13

Χαρακτηριστικές συναρτήσεις

13.1 Μετασχηματισμός Fourier μέτρου πιθανότητας στο $\mathbb{R}$

Έστω $(\Omega, \mathcal{F}, \mu)$ χώρος μέτρου και $f : \Omega \to \mathbb{C}$ Borel-μετρήσιμη συνάρτηση. Το πραγματικό και φανταστικό μέρος της f, που τα συμβολίζουμε με $\mathrm{Re}(f), \mathrm{Im}(f)$, είναι συναρτήσεις στο Ω με πραγματικές τιμές και είναι εύκολο να δει κανείς ότι είναι Borel-μετρήσιμες. Ορίζουμε το ολοκλήρωμα Lebesgue της f ως προς το μέτρο μ ως εξής:

$$\int f \,\mathrm{d}\mu := \int \mathrm{Re}(f)\,\mathrm{d}\mu + i \int \mathrm{Im}(f)\,\mathrm{d}\mu$$

με την προϋπόθεση ότι τα δύο πραγματικά ολοκληρώματα ορίζονται και είναι πραγματικοί αριθμοί. Τότε λέμε ότι η f είναι ολοκληρώσιμη και για κάθε τέτοια f ισχύει

$$\Big|\int f\,\mathrm{d}\mu\Big| \le \int |f|\,\mathrm{d}\mu, \tag{13.1}$$

$$\int \overline{f}\,\mathrm{d}\mu = \overline{\int f\,\mathrm{d}\mu}, \tag{13.2}$$

όπου $|\cdot|$ συμβολίζει το μέτρο μιγαδικού. Η δεύτερη ιδιότητα είναι προφανής, ενώ για την πρώτη αρκεί κανείς να παρατηρήσει ότι υπάρχει $\theta \in [0, 2\pi)$ έτσι ώστε

$$\Big|\int f\,\mathrm{d}\mu\Big| = e^{i\theta}\int f\,\mathrm{d}\mu.$$

Τότε

$$\Big|\int f\,\mathrm{d}\mu\Big| = e^{i\theta}\int f\,\mathrm{d}\mu = \int e^{i\theta} f\,\mathrm{d}\mu = \int \mathrm{Re}(e^{i\theta} f)\,\mathrm{d}\mu \le \int |e^{i\theta} f|\,\mathrm{d}\mu = \int |f|\,\mathrm{d}\mu, \tag{13.3}$$

από το οποίο προκύπτει το ζητούμενο. Η τρίτη ισότητα ισχύει γιατί ξέρουμε ότι το αριστερό της μέλος είναι πραγματικός αριθμός.

Ορισμός 13.1. Έστω μ μέτρο πιθανότητας στον $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$. **Μετασχηματισμό Fourier** του μ ονομάζουμε τη συνάρτηση $\hat{\mu} : \mathbb{R} \to \mathbb{C}$ που ορίζεται ως

$$\hat{\mu}(t) := \int e^{itx}\,\mathrm{d}\mu(x) = \int \cos(tx)\,\mathrm{d}\mu(x) + i\int \sin(tx)\,\mathrm{d}\mu(x)$$

για κάθε $t \in \mathbb{R}$.

Θεώρημα 13.2. *Έστω μ μέτρο πιθανότητας στον $\mathbb{R}$. Τότε*

(i) $|\hat{\mu}(t)| \le 1$ *για κάθε* $t \in \mathbb{R}$.

(ii) $\hat{\mu}(0) = 1$.

(iii) *Η $\hat{\mu}$ είναι ομοιόμορφα συνεχής.*

Απόδειξη. (i) Για $t \in \mathbb{R}$, εχουμε ότι

$$|\hat{\mu}(t)| = \left|\int e^{itx}\,\mathrm{d}\mu(x)\right| \le \int |e^{itx}|\,\mathrm{d}\mu(x) = \int 1\,\mathrm{d}\mu(x) = 1.$$

(ii) $\hat{\mu}(0) = \displaystyle\int e^0\,\mathrm{d}\mu(x) = 1.$

(iii) Αρκεί να δείξουμε ότι για κάθε ακολουθία $(\delta_k)_{k\in\mathbb{N}}$ στον $\mathbb{R}$ με $\delta_k \to 0$ ισχύει ότι

$$\lim_{k\to\infty}\sup_{t\in\mathbb{R}}|\hat{\mu}(t+\delta_k) - \hat{\mu}(t)| = 0.$$

Έστω $t, \zeta \in \mathbb{R}$. Τότε,

$$\begin{aligned}|\hat{\mu}(t+\zeta) - \hat{\mu}(t)| &= \left|\int (e^{i(t+\zeta)x} - e^{itx})\,\mathrm{d}\mu(x)\right| = \left|\int e^{itx}(e^{i\zeta x} - 1)\,\mathrm{d}\mu(x)\right| \\ &\le \int |e^{itx}||e^{i\zeta x} - 1|\,\mathrm{d}\mu(x) = \int |e^{i\zeta x} - 1|\,\mathrm{d}\mu(x).\end{aligned}$$

