $(-\varepsilon, \varepsilon)$ και

(ii) $M_X(t) = M_Y(t)$ για κάθε $t \in (-\varepsilon, \varepsilon)$,

τότε οι X, Y έχουν την ίδια κατανομή, δηλαδή $\mathbf{P}^X = \mathbf{P}^Y$.

Απόδειξη. Η υπόθεση (i) και η απόδειξη της Πρότασης 13.11 δίνουν ότι οι συναρτήσεις $g(z) = \mathbf{E}(e^{zX}), h(z) = \mathbf{E}(e^{zY})$ είναι αναλυτικές στο $A_\varepsilon := \{z \in \mathbb{C} : |\operatorname{Re}(z)| < \varepsilon\}$. Η υπόθεση (ii) και η αρχή αναλυτικής συνέχισης δίνουν ότι οι g, h ταυτίζονται στο A_ε . Άρα για $t \in \mathbb{R}$, έχουμε $\phi_X(t) = g(it) = h(it) = \phi_Y(t)$. Το συμπέρασμα έπεται από το Πόρισμα 13.6 (θεώρημα μοναδικότητας για χαρακτηριστικές συναρτήσεις). ■

Προσέξτε ότι το προηγούμενο θεώρημα ζητάει οι ροπογεννήτριες των X, Y να ταυτίζονται σε μια περιοχή του 0 (και να είναι πεπερασμένες σε αυτήν), ενώ το θεώρημα μοναδικότητας για χαρακτηριστικές συναρτήσεις (Πόρισμα 13.6) ζητάει ταύτισή τους σε όλο το \mathbb{R} . Ταύτιση των χαρακτηριστικών σε περιοχή του 0 δεν αρκεί για να δώσει ισότητα των κατανομών. Αντιπαραδείγματα δίνονται στην Παράγραφο 2α του Κεφαλαίου XV του Feller (1971).

13.6 Άθροισμα ανεξάρτητων τυχαίων μεταβλητών

Ορισμός 13.14. Έστω μ, ν μέτρα πιθανότητας στο \mathbb{R} . Συνέλιξη των μ, ν λέμε το μέτρο πιθανότητας $\mu * \nu$ στο \mathbb{R} που ορίζεται ως εξής:

$$\mu * \nu(A) := \int \int \mathbf{1}_A(x+y) d\mu(x) d\nu(y)$$

για κάθε $A \in \mathcal{B}(\mathbb{R})$.

Παρατήρηση 13.15. Εύκολα βλέπουμε ότι η συνέλιξη είναι συμμετρική, δηλαδή $\mu * \nu = \nu * \mu$. Επίσης, ισχύει ότι

$$\mu * \nu(A) = \int \int \mathbf{1}_{(A-y)}(x) d\mu(x) d\nu(y) = \int \mu(A-y) d\nu(y) = \int \nu(A-x) d\mu(x). \quad (13.11)$$

Θεώρημα 13.16. Έστω X, Y ανεξάρτητες τυχαίες μεταβλητές με τιμές στο \mathbb{R} και κατανομές $\mathbf{P}^X, \mathbf{P}^Y$ αντίστοιχα. Τότε $\mathbf{P}^{X+Y} = \mathbf{P}^X * \mathbf{P}^Y$.

Απόδειξη. Εφόσον οι X, Y είναι ανεξάρτητες, η κατανομή της (X, Y) είναι το μέτρο γινόμενο $\mathbf{P}^X \otimes \mathbf{P}^Y$ στον \mathbb{R}^2 (Πρόταση 10.12). Για $A \in \mathcal{B}(\mathbb{R})$ έχουμε

$$\begin{aligned} \mathbf{P}^{X+Y}(A) &= \mathbf{P}(X+Y \in A) = \mathbf{E}\{\mathbf{1}_A(X+Y)\} \\ &= \int \mathbf{1}_A(x+y) d(\mathbf{P}^X \otimes \mathbf{P}^Y)(x, y) = \int \int \mathbf{1}_A(x+y) d\mathbf{P}^X(x) d\mathbf{P}^Y(y) \\ &= \mathbf{P}^X * \mathbf{P}^Y(A). \end{aligned}$$

Στην πρώτη ισότητα της δεύτερης γραμμής χρησιμοποιήσαμε την Πρόταση 6.2 για τη συνάρτηση $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ με $g(x, y) = \mathbf{1}_A(x+y)$ και την τυχαία μεταβλητή (X, Y) . ■

Θεώρημα 13.17. Έστω X, Y ανεξάρτητες τυχαίες μεταβλητές με τιμές στο \mathbb{R} και $Z = X + Y$. Τότε:

(i) Αν η X έχει πυκνότητα f_X , τότε η Z έχει πυκνότητα και μια τέτοια είναι η

$$f_Z(z) = \int f_X(z-y) d\mathbf{P}^Y(y)$$

για κάθε $z \in \mathbb{R}$.

Χαρακτηριστικές συναρτήσεις

(ii) Αν οι X, Y έχουν αντίστοιχα πυκνότητες f_X, f_Y , τότε η

$$f_Z(z) = \int f_X(z-y)f_Y(y) dy = \int f_X(x)f_Y(z-x) dx$$

για κάθε $z \in \mathbb{R}$ είναι πυκνότητα της Z .

Απόδειξη. (i) Έστω $A \in \mathcal{B}(\mathbb{R})$. Τότε, από το Θεώρημα 13.16 και την (13.11), έχουμε

$$\begin{aligned} \mathbf{P}(Z \in A) &= \int \mathbf{P}^X(A-y) d\mathbf{P}^Y(y) = \int \int_{A-y} f_X(x) dx d\mathbf{P}^Y(y) \\ &\stackrel{x=z-y}{=} \int \int_A f_X(z-y) dz d\mathbf{P}^Y(y) = \int_A \int f_X(z-y) d\mathbf{P}^Y(y) dz, \end{aligned}$$

άρα η $f_Z(z) = \int f_X(z-y) d\mathbf{P}^Y(y)$ είναι πυκνότητα της Z .

(ii) Έστω $z \in \mathbb{R}$. Τότε, από το (i) και την Πρόταση 6.9, έχουμε

$$f_Z(z) = \int f_X(z-y) d\mathbf{P}^Y(y) = \int f_X(z-y)f_Y(y) dy.$$

Η δεύτερη ισότητα στην εκφώνηση προκύπτει με μια απλή αλλαγή μεταβλητής. ■

Το προηγούμενο θεώρημα συμπληρώνει την τεχνική προσδιορισμού κατανομής αθροίσματος που είδαμε στο Παράδειγμα 13.7. Το θεώρημα είναι χρήσιμο όταν η χαρακτηριστική συνάρτηση του αθροίσματος δεν είναι κάποια από τις γνωστές χαρακτηριστικές συναρτήσεις. Μια τέτοια περίπτωση περιγράφεται στην Άσκηση 13.14.

Άσκησης

13.1 Ναδειχθεί ότι ισχύει η ισότητα στην (13.1) αν και μόνο αν υπάρχει $a \in \mathbb{R}$ τέτοιο ώστε $f(x) = |f(x)|e^{ia}$ μ -σχεδόν παντού στο Ω .

13.2 Έστω X τυχαία μεταβλητή με τιμές στο \mathbb{R} .

(α) Αν υπάρχουν $a \in \mathbb{R}, b > 0$ ώστε η X να παίρνει τιμές στο $\{a + kb : k \in \mathbb{Z}\}$, ναδειχθεί ότι $|\phi_X(2\pi/b)| = 1$.

(β) Αν υπάρχει $t_0 \neq 0$ ώστε $|\phi_X(t_0)| = 1$, ναδειχθεί ότι υπάρχει $a \in \mathbb{R}$ ώστε με πιθανότητα 1 η X να παίρνει τιμές στο $\{a + k(2\pi/t_0) : k \in \mathbb{Z}\}$. Αν $\phi_X(t_0) = 1$, τότε μπορούμε να πάρουμε $a = 0$.

(γ)* Αν $|\phi_X(t)| = 1$ για κάθε t σε μια περιοχή του 0, τότε η X είναι σταθερή.

13.3 Έστω X τυχαία μεταβλητή με τιμές στο \mathbb{R} ώστε $\mathbf{E}|X| < \infty$. Ναδειχθεί ότι η ϕ_X είναι παραγωγίσιμη στο \mathbb{R} και $\phi'_X(t) = i\mathbf{E}(Xe^{itX})$ για κάθε $t \in \mathbb{R}$.

[Υπόδειξη: Θεώρημα κυριαρχημένης σύγκλισης.]

13.4 Έστω $X \sim \Gamma(a, \lambda)$ με $a, \lambda > 0$. Ναδειχθεί ότι η X έχει χαρακτηριστική συνάρτηση

$$\phi_X(t) = \frac{1}{(1 - \frac{it}{\lambda})^a}$$

για κάθε $t \in \mathbb{R}$.

13.5 Έστω $a > 0$ και X τυχαία μεταβλητή με πυκνότητα

$$f(x) = \frac{a}{2|x|^{a+1}} \mathbf{1}_{|x| \geq 1}.$$

Ναδειχθεί ότι $M_X(t) = \infty$ για κάθε $t \neq 0$.

Χαρακτηριστικές συναρτήσεις

13.6 Έστω X τυχαία μεταβλητή με κατανομή $N(0, 1)$ και $Y := e^X$. Ναδειχθεί ότι

(α) $\mathbf{E}(Y^k) < \infty$ για κάθε $k \in \mathbb{N}$.

(β) $M_Y(t) < \infty$ αν και μόνο αν $t \leq 0$.

Δηλαδή η ύπαρξη όλων των ροπών δεν συνεπάγεται ότι η ροπογεννήτρια είναι πεπερασμένη σε περιοχή του 0. Η κατανομή της Y λέγεται λογαριθμοκανονική (έτσι λέγεται κάθε κατανομή τυχαίας μεταβλητής Y της οποίας ο λογάριθμος, $\log Y$, ακολουθεί κάποια κανονική κατανομή).

13.7 Έστω X τυχαία μεταβλητή με τιμές στο \mathbb{R} . Να δείξετε ότι η X έχει συμμετρική κατανομή (δηλαδή $X \stackrel{d}{=} -X$) αν και μόνο αν $\phi_X(u) \in \mathbb{R}$ για κάθε $u \in \mathbb{R}$.

13.8 Έστω X, Y ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με τιμές στο \mathbb{R} . Να δείξετε ότι η $X - Y$ έχει συμμετρική κατανομή.

13.9 Έστω $n \in \mathbb{N}^+$ και $X_1, X_2, \dots, X_n \sim \text{Cauchy}$ ανεξάρτητες. Δείξτε ότι $\frac{X_1 + X_2 + \dots + X_n}{n} \sim \text{Cauchy}$.

13.10 Έστω $\alpha \in (0, 2]$. Είναι γνωστό³ ότι υπάρχει τυχαία μεταβλητή X με χαρακτηριστική συνάρτηση την $\phi_X(t) = e^{-|t|^\alpha}$ για κάθε $t \in \mathbb{R}$. Έστω τώρα n θετικός ακέραιος και X_1, X_2, \dots, X_n ανεξάρτητες και ισόνομες τυχαίες μεταβλητές καθεμία με κατανομή την ίδια με αυτήν της X . Ναδειχθεί ότι $X_1 + X_2 + \dots + X_n \stackrel{d}{=} n^{1/\alpha} X_1$.

13.11 Έστω $(X_n)_{n \geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με $\mathbf{P}(X_1 = -1) = \mathbf{P}(X_1 = 1) = 1/2$. Ναδειχθεί ότι η σειρά $\sum_{n=1}^{\infty} \frac{X_n}{2^n}$ συγκλίνει με πιθανότητα 1 σε μια τυχαία μεταβλητή με κατανομή την ομοιόμορφη στο $(-1, 1)$. Σημείωση: Εναλλακτική διατύπωση αυτού του αποτελέσματος είναι ότι αν οι $(X_n)_{n \geq 1}$ είναι ανεξάρτητες ισόνομες και καθεμία έχει την ομοιόμορφη κατανομή στο $\{0, 1\}$, τότε η σειρά συγκλίνει σχεδόν βέβαια σε μια τυχαία μεταβλητή με κατανομή την ομοιόμορφη στο $(0, 1)$. Αυτό ζητείται στην Άσκηση 12.8(α).

13.12 Έστω X τυχαία μεταβλητή όπως στην Άσκηση 13.5 με $\alpha \in (0, 2)$. Ναδειχθεί ότι υπάρχει σταθερά $C(\alpha) \in (0, \infty)$ ώστε

$$\lim_{t \rightarrow 0} \frac{1 - \phi_X(t)}{|t|^\alpha} = C(\alpha). \quad (13.12)$$

13.13* Έστω X τυχαία μεταβλητή με πυκνότητα $f(x) = |x|^{-3} \mathbf{1}_{|x| \geq 1}$. Ναδειχθεί ότι για κάθε $t \in [-1, 1]$ ισχύει

$$|\phi_X(t) - 1 - t^2 \log |t|| \leq 3t^2. \quad (13.13)$$

13.14 Έστω $X, Y \sim U(0, 1)$ ανεξάρτητες τυχαίες μεταβλητές. Να δείξετε ότι η $Z := X + Y$ έχει πυκνότητα

$$f_Z(z) = \begin{cases} z & \text{αν } z \in (0, 1), \\ 2 - z & \text{αν } z \in (1, 2), \\ 0 & \text{αν } z \in \mathbb{R} \setminus (0, 2). \end{cases}$$

13.15 (α) Έστω ότι η X έχει κατανομή με πυκνότητα $f(x) = e^{-|x|}/2$ για κάθε $x \in \mathbb{R}$. Η κατανομή αυτή λέγεται δίπλευρη εκθετική. Να υπολογιστεί η χαρακτηριστική συνάρτηση της X .

(β) Με χρήση του (α) και του Πορίσματος 13.8 να υπολογιστεί η χαρακτηριστική συνάρτηση της κατανομής Cauchy.

13.16 Έστω X τυχαία μεταβλητή με τιμές στο $[0, \infty)$ και με $\mu := \mathbf{E}(X) \in (0, \infty)$ και X^* η μεροληπτική με βάση το μέγεθος εκδοχή της X [δες Άσκηση 7.7].

(α) Ναδειχθεί ότι για κάθε $t \in \mathbb{R}$ ισχύει

$$\phi_{X^*}(t) = -\frac{i}{\mu} \phi'_X(t)$$

(β) Ναδειχθεί ότι $X^* \stackrel{d}{=} X + 1$ αν και μόνο αν $X \sim \text{Poisson}(\mu)$.

13.17 Έστω X τυχαία μεταβλητή με τιμές στο \mathbb{R} ώστε $M_X(t) < \infty$ για κάθε $t \in \mathbb{R}$. Με χρήση παραγώγων ναδειχθεί ότι η $\log M_X$ είναι κυρτή στο \mathbb{R} .

³Έπεται από την Άσκηση 15.10.

14

Σύγκλιση κατά κατανομή

14.1 Σύγκλιση κατά κατανομή

Στην παράγραφο αυτή θα μελετήσουμε μια ασθενέστερη, από όσες έχουμε δει έως τώρα, μορφή σύγκλισης, τη σύγκλιση κατά κατανομή. Θα θεωρήσουμε μέτρα πιθανότητας στο \mathbb{R} και τυχαίες μεταβλητές με τιμές στο \mathbb{R} .

Ορισμός 14.1. Έστω $\mu, (\mu_n)_{n \geq 1}$ μέτρα πιθανότητας στον $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Λέμε ότι η $(\mu_n)_{n \geq 1}$ **συγκλίνει ασθενώς** στο μ αν

$$\mu_n((-\infty, x]) \rightarrow \mu((-\infty, x]) \quad (14.1)$$

καθώς $n \rightarrow \infty$ για κάθε $x \in \mathbb{R}$ τέτοιο ώστε $\mu(\{x\}) = 0$. Γράφουμε τότε $\mu_n \Rightarrow \mu$.

Το κατά κατανομή όριο μιας ακολουθίας μέτρων που συγκλίνει κατά κατανομή είναι μοναδικό. Γιατί αν η ακολουθία συγκλίνει σε δύο μέτρα μ, ν , τότε οι συναρτήσεις κατανομής τους είναι ίσες στο $\{x \in \mathbb{R} : \mu(\{x\}) = \nu(\{x\}) = 0\}$ το οποίο έχει αριθμήσιμο συμπλήρωμα (το δείχνουμε με χρήση της Άσκησης 2.5), άρα είναι πυκνό. Και επειδή οι συναρτήσεις κατανομής είναι δεξιά συνεχείς, έπεται ότι ισούνται σε όλο το \mathbb{R} . Έτσι, το Θεώρημα 7.6 δίνει ότι $\mu = \nu$.

Ορισμός 14.2. Έστω $(X_n)_{n \geq 1}, X$ τυχαίες μεταβλητές με τιμές στο \mathbb{R} . Λέμε ότι η $(X_n)_{n \geq 1}$ **συγκλίνει κατά κατανομή** στη X και γράφουμε¹

$$X_n \Rightarrow X \text{ ή } X_n \xrightarrow{d} X \text{ ή } X_n \xrightarrow{\mathcal{L}} X,$$

αν η ακολουθία κατανομών $(\mathbf{P}^{X_n})_{n \geq 1}$ των X_n συγκλίνει ασθενώς στην κατανομή \mathbf{P}^X της X .

Το πιο πάνω σχόλιο για τη μοναδικότητα του ορίου κατά κατανομή, με όρους τυχαίων μεταβλητών, λέει ότι, αν η $(X_n)_{n \geq 1}$ συγκλίνει κατά κατανομή σε δύο τυχαίες μεταβλητές X, Y , τότε οι X, Y έχουν την ίδια κατανομή.

Θεώρημα 14.3. Έστω $(X_n)_{n \geq 1}, X$ όπως στον Ορισμό 14.2. Τότε $X_n \Rightarrow X$ αν και μόνο αν

$$F_{X_n}(x) \rightarrow F_X(x)$$

καθώς $n \rightarrow \infty$ για κάθε $x \in \mathbb{R}$ τέτοιο ώστε $F_X(x) = F_X(x-)$, δηλαδή για κάθε σημείο συνέχειας της F_X .

Απόδειξη. Προκύπτει από τον Ορισμό 14.1, τον ορισμό της συνάρτησης κατανομής, και το ότι $\mathbf{P}^X(\{x\}) = \mathbf{P}(X = x) = F_X(x) - F_X(x-)$ για κάθε $x \in \mathbb{R}$. ■

Μια συνάρτηση κατανομής F έχει αριθμήσιμο πλήθος σημείων ασυνέχειας (Άσκηση 7.2). Δηλαδή αυτά τα σημεία είναι λίγα. Σε οποιοδήποτε διάστημα θετικού μήκους μπορούμε να βρούμε σημείο συνέχειας της F .

¹d από το distribution, και \mathcal{L} από το law.

Παρατήρηση 14.4. (i) Στον Ορισμό 14.2, οι $X, \{X_n : n \geq 1\}$, δεν είναι απαραίτητο να ορίζονται στον ίδιο χώρο πιθανότητας. Κάθε μία ορίζεται σε χώρο πιθανότητας $(\Omega_n, \mathcal{F}_n, \mathbf{P}_n)$ και η X σε χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$. Αυτό θα δημιουργούσε πρόβλημα αν θέλαμε να θεωρήσουμε τη διαφορά $X_n(\omega) - X(\omega)$.

(ii) Αν οι $\{X_n : n \geq 1\}$ και X ορίζονται στον ίδιο χώρο πιθανότητας, έχει νόημα να εξετάσουμε πώς η σύγκλιση κατά κατανομή συνδέεται με τα υπόλοιπα είδη σύγκλισης που είδαμε στο Κεφάλαιο 8 (σχεδόν βέβαιη, στον \mathcal{L}^p , κατά πιθανότητα). Το Θεώρημα 14.14 πιο κάτω αφορά αυτό το ερώτημα.

Παράδειγμα 14.5. Έστω $(p_n)_{n \geq 1}$ ακολουθία στο $(0, 1)$ έτσι ώστε $p_n \rightarrow 0$ και $(X_n)_{n \geq 1}$ ακολουθία τυχαίων μεταβλητών έτσι ώστε $X_n \sim \text{Γεωμετρική}(p_n)$ για κάθε $n \geq 1$ (δηλαδή $\mathbf{P}(X_n = k) = (1 - p_n)^{k-1} p_n$ για $k = 1, 2, \dots$). Θα δείξουμε ότι

$$p_n X_n \Rightarrow X,$$

όπου X τυχαία μεταβλητή με κατανομή $\text{Exp}(1)$.

Η F_X είναι συνεχής στο \mathbb{R} . Συγκεκριμένα, έχουμε ότι

$$F_X(x) = \int_{-\infty}^x e^{-t} \mathbf{1}_{t>0} dt = \begin{cases} 0 & \text{αν } x \leq 0, \\ 1 - e^{-x} & \text{αν } x > 0. \end{cases}$$

Για $x \leq 0$,

$$F_{p_n X_n}(x) = \mathbf{P}(p_n X_n \leq 0) = 0,$$

αφού $p_n > 0$ και $X_n \geq 1$. Για $x > 0$,

$$\begin{aligned} F_{p_n X_n}(x) &= \mathbf{P}\left(X_n \leq \frac{x}{p_n}\right) = 1 - \mathbf{P}\left(X_n > \frac{x}{p_n}\right) \\ &= 1 - \mathbf{P}\left(X_n > \left\lceil \frac{x}{p_n} \right\rceil\right) = 1 - (1 - p_n)^{\lfloor \frac{x}{p_n} \rfloor}, \end{aligned}$$

εφόσον για $Y \sim \text{Γεωμετρική}(p)$ ισχύει $\mathbf{P}(Y \geq k) = (1 - p)^{k-1}$. Τώρα παρατηρούμε ότι

$$(1 - p_n)^{\lfloor \frac{x}{p_n} \rfloor} = e^{\lfloor \frac{x}{p_n} \rfloor p_n \frac{\log(1-p_n)}{p_n}} \rightarrow e^{-x}$$

για $n \rightarrow \infty$. Άρα, $F_{p_n X_n}(x) \rightarrow 1 - e^{-x}$. Από το Θεώρημα 14.3 έπεται το ζητούμενο.

Παρατηρήστε ότι η τυχαία μεταβλητή X_n παίρνει τιμές στο \mathbb{N}^+ και άρα η $p_n X_n$ στο $p_n \mathbb{N}^+$, το οποίο είναι ένα σύνολο που απλώνεται στο $[0, \infty)$ με όλο και πιο πυκνό τρόπο καθώς $n \rightarrow \infty$. Κατά μια έννοια, αυτό το σύνολο συγκλίνει στο $[0, \infty)$, το οποίο είναι το στήριγμα της κατανομής της X .

Παράδειγμα 14.6. Έστω $(X_n)_{n \geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με καθεμία να ακολουθεί την εκθετική κατανομή με παράμετρο 1. Για κάθε $n \geq 1$ θετικό ακέραιο, θέτουμε $W_n := \max\{X_1, X_2, \dots, X_n\}$ και $Z_n := W_n - \log n$. Θα δείξουμε ότι

$$W_n - \log n \Rightarrow Z \tag{14.2}$$

όπου Z είναι τυχαία μεταβλητή με συνάρτηση κατανομής $F_Z(t) := e^{-e^{-t}}$ ($t \in \mathbb{R}$).

Η F_Z είναι συνεχής στο \mathbb{R} . Έστω λοιπόν $t \in \mathbb{R}$. Για $n > e^{-t}$ έχουμε

$$\begin{aligned} F_{W_n - \log n}(t) &:= \mathbf{P}(W_n - \log n \leq t) = \mathbf{P}(\max\{X_1, X_2, \dots, X_n\} \leq t + \log n) \\ &= \mathbf{P}(X_1 \leq t + \log n, \dots, X_n \leq t + \log n) \\ &= \mathbf{P}(X_1 \leq t + \log n)^n = (1 - e^{-t - \log n})^n = \left(1 - \frac{e^{-t}}{n}\right)^n. \end{aligned}$$

Στην τέταρτη ισότητα χρησιμοποιούμε ότι οι X_i είναι ανεξάρτητες και ισόνομες. Στην πέμπτη ισότητα ότι $t + \log n > 0$ και ότι η X_1 ακολουθεί την εκθετική κατανομή με παράμετρο 1. Άρα $\lim_{n \rightarrow \infty} F_{W_n - \log n}(t) = e^{-e^{-t}} = F_Z(t)$ για κάθε $t \in \mathbb{R}$, που αποδεικνύει το ζητούμενο.

Αυτό που μας λέει η σύγκλιση (14.2) είναι ότι το μέγιστο n ανεξάρτητων εκθετικών τυχαίων μεταβλητών με παράμετρο 1 είναι της τάξης του $\log n$. Δηλαδή ισούται με $\log n$ συν μια διόρθωση που είναι περίπου μια τυχαία μεταβλητή όπως η Z . Αυτή η διόρθωση δεν αλλάζει την τάξη μεγέθους, γιατί, π.χ., με πιθανότητα 0.95 ισχύει $|Z| \leq 3$ (υπολογίζουμε $\mathbf{P}(|Z| \leq 3) = F_Z(3) - F_Z(-3) = e^{-e^{-3}} - e^{-e^3}$).

Παρατήρηση 14.7. [Ισοκατανεμημένες τυχαίες μεταβλητές.] Συνεχίζοντας από την Παρατήρηση 6.5, στο προηγούμενο παράδειγμα, αφού όλες οι X_1, X_2, \dots, X_n έχουν την ίδια κατανομή με τη X_1 , γιατί στην ποσότητα

$$\mathbf{P}(X_1 \leq t + \log n, \dots, X_n \leq t + \log n)$$

δεν αντικαθιστούμε όλες τις τυχαίες μεταβλητές με τη X_1 ; Και τότε η ποσότητα θα γινόταν $\mathbf{P}(X_1 \leq t + \log n)$. Αυτό βέβαια δεν είναι σωστό και δεν μπορεί να δικαιολογηθεί με χρήση της Πρότασης 6.2 γιατί, για να είναι η χρήση της σωστή, θα έπρεπε τα δύο διανύσματα (X_1, X_2, \dots, X_n) και (X_1, X_1, \dots, X_1) να έχουν την ίδια κατανομή, πράγμα το οποίο δεν ισχύει.

Όταν όμως με χρήση της ανεξαρτησίας γράψουμε την πιο πάνω πιθανότητα ως

$$\mathbf{P}(X_1 \leq t + \log n) \cdots \mathbf{P}(X_n \leq t + \log n),$$

τότε κάθε όρος του γινομένου ισούται με $\mathbf{P}(X_1 \leq t + \log n)$ επειδή κάθε μία από τις X_i έχει την ίδια κατανομή με τη X_1 .

Για την ανάπτυξη της θεωρίας είναι πιο βολικό αντί να δουλεύουμε με τον ορισμό της ασθενούς σύγκλισης να χρησιμοποιούμε τον χαρακτηρισμό της που δίνεται από το επόμενο θεώρημα. Η απόδειξή του δίνεται στο Παράρτημα Β'.

Θεώρημα 14.8. Έστω $\mu, (\mu_n)_{n \geq 1}$ μέτρα πιθανότητας στο \mathbb{R} . Τότε $\mu_n \Rightarrow \mu$ αν και μόνο αν

$$\int f(x) d\mu_n(x) \rightarrow \int f(x) d\mu(x) \quad (14.3)$$

για κάθε $f : \mathbb{R} \rightarrow \mathbb{R}$ συνεχή και φραγμένη.

Το ίδιο θεώρημα, με όρους τυχαίων μεταβλητών, γράφεται ως εξής.

Θεώρημα 14.9. Έστω $X, (X_n)_{n \geq 1}$ τυχαίες μεταβλητές με τιμές στο \mathbb{R} . Τότε $X_n \Rightarrow X$ αν και μόνο αν

$$\mathbf{E}\{f(X_n)\} \rightarrow \mathbf{E}\{f(X)\} \quad (14.4)$$

για κάθε $f : \mathbb{R} \rightarrow \mathbb{R}$ συνεχή και φραγμένη.

Πιο αναλυτικά, έχουμε $(\Omega, \mathcal{F}, \mathbf{P}), (\Omega_n, \mathcal{F}_n, \mathbf{P}_n), n \geq 1$ χώρους πιθανότητας και τυχαίες μεταβλητές $X : \Omega \rightarrow \mathbb{R}, X_n : \Omega_n \rightarrow \mathbb{R}$ για κάθε $n \geq 1$. Η μέση τιμή στο αριστερό μέλος της (14.4) είναι ως προς το μέτρο \mathbf{P}_n , ενώ στο δεξί ως προς το μέτρο \mathbf{P} .

Το προηγούμενο θεώρημα είναι περισσότερο χρήσιμο για την απόδειξη θεωρητικών αποτελεσμάτων και όχι για αποδείξεις σύγκλισης κατά κατανομή ακολουθιών τυχαίων μεταβλητών που εμφανίζονται συχνά. Μία από τις εξαιρέσεις είναι η ακόλουθη.

Παράδειγμα 14.10. Έστω $(X_n)_{n \geq 1}$ ακολουθία τυχαίων μεταβλητών ώστε η X_n να είναι ομοιόμορφη διακριτή στο σύνολο $\{1, 2, \dots, n\}$. Θα δείξουμε ότι

$$\frac{1}{n}X_n \Rightarrow U \quad (14.5)$$

καθώς $n \rightarrow \infty$, όπου $U \sim U(0, 1)$.

Για $f : \mathbb{R} \rightarrow \mathbb{R}$ συνεχή και φραγμένη και $n \rightarrow \infty$ έχουμε

$$\mathbf{E} \left\{ f \left(\frac{X_n}{n} \right) \right\} = \sum_{k=1}^n \mathbf{P}(X_n = k) f \left(\frac{k}{n} \right) = \frac{1}{n} \sum_{k=1}^n f \left(\frac{k}{n} \right) \rightarrow \int_0^1 f(x) dx = \mathbf{E}\{f(U)\}$$

Η σύγκλιση ισχύει γιατί έχουμε ένα άθροισμα Riemann για την f στο $[0, 1]$. Η τελευταία ισότητα ισχύει γιατί η U έχει πυκνότητα $1_{(0,1)}$.

Η σύγκλιση (14.5) είναι αναμενόμενη αφού η κατανομή της X_n/n δίνει μάζα $1/n$ σε καθένα από τα σημεία $\{1/n, 2/n, \dots, n/n\}$. Η μάζα ισομοιράζεται και τελικά, καθώς $n \rightarrow \infty$, κάθε υποδιάστημα του $[0, 1]$ παίρνει μάζα ανάλογη προς το μέγεθος του.

Αν η ακολουθία τυχαίων μεταβλητών $(X_n)_{n \geq 1}$ συγκλίνει στην τυχαία μεταβλητή X κατά πιθανότητα [ή σχεδόν βέβαια] και $g : \mathbb{R} \rightarrow \mathbb{R}$ είναι συνεχής συνάρτηση, έχουμε δει [Πρόταση 8.6] ότι η $g(X_n)$ συγκλίνει στην $g(X)$ κατά πιθανότητα [σχεδόν βέβαια αντίστοιχα]. Δηλαδή κάθε συνεχής συνάρτηση διατηρεί τη σχεδόν βέβαια σύγκλιση και τη σύγκλιση κατά πιθανότητα. Όμοια, διατηρεί και τη σύγκλιση κατά κατανομή.

Πόρισμα 14.11 (Θεώρημα συνεχούς απεικόνισης). Έστω $X, (X_n)_{n \geq 1}$ τυχαίες μεταβλητές με τιμές στο \mathbb{R} για τις οποίες ισχύει $X_n \Rightarrow X$. Αν η $g : \mathbb{R} \rightarrow \mathbb{R}$ είναι συνεχής, τότε $g(X_n) \Rightarrow g(X)$.

Απόδειξη. Με βάση το Θεώρημα 14.9, αρκεί να δείξουμε ότι για κάθε $f : \mathbb{R} \rightarrow \mathbb{R}$ συνεχή ισχύει

$$\lim_{n \rightarrow \infty} \mathbf{E}\{f(g(X_n))\} = \mathbf{E}\{f(g(X))\}.$$

Και πράγματι, αυτό ισχύει γιατί $X_n \Rightarrow X$ και η $f \circ g$ είναι συνεχής. ■

Δίνουμε ακόμη έναν χαρακτηρισμό της ασθενούς σύγκλισης. Αυτός είναι χρήσιμος στις εφαρμογές.

Θεώρημα 14.12. Έστω $\mu, (\mu_n)_{n \geq 1}$ μέτρα πιθανότητας στο \mathbb{R} . Τότε τα εξής είναι ισοδύναμα.

- (i) $\mu_n \Rightarrow \mu$.
- (ii) Για κάθε $A \in \mathcal{B}(\mathbb{R})$ με $\mu(\partial A) = 0$ ισχύει

$$\lim_{n \rightarrow \infty} \mu_n(A) = \mu(A).$$

Απόδειξη. (ii) \Rightarrow (i). Αυτή η κατεύθυνση είναι εύκολη. Αν το $x \in \mathbb{R}$ είναι σημείο με $\mu(\{x\}) = 0$, τότε εφαρμόζοντας την υπόθεση για το σύνολο $A = (-\infty, x]$, το οποίο έχει $\partial A = \{x\}$, παίρνουμε την (14.1), που είναι το ζητούμενο.

- (i) \Rightarrow (ii). Δίνεται στο Παράρτημα Β'.

Άμεση συνέπεια του προηγούμενου θεωρήματος είναι ο εξής εναλλακτικός χαρακτηρισμός για τη σύγκλιση κατά κατανομή.

Θεώρημα 14.13. Έστω $X, (X_n)_{n \geq 1}$ τυχαίες μεταβλητές όπως στο Θεώρημα 14.9. Τότε τα εξής είναι ισοδύναμα.

(i) $X_n \Rightarrow X$

(ii) Για κάθε $A \in \mathcal{B}(\mathbb{R})$ με $\mathbf{P}(X \in \partial A) = 0$ ισχύει

$$\lim_{n \rightarrow \infty} \mathbf{P}_n(X_n \in A) = \mathbf{P}(X \in A).$$

Το επόμενο θεώρημα δείχνει ότι η σύγκλιση κατά κατανομή είναι η ασθενέστερη μορφή σύγκλισης τυχαίων μεταβλητών από όσες έχουμε δει ως τώρα.

Θεώρημα 14.14. Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας και $(X_n)_{n \geq 1}, X$ τυχαίες μεταβλητές ορισμένες σε αυτόν και με τιμές στο \mathbb{R} . Αν $X_n \xrightarrow{\mathbf{P}} X$, τότε $X_n \Rightarrow X$.

Απόδειξη. Έστω ότι $X_n \not\Rightarrow X$. Τότε υπάρχουν $f : \mathbb{R} \rightarrow \mathbb{R}$ συνεχής και φραγμένη, $\varepsilon > 0$, και υπακολουθία $(X_{n_k})_{k \in \mathbb{N}}$ της $(X_n)_{n \geq 1}$ έτσι ώστε:

$$|\mathbf{E}(f(X_{n_k})) - \mathbf{E}(f(X))| \geq \varepsilon. \quad (14.6)$$

Επειδή $X_{n_k} \xrightarrow{\mathbf{P}} X$, από το Θεώρημα 8.4 υπάρχει υπακολουθία $(X_{n_{k_r}})_{r \in \mathbb{N}}$ της $(X_{n_k})_{k \in \mathbb{N}}$ έτσι ώστε $X_{n_{k_r}} \xrightarrow{\sigma.β.} X$. Αφού η f είναι συνεχής, από την Πρόταση 8.6 έχουμε ότι $f(X_{n_{k_r}}) \xrightarrow{\sigma.β.} f(X)$. Η f είναι φραγμένη, άρα υπάρχει $M > 0$ έτσι ώστε $|f(x)| \leq M$ για κάθε $x \in \mathbb{R}$. Επομένως, έχουμε $|f(X_{n_{k_r}})| \leq M$ για κάθε $r \in \mathbb{N}$ (και κάθε $\omega \in \Omega$). Από το θεώρημα φραγμένης σύγκλισης,

$$\mathbf{E}(f(X_{n_{k_r}})) \rightarrow \mathbf{E}(f(X)),$$

το οποίο συγκρούεται με την (14.6). Άρα $X_n \Rightarrow X$. ■

Μία περίπτωση κατά την οποία η σύγκλιση κατά κατανομή συνεπάγεται τη σύγκλιση κατά πιθανότητα είναι εκείνη κατά την οποία το όριο είναι μια σταθερή τυχαία μεταβλητή.

Θεώρημα 14.15. Έστω $(X_n)_{n \geq 1}$ ακολουθία τυχαίων μεταβλητών ορισμένων σε κοινό χώρο πιθανότητας, με τιμές στο \mathbb{R} και $C \in \mathbb{R}$. Αν $X_n \Rightarrow C$, τότε $X_n \xrightarrow{\mathbf{P}} C$.

Απόδειξη. Έστω $\varepsilon > 0$. Τότε

$$\begin{aligned} \mathbf{P}(|X_n - C| > \varepsilon) &= \mathbf{P}(X_n > C + \varepsilon) + \mathbf{P}(X_n < C - \varepsilon) \\ &= 1 - F_{X_n}(C + \varepsilon) + \mathbf{P}(X_n < C - \varepsilon) \\ &\leq 1 - F_{X_n}(C + \varepsilon) + F_{X_n}(C - \varepsilon). \end{aligned}$$

Τα $C - \varepsilon, C + \varepsilon$ είναι σημεία συνέχειας της F_C [αφού $F_C(x) = \mathbf{1}_{[C, \infty)}(x)$], άρα από το Θεώρημα 14.3,

$$\lim_{n \rightarrow \infty} F_{X_n}(C - \varepsilon) = F_C(C - \varepsilon) = 0$$

και

$$\lim_{n \rightarrow \infty} F_{X_n}(C + \varepsilon) = F_C(C + \varepsilon) = 1.$$

Συνεπώς $\lim_{n \rightarrow \infty} \mathbf{P}(|X_n - C| > \varepsilon) = 0$. ■

14.2 Η αναπαράσταση Skorokhod

Η σχεδόν βέβαιη σύγκλιση συνεπάγεται τη σύγκλιση κατά κατανομή (συνέπεια των Θεωρημάτων 8.2(ii), 14.14 και εναλλακτικά του Θεωρήματος 14.9 και του θεωρήματος φραγμένης σύγκλισης). Δηλαδή, αν $X, (X_n)_{n \geq 1}$ είναι τυχαίες μεταβλητές ορισμένες στον ίδιο χώρο πιθανότητας και με τιμές στο \mathbb{R} με $\lim_{n \rightarrow \infty} X_n = X$ με πιθανότητα 1, τότε $X_n \Rightarrow X$.

