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 12.1). Μας ενδιαφέρει να έχουμε μια καλή εκτίμηση του πόσο σύντομα συμβαίνει αυτό. Θα δούμε ότι για πολλά σύνολα A (τα οποία θα προσδιορίσουμε) ισχύει

$$\mathbf{P}\left(\frac{S_n}{n} \in A\right) \approx e^{-c(A)n}, \tag{17.1}$$

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);

Συμβολισμος: Για $(a_n)_{n\geq 1}, (b_n)_{n\geq 1}$ ακολουθίες θετικών πραγματικών αριθμών γράφουμε $a_n \approx b_n$ αν

$$\lim_{n\to\infty} \frac{\log a_n}{\log b_n} = 1.$$

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

$$\mathbf{E}(P_n) = \mathbf{E}(2^{X_1})^n = \left(\frac{2+2^{-1}}{2}\right)^n = e^{n\log(5/4)}.$$

Δηλαδή έχει *θετικό* εκθετικό ρυθμό ανάπτυξης 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 $U_n := \{k/n : k \in \mathbb{Z}, |k| \le n\}$. 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:

$$\mathbf{E}(2^{S_n}) = \sum_{x\in U_n} e^{nx\log 2}e^{-nI(x)}. \tag{17.2}$$

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 $e^{-nI(x)}$) 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 $\mathcal{X}$ μετρικός χώρος. **Συνάρτηση ρυθμού** στον $\mathcal{X}$ ονομάζουμε οποιαδήποτε συνάρτηση $I : \mathcal{X} \to [0,\infty]$ που είναι κάτω ημισυνεχής [δηλαδή το σύνολο $[I > a]$ είναι ανοιχτό για κάθε $a \in \mathbb{R}$].

Έστω τώρα $(\mu_n)_{n\ge1}$ ακολουθία μέτρων πιθανότητας στον $(\mathcal{X}, \mathscr{B}(\mathcal{X}))$ και $(a_n)_{n\ge1}$ αύξουσα ακολουθία θετικών αριθμών με $\lim_{n\to\infty} a_n = \infty$.

Ορισμός 17.1. Λέμε ότι η ακολουθία $(\mu_n)_{n\ge1}$ ικανοποιεί την **αρχή μεγάλων αποκλίσεων** με **ταχύτητα** a_n και **συνάρτηση ρυθμού** I αν για κάθε $A \in \mathscr{B}(\mathcal{X})$ ισχύει

$$-\inf_{x\in A^\circ} I(x) \le \varliminf_{n\to\infty} \frac{1}{n}\log\mu_n(A) \le \varlimsup_{n\to\infty} \frac{1}{n}\log\mu_n(A) \le -\inf_{x\in\bar{A}} I(x). \tag{17.3}$$

Στην πράξη, συνήθως έχουμε μια ακολουθία $(Y_n)_{n\ge1}$ τυχαίων μεταβλητών στον X που συγκλίνει κατά πιθανότητα σε ένα σημείο x_0 του X και εξετάζουμε αν η ακολουθία $(\mu_n)_{n\ge1}$ των *κατανομών* των Y_n ικανοποιεί την αρχή των μεγάλων αποκλίσεων. Αν την ικανοποιεί, λέμε ότι η ακολουθία $(Y_n)_{n\ge1}$ 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 (17.3) ισοδυναμεί με την απαίτηση το άνω φράγμα να ισχύει για A κλειστό και το κάτω φράγμα να ισχύει για A ανοιχτό. Δηλαδή

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(F) \leq -\inf_{x \in F} I(x) \tag{17.4}$$

για κάθε $F \subset X$ κλειστό και

$$\underline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(G) \geq -\inf_{x \in G} I(x) \tag{17.5}$$

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

$$\underline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(G) \geq -I(x). \tag{17.6}$$

Για την απόδειξη της αρχής μεγάλων αποκλίσεων, θα χρησιμοποιούμε αυτές τις ισοδύναμες μορφές του ορισμού.

17.4 Το Θεώρημα Cramer

Για $f : \mathbb{R} \to [-\infty, \infty]$ ορίζουμε τον μετασχηματισμό Legendre της f ως τη συνάρτηση $f^* : \mathbb{R} \to [-\infty, \infty]$ 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.

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^* το μετασχηματισμό Legendre της Λ. Δηλαδή

$$M(\lambda) := \mathbf{E}(e^{\lambda X}) = \int e^{\lambda x} \, d\mu(x), \tag{17.7}$$

$$\Lambda(\lambda) := \log M(\lambda), \tag{17.8}$$

$$\Lambda^*(x) := \sup_{\lambda \in \mathbb{R}} \{\lambda x - \Lambda(\lambda)\} \tag{17.9}$$

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

$$\Lambda^*(x) = \begin{cases} \frac{1}{2}\{(1+x)\log(1+x) + (1-x)\log(1-x)\} & \text{αν } x \in [-1,1], \\ \infty & \text{αν } x \in \mathbb{R}\setminus[-1,1], \end{cases} \tag{17.10}$$

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 ένα λήμμα που ουσιαστικά αποδεικνύει το άνω φράγμα της αρχής.

Λήμμα 17.5. *Υποθέτουμε ότι* $m := \mathbf{E}(X_1) \in \mathbb{R}$. *Τότε*

(i) $x \geq m \Rightarrow \Lambda^*(x) = \sup_{\lambda\geq 0}\{\lambda x - \Lambda(\lambda)\}$,
$x \leq m \Rightarrow \Lambda^*(x) = \sup_{\lambda\leq 0}\{\lambda x - \Lambda(\lambda)\}$,
$\Lambda^*(m) = 0$.

(ii) $x \geq m \Rightarrow \mathbf{P}\left(\frac{S_n}{n} \geq x\right) \leq e^{-n\Lambda^*(x)}$,
$x \leq m \Rightarrow \mathbf{P}\left(\frac{S_n}{n} \leq x\right) \leq e^{-n\Lambda^*(x)}$.

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 για να αποδείξουμε τον πρώτο ισχυρισμό, 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.

Ανάλογη είναι η απόδειξη του δεύτερου ισχυρισμού.

Όταν $x = m$, οι δύο παραπάνω ισχυρισμοί δίνουν ότι το $\Lambda^*(m)$ ισούται με την τιμή του $\lambda x - \Lambda(\lambda)$ για $\lambda = 0$, η οποία είναι 0.

(ii) Για $x \geq m$ και $\lambda \geq 0$ έχουμε

$$\begin{aligned}\mathbf{P}\left(\frac{S_n}{n} \geq x\right) &= \mathbf{P}(S_n \geq nx) = \mathbf{P}(\lambda S_n \geq \lambda nx) = \mathbf{P}(e^{\lambda S_n} \geq e^{\lambda nx}) \leq e^{-\lambda nx}\,\mathbf{E}(e^{\lambda S_n}) \\ &= e^{-\lambda nx} M(\lambda)^n = e^{n\Lambda(\lambda) - \lambda nx} = e^{-n\{\lambda x - \Lambda(\lambda)\}}.\end{aligned}$$

Η ανισότητα στην πρώτη γραμμή προέκυψε από την ανισότητα Markov. Επειδή το φράγμα ισχύει για κάθε $\lambda > 0$, 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

$$\inf_{\lambda\geq 0} e^{-n\{\lambda x - \Lambda(\lambda)\}} = e^{-n \sup_{\lambda\geq 0}\{\lambda x - \Lambda(\lambda)\}} = e^{-n\Lambda^*(x)}.$$

Στην τελευταία ισότητα χρησιμοποιούμε το μέρος (i) του λήμματος. Ο ισχυρισμός για $x \geq m$ αποδείχθηκε.

Για $x \leq m$ και $\lambda \leq 0$ έχουμε

$$\mathbf{P}\left(\frac{S_n}{n} \leq x\right) = \mathbf{P}(S_n \leq nx) = \mathbf{P}(\lambda S_n \geq \lambda nx) = \mathbf{P}(e^{\lambda S_n} \geq e^{\lambda nx}) \leq e^{-\lambda nx}\,\mathbf{E}(e^{\lambda S_n}) = e^{-n\{\lambda x - \Lambda(\lambda)\}}.$$

Και η απόδειξη συνεχίζεται όπως και στην περίπτωση $x \geq m$. ∎

Το επόμενο λήμμα είναι κρίσιμο για την απόδειξη του κάτω φράγματος της αρχής μεγάλων αποκλίσεων.

Λήμμα 17.6. *(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. Πρέπει όμως να δείξουμε ότι αυτό είναι επιτρεπτό. Έστω λ εσωτερικό σημείο του D_M και $\delta > 0$ με $[\lambda - \delta, \lambda + \delta] \subset D_M$. Τότε για $\varepsilon \in [-\delta, \delta], \varepsilon \neq 0$ έχουμε

$$\frac{M(\lambda + \varepsilon) - M(\lambda)}{\varepsilon} = \mathbf{E}\left(\frac{e^{(\lambda+\varepsilon)X_1} - e^{\lambda X_1}}{\varepsilon}\right) = \mathbf{E}\left(e^{\lambda X_1}\frac{e^{\varepsilon X_1} - 1}{\varepsilon}\right).$$

Το όριο για $\varepsilon \to 0$ της ποσότητας στη μέση τιμή είναι το επιθυμητό $X_1 e^{\lambda X_1}$ και η απόλυτή της τιμή φράσσεται από

$$e^{\lambda X_1}\frac{e^{\delta |X_1|} - 1}{\delta} \leq \delta^{-1}\{e^{(\lambda-\delta)X_1} + e^{(\lambda+\delta)X_1}\}.$$

Για να δούμε το πρώτο φράγμα, αναπτύσσουμε σε δυναμοσειρά την $e^{\varepsilon X_1}$. Το δεξί μέλος της τελευταίας ανισότητας δεν εξαρτάται από το ε και έχει πεπερασμένη μέση τιμή εξαιτίας του ότι $\lambda - \delta, \lambda + \delta \in D_M$. Το συμπέρασμα έπεται από το θεώρημα κυριαρχημένης σύγκλισης.

(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 17.7 (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 $I(x) := \Lambda^*(x)$.*

Απόδειξη. Έστω μ_n η κατανομή της τυχαίας μεταβλητής S_n/n. Ακολουθούμε τη μέθοδο της Παρατήρησης 17.3(γ).

Άνω φράγμα: Έστω $F \subset \mathbb{R}$ κλειστό μη κενό. Αν $I(F) = 0$, δεν έχουμε να αποδείξουμε τίποτα γιατί το αριστερό μέλος της (17.4) είναι μη θετικό πάντοτε.

Υποθέτουμε λοιπόν ότι $I(F) > 0$. Επειδή $I(m) = 0$ (Λήμμα 17.5), έπεται ότι το m είναι στοιχείο του ανοιχτού συνόλου $\mathbb{R}\backslash F$. Έστω (a, b) το μέγιστο υποδιάστημα του $\mathbb{R}\backslash F$ που περιέχει το m. Αυτό το υποδιάστημα είναι ανοιχτό (και άρα $a, b \in F$) γιατί το $\mathbb{R}\backslash F$ είναι ανοιχτό και ενδέχεται $a = -\infty$ ή $b = \infty$ (όχι όμως και τα δύο γιατί $F \neq \emptyset$). Επειδή $F \subset \mathbb{R}\backslash(a, b)$, όταν $a, b \in \mathbb{R}$, έχουμε

$$\mu_n(F) \leq \mu_n((-\infty, a]) + \mu_n([b, \infty)) \leq e^{-n\Lambda^*(a)} + e^{-n\Lambda^*(b)} \leq 2e^{-nI(F)}. \tag{17.11}$$

Η πρώτη ανισότητα έπεται από το Λήμμα 17.5, 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 (17.11).

Κάτω φράγμα: Με βάση την (17.6), επειδή η S_n/n παίρνει τιμές στο $\mathbb{R}$, αρκεί να δείξουμε ότι για κάθε $a \in \mathbb{R}$ και $\delta > 0$ ισχύει

$$\varliminf_{n\to\infty} \frac{1}{n} \log \mu_n((a - \delta, a + \delta)) \geq -\Lambda^*(a). \tag{17.12}$$

Περίπτωση 1. $\mu((-\infty, a)), \mu((a, \infty)) > 0$ και το μ έχει συμπαγή φορέα.