Άρα, αν $(\delta_k)_{k\in\mathbb{N}}$ μηδενική ακολουθία, για $k \in \mathbb{N}$ έχουμε

$$\sup_{t\in\mathbb{R}}|\hat{\mu}(t+\delta_k) - \hat{\mu}(t)| \le \int |e^{i\delta_k x} - 1|\,\mathrm{d}\mu(x). \tag{13.4}$$

Έστω $f_k(x) = |e^{i\delta_k x} - 1|$ για κάθε $k \in \mathbb{N}$. Τότε

(α) $\lim\limits_{k\to\infty} f_k(x) = 0$ για κάθε $x \in \mathbb{R}$.

(β) $|f_k(x)| \le 2 =: g(x)$ για κάθε $x \in \mathbb{R}$.

(γ) $\displaystyle\int g(x)\,\mathrm{d}\mu(x) = 2 < \infty$.

Συνεπώς, από το θεώρημα κυριαρχημένης σύγκλισης, το δεξί μέλος της (13.4) τείνει στο 0, και έτσι προκύπτει το ζητούμενο. ■

13.2 Χαρακτηριστικές συναρτήσεις

Ορισμός 13.3. Έστω X τυχαία μεταβλητή σε χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$ με τιμές στον $\mathbb{R}$. **Χαρακτηριστική συνάρτηση** της X λέμε τη συνάρτηση $\phi_X : \mathbb{R} \to \mathbb{C}$ με

$$\phi_X(t) = \mathbf{E}(e^{itX}).$$

Από την Πρόταση 7.2, $\phi_X(t) = \displaystyle\int e^{itx}\,\mathrm{d}\,\mathbf{P}^X(x)$, δηλαδή $\phi_X = \widehat{\mathbf{P}^X}$.

Πρόταση 13.4. *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 X, Y ανεξάρτητες, τότε $\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t)$.

Απόδειξη. (i) $\phi_X(-t) = \mathbf{E}(e^{i(-t)X}) = \mathbf{E}\left(\overline{e^{itX}}\right) = \overline{\mathbf{E}(e^{itX})} = \overline{\phi_X(t)}$.

(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 10.9. ■

Στο επόμενο παράδειγμα, υπολογίζουμε τη χαρακτηριστική συνάρτηση τυχαίων μεταβλητών που ακολουθούν κάποια από τις γνωστές κατανομές.

Παράδειγμα 13.5. (i) Έστω $X \sim \mathbf{Bin}(n,p)$. 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 $X \sim \mathbf{Poisson}(\lambda)$, $\lambda > 0$. Τότε, $\phi_X(t) = e^{\lambda(e^{it}-1)}$ [αποδεικνύεται όμοια με το (i)].

(iii) Έστω $X \sim \mathbf{U}(-a,a)$. Τότε

$$\phi_X(t) = \begin{cases} \frac{\sin(at)}{at} & \text{αν } t \in \mathbb{R}\backslash\{0\}, \\ 1 & \text{αν } t = 0. \end{cases}$$

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

$$\begin{aligned}\phi_X(t) = \mathbf{E}(e^{itX}) &= \int_{-a}^{a} e^{itx}\frac{1}{2a}\,\mathrm{d}x \\ &= \frac{1}{2a}\int_{-a}^{a}\cos(tx)\,\mathrm{d}x + i\frac{1}{2a}\int_{-a}^{a}\sin(tx)\,\mathrm{d}x \\ &= \frac{1}{2a}\frac{\sin(tx)}{t}\bigg|_{-a}^{a} + 0 = \frac{1}{2a}\Big(\frac{\sin(ta)}{t} - \frac{\sin(-ta)}{t}\Big) = \frac{\sin(ta)}{ta}.\end{aligned}$$

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 $X \sim \mathbf{N}(0,1)$. 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 13.18 της Παραγράφου 13.6. Ένας άλλος, όχι και τόσο προφανής τρόπος, είναι ο εξής:

$$\begin{aligned}\phi_X(t) = \mathbf{E}(e^{itx}) &= \int_{-\infty}^{\infty} e^{itx}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\cos(tx)e^{-x^2/2}\,\mathrm{d}x + i\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\sin(tx)e^{-x^2/2}\,\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\cos(tx)e^{-x^2/2}\,\mathrm{d}x.\end{aligned}$$

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

$$\phi_X'(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(-x)\sin(tx)e^{-x^2/2}\,\mathrm{d}x.$$

Η παραγώγιση κάτω από το ολοκλήρωμα απαιτεί δικαιολόγηση την οποία παραλείπουμε. Ολοκληρώνοντας κατά μέρη, έχουμε

$$\phi_X'(t) = -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t\cos(tx) e^{-x^2/2}\,\mathrm{d}x = -t\phi_X(t).$$

Έστι καταλήγουμε στη συνήθη διαφορική εξίσωση $\phi_X'(t) = -t\phi_X(t)$, η οποία έχει γενική λύση

$$\phi_X(t) = Ce^{-t^2/2}.$$

Και εφόσον $\phi_X(0) = 1$, έχουμε ότι $C = 1$. Άρα $\phi_X(t) = e^{-t^2/2}$.