Το θεώρημα αναπαράστασης του Skorokhod λέει ότι κάθε σύγκλιση κατά κατανομή προκύπτει από μια σχεδόν βέβαιη σύγκλιση. Σημειώστε όμως ότι στη σύγκλιση κατά κατανομή οι τυχαίες μεταβλητές δεν είναι απαραίτητο να ορίζονται στον ίδιο χώρο πιθανότητας.

Θεώρημα 14.16. Έστω $(X_n)_{n \geq 1}, X$ τυχαίες μεταβλητές με τιμές στο \mathbb{R} ώστε $X_n \Rightarrow X$. Τότε υπάρχει χώρος πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$ και τυχαίες μεταβλητές $Y_n, Y : \Omega \rightarrow \mathbb{R}$ ώστε

$$(i) \ Y \stackrel{d}{=} X \text{ και } Y_n \stackrel{d}{=} X_n \text{ για κάθε } n \geq 1.$$

$$(ii) \ \lim_{n \rightarrow \infty} Y_n = Y \text{ με πιθανότητα } 1.$$

Απόδειξη. Έστω $(F_n)_{n \geq 1}, F$ οι συναρτήσεις κατανομής των $(X_n)_{n \geq 1}, X$ και $(Q_n)_{n \geq 1}, Q$ οι μετασχηματισμοί ποσοστημοριών τους. Θεωρούμε τον χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P}) = ((0, 1), \mathcal{B}((0, 1)), \lambda)$ και την τυχαία μεταβλητή $U : \Omega \rightarrow \mathbb{R}$ με $U(x) = x$. Η U έχει την ομοιόμορφη κατανομή. Θεωρούμε τις τυχαίες μεταβλητές $Y_n, Y : \Omega \rightarrow \mathbb{R}$ με

$$Y_n := Q_n(U),$$

$$Y := Q(U).$$

Με βάση την Πρόταση 7.10, οι $(Y_n)_{n \geq 1}, Y$ ικανοποιούν το (i) του θεωρήματος.

Θα δείξουμε τώρα το εξής.

ΙΣΧΥΡΙΣΜΟΣ: $\lim_{n \rightarrow \infty} Q_n(t) = Q(t)$ για κάθε $t \in (0, 1) \setminus D$ όπου D είναι ένα αριθμήσιμο σύνολο.

Έστω

$$D := \{t \in (0, 1) : F(a) = F(b) = t \text{ για κάποια } a < b\}.$$

Το D είναι ακριβώς οι τιμές στο $(0, 1)$ τις οποίες παίρνει η F σε διαστήματα που είναι σταθερή και είναι αριθμήσιμο [για κάθε $t \in D$, επιλέγουμε έναν ρητό σε ένα διάστημα (a, b) που αντιστοιχεί στο t]. Επίσης είναι εύκολο να δει κανείς ότι είναι το σύνολο των σημείων ασυνέχειας της Q (Άσκηση), αλλά δεν θα το χρειαστούμε αυτό στην απόδειξη.

Έστω τώρα $t \in (0, 1) \setminus D$. Παρατηρούμε ότι $F(y) < t$ για $y < Q(t)$ και $F(y) > t$ για $y > Q(t)$ [για το δεύτερο, χρησιμοποιούμε τον ορισμό του D].

Αν $y > Q(t)$ είναι σημείο συνέχειας της F , τότε $F(y) > t$ και η $\lim_{n \rightarrow \infty} F_n(y) = F(y) > t$ δίνει ότι $Q_n(t) \leq y$ για όλα τα μεγάλα n . Άρα $\lim_{n \rightarrow \infty} Q_n(t) \leq y$. Και επειδή μπορούμε να επιλέξουμε το y αυθαίρετα κοντά στο $Q(t)$ [τα σημεία ασυνέχειας της F είναι αριθμήσιμο σύνολο], έχουμε $\lim_{n \rightarrow \infty} Q_n(t) \leq Q(t)$.

Αν $y < Q(t)$ είναι σημείο συνέχειας της F , τότε $F(y) < t$ και η $\lim_{n \rightarrow \infty} F_n(y) = F(y) < t$ δίνει ότι $Q_n(t) \geq y$ για όλα τα μεγάλα n . Άρα $\lim_{n \rightarrow \infty} Q_n(t) \geq y$. Και επειδή μπορούμε να επιλέξουμε το y αυθαίρετα κοντά στο $Q(t)$, έχουμε $\lim_{n \rightarrow \infty} Q_n(t) \geq Q(t)$. Ο ισχυρισμός αποδείχθηκε.

Επειδή $\mathbf{P}(D) = \lambda(D) = 0$ (αφού το D είναι αριθμήσιμο) και $\lim_{n \rightarrow \infty} Q_n(U(\omega)) = Q(U(\omega))$ για κάθε $\omega \in (0, 1) \setminus D$, δείχθηκε και το (ii) του θεωρήματος. ■

Είμαστε τώρα σε θέση να διατυπώσουμε μια ισχυροποίηση του Πορίσματος 14.11 η οποία είναι απαραίτητη πολύ συχνά.

Πόρισμα 14.17 (Θεώρημα συνεχούς απεικόνισης). Έστω $X, (X_n)_{n \geq 1}$ τυχαίες μεταβλητές με τιμές στο \mathbb{R} για τις οποίες ισχύει $X_n \Rightarrow X$. Αν $g : \mathbb{R} \rightarrow \mathbb{R}$ μετρήσιμη ώστε $\mathbf{P}(X \in D_g) = 0$, όπου $D_g := \{x \in \mathbb{R} : g \text{ ασυνεχής στο } x\}$, τότε $g(X_n) \Rightarrow g(X)$.

Απόδειξη. Θα χρησιμοποιήσουμε το Θεώρημα 14.9. Έστω $f : \mathbb{R} \rightarrow \mathbb{R}$ συνεχής και φραγμένη και $(Y_n)_{n \geq 1}, Y$ τυχαίες μεταβλητές όπως στο Θεώρημα 14.16. Τότε $\lim_{n \rightarrow \infty} f(g(Y_n)) = f(g(Y))$ σχεδόν βέβαια γιατί $Y_n \rightarrow Y$ σχεδόν βέβαια, $\mathbf{P}(Y \in D_g) = 0$, και η f είναι συνεχής. Έτσι, το θεώρημα φραγμένης σύγκλισης δίνει ότι

$$\lim_{n \rightarrow \infty} \mathbf{E}\{f(g(X_n))\} = \lim_{n \rightarrow \infty} \mathbf{E}\{f(g(Y_n))\} = \mathbf{E}\{f(g(Y))\} = \mathbf{E}\{f(g(X))\}. \quad \blacksquare$$

Αν $X_n \Rightarrow X$, τότε δεν είναι απαραίτητο να ισχύει $\lim_{n \rightarrow \infty} \mathbf{E}(X_n^2) = \mathbf{E}(X^2)$ [ή $\lim_{n \rightarrow \infty} \mathbf{E}(X_n) = \mathbf{E}(X)$] γιατί η συνάρτηση $x \mapsto x^2$ είναι συνεχής αλλά δεν είναι φραγμένη ώστε να εφαρμόσουμε το Θεώρημα 14.9. Επέκταση αυτού του θεωρήματος σε μη φραγμένες συναρτήσεις είναι το επόμενο πόρισμα της αναπαράστασης Skorokhod.

Πόρισμα 14.18. Έστω $(X_n)_{n \geq 1}, X$ τυχαίες μεταβλητές με τιμές στο \mathbb{R} για τις οποίες ισχύει $X_n \Rightarrow X$. Υποθέτουμε ότι οι $f, g : \mathbb{R} \rightarrow \mathbb{R}$ είναι συνεχείς, $\lim_{|x| \rightarrow \infty} f(x)/g(x) = 0$, και $\sup_{n \geq 1} \mathbf{E}|g(X_n)| = A < \infty$. Τότε

$$\lim_{n \rightarrow \infty} \mathbf{E}\{f(X_n)\} = \mathbf{E}\{f(X)\}.$$

Για παράδειγμα, αν $X_n \Rightarrow X$ και υποθέσουμε ότι $\sup_{n \geq 1} \mathbf{E}(|X_n|^k) < \infty$ για κάποιο $k \in \mathbb{N}^+$, τότε $\lim_{n \rightarrow \infty} \mathbf{E}(X_n^j) = \mathbf{E}(X^j)$ για κάθε $j = 1, \dots, k-1$.

Απόδειξη. Έστω $(Y_n)_{n \geq 1}, Y$ τυχαίες μεταβλητές όπως στο Θεώρημα 14.16. Τότε επειδή η g είναι συνεχής, το λήμμα Fatou δίνει $\mathbf{E}|g(X)| = \mathbf{E} \lim_{n \rightarrow \infty} |g(Y_n)| \leq \liminf_{n \rightarrow \infty} \mathbf{E}|g(Y_n)| \leq A$.

Για $\varepsilon > 0$ υπάρχει $M > 0$ ώστε $|f(x)| < \varepsilon|g(x)|$ για κάθε x με $|x| > M$ και $\mathbf{P}(Y = -M) = \mathbf{P}(Y = M) = 0$ [το σύνολο των M στα οποία $\mathbf{P}(Y = -M) + \mathbf{P}(Y = M) > 0$ είναι αριθμήσιμο]. Από το θεώρημα φραγμένης σύγκλισης και επειδή η f είναι συνεχής και $\mathbf{P}(Y = -M) = \mathbf{P}(Y = M) = 0$, έχουμε

$$\lim_{n \rightarrow \infty} \mathbf{E}\{f(Y_n)\mathbf{1}_{|Y_n| \leq M}\} = \mathbf{E}\{f(Y)\mathbf{1}_{|Y| \leq M}\}.$$

Έτσι, παίρνοντας $n \rightarrow \infty$ στην

$$\begin{aligned} |\mathbf{E}\{f(Y_n)\} - \mathbf{E}\{f(Y)\}| &\leq |\mathbf{E}\{f(Y_n)\mathbf{1}_{|Y_n| \leq M}\} - \mathbf{E}\{f(Y)\mathbf{1}_{|Y| \leq M}\}| + |\mathbf{E}\{f(Y_n)\mathbf{1}_{|Y_n| > M}\}| + |\mathbf{E}\{f(Y)\mathbf{1}_{|Y| > M}\}| \\ &\leq |\mathbf{E}\{f(Y_n)\mathbf{1}_{|Y_n| \leq M}\} - \mathbf{E}\{f(Y)\mathbf{1}_{|Y| \leq M}\}| + \varepsilon \mathbf{E}|g(Y_n)| + \varepsilon \mathbf{E}|g(Y)|, \end{aligned}$$

έχουμε $\overline{\lim}_{n \rightarrow \infty} |\mathbf{E}\{f(Y_n)\} - \mathbf{E}\{f(Y)\}| \leq 2\varepsilon A$. Επειδή το ε είναι αυθαίρετο, έπεται το συμπέρασμα. \blacksquare

14.3 Σφιχτότητα και υπακολουθιακά όρια

Ορισμός 14.19. Μια οικογένεια $\{\mu_i : i \in I\}$ μέτρων πιθανότητας στο \mathbb{R} λέγεται **σφιχτή** αν για κάθε $\varepsilon > 0$ υπάρχει $M > 0$ έτσι ώστε

$$\mu_i(\mathbb{R} \setminus [-M, M]) < \varepsilon \quad (14.7)$$

για κάθε $i \in I$.

Δηλαδή για μια σφιχτή οικογένεια υπάρχει ένα φραγμένο υποσύνολο K του \mathbb{R} ώστε κάθε στοιχείο της να δίνει «σχεδόν όλη» του τη μάζα στο K (το πολύ μάζα ε βρίσκεται εκτός του $K := [-M, M]$). Το σύνολο K είναι το ίδιο για όλα τα στοιχεία της οικογένειας.

Η απαίτηση του ορισμού μπορεί να διατυπωθεί και ως εξής:

$$\lim_{M \rightarrow \infty} \sup_{i \in I} \mu_i(\mathbb{R} \setminus [-M, M]) = 0. \quad (14.8)$$

Παρατήρηση 14.20. Αν το μ είναι μέτρο πιθανότητας στο \mathbb{R} , τότε εύκολα βλέπουμε ότι για κάθε $\varepsilon > 0$ υπάρχει $M > 0$ έτσι ώστε $\mu(\mathbb{R} \setminus [-M, M]) < \varepsilon$ (Άσκηση 14.7). Αυτό συνεπάγεται ότι κάθε πεπερασμένη οικογένεια μέτρων πιθανότητας είναι σφιχτή. Η σφιχτότητα μιας οικογένειας είναι μη τετριμμένο θέμα μόνο στην περίπτωση που το I είναι άπειρο.

Ορισμός 14.21. Έστω $\{X_i : i \in I\}$ οικογένεια τυχαίων μεταβλητών με τιμές στο \mathbb{R} . Η $\{X_i : i \in I\}$ λέγεται **σφιχτή** αν η οικογένεια κατανομών $\{\mathbf{P}^{X_i} : i \in I\}$ είναι σφιχτή.

Εννοείται εδώ ότι για κάθε $i \in I$ έχουμε έναν χώρο πιθανότητας $(\Omega_i, \mathcal{F}_i, \mathbf{P}_i)$ και $X_i : \Omega_i \rightarrow \mathbb{R}$. Επειδή $\mathbf{P}^{X_i}(\mathbb{R} \setminus [-M, M]) = \mathbf{P}_i(|X_i| > M)$, η οικογένεια $\{\mathbf{P}^{X_i} : i \in I\}$ είναι σφιχτή αν και μόνο αν

$$\lim_{M \rightarrow \infty} \sup_{i \in I} \mathbf{P}_i(|X_i| > M) = 0. \quad (14.9)$$

Παρατήρηση 14.22. Έστω $(X_n)_{n \geq 1}$ ακολουθία τυχαίων μεταβλητών έτσι ώστε $X_n \sim \text{Exp}(1/n)$. Τότε η $(X_n)_{n \geq 1}$ δεν είναι σφιχτή. Πράγματι, για $M > 0$,

$$\sup_{n \geq 1} \mathbf{P}(|X_n| > M) = \sup_{n \geq 1} e^{-\frac{M}{n}} = 1.$$

Αυτό συμβαίνει γιατί η X_n έχει μέση τιμή n και η κατανομή της δίνει την περισσότερή της μάζα γύρω από το n (δηλαδή η X_n πέφτει κοντά στο n με μεγάλη πιθανότητα). Καθώς όμως το $n \rightarrow \infty$ αυτό το σημείο συγκέντρωσης απομακρύνεται. Δεν μπορούμε να βρούμε ένα φραγμένο σύνολο ώστε όλες οι X_n να παίρνουν τιμές εκεί με πιθανότητα κοντά στο 1.

Η έννοια της σφιχτότητας στα μέτρα πιθανότητας είναι ανάλογη της έννοιας του φραγμένου συνόλου σε Ευκλείδειο χώρο (δηλαδή κάποιον \mathbb{R}^d εφοδιασμένο με την Ευκλείδεια μετρική). Οι όροι μιας συγκλίνουσας ακολουθίας σε οποιονδήποτε μετρικό χώρο ορίζουν ένα φραγμένο σύνολο. Το αντίστοιχο εδώ είναι το ακόλουθο αποτέλεσμα.

Πρόταση 14.23. Έστω $(\mu_n)_{n \geq 1}$, μ μέτρα πιθανότητας στο \mathbb{R} έτσι ώστε $\mu_n \Rightarrow \mu$. Τότε η $\{\mu_n : n \geq 1\}$ είναι σφιχτή.

Απόδειξη. Έστω $\varepsilon > 0$. Υπάρχει $M > 0$ έτσι ώστε $\mu(\mathbb{R} \setminus [-M, M]) < \varepsilon/2$ και $\mu(\{-M\}) = \mu(\{M\}) = 0$ (το τελευταίο επειδή η $\mu(\{x\}) > 0$ ισχύει για αριθμησιμο πλήθος $x \in \mathbb{R}$). Από την υπόθεση έχουμε

$$\begin{aligned} \mu_n(\mathbb{R} \setminus [-M, M]) &= \mu_n((-\infty, -M)) + \mu_n((M, \infty)) \\ &\leq \mu_n((-\infty, -M)) + 1 - \mu_n((-\infty, M]) \rightarrow \mu(\mathbb{R} \setminus [-M, M]) < \varepsilon/2 \end{aligned}$$

για $n \rightarrow \infty$. Άρα υπάρχει $n_0 \geq 1$ έτσι ώστε

$$\mu_n(\mathbb{R} \setminus [-M, M]) < \varepsilon$$

για κάθε $n \geq n_0$.

Έπειτα, για τα $\mu_1, \mu_2, \dots, \mu_{n_0-1}$ υπάρχει $\tilde{M} > 0$ έτσι ώστε $\mu_i(\mathbb{R} \setminus [-\tilde{M}, \tilde{M}]) < \varepsilon$ για κάθε $i = 1, 2, \dots, n_0 - 1$ (Άσκηση 14.8).

Έστω $L = \max\{M, \tilde{M}\}$. Τότε $\mu_n(\mathbb{R} \setminus [-L, L]) < \varepsilon$ για κάθε $n \geq 1$. ■

Σύγκλιση κατά κατανομή

Το ανάλογο του ότι κάθε φραγμένη ακολουθία σε έναν Ευκλείδειο χώρο έχει συγκλίνουσα υπακολουθία (Θεώρημα Bolzano-Weierstrass) είναι το επόμενο αποτέλεσμα. Η απόδειξή του είναι απαιτητική και την παραλείπουμε. Μπορεί να τη βρεί κανείς στο Jacod and Protter (2003) (Θεώρημα 18.6).

Θεώρημα 14.24 (Θεώρημα Prokhorov). Έστω $(\mu_n)_{n \geq 1}$ ακολουθία μέτρων πιθανότητας στο \mathbb{R} . Αν η $\{\mu_n : n \geq 1\}$ είναι σφιχτή, τότε υπάρχει υπακολουθία $(\mu_{k_n})_{n \geq 1}$ της $(\mu_n)_{n \geq 1}$ που συγκλίνει ασθενώς σε κάποιο μέτρο πιθανότητας στο \mathbb{R} .

Άμεση συνέπειά του είναι το εξής αποτέλεσμα για τυχαίες μεταβλητές.

Θεώρημα 14.25. Έστω $(X_n)_{n \geq 1}$ ακολουθία τυχαίων μεταβλητών με τιμές στο \mathbb{R} . Αν η $\{X_n : n \geq 1\}$ είναι σφιχτή, τότε υπάρχει υπακολουθία $(X_{k_n})_{n \geq 1}$ της $(X_n)_{n \geq 1}$ που συγκλίνει κατά κατανομή σε κάποια τυχαία μεταβλητή με τιμές στο \mathbb{R} .

Ασκήσεις

14.1 Να αποδειχθεί η (14.5) με χρήση του Θεωρήματος 14.3.

14.2 (α) Έστω X τυχαία μεταβλητή με κατανομή την Cauchy, δηλαδή με πυκνότητα $f(x) = \pi^{-1}(1+x^2)^{-1}$. Να δειχθεί ότι για κάθε $x > 0$ ισχύει

$$\frac{1}{1+x} < \pi \mathbf{P}(X > x) < \frac{1}{x}. \quad (14.10)$$

(β) Έστω $(X_n)_{n \geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών, καθεμία με κατανομή την Cauchy. Θέτουμε $M_n := \max\{X_1, X_2, \dots, X_n\}$ για κάθε $n \geq 1$. Να δειχθεί ότι

$$\frac{M_n}{n} \Rightarrow \frac{1}{W}$$

όπου η W ακολουθεί την εκθετική κατανομή με παράμετρο $1/\pi$.

14.3 Έστω $(X_n)_{n \geq 1}$ ακολουθία τυχαίων μεταβλητών που συγκλίνει κατά κατανομή σε μια τυχαία μεταβλητή X . Για καθένα από τα ακόλουθα ζεύγη κατανομής για τη X και συνόλου $A \subset \mathbb{R}$, συνεπάγεται η σύγκλιση κατά κατανομή $X_n \Rightarrow X$ την

$$\lim_{n \rightarrow \infty} \mathbf{P}(X_n \in A) = \mathbf{P}(X \in A);$$

	Κατανομή της X	Σύνολο A
(i)	<i>Poisson</i> (2)	$(2, 32.1) \cup \{100\}$
(ii)	<i>Poisson</i> (2)	\mathbb{Q}
(iii)	Γεωμετρική(1/3)	$(-1.5, 2.8)$
(iv)	$N(0, 1)$	$(-2, \pi)$
(v)	$U(0, 1)$	$(0, 1/3) \setminus \mathbb{Q}$
(vi)	Bernouli(2/5) στο $\{0, 1\}$	$(0, 1/2) \cup (2, 4)$

14.4 (Αυτή η άσκηση δείχνει πώς αντιμετωπίζουμε όρια της μορφής $\lim_{n \rightarrow \infty} \mathbf{P}(X_n \in A_n)$ όταν $X_n \Rightarrow X$. Το σύνολο A_n εξαρτάται από το n). Έστω $(X_n)_{n \geq 1}$ ακολουθία τυχαίων μεταβλητών που συγκλίνει κατά κατανομή σε μια τυχαία μεταβλητή X (όλες με τιμές στο \mathbb{R}).

(α) Να δειχθεί ότι $\lim_{n \rightarrow \infty} \mathbf{P}(X_n \leq n) = 1$.

(β) Για κάθε $x \in \mathbb{R}$ να δειχθεί ότι $\lim_{n \rightarrow \infty} \mathbf{P}(X_n \in [x - n^{-1}, x + n^{-1}]) \leq \mathbf{P}(X = x)$. Να δοθεί παράδειγμα $(X_n)_{n \geq 1}$, X , x που η προηγούμενη ανισότητα ισχύει ως $<$.

14.5 Έστω $(X_n)_{n \geq 1}$, $(Y_n)_{n \geq 1}$ τυχαίες μεταβλητές ορισμένες σε κοινό χώρο πιθανότητας και με τιμές στο \mathbb{R} . Υποθέτουμε ότι η $(X_n)_{n \geq 1}$ είναι σφιχτή και ότι $Y_n \Rightarrow 0$. Να δειχθεί ότι $X_n Y_n \Rightarrow 0$.

Σύγκλιση κατά κατανομή

14.6 (Θεώρημα Slutsky) Έστω $X, (X_n)_{n \geq 1}, (Y_n)_{n \geq 1}$ τυχαίες μεταβλητές σε χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$, με τιμές στο \mathbb{R} και $c \in \mathbb{R}$.

(α) Αν $X_n \Rightarrow X$ και $Y_n \rightarrow c$ κατά πιθανότητα, τότε $X_n + Y_n \Rightarrow X + c$.

(β) Αν $X_n \Rightarrow X$ και $Y_n \rightarrow c$ κατά πιθανότητα, τότε $X_n Y_n \Rightarrow cX$.

[Υπόδειξη: Υποθέτουμε πρώτα ότι $c = 0$. Στη γενική περίπτωση θεωρούμε την $Y'_n = Y_n - c$ και χρησιμοποιούμε την περίπτωση $c = 0$ και το (α).]

Παρατηρήστε ότι το (α) έχει την εξής άμεση συνέπεια:

(γ) Αν $X_n \Rightarrow X$ και $Y_n - X_n \rightarrow 0$ κατά πιθανότητα, τότε $Y_n \Rightarrow X$.

14.7 Έστω p μέτρο πιθανότητας στο \mathbb{R} . Να δείξετε ότι για κάθε $\varepsilon > 0$ υπάρχει $M > 0$ έτσι ώστε $p(\mathbb{R} \setminus [-M, M]) < \varepsilon$.

14.8 Έστω $\{p_i : i \in I\}$ οικογένεια μέτρων πιθανότητας στο \mathbb{R} με I πεπερασμένο. Να δείξετε ότι η $\{p_i : i \in I\}$ είναι σφικτή.

14.9 Έστω $(X_n)_{n \geq 1}$ ακολουθία τυχαίων μεταβλητών και $h : [0, \infty) \rightarrow [0, \infty)$ αύξουσα με $\lim_{x \rightarrow \infty} h(x) = \infty$ έτσι ώστε $\sup_{n \geq 1} \mathbf{E}\{h(|X_n|)\} < \infty$. Να δείξετε ότι η $(X_n)_{n \geq 1}$ είναι σφικτή.

14.10 Έστω $(X_n)_{n \geq 1}$ τυχαίες μεταβλητές με τιμές στο \mathbb{R} . Αν η $(X_n)_{n \geq 1}$ είναι σφικτή, $\mathbf{E}(X_n^2) < \infty$ για κάθε $n \geq 1$, και $\lim_{n \rightarrow \infty} \mathbf{E}(X_n^2) = \infty$, να δειχθεί ότι $\lim_{n \rightarrow \infty} \text{Var}(X_n) = \infty$.

14.11 Έστω $(\mu_n)_{n \geq 1}$ ακολουθία μέτρων πιθανότητας στο \mathbb{R} . Να δειχθεί ότι η $(\mu_n)_{n \geq 1}$ είναι σφικτή αν και μόνο αν κάθε υπακολουθία της, $(\mu_{n_k})_{k \geq 1}$, έχει υπακολουθία που συγκλίνει ασθενώς (σε ένα μέτρο πιθανότητας στο \mathbb{R}).

14.12 Έστω $(X_n)_{n \geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών καθεμία με πυκνότητα

$$f(x) = \begin{cases} \frac{1}{2\sqrt{x}} & \text{αν } x \in (0, 1), \\ 0 & \text{αν } x \in \mathbb{R} \setminus (0, 1). \end{cases}$$

(α) Να δειχθεί ότι για κάθε $t \in (0, 1)$ ισχύει $\mathbf{P}(X_1 > t) = 1 - \sqrt{t}$.

(β) Θέτουμε $A_n = \min\{X_1, X_2, \dots, X_n\}$ για κάθε $n \geq 1$. Να δειχθεί ότι $n^2 A_n \Rightarrow Y$ καθώς $n \rightarrow \infty$, όπου Y είναι τυχαία μεταβλητή με συνάρτηση κατανομής

$$F_Y(x) = \begin{cases} 0 & \text{αν } x \leq 0, \\ 1 - e^{-\sqrt{x}} & \text{αν } x > 0. \end{cases}$$

14.13 Δώστε παράδειγμα τυχαίων μεταβλητών $(X_n)_{n \geq 1}, X$ ώστε $X_n \Rightarrow X$ αλλά να μην ισχύει $\lim_{n \rightarrow \infty} \mathbf{E}(X_n) = \mathbf{E}(X)$.

14.14 Έστω $X, (X_n)_{n \in \mathbb{N}^+}$ τυχαίες μεταβλητές στο \mathbb{R} με πυκνότητες $f, (f_n)_{n \in \mathbb{N}^+}$ αντίστοιχα. Να δειχθεί ότι $X_n \Rightarrow X$ αν ισχύει ένα από τα εξής.

(i) $\lim_{n \rightarrow \infty} \int |f_n(x) - f(x)| d\lambda(x) = 0$.

(ii) $\lim_{n \rightarrow \infty} f_n = f$ λ -σχεδόν παντού.

14.15 Έστω $X, (X_n)_{n \in \mathbb{N}^+}$ τυχαίες μεταβλητές με τιμές σε ένα αριθμήσιμο σύνολο $S \subset \mathbb{R}$ (π.χ., το \mathbb{N}). Να δειχθεί ότι τα εξής είναι ισοδύναμα:

(i) $X_n \Rightarrow X$,

(ii) $\lim_{n \rightarrow \infty} \mathbf{P}(X_n = a) = \mathbf{P}(X = a)$ για κάθε $a \in S$.

14.16 [Μεροληψία με βάση το μέγεθος²] Αριθμούμε τις οικογένειες που ζουν σε μια πόλη ως 1, 2, 3, ... Ονομάζουμε X_i το πλήθος των παιδιών της οικογένειας i και θεωρούμε ότι οι $(X_i)_{i \in \mathbb{N}^+}$ είναι ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με $\mathbf{E}(X_1) \in (0, \infty)$. Για δεδομένο $n \in \mathbb{N}^+$, θεωρούμε τις οικογένειες 1, 2, ..., n και επιλέγουμε στην τύχη ένα από τα παιδιά τους. Έστω N_n το πλήθος των παιδιών που έχει η οικογένεια της οποίας παιδί έχουμε επιλέξει [παρατηρήστε ότι μια οικογένεια με πολλά παιδιά έχει μεγαλύτερη πιθανότητα να επιλεγεί]. Να δειχθεί

²Σε αυτή την άσκηση βλέπουμε πώς προκύπτει με φυσιολογικό τρόπο η μεροληπτική με βάση το μέγεθος εκδοχή μιας τυχαίας μεταβλητής με τιμές στο \mathbb{N} . Για τυχαίες μεταβλητές X με τιμές στο $[0, \infty)$ με πυκνότητα και πεπερασμένη μέση τιμή, η X^* εμφανίζεται φυσιολογικά στην ανανεωτική θεωρία.

Σύγκλιση κατά κατανομή

ότι για $n \rightarrow \infty$ η N_n συγκλίνει κατά κατανομή στη X_1^* , τη μεροληπτική με βάση το μέγεθος εκδοχή της X_1 [δες Άσκηση 7.7]. Δηλαδή στην τυχαία μεταβλητή με συνάρτηση πιθανότητας $k \mathbf{P}(X_1 = k) / \mathbf{E}(X_1)$ για κάθε $k \in \mathbb{R}$.

[Για αυτή την άσκηση θεωρούνται δεδομένες οι ιδιότητες της δεσμευμένης μέσης τιμής από τις στοιχειώδεις πιθανότητες. Ιδιαίτερα η σχέση $\mathbf{E}(X) = \mathbf{E}(\mathbf{E}(X|Y))$ για τυχαίες μεταβλητές $X : \Omega \rightarrow \mathbb{R}, Y : \Omega \rightarrow \mathbb{R}^n$ για κάποιο $n \in \mathbb{N}^+$ με $\mathbf{E}|X| < \infty$.]

15

Σύγκλιση κατά κατανομή και χαρακτηριστικές συναρτήσεις

Αν $X, (X_n)_{n \geq 1}$ είναι τυχαίες μεταβλητές με τιμές στο \mathbb{R} , για να δείξουμε την ασθενή σύγκλιση $X_n \Rightarrow X$ έχουμε δει δύο τρόπους (Παραδείγματα 14.5, 14.6 για τον πρώτο και Παράδειγμα 14.10 για τον δεύτερο). Σε αυτό το κεφάλαιο θα δούμε έναν τρίτο. Σύμφωνα με αυτόν,

η σύγκλιση $X_n \Rightarrow X$ ισοδυναμεί με τη συνθήκη: για κάθε $t \in \mathbb{R}$ ισχύει $\phi_{X_n}(t) \rightarrow \phi_X(t)$.

Καθένας από τους τρεις αυτούς τρόπους λειτουργεί καλά σε διαφορετικές περιπτώσεις. Ο πρώτος είναι χρήσιμος σε περιπτώσεις που η συνάρτηση κατανομής της X_n υπολογίζεται εύκολα (π.χ., αν η X_n αφορά μέγιστο ή ελάχιστο ακολουθίας τυχαίων μεταβλητών ή έχει γνωστή κατανομή). Ο τρίτος είναι χρήσιμος όταν η X_n εμπλέκει άθροισμα ανεξάρτητων τυχαίων μεταβλητών (τέτοια είναι η περίπτωση του κεντρικού οριακού θεωρήματος).

15.1 Το Θεώρημα Συνέχειας του Λένυ

Λήμμα 15.1. Έστω μ μέτρο πιθανότητας στο \mathbb{R} και ϕ_μ ο μετασχηματισμός Fourier του. Τότε

$$\mu\left(\left\{x : |x| > \frac{2}{u}\right\}\right) \leq \frac{1}{u} \int_{-u}^u (1 - \phi_\mu(t)) dt \quad (15.1)$$

για κάθε $u > 0$.

Η απόδειξη του λήμματος δίνεται στο Παράρτημα Β'. Είναι χρήσιμο γιατί αποδεικνύει σφιχτότητα για σύνολο μέτρων αν υπάρχει αρκετός έλεγχος στον μετασχηματισμό Fourier τους για t κοντά στο 0.

Θεώρημα 15.2. (Θεώρημα συνέχειας του Λένυ) Έστω $(\mu_n)_{n \geq 1}$ ακολουθία μέτρων πιθανότητας στο \mathbb{R} και $(\phi_{\mu_n})_{n \geq 1}$ η ακολουθία μετασχηματισμών Fourier τους.

(i) Αν μ μέτρο πιθανότητας στο \mathbb{R} έτσι ώστε $\mu_n \Rightarrow \mu$, τότε $\lim_{n \rightarrow \infty} \phi_{\mu_n}(t) = \phi_\mu(t)$ για κάθε $t \in \mathbb{R}$.

(ii) Αν το $\lim_{n \rightarrow \infty} \phi_{\mu_n}(t)$ υπάρχει για κάθε $t \in \mathbb{R}$ και η $f(t) := \lim_{n \rightarrow \infty} \phi_{\mu_n}(t)$ είναι συνεχής στο 0, τότε υπάρχει μέτρο πιθανότητας μ στο \mathbb{R} ώστε $\phi_\mu(t) = f(t)$ και $\mu_n \Rightarrow \mu$.

Απόδειξη. (i) Από την υπόθεση και το Θεώρημα 14.8, για κάθε συνάρτηση $f : \mathbb{R} \rightarrow \mathbb{R}$ συνεχή και φραγμένη έχουμε

$$\int f(x) d\mu_n(x) \rightarrow \int f(x) d\mu(x).$$

Εφόσον, για σταθερό $t \in \mathbb{R}$, οι συναρτήσεις $y \mapsto \cos(ty)$, $y \mapsto \sin(ty)$ είναι συνεχείς και φραγμένες, έχουμε

$$\begin{aligned}\phi_{\mu_n}(t) &= \int e^{itx} d\mu_n(x) = \int \cos(tx) d\mu_n(x) + i \int \sin(tx) d\mu_n(x) \\ &\rightarrow \int \cos(tx) d\mu(x) + i \int \sin(tx) d\mu(x) = \phi_{\mu}(t)\end{aligned}$$

για $n \rightarrow \infty$.

(ii) **Βήμα 1:** Η $\{\mu_n\}_{n \geq 1}$ είναι σφιχτή.

Έστω $\varepsilon > 0$. Για $u > 0$ και για κάθε $n \geq 1$, από το Λήμμα 15.1, έχουμε ότι

$$\mu_n \left(\left\{ x : |x| > \frac{2}{u} \right\} \right) \leq \frac{1}{u} \int_{-u}^u (1 - \phi_{\mu_n}(t)) dt. \quad (15.2)$$

Από την υπόθεση και το θεώρημα φραγμένης σύγκλισης, η τελευταία ανισότητα δίνει

$$\overline{\lim}_{n \rightarrow \infty} \mu_n \left(\left\{ x : |x| > \frac{2}{u} \right\} \right) \leq \frac{1}{u} \int_{-u}^u (1 - f(t)) dt. \quad (15.3)$$

Επειδή η f είναι συνεχής στο 0 και $f(0) = 1$ (εφόσον $\phi_{\mu_n}(0) = 1$ για κάθε $n \geq 1$ και η f είναι το σημειακό όριο των ϕ_{μ_n}) το όριο του δεξιού μέλους της ανισότητας για $u \rightarrow 0^+$ ισούται με $2(1 - f(0)) = 0$. Άρα υπάρχει $u_0 > 0$ και $n_0 \in \mathbb{N}$ έτσι ώστε

$$\mu_n \left(\left\{ x : |x| > \frac{2}{u_0} \right\} \right) < \varepsilon \text{ για κάθε } n \geq n_0. \quad (15.4)$$

Από την (15.4) και εφόσον η $\{\mu_k : 1 \leq k \leq n_0 - 1\}$ είναι σφιχτή (Άσκηση 14.8), έπεται ότι η $\{\mu_n : n \geq 1\}$ είναι σφιχτή.

Βήμα 2: Υπάρχει μέτρο μ ώστε $f(t) = \phi_{\mu}(t)$ για κάθε $t \in \mathbb{R}$.

Πράγματι, η $(\mu_n)_{n \geq 1}$ είναι σφιχτή και από το Θεώρημα 14.24 προκύπτει ότι υπάρχει υπακολουθία $(\mu_{n_k})_{k \geq 1}$ τέτοια ώστε να συγκλίνει ασθενώς σε ένα μέτρο πιθανότητας μ . Λόγω του (i) του θεωρήματος,

$$\lim_{k \rightarrow \infty} \phi_{\mu_{n_k}}(t) = \phi_{\mu}(t), \quad \forall t \in \mathbb{R}.$$

Όμως

$$\lim_{k \rightarrow \infty} \phi_{\mu_{n_k}}(t) = f(t), \quad \forall t \in \mathbb{R}.$$

Συνεπώς, $f(t) = \phi_{\mu}(t)$ για κάθε $t \in \mathbb{R}$.

Βήμα 3: Αν μια υπακολουθία της $(\mu_n)_{n \geq 1}$ συγκλίνει ασθενώς σε ένα μέτρο πιθανότητας ν , τότε $\nu = \mu$.