Τότε με βάση το Λήμμα 17.6 υπάρχει $\lambda_0 \in \mathbb{R}$ ώστε $\Lambda^*(a) = \lambda_0 a - \Lambda(\lambda_0)$. Ορίζουμε ένα νέο μέτρο $\tilde{\mu}$ από τη σχέση

$$\frac{d\tilde{\mu}}{d\mu}(x) = e^{\lambda_0 x - \Lambda(\lambda_0)}, x \in \mathbb{R}. \tag{17.13}$$

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 $\tilde{X}_1 + \cdots + \tilde{X}_n/n$ όταν οι $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_n$ 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

$$\begin{aligned}
\mu_n((a-\varepsilon, a+\varepsilon)) &= \mathbf{P}\left(\frac{S_n}{n} \in (a-\varepsilon, a+\varepsilon)\right) = \int_{|x_1+x_2+\cdots+x_n - na| < n\varepsilon} d\mu(x_1)\cdots d\mu(x_n) \\
&= \int_{|x_1+x_2+\cdots+x_n - na| < n\varepsilon} e^{n\Lambda(\lambda_0) - \lambda_0(x_1+\cdots+x_n)} d\tilde{\mu}(x_1)\cdots \tilde{d}\mu(x_n) \\
&\geq e^{n\Lambda(\lambda_0) - \lambda_0 na - |\lambda_0| n\varepsilon} \tilde{\mu}_n((a-\varepsilon, a+\varepsilon)) = e^{-n\Lambda^*(a) - n|\lambda_0|\varepsilon} \tilde{\mu}_n((a-\varepsilon, a+\varepsilon)).
\end{aligned}$$

Άρα

$$\varliminf_{n\to\infty} \frac{1}{n} \log \mu_n((a-\delta, a+\delta)) \geq -\Lambda^*(a) - |\lambda_0|\varepsilon - \varliminf_{n\to\infty} \frac{1}{n} \log \tilde{\mu}_n((a-\delta, a+\delta)). \tag{17.14}$$

Τώρα $\lim_{n\to\infty} \tilde{\mu}_n((a-\delta, a+\delta)) = 1$ από τον ασθενή νόμο των μεγάλων αριθμών γιατί

$$\tilde{\mu}_n((a-\delta, a+\delta)) = \mathbf{P}\left(\frac{\tilde{X}_1 + \cdots + \tilde{X}_n}{n} \in (a-\delta, a+\delta)\right)$$

και οι $\tilde{X}_1, \ldots, \tilde{X}_n$ είναι ανεξάρτητες ισόνομες με μέση τιμή a. Άρα το $\varliminf$ στο δεξί μέλος της ανισότητας (17.14) είναι 0 και παίρνοντας $\varepsilon \to 0$ έχουμε την (17.6).

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 $(\hat{X}_i)_{i\geq 1}$ ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με κατανομή αυτήν της X_1 με τη δέσμευση $|X_1| \leq R$. Δηλαδή

$$\mathbf{P}(\hat{X}_1 \in A) = \frac{\mathbf{P}(X_1 \in A, |X_1| \leq R)}{\mathbf{P}(|X_1| \leq R)}$$

για κάθε $A \in \mathscr{B}(\mathbb{R})$.
Θέτουμε $\hat{S}_n = \hat{X}_1 + \hat{X}_2 + \cdots + \hat{X}_n$. Τότε

$$\mathbf{P}\left(\frac{S_n}{n} \in (a-\varepsilon, a+\varepsilon)\right) \geq \mathbf{P}\left(\frac{S_n}{n} \in (a-\varepsilon, a+\varepsilon), |X_i| \leq R \text{ για κάθε } i = 1, 2, \ldots, n\right) \tag{17.15}$$

$$= \frac{\mathbf{P}\left(\frac{S_n}{n} \in (a-\varepsilon, a+\varepsilon), |X_i| \leq R \text{ για κάθε } i = 1, 2, \ldots, n\right)}{\mathbf{P}(|X_i| \leq R \text{ για κάθε } i = 1, 2, \ldots, n)} \tag{17.16}$$

$$\times\, \mathbf{P}(|X_i| \leq R \text{ για κάθε } i = 1, 2, \ldots, n) \tag{17.17}$$

$$= \mathbf{P}\left(\frac{\hat{S}_n}{n} \in (a-\varepsilon, a+\varepsilon)\right) \mathbf{P}(|X_1| \leq R)^n \tag{17.18}$$

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

$$\varliminf_{n\to\infty} \frac{1}{n} \log \mu_n((a-\delta, a+\delta)) \geq -I_R(a) + \log \mathbf{P}(|X_1| \leq R).$$

Αρχικά, θα βελτιώσουμε την έκφραση του κάτω φράγματος. Θέτουμε $C_R(\lambda) = \log \mathbf{E}(e^{\lambda X_1} \mathbf{1}_{|X_1| \le R})$. Η ροπογεννήτρια της $\hat{X}_1$ είναι $\mathbf{E}(e^{\lambda X_1} \mathbf{1}_{|X_1| \le R}) / \mathbf{P}(|X_1| \le R)$, με λογάριθμο $C_R(\lambda) - \log \mathbf{P}(|X_1| \le R)$, άρα $I_R(a) = \sup_{\lambda \in \mathbb{R}} \{\lambda a - C_R(\lambda)\} + \log \mathbf{P}(|X_1| \le R)$. 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.

Ισχυρισμός: $\underline{\lim}_{R \to \infty} C_R^*(a) \le \Lambda^*(a)$.

Η $C_R^*(a)$ είναι φθίνουσα ως προς R γιατί η $C_R(\lambda)$ είναι αύξουσα ως προς R. Άρα $\underline{\lim}_{R \to \infty} C_R^*(a) \le C_r^*(a)$ για κάθε $r > 0$. Έστω $u < \underline{\lim}_{R \to \infty} C_R^*(a)$. Θέτουμε

$$K_r := \{\lambda \in \mathbb{R} : \lambda a - C_r(\lambda) \ge u\}.$$

Για $r \ge R_0$, 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 17.6(β)). Επίσης η $(K_r)_{r \ge R_0}$ είναι φθίνουσα ως προς r, άρα η τομή $\cap_{r \ge R_0} K_r$ είναι μη κενή και έστω λ_0 ένα σημείο σε αυτήν. Τότε

$$\lambda_0 a - C_r(\lambda_0) \ge u$$

για κάθε $r \ge R_0$. Για $r \to \infty$ η τελευταία ανισότητα και το θεώρημα μονότονης σύγκλισης δίνουν $\lambda_0 a - \log \Lambda(\lambda_0) \ge u$, και άρα $\Lambda^*(a) \ge u$. 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 όριό της στο $-\infty$ όταν $\mu((-\infty, a)) = 0$ και με το όριο της στο ∞ όταν $\mu((a, -\infty)) = 0$. Τώρα το συμπέρασμα έπεται γιατί

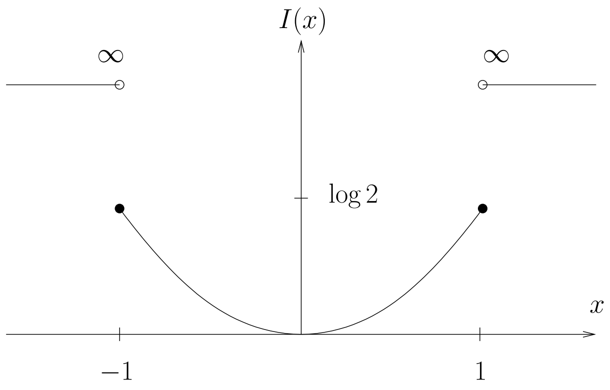
$$\mu_n((a - \delta, a + \delta)) \ge \mathbf{P}(X_1 = X_2 = \cdots = X_n = a) = \mathbf{P}(X_1 = a)^n.$$

■

Παρατήρηση 17.8. (α) (Η ιδέα της αλλαγής μέτρου) Το ουσιαστικό κομμάτι της απόδειξης του κάτω φράγματος είναι η Περίπτωση 1. Ας πάρουμε την περίπτωση $a \ne m$ 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 το νόμο των X_i με τέτοιο τρόπο ώστε το A_n να γίνει τυπικό για αυτό το 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.

(γ) Το Θεώρημα Cramer ισχύει ακόμα και χωρίς την υπόθεση ότι η $\mathbf{E}(X_1)$ ορίζεται και είναι πραγματικός αριθμός. Αυτό αποδεικνύεται με λίγες παρεμβάσεις στην απόδειξη πιο πάνω (δες Dembo and Zeitouni (1998), Θεώρημα 2.2.3).



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 (17.10). 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 $\{S_n/n$ είναι κοντά στο $x\}$ γίνεται ακριβότερο/πιο απίθανο.

Τώρα μπορούμε να επιστρέψουμε στην Παράγραφο 17.2 και να δούμε ότι πράγματι το x που μεγιστοποιεί τη διαφορά $x \log 2 - I(x)$ είναι το $x = 3/5$.

Ασκήσεις

17.1 Να υπολογιστεί ο μετασχηματισμός Λ^* στην περίπτωση που η X_1 ακολουθεί την κατανομή

(α) 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 Έστω $f : \mathbb{R} \to [-\infty, \infty]$ και $D_f := \{x \in \mathbb{R} : f(x) < \infty\}$.

(α) Αν $0 \in D_f^\circ$, τότε

$$\varliminf_{|x|\to\infty} \frac{f^*(x)}{|x|} > 0,$$

και άρα $\lim_{|x|\to\infty} f^*(x) = \infty$.

(β) Αν $D_f = \mathbb{R}$, τότε

$$\lim_{|x|\to\infty} \frac{f^*(x)}{|x|} = \infty.$$

Παραρτήματα

Α′

Αναλυτικά αποτελέσματα

Κύριος στόχος αυτού του παραρτήματος είναι η διατύπωση του Λήμματος Α′.2 παρακάτω που αφορά υπολογισμό ορίων απειρογινομένων και χρησιμοποιείται σε αποδείξεις ασθενούς σύγκλισης μέσω χαρακτηριστικών συναρτήσεων.

Για $z \in \mathbb{C} \backslash (-\infty, 0]$, 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

$$\begin{aligned} e^z &\simeq 1 + z, \\ \log(1+z) &\simeq z. \end{aligned}$$

Αυτές οι δύο προσεγγίσεις μας καθοδηγούν όταν κάνουμε ασυμπτωτική ανάλυση απειρογινομένων. Και έπειτα, για να δικαιολογήσουμε αυστηρά το αποτέλεσμα που μας υποδεικνύουν, χρησιμοποιούμε κάποια από τις ανισότητες στο παρακάτω λήμμα.

Λήμμα Α′.1. *(i)* $e^x \ge 1 + x$ *για κάθε* $x \in \mathbb{R}$.

(ii) $\log x \le x - 1$ *για κάθε* $x \in (0, \infty)$.

(iii) $|e^z - (1+z)| \le |z|^2$ *για κάθε* $z \in \mathbb{C}$ *με* $|z| \le 1$.

(iv) $|\log(1+z) - z| \le 2|z|^2$ *για κάθε* $z \in \mathbb{C}$ *με* $|z| \le 1/2$.