(v) Έστω $X \sim \mathbf{N}(\mu, \sigma^2)$. 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 $Z \sim \mathbf{N}(0, 1)$ και $X = \sigma Z + \mu$. Άρα, $\phi_X(t) = \phi_{\sigma Z+\mu}(t) = e^{it\mu}\phi_Y(t\sigma) = e^{it\mu}e^{-t^2\frac{\sigma^2}{2}}$.

(vi) Έστω $X \sim \mathbf{Exp}(\lambda)$, $\lambda > 0$. Τότε, $\phi_X(t) = \frac{\lambda}{\lambda - it}$ για κάθε $t \in \mathbb{R}$. Πράγματι,

$$\begin{aligned}
\phi_X(t) = \mathbf{E}(e^{itX}) &= \int_0^{\infty} e^{itx}\lambda e^{-\lambda x}\,\mathrm{d}x \\
&= \lambda \int_0^{\infty} e^{x(-\lambda+it)}\,\mathrm{d}x = \lim_{M\to\infty} \lambda \int_0^{M} e^{x(-\lambda+it)}\,\mathrm{d}x \\
&= \lim_{M\to\infty} \lambda \frac{e^{x(-\lambda+it)}}{-\lambda+it}\Bigg|_0^M = \lim_{M\to\infty} \lambda \frac{e^{M(-\lambda+it)} - 1}{-\lambda+it} \\
&= \frac{\lambda}{\lambda - it},
\end{aligned}$$

γιατί $\lim_{M\to\infty} e^{-\lambda M}e^{itM} = 0$ αφού $|e^{itM}| = 1$ και $\lambda > 0$.

13.3 Μετασχηματισμός Fourier στο $\mathbb{R}^n$

Έστω $n \geq 1$. Για $x, y \in \mathbb{R}^n$ με $x = (x_1, x_2, \ldots, x_n)$ και $y = (y_1, y_2, \ldots, y_n)$, το σύνηθες εσωτερικό γινόμενο των x και y ορίζεται ως:

$$\langle x, y\rangle := \sum_{k=1}^{n} x_k y_k.$$

Ορισμός 13.6. Έστω μ μέτρο πιθανότητας στον $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$. **Μετασχηματισμό Fourier** του μ ονομάζουμε τη συνάρτηση $\hat{\mu} : \mathbb{R}^n \to \mathbb{C}$ που ορίζεται ως

$$\hat{\mu}(u) := \int e^{i\langle u,x\rangle}\,\mathrm{d}\mu(x) = \int \cos(\langle u, x\rangle)\,\mathrm{d}\mu(x) + i\int \sin(\langle u, x\rangle)\,\mathrm{d}\mu(x)$$

για κάθε $u \in \mathbb{R}^n$.

Και σε αυτή την περίπτωση η συνάρτηση $\hat{\mu}$ είναι ομοιόμορφα συνεχής, έχει μέτρο φραγμένο από 1, και $\hat{\mu}(0) = 1$.

Ορισμός 13.7. Έστω X τυχαία μεταβλητή σε χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$ με τιμές στον $\mathbb{R}^n$. Χαρακτηριστική συνάρτηση της X λέμε τη συνάρτηση $\phi_X : \mathbb{R}^n \to \mathbb{C}$ με

$$\phi_X(u) = \mathbf{E}(e^{i\langle u,X\rangle}).$$

Το ανάλογο της Πρότασης 13.4 είναι το εξής.

Πρόταση 13.8. *Έστω X, Y τυχαίες μεταβλητές στον $(\Omega, \mathcal{F}, \mathbf{P})$ με τιμές στον $\mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$ και $b \in \mathbb{R}^m$. Τότε για κάθε $u \in \mathbb{R}^n, v \in \mathbb{R}^m$ έχουμε*

(i) $\phi_X(-u) = \overline{\phi_X(u)}$,

(ii) $\phi_{AX+b}(v) = e^{i\langle v,b\rangle}\phi_X(A^t v)$.

(iii) *Αν X, Y ανεξάρτητες, τότε* $\phi_{X+Y}(u) = \phi_X(u)\phi_Y(u)$.

A^t είναι ο ανάστροφος του πίνακα A.