Πράγματι, αν $(\lambda_n)_{n \geq 1}$ είναι γνήσια αύξουσα ακολουθία στο \mathbb{N} και ν μέτρο πιθανότητας στο \mathbb{R} έτσι ώστε $\mu_{\lambda_n} \Rightarrow \nu$, από το (i),

$$\lim_{n \rightarrow \infty} \phi_{\mu_{\lambda_n}}(t) = \phi_{\nu}(t), \quad \forall t \in \mathbb{R}.$$

Όμοια με πριν (Βήμα 2), $\phi_{\nu}(t) = f(t)$ για κάθε $t \in \mathbb{R}$ και λόγω μοναδικότητας του μετασχηματισμού Fourier [Πόρισμα 13.6], έχουμε $\nu = \mu$.

Βήμα 4:¹ Η $(\mu_n)_{n \geq 1}$ συγκλίνει ασθενώς στο μ .

¹Για την καλύτερη κατανόηση της απόδειξης, θεωρείστε την εξής ανάλογη άσκηση απειροστικού λογισμού: Έστω $\ell \in \mathbb{R}$ και $(x_n)_{n \in \mathbb{N}}$ ακολουθία πραγματικών αριθμών έτσι ώστε κάθε υπακολουθία της, $(x_{n_k})_{k \geq 1}$, να έχει υπακολουθία $(x_{n_{k_r}})_{r \geq 1}$ που συγκλίνει στο ℓ . Τότε $x_n \rightarrow \ell$.

Έστω ότι αυτό δεν ισχύει. Τότε υπάρχει $h : \mathbb{R} \rightarrow \mathbb{R}$ συνεχής και φραγμένη, $\varepsilon > 0$ και $(\lambda_n)_{n \geq 1}$ ακολουθία στο \mathbb{N} έτσι ώστε

$$\left| \int h(x) d\mu_{\lambda_n}(x) - \int h(x) d\mu(x) \right| \geq \varepsilon, \quad \forall n \geq 1. \quad (5)$$

Η $(\mu_{\lambda_n})_{n \geq 1}$ είναι σφικτή, άρα υπάρχει υπακολουθία της, έστω $(\mu_{\lambda_{n_k}})_{k \geq 1}$, και μέτρο πιθανότητας ν στο \mathbb{R} έτσι ώστε $\mu_{\lambda_{n_k}} \Rightarrow \nu$ (Θεώρημα 14.24). Από τα προηγούμενα (Βήμα 3) προκύπτει ότι $\nu = \mu$, δηλαδή $\mu_{\lambda_{n_k}} \Rightarrow \mu$. Συνεπώς,

$$\lim_{n \rightarrow \infty} \int h(x) d\mu_{\lambda_n}(x) = \int h(x) d\mu(x),$$

το οποίο είναι άτοπο λόγω της (5). ■

Πόρισμα 15.3. Έστω $(X_n)_{n \geq 1}$, X τυχαίες μεταβλητές με τιμές στο \mathbb{R} . Τότε $X_n \Rightarrow X$ αν και μόνο αν $\phi_{X_n}(t) \rightarrow \phi_X(t)$ για κάθε $t \in \mathbb{R}$ καθώς $n \rightarrow \infty$.

Απόδειξη. Ισχύει ότι $\phi_{X_n}(t) = \phi_{\mathbf{P}^{X_n}}(t)$ και $\phi_X(t) = \phi_{\mathbf{P}^X}(t)$, όπου \mathbf{P}^{X_n} , \mathbf{P}^X είναι κατανομές των X_n και X αντίστοιχα. Το συμπέρασμα έπεται με εφαρμογή του θεωρήματος συνέχειας του Λέβυ για τα μέτρα $(\mathbf{P}^{X_n})_{n \geq 1}$ και \mathbf{P}^X . ■

Το τελευταίο πόρισμα είναι ο βασικότερος τρόπος για να δείξει κανείς σύγκλιση κατά κατανομή. Θα το ονομάζουμε και αυτό θεώρημα συνέχειας του Λέβυ. Θα το χρησιμοποιήσουμε πολλές φορές στο εξής και ιδιαίτερα για να αποδείξουμε το κεντρικό οριακό θεώρημα.

15.2 Εφαρμογές

Σε υπολογισμούς ορίων της μορφής $\lim_{n \rightarrow \infty} \phi_{X_n}(t)$ χρήσιμο είναι το εξής απλό αποτέλεσμα.

Λήμμα 15.4. Αν $(c_n)_{n \geq 1}$ είναι ακολουθία στο \mathbb{C} τέτοια ώστε $c_n \rightarrow c$, με $c \in \mathbb{C}$, τότε

$$\left(1 + \frac{c_n}{n}\right)^n \rightarrow e^c.$$

Η απόδειξή του δίνεται στο Παράρτημα Α' (Πόρισμα Α'.3).

Παράδειγμα 15.5. Έστω $X_n \sim \text{Bin}(n, \frac{\lambda}{n})$ για κάθε $n \in \mathbb{N}$ και $X \sim \text{Poisson}(\lambda)$, όπου $\lambda > 0$ δεδομένη σταθερά. Τότε

$$\begin{aligned} \phi_{X_n}(t) &= \mathbf{E}(e^{itX_n}) = \sum_{k=0}^n e^{itk} \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \sum_{k=0}^n \binom{n}{k} \left(\frac{e^{it}\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \left(1 - \frac{\lambda}{n} + \frac{\lambda e^{it}}{n}\right)^n \\ &= \left(1 + \frac{\lambda(e^{it} - 1)}{n}\right)^n \end{aligned}$$

και

$$\lim_{n \rightarrow \infty} \left(1 + \frac{\lambda(e^{it} - 1)}{n}\right)^n = e^{\lambda(e^{it} - 1)}.$$

Σύγκλιση κατά κατανομή και χαρακτηριστικές συναρτήσεις

Όμως

$$\phi_X(t) = \mathbf{E}(e^{itX}) = \sum_{k=0}^{\infty} e^{itk} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^{it})^k}{k!} = e^{\lambda(e^{it}-1)}.$$

Άρα, $\phi_{X_n}(t) \rightarrow \phi_X(t)$. Συνεπώς, από το θεώρημα συνέχειας του Λένυ προκύπτει ότι $X_n \Rightarrow X$.

Παράδειγμα 15.6. Έστω $X_n \sim \text{Poisson}(n)$ και $Z \sim N(0, 1)$. Θέτουμε $Z_n = \frac{X_n - n}{\sqrt{n}}$. Τότε

$$\begin{aligned} \phi_{Z_n}(t) &= \mathbf{E}(e^{itZ_n}) = \mathbf{E}\left(e^{it\left(\frac{X_n - n}{\sqrt{n}}\right)}\right) \\ &= e^{-it\sqrt{n}} \mathbf{E}(e^{i\frac{t}{\sqrt{n}}X_n}) = e^{-it\sqrt{n}} \phi_{X_n}\left(\frac{t}{\sqrt{n}}\right) \\ &= e^{it\sqrt{n}} e^{n(e^{i\frac{t}{\sqrt{n}}}-1)} = e^{n(e^{i\frac{t}{\sqrt{n}}}-1)-it\sqrt{n}}. \end{aligned}$$

Θέτουμε $\varepsilon = \frac{1}{\sqrt{n}}$. Τότε ο εκθέτης στην τελευταία ποσότητα ισούται με

$$\frac{e^{i\varepsilon t} - 1}{\varepsilon^2} - \frac{it}{\varepsilon} = \frac{e^{i\varepsilon t} - 1 - i\varepsilon t}{\varepsilon^2}.$$

Χρησιμοποιώντας τον κανόνα του L' Hospital,

$$\lim_{\varepsilon \rightarrow 0} \frac{e^{i\varepsilon t} - 1 - i\varepsilon t}{\varepsilon^2} = \frac{-t^2}{2}.$$

Άρα $\phi_{Z_n}(t) \rightarrow \phi_Z(t) = e^{-\frac{t^2}{2}}$ και από το θεώρημα συνέχειας του Λένυ προκύπτει ότι $Z_n \Rightarrow Z$.

Παράδειγμα 15.7. Έστω X_n ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με

$$f_{X_1}(x) = \begin{cases} \frac{1}{|x|^3} & \text{αν } |x| \geq 1, \\ 0 & \text{αν } |x| < 1. \end{cases}$$

Θέτουμε $S_n = X_1 + \dots + X_n$ για κάθε $n \geq 1$. Θα δείξουμε ότι

$$\frac{S_n}{\sqrt{n \log n}} \Rightarrow Z$$

με $Z \sim N(0, 1)$.

Έστω $a_n := \sqrt{n \log n}$. Για $t \in \mathbb{R}$,

$$\phi_{S_n/a_n}(t) = (\phi_{X_1}(t/a_n))^n = \left(1 + \frac{C_n(t)}{n}\right)^n$$

με $C_n(t) := n(\phi_{X_1}(t/a_n) - 1)$. Από την Άσκηση 13.13 έχουμε ότι για μεγάλα n (αρκεί το n να ικανοποιεί την $|t| \leq a_n$) ισχύει

$$\phi_{X_1}(t/a_n) = 1 + \frac{t^2}{a_n^2} \log \frac{|t|}{a_n} + b(n, t)$$

με $|b(n, t)| \leq 3t^2/a_n^2$. Οπότε

$$C_n(t) = n \left(\frac{t^2}{a_n^2} \log |t| - \frac{t^2}{a_n^2} \log a_n + b(n, t) \right) = n \left(\frac{t^2}{n \log n} \log |t| - \frac{t^2}{2n \log n} \log(n \log n) + b(n, t) \right) \rightarrow -\frac{t^2}{2}$$

καθώς $n \rightarrow \infty$. Το συμπέρασμα έπεται από το Λήμμα 15.4 και το θεώρημα συνέχειας του Λένυ.

Ασκήσεις

15.1 Έστω μ μέτρο πιθανότητας στο $[0, \infty)$. Μετασχηματισμό Laplace του μ λέμε τη συνάρτηση $L : [0, \infty) \rightarrow [0, \infty)$ με $L(s) = \int_0^\infty e^{-sx} d\mu(x)$ για κάθε $s \geq 0$. Να δειχθεί ότι για κάθε $u > 0$ ισχύει

$$\mu([2/u, \infty)) \leq \frac{2}{u} \int_0^u \{1 - L(s)\} ds. \quad (15.5)$$

15.2 Έστω $(Y_n)_{n \geq 1}$ τυχαίες μεταβλητές ώστε $Y_n \sim N(\mu_n, \sigma_n^2)$ όπου $\mu_n \in \mathbb{R}, \sigma_n \in (0, \infty)$ για κάθε $n \geq 1$. Για δεδομένα $\mu \in \mathbb{R}, \sigma \in (0, \infty)$ θεωρούμε $Y \sim N(\mu, \sigma^2)$. Τότε $\lim_{n \rightarrow \infty} \mu_n = \mu$ και $\lim_{n \rightarrow \infty} \sigma_n = \sigma$ αν και μόνο αν $Y_n \Rightarrow Y$.

15.3 (α) Έστω Y τυχαία μεταβλητή που ακολουθεί τη γεωμετρική κατανομή με παράμετρο $p \in (0, 1]$. Δηλαδή $P(Y = k) = p(1-p)^{k-1}$ για $k = 1, 2, \dots$. Να υπολογιστεί η χαρακτηριστική συνάρτηση της Y .

(β) Έστω $a > 0$ και $(X_n)_{n \geq 1}$ ακολουθία τυχαίων μεταβλητών ώστε η X_n να ακολουθεί τη γεωμετρική κατανομή με παράμετρο $p_n = a/n$. Να δειχθεί ότι η ακολουθία $(X_n/n)_{n \geq 1}$ συγκλίνει κατά κατανομή σε μια τυχαία μεταβλητή $X \sim \text{Exp}(1)$ χρησιμοποιώντας

- (i) τον χαρακτηρισμό της σύγκλισης κατά κατανομή μέσω συναρτήσεων κατανομής,
- (ii) χαρακτηριστικές συναρτήσεις.

15.4 (α) Έστω X τυχαία μεταβλητή που ακολουθεί την κατανομή Poisson με παράμετρο $\lambda > 0$. Να υπολογιστεί η χαρακτηριστική συνάρτηση της X .

(β) Έστω $(X_n)_{n \geq 1}$ ακολουθία τυχαίων μεταβλητών ώστε η X_n να ακολουθεί τη διωνυμική κατανομή με παραμέτρους $n \in \mathbb{N}^+, p_n \in (0, 1)$. Αν $\lim_{n \rightarrow \infty} np_n = \lambda \in (0, \infty)$, να δειχθεί ότι η ακολουθία $(X_n)_{n \geq 1}$ συγκλίνει κατά κατανομή στην τυχαία μεταβλητή X του ερωτήματος (α).

15.5 Έστω $c > 0$, και $(X_n)_{n \geq 1}$ ακολουθία τυχαίων μεταβλητών ώστε $X_n \sim \Gamma(nc, n)$ για κάθε $n \geq 1$ και Z τυχαία μεταβλητή με $Z \sim N(0, c)$. Να δειχθεί ότι

$$\sqrt{n}(X_n - c) \Rightarrow Z$$

καθώς $n \rightarrow \infty$.

15.6 Έστω $(X_n)_{n \geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές ώστε η χαρακτηριστική συνάρτηση ϕ της X_1 να είναι διαφορίσιμη στο 0 και $\phi'(0) = ia$ για κάποιο $a \in \mathbb{R}$. Για κάθε $n \in \mathbb{N}^+$, θέτουμε $S_n := X_1 + X_2 + \dots + X_n$. Να δειχθεί ότι

$$\frac{S_n}{n} \rightarrow a$$

κατά πιθανότητα για $n \rightarrow \infty$.

15.7 Έστω $(X_k)_{k \geq 1}$ ακολουθία ανεξάρτητων τυχαίων μεταβλητών ώστε για κάθε $k \geq 1$ έχουμε

$$X_k = \begin{cases} -k & \text{με πιθανότητα } 1/2k, \\ k & \text{με πιθανότητα } 1/2k, \\ 0 & \text{με πιθανότητα } 1 - k^{-1}. \end{cases}$$

Να δειχθεί ότι η ακολουθία $(S_n/n)_{n \geq 1}$ συγκλίνει κατά κατανομή σε μια τυχαία μεταβλητή Y με χαρακτηριστική συνάρτηση

$$\phi_Y(t) = e^{-\int_0^1 \frac{1 - \cos(tx)}{x} dx}$$

για κάθε $t \in \mathbb{R}$.

[Υπόδειξη: Χρήσιμο είναι το Λήμμα Α'.2.]

Σύγκλιση κατά κατανομή και χαρακτηριστικές συναρτήσεις

15.8 Έστω $(X_k)_{k \geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με $\mathbf{E}(X_1) = 0$, $\text{Var}(X_1) = 1$. Για κάθε $n \geq 1$, θέτουμε $S_n := X_1 + X_2 + \dots + X_n$ και $T_n = S_1 + S_2 + \dots + S_n$. Να βρεθεί $\alpha > 0$ και η κατανομή τυχαίας μεταβλητής Y με τιμές στο \mathbb{R} και μη μηδενική ώστε

$$\frac{T_n}{n^\alpha} \Rightarrow Y$$

καθώς $n \rightarrow \infty$.

15.9 Έστω $(X_k)_{k \geq 1}$ ακολουθία ανεξάρτητων τυχαίων μεταβλητών ώστε για κάθε $k \geq 1$ έχουμε

$$X_k = \begin{cases} -k & \text{με πιθανότητα } 1/(2k^2), \\ k & \text{με πιθανότητα } 1/2k^2, \\ -1 & \text{με πιθανότητα } (1 - k^{-2})/2, \\ -1 & \text{με πιθανότητα } (1 - k^{-2})/2. \end{cases}$$

Θέτουμε $S_n := X_1 + \dots + X_n$ για κάθε $n \in \mathbb{N}^+$. Να δειχθεί ότι

(α) $\lim_{n \rightarrow \infty} \text{Var}(S_n)/n = 2$,

(β) η ακολουθία $(S_n/\sqrt{n})_{n \geq 1}$ συγκλίνει κατά κατανομή σε μια τυχαία μεταβλητή Z με $Z \sim N(0, 1)$.

[Υπόδειξη: Χρήσιμο είναι το Θεώρημα Slutsky (Άσκηση 14.6). Εναλλακτικά, μπορούμε να πάμε ευθέως με χαρακτηριστικές συναρτήσεις και να χρησιμοποιήσουμε το Λήμμα Α'.2.]

15.10 Έστω $(X_k)_{k \geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών, καθεμία με πυκνότητα όπως στην Άσκηση 13.5 με $a \in (0, 2)$. Για κάθε $n \geq 1$, θέτουμε $S_n := X_1 + X_2 + \dots + X_n$. Να δειχθεί ότι η ακολουθία $(S_n/n^{1/a})_{n \geq 1}$ συγκλίνει κατά κατανομή σε μια τυχαία μεταβλητή Y με χαρακτηριστική συνάρτηση $\phi_Y(t) = e^{-C|t|^a}$ για κάθε $t \in \mathbb{R}$ με $C > 0$ σταθερά.

15.11 Έστω $(U_i)_{i \geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με καθεμία να έχει κατανομή την ομοιόμορφη στο διάστημα $(0, 1)$. Θέτουμε $M_n := \max\{U_1, U_2, \dots, U_n\}$ για κάθε $n \in \mathbb{N}^+$.

(α) Να δειχθεί με χρήση συναρτήσεων κατανομής ότι η ακολουθία τυχαίων μεταβλητών $\{n(1 - M_n)\}_{n \in \mathbb{N}^+}$ συγκλίνει κατά κατανομή σε μια τυχαία μεταβλητή Y . Ποια η κατανομή της Y ;

(β) Ποια η πυκνότητα της M_n ; Να δειχθεί με χρήση του θεωρήματος συνέχειας του Λέβυ η σύγκλιση κατά κατανομή του ερωτήματος (α).

16

Το Κεντρικό Οριακό Θεώρημα

Ο ισχυρός νόμος των μεγάλων αριθμών λέει ότι ο μέσος όρος n ανεξάρτητων και ισόνομων τυχαίων μεταβλητών, S_n/n , για μεγάλο n είναι πολύ κοντά στην κοινή τους μέση τιμή $\mu := \mathbf{E}(X_1)$. Τώρα, με την επιπλέον υπόθεση $\sigma^2 := \text{Var}(X_1) \in (0, \infty)$, το κεντρικό οριακό θεώρημα συγκεκριμενοποιεί το πόσο κοντά. Λέει ότι μπορούμε να γράψουμε

$$\frac{S_n}{n} = \mu + \frac{Z_n}{\sqrt{n}} \sigma$$

με τη Z_n να ακολουθεί προσεγγιστικά την τυπική κανονική κατανομή $N(0, 1)$. Ισοδύναμα, $S_n = \mu n + \sigma \sqrt{n} Z_n$, δηλαδή τυπικά το S_n βρίσκεται σε απόσταση της τάξης του \sqrt{n} γύρω από τη μέση του τιμή, $n\mu$.

16.1 Προετοιμασία

Για την απόδειξη του κεντρικού οριακού θεωρήματος, κρίσιμη είναι η συμπεριφορά κοντά στο 0 της χαρακτηριστικής συνάρτησης $\phi_X(t)$ οποιασδήποτε τυχαίας μεταβλητής X με πεπερασμένη πρώτη και δεύτερη ροπή. Χρήσιμη είναι η ανισότητα

$$\left| e^{ix} - 1 - ix - \frac{1}{2}(ix)^2 \right| \leq 2x^2, \quad (16.1)$$

που ισχύει για κάθε $x \in \mathbb{R}$. Για την απόδειξή της παρατηρούμε ότι η ποσότητα μέσα στο μέτρο του αριστερού μέλους είναι $\cos x - 1 + x^2/2 + i(\sin x - x)$, οπότε το μέτρο της είναι φραγμένο από το

$$|\cos x - 1| + \frac{x^2}{2} + |\sin x - x|.$$

Για την πρώτη ποσότητα παρατηρούμε ότι από το θεώρημα Taylor ισχύει $\cos x = 1 - \frac{x^2}{2} \cos \xi$ για κάποιο ξ μεταξύ 0 και x . Άρα $|\cos x - 1| \leq x^2/2$. Όμοια δείχνουμε ότι $|\sin x - x| \leq x^2/2$.

Λήμμα 16.1. Έστω X τυχαία μεταβλητή με τιμές στο \mathbb{R} τέτοια ώστε $\mathbf{E}(X^2) < \infty$. Θέτουμε $\mathbf{E}(X) = \mu$ και $\mathbf{E}(X^2) = \beta$. Τότε

$$\phi_X(t) = 1 + it\mu - \frac{t^2\beta}{2} + \nu(t), \quad (16.2)$$

όπου η ν ικανοποιεί $\lim_{t \rightarrow 0} \frac{\nu(t)}{t^2} = 0$.

Δηλαδή η συνάρτηση ν που ορίζεται ως $\nu(t) := \phi_X(t) - 1 - it\mu + (1/2)t^2\beta$ τείνει στο 0 γρηγορότερα από το t^2 .

Οι πρώτοι τρεις όροι του αναπτύγματος (16.2) προκύπτουν αν στη $\phi_X(t) = \mathbf{E}(e^{itX})$ αναπτύξουμε την e^{itX} σε δυναμοσειρά, κρατήσουμε τους 3 πρώτους όρους

$$1 + itX + \frac{1}{2}(itX)^2,$$

Το Κεντρικό Οριακό Θεώρημα

και πάρουμε τη μέση τιμή τους. Η $v(t)$ είναι η μέση τιμή της υπόλοιπης δυναμοσειράς.

Απόδειξη. Έστω $A(x) := e^{ix} - 1 - ix - (ix)^2/2$ για κάθε $x \in \mathbb{R}$.

$$|v(t)| = \left| \phi_X(t) - 1 - it\mu + \frac{t^2\beta}{2} \right| = \left| \mathbf{E} \left(e^{itX} - 1 - itX - \frac{(itX)^2}{2} \right) \right| \leq \mathbf{E} |A(tX)|$$

Για την απόδειξη του λήμματος αρκεί να δείξουμε ότι $\lim_{t \rightarrow 0} \mathbf{E}(|A(tX)|/t^2) = 0$. Έστω ακολουθία $(t_n)_{n \in \mathbb{N}}$ στο $\mathbb{R} \setminus \{0\}$ τέτοια ώστε $t_n \rightarrow 0$. Τότε

- $A(t_n X)/t_n^2 \rightarrow 0$ για $n \rightarrow \infty$ λόγω του αναπτύγματος της e^{ix} σε δυναμοσειρά.
- $|A(t_n X)/t_n^2| \leq 2X^2$ για κάθε $n \in \mathbb{N}$ λόγω της (16.1) και $\mathbf{E}(X^2) < \infty$.

Από το θεώρημα κυριαρχημένης σύγκλισης, $\lim_{n \rightarrow \infty} \mathbf{E}(|A(t_n X)/t_n^2|) = \mathbf{E}(\lim_{n \rightarrow \infty} |A(t_n X)/t_n^2|) = 0$. ■

16.2 Το Κεντρικό Οριακό Θεώρημα

Θεώρημα 16.2 (Το Κεντρικό Οριακό Θεώρημα). Έστω $(X_n)_{n \geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με $\mathbf{E}(X_1) = \mu$ και $\text{Var}(X_1) = \sigma^2 \in (0, \infty)$. Θέτουμε $S_n = X_1 + X_2 + \dots + X_n$ για κάθε $n \geq 1$. Τότε

$$\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \Rightarrow Z, \text{ όπου } Z \sim N(0, 1).$$

Απόδειξη. Θα αποδείξουμε αρχικά την περίπτωση όπου $\mu = 0$, οπότε $\sigma^2 = \mathbf{E}(X^2)$. Θα χρησιμοποιήσουμε το θεώρημα συνέχειας του Λένυ (Πόρισμα 15.3). Υπολογίζουμε τη χαρακτηριστική συνάρτηση της $S_n/\sqrt{n\sigma^2}$. Για $t \in \mathbb{R} \setminus \{0\}$,

$$\begin{aligned} \phi_{\frac{S_n}{\sqrt{n\sigma^2}}}(t) &= \mathbf{E} \left(e^{it \frac{S_n}{\sqrt{n\sigma^2}}} \right) = \mathbf{E} \left(e^{it \frac{X_1}{\sqrt{n\sigma^2}}} \dots e^{it \frac{X_n}{\sqrt{n\sigma^2}}} \right) \\ &= \mathbf{E} \left(e^{it \frac{X_1}{\sqrt{n\sigma^2}}} \right) \dots \mathbf{E} \left(e^{it \frac{X_n}{\sqrt{n\sigma^2}}} \right) && \text{(λόγω ανεξαρτησίας)} \\ &= \phi_{X_1} \left(\frac{t}{\sqrt{n\sigma^2}} \right) \dots \phi_{X_n} \left(\frac{t}{\sqrt{n\sigma^2}} \right) \\ &= \left(\phi_{X_1} \left(\frac{t}{\sqrt{n\sigma^2}} \right) \right)^n && \text{(λόγω ισονομίας)}. \end{aligned}$$

Από το Λήμμα 16.1,

$$\begin{aligned} \phi_{X_1} \left(\frac{t}{\sqrt{n\sigma^2}} \right) &= 1 + \frac{it}{\sqrt{n\sigma^2}} \mathbf{E}(X_1) - \frac{t^2}{2n\sigma^2} \mathbf{E}((X_1)^2) + v \left(\frac{t}{\sqrt{n\sigma^2}} \right) = 1 - \frac{t^2}{2n} + v \left(\frac{t}{\sqrt{n\sigma^2}} \right) \\ &= 1 + \frac{C_n}{n} \end{aligned}$$

με $C_n = -\frac{t^2}{2} + nv \left(\frac{t}{\sqrt{n\sigma^2}} \right) \rightarrow -t^2/2$ για $n \rightarrow \infty$ λόγω του ότι $\lim_{s \rightarrow 0} v(s)/s^2 = 0$. Συνεπώς, το Λήμμα 15.4 δίνει

$$\lim_{n \rightarrow \infty} \phi_{\frac{S_n}{\sqrt{n\sigma^2}}}(t) = e^{-t^2/2}.$$

Η ίδια σχέση ισχύει προφανώς και για $t = 0$. Μια $Z \sim N(0, 1)$ έχει χαρακτηριστική συνάρτηση $e^{-t^2/2}$. Από το θεώρημα συνέχειας του Λένυ, $S_n/\sqrt{n\sigma^2} \Rightarrow Z$.

Στην περίπτωση όπου $\mathbf{E}(X_1) = \mu \neq 0$, θεωρούμε τις ανεξάρτητες και ισόνομες τυχαίες μεταβλητές $Y_n = X_n - \mu$ για κάθε $n \in \mathbb{N}^+$, που καθεμία έχει μέση τιμή 0, και εφαρμόζουμε τα προηγούμενα. ■

Το Κεντρικό Οριακό Θεώρημα

Ας επιστρέψουμε στο Παράδειγμα 15.7. Αυτό που συμβαίνει εκεί είναι ότι $\mathbf{E}(X_1) = 0$ αλλά $\text{Var}(X_1) = \infty$. Ο νόμος των μεγάλων αριθμών εφαρμόζεται και δίνει ότι $S_n/n \rightarrow 0$. Όμως το ότι $\text{Var}(X_1) = \infty$ σημαίνει ότι το άθροισμα S_n έχει μεγαλύτερη μεταβλητότητα από την περίπτωση που $\text{Var}(X_1) < \infty$. Έτσι, χρειάζεται να διαιρέσουμε το S_n με το $\sqrt{n} \sqrt{\log n}$ (που είναι κατά πολύ μεγαλύτερο του \sqrt{n}) ώστε να πάρουμε ποσότητα που συγκλίνει κατά κατανομή. Το παράδειγμα αυτό δείχνει επίσης ότι σύγκλιση σε κανονική κατανομή μπορεί να πάρει κανείς από την S_n και με παράγοντα κανονικοποίησης διαφορετικό από τον \sqrt{n} .

16.3 Εφαρμογές

Το κεντρικό οριακό θεώρημα δίνει προσεγγίσεις για πιθανότητες που αφορούν άθροισμα ανεξάρτητων και ισόνομων τυχαίων μεταβλητών. Έχει εφαρμογές στη Στατιστική, για παράδειγμα στην κατασκευή προσεγγιστικών διαστημάτων εμπιστοσύνης. Το χρησιμοποιούμε θεωρώντας ότι μια ποσότητα της μορφής

$$\mathbf{P}\left(\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \in A\right) \quad (16.3)$$

προσεγγίζεται (για n μεγάλο) από την

$$\mathbf{P}(Z \in A) \quad (16.4)$$

όπου $Z \sim N(0, 1)$ και A είναι ένα διάστημα ή μια πεπερασμένη ένωση διαστημάτων (αφού το όριο της πρώτης ποσότητας για $n \rightarrow \infty$ είναι η δεύτερη).

Παράδειγμα 16.3. Έχουμε ένα νόμισμα και θέλουμε να ελέγξουμε αν είναι αμερόληπτο. Το ρίχνουμε 100 φορές και έρχεται «Κεφαλή» 38 φορές. Πρέπει να μπούμε σε σκέψεις ότι δεν είναι αμερόληπτο;

Θεωρούμε τις τυχαίες μεταβλητές $(X_k)_{k \geq 1}$ με $X_k = 1$ όταν η k ρίψη φέρνει «Κεφαλή» και $X_k = 0$ όταν η k ρίψη φέρνει γράμματα. Ο συνολικός αριθμός κεφαλών σε n δοκιμές είναι $S_n := X_1 + \dots + X_n$. Υποθέτουμε ότι πράγματι το νόμισμα είναι αμερόληπτο (δηλαδή φέρνει «Κεφαλή» με πιθανότητα $p = 1/2$) και υπολογίζουμε την πιθανότητα $\mathbf{P}(S_{100} \leq 38)$. Έχουμε $\mu = \mathbf{E}(X_1) = 1/2$, $\sigma^2 := \text{Var}(X_1) = p(1-p) = 1/4$ και

$$\begin{aligned} \mathbf{P}(S_{100} \leq 38) &= \mathbf{P}\left(\frac{S_{100} - 100 \times 0.5}{\sqrt{100/4}} \leq \frac{38 - 100 \times 0.5}{\sqrt{100/4}}\right) \approx \mathbf{P}(Z \leq -12/5) \\ &= \Phi(-2.4) = 1 - \Phi(2.4) \approx 0.0082 \end{aligned}$$

Η πιθανότητα είναι πολύ μικρή, μικρότερη του 1%. Μπορούμε να πούμε με μεγάλη βεβαιότητα ότι το νόμισμα δεν είναι αμερόληπτο.

Αν ερχόταν «Κεφαλή» 43 φορές, η ίδια διαδικασία δίνει την προσέγγιση $\mathbf{P}(S_{100} \leq 43) \approx 0.0887$, περίπου 9%. Είναι ένα ενδεχόμενο που έχει σημαντική πιθανότητα να συμβεί. Δεν θα μας έκανε εντύπωση.

Παράδειγμα 16.4. Θεωρούμε ότι το βάρος έλξης κατά το οποίο σπάει ένα συγκεκριμένο συρματόσχοινο από μια πολύ μεγάλη ποσότητα που ένα εργοστάσιο μόλις παρήγαγε είναι μια τυχαία μεταβλητή με άγνωστη μέση τιμή μ και διασπορά $\sigma = 1/10$. Θέλουμε να προσδιορίσουμε προσεγγιστικά τη μέση τιμή μ . Πραγματοποιούμε n ανεξάρτητες μετρήσεις X_1, X_2, \dots, X_n και θα χρησιμοποιήσουμε τον μέσο όρο $\bar{X}_n := (X_1 + \dots + X_n)/n$ ως προσέγγιση της μ . Πόσο μεγάλο πρέπει να είναι το n ώστε να ξέρουμε ότι θα συμβεί $|\bar{X}_n - \mu| \leq 1/100$ με πιθανότητα τουλάχιστον 0.95;

Το Κεντρικό Οριακό Θεώρημα

Για την πιθανότητα του γεγονότος που μας ενδιαφέρει έχουμε την εξής προσέγγιση

$$\mathbf{P}\left(\left|\frac{S_n}{n} - \mu\right| \leq \frac{1}{100}\right) = \mathbf{P}\left(\frac{|S_n - n\mu|}{\sqrt{n\sigma^2}} \leq \sqrt{n}\frac{1}{100\sigma}\right) \approx \mathbf{P}\left(|Z| \leq \sqrt{n}\frac{1}{100\sigma}\right) \quad (16.5)$$

$$= \Phi\left(\sqrt{n}\frac{1}{100\sigma}\right) - \Phi\left(-\sqrt{n}\frac{1}{100\sigma}\right) = 2\Phi\left(\sqrt{n}\frac{1}{100\sigma}\right) - 1. \quad (16.6)$$

Z είναι μια $N(0, 1)$ τυχαία μεταβλητή. Η τελευταία ποσότητα είναι τουλάχιστον 0.95 αν

$$\Phi\left(\sqrt{n}\frac{1}{100\sigma}\right) \geq 1 - 0.025.$$

Από πίνακες της Φ βρίσκουμε τον μοναδικό αριθμό $z_{0.025}$ που ικανοποιεί $\Phi(z_{0.025}) = 1 - 0.025$. Ισχύει $z_{0.025} \approx 1.96$. Πρέπει $\sqrt{n}\frac{1}{100\sigma} \geq z_{0.025}$ οπότε

$$n \geq (100\sigma z_{0.025})^2 \approx 384.16$$

Άρα παίρνουμε οποιονδήποτε φυσικό n με $n \geq 385$.

Σχόλιο: Στην προσέγγιση της γραμμής (16.5) φαίνεται να επικαλούμαστε την προσέγγιση της (16.3) από την (16.4) για ένα σύνολο A που εξαρτάται από το n . Αυτό σαφώς δεν έπεται από το κεντρικό οριακό θεώρημα. Το θεώρημα ζητάει το σύνολο A να είναι σταθερό. Δικαιολογούμε την (16.5) λέγοντας ότι για σταθερό n , αφού

$$\lim_{k \rightarrow \infty} \mathbf{P}\left(\frac{|S_k - k\mu|}{\sqrt{k\sigma^2}} \leq \sqrt{n}\frac{1}{100\sigma}\right) = \mathbf{P}\left(|Z| \leq \sqrt{n}\frac{1}{100\sigma}\right),$$

ήδη για $k = n$ οι δύο αριθμοί

$$\mathbf{P}\left(\frac{|S_n - n\mu|}{\sqrt{n\sigma^2}} \leq \sqrt{n}\frac{1}{100\sigma}\right), \mathbf{P}\left(|Z| \leq \sqrt{n}\frac{1}{100\sigma}\right) \quad (16.7)$$

είναι πολύ κοντά. Έχουμε στο μυαλό μας ότι το n είναι μεγάλο, επομένως η επιλογή $k = n$ δίνει κάτι κοντά στην οριακή συμπεριφορά.

16.4 Ισχύς των προσεγγίσεων*

Έστω μ, σ, S_n όπως στο κεντρικό οριακό θεώρημα. Για $A \in \mathcal{B}(\mathbb{R})$ με $\mathbf{P}(Z \in \partial A) = 0$, το ότι

$$\lim_{n \rightarrow \infty} \mathbf{P}\left(\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \in A\right) = \mathbf{P}(Z \in A), \quad (16.8)$$

σημαίνει ότι για μεγάλο n οι δύο ποσότητες $\mathbf{P}\left(\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \in A\right), \mathbf{P}(Z \in A)$ είναι πολύ κοντά. Πόσο κοντά όμως; Το πότε η δεύτερη ποσότητα είναι αποδεκτή ως προσέγγιση για την πρώτη εξαρτάται από το τι θέλει να κάνει κανείς. Ένα αυστηρό κριτήριο είναι να ζητήσουμε το πηλίκο τους να βρίσκεται κοντά στο 1 ή ασ πούμε στο διάστημα $[1/2, 2]$ (και να έχουμε τρόπο να το διασφαλίσουμε αυτό). Το κεντρικό οριακό θεώρημα δεν μας δίνει κάποια πληροφορία προς αυτή την κατεύθυνση. Έλεγχος στο λάθος της προσέγγισης δίνουν αποτελέσματα όπως το Θεώρημα Berry-Esseen το οποίο λέει ότι αν επιπλέον για τις X_n υποθέσουμε ότι έχουν $\mathbf{E}|X_1 - \mu|^3 = \rho < \infty$, τότε για κάθε $n \geq 1$ και $x \in \mathbb{R}$ ισχύει

$$\left|\mathbf{P}\left(\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \leq x\right) - \Phi(x)\right| \leq \frac{3\rho}{\sigma^2 \sqrt{n}}.$$

Το Κεντρικό Οριακό Θεώρημα

Και πάλι αυτό το φράγμα είναι μικρής αξίας αν, π.χ., αυτό είναι της τάξης του 10^{-3} και το $\Phi(x)$ είναι της ίδιας τάξης ή μικρότερο (Για θετικό αριθμό a , η ανισότητα $|a - 10^{-3}| \leq 10^{-3}$ δίνει μόνο το άνω φράγμα $a \leq 2 \times 10^{-3}$).

Πειραματιζόμαστε στο παρακάτω παράδειγμα με την ποιότητα της προσέγγισης.

Παράδειγμα 16.5. Έστω ότι οι $(X_n)_{n \geq 1}$ έχουν καθεμία την εκθετική κατανομή με παράμετρο 1. Καταγράφουμε στον πιο κάτω πίνακα την ακριβή τιμή για τις πιθανότητες $\mathbf{P}(S_{100} \geq 125)$, $\mathbf{P}(S_{100} \geq 150)$, $\mathbf{P}(S_{100} \geq 180)$, καθώς και την προσέγγιση που δίνει το κεντρικό οριακό θεώρημα. Έχουμε $\mathbf{E}(S_{100}) = 100$, $\sqrt{\text{Var}(S_{100})} = 10$. Άρα τα γεγονότα $S_{100} \geq 125$, $S_{100} \geq 150$, $S_{100} \geq 180$ είναι σημαντικές αποκλίσεις από τη μέση τιμή (2.5, 5, και 8 φορές αντίστοιχα το μέγεθος της τυπικής απόκλισης).