Απόδειξη. Τα (i), (ii) είναι γνωστά από το λύκειο.
(iii) Με χρήση της δυναμοσειράς για την e^z έχουμε

$$|e^z - (1+z)| = \left| \sum_{k=2}^{\infty} \frac{z^k}{k!} \right| \le |z|^2 \sum_{k=2}^{\infty} \frac{1}{k!} \le |z|^2 \sum_{k=2}^{\infty} \frac{1}{2^k} = |z|^2$$

Για την πρώτη ανισότητα, βγάλαμε κοινό παράγοντα το $|z|^2$ και χρησιμοποιήσαμε το ότι $|z| \le 1$.
(iv) Για $z \in \mathbb{C}$ με $|z| \le 1/2$ ισχύει

$$|\log(1+z) - z| = \left| \sum_{k=2}^{\infty} \frac{(-1)^{k-1}}{k} z^k \right| \le \sum_{k=2}^{\infty} |z|^k = |z|^2 \frac{1}{1-|z|} \le 2|z|^2.$$

Η πρώτη ισότητα ισχύει για κάθε z με $|z| < 1$, ενώ η τελευταία ανισότητα για $|z| \le 1/2$. ■

Λήμμα Α′.2. *Έστω* $(k_n)_{n\ge1}$ *ακολουθία φυσικών αριθμών και* $\{a_{n,j} : n \ge 1, 1 \le j \le k_n\}$ *μιγαδικοί αριθμοί. Υποθέτουμε ότι*

(i) $\lim_{n\to\infty} \sum_{j=1}^{k_n} a_{n,j} = A$ *με* $A \in \mathbb{C}$ *και*

(ii) $\lim_{n\to\infty} \sum_{j=1}^{k_n} |a_{n,j}|^2 = 0$.

Τότε

$$\lim_{n\to\infty}\prod_{j=1}^{k_n}(1+a_{n,j})=e^A.$$

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 $|a_{n,j}|\le 1/2$ για κάθε $n\ge n_0$ $1\le j\le k_n$. Έπειτα, έχουμε

$$\frac{\prod_{j=1}^{k_n}(1+a_{n,j})}{e^{\sum_{j=1}^{k_n}a_{n,j}}}=e^{\sum_{j=1}^{k_n}\{\log(1+a_{n,j})-a_{n,j}\}}\tag{Α$'$.1}$$

Ο εκθέτης με χρήση του Λήμματος Α$'$.1(iv) φράσσεται ως εξής

$$\left|\sum_{j=1}^{k_n}\{\log(1+a_{n,j})-a_{n,j}\}\right|\le\sum_{j=1}^{k_n}|\log(1+a_{n,j})-a_{n,j}|\le 2\sum_{j=1}^{k_n}|a_{n,j}|^2.$$

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 $c_n\to c$, με $c\in\mathbb{C}$, τότε*

$$\lim_{n\to\infty}\left(1+\frac{c_n}{n}\right)^n=e^c.$$

Απόδειξη. Εφαρμόζουμε το Λήμμα Α$'$.2 με $k_n=n, a_{n,j}=c_n/n$ για κάθε $n\ge 1$ και $1\le j\le n$. Έχουμε

$$\sum_{j=1}^{n}\frac{c_n}{n}=c_n\to c\quad\text{και}\quad\sum_{j=1}^{n}\left(\frac{c_n}{n}\right)^2=\frac{c_n^2}{n}\to 0$$

∎

Τέλος, 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*

$$\Gamma(x)=\left(\frac{x}{e}\right)^x\frac{\sqrt{2\pi}}{\sqrt{x}}\,e^{\theta_x/(12x)}.\tag{Α$'$.2}$$

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 $n \to \infty$.

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 10.11: Επειδή η οικογένεια $(\mathcal{G}_j)_{j\in J}$ είναι ανεξάρτητη αν και μόνο αν κάθε πεπερασμένη υποοιογένεια της είναι ανεξάρτητη, αρκεί να δείξουμε το θεώρημα στην περίπτωση που το J είναι πεπερασμένο. Υποθέτουμε λοιπόν ότι $J = \{1, 2, \ldots, n\}$ για κάποιο $n \geq 2$. Για κάθε $k \in J$, ονομάζουμε $\mathcal{C}_k$ το σύνολο των συνόλων της μορφής $A_{i_1} \cap A_{i_2} \cdots \cap A_{i_r}$ όπου $r \geq 1$ και $A_{i_1}, A_{i_2}, \ldots, A_{i_r} \in \cup_{i\in I_k}\mathcal{F}_i$. Παρατηρούμε ότι $\sigma(\mathcal{C}_k) = \mathcal{G}_k$. Θέτουμε

$$\mathcal{D}_1 := \{A \in \mathcal{G}_1 : \mathbf{P}(A \cap A_2 \cap \cdots \cap A_n) = \mathbf{P}(A)\,\mathbf{P}(A_2)\cdots\mathbf{P}(A_n) \text{ για κάθε } A_2 \in \mathcal{C}_2, \ldots, A_n \in \mathcal{C}_n\}$$

Από υπόθεση, $\mathcal{C}_1 \subset \mathcal{D}_1$. Εύκολα δείχνουμε ότι η $\mathcal{D}_1$ είναι κλάση Dynkin, άρα $\delta(\mathcal{C}_1) \subset \mathcal{D}_1$. Όμως η $\mathcal{C}_1$ είναι κλειστή στις πεπερασμένες τομές, οπότε το θεώρημα μονότονης κλάσης δίνει ότι $\delta(\mathcal{C}_1) = \sigma(\mathcal{C}_1) = \mathcal{G}_1$. Άρα

$$\text{τα } \mathcal{G}_1, \mathcal{C}_2, \mathcal{C}_3, \ldots, \mathcal{C}_n \text{ είναι ανεξάρτητα.}$$

Με ανάλογο επιχείρημα δείχνουμε ότι

$$\text{τα } \mathcal{G}_1, \mathcal{G}_2, \mathcal{C}_3, \ldots, \mathcal{C}_n \text{ είναι ανεξάρτητα,}$$

και τελικά το θεώρημα. ■

Απόδειξη του Θεωρήματος 10.12: 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 $(\mathcal{G}_j)_{i\in J}$ είναι ανεξάρτητες. Αρκεί επομένως να δείξουμε ότι για κάθε $j \in J$, η Y_j είναι $\mathcal{G}_j$-μετρήσιμη. Έστω $W_j := (X_i)_{i\in I_j} : \Omega \to \mathbb{R}^{I_j}$, 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 τους μετρήσιμους κυλίνδρους γιατί αν πάρουμε έναν 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 τους μετρήσιμους κυλίνδρους, έπεται ότι $\otimes_{i\in I_j}\mathscr{B}(\mathbb{R}) \subset \mathcal{B}_j$. Ο ισχυρισμός αποδείχθηκε. ■

Κεφάλαιο 11

Απόδειξη του Θεωρήματος 11.10: 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 $\mathscr{D}_n = \sigma(\{X_k : k \leq n\})$, $\mathscr{C}_n$ είναι ανεξάρτητες.

Αυτό έπεται από το ότι οι $(X_n)_{n\geq 1}$ είναι ανεξάρτητες, τους Ορισμούς 10.3, 5.17, και το Θεώρημα 10.11 για τη διαμέριση $\{\{1, 2, \dots, n\}, \{n+1, n+2, \dots\}\}$ του $\mathbb{N}^+$.

Θέτουμε τώρα $\mathscr{D} = \cup_{n\in\mathbb{N}}\mathscr{D}_n$.

Ισχυρισμος 2: Το C είναι ανεξάρτητο από κάθε στοιχείο της $\sigma(\mathscr{D})$.

Επειδή το C είναι στοιχείο της $\mathscr{C}_n$ για κάθε $n \geq 1$, έπεται ότι το C είναι ανεξάρτητο από κάθε $\mathscr{D}_n$ και άρα από κάθε στοιχείο της ένωσής τους, που είναι το $\mathscr{D}$. Το σύνολο $\mathcal{E}$ των στοιχείων της $\sigma(\mathscr{D})$ που είναι ανεξάρτητα από το C είναι μια κλάση Dynkin (Άσκηση 3.1) που περιέχει την $\mathscr{D}$ και η $\mathscr{D}$ είναι κλειστή στις πεπερασμένες τομές. Άρα, από το θεώρημα π-λ, $\sigma(\mathscr{D}) = \delta(\mathscr{D})$. Όμως $\delta(\mathscr{D}) \subset \mathcal{E} \subset \sigma(\mathscr{D})$, οπότε $\mathcal{E} = \sigma(\mathscr{D})$.

Τώρα $\mathscr{C}_\infty \subset \sigma(\mathscr{D})$ γιατί εύκολα βλέπουμε ότι $\sigma(\mathscr{D}) = \sigma(\{X_n : n \geq 1\})$. Άρα από τον Ισχυρισμό 2 έχουμε ότι το C είναι ανεξάρτητο από τον εαυτό του. ∎

Κεφάλαιο 14

Απόδειξη του Θεωρήματος 14.8: Έστω F και F_n η συνάρτηση κατανομής των μέτρων μ, μ_n αντίστοιχα. «⇒» Αν τα a, b είναι σημεία συνέχειας της F, τότε για την $f := \mathbf{1}_{(a,b]}$ ισχύει η (14.3). Πράγματι

$$\int \mathbf{1}_{(a,b]}(x)\, d\mu_n(x) = \mu_n((a,b]) = F_n(b) - F_n(a) \to F(b) - F(a) = \int \mathbf{1}_{(a,b]}(x)\, d\mu(x) \tag{B$'$.1}$$

για $n \to \infty$.

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

$$x, y \in [-K, K], |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon_1. \tag{B$'$.2}$$

Βρίσκουμε στο $[-K, K]$ σημεία $-K = a_0 < a_1 < a_2 < \cdots < a_{N-1} < a_N = K$ 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

- $\lim_{n\to\infty} \int s(x)\, d\mu_n(x) = \int s(x)\, d\mu(x)$ λόγω της (B$'$.1) 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

$$\left| \int f(x)\, d\mu_n(x) - \int s(x) d\mu_n(x) \right| \leq \int |f(x) - s(x)|\, d\mu_n(x) \tag{B$'$.3}$$

$$\leq \int |f(x) - s(x)| \mathbf{1}_{x\leq -K}\, d\mu_n(x) + \int |f(x) - s(x)| \mathbf{1}_{x>K}\, d\mu_n(x) + \int |f(x) - s(x)| \mathbf{1}_{x\in(-K,K]}\, d\mu_n(x) \tag{B$'$.4}$$

$$\leq \|f\|_\infty \{\mu_n((-\infty, -K]) + \mu_n((K, \infty))\} + \varepsilon_1 \mu_n((-K, K]) \tag{B$'$.5}$$

$$\leq \|f\|_\infty \{F_n(-K) + 1 - F_n(K)\} + \varepsilon_1 \tag{B$'$.6}$$

Άρα $\overline{\lim}_{n\to\infty} \left| \int f(x)\, d\mu_n(x) - \int s(x) d\mu_n(x) \right| \leq 2\varepsilon \|f\|_\infty + \varepsilon_1$.

- Όμοια, $\left| \int f(x)\, d\mu(x) - \int s(x) d\mu(x) \right| \leq 2\varepsilon \|f\|_\infty + 2\varepsilon_1$.

Άρα, από την τριγωνική ανισότητα,

$$\overline{\lim_{n\to\infty}}\left|\int f(x)\,d\mu_n(x) - \int f(x)\,d\mu(x)\right| \le 4\varepsilon\|f\|_\infty + \varepsilon_1.$$

Το αριστερό μέλος δεν εξαρτάται από τα $\varepsilon, \varepsilon_1$. Θεωρούμε λοιπόν $\varepsilon, \varepsilon_1 \to 0^+$ και το ζητούμενο έπεται.

«$\Leftarrow$» Έστω $x_0 \in \mathbb{R}$ σημείο συνέχειας της F. Για $\varepsilon > 0$, θεωρούμε τη συνεχή και φραγμένη συνάρτηση

$$f(x) = \begin{cases} 1 & \text{αν } x \le x_0, \\ -(x - x_0 - \varepsilon)/\varepsilon & \text{αν } x \in (x_0, x_0 + \varepsilon], \\ 0 & \text{αν } x \ge x_0 + \varepsilon. \end{cases} \tag{Β΄.7}$$

η οποία ικανοποιεί $\mathbf{1}_{(-\infty,x_0]}(x) \le f(x) \le \mathbf{1}_{(-\infty,x_0+\varepsilon]}(x)$. Παίρνοντας ολοκλήρωμα ως προς μ_n στην πρώτη ανισότητα, ως προς μ στη δεύτερη ανισότητα, και χρησιμοποιώντας την υπόθεση παίρνουμε

$$\overline{\lim_{n\to\infty}}\, F_n(x_0) \le \overline{\lim_{n\to\infty}} \int f(x)\,d\mu_n(x) = \int f(x)\,d\mu(x) \le F(x_0 + \varepsilon).$$

Όμως το ε είναι αυθαίρετο. Και επειδή η F είναι δεξιά συνεχής στο x_0, για $\varepsilon \to 0$ παίρνουμε

$$\overline{\lim_{n\to\infty}}\, F_n(x_0) \le F(x_0). \tag{Β΄.8}$$

Για το κάτω φράγμα, παίρνουμε $\varepsilon > 0$ και θεωρούμε τη συνάρτηση

$$g(x) = \begin{cases} 1 & \text{αν } x \le x_0 - \varepsilon, \\ -(x - x_0)/\varepsilon & \text{αν } x \in (x_0 - \varepsilon, x_0], \\ 0 & \text{αν } x \ge x_0, \end{cases} \tag{Β΄.9}$$

η οποία ικανοποιεί $\mathbf{1}_{(-\infty,x_0-\varepsilon]}(x) \le g(x) \le \mathbf{1}_{(-\infty,x_0]}(x)$. Όπως πριν παίρνουμε

$$\underline{\lim_{n\to\infty}}\, F_n(x_0) \ge \underline{\lim_{n\to\infty}} \int g(x)\,d\mu_n(x) = \int g(x)\,d\mu(x) \ge F(x_0 - \varepsilon).$$

Επειδή η F είναι αριστερά συνεχής στο x_0 (εδώ μόνο χρησιμοποιούμε ότι το x_0 είναι σημείο συνέχειας της F), για $\varepsilon \to 0$ παίρνουμε

$$\underline{\lim_{n\to\infty}}\, F_n(x_0) \ge F(x_0). \tag{Β΄.10}$$

Η τελευταία σχέση μαζί με την (Β΄.8) δίνουν το ζητούμενο. ∎

Απόδειξη του Θεωρήματος 14.11 (i) ⇒ (ii): Ισχυρισμος: Αν το A είναι κλειστό, τότε

$$\overline{\lim_{n\to\infty}}\, \mathbf{P}(X_n \in A) \le \mathbf{P}(X \in A).$$

Σταθεροποιούμε $c > 0$ και θεωρούμε τη συνάρτηση $f_c(x) := 1/(1 + d(x, A))^c$, όπου $d(x, A) := \inf\{|x - y| : y \in A\}$ είναι η απόσταση του x από το A. Η f_c είναι συνεχής, φραγμένη, και ικανοποιεί $\mathbf{1}_A \le f_c$. Άρα

$$\overline{\lim_{n\to\infty}}\, \mathbf{P}(X_n \in A) \le \overline{\lim_{n\to\infty}}\, \mathbf{E}\, f_c(X_n) = \mathbf{E}\, f_c(X). \tag{Β΄.11}$$

Τώρα, $\lim_{c\to\infty} f_c(x) = \mathbf{1}_A(x)$ για κάθε x επειδή κάθε $x \in \mathbb{R}\setminus A$ έχει $d(x, A) > 0$ (το A είναι κλειστό). Οπότε, παίρνοντας $c \to \infty$ στην (Β΄.11) κατά μήκος μιας ακολουθίας (π.χ. $c = k$ φυσικός) το θεώρημα φραγμένης σύγκλισης δίνει ότι $\lim_{k\to\infty} \mathbf{E}\, f_k(X) = \mathbf{E}\, \mathbf{1}_A(X) = \mathbf{P}(X \in A)$. Και ο ισχυρισμός αποδείχθηκε.

Αν το A είναι ανοιχτό, τότε εφαρμόζοντας τον ισχυρισμό για το κλειστό $\mathbb{R}\setminus A$ παίρνουμε

$$\underline{\lim_{n\to\infty}}\, \mathbf{P}(X_n \in A) \ge \mathbf{P}(X \in A).$$

Τώρα για ένα A όπως στην εκφώνηση έχουμε $\mathbf{P}(X \in \bar{A}) = \mathbf{P}(X \in A^\circ) + \mathbf{P}(X \in \partial A) = \mathbf{P}(X \in A^\circ)$. Και από τα πιο πάνω

$$\mathbf{P}(X \in A^\circ) \le \varliminf_{n\to\infty} \mathbf{P}(X_n \in A^\circ) \le \varliminf_{n\to\infty} \mathbf{P}(X_n \in A) \le \varlimsup_{n\to\infty} \mathbf{P}(X_n \in A) \le \varlimsup_{n\to\infty} \mathbf{P}(X_n \in \bar{A}) \le \mathbf{P}(X \in \bar{A}).$$

Το ζητούμενο έπεται. ■

Κεφάλαιο 15

Απόδειξη του Λήμματος 15.1: Από τον ορισμό του μετασχηματισμού Fourier του μέτρου μ έχουμε

$$\int_{-u}^{u} (1 - \hat{\mu}(t))\,\mathrm{d}t = \int_{-u}^{u} \int (1 - e^{itx})\,\mathrm{d}\mu(x)\,\mathrm{d}t = \int \int_{-u}^{u} (1 - \cos(tx) + i\sin(tx))\,\mathrm{d}t\,\mathrm{d}\mu(x).$$

Η δεύτερη ισότητα προκύπτει από το Θεώρημα 9.4 (Tonelli-Fubini). Εφόσον η συνάρτηση $1 - \cos(tx)$ είναι άρτια και η συνάρτηση $\sin(tx)$ είναι περιττή, το τελευταίο ολοκλήρωμα ισούται με

$$2 \int \int_{-u}^{u} (1 - \cos(tx))\,\mathrm{d}t\,\mathrm{d}\mu(x) = 2 \int \left(u - \frac{\sin(ux)}{x} \right) \mathrm{d}\mu(x) = 2u \int \left(1 - \frac{\sin(ux)}{ux} \right) \mathrm{d}\mu(x).$$

Παρατηρούμε τώρα ότι η συνάρτηση στο τελευταίο ολοκλήρωμα είναι μη αρνητική ($1 - \frac{\sin(ux)}{ux} \ge 0$ για κάθε $x \in \mathbb{R}$. Άρα, αν ολοκληρώσουμε σε μικρότερο χωρίο, το ολοκλήρωμα μικραίνει) και για $|ux| > 2$ έχουμε

$$\left| \frac{\sin(ux)}{ux} \right| \le \frac{1}{ux} \le \frac{1}{2}.$$

Συνεπώς,

$$\int_{-u}^{u} \{1 - \hat{\mu}(t)\}\,\mathrm{d}t \ge 2u \int_{\{x:|x|>2/u\}} \frac{1}{2}\,\mathrm{d}\mu(x) = u\mu\left(\left\{ x : |x| > \frac{2}{u} \right\} \right), \tag{B'.12}$$

που είναι το ζητούμενο. ■

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 $B_1 = A_1, B_2 = A_2 \backslash A_1, B_3 = A_3 \backslash (A_1 \cup A_2)$, ...

1.7 (α) $\emptyset \in \mathcal{B}$ γιατί $f^{-1}(\emptyset) = \emptyset$ και $\emptyset \in \mathcal{F}$. Έπειτα, αν $A \in \mathcal{B}$ τότε $f^{-1}(Y \backslash A) = X \backslash f^{-1}(A)$ και εφόσον $f^{-1}(A) \in \mathcal{A}$ και η $\mathcal{A}$ είναι σ-άλγεβρα, έχουμε ότι $X \backslash f^{-1}(A) \in \mathcal{F}$. Άρα $Y \backslash A \in \mathcal{B}$. Τέλος, αν $(A_n)_{n\in\mathbb{N}}$ είναι ακολουθία στη $\mathcal{B}$, τότε

$$f^{-1}\left(\cup_{n\in\mathbb{N}} A_n\right) = \cup_{n\in\mathbb{N}} f^{-1}(A_n) \in \mathcal{A}.$$

1.8 Θα αποδείξουμε ότι η $\mathcal{A}$ είναι σ-άλγεβρα στην περίπτωση που ισχύει το 1. Με παρόμοιο τρόπο αποδεικνύονται και οι άλλες δυο περιπτώσεις. Αφού η $\mathcal{A}$ είναι άλγεβρα, αρκεί να δείξουμε ότι είναι κλειστή στις αριθμήσιμες ενώσεις. Έστω $\{B_n\}_{n\in\mathbb{N}}$ ακολουθία στην $\mathcal{A}$. Θέτουμε $A_n = \cup_{i=1}^n B_i$ $n = 1, 2, \ldots$. Τότε, η A_n είναι αύξουσα και $\cup_{n\in\mathbb{N}} A_n = \cup_{n\in\mathbb{N}} B_n$. Άρα, λόγω του 1 έχουμε $\cup_{n\in\mathbb{N}} B_n \in \mathcal{A}$.

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, $\lim_{n\to\infty} \mathbf{P}(A_n) = 0$.

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 $\Omega = (0,1), \mathcal{F} = \mathcal{B}((0,1)), \mathbf{P} = \lambda_1$ το μέτρο Lebesgue, $I = I' = (0,1)$, $A_x := \{x\}, B_x := (0,1) \backslash \{x\}$ 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 $\mathbf{P}(\cap_{x\in(0,1)} B_x) = \mathbf{P}(\emptyset) = 0$.

2.5 Για $n \geq 1$, 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

$$\mathbf{P}(\liminf A_n) = \mathbf{P}(\cup_{n=1}^\infty \cap_{k=n}^\infty A_k) = \lim_{n\to\infty} \mathbf{P}(\cap_{k=n}^\infty A_k) \leq \varliminf_{n\to\infty} \mathbf{P}(A_n).$$

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

$$\varlimsup_{n\to\infty} \mathbf{P}(A_n) \leq \mathbf{P}(\limsup A_n)$$

αποδεικνύεται όμοια.

Κεφάλαιο 3

3.1 Χρήσιμη είναι η Πρόταση 2.12

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 Έστω $A(F) := \{x \in \mathbb{R} :$ Η F 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, θεωρούμε το σύνολο $B = \{x \in \mathbb{R} :$ η $\mathbf{F}$ είναι ασυνεχής στο $x\}$ και θέτουμε $A_\beta = \{\beta\}, \beta \in B$. Τότε $\mathbf{P}(A_\beta) = \mathbf{F}(\beta) - \mathbf{F}(\beta-) > 0$, αφού η $\mathbf{F}$ έχει άλμα στο x, και εφαρμόζουμε το συμπέρασμα της Άσκησης 2.5.

4.2 (α) Έχουμε ότι

$$\begin{aligned}\mathbf{P}((0,4)) &= \lambda\,\mathbf{P}_1((0,4)) + (1-\lambda)\,\mathbf{P}_2((0,4)) = \lambda\int_0^4 e^{-x}\,\mathrm{d}x + (1-\lambda)\frac{1}{2}\\ &= \lambda(1-e^{-4}) + (1-\lambda)\frac{1}{2}.\end{aligned}$$

(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

$$F(x) = \begin{cases} 0 & \text{αν } x < -2,\\ (1-\lambda)\frac{1}{2} & \text{αν } -2 \le x \le 0,\\ \lambda(1-e^{-x}) + (1-\lambda)\frac{1}{2} & \text{αν } 0 < x < 3,\\ 1-\lambda e^{-x} & \text{αν } x \ge 3.\end{cases}$$

4.3 Για το (α), έχουμε

$$\mathbf{P}(\{x\}) = \mathbf{P}\left(\bigcap_{n\in\mathbb{N}^+}\left(x-\frac{1}{n}, x\right]\right) = \lim_{n\to\infty}\mathbf{P}\left(\left(x-\frac{1}{n}, x\right]\right) = \lim_{n\to\infty}\left(F(y) - F\left(x-\frac{1}{n}\right)\right) \overset{(i)}{=} F(y) - F(x-).$$

Για το (β),

$$\mathbf{P}([x,y]) = \mathbf{P}(\{x\}) + \mathbf{P}((x,y]) \overset{(\alpha),(4.1)}{=} F(x) - F(x-) + F(y) - F(x) = F(y) - F(x-).$$

Τα (γ) και (δ) προκύπτουν με τον ίδιο τρόπο.