Πιθανότητα	Προσέγγιση μέσω ΚΟΘ	Ακριβής τιμή
$\mathbf{P}(S_{100} \geq 125)$	0.0062	0.0093
$\mathbf{P}(S_{100} \geq 150)$	2.86×10^{-7}	5.92×10^{-6}
$\mathbf{P}(S_{100} \geq 180)$	6.22×10^{-16}	2.94×10^{-11}

Παρατηρήστε ότι η προσέγγιση για τις πιθανότητες $\mathbf{P}(S_{100} \geq 150)$, $\mathbf{P}(S_{100} \geq 180)$ είναι περίπου το $1/20$ και το $1/10^5$ αντίστοιχα της αληθινής τιμής.

Γενικά η σύγκλιση (16.8) συμβαίνει γρήγορα όταν

- το σύνολο A έχει μεγάλη πιθανότητα $\mathbf{P}(Z \in A)$ (για παράδειγμα, είναι ένα διάστημα όχι πολύ μικρό γύρω από το 0) και
- οι τυχαίες μεταβλητές X_i έχουν κατανομή που δίνει μικρή μάζα σε σύνολα μακριά από το μηδέν. Δηλαδή η $\mathbf{P}(|X_1| > t)$ φθίνει γρήγορα για $t \rightarrow \infty$ όσο όταν $X_1 \sim \mathbf{N}(0, 1)$ ή συντομότερα.

Οι πιθανότητες που μας απασχόλησαν στο προηγούμενο παράδειγμα ήταν ενδεχομένων με πολύ μικρή πιθανότητα (μη τυπική συμπεριφορά της S_n). Για την εκτίμηση της πιθανότητας τέτοιων, απίθανων, ενδεχομένων καταλληλότερη είναι η θεωρία των μεγάλων αποκλίσεων, στοιχεία της οποίας εκτίθενται στο επόμενο κεφάλαιο.

16.5 Η σύγκλιση στο κεντρικό οριακό θεώρημα*

Ας υποθέσουμε ότι οι $(X_n)_{n \geq 1}$ είναι ανεξάρτητες ισόνομες με $\mathbf{E}(X_1) = 0$, $\text{Var}(X_1) = 1$. Το κεντρικό οριακό θεώρημα λέει ότι η ακολουθία $R_n := S_n / \sqrt{n}$ συγκλίνει κατά κατανομή σε μια τυχαία μεταβλητή με κατανομή την $\mathbf{N}(0, 1)$. Μήπως όμως συγκλίνει με πιθανότητα 1 ή έστω κατά πιθανότητα σε κάποια τυχαία μεταβλητή; Η απάντηση είναι όχι. Γιατί, αν συνέκλινε κατά πιθανότητα σε μια τυχαία μεταβλητή W , τότε θα υπήρχε υπακολουθία $(R_{n_k})_{k \geq 1}$ που να συγκλίνει στην W με πιθανότητα 1. Εφαρμόζοντας τα επιχειρήματα της Άσκησης 16.7 [τόρα για την ακολουθία $(R_{n_k})_{k \geq 1}$] δείχνουμε ότι $\underline{\lim}_{k \rightarrow \infty} R_{n_k} = -\infty$, $\overline{\lim}_{k \rightarrow \infty} R_{n_k} = \infty$ με πιθανότητα 1. Άρα η $(R_{n_k})_{k \geq 1}$ δεν μπορεί να συγκλίνει με πιθανότητα 1.

Ασκήσεις

Στις ασκήσεις πιο κάτω, αν $(X_i)_{i \geq 1}$ είναι ακολουθία ανεξάρτητων τυχαίων μεταβλητών, συμβολίζουμε με S_n το άθροισμα των πρώτων n από αυτές. Δηλαδή $S_n := X_1 + X_2 + \dots + X_n$ για κάθε $n \in \mathbb{N}^+$.

Το Κεντρικό Οριακό Θεώρημα

16.1 Έστω $(X_n)_{n \geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με $\mathbf{E}(X_1) = \mu \in \mathbb{R}$ και διασπορά $\sigma^2 \in (0, \infty)$. Να δειχθεί ότι

$$\lim_{n \rightarrow \infty} \mathbf{P}(S_n > n\mu) = \frac{1}{2}.$$

16.2 Έστω $(X_i)_{i \geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με $\mathbf{E}X_1 = 2, \text{Var}(X_1) = 1$. Να υπολογιστούν τα όρια

(α) $\lim_{n \rightarrow \infty} \mathbf{P}(S_n > 2.1n)$,

(β) $\lim_{n \rightarrow \infty} \mathbf{P}(S_n > 2n + \sqrt{n})$,

(γ) $\lim_{n \rightarrow \infty} \mathbf{P}(S_n > 10\sqrt{n})$,

(δ) $\lim_{n \rightarrow \infty} \mathbf{P}(S_n < 3n)$,

(ε) $\lim_{n \rightarrow \infty} \mathbf{P}(S_n > 10^{10})$.

16.3 Έστω $(X_n)_{n \geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές έτσι ώστε $\mathbf{E}(X_1) = 0$ και $\mathbf{E}(X_1^2) = 1$. Να δειχθεί ότι

$$\frac{S_n}{\sqrt{\sum_{i=1}^n X_i^2}} \Rightarrow Z, \text{ όπου } Z \sim \mathbf{N}(0, 1).$$

16.4 Έστω $p \in (0, 1)$ και $A \in \mathbb{R}$. Να υπολογιστεί το όριο

$$\lim_{n \rightarrow \infty} \sum_{k=0}^{np+A\sqrt{n}} \binom{n}{k} p^k (1-p)^{n-k}.$$

16.5 Να δειχθεί ότι

$$\lim_{n \rightarrow \infty} e^{-n} \sum_{0 \leq k \leq nx} \frac{n^k}{k!} = \begin{cases} 0 & \text{αν } 0 \leq x < 1, \\ \frac{1}{2} & \text{αν } x = 1, \\ 1 & \text{αν } x > 1. \end{cases}$$

16.6* Έστω $(X_n)_{n \geq 1}$ ακολουθία ανεξάρτητων ισόνομων τυχαίων μεταβλητών με $\mathbf{E}(X_1) \in \mathbb{R}$ και $\text{Var}(X_1) = 1$. Να υπολογιστεί το όριο

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\left| \frac{X_1 + X_2 + \dots + X_n - (X_{n+1} + \dots + X_{2n})}{\sqrt{n}} \right| \leq 1 \right).$$

16.7 Έστω $(X_n)_{n \geq 1}$ ακολουθία ανεξάρτητων ισόνομων τυχαίων μεταβλητών με $\mathbf{E}(X_1) = 0, \text{Var}(X_1) = 1$. Να δειχθεί ότι

(α)

$$\overline{\lim}_{n \rightarrow \infty} \frac{S_n}{\sqrt{n}} = \infty$$

με πιθανότητα 1.

[Υπόδειξη: Η στρατηγική της Άσκησης 11.16 λειτουργεί με μια μικρή τροποποίηση του μέρους 11.16 (α).]

(β)

$$\underline{\lim}_{n \rightarrow \infty} \frac{|S_n|}{\sqrt{n}} = 0$$

με πιθανότητα 1.

(γ) Το σύνολο των οριακών σημείων της $(S_n / \sqrt{n})_{n \geq 1}$ είναι το $[-\infty, \infty]$ με πιθανότητα 1.

17

Μεγάλες αποκλίσεις*

17.1 Η έννοια της μεγάλης απόκλισης

Έστω $(X_n)_{n \geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές ώστε $\mathbf{P}(X_1 = -1) = \mathbf{P}(X_1 = 1) = 1/2$ και $S_n = \sum_{k=1}^n X_k$ το άθροισμα των πρώτων n από αυτές. Ο νόμος των μεγάλων αριθμών λέει ότι με πιθανότητα 1 ο μέσος όρος S_n/n συγκλίνει στο 0. Μεγάλη απόκλιση για τον μέσο όρο λέμε ένα ενδεχόμενο της μορφής

$$\left\{ \frac{S_n}{n} \in A \right\}$$

όπου $A \subset \mathbb{R}$ είναι ένα σύνολο «μακριά» από το 0, δηλαδή με $0 \notin \bar{A}$. Για παράδειγμα, το A μπορεί να είναι ένα από τα $(1, \infty)$, $(-4, -1) \cup (0.5, 10)$ αλλά όχι το $\{1/n : n \in \mathbb{N}^+\}$. Επειδή η S_n/n συγκλίνει στο 0 με πιθανότητα 1, μια μεγάλη απόκλιση ζητάει από την S_n/n να κάνει κάτι που η ακολουθία δεν θέλει να κάνει. Και η πιθανότητα μιας μεγάλης απόκλισης τείνει στο 0 εξαιτίας του ασθενούς νόμου των μεγάλων αριθμών (Πόρισμα 12.2). Μας ενδιαφέρει να έχουμε μια καλή εκτίμηση του πόσο σύντομα συμβαίνει αυτό. Θα δούμε ότι για πολλά σύνολα A (τα οποία θα προσδιορίσουμε) ισχύει

$$\mathbf{P}\left(\frac{S_n}{n} \in A\right) \approx e^{-c(A)n}, \quad (17.1)$$

όπου $c(A)$ είναι μια θετική σταθερά που εξαρτάται από το σύνολο A . Θα διευκρινίσουμε τη σημασία του \approx και θα υπολογίσουμε αυτή τη σταθερά $c(A)$.

Επίσης, δεν θα περιοριστούμε μόνο στην πιο πάνω ακολουθία $(X_n)_{n \geq 1}$ αλλά θα θεωρήσουμε οποιαδήποτε ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με τιμές στο \mathbb{R} .

Προηγουμένως όμως θα εξηγήσουμε γιατί είναι σημαντικό να ξέρουμε τον ακριβή ρυθμό με τον οποίο φθίνει η πιθανότητα μιας μεγάλης απόκλισης. Γιατί ασχολούμαστε με την πιθανότητα ενός ενδεχομένου που εκ των προτέρων ξέρουμε ότι είναι ελάχιστη (και επομένως δεν περιμένουμε το ενδεχόμενο να συμβεί);

Συμβολισμός: Για $(a_n)_{n \geq 1}, (b_n)_{n \geq 1}$ ακολουθίες θετικών πραγματικών αριθμών γράφουμε $a_n \approx b_n$ αν

$$\lim_{n \rightarrow \infty} \frac{\log a_n}{\log b_n} = 1.$$

Αυτή είναι μια σχέση ισοδυναμίας για ακολουθίες πιο ασθενής από την $\lim_{n \rightarrow \infty} (a_n/b_n) = 1$. Για παράδειγμα, ισχύει $n^3 e^n \approx e^{n+\sqrt{n}}/(n \log n)$.

17.2 Γιατί οι μεγάλες αποκλίσεις είναι σημαντικές

Θεωρούμε το εξής παιχνίδι. Ξεκινάμε με αρχική περιουσία $P_0 = 1$ Ευρώ και πραγματοποιούμε μια ακολουθία ρίψεων ενός αμερόληπτου νομίσματος. Όποτε το νόμισμα φέρνει «Κεφαλή», η περιουσία μας διπλασιάζεται, όποτε φέρνει «Γράμματα», η περιουσία μας υποδιπλασιάζεται.

Ερώτημα: Ποια είναι η μέση τιμή της περιουσίας μετά από n βήματα;

Η περιουσία μετά n βήματα είναι $P_n = 2^{S_n}$, όπου S_n είναι η ακολουθία της προηγούμενης ενότητας.

Μια διαισθητική προσέγγιση: Έστω $\varepsilon_n := S_n/n$, που ξέρουμε ότι τείνει στο μηδέν με πιθανότητα 1. Τότε

$$\mathbf{E}(P_n) = \mathbf{E}(e^{S_n}) = \mathbf{E}(e^{n\varepsilon_n}) \doteq e^{na_n}$$

με a_n ακολουθία που τείνει στο 0. Η τελευταία ισότητα είναι μια εικασία. Παίρνουμε μέση τιμή μιας ποσότητας με ρυθμό εκθετικής αύξησης $\frac{1}{n} \log e^{n\varepsilon_n} (= \varepsilon_n)$ που είναι περίπου 0. Αναμένουμε η συνολική μέση τιμή να έχει ρυθμό εκθετικής αύξησης επίσης περίπου 0.

Τι πραγματικά συμβαίνει: Η μέση τιμή $\mathbf{E}(P_n)$ υπολογίζεται άμεσα ως

$$\mathbf{E}(P_n) = \mathbf{E}(2^{X_1})^n = \left(\frac{2 + 2^{-1}}{2} \right)^n = e^{n \log(5/4)}. \quad (17.2)$$

Δηλαδή έχει θετικό εκθετικό ρυθμό αύξησης ίσο με $\log(5/4)$.

Εξήγηση: Ποιο είναι το πρόβλημα με τη διαισθητική προσέγγιση πιο πάνω; Το κλάσμα $\varepsilon_n := S_n/n$ παίρνει τιμές στο $U_n := \{k/n : k \in \mathbb{Z}, |k| \leq n\}$. Προσεγγιστικά ισχύει

$$\mathbf{P}(\varepsilon_n = x) \approx e^{-nI(x)},$$

με I μια συνεχή συνάρτηση στο $[-1, 1]$ περίπου της μορφής x^2 . Δηλαδή τιμές του x μακριά από το 0 είναι δύσκολο να ληφθούν από την S_n/n .

Ο υπολογισμός της $\mathbf{E}(P_n)$ γίνεται ως εξής:

$$\mathbf{E}(2^{S_n}) = \sum_{x \in U_n} e^{nx \log 2} \mathbf{P}(\varepsilon_n = x). \quad (17.3)$$

Η διαισθητική προσέγγιση πρότεινε να αγνοήσουμε όλους τους όρους με $x \neq 0$ γιατί έχουν πολύ μικρή πιθανότητα. Βέβαια κάθε τέτοιος όρος δεν έχει μόνο κόστος (συγκεκριμένα $\approx e^{-nI(x)}$) αλλά και όφελος (συγκεκριμένα $e^{nx \log 2}$) το οποίο ίσως να ισοσκελίζει το κόστος. Κυρίαρχος όρος στο άθροισμα είναι αυτός που μεγιστοποιεί τη διαφορά $x \log 2 - I(x)$ (όφελος μείον κόστος). Πιο κάτω που θα έχουμε την ακριβή μορφή της συνάρτησης I (Παράδειγμα 17.9), θα δούμε ότι το καλύτερο x είναι το $x = 3/5$. Η μέγιστη συνεισφορά στη μέση τιμή προέρχεται από μια μεγάλη απόκλιση. Η τυπική συμπεριφορά του μέσου S_n/n είναι αδιάφορη στον υπολογισμό.

Στο πιο πάνω πρόβλημα η επίκληση των μεγάλων αποκλίσεων δεν ήταν απαραίτητη αφού υπάρχει απλούστερος τρόπος αντιμετώπισης. Υπάρχουν όμως άλλα προβλήματα στα οποία μια μεγάλη απόκλιση παίζει κεντρικό ρόλο και η θεωρία των μεγάλων αποκλίσεων είναι το μόνο διαθέσιμο εργαλείο.

17.3 Η αρχή μεγάλων αποκλίσεων

Έστω X μετρικός χώρος. **Συνάρτηση ρυθμού** στον X ονομάζουμε οποιαδήποτε συνάρτηση $I : X \rightarrow [0, \infty]$ που είναι κάτω ημισυνεχής [δηλαδή το σύνολο $[I > a]$ είναι ανοιχτό για κάθε $a \in \mathbb{R}$].

Έστω τώρα $(\mu_n)_{n \geq 1}$ ακολουθία μέτρων πιθανότητας στον $(X, \mathcal{B}(X))$ και $(a_n)_{n \geq 1}$ αύξουσα ακολουθία θετικών αριθμών με $\lim_{n \rightarrow \infty} a_n = \infty$.

Ορισμός 17.1. Λέμε ότι η ακολουθία $(\mu_n)_{n \geq 1}$ ικανοποιεί την **αρχή μεγάλων αποκλίσεων με ταχύτητα** a_n και **συνάρτηση ρυθμού** I αν για κάθε $A \in \mathcal{B}(X)$ ισχύει

$$-\inf_{x \in A^\circ} I(x) \leq \liminf_{n \rightarrow \infty} \frac{1}{a_n} \log \mu_n(A) \leq \overline{\lim}_{n \rightarrow \infty} \frac{1}{a_n} \log \mu_n(A) \leq -\inf_{x \in \bar{A}} I(x). \quad (17.4)$$

Μεγάλες αποκλίσεις*

Στην πράξη, συνήθως έχουμε μια ακολουθία $(Y_n)_{n \geq 1}$ τυχαίων μεταβλητών στον \mathcal{X} που συγκλίνει κατά πιθανότητα σε ένα σημείο x_0 του \mathcal{X} και εξετάζουμε αν η ακολουθία $(\mu_n)_{n \geq 1}$ των κατανομών των Y_n ικανοποιεί την αρχή των μεγάλων αποκλίσεων. Αν την ικανοποιεί, λέμε ότι η ακολουθία $(Y_n)_{n \geq 1}$ ικανοποιεί την αρχή μεγάλων αποκλίσεων.

Παράδειγμα 17.2. Έστω Y_n ακολουθία τυχαίων μεταβλητών (στον ίδιο χώρο πιθανότητας) με την Y_n να ακολουθεί την εκθετική κατανομή με παράμετρο n (και άρα μέση τιμή $1/n$). Η Y_n συγκλίνει κατά πιθανότητα στο 0. Η ακολουθία (των κατανομών) των Y_n ικανοποιεί την αρχή των μεγάλων αποκλίσεων με ταχύτητα n και συνάρτηση ρυθμού

$$I(x) = \begin{cases} \infty & \text{αν } x < 0, \\ x & \text{αν } x \geq 0. \end{cases}$$

Η απόδειξη αφήνεται ως άσκηση.

Παρατήρηση 17.3. (α) Για κάθε σύνολο $A \subset \mathcal{X}$, εισάγουμε τη συντομογραφία

$$I(A) = \inf_{x \in A} I(x).$$

(β) Όταν για ένα σύνολο Borel $A \subset X$ ισχύει $I(A^\circ) = I(\bar{A})$, τότε έχουμε ότι η $\frac{1}{n} \log \mu_n(A)$ συγκλίνει στην τιμή $I(A^\circ) = I(A) = I(\bar{A})$. Δηλαδή

$$\mu_n(A) \approx e^{-nI(A)}.$$

(γ) Η (17.4) ισοδυναμεί με την απαίτηση το άνω φράγμα να ισχύει για A κλειστό και το κάτω φράγμα να ισχύει για A ανοιχτό. Δηλαδή

$$\overline{\lim}_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(F) \leq -\inf_{x \in F} I(x) \quad (17.5)$$

για κάθε $F \subset X$ κλειστό και

$$\underline{\lim}_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(G) \geq -\inf_{x \in G} I(x) \quad (17.6)$$

για κάθε $G \subset X$ ανοιχτό. Επιπλέον, το κάτω φράγμα ισοδυναμεί με το εξής: Για κάθε $x \in X$ και ανοιχτό σύνολο $G \subset X$ που περιέχει το x ισχύει

$$\underline{\lim}_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(G) \geq -I(x). \quad (17.7)$$

Για την απόδειξη της αρχής μεγάλων αποκλίσεων, θα χρησιμοποιούμε αυτές τις ισοδύναμες μορφές του ορισμού.

17.4 Το Θεώρημα Cramer

Για $f : \mathbb{R} \rightarrow [-\infty, \infty]$, ορίζουμε τον μετασχηματισμό Legendre της f ως τη συνάρτηση $f^* : \mathbb{R} \rightarrow [-\infty, \infty]$ με

$$f^*(x) := \sup_{t \in \mathbb{R}} \{xt - f(t)\}$$

για κάθε $x \in \mathbb{R}$, όπου υπενθυμίζουμε ότι $\sup \emptyset = -\infty$ και $\sup A = \infty$ αν το $A \subset \mathbb{R}$ είναι μη φραγμένο.

Μεγάλες αποκλίσεις*

Έστω τώρα $(X_n)_{n \geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με τιμές στο \mathbb{R} και μ η κατανομή καθεμιάς. Συμβολίζουμε με M τη ροπογεννήτρια της X_1 , με Λ τον λογάριθμο της M και με Λ^* τον μετασχηματισμό Legendre της Λ . Δηλαδή

$$M(\lambda) := \mathbf{E}(e^{\lambda X}) = \int e^{\lambda x} d\mu(x), \quad (17.8)$$

$$\Lambda(\lambda) := \log M(\lambda), \quad (17.9)$$

$$\Lambda^*(x) := \sup_{\lambda \in \mathbb{R}} \{\lambda x - \Lambda(\lambda)\} \quad (17.10)$$

για κάθε $\lambda, x \in \mathbb{R}$.

Παράδειγμα 17.4. Ας δούμε την περίπτωση που η X_1 είναι η ομοιόμορφη στο $\{-1, 1\}$. Τότε

$$\Lambda(\lambda) = \log \left(\frac{e^{-\lambda} + e^{\lambda}}{2} \right)$$

για κάθε $\lambda \in \mathbb{R}$ και είναι άσκηση απειροστικού λογισμού (μεγιστοποίησης) να δείξει κανείς ότι

$$\Lambda^*(x) = \begin{cases} \frac{1}{2} \{(1+x) \log(1+x) + (1-x) \log(1-x)\} & \text{αν } x \in [-1, 1], \\ \infty & \text{αν } x \in \mathbb{R} \setminus [-1, 1], \end{cases} \quad (17.11)$$

με τη σύμβαση $0 \log 0 = 0$.

Το θεώρημα Cramer λέει ότι η ακολουθία $(S_n/n)_{n \geq 1}$ ικανοποιεί την αρχή μεγάλων αποκλίσεων με ταχύτητα n και συνάρτηση ρυθμού Λ^* . Ξεκινάμε με δύο λήμματα που ουσιαστικά αποδεικνύουν το άνω φράγμα της αρχής.

Λήμμα 17.5. Για κάθε $x \in \mathbb{R}$ ισχύει

$$\mathbf{P} \left(\frac{S_n}{n} \geq x \right) \leq e^{-n \sup_{\lambda \geq 0} \{\lambda x - \Lambda(\lambda)\}}, \quad (17.12)$$

$$\mathbf{P} \left(\frac{S_n}{n} \leq x \right) \leq e^{-n \sup_{\lambda \leq 0} \{\lambda x - \Lambda(\lambda)\}}. \quad (17.13)$$

Δηλαδή μια απόκλιση προς τα πάνω ελέγχεται από τις τιμές της ροπογεννήτριας $M(\lambda)$ για $\lambda \geq 0$ ενώ μια απόκλιση προς τα κάτω ελέγχεται από τις τιμές της $M(\lambda)$ για $\lambda \leq 0$.

Απόδειξη. Για $\lambda \geq 0$, εφαρμόζοντας την ανισότητα Markov, έχουμε

$$\begin{aligned} \mathbf{P} \left(\frac{S_n}{n} \geq x \right) &= \mathbf{P}(S_n \geq nx) = \mathbf{P}(\lambda S_n \geq \lambda nx) = \mathbf{P}(e^{\lambda S_n} \geq e^{\lambda nx}) \leq e^{-\lambda nx} \mathbf{E}(e^{\lambda S_n}) \\ &= e^{-\lambda nx} M(\lambda)^n = e^{n\Lambda(\lambda) - \lambda nx} = e^{-n\{\lambda x - \Lambda(\lambda)\}}. \end{aligned}$$

Επειδή το φράγμα ισχύει για κάθε $\lambda \geq 0$, η ιδέα είναι να διαλέξουμε το λ που δίνει το καλύτερο/μικρότερο φράγμα. Συγκεκριμένα παίρνουμε ότι η πιθανότητα $\mathbf{P}(S_n/n \geq x)$ φράσσεται πάνω από την ποσότητα

$$\inf_{\lambda \geq 0} e^{-n\{\lambda x - \Lambda(\lambda)\}} = e^{-n \sup_{\lambda \geq 0} \{\lambda x - \Lambda(\lambda)\}}.$$

Η πρώτη ανισότητα αποδείχθηκε.

Για την απόδειξη της δεύτερης, παρατηρούμε ότι για $\lambda \leq 0$ ισχύει

$$\mathbf{P} \left(\frac{S_n}{n} \leq x \right) = \mathbf{P}(S_n \leq nx) = \mathbf{P}(\lambda S_n \geq \lambda nx) = \mathbf{P}(e^{\lambda S_n} \geq e^{\lambda nx}) \leq e^{-\lambda nx} \mathbf{E}(e^{\lambda S_n}) = e^{-n\{\lambda x - \Lambda(\lambda)\}}.$$

Και η απόδειξη συνεχίζεται όπως και για την πρώτη ανισότητα. ■

Λήμμα 17.6. Υποθέτουμε ότι $m := \mathbf{E}(X_1) \in \mathbb{R}$. Τότε

$$(i) \quad x \geq m \Rightarrow \Lambda^*(x) = \sup_{\lambda \geq 0} \{\lambda x - \Lambda(\lambda)\}.$$

$$(ii) \quad x \leq m \Rightarrow \Lambda^*(x) = \sup_{\lambda \leq 0} \{\lambda x - \Lambda(\lambda)\}.$$

$$(iii) \quad \Lambda^*(m) = 0.$$

Απόδειξη. (i) Για κάθε $\lambda \in \mathbb{R}$, εφαρμόζοντας την ανισότητα Jensen έχουμε

$$\Lambda(\lambda) = \log \mathbf{E}(e^{\lambda X_1}) \geq \mathbf{E}(\lambda X_1) = \lambda m,$$

επομένως $\lambda m - \Lambda(\lambda) \leq 0$. Τώρα θέλουμε να δείξουμε ότι στο supremum που ορίζει το $\Lambda^*(x)$ μπορούμε να αγνοήσουμε τους αριθμούς $\lambda x - \Lambda(\lambda)$ που έχουν $\lambda < 0$. Πράγματι, για $x \geq m$ και $\lambda < 0$ έχουμε $\lambda x \leq \lambda m \leq \Lambda(\lambda)$ όπως δείξαμε πιο πάνω, οπότε $\lambda x - \Lambda(\lambda) \leq 0$. Όμως 0 είναι η τιμή του $\lambda x - \Lambda(\lambda)$ όταν $\lambda = 0$. Άρα οι όροι με $\lambda < 0$ δεν μπορούν να αυξήσουν το supremum.

(ii) Η απόδειξη είναι ανάλογη με αυτήν στο (i).

(iii) Όταν $x = m$, οι (i), (ii) δίνουν ότι το $\Lambda^*(m)$ ισούται με την τιμή του $\lambda x - \Lambda(\lambda)$ για $\lambda = 0$, η οποία είναι 0. ■

Το επόμενο λήμμα είναι κρίσιμο για την απόδειξη του κάτω φράγματος της αρχής μεγάλων αποκλίσεων.

Λήμμα 17.7. (α) Η M είναι διαφορίσιμη στο εσωτερικό του $D_M := \{\lambda \in \mathbb{R} : M(\lambda) < \infty\}$ με παράγωγο $M'(\lambda) = \mathbf{E}(X_1 e^{\lambda X_1})$.

(β) Αν $\mu((-\infty, a)), \mu((a, \infty)) > 0$ και το μ έχει συμπαγή φορέα τότε $D_M = \mathbb{R}$ και υπάρχει $\lambda_0 \in \mathbb{R}$ ώστε $\Lambda^*(a) = \lambda_0 a - \Lambda(\lambda_0)$. Για αυτό το λ_0 ισχύει $\Lambda'(\lambda_0) = a$

Απόδειξη. (α) Ο τύπος για την παράγωγο προκύπτει διαφορίζοντας την $\mathbf{E}(e^{\lambda X_1})$ μέσα από την μέση τιμή. Εφαρμόζουμε την Πρόταση Α'5 από το Παράρτημα Α'. Έστω λ εσωτερικό σημείο του D_M και $\delta > 0$ με $[\lambda - 2\delta, \lambda + 2\delta] \subset D_M$. Θέτουμε $f(\omega, t) = e^{tX(\omega)}$ για κάθε $(\omega, t) \in \Omega \times [\lambda - \delta, \lambda + \delta]$. Οι υποθέσεις (i), (ii) της πρότασης ισχύουν προφανώς. Για την (iii), παρατηρούμε ότι

$$\begin{aligned} \sup_{t \in [\lambda - \delta, \lambda + \delta]} |\partial_t f(\omega, t)| &\leq |X(\omega)| \{e^{(\lambda + \delta)X(\omega)} + e^{(\lambda - \delta)X(\omega)}\} \leq \frac{e^{\delta|X(\omega)|}}{\delta} \{e^{(\lambda + \delta)X(\omega)} + e^{(\lambda - \delta)X(\omega)}\} \\ &\leq \frac{1}{\delta} \{e^{(\lambda - 2\delta)X(\omega)} + e^{\lambda X(\omega)} + e^{(\lambda + 2\delta)X(\omega)}\}. \end{aligned}$$

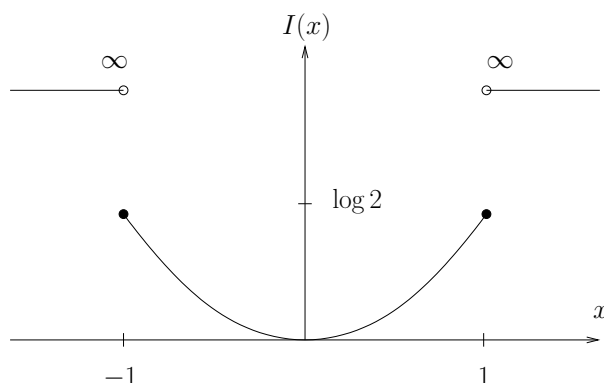
Η τυχαία μεταβλητή στην τελευταία γραμμή έχει πεπερασμένη μέση τιμή εξαιτίας του ότι $\lambda - 2\delta, \lambda, \lambda + 2\delta \in D_M$. Έτσι, η Πρόταση Α'5 δίνει το συμπέρασμα.

(β) Έχουμε $\Lambda^*(a) = \sup_{\lambda \in \mathbb{R}} A(\lambda)$ με $A(\lambda) := \lambda a - \Lambda(\lambda) = -\log \mathbf{E}(e^{\lambda(X_1 - a)})$. Η A είναι πεπερασμένη και διαφορίσιμη στο \mathbb{R} με όρια $A(-\infty) = A(\infty) = -\infty$ εξαιτίας της $\mu((-\infty, a)), \mu((a, \infty)) > 0$. Άρα παίρνει μέγιστο σε ένα σημείο $\lambda_0 \in \mathbb{R}$ και $0 = A'(\lambda_0) = a - \Lambda'(\lambda_0)$. Ο ισχυρισμός αποδείχθηκε. ■

Θεώρημα 17.8 (Θεώρημα Cramer). Υποθέτουμε ότι $m := \mathbf{E}(X_1) \in \mathbb{R}$. Η ακολουθία $(S_n/n)_{n \geq 1}$ ικανοποιεί την αρχή μεγάλων αποκλίσεων με ταχύτητα n και συνάρτηση ρυθμού $I(x) := \Lambda^*(x)$.

Η απόδειξη του θεωρήματος δίνεται στο Παράρτημα Β'. Το συμπέρασμα ισχύει ακόμα και χωρίς την υπόθεση ότι η $\mathbf{E}(X_1)$ ορίζεται και είναι πραγματικός αριθμός. Αυτό αποδεικνύεται με λίγες παρεμβάσεις στην απόδειξη που περιγράφουμε στο παράρτημα (δες Dembo and Zeitouni (1998), Θεώρημα 2.2.3).

Μεγάλες αποκλίσεις*



Σχήμα 17.1: Η συνάρτηση ρυθμού της αρχής μεγάλων αποκλίσεων για τον μέσο όρο ομοιόμορφων στο $\{-1, 1\}$.

Παράδειγμα 17.9. Το θεώρημα Cramer εφαρμόζεται στην ακολουθία $(S_n/n)_{n \geq 1}$ της Παραγράφου 17.1 και δίνει ότι αυτή ικανοποιεί την αρχή μεγάλων αποκλίσεων με συνάρτηση ρυθμού $I(x)$ τη $\Lambda^*(x)$ της (17.11). Το γράφημά της δίνεται στο Σχήμα 17.1. Να παρατηρήσουμε τα εξής:

- Η I έχει την τιμή 0 στη μέση τιμή $\mathbf{E}(X_1) = 0$.
- Η I έχει την τιμή ∞ για $x \notin [-1, 1]$, που είναι αναμενόμενο αφού η S_n/n παίρνει τιμές στο $[-1, 1]$.
- Όσο απομακρυνόμαστε από το 0 (τη μέση τιμή των X_i), η $I(x)$ αυξάνει. Το γεγονός $\{S_n/n \text{ είναι κοντά στο } x\}$ γίνεται ακριβότερο/πιο απίθανο.

Τώρα μπορούμε να επιστρέψουμε στην Παράγραφο 17.2 και να δούμε ότι πράγματι η διαφορά $x \log 2 - I(x)$ λαμβάνει τη μέγιστη τιμή της όταν $x = 3/5$ και αυτή η τιμή είναι η $\log(5/4)$, σε συμφωνία με την (17.2).

Ασκήσεις

17.1 Να υπολογιστεί ο μετασχηματισμός Λ^* στην περίπτωση που η X_1 ακολουθεί την κατανομή:

(α) $\text{Poisson}(a)$,

(β) $\exp(a)$,

(γ) $N(0, \sigma^2)$,

όπου $a, \sigma > 0$. Επίσης, με χρήση Mathematica ή άλλου προγράμματος να γίνει σε κάθε περίπτωση η γραφική παράσταση του Λ^* .

17.2 Να υπολογιστεί ο μετασχηματισμός Λ^* στην περίπτωση που η X_1 έχει πυκνότητα $f(x) = (3/2)|x|^{-4} \mathbf{1}_{|x| \geq 1}$. Τι πληροφορίες δίνει το άνω και το κάτω φράγμα της αρχής μεγάλων αποκλίσεων για την ακολουθία S_n/n ;

17.3 Εκτελούμε μια ακολουθία ανεξάρτητων ρίψεων ενός αμερόληπτου νομίσματος και ονομάζουμε S_n το πλήθος των φορών που ήρθε η ένδειξη «Κεφαλή» στις πρώτες n ρίψεις.

(α) Να δειχθεί ότι η ακολουθία $(S_n/n)_{n \geq 1}$ ικανοποιεί την αρχή μεγάλων αποκλίσεων με ταχύτητα n και συνάρτηση ρυθμού

$$I(x) := \begin{cases} \log 2 + x \log x + (1-x) \log(1-x) & \text{αν } x \in [0, 1], \\ \infty & \text{αν } x \in \mathbb{R} \setminus [0, 1]. \end{cases}$$

Μεγάλες αποκλίσεις*

(β) Για την πιθανότητα $\mathbf{P}(S_{1000} \geq 700)$ να προσδιοριστεί το άνω φράγμα που δίνουν τα Λήμματα 17.5 και 17.6 και η προσέγγιση που δίνει το κεντρικό οριακό θεώρημα.

17.4 Έστω ακολουθία μέτρων πιθανότητας $(\mu_n)_{n \geq 1}$ σε έναν μετρικό χώρο \mathcal{X} η οποία ικανοποιεί την αρχή μεγάλων αποκλίσεων με ταχύτητα a_n και συνάρτηση ρυθμού I . Να δειχθεί ότι $\inf\{I(x) : x \in \mathcal{X}\} = 0$.

17.5 Για κάθε $n \in \mathbb{N}^+$, έστω Y_n τυχαία μεταβλητή που ακολουθεί την κατανομή $N(0, 1/n)$. Να δειχθεί ότι η ακολουθία $(Y_n)_{n \geq 1}$ ικανοποιεί την αρχή μεγάλων αποκλίσεων με ταχύτητα n και συνάρτηση ρυθμού $I(x) = x^2/2$, $x \in \mathbb{R}$.

17.6 Να αποδειχθεί ο ισχυρισμός του Παραδείγματος 17.2.

17.7 Έστω $f : \mathbb{R} \rightarrow [-\infty, \infty]$ και $D_f := \{x \in \mathbb{R} : f(x) < \infty\}$.

(α) Αν $0 \in D_f^\circ$, τότε

$$\liminf_{|x| \rightarrow \infty} \frac{f^*(x)}{|x|} > 0,$$

και άρα $\lim_{|x| \rightarrow \infty} f^*(x) = \infty$.

(β) Αν $D_f = \mathbb{R}$, τότε

$$\lim_{|x| \rightarrow \infty} \frac{f^*(x)}{|x|} = \infty.$$

17.8 Έστω ότι η τυχαία μεταβλητή X_1 (με τιμές στο \mathbb{R}) έχει μέση τιμή $m = \mathbf{E}(X_1)$ και ροπογεννήτρια M η οποία είναι πεπερασμένη για όλα τα λ σε μια περιοχή του μηδενός. Να δειχθεί ότι για τη συνάρτηση ρυθμού της αρχής μεγάλων αποκλίσεων που δίνει το θεώρημα Cramer ισχύει $I(x) > 0$ για κάθε $x \neq m$. Οπότε, παίρνοντας υπόψιν το Λήμμα 17.6(iii), έχουμε ότι το m είναι το μοναδικό μηδενικό της I .

17.9 Έστω $(X_n)_{n \geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με τιμές στο $[0, \infty)$ ώστε η ροπογεννήτρια της X_1 να είναι η

$$M(t) = \begin{cases} e^{-C|t|^a} & \text{αν } t \leq 0, \\ \infty & \text{αν } t > 0, \end{cases}$$

όπου $C > 0$ και $a \in (0, 1)$. Θέτουμε $S_k := X_1 + X_2 + \dots + X_k$ για κάθε $k \in \mathbb{N}^+$. Να δειχθεί ότι για κάθε $t > 0$ ισχύει

$$\mathbf{P}(S_k < tk^{1/a}) \leq e^{-C_1 t^{-\frac{a}{1-a}}} \quad (17.14)$$

με $C_1 := (1-a)(Ca^a)^{(1-a)^{-1}}$.