Κεφάλαιο 5

5.1 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 $\sigma(\mathcal{A}_4) = \mathscr{B}(\mathbb{R})$.

5.2 Τα σύνολα $\{-\infty\}, \{\infty\}$ είναι κλειστά.

5.3 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 1.1 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 $(\Omega, \mathcal{F})$.

5.4 (α) $\{\omega \in \Omega : \lim_n X_n(\omega) = \infty\} = \cap_{k=1}^\infty \cup_{j=1}^\infty \cap_{r=j}^\infty \{X_r > k\}$.

(β) Το $\lim_n X_n(\omega)$ υπάρχει στο $\mathbb{R}$ αν και μόνο αν η ακολουθία $(X_n(\omega))_{n\ge1}$ είναι βασική. Δηλαδή για κάθε $k \ge 1$ υπάρχει $j \ge 1$ ώστε $|X_r(\omega) - X_s(\omega)| < 1/k$ για κάθε $r, s \ge j$. Άρα, το δοσμένο σύνολο γράφεται ως ...

Εναλλακτικά, το δοσμένο σύνολο γράφεται ως $\{\underline{\lim}_{n\to\infty} X_n(\omega) = \overline{\lim}_{n\to\infty} X_n(\omega)\} \cap \{\underline{\lim}_{n\to\infty} X_n(\omega) \in \mathbb{R}\}$. Οι $\underline{\lim}_{n\to\infty} X_n(\omega), \overline{\lim}_{n\to\infty} X_n(\omega)$ είναι μετρήσιμες.

5.5 Χρησιμοποιούμε το Πόρισμα 5.4. Η αντίστροφη εικόνα κάθε διαστήματος είναι διάστημα (και άρα Borel-μετρήσιμο) αφού η f είναι μονότονη.

5.6 Θεωρούμε τη συνάρτηση $h = f - g$. Τότε η h είναι μετρήσιμη και ισχύει ότι $\{f = g\} = h^{-1}(\{0\}) \in \mathcal{A}$, εφόσον $\{0\} \in \mathscr{B}(\mathbb{R})$.

Εναλλακτικά,

$$\begin{aligned}\Omega\backslash\{f=g\} &= \{f\neq g\} = \{f>g\}\cup\{g>f\}\\ &= \Big(\cup_{q\in\mathbb{Q}}(\{f>q\}\cap\{g<q\})\Big)\cup\Big(\cup_{q\in\mathbb{Q}}(\{g>q\}\cap\{f<q\})\Big)\end{aligned}$$

από το οποίο έπεται το ζητούμενο.

5.11 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 6

6.1 Στην ισότητα $1-\mathbf{1}_{\cup_{i=1}^n A_i}=(1-\mathbf{1}_{A_1})(1-\mathbf{1}_{A_2})\cdots(1-\mathbf{1}_{A_n})$, αναπτύσσουμε το δεξί μέλος και παίρνουμε μέση τιμή.

6.2 Θεωρούμε την τυχαία μεταβλητή $X=\sum_{k=1}^n \mathbf{1}_{A_k}$.

6.3 $X=\sum_{k=1}^X 1=\sum_{k=1}^\infty \mathbf{1}_{k\le X}$ και θεώρημα Beppo Levi.

6.4 $[X]\le X\le [X]+1$.

6.6 (α) Δουλεύουμε όπως στην αποδειξη της ανισότητας Chebyshev.

$$\begin{aligned}P(X\le a\,\mathbf{E}\,X) &= \mathbf{P}(X-\mathbf{E}\,X\le -(1-a)\,\mathbf{E}\,X)\\ &\le \mathbf{P}(|X-\mathbf{E}\,X|\ge (1-a)\,\mathbf{E}\,X)\le \frac{\mathrm{Var}(X)}{(1-a)^2(\mathbf{E}\,X)^2}.\end{aligned}$$

Χρησιμοποιήσαμε το ότι $(1-a)EX>0$.

(β) Έστω $A:=\{\omega: X(\omega)>a\,\mathbf{E}\,X\}$.

$$\begin{aligned}&\mathbf{E}\,X=\mathbf{E}(X1_{A^c})+\mathbf{E}(X1_A)\le a\,\mathbf{E}\,X+\mathbf{E}(X^2)^{1/2}\,\mathbf{P}(A)^{1/2}\Rightarrow\\ &(1-a)\,\mathbf{E}\,X\le \mathbf{E}(X^2)^{1/2}\,\mathbf{P}(A)^{1/2}\Rightarrow \mathbf{P}(A)\ge (1-a)(\mathbf{E}\,X)^2/\,\mathbf{E}(X^2)\end{aligned}$$

Στην πρώτη γραμμή χρησιμοποιήσαμε την ανισότητα Cauchy-Schwarz και το ότι $X\le a\,\mathbf{E}\,X$ στο A^c.

6.7 $\mathbf{P}(X>t)=\mathbf{P}(aX>at)=\mathbf{P}(e^{aX}>e^{at})\le \mathbf{E}(e^{aX})/e^{at}$. Παίρνουμε $C=\mathbf{E}(e^{aX})\in(0,\infty)$.

6.8 Χρησιμοποιούμε την ανισότητα Cauchy-Schwarz.

$$1\le \mathbf{E}(\sqrt{XY})\le (\mathbf{E}(X))^{1/2}(\mathbf{E}(Y))^{1/2}.$$

6.10 Χρησιμοποιούμε την προηγούμενη άσκηση.

6.13 Η ακολουθία $A_n:=\{|X|>n\}, n\ge 1$ είναι φθίνουσα με τομή το $\emptyset$ αφού η X παίρνει τιμές στο $\mathbb{R}$. Άρα $\lim_{n\to\infty}\mathbf{P}(A_n)=\mathbf{P}(\emptyset)=0$.

6.14 Έστω $a_t:=\int f_t(x)d\mu(x)$. Για οποιοδήποτε $\ell\in\bar{\mathbb{R}}$ ισχύει $\lim_{t\to\infty}a_t=\ell$ αν και μόνο αν για κάθε ακολουθία $(t_n)_{n\ge1}$ με $t_n\to\infty$ ισχύει $\lim_{n\to\infty}a_{t_n}=\ell$. Έτσι, αναγόμαστε στο γνωστό θεώρημα κυριαρχημένης σύγκλισης.

6.15 Στην απόδειξη της ανισότητας Markov μπορούμε να είμαστε λιγότερο γενναιόδωροι και να παρατηρήσουμε ότι η ποσότητα $n\,\mathbf{P}(|X|\ge n)$ είναι μικρότερη από τη μέση τιμή $\mathbf{E}(|X|\mathbf{1}_{|X|\ge n})$.

6.17 Για το (α), αρκεί να το δείξουμε για κάθε ακολουθία $(\varepsilon_n)_{n\ge1}$ θετικών αριθμών με $\varepsilon_n\to 0$. Χρησιμοποιούμε το θεώρημα κυριαρχημένης σύγκλισης με κυριαρχούσα συνάρτηση την 1 αφού

$$\left|\frac{X}{\varepsilon}\mathbf{1}_{X<\varepsilon}\right|\le 1.$$

Κεφάλαιο 7

7.3 Εύκολα ελέγχουμε την ισότητα για $Y = \sum_{i=1}^n a_i \mathbf{1}_{A_i}$, $A_i \in \mathcal{F}$, απλή αφού $\int \mathbf{1}_A \, d\mathbf{Q} = \mathbf{Q}(A) = \mathbf{E_P}(X\mathbf{1}_A)$ 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 $\lim_{n\to\infty} \int Y_n \, d\mathbf{Q} = \int Y \, d\mathbf{Q}$ και $\lim_{n\to\infty} \mathbf{E_P}(Y_n X) = \mathbf{E_P}(YX)$. Όμως, $\int Y_n \, d\mathbf{Q} = \mathbf{E_P}(Y_n X)$ για κάθε $n \in \mathbb{N}$, άρα και πάλι το ζητούμενο ισχύει. Τέλος, αν $\mathbf{E_P}(|Y|X) < \infty$, από τα προηγούμενα έχουμε ότι $\int Y^- \, d\mathbf{Q} = \mathbf{E_P}(Y^- X)$, $\int Y^+ \, d\mathbf{Q} = \mathbf{E_P}(Y^+ X)$ όπου και οι τέσσερις αυτοί αριθμοί είναι πεπερασμένοι εφόσον $\mathbf{E_P}(Y^- X) + \mathbf{E_P}(Y^+ X) = \mathbf{E_P}(|Y|X) < \infty$, $\int Y^- \, d\mathbf{Q} + \int Y^+ \, d\mathbf{Q} = \int |Y| \, d\mathbf{Q} = \mathbf{E_P}(|Y|X)$ (η τελευταία ισότητα ισχύει αφού η $|Y|$ είναι θετική). Συνεπώς,

$$\int Y \, d\mathbf{Q} = \int Y^+ \, d\mathbf{Q} - \int Y^- \, d\mathbf{Q} = \mathbf{E_P}(Y^+ X) - \mathbf{E_P}(Y^- X) = \mathbf{E_P}(YX).$$

7.4 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 8

8.1 Για τα ω στο σύνολο $\limsup_n A_n^\varepsilon$ έχουμε $|X_n| \geq \varepsilon$ για άπειρα n και άρα $\limsup_n A_n^\varepsilon \subset \Omega \setminus \{\lim_{n\to\infty} X_n = 0\}$. Για το ότι το (β) δίνει το (α), παρατηρούμε ότι $\Omega \setminus \{\lim_{n\to\infty} X_n = 0\} = \cup_{k=1}^\infty \limsup_n A_n^{1/k}$ 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 $[0,\infty)$.

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 6.12(iii)] έπεται ότι το σύνολο των $x \in (0,1)$ με $f(x) = \infty$ έχει μέτρο Lebesgue 0.

9.2 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.

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. Έστω ότι δεν υπάρχει τέτοιο c. Τότε υπάρχει $a \in \mathbb{R}$ ώστε $\mathbf{P}(X < a) = p \in (0,1)$ και επομένως $\mathbf{P}(X \geq a) = 1 - p \in (0,1)$. Η ανεξαρτησία δίνει

$$\mathbf{P}(X < a, Y \geq a) = \mathbf{P}(X < a)\mathbf{P}(Y \geq a) > 0,$$

ενώ το αριστερό μέλος είναι μικρότερο απο $\mathbf{P}(X \neq Y) = 0$ από υπόθεση. Άτοπο.

10.7 (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,

$$\begin{aligned} \mathbf{P}\left(\left|\frac{S_n}{n}\right| > \epsilon\right) &= \mathbf{P}(|S_n| > n\epsilon) = \mathbf{P}(S_n^2 > n^2\epsilon^2) \\ &\leq \frac{\mathbf{E}(S_n^2)}{n^2\epsilon^2} = \frac{n}{n^2\epsilon^2} = \frac{1}{n\epsilon^2} \to 0 \end{aligned}$$

για $n \to \infty$.

10.8. Για $\varepsilon > 0$, χρησιμοποιώντας την ανεξαρτησία των $X_1, X_2, \ldots,$ βρίσκουμε

$$\mathbf{P}(|m_n| > \varepsilon) = \mathbf{P}(m_n > \varepsilon) = \mathbf{P}(X_1 > \varepsilon, \ldots, X_n > \varepsilon) = (1-\varepsilon)^n \to 0,$$

και όμοια

$$\mathbf{P}(|M_n - 1| > \varepsilon) = \mathbf{P}(M_n < 1-\varepsilon) = \mathbf{P}(X_1 < 1-\varepsilon, \ldots, X_n < 1-\varepsilon) = (1-\varepsilon)^n \to 0.$$

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 $A_n^\varepsilon := \{M_n \le 1-\varepsilon\}$. Επειδή $\mathbf{P}(A_n^\varepsilon) \le (1-\varepsilon)^n$ και $\sum_{n\ge1} \mathbf{P}(A_n) < \infty$, με πιθανότητα 1 για όλα τα μεγάλα n ισχύει $M_n \ge 1-\varepsilon$. Άρα το σύνολο $B_\varepsilon := \{\underline{\lim}_{n\to\infty} M_n \ge 1-\varepsilon\}$ έχει πιθανότητα 1 και επομένως το ίδιο ισχύει και για το $\cap_{k=1}^\infty B_{1/k} = \{\underline{\lim}_{n\to\infty} M_n \ge 1\}$.