Σχόλια: 1) Αποδεικνύεται ότι για κάθε $C > 0$ και $a \in (0, 1)$ υπάρχει τυχαία μεταβλητή με ροπογεννήτρια όπως πιο πάνω. Μάλιστα αυτή η τυχαία μεταβλητή έχει μέση τιμή ∞ .

2) Μπορούμε να δείξουμε ότι η $S_k/k^{1/a}$ συγκλίνει κατά κατανομή σε μια μη σταθερή τυχαία μεταβλητή Y με πυκνότητα. Άρα το όριο για $k \rightarrow \infty$ της πιθανότητας στην (17.14) είναι $\mathbf{P}(Y < t)$. Προσέξτε ότι η (17.14) ισχύει για όλα τα k και όχι απλώς για τα μεγάλα k .

Παραρτήματα

Α'

Αναλυτικά αποτελέσματα

Α'.1 Αθροίσματα

Στόχος αυτής της παραγράφου είναι να δώσει νόημα στο σύμβολο

$$\sum_{i \in I} a_i,$$

όπου I είναι οποιοδήποτε σύνολο και $(a_i)_{i \in I}$ είναι πραγματικοί αριθμοί.

Περίπτωση 1. Αν $a_i \geq 0$ για κάθε $i \in I$.

Τότε θέτουμε

$$\sum_{i \in I} a_i := \sup \left\{ \sum_{i \in J} a_i : J \subset I \text{ πεπερασμένο} \right\}.$$

Το νόημα του $\sum_{i \in J} a_i$ για $J \subset I$ πεπερασμένο και μη κενό είναι ξεκάθαρο, ενώ όταν $J = \emptyset$, ορίζεται ως 0.

Δεν είναι απαραίτητο το I να είναι αριθμήσιμο. Όμως εύκολα δείχνει κανείς ότι αν το άθροισμα $\sum_{i \in I} a_i$ είναι πεπερασμένο, τότε το $\{i \in I : a_i > 0\}$ είναι αριθμήσιμο.

Αν το I είναι άπειρο αριθμήσιμο, τότε το άθροισμα συμπίπτει με το όριο της σειράς των a_i . Πιο συγκεκριμένα, αν το $I = \mathbb{N}^+$, τότε

$$\sum_{i \in \mathbb{N}^+} a_i = \sum_{i=1}^{\infty} a_i.$$

Η απόδειξη αυτής της ισότητας είναι απλή και αφήνεται ως άσκηση.

Περίπτωση 2. Αν $a_i \in \mathbb{R}$ για κάθε $i \in I$.

Τότε θέτουμε

$$\sum_{i \in I} a_i := \sum_{i \in I} a_i^+ - \sum_{i \in I} a_i^-$$

όποτε η διαφορά ορίζεται. Δηλαδή όποτε δεν έχουμε την μορφή $\infty - \infty$.

Α'.2 Όρια γινομένων

Κύριος στόχος αυτής της παραγράφου είναι η διατύπωση του Λήμματος Α'.2 παρακάτω, που αφορά υπολογισμό ορίων απειρογινόμενων και χρησιμοποιείται σε αποδείξεις σύγκλισης κατά κατανομή μέσω χαρακτηριστικών συναρτήσεων.

Για $z \in \mathbb{C} \setminus (-\infty, 0]$, με $\log z$ συμβολίζουμε τον κύριο κλάδο του λογαρίθμου του z . Δηλαδή

$$\log z = \log |z| + i \operatorname{Arg}(z),$$

όπου $\operatorname{Arg}(z) \in (-\pi, \pi)$ είναι το όρισμα του z .

Για $z \in \mathbb{C}$ κοντά στο 0, έχουμε $e^z \simeq 1 + z$ και $\log(1 + z) \simeq z$.

Αυτές οι δύο προσεγγίσεις μας καθοδηγούν όταν κάνουμε ασυμπτωτική ανάλυση απειρογινόμενων. Και έπειτα, για να δικαιολογήσουμε αυστηρά το αποτέλεσμα που μας υποδεικνύουν, χρησιμοποιούμε κάποια από τις ανισότητες στο παρακάτω λήμμα.

Λήμμα Α'.1. (i) $e^x \geq 1 + x$ για κάθε $x \in \mathbb{R}$.

(ii) $\log x \leq x - 1$ για κάθε $x \in (0, \infty)$.

(iii) $|e^z - (1 + z)| \leq |z|^2$ για κάθε $z \in \mathbb{C}$ με $|z| \leq 1$.

(iv) $|\log(1 + z) - z| \leq |z|^2$ για κάθε $z \in \mathbb{C}$ με $|z| \leq 1/2$.

Απόδειξη. Τα (i), (ii) είναι γνωστά από το λύκειο.

(iii) Με χρήση της δυναμοσειράς για την e^z έχουμε

$$|e^z - (1 + z)| = \left| \sum_{k=2}^{\infty} \frac{z^k}{k!} \right| \leq |z|^2 \sum_{k=2}^{\infty} \frac{1}{k!} = (e - 2)|z|^2 \leq |z|^2.$$

Για την πρώτη ανισότητα, βγάλαμε κοινό παράγοντα το $|z|^2$ και χρησιμοποιήσαμε το ότι $|z| \leq 1$.

(iv) Για $z \in \mathbb{C}$ με $|z| \leq 1/2$ ισχύει

$$|\log(1 + z) - z| = \left| \sum_{k=2}^{\infty} \frac{(-1)^{k-1}}{k} z^k \right| \leq \frac{1}{2} \sum_{k=2}^{\infty} |z|^k = \frac{|z|^2}{2} \frac{1}{1 - |z|} \leq |z|^2.$$

Η πρώτη ισότητα ισχύει για κάθε z με $|z| < 1$, ενώ η τελευταία ανισότητα για $|z| \leq 1/2$. ■

Λήμμα Α'.2. Έστω $(k_n)_{n \geq 1}$ ακολουθία φυσικών αριθμών και $\{a_{n,j} : n \geq 1, 1 \leq j \leq k_n\}$ μιγαδικοί αριθμοί. Υποθέτουμε ότι

(i) $\lim_{n \rightarrow \infty} \sum_{j=1}^{k_n} a_{n,j} = A$ με $A \in \mathbb{C}$ και

(ii) $\lim_{n \rightarrow \infty} \sum_{j=1}^{k_n} |a_{n,j}|^2 = 0$.

Τότε

$$\lim_{n \rightarrow \infty} \prod_{j=1}^{k_n} (1 + a_{n,j}) = e^A.$$

[Το συμπέρασμα είναι αναμενόμενο αφού λόγω του (ii) όλα τα $a_{n,j}$ είναι κοντά στο 0 και άρα

$$\prod_{j=1}^{k_n} (1 + a_{n,j}) \simeq \prod_{j=1}^{k_n} e^{a_{n,j}} = e^{\sum_{j=1}^{k_n} a_{n,j}}.$$

Ο εκθέτης στην τελευταία έκφραση τείνει στο A . Πρέπει να δείξουμε βέβαια αυστηρά ότι το \simeq στο όριο γίνεται =.]

Απόδειξη. Από την υπόθεση (ii), υπάρχει $n_0 \in \mathbb{N}^+$ ώστε $|a_{n,j}| \leq 1/2$ για κάθε $n \geq n_0$ και $1 \leq j \leq k_n$. Επομένως $1 + a_{n,j} \in \mathbb{C} \setminus (-\infty, 0]$ και

$$\frac{\prod_{j=1}^{k_n} (1 + a_{n,j})}{e^{\sum_{j=1}^{k_n} a_{n,j}}} = e^{\sum_{j=1}^{k_n} \{\log(1 + a_{n,j}) - a_{n,j}\}} \quad (\text{A'.1})$$

Ο εκθέτης με χρήση του Λήμματος Α'.1(iv) φράσσεται ως εξής

$$\left| \sum_{j=1}^{k_n} \{\log(1 + a_{n,j}) - a_{n,j}\} \right| \leq \sum_{j=1}^{k_n} |\log(1 + a_{n,j}) - a_{n,j}| \leq \sum_{j=1}^{k_n} |a_{n,j}|^2.$$

Η τελευταία ποσότητα τείνει στο 0 από υπόθεση. Άρα το πηλίκο στην (Α'.1) τείνει στο 1 και το λήμμα αποδείχθηκε. ■

Πόρισμα Α'.3. Αν $(c_n)_{n \in \mathbb{N}}$ ακολουθία στο \mathbb{C} τέτοια ώστε $c_n \rightarrow c$, με $c \in \mathbb{C}$, τότε

$$\lim_{n \rightarrow \infty} \left(1 + \frac{c_n}{n}\right)^n = e^c.$$

Απόδειξη. Εφαρμόζουμε το Λήμμα Α'.2 με $k_n = n$, $a_{n,j} = c_n/n$ για κάθε $n \geq 1$ και $1 \leq j \leq n$. Έχουμε $\sum_{j=1}^n \frac{c_n}{n} = c_n \rightarrow c$ και $\sum_{j=1}^n \left(\frac{c_n}{n}\right)^2 = \frac{c_n^2}{n} \rightarrow 0$ για $n \rightarrow \infty$. ■

Α'.3 Η συνάρτηση Γάμμα

Διατυπώνουμε την προσέγγιση Stirling για τη συνάρτηση Γ. Ο ορισμός της Γ είναι

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$$

για κάθε $x > 0$. Αποδεικνύονται εύκολα οι εξής βασικές ιδιότητές της.

- (i) $\Gamma(1) = 1$.
- (ii) $\Gamma(1/2) = \sqrt{\pi}$.
- (iii) $\Gamma(x+1) = x\Gamma(x)$ για κάθε $x > 0$.
- (iv) $\Gamma(n) = (n-1)!$ για κάθε $n \in \mathbb{N}^+$.

Η προσέγγιση Stirling είναι η εξής.

Θεώρημα Α'.4. Για κάθε $x > 0$, υπάρχει $\theta_x \in (0, 1)$ έτσι ώστε

$$\Gamma(x) = \left(\frac{x}{e}\right)^x \frac{\sqrt{2\pi}}{\sqrt{x}} e^{\theta_x/(12x)}. \quad (\text{Α'.2})$$

Για την απόδειξη του θεωρήματος, δες την Παράγραφο 12.33 στο Whittaker and Watson (1965). Για $n \in \mathbb{N}^+$, έχουμε την εξής ειδική περίπτωση:

$$n! = \Gamma(n+1) = n\Gamma(n) \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n},$$

καθώς $n \rightarrow \infty$.

Α'4 Παραγωγήιση ολοκληρώματος που εξαρτάται από παράμετρο

Αν X είναι τυχαία μεταβλητή με πραγματικές τιμές, τότε η ροπογεννήτρια της, $M_X(t) := \mathbf{E}(e^{tX})$, και η χαρακτηριστική συνάρτησή της, $\phi_X(t) := \mathbf{E}(e^{itX})$, είναι ολοκληρώματα που εξαρτώνται από μια παράμετρο, το t . Φαίνεται φυσιολογικό οι παράγωγοί τους σε οποιοδήποτε $t \in \mathbb{R}$ να ισούνται αντίστοιχα με $\mathbf{E}(Xe^{tX})$, $\mathbf{E}(iXe^{itX})$. Η επόμενη πρόταση, που διατυπώνεται σε γενικότερο πλαίσιο, δίνει συνθήκες ώστε αυτό να ισχύει.

Πρόταση Α'5. Έστω $(\Omega, \mathcal{A}, \mu)$ χώρος μέτρου, $a < b$ πραγματικοί αριθμοί, και $f : \Omega \times (a, b) \rightarrow \mathbb{R}$ συνάρτηση ώστε

- (i) Για κάθε $t \in (a, b)$, η $x \mapsto f(x, t)$ είναι μετρήσιμη και $\int |f(x, t)| d\mu(x) < \infty$.
- (ii) Η $\partial_t f(x, t)$ υπάρχει για κάθε $(x, t) \in \Omega \times (a, b)$.
- (iii) Υπάρχει μετρήσιμη συνάρτηση $A : \Omega \rightarrow \mathbb{R}$ ώστε $\sup_{s \in (a, b)} |\partial_t f(x, s)| \leq A(x)$ για κάθε $x \in \Omega$ και $\int A(x) d\mu(x) < \infty$.

Τότε η συνάρτηση $I : (a, b) \rightarrow \mathbb{R}$ με

$$I(t) := \int f(x, t) d\mu(x)$$

για κάθε $t \in (a, b)$ είναι καλά ορισμένη και διαφορίσιμη με

$$I'(t) = \int \partial_t f(x, t) d\mu(x)$$

για κάθε $t \in (a, b)$.

Η πρόταση ισχύει και στην περίπτωση που η f παίρνει τιμές στο \mathbb{C} . Η ισχύς της τότε έπεται από την περίπτωση που έχουμε διατυπώσει.

Απόδειξη. Έστω $t_0 \in (a, b)$. Θέτουμε $h_0 := \min\{t_0 - a, b - t_0\} > 0$. Για $h \in (-h_0, h_0) \setminus \{0\}$ και $x \in \Omega$ θέτουμε $g_h(x) := \{f(x, t_0 + h) - f(x, t_0)\}/h$. Τότε

$$\frac{1}{h}\{I(t_0 + h) - I(t_0)\} = \int g_h(x) d\mu(x).$$

Παρατηρούμε ότι

- $\lim_{h \rightarrow 0} g_h(x) = \partial_t f(x, t_0)$ για κάθε $x \in \Omega$.
- $|g_h(x)| \leq A(x)$ για κάθε $x \in \Omega, h \in (-h_0, h_0) \setminus \{0\}$ και $\int A(x) d\mu(x) < \infty$. [Αυτό γιατί από το θεώρημα μέση τιμής $g_h(x) = \partial_t f(x, s)$ για κάποιο s ανάμεσα στα $t_0, t_0 + h$.]

Οπότε το συμπέρασμα προκύπτει από το θεώρημα κυριαρχημένης σύγκλισης. ■

Παράδειγμα Α'6. (α) Παραγωγίζοντας κατάλληλο ολοκλήρωμα θα δείξουμε ότι $C := \int_0^\infty e^{-x^2} dx = \sqrt{\pi}/2$. Για κάθε $t \in [0, \infty)$ θέτουμε

$$I(t) = \int_0^\infty \frac{1}{1+x^2} e^{-t^2(1+x^2)} dx.$$

Αναλυτικά αποτελέσματα

Με χρήση του θεωρήματος κυριαρχημένης σύγκλισης δείχνουμε ότι η I είναι συνεχής στο $[0, \infty)$. Έπειτα, για $t \in (0, \infty)$ ισχύει

$$I'(t) = \int_0^\infty \frac{1}{1+x^2} e^{-t^2(1+x^2)} \{-2t(1+x^2)\} dx = -2te^{-t^2} \int_0^\infty e^{-t^2x^2} dx$$

$$\stackrel{y:=tx}{=} -2e^{-t^2} \int_0^\infty e^{-y^2} dy = -2e^{-t^2} C.$$

[Για τη δικαιολόγηση της παραγώγισης του ολοκληρώματος, επιλέγουμε a, b με $0 < a < t < b$, και θέτουμε $f(x, s) = e^{-s^2(1+x^2)}/(1+x^2)$ για κάθε $s \in (a, b)$, $x \in \mathbb{R}$. Ισχύει

$$|\partial f(x, s)| = 2t(1+x^2)e^{-t^2(1+x^2)} \leq 2b(1+x^2)e^{-a^2(1+x^2)} =: A(x)$$

για κάθε $s \in (a, b)$ και η A έχει $\int_0^\infty A(x) dx < \infty$. Έτσι η πιο πάνω πρόταση εφαρμόζεται.]

Επειδή το όριο $\lim_{t \rightarrow 0} I'(t)$ υπάρχει και είναι πραγματικός αριθμός και η I είναι συνεχής σε περιοχή του 0, έχουμε ότι η $I'(0)$ υπάρχει με τιμή $\lim_{t \rightarrow 0^+} I'(t)$. Σαφώς η I' είναι συνεχής στο $[0, \infty)$. Έτσι, ολοκληρώνοντας την $I'(t) = -2e^{-t^2} C$ στο $[0, \infty)$ παίρνουμε $I(\infty) - I(0) = -2C^2$. Αλλά $I(\infty) = 0$ (θεώρημα κυριαρχημένης σύγκλισης) και $I(0) = \tan^{-1}(\infty) - \tan^{-1}(0) = \pi/2$. Το συμπέρασμα έπεται.

(β) Αν $\mathbf{E}|X| < \infty$, θα δείξουμε ότι $\phi'_X(t) = i \mathbf{E}(Xe^{itX})$ για κάθε $t \in \mathbb{R}$. Όπως σημειώσαμε πιο πάνω, η Πρόταση Α'5 ισχύει και για f με τιμές στο \mathbb{C} . Θέτουμε $f(\omega, t) = e^{itX(\omega)}$ για κάθε $(\omega, t) \in \Omega \times \mathbb{R}$. Οι υποθέσεις (i), (ii) της πρότασης ισχύουν ενώ για την (iii) παρατηρούμε ότι $\sup_{t \in \mathbb{R}} |\partial_t f(\omega, t)| = |X(\omega)|$ και $\mathbf{E}|X| < \infty$. Άρα η παραγώγιση περνάει μέσα από τη μέση τιμή.

Β'

Τεχνικές αποδείξεις

Το παράρτημα αυτό περιέχει αποδείξεις κάποιων αποτελεσμάτων οι οποίες είναι πέρα από τον στόχο των σημειώσεων. Καταγράφονται εδώ για τον ενδιαφερόμενο αναγνώστη.

Κεφάλαιο 10

Απόδειξη του Θεωρήματος 10.10: Επειδή η οικογένεια $(\mathcal{G}_j)_{j \in J}$ είναι ανεξάρτητη αν και μόνο αν κάθε πεπερασμένη υποοικογένεια της είναι ανεξάρτητη, αρκεί να δείξουμε το θεώρημα στην περίπτωση που το J είναι πεπερασμένο. Υποθέτουμε λοιπόν ότι $J = \{1, 2, \dots, n\}$ για κάποιο $n \geq 2$. Για κάθε $k \in J$, ονομάζουμε C_k το σύνολο των συνόλων της μορφής $A_{i_1} \cap A_{i_2} \cdots \cap A_{i_r}$, όπου $r \geq 1$ και $A_{i_1}, A_{i_2}, \dots, A_{i_r} \in \cup_{i \in I_k} \mathcal{F}_i$. Παρατηρούμε ότι $\sigma(C_k) = \mathcal{G}_k$ γιατί σαφώς $C_k \subset \mathcal{G}_k$ και $C_k \supset \cup_{i \in I_k} \mathcal{F}_i$. Θέτουμε

$$\mathcal{D}_1 := \{A \in \mathcal{G}_1 : \mathbf{P}(A \cap A_2 \cap \cdots \cap A_n) = \mathbf{P}(A)\mathbf{P}(A_2) \cdots \mathbf{P}(A_n) \text{ για κάθε } A_2 \in C_2, \dots, A_n \in C_n\}.$$

Από υπόθεση, $C_1 \subset \mathcal{D}_1$. Εύκολα δείχνουμε ότι η \mathcal{D}_1 είναι κλάση Dynkin, άρα $\delta(C_1) \subset \mathcal{D}_1$. Όμως η C_1 είναι κλειστή στις πεπερασμένες τομές, οπότε το θεώρημα π-λ δίνει ότι $\delta(C_1) = \sigma(C_1) = \mathcal{G}_1$. Άρα

τα $\mathcal{G}_1, C_2, C_3, \dots, C_n$ είναι ανεξάρτητα.

Με ανάλογο επιχείρημα δείχνουμε ότι

τα $\mathcal{G}_1, \mathcal{G}_2, C_3, \dots, C_n$ είναι ανεξάρτητα,

και τελικά το θεώρημα. (Η τυπική απόδειξη γίνεται με επαγωγή.) ■

Απόδειξη του Θεωρήματος 10.11: Έστω $\mathcal{G}_j := \sigma(\cup_{i \in I_j} \sigma(X_i))$. Από την υπόθεση ανεξαρτησίας των $(X_i)_{i \in I}$ και το προηγούμενο θεώρημα, οι σ -άλγεβρες $(\mathcal{G}_j)_{j \in J}$ είναι ανεξάρτητες. Αρκεί επομένως να δείξουμε ότι για κάθε $j \in J$, η Y_j είναι \mathcal{G}_j -μετρήσιμη. Έστω $W_j := (X_i)_{i \in I_j} : \Omega \rightarrow \mathbb{R}^{I_j}$, οπότε $Y_j = f_j \circ W_j$, και για $A \subset \mathbb{R}$ σύνολο Borel, έχουμε $Y_j^{-1}(A) = W_j^{-1}(f_j^{-1}(A))$. Δεδομένου ότι η f_j είναι μετρήσιμη, μένει να δείξουμε τον εξής ισχυρισμό.

ΙΣΧΥΡΙΣΜΟΣ: Η W_j είναι \mathcal{G}_j μετρήσιμη.

Πράγματι, η οικογένεια $\mathcal{B}_j := \{B \subset \mathbb{R}^{I_j} : W_j^{-1}(B) \in \mathcal{G}_j\}$ είναι σ -άλγεβρα [Άσκηση 1.7(α)] και περιέχει τα μετρήσιμα ορθογώνια γιατί αν πάρουμε ένα τέτοιο $B = \prod_{i \in I_j} B_i$, θα έχουμε

$$W_j^{-1}(B) = \cap_{i \in I_j} X_i^{-1}(B_i).$$

Σε αυτή την τομή, μόνο πεπερασμένα σύνολα είναι διαφορετικά από το Ω αφού το $\{i \in I_j : B_i \neq \mathbb{R}\}$ είναι πεπερασμένο. Άρα, ως αριθμήσιμη (πεπερασμένη μάλιστα) τομή στοιχείων της σ -άλγεβρας \mathcal{G}_j είναι στοιχείο της \mathcal{G}_j . Και επειδή η σ -άλγεβρα γινόμενο παράγεται από τα μετρήσιμα ορθογώνια, έπεται ότι $\otimes_{i \in I_j} \mathcal{B}(\mathbb{R}) \subset \mathcal{B}_j$. Ο ισχυρισμός αποδείχθηκε. ■

Κεφάλαιο 11

Απόδειξη του Θεωρήματος 11.12: Θα δείξουμε ότι το C είναι ανεξάρτητο από τον εαυτό του. Γιατί αυτό δίνει $\mathbf{P}(C \cap C) = \mathbf{P}(C)\mathbf{P}(C)$, δηλαδή $\mathbf{P}(C) = \mathbf{P}^2(C)$, που γράφεται $\mathbf{P}(C)\{1 - \mathbf{P}(C)\} = 0$, από το οποίο προκύπτει το ζητούμενο.

ΙΣΧΥΡΙΣΜΟΣ 1: Για κάθε $n \geq 1$, οι σ-άλγεβρες $\mathcal{D}_n = \sigma(\{X_k : k \leq n\})$, \mathcal{C}_n είναι ανεξάρτητες.

Αυτό έπεται από το ότι οι $(X_n)_{n \geq 1}$ είναι ανεξάρτητες, τους Ορισμούς 10.3, 4.18, και το Θεώρημα 10.10 για τη διαμέριση $\{\{1, 2, \dots, n\}, \{n+1, n+2, \dots\}\}$ του \mathbb{N}^+ .

Θέτουμε τώρα $\mathcal{D} = \cup_{n \in \mathbb{N}} \mathcal{D}_n$.

ΙΣΧΥΡΙΣΜΟΣ 2: Το C είναι ανεξάρτητο από κάθε στοιχείο της $\sigma(\mathcal{D})$.

Επειδή το C είναι στοιχείο της \mathcal{C}_n για κάθε $n \geq 1$, έπεται ότι το C είναι ανεξάρτητο από κάθε \mathcal{D}_n και άρα από κάθε στοιχείο της ένωσής τους, που είναι το \mathcal{D} . Το σύνολο \mathcal{E} των στοιχείων της $\sigma(\mathcal{D})$ που είναι ανεξάρτητα από το C είναι μια κλάση Dynkin (Άσκηση 3.1) που περιέχει την \mathcal{D} και η \mathcal{D} είναι κλειστή στις πεπερασμένες τομές. Άρα, από το θεώρημα π-λ, $\sigma(\mathcal{D}) = \delta(\mathcal{D})$. Όμως $\delta(\mathcal{D}) \subset \mathcal{E} \subset \sigma(\mathcal{D})$, οπότε $\mathcal{E} = \sigma(\mathcal{D})$.

Τώρα $\mathcal{C}_\infty \subset \sigma(\mathcal{D})$ γιατί εύκολα βλέπουμε ότι $\sigma(\mathcal{D}) = \sigma(\{X_n : n \geq 1\})$. Άρα από τον Ισχυρισμό 2 έχουμε ότι το C είναι ανεξάρτητο από τον εαυτό του. ■

Κεφάλαιο 13

Απόδειξη του Θεωρήματος 13.5: Η ποσότητα στο όριο ισούται με

$$\frac{1}{2\pi} \int_{-T}^T \frac{e^{-ita} - e^{-itb}}{it} \int_{\mathbb{R}} e^{itx} d\mu(x) dt = \frac{1}{2\pi} \int_{\mathbb{R}} \int_{-T}^T \frac{e^{it(x-a)} - e^{it(x-b)}}{it} dt d\mu(x) \quad (\text{B'.1})$$

$$= \frac{1}{\pi} \int_{\mathbb{R}} \int_0^T \left(\frac{\sin\{t(x-a)\}}{t} - \frac{\sin\{t(x-b)\}}{t} \right) dt d\mu(x) = \frac{1}{\pi} \int_{\mathbb{R}} \{A(T, x-a) - A(T, x-b)\} d\mu(x) \quad (\text{B'.2})$$

όπου

$$A(T, y) := \int_0^T \frac{\sin(ty)}{t} dt = \int_0^{Ty} \frac{\sin(z)}{z} dz$$

για κάθε $T, y \in \mathbb{R}$.

Στην ισότητα της (B'.1), η εφαρμογή του θεωρήματος Fubini είναι επιτρεπτή γιατί ο ολοκληρωτέος είναι φραγμένη συνάρτηση, αφού

$$\left| \frac{e^{-ita} - e^{-itb}}{it} e^{itx} \right| = \left| \int_a^b e^{-itz} dz \right| \leq (b-a),$$

και τα μέτρα λ, μ είναι πεπερασμένα στους χώρους $[-T, T], \mathbb{R}$ αντίστοιχα. Στην πρώτη ισότητα της (B'.2) χρησιμοποιούμε το ότι η $\sin(ct)/t$ είναι άρτια συνάρτηση ενώ η $\cos(ct)/t$ είναι περιττή ($c \in \mathbb{R}$ είναι οποιαδήποτε σταθερά).

Στόχος μας τώρα είναι να πάρουμε $T \rightarrow \infty$ στο τελευταίο ολοκλήρωμα της (B'.2). Επειδή $\lim_{M \rightarrow \infty} \int_0^M z^{-1} \sin(z) dz = \pi/2$, έχουμε ότι

$$\lim_{T \rightarrow \infty} A(T, y) = \begin{cases} 0 & \text{αν } y = 0, \\ -\frac{\pi}{2} & \text{αν } y < 0, \\ \frac{\pi}{2} & \text{αν } y > 0, \end{cases}$$

και υπάρχει $C \in (0, \infty)$ ώστε $|A(T, y)| \leq C$ για κάθε $T, y \in \mathbb{R}$. Άρα ο ολοκληρωτέος στο τελευταίο ολοκλήρωμα της (B'.2) είναι απόλυτα φραγμένος από το $2C$ και έχει όριο

$$\lim_{T \rightarrow \infty} \{A(T, x-a) - A(T, x-b)\} = \begin{cases} 0 & \text{αν } x < a, \\ \frac{\pi}{2} & \text{αν } x = a, \\ \pi & \text{αν } a < x < b, \\ \frac{\pi}{2} & \text{αν } x = b, \\ 0 & \text{αν } x > b. \end{cases}$$

Το θεώρημα φραγμένης σύγκλισης δίνει το συμπέρασμα. ■

Κεφάλαιο 14

Απόδειξη του Θεωρήματος 14.8: Έστω F και F_n η συνάρτηση κατανομής των μέτρων μ, μ_n αντίστοιχα.

« \Rightarrow » Αν τα a, b είναι σημεία συνέχειας της F , τότε για την $f := \mathbf{1}_{(a,b)}$ ισχύει η (14.3). Πράγματι

$$\int \mathbf{1}_{(a,b)}(x) d\mu_n(x) = \mu_n((a, b]) = F_n(b) - F_n(a) \rightarrow F(b) - F(a) = \int \mathbf{1}_{(a,b)}(x) d\mu(x) \quad (\text{B'.3})$$

για $n \rightarrow \infty$.

Έστω τώρα f συνεχής και φραγμένη. Θέτουμε $\|f\|_\infty := \sup_{x \in \mathbb{R}} |f(x)|$. Παίρνουμε $\varepsilon > 0$. Βρίσκουμε $K > 0$ ώστε τα $-K, K$ να είναι σημεία συνέχειας της F και $F(-K) \leq \varepsilon, 1 - F(K) \leq \varepsilon$. Σταθεροποιούμε $\varepsilon_1 > 0$. Επειδή η f είναι ομοιόμορφα συνεχής στο $[-K, K]$, υπάρχει $\delta > 0$ ώστε

$$x, y \in [-K, K], |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon_1. \quad (\text{B'.4})$$

Βρίσκουμε στο $[-K, K]$ σημεία $-K = a_0 < a_1 < a_2 < \dots < a_{N-1} < a_N = K$ ώστε $0 < a_i - a_{i-1} < \delta$ για κάθε $i = 1, 2, \dots, N$ και η F να είναι συνεχής σε καθένα από τα a_1, a_2, \dots, a_{N-1} . Έστω $I_i := (a_{i-1}, a_i]$ για $i = 1, 2, \dots, N$ και

$$s(x) := \sum_{i=1}^N f(a_{i-1}) \mathbf{1}_{(a_{i-1}, a_i]}(x).$$

Τότε

- $\lim_{n \rightarrow \infty} \int s(x) d\mu_n(x) = \int s(x) d\mu(x)$ λόγω της (B'.3) και του ότι τα a_0, a_1, \dots, a_N είναι σημεία συνέχειας της F .
- $|f(x) - s(x)| < \varepsilon_1$ για κάθε $x \in (-K, K]$. Άρα

$$\left| \int f(x) d\mu_n(x) - \int s(x) d\mu_n(x) \right| \leq \int |f(x) - s(x)| d\mu_n(x) \quad (\text{B'.5})$$

$$\leq \int |f(x) - s(x)| \mathbf{1}_{x \leq -K} d\mu_n(x) + \int |f(x) - s(x)| \mathbf{1}_{x > K} d\mu_n(x) + \int |f(x) - s(x)| \mathbf{1}_{x \in (-K, K]} d\mu_n(x) \quad (\text{B'.6})$$

$$\leq \|f\|_\infty \{\mu_n((-\infty, -K]) + \mu_n((K, \infty))\} + \varepsilon_1 \mu_n((-K, K]) \quad (\text{B'.7})$$

$$\leq \|f\|_\infty \{F_n(-K) + 1 - F_n(K)\} + \varepsilon_1 \quad (\text{B'.8})$$

$$\text{Άρα } \overline{\lim}_{n \rightarrow \infty} \left| \int f(x) d\mu_n(x) - \int s(x) d\mu_n(x) \right| \leq 2\varepsilon \|f\|_\infty + \varepsilon_1.$$

• Όμοια, $|\int f(x) d\mu(x) - \int s(x) d\mu(x)| \leq 2\varepsilon \|f\|_\infty + 2\varepsilon_1$.

Άρα, από την τριγωνική ανισότητα,

$$\overline{\lim}_{n \rightarrow \infty} \left| \int f(x) d\mu_n(x) - \int f(x) d\mu(x) \right| \leq 4\varepsilon \|f\|_\infty + \varepsilon_1.$$

Το αριστερό μέλος δεν εξαρτάται από τα $\varepsilon, \varepsilon_1$. Θεωρούμε λοιπόν $\varepsilon, \varepsilon_1 \rightarrow 0^+$ και το ζητούμενο έπεται.

« \Leftarrow » Έστω $x_0 \in \mathbb{R}$ σημείο συνέχειας της F . Για $\varepsilon > 0$, θεωρούμε τη συνεχή και φραγμένη συνάρτηση

$$f(x) = \begin{cases} 1 & \text{αν } x \leq x_0, \\ -(x - x_0 - \varepsilon)/\varepsilon & \text{αν } x \in (x_0, x_0 + \varepsilon], \\ 0 & \text{αν } x \geq x_0 + \varepsilon. \end{cases} \quad (\text{B'.9})$$

η οποία ικανοποιεί $\mathbf{1}_{(-\infty, x_0]}(x) \leq f(x) \leq \mathbf{1}_{(-\infty, x_0 + \varepsilon]}(x)$. Παίρνοντας ολοκλήρωμα ως προς μ_n στην πρώτη ανισότητα, ως προς μ στη δεύτερη ανισότητα, και χρησιμοποιώντας την υπόθεση παίρνουμε

$$\overline{\lim}_{n \rightarrow \infty} F_n(x_0) \leq \overline{\lim}_{n \rightarrow \infty} \int f(x) d\mu_n(x) = \int f(x) d\mu(x) \leq F(x_0 + \varepsilon).$$

Όμως το ε είναι αυθαίρετο. Και επειδή η F είναι δεξιά συνεχής στο x_0 , για $\varepsilon \rightarrow 0^+$ παίρνουμε

$$\overline{\lim}_{n \rightarrow \infty} F_n(x_0) \leq F(x_0). \quad (\text{B'.10})$$

Για το κάτω φράγμα, παίρνουμε $\varepsilon > 0$ και θεωρούμε τη συνάρτηση

$$g(x) = \begin{cases} 1 & \text{αν } x \leq x_0 - \varepsilon, \\ -(x - x_0)/\varepsilon & \text{αν } x \in (x_0 - \varepsilon, x_0], \\ 0 & \text{αν } x \geq x_0, \end{cases} \quad (\text{B'.11})$$

η οποία ικανοποιεί $\mathbf{1}_{(-\infty, x_0 - \varepsilon]}(x) \leq g(x) \leq \mathbf{1}_{(-\infty, x_0]}(x)$. Όπως πριν παίρνουμε

$$\underline{\lim}_{n \rightarrow \infty} F_n(x_0) \geq \underline{\lim}_{n \rightarrow \infty} \int g(x) d\mu_n(x) = \int g(x) d\mu(x) \geq F(x_0 - \varepsilon).$$

Επειδή η F είναι αριστερά συνεχής στο x_0 (εδώ μόνο χρησιμοποιούμε ότι το x_0 είναι σημείο συνέχειας της F), για $\varepsilon \rightarrow 0^+$ παίρνουμε

$$\underline{\lim}_{n \rightarrow \infty} F_n(x_0) \geq F(x_0). \quad (\text{B'.12})$$

Η τελευταία σχέση μαζί με την (B'.10) δίνουν το ζητούμενο. ■

Απόδειξη του Θεωρήματος 14.12 (i) \Rightarrow (ii): ΙΣΧΥΡΙΣΜΟΣ: Αν το A είναι κλειστό, τότε

$$\overline{\lim}_{n \rightarrow \infty} \mu_n(A) \leq \mu(A).$$

Σταθεροποιούμε $r > 0$ και θεωρούμε τη συνάρτηση $f_r(x) := 1/(1 + d(x, A))^r$, όπου $d(x, A) := \inf\{|x - y| : y \in A\}$ είναι η απόσταση του x από το A . Η f_r είναι συνεχής, φραγμένη, και ικανοποιεί $\mathbf{1}_A \leq f_r$. Άρα

$$\overline{\lim}_{n \rightarrow \infty} \mu_n(A) = \overline{\lim}_{n \rightarrow \infty} \int \mathbf{1}_A(x) d\mu_n(x) \leq \overline{\lim}_{n \rightarrow \infty} \int f_r(x) d\mu_n(x) = \int f_r(x) d\mu(x). \quad (\text{B'.13})$$

Τεχνικές αποδείξεις

Τώρα, $\lim_{r \rightarrow \infty} f_r(x) = \mathbf{1}_A(x)$ για κάθε x επειδή κάθε $x \in \mathbb{R} \setminus A$ έχει $d(x, A) > 0$ (το A είναι κλειστό). Οπότε, παίρνοντας $r \rightarrow \infty$ στην (B'.13) κατά μήκος μιας ακολουθίας (π.χ. $r = k$ φυσικός) το θεώρημα φραγμένης σύγκλισης δίνει ότι $\lim_{k \rightarrow \infty} \int f_k(x) d\mu(x) = \int \mathbf{1}_A(x) d\mu(x) = \mu(A)$. Και ο ισχυρισμός αποδείχθηκε.