Εναλλακτικά, μπορεί να παρατηρήσει κανείς ότι $\mathbf{P}(A_n^{1/\sqrt{n}}) = (1-n^{-1/2})^n \le e^{-\sqrt{n}}$ (με χρήση της $1-x \le e^{-x}$) και $\sum_{n\ge1} e^{-\sqrt{n}} < \infty$. Το συμπέρασμα έπεται από το πρώτο λήμμα Borel-Cantelli.

11.6 Αρκεί για κάθε $n \ge 1$ να βρούμε σταθερά M_n ώστε $\mathbf{P}(|X_n| > M_n) \le n^{-2}$. 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 $X_n \le M$ για όλα τα μεγάλα n και έπεται το συμπέρασμα.

[Να το γράψουμε και τυπικά. Το σύνολο $A := \limsup_{n\ge1}\{X_n > M\}$ έχει πιθανότητα 0 και για κάθε $\omega \in \Omega\backslash A$ υπάρχει φυσικός $n_0(\omega)$ ώστε $X_n \le M$ για κάθε $n \ge n_0(\omega)$. Άρα

$$X^*(\omega) \le \max\{X_1, X_2, \ldots, X_{n_0(\omega)-1}, M\} < \infty,$$

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 $C_K := \{X^* \ge K\}$ έχει πιθανότητα 1 και άρα και το $\cap_{K=1}^\infty C_K$ (*αριθμήσιμη* τομή συνόλων με πιθανότητα 1). Όμως $\cap_{K=1}^\infty C_K = \{X^* = \infty\}$, το οποίο από υπόθεση έχει πιθανότητα 0, και έχουμε άτοπο.

11.8 Χρήσιμη είναι η Άσκηση 7.4.

11.9. Το όριο είναι το πολύ 1 λόγω της Άσκησης 2.6. Για το κάτω φράγμα, εφαρμόζουμε το πρώτο Λήμμα Borel-Cantelli. Για $\varepsilon > 0$

$$\mathbf{P}\left(\frac{M_n}{\log n} \le 1-\varepsilon\right) = \mathbf{P}(X_1 \le (1-\varepsilon)\log n)^n = (1-n^{-(1-\varepsilon)})^n \le (e^{-n^{\varepsilon-1}})^n = e^{-n^\varepsilon}$$

και συνεχίζουμε όπως στην προηγούμενη άσκηση.

11.10. (α) Έστω $\epsilon > 0$. Θέτουμε $B_\epsilon = \left\{\overline{\lim_{n\to\infty}} \dfrac{L_n}{\log_2 n} \le 1+\epsilon\right\}$. Θα δείξουμε ότι $\mathbf{P}(B_\epsilon) = 1$.

Έστω $A_n = \left\{\frac{L_n}{\log_2 n} > 1+\epsilon\right\}$. Τότε

$$\begin{aligned}\mathbf{P}(A_n) &= \mathbf{P}(L_n > (1+\epsilon)\log_2 n) = \mathbf{P}(X_n = X_{n+1} = \ldots = X_{n+[(1+\epsilon)\log_2 n]-1} = K \text{ ή } \Gamma)\\ &= 2\,\mathbf{P}(X_n = X_{n+1} = \ldots = X_{n+[(1+\epsilon)\log_2 n]-1} = K)\\ &= 2\left(\frac{1}{2}\right)^{[(1+\epsilon)\log_2 n]} \le 2\left(\frac{1}{2}\right)^{(1+\epsilon)\log_2 n-1} = \frac{4}{2^{(1+\epsilon)\log_2 n}} = \frac{4}{n^{1+\epsilon}}.\end{aligned}$$

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 $\mathbf{P}(\{\limsup_{n\geq 1} A_n\}^c) = 1$.

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

$$\overline{\lim_{n\to\infty}} \frac{L_n(\omega)}{\log_2 n} \leq 1 + \epsilon.$$

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

$$\left\{\overline{\lim_{n\to\infty}} \frac{L_n}{\log_2 n} \leq 1\right\} = \cap_{k=1}^{\infty} B_{\frac{1}{k}}$$

και $\mathbf{P}(B_{1/k}) = 1$ για κάθε $k \geq 1$, από την Άσκηση 2.3 (β) έχουμε

$$\mathbf{P}\left(\cap_{k=1}^{\infty} B_{\frac{1}{k}}\right) = 1.$$

Έτσι, το (α) αποδείχθηκε.

(β) Επειδή κάθε L_n παίρνει τιμή που είναι ένας θετικός ακέραιος ή ∞, το ζητούμενο ισοδυναμεί με $\mathbf{P}(L_n = 1$ 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 $\mathbf{P}(B_n) = \frac{1}{2}\frac{1}{2} = \frac{1}{4}$. 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 $\sum_{n=1} \mathbf{P}(|X_n| \geq \varepsilon n) < \infty$ για κάθε $\varepsilon > 0$. Έπειτα χρησιμοποιούμε ότι οι X_n είναι ισόνομες και την Άσκηση 6.4.

11.12 Χρησιμοποιούμε την Άσκηση 6.4 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 $\varepsilon > 0$, με πιθανότητα 1, η ανισότητα

$$\frac{1}{n} \log X_n > a + \varepsilon$$

ισχύει μόνο για πεπερασμένα n.

11.17. (β) Από το προηγούμενο ερώτημα, οι $(X_i)_{i\in I}$ είναι ανεξάρτητες. Άρα ο νόμος 0-1 του Kolmogorov εφαρμόζεται για την τελική σ-αλγεβρα τους

$$C_\infty := \cap_{n=1}^{\infty} \sigma(X_n, X_{n+1}, \ldots).$$

Βέβαια το αν ένα $\omega \in \Omega$ ανήκει σε ένα από τα $\liminf A_i, \limsup A_i$ δεν εξαρτάται από οποιαδήποτε πεπερασμένο πλήθος $X_1(\omega), \ldots, X_k(\omega)$, οπότε και τα δύο σύνολα ανήκουν στη C_∞. Τυπικά το αποδεικνύουμε ως εξής. Για κάθε $n \geq 1$,

$$\liminf_i A_i = \cup_{j=1}^{\infty} \cap_{k=j}^{\infty} A_k = \cup_{j=n}^{\infty} \cap_{k=j}^{\infty} A_k = \cup_{j=n}^{\infty} \cap_{k=j}^{\infty} X_k^{-1}(\{1\}) \in \sigma(X_n, X_{n+1}, \ldots).$$

Η δεύτερη ισότητα ισχύει γιατί έχουμε ένωση μιας αύξουσας ακολουθίας συνόλων. Όμοια

$$\limsup_i A_i = \cap_{j=1}^{\infty} \cup_{k=j}^{\infty} A_k = \cap_{j=n}^{\infty} \cup_{k=j}^{\infty} A_k = \cap_{j=n}^{\infty} \cup_{k=j}^{\infty} X_k^{-1}(\{1\}) \in \sigma(X_n, X_{n+1}, \ldots).$$

11.18. (α) Από τον απειροστικό λογισμό ξέρουμε ότι $R^{-1} = \overline{\lim}_{n\to\infty} |X_n|^{1/n}$.

(β) Δείχνουμε όπως στην Άσκηση 11.8 ότι $\overline{\lim}_{n\to\infty} |X_n|^{1/n} = 1$ με πιθανότητα 1. Μάλιστα εδώ αρκεί η χρήση του πρώτου λήμματος Borel-Cantelli για να δείξουμε ότι $\lim_{n\to\infty} |X_n|^{1/n} = 1$.

11.20. Δουλεύουμε όπως στην προηγούμενη άσκηση.

Κεφάλαιο 12

12.1 Έστω $Y = \frac{S_n}{n}$. Τότε $\mathbf{E}(Y) = \frac{1}{n}\mathbf{E}(X_1 + X_2 + ... + X_n) = \frac{1}{n} n\,\mathbf{E}(X_1) = \mu$ και $\mathrm{Var}(Y) = \mathrm{Var}(\frac{1}{n}S_n) = \frac{1}{n^2}\mathrm{Var}(X_1 + X_2 + ... + X_n) = \frac{1}{n}\mathrm{Var}(X_1)$. Έστω $\epsilon > 0$. Τότε,

$$\mathbf{P}\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) = \mathbf{P}(|Y - \mathbf{E}(Y)| \geq \epsilon) \leq \frac{1}{\epsilon^2}\mathrm{Var}(Y) = \frac{1}{\epsilon^2}\frac{\mathrm{Var}(X_1)}{n} \to 0$$

για $n \to \infty$, 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 $S_n^M = Y_1^M + Y_2^M + ... + Y_n^M$ για κάθε $n \geq 1$. Τότε, με πιθανότητα 1,

$$\varliminf_{n\to\infty} \frac{S_n}{n} \geq \varliminf_{n\to\infty} \frac{S_n^M}{n} = \mathbf{E}(Y_1^M),$$

όπου η ισότητα προκύπτει από το νόμο των μεγάλων αριθμών. Έστω τώρα

$$\Omega_M = \left\{\omega \in \Omega : \varliminf_{n\to\infty} \frac{S_n(\omega)}{n} \geq \mathbf{E}(X_1 \wedge M)\right\}.$$

Δείξαμε ότι $\mathbf{P}(\Omega_M) = 1$ για κάθε $M > 0$. Θεωρώντας το σύνολο $A = \cap_{k=1}^{\infty}\Omega_k$, έχουμε ότι $\mathbf{P}(A) = 1$ και για $\omega \in A$, ισχύει ότι

$$\varliminf_{n\to\infty} \frac{S_n(\omega)}{n} \geq \mathbf{E}(X_1 \wedge k) \quad \text{για κάθε } k \geq 1. \tag{Β΄.13}$$

Από το θεώρημα μονότονης σύγκλισης, $\mathbf{E}(Y_1^k) = \mathbf{E}(X_1^+ \wedge k) - \mathbf{E}(X_1^-) \to \mathbf{E}(X_1^+) - \mathbf{E}(X_1^-) = \infty$ για $k \to \infty$. Άρα, για $\omega \in A$, παίρνοντας $k \to \infty$ στην (Β΄.13) έχουμε ότι $\varliminf_{n\to\infty} S_n(\omega)/n = \infty$, οπότε $\lim_{n\to\infty} S_n(\omega)/n = \infty$, που είναι το ζητούμενο.

12.3 Παρατηρούμε ότι

$$\frac{X_n}{n} = \frac{S_n}{n} - \frac{S_{n-1}}{n-1} = \frac{S_n}{n} - \frac{n-1}{n}\frac{S_{n-1}}{n-1} \to \mu - \mu = 0 \tag{Β΄.14}$$

για $n \to \infty$ 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 (Β΄.14) 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 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΄.14). Τέλος, ισχύει ότι (Άσκηση 6.4)

$$\sum_{n=1}^{\infty}\mathbf{P}\left(|X_n| \geq n\right) \leq \mathbf{E}(|X_1|) \leq \sum_{n=1}^{\infty}\mathbf{P}\left(|X_n| \geq n\right) + 1.$$

Άρα $\mathbf{E}(X_1) \in \mathbb{R}$ και από το νόμο των μεγάλων αριθμών $\frac{S_n}{n} \to \mathbf{E}(X_1)$. Όμως, από υπόθεση $\frac{S_n}{n} \to \mu$, επομένως $\mathbf{E}(X_1) = \mu$.

12.4 Έχουμε ότι $\mathbf{E}(X_1) = 1$ και $\mathbf{E}(X_1^2) = \mathrm{Var}(X_1) + \mathbf{E}(X_1)^2 = 4$, καθώς επίσης και ότι οι $\{X_i^2, i \geq 1\}$ είναι ανεξάρτητες και ισόνομες. Γράφουμε

$$\frac{X_1 + X_2 + \ldots + X_n}{(X_1)^2 + (X_2)^2 + \ldots + (X_n)^2} = \frac{\frac{X_1+X_2+\ldots+X_n}{n}}{\frac{(X_1)^2+(X_2)^2+\ldots+(X_n)^2}{n}}.$$

Από το νόμο των μεγάλων αριθμών ο αριθμητής συγκλίνει στο 1 με πιθανότητα 1 και αντίστοιχα ο παρονομαστής στο 2. Τυπικά, για τα $A_1 = \{\omega \in \Omega : \lim_{n\to\infty} \frac{X_1+X_2+\ldots+X_n}{n} = 1\}$ και $A_2 = \{\omega \in \Omega : \lim_{n\to\infty} \frac{(X_1)^2+(X_2)^2+\ldots+(X_n)^2}{n} = 1\}$ 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

$$\lim_{n\to\infty} \frac{X_1(\omega) + X_2(\omega) + \ldots + X_n(\omega)}{(X_1)^2(\omega) + (X_2)^2(\omega) + \ldots + (X_n)^2(\omega)} = \frac{1}{4},$$

από το οποίο προκύπτει το ζητούμενο.

12.5. $S_n = n(S_n/n) \to \infty \times \mathbf{E}(X_1) = \infty$ αφού $\mathbf{E}(X_1) > 0$.

12.6. (α) $(U_1 U_2 \cdots U_n)^{1/n} = e^{\frac{1}{n}(\log U_1 + \cdots + \log U_n)}$ Επειδή $\mathbf{E}(\log U_1) = \int_0^1 \log x\, dx = \ldots = -1$, ο ισχυρός νόμος των μεγάλων αριθμών για την ακολουθία $(\log U_i)_{i\geq 1}$ δίνει ότι

$$\lim_{n\to\infty} \frac{\log U_1 + \cdots + \log U_n}{n} = -1 \text{ με πιθανότητα } 1.$$

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 $(U_1 U_2 \cdots U_n)^{1/n} < \theta$. Άρα για $n > n_0$

$$0 < U_1 U_2 \cdots U_n < \theta^n \to 0$$

καθώς $n \to \infty$ επειδή $0 < \theta < 1$.

Εναλλακτικά, $U_1 U_2 \cdots U_n = e^{n \frac{\log U_1 + \cdots + \log U_n}{n}} \to 0$ 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.

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 $f : \mathbb{C}\backslash[\lambda, \infty) \to \mathbb{C}$ με

$$f(z) = \left(1 - \frac{z}{\lambda}\right)^{-a} = e^{-a \log\left(1 - \frac{z}{\lambda}\right)}$$

είναι αναλυτική στο πεδίο ορισμού της, το οποίο περιέχει το $\{z \in \mathbb{C} : |\mathrm{Re}(z)| < \lambda\}$, και $M_X(t) = f(t)$ για κάθε $t \in (-\lambda, \lambda)$. Το συμπέρασμα έπεται από την Πρόταση 13.17.

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 (α) Η χαρακτηριστική συνάρτηση της τυχαίας μεταβλητής $\frac{X+Y}{2}$ είναι

$$\phi_{\frac{X+Y}{2}}(t) = \mathbf{E}(e^{iXt/2}e^{iYt/2}) = \mathbf{E}(e^{iXt/2})\,\mathbf{E}(e^{iYt/2}) = \phi_X\left(\frac{t}{2}\right)\phi_Y\left(\frac{t}{2}\right) = e^{-|t|/2}e^{-|t|/2} = e^{-|t|},$$

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 13.9 έπεται ότι $\frac{X+Y}{2} \sim$ **Cauchy**.

(β) Εργαζόμαστε όμοια με το (α) βασιζόμενοι στην ανεξαρτησία των $X_1, X_2, \ldots, X_n$.

13.11

$$\phi_X(t) - 1 = 2\int_1^\infty \frac{1}{|x|^3}(\cos(tx) - 1)\,dx$$

Το $|\cos(x) - 1|$ είναι περίπου $x^2/2$ όταν το x είναι κοντά στο 0 και φράσσεται από το 2 για οποιοδήποτε x.

13.12 Έχουμε ότι $f_X(x) = f_Y(y) = \mathbf{1}_{(0,1)}(x)$ για κάθε $x \in \mathbb{R}$. Από το Θεώρημα 13.23 η $X + Y$ έχει πυκνότητα $f_{X+Y}(z) = \displaystyle\int_{\mathbb{R}} f_X(x) f_Y(z - x)\,dx$. Ο ολοκληρωτέος είναι μη μηδενικός για (x, z) τέτοια ώστε $0 < x < 1$ και $0 < z - x < 1$, δηλαδή,

$$\begin{aligned} 0 < x < 1, &\qquad \text{και}\\ z - 1 < x < z. &\end{aligned} \tag{Β$'$.15}$$

- Αν $z \leq 0$ ή $z \geq 2$, τότε δεν υπάρχουν x που ικανοποιούν την (Β$'$.15).
- Αν $z \in (0, 1)$, τότε από (Β$'$.15), $0 < x < z$ και $f_{X+Y}(z) = \int_0^z 1\,\mathrm{d}x = z$.
- Αν $z \in (1, 2)$ τότε από (Β$'$.15), $z - 1 < x < 1$ και $f_{X+Y}(z) = \int_{z-1}^1 1\,\mathrm{d}x = 2 - z$.

Συνδυάζοντας τα παραπάνω, προκύπτει το ζητούμενο.

Κεφάλαιο 14

14.1. Η συνάρτηση κατανομής της Z είναι

$$F_Z(x) = \begin{cases} 0 & \text{αν } x < 0,\\ x & \text{αν } x \in [0,1],\\ 1 & \text{αν } x > 1,\end{cases}$$

και τα σημεία συνέχειάς της είναι όλο το $\mathbb{R}$. Για $x \in [0, 1]$,

$$F_{Y_n}(x) = \mathbf{P}(X_n \leq nx) = \frac{[nx]}{n} \to x = F_Z(x)$$

για $n \to \infty$ (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 14.11, η σύγκλιση $X_n \Rightarrow X$ συνεπάγεται (μάλιστα ισοδυναμεί με) την

$$\mathbf{P}(X_n \in A) \to \mathbf{P}(X \in A) \text{ για όλα τα } A \subset \mathbb{R} \text{ Borel με } \mathbf{P}(X \in \partial A) = 0.$$

Είναι δυνατόν βέβαια η σχέση $\mathbf{P}(X_n \in A) \to \mathbf{P}(X \in A)$ 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 $\partial A = \bar{A} \backslash A^o = \mathbb{R} \backslash \emptyset = \mathbb{R}$, 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 $\partial A = \bar{A} \backslash A^o = [0, 1/3] \backslash \emptyset = [0, 1/3]$ 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

$$\varliminf_{n\to\infty} \mathbf{P}(X_n \leq n) \geq \varliminf_{n\to\infty} \mathbf{P}(X_n \leq x_M) = \mathbf{P}(X \leq x_M) \geq \mathbf{P}(X \leq M).$$

Για $M \to \infty$ 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

$$\varlimsup_{n\to\infty} \mathbf{P}(X_n \in [x - n^{-1}, x + n^{-1}]) \leq \varlimsup_{n\to\infty} \mathbf{P}(X_n \in [x_M, y_M]) = \mathbf{P}(X \in [x_M, y_M]) = F_X(y_M) - F_X(x_M)$$

Για $M \to \infty$ 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 $n \to \infty$ 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

$$\varlimsup_{n\to\infty} \mathbf{P}(|X_n Y_n| > \varepsilon) \leq \sup_{n\geq 1} \mathbf{P}(|X_n| \geq M)$$

Το αριστερό μέλος δεν εξαρτάται από το M, ενώ το δεξί, για $M \to \infty$, τείνει στο 0.

14.6 (α) Έστω x σημείο συνέχειας της F_X. Θα δείξουμε ότι $\lim_{n\to\infty} F_{Y_n}(x) = F_X(x)$.
Πρώτα δείχνουμε ότι $\varlimsup_{n\to\infty} F_{Y_n}(x) \leq F_X(x)$.
Έστω $\varepsilon > 0$. Τότε

$$\begin{aligned} F_{Y_n}(x) = \mathbf{P}(Y_n \leq x) &= \mathbf{P}(Y_n \leq x, X_n \leq x + \varepsilon) + \mathbf{P}(Y_n \leq x, X_n > x + \varepsilon) \\ &\leq \mathbf{P}(X_n \leq x + \varepsilon) + \mathbf{P}(X_n - Y_n > \varepsilon) \\ &= F_{X_n}(x + \varepsilon) + \mathbf{P}(X_n - Y_n > \varepsilon). \end{aligned}$$

Αν το $x + \varepsilon$ είναι σημείο συνέχειας της F_X, τότε

$$\lim_{n\to\infty} F_{X_n}(x + \varepsilon) = F_X(x + \varepsilon)$$

και

$$\lim_{n\to\infty} \mathbf{P}(X_n - Y_n > \varepsilon) = 0$$

εφόσον $X_n - Y_n \xrightarrow{\mathbf{P}} 0$. Άρα

$$\varlimsup_{n\to\infty} F_{Y_n}(x) \leq F_X(x + \varepsilon).$$

Τα σημεία ασυνέχειας της F_X είναι αριθμήσιμα, άρα υπάρχει 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

$$\overline{\lim_{n\to\infty}} F_{Y_n}(x) \le F_X(x+\varepsilon_k).$$

Εφόσον η F_X είναι συνεχής στο x, έχουμε $\lim_{k\to\infty} F_X(x+\varepsilon_k) = F(x)$, και άρα

$$\overline{\lim_{n\to\infty}} F_{Y_n}(x) \le F_X(x).$$

Δείχνουμε τώρα ότι $\underline{\lim}_{n\to\infty} F_{Y_n}(x) \ge F_X(x)$.

Για $\varepsilon > 0$,

$$F_{Y_n}(x) = \mathbf{P}(Y_n \le x) \ge \mathbf{P}(X_n \le x-\varepsilon) - \mathbf{P}(Y_n - X_n \ge \varepsilon).$$

Συνεπώς $\underline{\lim}_{n\to\infty} F_{Y_n}(x) \ge F_X(x-\varepsilon)$. Ακολουθώντας την ίδια διαδικασία όπως προηγουμένως βλέπουμε ότι $\underline{\lim}_{n\to\infty} F_{Y_n}(x) \ge F_X(x)$.

14.7 Έστω $A_n = \mathbb{R}\backslash[-n,n]$. Τότε η A_n είναι φθίνουσα ακολουθία, άρα,

$$\lim_{n\to\infty} p(A_n) = p(\cap_{n=1}^{\infty} A_n) = p(\emptyset) = 0.$$

Επομένως, για κάθε $\epsilon > 0$ υπάρχει n_0 έτσι ώστε $\mu(R\backslash[-n_0,n_0]) < \epsilon$.

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 $p_i([-M,M]) \ge p_i([-M_i,M_i]) > 1-\epsilon$, 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,

$$\mathbf{P}(|X_n| > M) = \mathbf{P}(h(|X_n|) > h(M)) \le \frac{1}{h(M)}\mathbf{E}(h(|X_n|)) \le \frac{1}{h(M)}C,$$

όπου $C = \sup_{n\ge1}\mathbf{E}(h(|X_n|)) < \infty$. Άρα $\sup_{n\ge1}\mathbf{P}(|X_n| > M) \le \frac{C}{h(M)} \to 0$ για $M\to\infty$.

Κεφάλαιο 15

15.1. Μιμούμαστε την απόδειξη του Λήμματος 15.1.

15.2. (α) Παράδειγμα 13.5(vi).

(β) Για $t\in\mathbb{R}$ έχουμε

$$\phi_X(t) = \mathbf{E}(e^{itX}) = \sum_{k=1}^{\infty} e^{itk}(1-p)^{k-1}p = pe^{it}\sum_{j=0}^{\infty}(e^{it}(1-p))^j = \frac{pe^{it}}{1-(1-p)e^{it}}.$$

Αθροίσαμε μια γεωμετρική πρόοδο της οποίας ο λόγος έχει μέτρο $|(1-p)e^{it}| = 1-p < 1$ αφού $p > 0$.