Αν το A είναι ανοιχτό, τότε εφαρμόζοντας τον ισχυρισμό για το κλειστό $\mathbb{R} \setminus A$ παίρνουμε

$$\varliminf_{n \rightarrow \infty} \mu_n(A) \geq \mu(A).$$

Τώρα για ένα A όπως στην εκφώνηση έχουμε $\mu(\bar{A}) = \mu(A^\circ) + \mu(\partial A) = \mu(A^\circ)$. Και από τα πιο πάνω

$$\mu(A^\circ) \leq \varliminf_{n \rightarrow \infty} \mu_n(A^\circ) \leq \varliminf_{n \rightarrow \infty} \mu_n(A) \leq \overline{\lim}_{n \rightarrow \infty} \mu_n(A) \leq \overline{\lim}_{n \rightarrow \infty} \mu_n(\bar{A}) \leq \mu(\bar{A}).$$

Το ζητούμενο έπεται. ■

Κεφάλαιο 15

Απόδειξη του Λήμματος 15.1: Από τον ορισμό του μετασχηματισμού Fourier του μέτρου μ έχουμε

$$\int_{-u}^u (1 - \phi_\mu(t)) dt = \int_{-u}^u \int (1 - e^{itx}) d\mu(x) dt = \int \int_{-u}^u (1 - \cos(tx) + i \sin(tx)) dt d\mu(x).$$

Η δεύτερη ισότητα προκύπτει από το Θεώρημα 9.4 (Fubini). Εφόσον η συνάρτηση $1 - \cos(tx)$ είναι άρτια και η συνάρτηση $\sin(tx)$ είναι περιττή, το τελευταίο ολοκλήρωμα ισούται με

$$2 \int \int_0^u (1 - \cos(tx)) dt d\mu(x) = 2 \int \left(u - \frac{\sin(ux)}{x} \right) d\mu(x) = 2u \int \left(1 - \frac{\sin(ux)}{ux} \right) d\mu(x).$$

Παρατηρούμε τώρα ότι η συνάρτηση στο τελευταίο ολοκλήρωμα είναι μη αρνητική ($1 - \frac{\sin(ux)}{ux} \geq 0$ για κάθε $x \in \mathbb{R}$. Άρα, αν ολοκληρώσουμε σε μικρότερο χωρίο, το ολοκλήρωμα μικραίνει.) και για $|ux| > 2$ έχουμε

$$\left| \frac{\sin(ux)}{ux} \right| \leq \frac{1}{ux} \leq \frac{1}{2}.$$

Συνεπώς,

$$\int_{-u}^u \{1 - \phi_\mu(t)\} dt \geq 2u \int_{\{|x| > 2/u\}} \frac{1}{2} d\mu(x) = u\mu \left(\left\{ x : |x| > \frac{2}{u} \right\} \right), \quad (\text{B'.14})$$

που είναι το ζητούμενο. ■

Κεφάλαιο 17

Απόδειξη του Θεωρήματος 17.8. Έστω μ_n η κατανομή της τυχαίας μεταβλητής S_n/n . Ακολουθούμε τη μέθοδο της Παρατήρησης 17.3(γ).

ΑΝΩ ΦΡΑΓΜΑ: Έστω $F \subset \mathbb{R}$ κλειστό μη κενό. Αν $I(F) = 0$, δεν έχουμε να αποδείξουμε τίποτα γιατί το αριστερό μέλος της (17.5) είναι μη θετικό πάντοτε.

Υποθέτουμε λοιπόν ότι $I(F) > 0$. Επειδή $I(m) = 0$ (Λήμμα 17.6), έπεται ότι το m είναι στοιχείο του ανοιχτού συνόλου $\mathbb{R} \setminus F$. Έστω (a, b) το μέγιστο υποδιάστημα του $\mathbb{R} \setminus F$ που περιέχει το m . Αυτό το υποδιάστημα είναι ανοιχτό (και άρα $a, b \in F$) γιατί το $\mathbb{R} \setminus F$ είναι ανοιχτό και ενδέχεται $a = -\infty$ ή $b = \infty$ (όχι όμως και τα δύο γιατί $F \neq \emptyset$). Επειδή $F \subset \mathbb{R} \setminus (a, b)$, όταν $a, b \in \mathbb{R}$, έχουμε

$$\mu_n(F) \leq \mu_n((-\infty, a]) + \mu_n([b, \infty)) \leq e^{-n\Lambda^*(a)} + e^{-n\Lambda^*(b)} \leq 2e^{-nI(F)}. \quad (\text{B'.15})$$

Η πρώτη ανισότητα έπεται από τα Λήμματα 17.5, 17.6, ενώ η δεύτερη από το ότι $a, b \in F$. Αν $a = -\infty$, οι ανισότητες ισχύουν αν παραλείψουμε τους όρους $\mu_n((-\infty, a])$, $e^{-n\Lambda^*(a)}$. Ανάλογα και όταν $b = \infty$. Τώρα το άνω φράγμα έπεται από την (B'.15).

ΚΑΤΩ ΦΡΑΓΜΑ: Με βάση την (17.7), επειδή η S_n/n παίρνει τιμές στο \mathbb{R} , αρκεί να δείξουμε ότι για κάθε $a \in \mathbb{R}$ και $\delta > 0$ ισχύει

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n((a - \delta, a + \delta)) \geq -\Lambda^*(a). \quad (\text{B'.16})$$

Περίπτωση 1. $\mu((-\infty, a)), \mu((a, \infty)) > 0$ και το μ έχει συμπαγή φορέα.

Τότε με βάση το Λήμμα 17.7 υπάρχει $\lambda_0 \in \mathbb{R}$ ώστε $\Lambda^*(a) = \lambda_0 a - \Lambda(\lambda_0)$. Ορίζουμε ένα νέο μέτρο $\tilde{\mu}$ από τη σχέση (δες Παράδειγμα 5.33)

$$\frac{d\tilde{\mu}}{d\mu}(x) = e^{\lambda_0 x - \Lambda(\lambda_0)}, x \in \mathbb{R}. \quad (\text{B'.17})$$

Το $\tilde{\mu}$ είναι μέτρο πιθανότητας γιατί

$$\tilde{\mu}(\mathbb{R}) = \int_{\mathbb{R}} e^{\lambda_0 x - \Lambda(\lambda_0)} d\mu(x) = \frac{1}{M(\lambda_0)} \int_{\mathbb{R}} e^{\lambda_0 x} d\mu(x) = 1$$

και έχει μέση τιμή a γιατί

$$\int_{\mathbb{R}} x d\tilde{\mu}(x) = \frac{\int_{\mathbb{R}} x e^{\lambda_0 x} d\mu(x)}{M(\lambda_0)} = \frac{M'(\lambda_0)}{M(\lambda_0)} = \Lambda'(\lambda_0) = a$$

Επίσης, συμβολίζουμε με $\tilde{\mu}_n$ την κατανομή του μέσου όρου $\tilde{X}_1 + \dots + \tilde{X}_n/n$ όταν οι $\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_n$ είναι ανεξάρτητες ισόνομες καθεμία με κατανομή $\tilde{\mu}$. Και τώρα είμαστε σε θέση να δείξουμε το ζητούμενο κάτω φράγμα. Για οποιοδήποτε $\varepsilon \in (-\delta, \delta)$ υπολογίζουμε

$$\begin{aligned} \mu_n((a - \varepsilon, a + \varepsilon)) &= \mathbf{P}\left(\frac{S_n}{n} \in (a - \varepsilon, a + \varepsilon)\right) = \int_{|x_1 + x_2 + \dots + x_n - na| < n\varepsilon} d\mu(x_1) \dots d\mu(x_n) \\ &= \int_{|x_1 + x_2 + \dots + x_n - na| < n\varepsilon} e^{n\Lambda(\lambda_0) - \lambda_0(x_1 + \dots + x_n)} d\tilde{\mu}(x_1) \dots d\tilde{\mu}(x_n) \\ &\geq e^{n\Lambda(\lambda_0) - \lambda_0 na - |\lambda_0| n\varepsilon} \tilde{\mu}_n((a - \varepsilon, a + \varepsilon)) = e^{-n\Lambda^*(a) - n|\lambda_0|\varepsilon} \tilde{\mu}_n((a - \varepsilon, a + \varepsilon)). \end{aligned}$$

Άρα

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n((a - \delta, a + \delta)) \geq -\Lambda^*(a) - |\lambda_0|\varepsilon - \liminf_{n \rightarrow \infty} \frac{1}{n} \log \tilde{\mu}_n((a - \delta, a + \delta)). \quad (\text{B'.18})$$

Τώρα $\lim_{n \rightarrow \infty} \tilde{\mu}_n((a - \delta, a + \delta)) = 1$ από τον ασθενή νόμο των μεγάλων αριθμών γιατί

$$\tilde{\mu}_n((a - \delta, a + \delta)) = \mathbf{P}\left(\frac{\tilde{X}_1 + \dots + \tilde{X}_n}{n} \in (a - \delta, a + \delta)\right)$$

και οι $\tilde{X}_1, \dots, \tilde{X}_n$ είναι ανεξάρτητες ισόνομες με μέση τιμή a . Άρα το \lim στο δεξί μέλος της ανισότητας (B'.18) είναι 0 και παίρνοντας $\varepsilon \rightarrow 0$ έχουμε την (17.7).

Περίπτωση 2. $\mu((-\infty, a)), \mu((a, \infty)) > 0$ και το μ δεν έχει συμπαγή φορέα.

Υπάρχει $R_0 > 0$ μεγάλο ώστε $\mu((-R_0, a)), \mu((a, R_0)) > 0$. Θεωρούμε τώρα οποιοδήποτε $R > R_0$ και ακολουθία $(\tilde{X}_i)_{i \geq 1}$ ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με κατανομή αυτήν της X_1 με τη δέσμευση $|X_1| \leq R$. Δηλαδή

$$\mathbf{P}(\tilde{X}_1 \in A) = \frac{\mathbf{P}(X_1 \in A, |X_1| \leq R)}{\mathbf{P}(|X_1| \leq R)}$$

για κάθε $A \in \mathcal{B}(\mathbb{R})$.

Θέτουμε $\hat{S}_n = \hat{X}_1 + \hat{X}_2 + \dots + \hat{X}_n$. Τότε

$$\mathbf{P}\left(\frac{S_n}{n} \in (a - \varepsilon, a + \varepsilon)\right) \geq \mathbf{P}\left(\frac{S_n}{n} \in (a - \varepsilon, a + \varepsilon), |X_i| \leq R \text{ για κάθε } i = 1, 2, \dots, n\right) \quad (\text{B'.19})$$

$$= \frac{\mathbf{P}\left(\frac{S_n}{n} \in (a - \varepsilon, a + \varepsilon), |X_i| \leq R \text{ για κάθε } i = 1, 2, \dots, n\right)}{\mathbf{P}(|X_i| \leq R \text{ για κάθε } i = 1, 2, \dots, n)} \quad (\text{B'.20})$$

$$\times \mathbf{P}(|X_i| \leq R \text{ για κάθε } i = 1, 2, \dots, n) \quad (\text{B'.21})$$

$$= \mathbf{P}\left(\frac{\hat{S}_n}{n} \in (a - \varepsilon, a + \varepsilon)\right) \mathbf{P}(|X_1| \leq R)^n \quad (\text{B'.22})$$

Τώρα για την ακολουθία \hat{S}_n/n εφαρμόζεται η περίπτωση 1 του κάτω φράγματος. Συμβολίζουμε με I_R τη συνάρτηση ρυθμού της αρχής μεγάλων αποκλίσεων που ικανοποιεί η ακολουθία. Άρα

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n((a - \delta, a + \delta)) \geq -I_R(a) + \log \mathbf{P}(|X_1| \leq R).$$

Αρχικά, θα βελτιώσουμε την έκφραση του κάτω φράγματος. Θέτουμε $C_R(\lambda) = \log \mathbf{E}(e^{\lambda X_1} \mathbf{1}_{|X_1| \leq R})$. Η ροπογεννήτρια της \hat{X}_1 είναι $\mathbf{E}(e^{\lambda X_1} \mathbf{1}_{|X_1| \leq R}) / \mathbf{P}(|X_1| \leq R)$, με λογάριθμο $C_R(\lambda) - \log \mathbf{P}(|X_1| \leq R)$, άρα $I_R(a) = \sup_{\lambda \in \mathbb{R}} \{\lambda a - C_R(\lambda)\} + \log \mathbf{P}(|X_1| \leq R)$. Επομένως το προηγούμενο κάτω φράγμα είναι απλώς

$$- \sup_{\lambda \in \mathbb{R}} \{\lambda a - C_R(\lambda)\},$$

το οποίο είναι το αντίθετο του μετασχηματισμού Legendre $C_R^*(a)$ της C_R στο a . Έτσι, το ζητούμενο κάτω φράγμα έπεται από τον εξής ισχυρισμό.

ΙΣΧΥΡΙΣΜΟΣ: $\lim_{R \rightarrow \infty} C_R^*(a) \leq \Lambda^*(a)$.

Η $C_R^*(a)$ είναι φθίνουσα ως προς R γιατί η $C_R(\lambda)$ είναι αύξουσα ως προς R . Άρα $\lim_{R \rightarrow \infty} C_R^*(a) \leq C_r^*(a)$ για κάθε $r > 0$. Έστω $u < \lim_{R \rightarrow \infty} C_R^*(a)$. Θέτουμε

$$K_r := \{\lambda \in \mathbb{R} : \lambda a - C_r(\lambda) \geq u\}.$$

Για $r \geq R_0$, το K_r είναι μη κενό αφού $u < C_r^*(a)$ και συμπαγές γιατί η $A_r(\lambda) := \lambda a - C_r(\lambda)$ είναι πεπερασμένη παντού και συνεχής ως προς λ με $A_r(-\infty) = A_r(\infty) = -\infty$ (απόδειξη όπως στο Λήμμα 17.7(β)). Επίσης η $(K_r)_{r \geq R_0}$ είναι φθίνουσα ως προς r , άρα η τομή $\bigcap_{r \geq R_0} K_r$ είναι μη κενή και έστω λ_0 ένα σημείο σε αυτήν. Τότε

$$\lambda_0 a - C_r(\lambda_0) \geq u$$

για κάθε $r \geq R_0$. Για $r \rightarrow \infty$ η τελευταία ανισότητα και το θεώρημα μονότονης σύγκλισης δίνουν $\lambda_0 a - \log \Lambda(\lambda_0) \geq u$, και άρα $\Lambda^*(a) \geq u$. Ο ισχυρισμός αποδείχθηκε.

Περίπτωση 3. Κανένας περιορισμός στο μ (πέραν του $\mathbf{E}(X_1) \in \mathbb{R}$).

Μένει να εξετάσουμε την περίπτωση που ένα τουλάχιστον από τα $\mu((-\infty, a)), \mu((a, \infty))$ είναι 0. Τότε

$$\Lambda^*(a) = \sup_{\lambda \in \mathbb{R}} \{\lambda a - \log \mathbf{E}(e^{\lambda X_1})\} = \sup_{\lambda \in \mathbb{R}} \{-\log \mathbf{E}(e^{\lambda(X_1 - a)})\} = -\log \mathbf{P}(X_1 = a).$$

Η τελευταία ισότητα ισχύει γιατί κάτω από τις υποθέσεις μας, η $\mathbf{E}(e^{\lambda(X_1 - a)})$ είναι μονότονη ως προς λ και άρα το infimum της ισούται με το όριο της στο $-\infty$ όταν $\mu((-\infty, a)) = 0$ και με το όριο της στο ∞ όταν $\mu((a, -\infty)) = 0$. Τώρα το συμπέρασμα έπεται γιατί

$$\mu_n((a - \delta, a + \delta)) \geq \mathbf{P}(X_1 = X_2 = \dots = X_n = a) = \mathbf{P}(X_1 = a)^n. \quad \blacksquare$$

Παρατήρηση Β'.1. (α) (Η ιδέα της αλλαγής μέτρου) Το ουσιαστικό κομμάτι της απόδειξης του κάτω φράγματος είναι η Περίπτωση 1. Ας πάρουμε την περίπτωση $a \neq m$ και ε μικρό. Το γεγονός $A_n = \{S_n/n \in (a - \varepsilon, a + \varepsilon)\}$ είναι μη τυπικό όταν οι X_i έχουν κατανομή μ , και δυσκολευόμαστε να εκτιμήσουμε την πιθανότητά του. Αυτό που κάνουμε είναι να αλλάξουμε τον νόμο των X_i με τέτοιο τρόπο ώστε το A_n να γίνει τυπικό για αυτόν τον νέο νόμο. Και πράγματι, ο νόμος $\tilde{\mu}$ έχει μέση τιμή a , οπότε, όταν οι X_i είναι ανεξάρτητες, καθεμία με κατανομή $\tilde{\mu}$, το A_n έχει πιθανότητα που τείνει στο 1. Το κόστος για την αλλαγή νόμου (μέτρου) είναι η παράγωγος Radon-Nikodym, για την οποία ευτυχώς έχουμε καλό έλεγχο. Στο σύνολο A_n αυτή έχει τιμή περίπου $e^{n(\Lambda(\lambda_0) - \lambda_0 a)}$.

(β) Προσέξτε ότι για την Περίπτωση 2 του κάτω φράγματος εφαρμόσαμε την τεχνική της περικοπής ώστε να αναχθούμε στην Περίπτωση 1. Με τον ίδιο τρόπο αποδείξαμε την επέκταση του νόμου των μεγάλων αριθμών στην Άσκηση 12.2.

Υποδείξεις για επιλεγμένες ασκήσεις

Κεφάλαιο 1

1.5 Αν παραγόταν, τότε η διαμέριση θα ήταν αναγκαστικά η $C := \{\{x\} : x \in \mathbb{R}\}$. Έπειτα, χρησιμοποιούμε την Άσκηση 1.3.

1.6 $B_1 = A_1, B_2 = A_2 \setminus A_1, B_3 = A_3 \setminus (A_1 \cup A_2), \dots$

1.7 (α) $\emptyset \in \mathcal{B}$ γιατί $f^{-1}(\emptyset) = \emptyset$ και $\emptyset \in \mathcal{F}$. Έπειτα, αν $A \in \mathcal{B}$ τότε $f^{-1}(Y \setminus A) = X \setminus f^{-1}(A)$ και εφόσον $f^{-1}(A) \in \mathcal{A}$ και η \mathcal{A} είναι σ-άλγεβρα, έχουμε ότι $X \setminus f^{-1}(A) \in \mathcal{A}$. Άρα $Y \setminus A \in \mathcal{B}$. Τέλος, αν $(A_n)_{n \in \mathbb{N}}$ είναι ακολουθία στη \mathcal{B} , τότε

$$f^{-1}(\cup_{n \in \mathbb{N}} A_n) = \cup_{n \in \mathbb{N}} f^{-1}(A_n) \in \mathcal{A}.$$

Που σημαίνει ότι $\cup_{n \in \mathbb{N}} A_n \in \mathcal{B}$.

1.9 (α) \subset . (β) \supset . (γ) Δεν συγκρίνονται. (δ) \supset . (ε) \subset .

Κεφάλαιο 2

2.2 Έχουμε ότι $\sum_{n=1}^{\infty} \mathbf{P}(A_n) = \mathbf{P}(\cup_{n=1}^{\infty} A_n) \leq 1$. Αφού η σειρά συγκλίνει, $\lim_{n \rightarrow \infty} \mathbf{P}(A_n) = 0$.

2.3 (α) Ισχύει ότι $\mathbf{P}(\cup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mathbf{P}(A_n) = 0$.

(β) Ισχύει ότι $\mathbf{P}(\cap_{n=1}^{\infty} A_n) = 1 - \mathbf{P}(\cup_{n=1}^{\infty} A_n^c) = 1$ αφού $\mathbf{P}(\cup_{n=1}^{\infty} A_n^c) = 0$ λόγω του (α).

2.4 Λόγω της προηγούμενης άσκησης, πρέπει τα I, I' να είναι υπεραριθμήσιμα. Έστω $\Omega = (0, 1), \mathcal{F} = \mathcal{B}((0, 1)), \mathbf{P} = \lambda$ [το μέτρο Lebesgue περιορισμένο στο $(0, 1)$], $I = I' = (0, 1), A_x := \{x\}, B_x := (0, 1) \setminus \{x\}$ για κάθε $x \in (0, 1)$. Τότε

(α) $\mathbf{P}(A_x) = \mathbf{P}(\{x\}) = 0$ για κάθε $x \in (0, 1)$, όμως $\mathbf{P}(\cup_{x \in (0, 1)} A_x) = \mathbf{P}((0, 1)) = 1$.

(β) $\mathbf{P}(B_x) = 1$ για κάθε $x \in (0, 1)$, όμως $\cap_{x \in (0, 1)} B_x = \emptyset$.

2.5 Για $n \in \mathbb{N}^+$, θέτουμε $B_n = \{\beta \in B : \mathbf{P}(A_\beta) \geq \frac{1}{n}\}$. Τότε $|B_n| \leq n$, γιατί $\mathbf{P}(\cup_{\beta \in B_n} A_\beta) \leq 1$ και $\mathbf{P}(\cup_{\beta \in B_n} A_\beta) = \sum_{\beta \in B_n} \mathbf{P}(A_\beta) \geq \frac{1}{n} |B_n|$. Αφού $B = \cup_{n=1}^{\infty} B_n$, με $|B_n|$ πεπερασμένο για κάθε $n \geq 1$, έχουμε το ζητούμενο.

2.6 Για την πρώτη ανισότητα έχουμε

$$\mathbf{P}(\liminf_{n \geq 1} A_n) = \mathbf{P}(\cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k) = \lim_{n \rightarrow \infty} \mathbf{P}(\cap_{k=n}^{\infty} A_k) \leq \varliminf_{n \rightarrow \infty} \mathbf{P}(A_n).$$

Η δεύτερη ισότητα ισχύει γιατί η ακολουθία $B_n := \cap_{k=n}^{\infty} A_k$ είναι αύξουσα, ενώ η ανισότητα ισχύει γιατί $B_n \subset A_n$. Η ανισότητα

$$\varlimsup_{n \rightarrow \infty} \mathbf{P}(A_n) \leq \mathbf{P}(\limsup A_n)$$

αποδεικνύεται όμοια.

2.7 Το στήριγμα του μέτρου είναι το $[2, 5] \cup \mathbb{N}$.

2.8 Το F ως υπόχωρος διαχωρίσιμου μετρικού χώρου είναι διαχωρίσιμος (Αυτό αποδεικνύεται ως εξής: Έστω A ένα αριθμήσιμο πυκνό υποσύνολο του \mathbb{R} . Για κάθε $x \in A$ και θετικό ακέραιο n επιλέγουμε $y \in F$ με $|y - x| < 1/n$ αν τέτοιο y υπάρχει. Το σύνολο όλων αυτών των y που συλλέγουμε έτσι είναι αριθμήσιμο και πυκνό υποσύνολο του F). Έστω $D := \{x_k : k \in \mathbb{N}\}$ ένα αριθμήσιμο πυκνό υποσύνολο του F . Θέτουμε $\mu = \sum_{k=1}^{\infty} 2^{-k} \delta_{x_k}$. δ_x είναι το μέτρο Dirac στο σημείο x . Το μ έχει στήριγμα F .

Κεφάλαιο 3

3.1 Χρήσιμη είναι η Πρόταση 2.5

3.2 Η \mathcal{A} δεν περιέχει το $\{2\}$ που είναι τομή των $\{1, 2\}, \{2, 3\}$.

3.3 Από το Θεώρημα 3.6 έχουμε ότι $\sigma(C_1) = \delta(C_1)$.

Κεφάλαιο 4

4.1 Το σύνολο $\mathcal{A} := \{A \subset \mathbb{R} : f^{-1}(A) \in \mathcal{F}\}$ είναι σ-άλγεβρα [Άσκηση 1.7(i)]. Έστω \mathcal{T} η οικογένεια των ανοιχτών υποσυνόλων του \mathbb{R} . Προφανώς το (α) συνεπάγεται τα (β), (γ). Αν υποθέσουμε το (β), δηλαδή $\mathcal{T} \subset \mathcal{A}$, τότε $\sigma(\mathcal{T}) \subset \mathcal{A}$, που είναι το (α). Έπειτα θέτουμε $\mathcal{A}_4 := \{[a, b] : a < b, a, b \in \mathbb{R}\}$. Αν ισχύει το (γ), δηλαδή $\mathcal{A}_4 \subset \mathcal{A}$, τότε $\sigma(\mathcal{A}_4) \subset \mathcal{A}$. Μένει να δείξουμε ότι $\sigma(\mathcal{A}_4) = \mathcal{B}(\mathbb{R})$, το οποίο κάνουμε όπως στην απόδειξη του Θεωρήματος 1.14.

4.2 Τα σύνολα $\{-\infty\}, \{\infty\}$ είναι κλειστά.

4.3 Έστω f μετρήσιμη και $i_0 \in I$. Ας υποθέσουμε ότι η f παίρνει δύο διαφορετικές τιμές $a < b$ στο A_{i_0} . Θα έπρεπε λοιπόν το σύνολο $A_{i_0} \cap \{f < b\}$ να ανήκει στην \mathcal{F} . Όμως, αυτό το σύνολο είναι μη κενό και γνήσιο υποσύνολο του A_{i_0} . Τέτοιο σύνολο δεν υπάρχει στην \mathcal{F} (δες στο Παράδειγμα 1.10 για την περιγραφή της \mathcal{F}). Επίσης, είναι εύκολο να δείξει κανείς ότι μια συνάρτηση που είναι σταθερή σε κάθε σύνολο της διαμέρισης είναι μετρήσιμη. Άρα, αυτές είναι ακριβώς όλες οι μετρήσιμες συναρτήσεις στον (Ω, \mathcal{F}) .

4.4 (α) $\{\omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) = \infty\} = \bigcap_{k=1}^{\infty} \bigcup_{j=1}^{\infty} \bigcap_{r=j}^{\infty} \{X_r > k\}$.

(β) Το $\lim_{n \rightarrow \infty} X_n(\omega)$ υπάρχει στο \mathbb{R} αν και μόνο αν η ακολουθία $(X_n(\omega))_{n \geq 1}$ είναι βασική. Δηλαδή για κάθε $k \geq 1$ υπάρχει $j \geq 1$ ώστε $|X_r(\omega) - X_s(\omega)| < 1/k$ για κάθε $r, s \geq j$. Άρα, το δοσμένο σύνολο γράφεται ως ...

Για μια άλλη λύση, παρατηρούμε ότι το δοσμένο σύνολο γράφεται ως

$$\{\underline{\lim}_{n \rightarrow \infty} X_n(\omega) = \overline{\lim}_{n \rightarrow \infty} X_n(\omega)\} \cap \{\underline{\lim}_{n \rightarrow \infty} X_n(\omega) \in \mathbb{R}\}.$$

Οι $\underline{\lim}_{n \rightarrow \infty} X_n(\omega), \overline{\lim}_{n \rightarrow \infty} X_n(\omega)$ είναι μετρήσιμες.

4.5 Χρησιμοποιούμε το Πρόσχημα 4.4. Η αντίστροφη εικόνα κάθε διαστήματος είναι διάστημα (και άρα Borel-μετρήσιμο) αφού η f είναι μονότονη.

4.6 Θεωρούμε τη συνάρτηση $h = f - g$. Τότε η h είναι μετρήσιμη και ισχύει ότι $\{f = g\} = h^{-1}(\{0\}) \in \mathcal{A}$, εφόσον $\{0\} \in \mathcal{B}(\mathbb{R})$.

Εναλλακτικά,

$$\begin{aligned} \Omega \setminus \{f = g\} &= \{f \neq g\} = \{f > g\} \cup \{g > f\} \\ &= \left(\bigcup_{q \in \mathbb{Q}} (\{f > q\} \cap \{g < q\}) \right) \cup \left(\bigcup_{q \in \mathbb{Q}} (\{g > q\} \cap \{f < q\}) \right), \end{aligned}$$

από το οποίο έπεται το ζητούμενο.

4.11 Επειδή το $\{X > 1\}$ έχει θετική πιθανότητα και ισούται με την αριθμήσιμη ένωση

$$\{X > 1\} = \bigcup_{n=1}^{\infty} \left\{ X > 1 + \frac{1}{n} \right\},$$

κάποιο από τα σύνολα της ένωσης πρέπει να έχει θετική πιθανότητα.

Κεφάλαιο 5

5.1 Στην ισότητα $1 - \mathbf{1}_{\bigcup_{i=1}^n A_i} = (1 - \mathbf{1}_{A_1})(1 - \mathbf{1}_{A_2}) \cdots (1 - \mathbf{1}_{A_n})$, αναπτύσσουμε το δεξιό μέλος και παίρνουμε μέση τιμή.

5.2 Θεωρούμε την τυχαία μεταβλητή $X = \sum_{k=1}^n \mathbf{1}_{A_k}$.

5.4 Χρησιμοποιούμε την προηγούμενη άσκηση.

5.5 $\mathbf{P}(X > t) = \mathbf{P}(aX > at) = \mathbf{P}(e^{aX} > e^{at}) \leq \mathbf{E}(e^{aX})/e^{at}$. Παίρνουμε $C = \mathbf{E}(e^{aX}) \in (0, \infty)$.

5.6 (α) Δουλεύουμε όπως στην αποδείξη της ανισότητας Chebyshev.

$$\begin{aligned} \mathbf{P}(X \leq a \mathbf{E} X) &= \mathbf{P}(X - \mathbf{E} X \leq -(1-a) \mathbf{E} X) \\ &\leq \mathbf{P}(|X - \mathbf{E} X| \geq (1-a) \mathbf{E} X) \leq \frac{\text{Var}(X)}{(1-a)^2 (\mathbf{E} X)^2}. \end{aligned}$$

Χρησιμοποιήσαμε το ότι $(1-a) \mathbf{E} X > 0$.

(β) Έστω $A := \{\omega : X(\omega) > a \mathbf{E} X\}$.

$$\begin{aligned} \mathbf{E} X &= \mathbf{E}(X \mathbf{1}_{A^c}) + \mathbf{E}(X \mathbf{1}_A) \leq a \mathbf{E} X + \mathbf{E}(X^2)^{1/2} \mathbf{P}(A)^{1/2} \Rightarrow \\ (1-a) \mathbf{E} X &\leq \mathbf{E}(X^2)^{1/2} \mathbf{P}(A)^{1/2} \Rightarrow \mathbf{P}(A) \geq (1-a)^2 (\mathbf{E} X)^2 / \mathbf{E}(X^2) \end{aligned}$$

Στην πρώτη γραμμή χρησιμοποιήσαμε την ανισότητα Cauchy-Schwarz και το ότι $X \leq a \mathbf{E} X$ στο A^c .

5.7 Χρησιμοποιούμε την ανισότητα Cauchy-Schwarz.

$$1 \leq \mathbf{E}(\sqrt{XY}) \leq (\mathbf{E}(X))^{1/2} (\mathbf{E}(Y))^{1/2}.$$

5.10 Η ακολουθία $A_n := \{|X| > n\}$, $n \geq 1$ είναι φθίνουσα με τομή το \emptyset αφού η X παίρνει τιμές στο \mathbb{R} . Άρα $\lim_{n \rightarrow \infty} \mathbf{P}(A_n) = \mathbf{P}(\emptyset) = 0$.

5.13 Έστω $a_t := \int f_t(x) d\mu(x)$. Για οποιοδήποτε $\ell \in \bar{\mathbb{R}}$ ισχύει $\lim_{t \rightarrow \infty} a_t = \ell$ αν και μόνο αν για κάθε ακολουθία $(t_n)_{n \geq 1}$ με $t_n \rightarrow \infty$ ισχύει $\lim_{n \rightarrow \infty} a_{t_n} = \ell$. Έτσι, αναγώμαστε στο γνωστό θεώρημα κυριαρχημένης σύγκλισης.

5.14 Στην απόδειξη της ανισότητας Markov μπορούμε να είμαστε λιγότερο γεναιόδωροι και να παρατηρήσουμε ότι η ποσότητα $n \mathbf{P}(|X| \geq n)$ είναι μικρότερη από τη μέση τιμή $\mathbf{E}(|X| \mathbf{1}_{|X| \geq n})$.

5.16 Για το (α), αρκεί να το δείξουμε για κάθε ακολουθία $(\varepsilon_n)_{n \geq 1}$ θετικών αριθμών με $\varepsilon_n \rightarrow 0$. Χρησιμοποιούμε το θεώρημα κυριαρχημένης σύγκλισης με κυριαρχούσα συνάρτηση την 1 αφού

$$\left| \frac{X}{\varepsilon} \mathbf{1}_{|X| < \varepsilon} \right| \leq 1.$$

5.18 $X = \sum_{k=1}^X 1 = \sum_{k=1}^{\infty} \mathbf{1}_{k \leq X}$ και θεώρημα Beppo Levi.

5.19 $[X] \leq X \leq [X] + 1$.

Κεφάλαιο 6

6.4 Εύκολα ελέγχουμε την ισότητα για $Y = \sum_{i=1}^n a_i \mathbf{1}_{A_i}$, $A_i \in \mathcal{F}$, απλή αφού $\int \mathbf{1}_A d\mathbf{Q} = \mathbf{Q}(A) = \mathbf{E}_{\mathbf{P}}(X \mathbf{1}_A)$ για κάθε $A \in \mathcal{F}$ και λόγω γραμμικότητας. Αν τώρα η Y είναι μη αρνητική, γνωρίζουμε ότι υπάρχει ακολουθία $(Y_n)_{n \in \mathbb{N}}$ απλών μη αρνητικών συναρτήσεων έτσι ώστε $Y_n \nearrow Y$. Από το θεώρημα μονότονης σύγκλισης, έχουμε ότι $\lim_{n \rightarrow \infty} \int Y_n d\mathbf{Q} = \int Y d\mathbf{Q}$ και $\lim_{n \rightarrow \infty} \mathbf{E}_{\mathbf{P}}(Y_n X) = \mathbf{E}_{\mathbf{P}}(Y X)$. Όμως, $\int Y_n d\mathbf{Q} = \mathbf{E}_{\mathbf{P}}(Y_n X)$ για κάθε $n \in \mathbb{N}$, άρα και πάλι το ζητούμενο ισχύει. Τέλος, αν $\mathbf{E}_{\mathbf{P}}(|Y|X) < \infty$, από τα προηγούμενα έχουμε ότι $\int Y^- d\mathbf{Q} = \mathbf{E}_{\mathbf{P}}(Y^- X)$, $\int Y^+ d\mathbf{Q} = \mathbf{E}_{\mathbf{P}}(Y^+ X)$ όπου και οι τέσσερις αυτοί αριθμοί είναι πεπερασμένοι εφόσον $\mathbf{E}_{\mathbf{P}}(Y^- X) + \mathbf{E}_{\mathbf{P}}(Y^+ X) = \mathbf{E}_{\mathbf{P}}(|Y|X) < \infty$, $\int Y^- d\mathbf{Q} + \int Y^+ d\mathbf{Q} = \int |Y| d\mathbf{Q} = \mathbf{E}_{\mathbf{P}}(|Y|X)$ (η τελευταία ισότητα ισχύει αφού η $|Y|$ είναι θετική). Συνεπώς,

$$\int Y d\mathbf{Q} = \int Y^+ d\mathbf{Q} - \int Y^- d\mathbf{Q} = \mathbf{E}_{\mathbf{P}}(Y^+ X) - \mathbf{E}_{\mathbf{P}}(Y^- X) = \mathbf{E}_{\mathbf{P}}(Y X).$$

6.5 Για την πρώτη ανισότητα, παρατηρούμε ότι η συνάρτηση $x \mapsto \mathbf{P}(X > x)$ έχει παράγωγο $-e^{-x^2/2}/\sqrt{2\pi}$ και μελετούμε τη συνάρτηση της διαφοράς των δύο μελών.

Κεφάλαιο 7

7.1 Για το (α), έχουμε

$$\mathbf{P}(\{x\}) = \mathbf{P}\left(\bigcap_{n \in \mathbb{N}^+} \left(x - \frac{1}{n}, x\right]\right) = \lim_{n \rightarrow \infty} \mathbf{P}\left(\left(x - \frac{1}{n}, x\right]\right) \stackrel{(7.1)}{=} \lim_{n \rightarrow \infty} \left(F(x) - F\left(x - \frac{1}{n}\right)\right) = F(x) - F(x-).$$

Το όριο $F(x-)$ υπάρχει γιατί η F είναι αύξουσα. Για το (β),

$$\mathbf{P}([x, y]) = \mathbf{P}(\{x\}) + \mathbf{P}((x, y]) \stackrel{(α),(7.1)}{=} F(x) - F(x-) + F(y) - F(x) = F(y) - F(x-).$$

Τα (γ) και (δ) προκύπτουν με τον ίδιο τρόπο.

7.2 Έστω $A(F) := \{x \in \mathbb{R} : \text{Η } F \text{ είναι ασυνεχής στο } x\}$. Επειδή η F είναι αύξουσα, σε κάθε σημείο ασυνέχειας, η F έχει άλμα προς τα πάνω, δηλαδή, $F(x-) < F(x+)$ (βέβαια, $F(x+) = F(x)$, αλλά δεν το χρειαζόμαστε). Για $x \in A(F)$ επιλέγουμε έναν ρητό $q_x \in (F(x-), F(x+))$. Επειδή η F είναι αύξουσα, η απεικόνιση $x \mapsto q_x$ είναι 1-1 από το $A(F)$ στο \mathbb{Q} .

Εναλλακτικά, εφαρμόζουμε την Άσκηση 2.5 για $B = A(F)$ και $A_\beta = \{\beta\}$ για κάθε $\beta \in B$. Ισχύει $\mathbf{P}(A_\beta) = \mathbf{F}(\beta) - \mathbf{F}(\beta-) > 0$ αφού η \mathbf{F} έχει άλμα στο β .

7.3 (α) Έχουμε ότι

$$\begin{aligned} \mathbf{P}((0, 4)) &= \lambda \mathbf{P}_1((0, 4)) + (1 - \lambda) \mathbf{P}_2((0, 4)) = \lambda \int_0^4 e^{-x} dx + (1 - \lambda) \frac{1}{2} \\ &= \lambda(1 - e^{-4}) + (1 - \lambda) \frac{1}{2}. \end{aligned}$$

(β) Γνωρίζουμε ότι $F(x) = \mathbf{P}((-\infty, x]) = \lambda \mathbf{P}_1((-\infty, x]) + (1 - \lambda) \mathbf{P}_2((-\infty, x])$. Εύκολα μπορεί κανείς να ελέγξει ότι

$$F(x) = \begin{cases} 0 & \text{αν } x < -2, \\ (1 - \lambda) \frac{1}{2} & \text{αν } -2 \leq x \leq 0, \\ \lambda(1 - e^{-x}) + (1 - \lambda) \frac{1}{2} & \text{αν } 0 < x < 3, \\ 1 - \lambda e^{-x} & \text{αν } x \geq 3. \end{cases}$$

Κεφάλαιο 8

8.1 Για τα ω στο σύνολο $\limsup_n A_n^\varepsilon$ έχουμε $|X_n| \geq \varepsilon$ για άπειρα n και άρα $\limsup_n A_n^\varepsilon \subset \Omega \setminus \{\lim_{n \rightarrow \infty} X_n = 0\}$. Για το ότι το (β) δίνει το (α), παρατηρούμε ότι $\Omega \setminus \{\lim_{n \rightarrow \infty} X_n = 0\} = \bigcup_{k=1}^{\infty} \limsup_n A_n^{1/k}$ και χρησιμοποιούμε την Άσκηση 2.3 (α).