(γ) (i) Τα σημεία συνέχειας της συνάρτησης κατανομής F_X της X είναι όλο το $\mathbb{R}$. Έστω $x > 0$. Τότε

$$\begin{aligned} F_{X_n/n}(x) &= \mathbf{P}\left(\frac{X_n}{n} \le x\right) = \mathbf{P}(X_n \le [nx]) = 1 - \mathbf{P}(X_n > [nx]) \\ &= 1-(1-p_n)^{[nx]} = 1 - \left(\left(1-\frac{a}{n}\right)^n\right)^{[nx]/n}. \end{aligned}$$

Επειδή $[nx]/n \to x$ και $(1-an^{-1})^n \to e^{-a}$, έπεται ότι $\lim_{n\to\infty} F_{X_n/n}(x) = 1-e^{-ax} = F_X(x)$. Προφανώς το ίδιο ισχύει και για $x\le 0$.

(ii) Για $t\in\mathbb{R}$ έχουμε

$$\phi_{X_n/n}(t) = \phi_{X_n}(t/n) = \frac{p_n e^{it/n}}{1-(1-p_n)e^{it/n}} = \frac{e^{it/n}}{\frac{1-e^{it/n}}{a/n} + e^{it/n}} \to \frac{1}{1-\frac{it}{a}} = \phi_X(t)$$

καθώς $n \to \infty$ αφού

$$\lim_{n\to\infty} \frac{1-e^{it/n}}{a/n} = \lim_{n\to\infty} \frac{it}{a} \frac{1-e^{it/n}}{it/n} = -\frac{it}{a}.$$

Άρα η σύγκλιση $X_n/n \Rightarrow X$ έπεται από το θεώρημα συνέχειας του Lévy.

15.3. (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

$$\phi_Y(t) = \mathbf{E}(e^{itY}) = \sum_{k=0}^{n} e^{itk} \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^{n} \binom{n}{k} (pe^{it})^k (1-p)^{n-k} = (pe^{it} + 1 - p)^n.$$

(γ) Για $t \in \mathbb{R}$ έχουμε

$$\phi_{X_n}(t) = (p_n e^{it} + 1 - p_n)^n = (1 + p_n(e^{it} - 1))^n \to e^{\lambda(e^{it}-1)} = \phi_X(t)$$

αφού $\lim_{n\to\infty} p_n(e^{it}-1)n = \lambda(e^{it}-1)$. Άρα η σύγκλιση $X_n \Rightarrow X$ έπεται από το θεώρημα συνέχειας του Lévy.

15.4. 1ος τρόπος. Από ιδιότητες της κατανομής Γάμμα, η X_n έχει την ίδια κατανομή με την $(W_1 + W_2 + \ldots + W_n)/n$ όπου οι $W_1, \ldots, W_n$ είναι ανεξάρτητες, ισόνομες με $W_1 \sim \Gamma(c, 1)$. Το συμπέρασμα έπεται από το κεντρικό οριακό θεώρημα γιατί $\mathbf{E}(W_1) = \mathrm{Var}(W_1) = c$.

2ος τρόπος. Χρησιμοποιούμε το θεώρημα συνέχειας του Lévy και τον τύπο για τη χαρακτηριστική συνάρτηση της κατανομής Γάμμα (Άσκηση 13.4). Για $t \in \mathbb{R}$,

$$\begin{aligned}\phi_{\sqrt{n}(X_n - c)} &= e^{-c\sqrt{n}ti} \phi_{X_n}(\sqrt{n}t) = e^{-c\sqrt{n}ti} \frac{1}{\left(1 - \frac{i\sqrt{n}t}{n}\right)^{nc}} \\ &= e^{-c\sqrt{n}ti - nc\mathrm{Log}\left(1 - \frac{it}{\sqrt{n}}\right)}\end{aligned}$$

Log είναι ο κλάδος του λογαρίθμου που είναι ολόμορφος στο $\mathbb{C}$ εκτός των αρνητικών πραγματικών και $\mathrm{Log}(1) = 0$. Χρησιμοποιώντας το ανάπτυγμα σε δυναμοσειρά της $\mathrm{Log}(1-z)$ με κέντρο το 0 (στο δίσκο $\{z : |z| < 1\}$) βρίσκουμε ότι ο εκθέτης στην τελευταία ποσότητα συγκλίνει στο $-ct^2/2$ και έτσι έχουμε το ζητούμενο.

15.9. Χρήσιμη είναι η Άσκηση 13.10.

Κεφάλαιο 16

16.2. Για την ακολουθία $(S_n)_{n\geq 1}$ έχουμε ότι $S_n/n \to 2$ 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 $S_n/n \to 2$ κατά πιθανότητα δίνει

$$\mathbf{P}(S_n > 2.1n) = \mathbf{P}\left(\frac{S_n}{n} > 2.1\right) \leq \mathbf{P}\left(\left|\frac{S_n}{n} - 2\right| > 0.1\right) \to 0$$

καθώς $n \to \infty$.

(β) Έχουμε

$$\mathbf{P}(S_n > 2n + \sqrt{n}) = \mathbf{P}\left(\frac{S_n - 2n}{\sqrt{n}} > 1\right) \to \mathbf{P}(Z > 1) = 1 - \Phi(1)$$

καθώς $n \to \infty$ 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.

$$\mathbf{P}(S_n > 10\sqrt{n}) = P\left(\frac{S_n}{n} > \frac{10}{\sqrt{n}}\right) = 1 - P\left(\frac{S_n}{n} \le \frac{10}{\sqrt{n}}\right).$$

Για $n > 100$,

$$\mathbf{P}\left(\frac{S_n}{n} \le \frac{10}{\sqrt{n}}\right) \le \mathbf{P}\left(\left|\frac{S_n}{n} - 2\right| > 1\right) \to 0$$

Άρα $\lim_{n\to\infty} \mathbf{P}(S_n > 10\sqrt{n}) = 1$.

(δ)

$$\mathbf{P}(S_n \ge 3n) = \mathbf{P}\left(\frac{S_n}{n} \ge 3\right) \le \mathbf{P}\left(\left|\frac{S_n}{n} - 2\right| > \frac{1}{2}\right) \to 0$$

καθώς $n \to \infty$. Άρα $\lim_{n\to\infty} \mathbf{P}(S_n < 3n) = 1$.

(ε) Για $n > 10^{10}$ έχουμε

$$\mathbf{P}(S_n \le 10^{10}) = \mathbf{P}\left(\frac{S_n}{n} \le \frac{10^{10}}{n}\right) \le \mathbf{P}\left(\left|\frac{S_n}{n} - 2\right| > \frac{1}{2}\right) \to 0$$

καθώς $n \to \infty$. Άρα $\lim_{n\to\infty} \mathbf{P}(S_n > 10^{10}) = 1$.

16.3 Παρατηρούμε ότι $\mathbf{E}(X_1) = \operatorname{Var} X_1 = 1$ και εφαρμόζουμε το κεντρικό οριακό θεώρημα.

16.5 Παρατηρούμε ότι $\mathbf{P}(X = k) = e^{-n}\frac{n^k}{k!}$, όπου $X \sim \mathbf{Poisson}(n)$. Έστω $(X_k)_{k\in\mathbb{N}}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με $X_1 \sim \mathbf{Poisson}(1)$. Γνωρίζουμε ότι η $S_n = \sum_{k=1}^n X_k \sim \mathbf{Poisson}(n)$. Άρα,

$$\begin{aligned} e^{-n}\sum_{k=0}^{n}\frac{n^k}{k!} &= \sum_{k=0}^{n}\mathbf{P}(S_n \le n) \\ &= \mathbf{P}\left(\frac{S_n - n}{\sqrt{n}} \le 0\right) \\ &= \mathbf{F}_{\frac{S_n-n}{\sqrt{n}}}(0). \end{aligned}$$

Από την Άσκηση 16.3, αφού $\frac{S_n-n}{\sqrt{n}} \Rightarrow Z$, με $Z \sim \mathbf{N}(0,1)$, έχουμε ότι $\mathbf{F}_{\frac{S_n-n}{\sqrt{n}}}(0) \to \mathbf{F}_Z(0) = \frac{1}{2}$ για $n \to \infty$.

16.6. Έστω ακολουθίες $(Y_n)_{n\ge1}, (Z_n)_{n\ge1}$ τυχαίες μεταβλητές στον ίδιο χώρο πιθανότητας ώστε οι $\{Y_n, Z_n : n \ge 1\}$ να είναι ανεξάρτητες και ισόνομες και καθεμία να έχει την ίδια κατανομή με τη X_1.

Τότε επειδή το διάνυσμα $(X_1, \cdots, X_{2n})$ έχει την ίδια κατανομή με το $(Y_1, Y_2, \ldots, Y_n, Z_1, Z_2, \ldots, Z_n)$ (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

$$\begin{aligned} &\mathbf{P}\left(\left|\frac{X_1 + X_2 + \cdots + X_n - (X_{n+1} + \cdots + X_{2n})}{\sqrt{n}}\right| \le 1\right) \\ &\qquad = \mathbf{P}\left(\left|\frac{Y_1 + Y_2 + \cdots + Y_n - (Z_1 + \cdots + Z_n)}{\sqrt{n}}\right| \le 1\right) = \mathbf{P}\left(\left|\frac{W_1 + W_2 + \cdots + W_n}{\sqrt{n}}\right| \le 1\right), \end{aligned}$$

όπου θέσαμε $W_i = Y_i - Z_i$ για κάθε $i \ge 1$. Από την υπόθεση, οι $\{W_i : i \ge 1\}$ είναι ανεξάρτητες και ισόνομες, καθεμία με μέση τιμή $\mathbf{E}(W_1) = 0$ και διασπορά $V(W_1) = V(X_1) + V(Y_1) - 2\mathrm{Cov}(X_1, Y_1) = 1 + 1 - 0 = 2$. Άρα, εφαρμόζοντας το κεντρικό οριακό θεώρημα, βρίσκουμε

$$\begin{aligned} \lim_{n\to\infty}\mathbf{P}\left(\left|\frac{W_1 + W_2 + \cdots + W_n}{\sqrt{n}}\right| \le 1\right) &= \lim_{n\to\infty}\mathbf{P}\left(\left|\frac{W_1 + W_2 + \cdots + W_n}{\sqrt{n2}}\right| \le \frac{1}{\sqrt{2}}\right) \\ &= \mathbf{P}(|Z| \le 1/\sqrt{2}) = \Phi(1/\sqrt{2}) - \Phi(-1/\sqrt{2}) \\ &= 2\Phi(1/\sqrt{2}) - 1. \end{aligned}$$

Βιβλιογραφία

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).

Κουμουλλής Γ. και Νεγρεπόντης Σ. *Θεωρία Μέτρου*. Εκδόσεις Συμμετρία, Αθήνα, (1991).

Νεγρεπόντης Σ., Γιωτόπουλος Σ., Γιαννακούλιας Ε. *Απειροστικός Λογισμός*. Εκδόσεις Αίθρα, Αθήνα, (1992).

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

Ευρετήριο ξενόγλωσσων όρων

Μετάφραση ορολογίας

Άλγεβρα	Algebra
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
Μετασχηματισμός ποσοστημορίων	Quantile 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
Πυκνότητα μέτρου	Density of a measure
Ροπογεννήτρια	Moment generating function
Σύγκλιση ασθενής	Weak convergence
Σύγκλιση κατά κατανομή	Convergence in distribution
Σύγκλιση κατά πιθανότητα	Convergence in probability
Σύγκλιση σχεδόν βέβαιη	Almost sure convergence
Σ-άλγεβρα	Σ-algebra
Σ-πεπερασμένο	Σ-finite
Σφιχτή οικογένεια τυχαίων μεταβλητών	Tight family of random variables
Σφιχτή οικογένεια μέτρων	Tight family of measures
Στήριγμα μέτρου	Support of a measure
Συμμετρική κατανομή	Symmetric distribution
Συνάρτηση κατανομής	Distribution function
Συνάρτηση ρυθμού	Rate function
Συνεχής κατανομή	Continuous distribution
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