8.2 (α) Για την τριγωνική ανισότητα. Επειδή $|X - Z| \leq |X - Y| + |Y - Z|$ και η $f(x) = x/(1+x) = 1 - (1+x)^{-1}$ είναι αύξουσα [στο $[0, \infty)$], αρκεί να δείξει κανείς ότι για $x, y \geq 0$ ισχύει $f(x+y) \leq f(x) + f(y)$. Αυτό προκύπτει με πράξεις ή παρατηρώντας ότι η $x \mapsto f(x+y) - f(x) - f(y)$ είναι φθίνουσα στο $[0, \infty)$.

8.5 Για κάθε $\varepsilon > 0$ και $n \geq 1$ ισχύει $\mathbf{P}(|X - Y| > \varepsilon) \leq \mathbf{P}(|X - X_n| > \varepsilon/2) + \mathbf{P}(|X_n - Y| > \varepsilon/2)$.

Κεφάλαιο 9

9.1

$$\int_0^1 f(x) dx = \sum_{n=1}^{\infty} \frac{1}{n^2} \int_0^1 \frac{1}{\sqrt{|x - r_n|}} dx \leq 2 \sum_{n=1}^{\infty} \frac{1}{n^2} \int_0^1 \frac{1}{\sqrt{x}} dx < \infty$$

Από γνωστή πρόταση [Πρόταση 5.12(iii)] έπεται ότι το σύνολο των $x \in (0, 1)$ με $f(x) = \infty$ έχει μέτρο Lebesgue 0.

9.2 (α) Ένας τρόπος. Γράφουμε

$$g(X) = g(0) + \int_0^X g'(t) dt = g(0) + \int_0^\infty g'(t) \mathbf{1}_{t < X} dt.$$

Παίρνουμε μέση τιμή και εφαρμόζουμε το Θεώρημα Fubini.

9.4 (β) Σχόλιο: Η υπόθεση ότι η συνάρτηση κατανομής της X είναι συνεχής μπορεί να παραληφθεί αλλά τότε η απόδειξη είναι λίγο πιο απαιτητική.

Κεφάλαιο 10

10.1 Θέλουμε $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$ για κάθε $A \in \mathcal{F}_1, B \in \mathcal{F}_2$. Θεωρούμε τα δύο σενάρια $\mathbf{P}(A) = 0$ ή $\mathbf{P}(A) = 1$.

10.2. (α) \Rightarrow (β) Έστω $A, B \in \mathcal{B}(\mathbb{R})$. Τότε,

$$\begin{aligned} \mathbf{P}(f(X) \in A, g(Y) \in B) &= \mathbf{P}(X \in f^{-1}(A), Y \in g^{-1}(B)) = \mathbf{P}(X \in f^{-1}(A))\mathbf{P}(Y \in g^{-1}(B)) \\ &= \mathbf{P}(f(X) \in A)\mathbf{P}(g(Y) \in B). \end{aligned}$$

Η δεύτερη ισότητα προκύπτει από την ανεξαρτησία των X, Y . Το ζητούμενο δείχθηκε.

(β) \Rightarrow (α) Έστω $\Gamma \in \mathcal{E}, \Delta \in \mathcal{G}$. Το (β) για τις $f = \mathbf{1}_\Gamma$ και $g = \mathbf{1}_\Delta$ δίνει

$$\mathbf{P}(\{f(X) \in \{1\}\} \cap \{g(X) \in \{1\}\}) = \mathbf{P}(\{f(X) \in \{1\}\})\mathbf{P}(\{g(X) \in \{1\}\})$$

δηλαδή

$$\mathbf{P}(X \in \Gamma, Y \in \Delta) = \mathbf{P}(X \in \Gamma)\mathbf{P}(Y \in \Delta),$$

που είναι το ζητούμενο.

10.3. Έστω ότι δεν υπάρχει τέτοιο c . Τότε υπάρχει $a \in \mathbb{R}$ ώστε $\mathbf{P}(X \leq a) = p \in (0, 1)$ και επομένως $\mathbf{P}(X > a) = 1 - p \in (0, 1)$. Η ανεξαρτησία δίνει

$$\mathbf{P}(X \leq a, Y > a) = \mathbf{P}(X \leq a)\mathbf{P}(Y > a) > 0,$$

ενώ το αριστερό μέλος είναι μικρότερο από $\mathbf{P}(X \neq Y) = 0$ από υπόθεση. Άτοπο.

10.6

$$\mathbf{E}|X + Y| = \int_{\mathbb{R}^2} |x + y| d\mathbf{P}^{(X,Y)}(x, y) = \int_{\mathbb{R}^2} |x + y| d(\mathbf{P}^X \otimes \mathbf{P}^Y)(x, y) = \int_{\mathbb{R}} \int_{\mathbb{R}} |x + y| d\mathbf{P}^X(x) d\mathbf{P}^Y(y).$$

Η Πρόταση 5.14(iii) δίνει ότι το σύνολο των $y \in \mathbb{R}$ με $\int_{\mathbb{R}} |x + y| d\mathbf{P}^X(x) < \infty$ έχει \mathbf{P}^Y -μέτρο 1. Παίρνουμε ένα y_0 σε εκείνο το σύνολο. Θα έχουμε λοιπόν $\mathbf{E}|X + y_0| < \infty$, άρα $\mathbf{E}|X| = \mathbf{E}|X + y_0 - y_0| \leq \mathbf{E}|X + y_0| + |y_0| < \infty$.

10.9 (γ) Έχουμε ότι $\mathbf{E}(S_n) = n\mathbf{E}(X_1) = 0$ και λόγω του (α) ότι $\mathbf{E}(S_n^2) = n$. Έστω $\epsilon > 0$. Τότε,

$$\begin{aligned} \mathbf{P}\left(\left|\frac{S_n}{n}\right| > \epsilon\right) &= \mathbf{P}(|S_n| > n\epsilon) = \mathbf{P}(S_n^2 > n^2\epsilon^2) \\ &\leq \frac{\mathbf{E}(S_n^2)}{n^2\epsilon^2} = \frac{n}{n^2\epsilon^2} = \frac{1}{n\epsilon^2} \rightarrow 0 \end{aligned}$$

για $n \rightarrow \infty$.

10.10. Για $\epsilon > 0$, χρησιμοποιώντας την ανεξαρτησία των X_1, X_2, \dots , βρίσκουμε

$$\mathbf{P}(|m_n| > \epsilon) = \mathbf{P}(m_n > \epsilon) = \mathbf{P}(X_1 > \epsilon, \dots, X_n > \epsilon) = (1 - \epsilon)^n \rightarrow 0,$$

και όμοια

$$\mathbf{P}(|M_n - 1| > \varepsilon) = \mathbf{P}(M_n < 1 - \varepsilon) = \mathbf{P}(X_1 < 1 - \varepsilon, \dots, X_n < 1 - \varepsilon) = (1 - \varepsilon)^n \rightarrow 0.$$

Κεφάλαιο 11

11.3. Αποδεικνύουμε το ζητούμενο για τη M_n . Το όριο είναι το πολύ 1 αφού κάθε μια X_i είναι το πολύ 1. Για το κάτω φράγμα, εφαρμόζουμε το πρώτο λήμμα Borel-Cantelli. Για $\varepsilon > 0$, θέτουμε $A_n^\varepsilon := \{M_n \leq 1 - \varepsilon\}$. Επειδή $\mathbf{P}(A_n^\varepsilon) \leq (1 - \varepsilon)^n$ και $\sum_{n \geq 1} \mathbf{P}(A_n^\varepsilon) < \infty$, με πιθανότητα 1 για όλα τα μεγάλα n ισχύει $M_n \geq 1 - \varepsilon$. Άρα το σύνολο $B_\varepsilon := \{\liminf_{n \rightarrow \infty} M_n \geq 1 - \varepsilon\}$ έχει πιθανότητα 1 και επομένως το ίδιο ισχύει και για το $\bigcap_{k=1}^{\infty} B_{1/k} = \{\liminf_{n \rightarrow \infty} M_n \geq 1\}$.

Εναλλακτικά, μπορεί να παρατηρήσει κανείς ότι $\mathbf{P}(A_n^{1/\sqrt{n}}) = (1 - n^{-1/2})^n \leq e^{-\sqrt{n}}$ (με χρήση της $1 - x \leq e^{-x}$) και $\sum_{n \geq 1} e^{-\sqrt{n}} < \infty$. Το συμπέρασμα έπεται από το πρώτο λήμμα Borel-Cantelli.

11.6 Αρκεί για κάθε $n \geq 1$ να βρούμε σταθερά M_n ώστε $\mathbf{P}(|X_n| > M_n) \leq n^{-2}$. Τότε η $a_n = nM_n$ ικανοποιεί το ζητούμενο (πρώτο λήμμα Borel-Cantelli).

11.7 Η κατεύθυνση \Leftarrow είναι πιο εύκολη. Αν υπάρχει τέτοιο M τότε (από το πρώτο λήμμα Borel-Cantelli) με πιθανότητα 1, ισχύει $X_n \leq M$ για όλα τα μεγάλα n και έπεται το συμπέρασμα.

[Να το γράψουμε και τυπικά. Το σύνολο $A := \limsup_{n \geq 1} \{X_n > M\}$ έχει πιθανότητα 0 και για κάθε $\omega \in \Omega \setminus A$ υπάρχει φυσικός $n_0(\omega)$ ώστε $X_n \leq M$ για κάθε $n \geq n_0(\omega)$. Άρα

$$X^*(\omega) \leq \max\{X_1, X_2, \dots, X_{n_0(\omega)-1}, M\} < \infty,$$

ως μέγιστο πεπερασμένου αριθμού πραγματικών αριθμών.]

Για την άλλη κατεύθυνση, έστω ότι δεν υπάρχει τέτοιο M , τότε για κάθε $K \in \mathbb{N}$, το δεύτερο λήμμα Borel-Cantelli δίνει ότι $\mathbf{P}(\limsup_n \{X_n > K\}) = 1$ (εδώ χρησιμοποιούμε την ανεξαρτησία των X_n). Επομένως το $C_K := \{X^* \geq K\}$ έχει πιθανότητα 1 και άρα και το $\bigcap_{K=1}^{\infty} C_K$ (αριθμησίμη τομή συνόλων με πιθανότητα 1). Όμως $\bigcap_{K=1}^{\infty} C_K = \{X^* = \infty\}$, το οποίο από υπόθεση έχει πιθανότητα 0, και έχουμε άτοπο.

11.8 Χρήσιμη είναι η Άσκηση 6.5.

11.9. Το όριο είναι το πολύ 1 λόγω του Παραδείγματος 11.8. Για το κάτω φράγμα, εφαρμόζουμε το πρώτο λήμμα Borel-Cantelli. Για $\varepsilon > 0$

$$\mathbf{P}\left(\frac{M_n}{\log n} \leq 1 - \varepsilon\right) = \mathbf{P}(X_1 \leq (1 - \varepsilon) \log n)^n = (1 - n^{-(1-\varepsilon)})^n \leq (e^{-n^{\varepsilon-1}})^n = e^{-n^\varepsilon}$$

και συνεχίζουμε όπως στο Παράδειγμα 11.8.

11.10. (α) Έστω $\varepsilon > 0$. Θέτουμε $B_\varepsilon = \left\{\overline{\lim}_{n \rightarrow \infty} \frac{L_n}{\log_2 n} \leq 1 + \varepsilon\right\}$. Θα δείξουμε ότι $\mathbf{P}(B_\varepsilon) = 1$.

Έστω $A_n = \left\{\frac{L_n}{\log_2 n} > 1 + \varepsilon\right\}$. Τότε

$$\begin{aligned} \mathbf{P}(A_n) &= \mathbf{P}(L_n > (1 + \varepsilon) \log_2 n) = \mathbf{P}(X_n = X_{n+1} = \dots = X_{n+[(1+\varepsilon)\log_2 n]-1} = K \text{ ή } \Gamma) \\ &= 2 \mathbf{P}(X_n = X_{n+1} = \dots = X_{n+[(1+\varepsilon)\log_2 n]-1} = K) \\ &= 2 \left(\frac{1}{2}\right)^{[(1+\varepsilon)\log_2 n]} \leq 2 \left(\frac{1}{2}\right)^{(1+\varepsilon)\log_2 n-1} = \frac{4}{2^{(1+\varepsilon)\log_2 n}} = \frac{4}{n^{1+\varepsilon}}. \end{aligned}$$

Συνεπώς,

$$\sum_{n=1}^{\infty} \mathbf{P}(A_n) < \infty$$

και από το πρώτο Λήμμα Borel-Cantelli, $\mathbf{P}(\limsup_{n \geq 1} A_n) = 0$, δηλαδή $\mathbf{P}((\limsup_{n \geq 1} A_n)^c) = 1$.

Υποδείξεις για επιλεγμένες ασκήσεις

Έστω τώρα $\omega \in (\limsup_{n \geq 1} A_n)^c$. Τότε υπάρχει $n_0(\omega) \in \mathbb{N}$ τέτοιο ώστε για κάθε $n \geq n_0(\omega)$ να ισχύει

$$\frac{L_n(\omega)}{\log_2 n} \leq 1 + \epsilon,$$

άρα

$$\overline{\lim}_{n \rightarrow \infty} \frac{L_n(\omega)}{\log_2 n} \leq 1 + \epsilon.$$

Προκύπτει, λοιπόν, ότι $\omega \in B_\epsilon$. Άρα $(\limsup_{n \geq 1} A_n)^c \subset B_\epsilon$, οπότε $\mathbf{P}(B_\epsilon) = 1$.

Επειδή

$$\left\{ \overline{\lim}_{n \rightarrow \infty} \frac{L_n}{\log_2 n} \leq 1 \right\} = \bigcap_{k=1}^{\infty} B_{\frac{1}{k}}$$

και $\mathbf{P}(B_{1/k}) = 1$ για κάθε $k \geq 1$, από την Άσκηση 2.3 (β) έχουμε

$$\mathbf{P}\left(\bigcap_{k=1}^{\infty} B_{\frac{1}{k}}\right) = 1.$$

Έτσι, το (α) αποδείχθηκε.

(β) Επειδή κάθε L_n παίρνει τιμή που είναι ένας θετικός ακέραιος ή ∞ , το ζητούμενο ισοδυναμεί με $\mathbf{P}(L_n = 1 \text{ άπειρες φορές}) = 1$ (δηλαδή ο μόνος τρόπος να πλησιάσει η L_n το 1 είναι να πέσει πάνω του).

Έστω

$$B_n = \{X_{2n} = K, X_{2n+1} = \Gamma\}$$

για κάθε $n \geq 1$. Εύκολα βλέπουμε ότι τα $(B_n)_{n \geq 1}$ είναι ανεξάρτητα και ισχύει ότι $\mathbf{P}(B_n) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$. Άρα,

$$\sum_{n=1}^{\infty} \mathbf{P}(B_n) = \infty.$$

Από το δεύτερο Λήμμα Borel-Cantelli, έχουμε ότι $\mathbf{P}(\limsup_{n \geq 1} B_n) = 1$. Όμως

$$\limsup_{n \geq 1} B_n \subset \{L_n = 1 \text{ άπειρες φορές}\}.$$

Άρα και το τελευταίο ενδεχόμενο έχει πιθανότητα 1.

11.11 Χρησιμοποιώντας την Άσκηση 8.1 και ότι οι X_n είναι ανεξάρτητες, δείχνουμε πρώτα ότι ο ισχυρισμός για το όριο είναι ισοδύναμος με τη $\sum_{n=1}^{\infty} \mathbf{P}(|X_n| \geq \epsilon n) < \infty$ για κάθε $\epsilon > 0$. Έπειτα χρησιμοποιούμε ότι οι X_n είναι ισόνομες και την Άσκηση 5.19.

11.12 Χρησιμοποιούμε την Άσκηση 5.19 και τα δύο λήμματα Borel-Cantelli.

11.14. Αν το δεξί μέλος είναι ∞ , δεν έχουμε τίποτα να αποδείξουμε. Αν είναι πεπερασμένο, ας το ονομάσουμε a . Έπεται ότι υπάρχει ένας θετικός ακέραιος n_0 ώστε $\mathbf{E}(X_n) < \infty$ για κάθε $n \geq n_0$. Ως πρώτο βήμα δείχνουμε ότι για κάθε $\epsilon > 0$, με πιθανότητα 1, η ανισότητα

$$\frac{1}{n} \log X_n > a + \epsilon$$

ισχύει μόνο για πεπερασμένα n .

11.17. (β) Από το προηγούμενο ερώτημα, οι $(X_i)_{i \in \mathbb{I}}$ είναι ανεξάρτητες. Άρα ο νόμος 0-1 του Kolmogorov εφαρμόζεται για την τελική σ -αλγεβρα τους

$$C_\infty := \bigcap_{n=1}^{\infty} \sigma(X_n, X_{n+1}, \dots).$$

Βέβαια το αν ένα $\omega \in \Omega$ ανήκει σε ένα από τα $\liminf A_i, \limsup A_i$ δεν εξαρτάται από οποιαδήποτε πεπερασμένο πλήθος $X_1(\omega), \dots, X_k(\omega)$, οπότε και τα δύο σύνολα ανήκουν στη C_∞ . Τυπικά το αποδεικνύουμε ως εξής. Για κάθε $n \geq 1$,

$$\liminf_i A_i = \bigcup_{j=1}^{\infty} \bigcap_{k=j}^{\infty} A_k = \bigcup_{j=n}^{\infty} \bigcap_{k=j}^{\infty} A_k = \bigcup_{j=n}^{\infty} \bigcap_{k=j}^{\infty} X_k^{-1}(\{1\}) \in \sigma(X_n, X_{n+1}, \dots).$$

Η δεύτερη ισότητα ισχύει γιατί έχουμε ένωση μιας άξουσας ακολουθίας συνόλων. Όμοια

$$\limsup_i A_i = \bigcap_{j=1}^{\infty} \bigcup_{k=j}^{\infty} A_k = \bigcap_{j=n}^{\infty} \bigcup_{k=j}^{\infty} A_k = \bigcap_{j=n}^{\infty} \bigcup_{k=j}^{\infty} X_k^{-1}(\{1\}) \in \sigma(X_n, X_{n+1}, \dots).$$

11.18. (α) Από τον απειροστικό λογισμό ξέρουμε ότι $R^{-1} = \overline{\lim_{n \rightarrow \infty} |X_n|^{1/n}}$.

(β) Δείχνουμε όπως στην Άσκηση 11.8 ότι $\overline{\lim_{n \rightarrow \infty} |X_n|^{1/n}} = 1$ με πιθανότητα 1. Μάλιστα εδώ αρκεί η χρήση του πρώτου λήμματος Borel-Cantelli για να δείξουμε ότι $\lim_{n \rightarrow \infty} |X_n|^{1/n} = 1$.

11.20. Δουλεύουμε όπως στην προηγούμενη άσκηση.

Κεφάλαιο 12

12.1 Έστω $Y = \frac{S_n}{n}$. Τότε $\mathbf{E}(Y) = \frac{1}{n} \mathbf{E}(X_1 + X_2 + \dots + X_n) = \frac{1}{n} n \mathbf{E}(X_1) = \mu$ και $\text{Var}(Y) = \text{Var}(\frac{1}{n} S_n) = \frac{1}{n^2} \text{Var}(X_1 + X_2 + \dots + X_n) = \frac{1}{n} \text{Var}(X_1)$. Έστω $\epsilon > 0$. Τότε,

$$\mathbf{P}\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) = \mathbf{P}(|Y - \mathbf{E}(Y)| \geq \epsilon) \leq \frac{1}{\epsilon^2} \text{Var}(Y) = \frac{1}{\epsilon^2} \frac{\text{Var}(X_1)}{n} \rightarrow 0$$

για $n \rightarrow \infty$, από το οποίο προκύπτει το ζητούμενο.

12.2 Έστω $M > 0$. Για κάθε $i \geq 1$ θέτουμε $Y_i^M = X_i \wedge M$. Οι Y_i^M είναι ανεξάρτητες και ισόνομες, και ισχύει ότι $(Y_i^M)^- = X_i^-$, $(Y_i^M)^+ \leq M$. Συνεπώς $\mathbf{E}(|Y_i^M|) \leq \mathbf{E}(X_i^-) + M < \infty$. Θέτουμε $S_n^M = Y_1^M + Y_2^M + \dots + Y_n^M$ για κάθε $n \geq 1$. Τότε, με πιθανότητα 1,

$$\underline{\lim}_{n \rightarrow \infty} \frac{S_n}{n} \geq \underline{\lim}_{n \rightarrow \infty} \frac{S_n^M}{n} = \mathbf{E}(Y_1^M),$$

όπου η ισότητα προκύπτει από τον νόμο των μεγάλων αριθμών. Έστω τώρα

$$\Omega_M = \left\{ \omega \in \Omega : \underline{\lim}_{n \rightarrow \infty} \frac{S_n(\omega)}{n} \geq \mathbf{E}(X_1 \wedge M) \right\}.$$

Δείξαμε ότι $\mathbf{P}(\Omega_M) = 1$ για κάθε $M > 0$. Θεωρώντας το σύνολο $A = \bigcap_{k=1}^{\infty} \Omega_k$, έχουμε ότι $\mathbf{P}(A) = 1$ και για $\omega \in A$, ισχύει ότι

$$\underline{\lim}_{n \rightarrow \infty} \frac{S_n(\omega)}{n} \geq \mathbf{E}(X_1 \wedge k) \quad \text{για κάθε } k \geq 1. \quad (\text{B'.23})$$

Από το θεώρημα μονότονης σύγκλισης, $\mathbf{E}(Y_1^k) = \mathbf{E}(X_1^+ \wedge k) - \mathbf{E}(X_1^-) \rightarrow \mathbf{E}(X_1^+) - \mathbf{E}(X_1^-) = \infty$ για $k \rightarrow \infty$. Άρα, για $\omega \in A$, παίρνοντας $k \rightarrow \infty$ στην (B'.23) έχουμε ότι $\underline{\lim}_{n \rightarrow \infty} S_n(\omega)/n = \infty$, οπότε $\lim_{n \rightarrow \infty} S_n(\omega)/n = \infty$, που είναι το ζητούμενο.

12.3 Παρατηρούμε ότι

$$\frac{X_n}{n} = \frac{S_n}{n} - \frac{S_{n-1}}{n-1} = \frac{S_n}{n} - \frac{n-1}{n} \frac{S_{n-1}}{n-1} \rightarrow \mu - \mu = 0 \quad (\text{B'.24})$$

για $n \rightarrow \infty$ με πιθανότητα 1. Αυτό δίνει ότι $\mathbf{E}|X_1| < \infty$ λόγω της Άσκησης 11.11, αλλά θα το αποδείξουμε και εδώ (λύνοντας τη μισή άσκηση). Η (B'.24) συνεπάγεται ότι $\sum_{n=1}^{\infty} \mathbf{P}\left(\frac{|X_n|}{n} \geq 1\right) < \infty$ γιατί αν $\sum_{n=1}^{\infty} \mathbf{P}\left(\frac{|X_n|}{n} \geq 1\right) = \infty$, τότε το 2ο λήμμα Borel-Cantelli, εφαρμοζόμενο στην ακολουθία ανεξάρτητων

Υποδείξεις για επιλεγμένες ασκήσεις

ενδεχομένων $A_n = \left\{ \frac{|X_n|}{n} \geq 1 \right\}$, $n \geq 1$ θα έδινε ότι $\overline{\lim} |X_n|/n \geq 1$ με πιθανότητα 1, το οποίο συγκρούεται με την (Β'.24). Τέλος, ισχύει ότι (Άσκηση 5.19)

$$\sum_{n=1}^{\infty} \mathbf{P}(|X_1| \geq n) \leq \mathbf{E}(|X_1|) \leq \sum_{n=1}^{\infty} \mathbf{P}(|X_1| \geq n) + 1.$$

και επιπλέον, επειδή οι $(X_n)_{n \geq 1}$ είναι ισόνομες, $\sum_{n=1}^{\infty} \mathbf{P}(|X_n| \geq n) = \sum_{n=1}^{\infty} \mathbf{P}(|X_1| \geq n)$. Άρα $\mathbf{E}(X_1) \in \mathbb{R}$ και από τον νόμο των μεγάλων αριθμών $\frac{S_n}{n} \rightarrow \mathbf{E}(X_1)$. Όμως, από υπόθεση $\frac{S_n}{n} \rightarrow \mu$, επομένως $\mathbf{E}(X_1) = \mu$.

12.4 Έχουμε ότι $\mathbf{E}(X_1) = 1$ και $\mathbf{E}(X_1^2) = \text{Var}(X_1) + \mathbf{E}(X_1)^2 = 4$, καθώς επίσης και ότι οι $\{X_i^2, i \geq 1\}$ είναι ανεξάρτητες και ισόνομες. Γράφουμε

$$\frac{X_1 + X_2 + \dots + X_n}{(X_1)^2 + (X_2)^2 + \dots + (X_n)^2} = \frac{\frac{X_1 + X_2 + \dots + X_n}{n}}{\frac{(X_1)^2 + (X_2)^2 + \dots + (X_n)^2}{n}}.$$

Από τον νόμο των μεγάλων αριθμών ο αριθμητής συγκλίνει στο 1 με πιθανότητα 1 και αντίστοιχα ο παρονομαστής στο 2. Τυπικά, για τα

$$A_1 = \left\{ \omega \in \Omega : \lim_{n \rightarrow \infty} \frac{X_1 + X_2 + \dots + X_n}{n} = 1 \right\}$$

$$A_2 = \left\{ \omega \in \Omega : \lim_{n \rightarrow \infty} \frac{(X_1)^2 + (X_2)^2 + \dots + (X_n)^2}{n} = 1 \right\}$$

έχουμε ότι $\mathbf{P}(A_1) = \mathbf{P}(A_2) = 1$. Άρα $\mathbf{P}(A_1 \cap A_2) = 1$ και για $\omega \in A_1 \cap A_2$ έχουμε

$$\lim_{n \rightarrow \infty} \frac{X_1(\omega) + X_2(\omega) + \dots + X_n(\omega)}{(X_1)^2(\omega) + (X_2)^2(\omega) + \dots + (X_n)^2(\omega)} = \frac{1}{4},$$

από το οποίο προκύπτει το ζητούμενο.

12.6. (α) $(U_1 U_2 \dots U_n)^{1/n} = e^{\frac{1}{n}(\log U_1 + \dots + \log U_n)}$ Επειδή $\mathbf{E}(\log U_1) = \int_0^1 \log x \, dx = \dots = -1$, ο ισχυρός νόμος των μεγάλων αριθμών για την ακολουθία $(\log U_i)_{i \geq 1}$ δίνει ότι

$$\lim_{n \rightarrow \infty} \frac{\log U_1 + \dots + \log U_n}{n} = -1 \text{ με πιθανότητα 1.}$$

Το συμπέρασμα έπεται.

(β) Έπεται από το (α). Επιλέγουμε θ ώστε $e^{-1} < \theta < 1$. Με πιθανότητα 1, υπάρχει $n_0 \in \mathbb{N}$ ώστε για $n > n_0$ να ισχύει $(U_1 U_2 \dots U_n)^{1/n} < \theta$. Άρα για $n > n_0$

$$0 < U_1 U_2 \dots U_n < \theta^n \rightarrow 0$$

καθώς $n \rightarrow \infty$ επειδή $0 < \theta < 1$.

Εναλλακτικά, $U_1 U_2 \dots U_n = e^{n \frac{\log U_1 + \dots + \log U_n}{n}} \rightarrow 0$ αφού το όριο του εκθέτη είναι $\infty \times (-1) = -\infty$.

(γ) Η ακολουθία $(U_i^a)_{i \geq 1}$ αποτελείται από ανεξάρτητες και ισόνομες τυχαίες μεταβλητές, καθεμία με μέση τιμή

$$\mathbf{E}(U_1^a) = \int_0^1 x^a \, dx = \begin{cases} \frac{1}{1+a} & \text{αν } a > -1, \\ \infty & \text{αν } a \leq -1. \end{cases}$$

Το συμπέρασμα έπεται από τον ισχυρό νόμο των μεγάλων αριθμών και την Άσκηση 12.2.

12.7. Οι όροι της ακολουθίας $((X_i - \mu)^2)_{i \geq 1}$ είναι ανεξάρτητες ισόνομες τυχαίες μεταβλητές, καθεμία με μέση τιμή $\mathbf{E}((X_1 - \mu)^2) = V(X_1) = \sigma^2$. Ο ισχυρός νόμος των μεγάλων αριθμών δίνει το συμπέρασμα.

12.8. (β) Γιατί βρίσκεται αυτή η άσκηση σε αυτό το κεφάλαιο;

Κεφάλαιο 13

13.4 Η ροπογεννήτρια της X είναι

$$M_X(t) = \begin{cases} \left(1 - \frac{t}{\lambda}\right)^{-a} & \text{αν } t < \lambda, \\ \infty & \text{αν } t \geq \lambda. \end{cases}$$

Η συνάρτηση $f : \mathbb{C} \setminus [\lambda, \infty) \rightarrow \mathbb{C}$ με

$$f(z) = \left(1 - \frac{z}{\lambda}\right)^{-a} = e^{-a \log(1 - \frac{z}{\lambda})}$$

είναι αναλυτική στο πεδίο ορισμού της, το οποίο περιέχει το $\{z \in \mathbb{C} : |\operatorname{Re}(z)| < \lambda\}$, και $M_X(t) = f(t)$ για κάθε $t \in (-\lambda, \lambda)$. Το συμπέρασμα έπεται από την Πρόταση 13.11.

13.7 Έστω $u \in \mathbb{R}$. Τότε $\phi_{-X}(u) = \mathbf{E}(e^{iu(-X)}) = \phi_X(-u) = \overline{\phi_X(u)}$. Συνεπώς, αν $X \stackrel{d}{=} -X$, έχουμε ότι $\phi_X(u) = \phi_{-X}(u) = \overline{\phi_X(u)}$ για κάθε $u \in \mathbb{R}$, δηλαδή $\phi_X(u) \in \mathbb{R}$. Αντίστροφα, αν $\phi_X(u) \in \mathbb{R}$, έχουμε ότι $\phi_X(u) = \overline{\phi_X(u)}$, όμως από τα παραπάνω $\overline{\phi_X(u)} = \phi_{-X}(u)$, άρα $X \stackrel{d}{=} -X$.

13.8 1η λύση (στοιχειώδης): Το ζευγάρι (X, Y) έχει την ίδια κατανομή με το (Y, X) (η οποία είναι το μέτρο γινόμενο $\mathbf{P}^X \otimes \mathbf{P}^X$). Θεωρούμε τη συνάρτηση $f(x, y) = x - y$. Τότε $X - Y = f(X, Y) \stackrel{d}{=} f(Y, X) = Y - X$.

2η λύση: Έστω $u \in \mathbb{R}$. Τότε,

$$\begin{aligned} \phi_{X-Y}(u) &= \mathbf{E}(e^{iuX} e^{-iuY}) = \mathbf{E}(e^{iuX}) \mathbf{E}(e^{-iuY}) \\ &= \phi_X(u) \phi_Y(-u) = \phi_X(u) \phi_X(-u) \\ &= \phi_X(u) \overline{\phi_X(u)} = |\phi_X(u)|^2 \in \mathbb{R}, \end{aligned}$$

όπου η δεύτερη ισότητα προκύπτει λόγω ανεξαρτησίας. Από την Άσκηση 13.7 προκύπτει ότι η $X - Y$ έχει συμμετρική κατανομή.

13.9 Η χαρακτηριστική συνάρτηση της τυχαίας μεταβλητής $(X_1 + X_2 + \dots + X_n)/n$ είναι

$$\phi_{\frac{X_1 + \dots + X_n}{n}}(t) = \mathbf{E}(e^{iX_1 t/n} \dots e^{iX_n t/n}) = \mathbf{E}(e^{iX_1 t/n}) \dots \mathbf{E}(e^{iX_n t/n}) = \{\phi_{X_1}(t/n)\}^n = (e^{-|t|/n})^n = e^{-|t|},$$

όπου η δεύτερη ισότητα προκύπτει λόγω ανεξαρτησίας. Γνωρίζουμε ότι η $e^{-|t|}$ είναι χαρακτηριστική συνάρτηση μιας τυχαίας μεταβλητής με κατανομή Cauchy. Από το θεώρημα μοναδικότητας 13.6 έπεται ότι $\frac{X_1 + \dots + X_n}{n} \sim \mathbf{Cauchy}$.

13.11 Η σειρά συγκλίνει απολύτως σε μια τυχαία μεταβλητή X . Υπολογίζουμε τη χαρακτηριστική συνάρτηση της X και δείχνουμε ότι είναι αυτή της ομοιόμορφης στο $(-1, 1)$.

13.13

$$\phi_X(t) - 1 = 2 \int_1^\infty \frac{1}{|x|^3} (\cos(tx) - 1) dx$$

Το $|\cos(x) - 1|$ είναι περίπου $x^2/2$ όταν το x είναι κοντά στο 0 και φράσσεται από το 2 για οποιοδήποτε x .

13.14 Έχουμε ότι $f_X(x) = f_Y(x) = \mathbf{1}_{(0,1)}(x)$ για κάθε $x \in \mathbb{R}$. Από το Θεώρημα 13.17 η $X+Y$ έχει πυκνότητα $f_{X+Y}(z) = \int_{\mathbb{R}} f_X(x) f_Y(z-x) dx$. Ο ολοκληρωτέος είναι μη μηδενικός για (x, z) τέτοια ώστε $0 < x < 1$ και $0 < z-x < 1$, δηλαδή,

$$\begin{aligned} 0 < x < 1, & \quad \text{και} \\ z-1 < x < z. & \end{aligned} \tag{B'.25}$$

Υποδείξεις για επιλεγμένες ασκήσεις

- Αν $z \leq 0$ ή $z \geq 2$, τότε δεν υπάρχουν x που ικανοποιούν την (B'.25).
- Αν $z \in (0, 1)$, τότε οι σχέσεις της (B'.25) ικανοποιούνται ακριβώς για $0 < x < z$ και έτσι $f_{X+Y}(z) = \int_0^z 1 dx = z$.
- Αν $z \in (1, 2)$ τότε οι σχέσεις της (B'.25) ικανοποιούνται ακριβώς για $z - 1 < x < 1$ και έτσι $f_{X+Y}(z) = \int_{z-1}^1 1 dx = 2 - z$.

Συνδυάζοντας τα παραπάνω, προκύπτει το ζητούμενο.

Κεφάλαιο 14

14.1. Η συνάρτηση κατανομής της U είναι

$$F_U(x) = \begin{cases} 0 & \text{αν } x < 0, \\ x & \text{αν } x \in [0, 1], \\ 1 & \text{αν } x > 1, \end{cases}$$

και τα σημεία συνέχειάς της είναι όλο το \mathbb{R} . Για $x \in [0, 1]$,

$$F_{X_n/n}(x) = \mathbf{P}(X_n \leq nx) = \frac{[nx]}{n} \rightarrow x = F_U(x)$$

για $n \rightarrow \infty$ (π.χ., $nx - 1 < [nx] \leq nx$, κ.λπ). Η σύγκλιση για τα υπόλοιπα $x \in \mathbb{R}$ είναι πιο εύκολη.

14.3. Με βάση το Θεώρημα 14.13, η σύγκλιση $X_n \Rightarrow X$ συνεπάγεται (μάλιστα ισοδυναμεί με) την

$$\mathbf{P}(X_n \in A) \rightarrow \mathbf{P}(X \in A) \text{ για όλα τα } A \subset \mathbb{R} \text{ Borel με } \mathbf{P}(X \in \partial A) = 0.$$

Είναι δυνατόν βέβαια η σχέση $\mathbf{P}(X_n \in A) \rightarrow \mathbf{P}(X \in A)$ να ισχύει για ένα σύνολο Borel A συμπτωματικά, ίσως εξαιτίας της φύσης της ακολουθίας $(X_n)_{n \geq 1}$. Πάντως δεν μας την εγγυάται η $X_n \Rightarrow X$.

(i) Όχι. Γιατί $\partial A = \{2, 32.1, 100\}$, στο οποίο η κατανομή της X δίνει θετική πιθανότητα αφού περιέχει τους θετικούς ακέραιους 2, 100.

(ii) Όχι. Γιατί $\partial A = \bar{A} \setminus A^\circ = \mathbb{R} \setminus \emptyset = \mathbb{R}$ και $\mathbf{P}(X \in \mathbb{R}) = 1 > 0$.

(iii) Ναι. Γιατί $\partial A = \{-1.5, 2.8\}$ και $\mathbf{P}(X \in \{-1.5, 2.8\}) = 0$ αφού μια γεωμετρική τυχαία μεταβλητή παίρνει μόνο ακέραιες θετικές τιμές.

(iv) Ναι. Γιατί $\partial A = \{-2, \pi\}$ και $\mathbf{P}(X \in \partial A) = 0$ αφού η X είναι συνεχής τυχαία μεταβλητή και το ∂A είναι πεπερασμένο.

(v) Όχι. Γιατί $\partial A = \bar{A} \setminus A^\circ = [0, 1/3] \setminus \emptyset = [0, 1/3]$ και $\mathbf{P}(X \in [0, 1/3]) = 1/3 > 0$.

(vi) Όχι. Γιατί $\partial A = \{0, 1/2, 2, 4\}$ και $\mathbf{P}(X \in \{0, 1/2, 2, 4\}) = \mathbf{P}(X = 0) = 3/5 > 0$.

14.4 (α) Έστω $M > 0$. Υπάρχει $x_M > M$ σημείο συνέχειας της F_X . Τότε

$$\liminf_{n \rightarrow \infty} \mathbf{P}(X_n \leq n) \geq \liminf_{n \rightarrow \infty} \mathbf{P}(X_n \leq x_M) = \mathbf{P}(X \leq x_M) \geq \mathbf{P}(X \leq M).$$

Για $M \rightarrow \infty$ το δεξί μέλος τείνει στο 1.

(β) Για κάθε $M \geq 1$ υπάρχουν $x_M \in (x - 1/M, x)$, $y_M \in (x, x + 1/M)$ σημεία συνέχειας της F_X . Τότε

$$\overline{\lim}_{n \rightarrow \infty} \mathbf{P}(X_n \in [x - n^{-1}, x + n^{-1}]) \leq \overline{\lim}_{n \rightarrow \infty} \mathbf{P}(X_n \in [x_M, y_M]) = \mathbf{P}(X \in [x_M, y_M]) = F_X(y_M) - F_X(x_M)$$

Για $M \rightarrow \infty$ το δεξί μέλος τείνει στο $F(x) - F(x-) = \mathbf{P}(X = x)$. Για αντιπαράδειγμα, παίρνουμε $X_n = 2/n$, $X = 0$ (σταθερές συναρτήσεις) και $x = 0$.

14.5 Έστω $\varepsilon > 0$. Για οποιοδήποτε $M > 0$ έχουμε

$$\mathbf{P}(|X_n Y_n| > \varepsilon) \leq \mathbf{P}(|Y_n| \geq \varepsilon/M) + \mathbf{P}(|X_n| \geq M)$$

Για $n \rightarrow \infty$ η πρώτη πιθανότητα στο δεύτερο μέλος τείνει στο 0, ενώ ο δεύτερος όρος είναι φραγμένος από το $\sup_{n \geq 1} \mathbf{P}(|X_n| \geq M)$. Άρα

$$\overline{\lim}_{n \rightarrow \infty} \mathbf{P}(|X_n Y_n| > \varepsilon) \leq \sup_{n \geq 1} \mathbf{P}(|X_n| \geq M)$$

Το αριστερό μέλος δεν εξαρτάται από το M , ενώ το δεξί, για $M \rightarrow \infty$, τείνει στο 0.

14.6 (α) Θα δείξουμε το ζητούμενο για $c = 0$. Έπειτα, για οποιοδήποτε c , θα έχουμε $X_n \Rightarrow X$ και $Y_n - c \xrightarrow{\mathbf{P}} 0$. Η περίπτωση $c = 0$ θα δίνει $X_n + Y_n - c \Rightarrow X$, και αυτό συνεπάγεται άμεσα ότι $X_n + Y_n \Rightarrow X + c$.

Υποθέτουμε λοιπόν ότι $c = 0$. Έστω x σημείο συνέχειας της F_X . Θα δείξουμε ότι $\lim_{n \rightarrow \infty} F_{X_n + Y_n}(x) = F_X(x)$.

ΙΣΧΥΡΙΣΜΟΣ 1: $\overline{\lim}_{n \rightarrow \infty} F_{X_n + Y_n}(x) \leq F_X(x)$.

Έστω $\varepsilon > 0$. Τότε

$$\begin{aligned} F_{X_n + Y_n}(x) &= \mathbf{P}(X_n + Y_n \leq x) = \mathbf{P}(X_n + Y_n \leq x, X_n \leq x + \varepsilon) + \mathbf{P}(X_n + Y_n \leq x, X_n > x + \varepsilon) \\ &\leq \mathbf{P}(X_n \leq x + \varepsilon) + \mathbf{P}(Y_n < -\varepsilon) \\ &= F_{X_n}(x + \varepsilon) + \mathbf{P}(Y_n < -\varepsilon). \end{aligned}$$

Τότε

$$\lim_{n \rightarrow \infty} \mathbf{P}(Y_n < -\varepsilon) = 0$$

εφόσον $Y_n \xrightarrow{\mathbf{P}} 0$, και αν το $x + \varepsilon$ είναι σημείο συνέχειας της F_X ,

$$\lim_{n \rightarrow \infty} F_{X_n}(x + \varepsilon) = F_X(x + \varepsilon)$$

Άρα

$$\overline{\lim}_{n \rightarrow \infty} F_{X_n + Y_n}(x) \leq F_X(x + \varepsilon).$$

Τα σημεία ασυνέχειας της F_X είναι αριθμήσιμα, άρα υπάρχει γνήσια φθίνουσα μηδενική ακολουθία $(\varepsilon_k)_{k \in \mathbb{N}}$ έτσι ώστε το $x + \varepsilon_k$ να είναι σημείο συνέχειας της F_X . Από τα παραπάνω, για κάθε $k \in \mathbb{N}$ ισχύει

$$\overline{\lim}_{n \rightarrow \infty} F_{X_n + Y_n}(x) \leq F_X(x + \varepsilon_k).$$

Εφόσον η F_X είναι συνεχής στο x , έχουμε $\lim_{k \rightarrow \infty} F_X(x + \varepsilon_k) = F_X(x)$, και άρα

$$\overline{\lim}_{n \rightarrow \infty} F_{X_n + Y_n}(x) \leq F_X(x).$$

ΙΣΧΥΡΙΣΜΟΣ 2: $\underline{\lim}_{n \rightarrow \infty} F_{X_n + Y_n}(x) \geq F_X(x)$.

Για $\varepsilon > 0$,

$$F_{X_n + Y_n}(x) = \mathbf{P}(X_n + Y_n \leq x) \geq \mathbf{P}(X_n \leq x - \varepsilon) - \mathbf{P}(Y_n > \varepsilon).$$

Αν το $x - \varepsilon$ είναι σημείο συνέχειας της F_X , η τελευταία ανισότητα δίνει $\underline{\lim}_{n \rightarrow \infty} F_{X_n + Y_n}(x) \geq F_X(x - \varepsilon)$. Η απόδειξη του ισχυρισμού συνεχίζεται όπως προηγουμένως.

Συνδυάζοντας τους δύο ισχυρισμούς έχουμε το ζητούμενο.

Υποδείξεις για επιλεγμένες ασκήσεις

14.7 Έστω $A_n = \mathbb{R} \setminus [-n, n]$. Τότε η A_n είναι φθίνουσα ακολουθία, άρα,

$$\lim_{n \rightarrow \infty} p(A_n) = p(\bigcap_{n=1}^{\infty} A_n) = p(\emptyset) = 0.$$

Επομένως, για κάθε $\epsilon > 0$ υπάρχει n_0 έτσι ώστε $p(\mathbb{R} \setminus [-n_0, n_0]) < \epsilon$.

14.8 Από την Άσκηση 14.7, για κάθε $i \in I$ υπάρχει $M_i > 0$ έτσι ώστε $p_i([-M_i, M_i]) > 1 - \epsilon$. Θέτοντας $M = \max\{M_i : i \in I\}$ έχουμε ότι $p_i([-M, M]) \geq p_i([-M_i, M_i]) > 1 - \epsilon$, από το οποίο προκύπτει ότι η $(p_i)_{i \in I}$ είναι σφιχτή.

14.9 Χρησιμοποιώντας ότι η h είναι αύξουσα και την ανισότητα του Markov έχουμε,

$$\mathbf{P}(|X_n| > M) = \mathbf{P}(h(|X_n|) > h(M)) \leq \frac{1}{h(M)} \mathbf{E}(h(|X_n|)) \leq \frac{1}{h(M)} C,$$

όπου $C = \sup_{n \geq 1} \mathbf{E}(h(|X_n|)) < \infty$. Άρα $\sup_{n \geq 1} \mathbf{P}(|X_n| > M) \leq \frac{C}{h(M)} \rightarrow 0$ για $M \rightarrow \infty$.

14.11 Για το αντίστροφο. Αν η ακολουθία δεν είναι σφιχτή, τότε υπάρχουν $\epsilon > 0$ και γνήσια αύξουσες ακολουθίες φυσικών $(n_k)_{k \geq 1}, (M_k)_{k \geq 1}$ ώστε $\mu_{n_k}(\mathbb{R} \setminus [-M_k, M_k]) \geq \epsilon$. Η $(\mu_{n_k})_{k \geq 1}$ έχει υπακολουθία $(\mu_{n_k})_{k \geq 1}$ που συγκλίνει ασθενώς και άρα αυτή η υπακολουθία είναι σφιχτή. Προκύπτει έτσι άτοπο.

Κεφάλαιο 15

15.1. Μιμούμαστε την απόδειξη του Λήμματος 15.1. Χρήσιμη είναι η ανισότητα $e^{-y} - 1 + y \geq y/2$ για κάθε $y \geq 2$ (χρειάζεται απόδειξη).

15.3.(α) Για $t \in \mathbb{R}$ έχουμε

$$\phi_Y(t) = \mathbf{E}(e^{itX}) = \sum_{k=1}^{\infty} e^{itk} (1-p)^{k-1} p = p e^{it} \sum_{j=0}^{\infty} (e^{it} (1-p))^j = \frac{p e^{it}}{1 - (1-p)e^{it}}.$$

Αθροίσαμε μια γεωμετρική πρόοδο της οποίας ο λόγος έχει μέτρο $|(1-p)e^{it}| = 1-p < 1$ αφού $p > 0$.

(β) (i) Τα σημεία συνέχειας της συνάρτησης κατανομής, F_X , της X είναι όλο το \mathbb{R} . Έστω $x > 0$. Τότε

$$\begin{aligned} F_{X_n/n}(x) &= \mathbf{P}\left(\frac{X_n}{n} \leq x\right) = \mathbf{P}(X_n \leq [nx]) = 1 - \mathbf{P}(X_n > [nx]) \\ &= 1 - (1-p_n)^{[nx]} = 1 - \left(1 - \frac{a}{n}\right)^{[nx]/n}. \end{aligned}$$

Επειδή $[nx]/n \rightarrow x$ και $(1 - an^{-1})^n \rightarrow e^{-a}$, έπεται ότι $\lim_{n \rightarrow \infty} F_{X_n/n}(x) = 1 - e^{-ax} = F_X(x)$. Προφανώς το ίδιο ισχύει και για $x \leq 0$.

(ii) Για $t \in \mathbb{R}$ έχουμε

$$\phi_{X_n/n}(t) = \phi_{X_n}(t/n) = \frac{p_n e^{it/n}}{1 - (1-p_n)e^{it/n}} = \frac{e^{it/n}}{\frac{1-e^{it/n}}{a/n} + e^{it/n}} \rightarrow \frac{1}{1 - \frac{it}{a}} = \phi_X(t)$$

καθώς $n \rightarrow \infty$ αφού

$$\lim_{n \rightarrow \infty} \frac{1 - e^{it/n}}{a/n} = \lim_{n \rightarrow \infty} \frac{it/n - e^{it/n}}{it/n} = -\frac{it}{a}.$$

Άρα η σύγκλιση $X_n/n \Rightarrow X$ έπεται από το θεώρημα συνέχειας του Lévy.

15.4. (α) Για $t \in \mathbb{R}$ έχουμε

$$\phi_X(t) = \mathbf{E}(e^{itX}) = \sum_{k=0}^{\infty} e^{itk} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^{it}\lambda)^k}{k!} = e^{\lambda(e^{it}-1)}.$$

(β) Για $t \in \mathbb{R}$ έχουμε

$$\phi_{X_n}(t) = (p_n e^{it} + 1 - p_n)^n = (1 + p_n(e^{it} - 1))^n \rightarrow e^{\lambda(e^{it} - 1)} = \phi_X(t)$$

αφού $\lim_{n \rightarrow \infty} p_n(e^{it} - 1)n = \lambda(e^{it} - 1)$. Άρα η σύγκλιση $X_n \Rightarrow X$ έπεται από το θεώρημα συνέχειας του Lévy.

15.5. Χρησιμοποιούμε το θεώρημα συνέχειας του Lévy και τον τύπο για τη χαρακτηριστική συνάρτηση της κατανομής Γάμμα (Άσκηση 13.4). Για $t \in \mathbb{R}$,

$$\begin{aligned} \phi_{\sqrt{n}(X_n - c)} &= e^{-c\sqrt{ni}} \phi_{X_n}(\sqrt{nt}) = e^{-c\sqrt{ni}} \frac{1}{\left(1 - \frac{i\sqrt{nt}}{n}\right)^{nc}} \\ &= e^{-c\sqrt{ni} - nc \operatorname{Log}\left(1 - \frac{i}{\sqrt{n}}\right)} \end{aligned}$$

Log είναι ο κλάδος του λογαρίθμου που είναι ολόμορφος στο \mathbb{C} εκτός των αρνητικών πραγματικών και $\operatorname{Log}(1) = 0$. Χρησιμοποιώντας το ανάπτυγμα σε δυναμοσειρά της $\operatorname{Log}(1 - z)$ με κέντρο το 0 (στον δίσκο $\{z : |z| < 1\}$) βρίσκουμε ότι ο εκθέτης στην τελευταία ποσότητα συγκλίνει στο $-ct^2/2$ και έτσι έχουμε το ζητούμενο.

Εναλλακτικά, με χρήση του κεντρικού οριακού θεωρήματος (το οποίο καλύπτεται στο επόμενο κεφάλαιο). Από ιδιότητες της κατανομής Γάμμα, η X_n έχει την ίδια κατανομή με την $(W_1 + W_2 + \dots + W_n)/n$ όπου οι W_1, \dots, W_n είναι ανεξάρτητες, ισόνομες με $W_1 \sim \Gamma(c, 1)$. Το συμπέρασμα έπεται από το κεντρικό οριακό θεώρημα γιατί $\mathbf{E}(W_1) = \operatorname{Var}(W_1) = c$.

15.10. Χρήσιμη είναι η Άσκηση 13.12.

Κεφάλαιο 16

16.2. Για την ακολουθία $(S_n)_{n \geq 1}$ έχουμε ότι $S_n/n \rightarrow 2$ κατά πιθανότητα (ασθενής νόμος των μεγάλων αριθμών) και

$$\frac{S_n - 2n}{\sqrt{n}} \Rightarrow Z \sim N(0, 1)$$

(κεντρικό οριακό θεώρημα).

(α) Η σύγκλιση $S_n/n \rightarrow 2$ κατά πιθανότητα δίνει

$$\mathbf{P}(S_n > 2.1n) = \mathbf{P}\left(\frac{S_n}{n} > 2.1\right) \leq \mathbf{P}\left(\left|\frac{S_n}{n} - 2\right| > 0.1\right) \rightarrow 0$$

καθώς $n \rightarrow \infty$.

(β) Έχουμε

$$\mathbf{P}(S_n > 2n + \sqrt{n}) = \mathbf{P}\left(\frac{S_n - 2n}{\sqrt{n}} > 1\right) \rightarrow \mathbf{P}(Z > 1) = 1 - \Phi(1)$$

καθώς $n \rightarrow \infty$ λόγω του κεντρικού οριακού θεωρήματος.

(γ) Ξέρουμε ότι $S_n \sim 2n$, άρα το ενδεχόμενο $S_n > 10\sqrt{n}$ είναι πολύ πιθανό. Τυπικά προχωρούμε ως εξής.

$$\mathbf{P}(S_n > 10\sqrt{n}) = P\left(\frac{S_n}{n} > \frac{10}{\sqrt{n}}\right) = 1 - P\left(\frac{S_n}{n} \leq \frac{10}{\sqrt{n}}\right).$$

Για $n > 100$,

$$\mathbf{P}\left(\frac{S_n}{n} \leq \frac{10}{\sqrt{n}}\right) \leq \mathbf{P}\left(\left|\frac{S_n}{n} - 2\right| > 1\right) \rightarrow 0$$

Άρα $\lim_{n \rightarrow \infty} \mathbf{P}(S_n > 10\sqrt{n}) = 1$.

(δ)

$$\mathbf{P}(S_n \geq 3n) = \mathbf{P}\left(\frac{S_n}{n} \geq 3\right) \leq \mathbf{P}\left(\left|\frac{S_n}{n} - 2\right| > \frac{1}{2}\right) \rightarrow 0$$

καθώς $n \rightarrow \infty$. Άρα $\lim_{n \rightarrow \infty} \mathbf{P}(S_n < 3n) = 1$.

(ε) Για $n > 10^{10}$ έχουμε

$$\mathbf{P}(S_n \leq 10^{10}) = \mathbf{P}\left(\frac{S_n}{n} \leq \frac{10^{10}}{n}\right) \leq \mathbf{P}\left(\left|\frac{S_n}{n} - 2\right| > \frac{1}{2}\right) \rightarrow 0$$

καθώς $n \rightarrow \infty$. Άρα $\lim_{n \rightarrow \infty} \mathbf{P}(S_n > 10^{10}) = 1$.

16.3 Εφαρμόζουμε το θεώρημα Slutsky [Άσκηση 14.6].

16.5 Παρατηρούμε ότι $e^{-n} \frac{n^k}{k!} = \mathbf{P}(X = k)$, όπου $X \sim \text{Poisson}(n)$. Έστω $(X_k)_{k \in \mathbb{N}}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με $X_1 \sim \text{Poisson}(1)$. Γνωρίζουμε ότι η $S_n = \sum_{k=1}^n X_k \sim \text{Poisson}(n)$. Άρα,

$$e^{-n} \sum_{0 \leq k \leq nx} \frac{n^k}{k!} = \sum_{0 \leq k \leq nx} \mathbf{P}(S_n = k) = \mathbf{P}(S_n \leq nx)$$

Όταν $x \in [0, \infty) \setminus \{1\}$, γράφουμε αυτή την πιθανότητα ως $\mathbf{P}(S_n/n \leq x)$, και επειδή $S_n/n \rightarrow \mathbf{E}(X_1) = 1$ κατά πιθανότητα (από τον ασθενή νόμο των μεγάλων αριθμών), παίρνουμε ότι το όριο $\lim_{n \rightarrow \infty} \mathbf{P}(S_n/n \leq x)$, ισούται με 0 για $x < 1$ και με 1 για $x > 1$ [κάνουμε την τυπική δικαιολόγηση όπως στην Άσκηση 16.2].

Όταν $x = 1$, γράφουμε

$$\mathbf{P}(S_n \leq n) = \mathbf{P}\left(\frac{S_n - n}{\sqrt{n}} \leq 0\right) = F_{\frac{S_n - n}{\sqrt{n}}}(0).$$

Από το κεντρικό οριακό θεώρημα έχουμε $\frac{S_n - n}{\sqrt{n}} \Rightarrow Z$, με $Z \sim \text{N}(0, 1)$, και άρα $F_{\frac{S_n - n}{\sqrt{n}}}(0) \rightarrow F_Z(0) = \frac{1}{2}$ για $n \rightarrow \infty$.

16.6. Έστω ακολουθίες $(Y_n)_{n \geq 1}, (Z_n)_{n \geq 1}$ τυχαίες μεταβλητές στον ίδιο χώρο πιθανότητας ώστε οι $\{Y_n, Z_n : n \geq 1\}$ να είναι ανεξάρτητες και ισόνομες και καθεμία να έχει την ίδια κατανομή με τη X_1 .

Τότε επειδή το διάνυσμα (X_1, \dots, X_{2n}) έχει την ίδια κατανομή με το $(Y_1, Y_2, \dots, Y_n, Z_1, Z_2, \dots, Z_n)$ (το μέτρο γινόμενο $\otimes_{i=1}^{2n} \mathbf{P}^{X_1}$ της κατανομής \mathbf{P}^{X_1} $2n$ φορές με τον εαυτό της), έχουμε ότι

$$\begin{aligned} \mathbf{P}\left(\left|\frac{X_1 + X_2 + \dots + X_n - (X_{n+1} + \dots + X_{2n})}{\sqrt{n}}\right| \leq 1\right) \\ = \mathbf{P}\left(\left|\frac{Y_1 + Y_2 + \dots + Y_n - (Z_1 + \dots + Z_n)}{\sqrt{n}}\right| \leq 1\right) = \mathbf{P}\left(\left|\frac{W_1 + W_2 + \dots + W_n}{\sqrt{n}}\right| \leq 1\right), \end{aligned}$$

όπου θέσαμε $W_i = Y_i - Z_i$ για κάθε $i \geq 1$. Από την υπόθεση, οι $\{W_i : i \geq 1\}$ είναι ανεξάρτητες και ισόνομες, καθεμία με μέση τιμή $\mathbf{E}(W_1) = 0$ και διασπορά $V(W_1) = V(X_1) + V(Y_1) - 2\text{Cov}(X_1, Y_1) = 1 + 1 - 0 = 2$. Άρα, εφαρμόζοντας το κεντρικό οριακό θεώρημα, βρίσκουμε

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbf{P}\left(\left|\frac{W_1 + W_2 + \dots + W_n}{\sqrt{n}}\right| \leq 1\right) &= \lim_{n \rightarrow \infty} \mathbf{P}\left(\left|\frac{W_1 + W_2 + \dots + W_n}{\sqrt{n2}}\right| \leq \frac{1}{\sqrt{2}}\right) \\ &= \mathbf{P}(|Z| \leq 1/\sqrt{2}) = \Phi(1/\sqrt{2}) - \Phi(-1/\sqrt{2}) \\ &= 2\Phi(1/\sqrt{2}) - 1. \end{aligned}$$

Κεφάλαιο 17

17.1. Λύνουμε τα αντίστοιχα προβλήματα μεγιστοποίησης με χρήση παραγώγων. Οι απαντήσεις είναι ως εξής.

(α) $\Lambda^*(x) = a - x + x \log(x/a)$ για $x \geq 0$ και $\Lambda^*(x) = \infty$ για $x < 0$.

(β) $\Lambda^*(x) = ax - 1 - \log(ax)$ για $x > 0$ και $\Lambda^*(x) = \infty$ για $x \leq 0$.

(γ) $\Lambda^*(x) = x^2/(2\sigma)^2$ για κάθε $x \in \mathbb{R}$.

17.2. $M(\lambda) = \infty$ για κάθε $\lambda \neq 0$ και $M(0) = 1$. Έτσι $\Lambda^*(x) = 0$ για κάθε $x \in \mathbb{R}$. Το άνω φράγμα της αρχής για ένα μη κενό σύνολο Borel A λέει απλώς ότι το όριο $\overline{\lim}_{n \rightarrow \infty} a_n^{-1} \log \mu_n(A)$ είναι ≤ 0 , κάτι που το γνωρίζουμε από πριν αφού $\mu_n(A) \leq 1$ (άρα το άνω φράγμα είναι άχρηστο). Το κάτω φράγμα όμως λέει κάτι χρήσιμο. Δηλαδή ότι αν το A έχει μη κενό εσωτερικό τότε η πιθανότητα $\mu_n(A)$ δεν είναι εκθετικά μικρή. Π.χ., μπορεί να είναι της τάξης του n^{-100} αλλά όχι της τάξης του e^{-n} .

17.3. Θέτουμε $X_i = 1$ αν η i ρίψη του νομίσματος φέρει κεφαλή και $X_i = 0$ αν αυτή η ρίψη φέρει γράμματα. Οι $(X_i)_{i \geq 1}$ είναι ανεξάρτητες και ισόνομες, και $S_n = X_1 + X_2 + \dots + X_n$.

(α) Δουλεύουμε όπως στο Παράδειγμα 17.4. Εναλλακτικά, θεωρούμε τις τυχαίες μεταβλητές $Y_i := 2X_i - 1$ και θέτουμε $\Sigma_n := Y_1 + Y_2 + \dots + Y_n$. Η $(\Sigma_n/n)_{n \geq 1}$ ικανοποιεί την αρχή μεγάλων αποκλίσεων με συνάρτηση ρυθμού \tilde{I} όπως στο Παράδειγμα 17.4. Από αυτό προκύπτει η αρχή μεγάλων αποκλίσεων για την $(S/n)_{n \geq 1}$ με συνάρτηση ρυθμού $I(x) = \tilde{I}(2x - 1)$.

(β) Τα λήμματα δίνουν το φράγμα $e^{-1000I(7/10)} \approx 1.84073 \times 10^{-36}$ [αφού $7/10 > 1/2 = \mathbf{E}(X_1)$]. Το κεντρικό οριακό θεώρημα δίνει την προσέγγιση

$$\mathbf{P}(S_{1000} \geq 700) \approx 1 - \Phi\left(\frac{20}{\sqrt{10}}\right) \approx 1.26981 \times 10^{-10}.$$

Προφανώς αυτή είναι λάθος. Δεν είναι αρμόζουσα η εφαρμογή του κεντρικού οριακού θεωρήματος σε αυτή την περίπτωση.

17.4. Εφαρμόζουμε τον ορισμό της αρχής για το σύνολο \mathcal{X} , το οποίο έχει $\mu_n(\mathcal{X}) = 1$.

17.5. Γράφουμε $Y_n = Z/\sqrt{n}$ με $Z \sim N(0, 1)$. Έστω f_Z η πυκνότητα της f_Z . Για $A \in \mathcal{B}(\mathbb{R})$ έχουμε

$$\mathbf{P}(Y_n \in A) = \mathbf{P}(Z \in \sqrt{n}A) = \int_{\sqrt{n}A} f_Z(x) dx = \sqrt{n} \int_A f_Z(y \sqrt{n}) dy.$$

Για το κάτω φράγμα της αρχής των μεγάλων αποκλίσεων παίρνουμε $x \in A^o$. Για $\varepsilon > 0$ αρκετά μικρό, η πιο πάνω πιθανότητα φράσσεται από κάτω από την ποσότητα

$$\sqrt{n} \int_{x-\varepsilon}^{x+\varepsilon} f_Z(y \sqrt{n}) dy \geq \frac{\sqrt{n}2\varepsilon}{\sqrt{2\pi}} \min\{e^{-(x-\varepsilon)^2 n/2}, e^{-(x+\varepsilon)^2 n/2}\}.$$

Η συνέχεια αφήνεται στον αναγνώστη. Για το άνω φράγμα, θέτουμε $c := \inf\{|x| : x \in \bar{A}\}$. Τότε $A \subset (-\infty, -c] \cup [c, \infty)$ και άρα με χρήση της Άσκησης 7.4 έχουμε

$$\mathbf{P}(Y_n \in A) \leq \mathbf{P}(|Z| \geq c \sqrt{n}) \leq 2 \frac{1}{c \sqrt{n} \sqrt{2\pi}} e^{-nc^2/2}.$$

Άρα $\overline{\lim}_{n \rightarrow \infty} n^{-1} \log \mathbf{P}(Y_n \in A) \leq -c^2/2 = -\inf\{I(x) : x \in \bar{A}\}$.

17.6. Η απόδειξη είναι ανάλογη της προηγούμενης άσκησης. Μάλιστα οι εκτιμήσεις είναι πιο εύκολες.

Υποδείξεις για επιλεγμένες ασκήσεις

17.7. Για κάθε t έχουμε $f^*(x) \geq tx - f(t)$. Παίρνοντας $t > 0$ με $f(t) < \infty$ έχουμε

$$\underline{\lim}_{x \rightarrow \infty} \frac{f^*(x)}{x} \geq t,$$

ενώ αν πάρουμε $t < 0$ με $f(t) < \infty$ έχουμε

$$\underline{\lim}_{x \rightarrow -\infty} \frac{f^*(x)}{|x|} \geq |t|.$$

Τα (α), (β) έπονται εύκολα από αυτές τις ανισότητες.

17.8. Για κάθε $\lambda \neq 0$ αρκετά μικρό έχουμε

$$I(x) \geq \lambda x - \log M(\lambda) = \lambda \left(x - \frac{\log M(\lambda)}{\lambda} \right). \quad (\text{B'.26})$$

Όμως το όριο της $\log M(\lambda)/\lambda$ για $\lambda \rightarrow 0$ ισούται με την παράγωγο της $\log M(\lambda)$ στο 0, η οποία ισούται με $M'(0)/M(0) = m$ με βάση το Λήμμα 17.6. Άρα αν $x > m$, τότε παρατηρούμε ότι για $\lambda > 0$ κοντά στο 0 το δεξί μέλος της (B'.26) είναι θετικό. Αν $x < m$ παίρνουμε $\lambda < 0$ κοντά στο 0.

17.9. Για κάθε $\lambda < 0$ έχουμε

$$\mathbf{P}(S_k < tk^{1/a}) = \mathbf{P}(\lambda S_k > \lambda tk^{1/a}) \leq e^{-\lambda tk^{1/a}} M(\lambda)^k.$$

Παίρνουμε το ελάχιστο του δεξιού μέλους ως προς λ .

Βιβλιογραφία

- L. Breiman. *Probability*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, (1992).
- T. Cover and J. Thomas. *Στοιχεία της θεωρίας πληροφορίας*. Πανεπιστημιακές Εκδόσεις Κρήτης, (2014).
- A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38. Springer Science & Business Media, (1998).
- R. Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, (2010).
- K. B. Erickson (1973). The strong law of large numbers when the mean is undefined. *Transactions of the American Mathematical Society*, 185:371–381.
- W. Feller. *An introduction to probability and its applications, Vol. II*. Wiley, New York, (1971).
- J. Jacod and P. Protter. *Probability essentials*. Universitext. Springer-Verlag, Berlin, second edition, (2003).
- O. Kallenberg. *Foundations of modern probability*. Springer Science & Business Media, (2006).
- Κουμουλλής Γ. και Νεγρεπόντης Σ. *Θεωρία Μέτρου*. Εκδόσεις Συμμετρία, Αθήνα, (1991).
- Νεγρεπόντης Σ., Γιωτόπουλος Σ., Γιαννακούλιας Ε. *Απειροστικός Λογισμός*. Εκδόσεις Αίθρα, Αθήνα, (1992).
- D. Pollard. *A user's guide to measure theoretic probability*. Cambridge University Press, (2002).
- E. M. Stein and R. Shakarchi. *Real analysis: measure theory, integration, and Hilbert spaces*. Princeton University Press, (2005).
- Παπαδάτος Ν. *Θεωρία Πιθανοτήτων*. Αθήνα, (2006). Διατίθεται από τον συγγραφέα.
- S. R. Varadhan. *Probability theory, volume 7 of Courant Lecture Notes in Mathematics*. American Mathematical Society, (2001).
- E. T. Whittaker and G. N. Watson. *A course of modern analysis*. Cambridge university press, fourth edition, (1965).
- D. Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, (1991).

Ευρετήριο ελληνικών όρων

- Ανανεωτική θεωρία, 94
Ανεξάρτητες οικογένειες συνόλων, 73
Ανεξάρτητες τυχαίες μεταβλητές, 73
Απλή συνάρτηση, 22
Απολύτως συνεχείς κατανομές, 58
Αρχή μεγάλων αποκλίσεων, 136
Ασθενής νόμος των μεγάλων αριθμών, 94
Άτομο, 10
- Διακριτές κατανομές, 50
Διακριτή τυχαία μεταβλητή, 51
- Εντροπία, 95
- Θεώρημα κυριαρχημένης σύγκλισης, 38
Θεώρημα μοναδικότητας για ροπογεννήτριες, 107
Θεώρημα μοναδικότητας για χαρακτηριστικές συναρτήσεις, 102
Θεώρημα μονότονης σύγκλισης, 38
Θεώρημα π-λ, 17
Θεώρημα συνέχειας του Lévy, 122
Θεώρημα συνεχούς απεικόνισης, 114, 117
Θεώρημα φραγμένης σύγκλισης, 39
- Ιδιάζουσες κατανομές, 58
Ισόνομες τυχαίες μεταβλητές, 47
Ισχυρός νόμος των μεγάλων αριθμών, 92
- Κανονική μορφή απλής συνάρτησης, 22
Κατανομή τυχαίας μεταβλητής, 45
- λ-σύστημα, 17
- Μέτρο, 8
Μέτρο Lebesgue, 10
Μέτρο γινόμενο, 67
Μέτρο πεπερασμένο, 10
Μέτρο πιθανότητας, 10
Μεικτή κατανομή, 59
Μεροληψία με βάση το μέγεθος, 60
Μετασχηματισμός Fourier μέτρου, 102
Μετασχηματισμός Legendre, 136
Μετρήσιμη συνάρτηση, 19
Μετρήσιμο ορθογώνιο, 67, 70
- Μετρήσιμο σύνολο, 8
Μετρήσιμος χώρος, 8
- Νόμος 0-1 του Kolmogorov, 88
- Ολοκλήρωμα Lebesgue, 27
Ολοκληρώσιμη συνάρτηση, 28
Ομοιόμορφο μέτρο πιθανότητας, 12
Ουσιώδες infimum, 38
Ουσιώδες supremum, 38
- π-σύστημα, 17
Περικοπή συνάρτησης, 40
Πυκνότητα μέτρου, 48
Πυκνότητα τυχαίας μεταβλητής, 49
- Ροπογεννήτρια, 104
- Σ-άλγεβρα, 1
σ-άλγεβρα παραγόμενη από οικογένεια συνόλων, 2
σ-άλγεβρα παραγόμενη από συναρτήσεις, 24
Σ-πεπερασμένο μέτρο, 67
Στήριγμα μέτρου, 13
Στοχαστικά μικρότερη, 59
Σύγκλιση ασθενής, 111
Σύγκλιση κατά κατανομή, 111
Σύγκλιση κατά πιθανότητα, 61
Σύγκλιση με πιθανότητα 1, 61
Σύγκλιση στον \mathcal{L}^p , 61
Σύγκλιση σχεδόν βέβαιη, 61
Συμμετρική κατανομή, 110
Συνάρτηση κατανομής μέτρου πιθανότητας, 53
Συνάρτηση κατανομής τυχαίας μεταβλητής, 56
Συνάρτηση πιθανότητας, 51
Συνάρτηση ποσοστημορίων, 57
Συνάρτηση ρυθμού, 136
Συνδιακύμανση, 37
Συνεχείς κατανομές, 58
Σφιχτή οικογένεια μέτρων, 117
Σφιχτή οικογένεια τυχαίων μεταβλητών, 118
Σχεδόν βέβαια, 33
Σχεδόν παντού, 33

Τελική σ-άλγεβρα, 87
Τυπική μηχανή, 46
Τυχαία μεταβλητή, 20

Φορέας μέτρου, 13

Χαρακτηριστική συνάρτηση, 99
Χώρος γινόμενο, 67
Χώρος μέτρου, 8

Ευρετήριο ξενόγλωσσων όρων

Bepo-Levi θεώρημα, 40
Berry-Esseen θεώρημα , 131
Borel-Cantelli λήμμα, δεύτερο, 83
Borel-Cantelli λήμμα, πρώτο, 82
Borel σύνολο, 4
Borel-μετρήσιμη συνάρτηση, 19
Cauchy-Schwarz ανισότητα, 36
Chebyshev ανισότητα, 34
Cramer θεώρημα, 138
Dynkin κλάση, 16
FKG ανισότητα, 80
Fatou λήμμα , 38
Fubini θεώρημα , 69
Hölder ανισότητα, 36
Jensen ανισότητα, 35
Lebesgue-μετρήσιμα σύνολα, 10
Lebesgue-μετρήσιμη συνάρτηση, 19
Liminf, limsup ακολουθίας συνόλων, 5
Markov ανισότητα, 34
Paley-Zygmund ανισότητα, 42
Radon-Nikodym παράγωγος , 41
Scheffe λήμμα, 43
Stirling τύπος, 144
Tonelli θεώρημα, 69

Μετάφραση ορολογίας

Ανανεωτική θεωρία	Renewal theory
Ανεξάρτητες οικογένειες συνόλων	Independent families of sets
Ανεξάρτητες τυχαίες μεταβλητές	Independent random variables
Απλή συνάρτηση	Simple function
Απόλυτα συνεχής κατανομή	Absolutely continuous distribution
Αρχή μεγάλων αποκλίσεων	Large deviations principle
Ασθενής νόμος των μεγάλων αριθμών	Weak law of large numbers
Διακριτή κατανομή	Discrete distribution
Εντροπία	Entropy
Ιδιάζουσα κατανομή	Singular distribution
Ισχυρός νόμος των μεγάλων αριθμών	Strong law of large numbers
Κατανομή τυχαίας μεταβλητής	Distribution of a random variable
Μετασχηματισμός Legendre	Legendre transform
Μετρήσιμη συνάρτηση	Measurable function
Μετρήσιμο ορθογώνιο	Measurable rectangle
Μετρήσιμο σύνολο	Measurable set
Μετρήσιμος χώρος	Measurable space
Μέτρο	Measure
Μέτρο γινόμενο	Product measure
Ολοκλήρωμα Lebesgue	Lebesgue integral
Ολοκληρώσιμη συνάρτηση	Integrable function
Παράγωγος Radon-Nikodym	Radon-Nikodym derivative
Περικοπή	Truncation
Πυκνότητα μέτρου	Density of a measure
Ροπογεννήτρια	Moment generating function
Σ-άλγεβρα	Σ-algebra
Σ-πεπερασμένο	Σ-finite
Στήριγμα μέτρου	Support of a measure
Σύγκλιση ασθενώς	Weak convergence
Σύγκλιση κατά κατανομή	Convergence in distribution
Σύγκλιση κατά πιθανότητα	Convergence in probability
Σύγκλιση σχεδόν βέβαιη	Almost sure convergence
Συμμετρική κατανομή	Symmetric distribution
Συνάρτηση κατανομής	Distribution function
Συνάρτηση ποσοστημορίων	Quantile function
Συνάρτηση ρυθμού	Rate function
Συνεχής κατανομή	Continuous distribution
Σφιχτή οικογένεια τυχαίων μεταβλητών	Tight family of random variables
Σφιχτή οικογένεια μέτρων	Tight family of measures
Σχεδόν βέβαια	Almost surely
Σχεδόν παντού	Almost everywhere
Τελική σ-άλγεβρα	Final σ-algebra
Τυπική μηχανή	Standard machine
Τυχαία μεταβλητή	Random variable
Φορέας μέτρου	Support of a measure

Χαρακτηριστική συνάρτηση	Characteristic function
Χώρος γινόμενο	Product space
Χώρος μέτρου	Measure space