Η απόδειξη της πρότασης αφήνεται ως άσκηση.

13.4 Θεώρημα μοναδικότητας και εφαρμογές

Στην παράγραφο αυτή βασιζόμαστε στο επόμενο θεώρημα, το οποίο είναι απόρροια του θεωρήματος μοναδικότητας του μετασχηματισμού Fourier και η απόδειξή του παραλείπεται γιατί ξεφεύγει από τα πλαίσια του σκοπού μας. Ο αναγνώστης μπορεί να αναζητήσει την απόδειξη του σε βιβλία Ανάλυσης Fourier ή Πιθανοτήτων [π.χ. Θεώρημα 14.1 στο Jacod and Protter (2003)].

Θεώρημα 13.9 (Θεώρημα Μοναδικότητας). *Έστω μ, ν μέτρα πιθανότητας στον $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$ ώστε $\hat{\mu}(u) = \hat{\nu}(u)$ για κάθε $u \in \mathbb{R}^n$. Τότε, $\mu = \nu$.*

Το θεώρημα μεταφέρει τον έλεγχο $\mu(A) = \nu(A)$ για κάθε $A \in \mathscr{B}(\mathbb{R}^n)$ στον $\hat{\mu}(u) = \hat{\nu}(u)$ για κάθε $u \in \mathbb{R}^n$, που είναι ένας έλεγχος πάνω σε αριθμούς.

Αναδιατύπωση του θεωρήματος είναι το ακόλουθο πόρισμα το οποίο αφορά τυχαίες μεταβλητές.

Πόρισμα 13.10. *Έστω X, Y τυχαίες μεταβλητές με τιμές στον $\mathbb{R}^n$. Αν $\phi_X(u) = \phi_Y(u)$ για κάθε $u \in \mathbb{R}^n$, τότε οι X, Y έχουν την ίδια κατανομή, δηλαδή $\mathbf{P}^X = \mathbf{P}^Y$.*

Απόδειξη. Επειδή $\phi_X(u) = \widehat{\mathbf{P}^X}(u)$, από το θεώρημα μοναδικότητας προκύπτει το ζητούμενο. ∎

Πόρισμα 13.11. *Έστω $X = (X_1, X_2, \dots, X_n)$ τυχαία μεταβλητή με τιμές στον $\mathbb{R}^n$. Τότε οι $\{X_j : 1 \le j \le n\}$ είναι ανεξάρτητες αν και μόνο αν $\phi_X(u_1, u_2, \dots, u_n) = \phi_{X_1}(u_1)\phi_{X_2}(u_2)\cdots\phi_{X_n}(u_n)$ για κάθε $(u_1, u_2, \dots, u_n) \in \mathbb{R}^n$.*

Απόδειξη. $\Rightarrow$ Έστω ότι οι $X_1, X_2, \dots, X_n$ είναι ανεξάρτητες. Τότε

$$\phi_X(u_1, u_2, \dots, u_n) = \mathbf{E}\left(e^{i\sum_{j=1}^n u_j X_j}\right) = \mathbf{E}(e^{iu_1X_1}e^{iu_2X_2}\cdots e^{iu_nX_n}).$$

Εφόσον οι $X_1, X_2, \dots, X_n$ είναι ανεξάρτητες, από προφανή γενίκευση της Πρότασης 10.8, οι $e^{iu_1X_1}, e^{iu_2X_2}, \dots, e^{iu_nX_n}$ είναι ανεξάρτητες (και προφανώς φραγμένες), άρα

$$\mathbf{E}(e^{iu_1X_1}e^{iu_2X_2}\cdots e^{iu_nX_n}) = \mathbf{E}(e^{iu_1X_1})\,\mathbf{E}(e^{iu_2X_2})\cdots\mathbf{E}(e^{iu_nX_n}),$$

από το οποίο προκύπτει το ζητούμενο.

$\Leftarrow$ Από την υπόθεση, για $u = (u_1, u_2, \dots, u_n)$ έχουμε

$$\begin{aligned}
\widehat{P^X}(u) = \phi_X(u) &= \phi_{X_1}(u_1)\phi_{X_2}(u_2)\cdots\phi_{X_n}(u_n)\\
&= \widehat{\mathbf{P}^{X_1}}(u_1)\widehat{\mathbf{P}^{X_2}}(u_2)\cdots\widehat{\mathbf{P}^{X_n}}(u_n)\\
&= \int e^{iu_1x_1}\,\mathrm{d}\,\mathbf{P}^{X_1}(x_1)\int e^{iu_2x_2}\,\mathrm{d}\,\mathbf{P}^{X_2}(x_2)\cdots\int e^{iu_nx_n}\,\mathrm{d}\,\mathbf{P}^{X_n}(x_n)\\
&= \int e^{i\sum_{j=0}^n u_jx_j}\,\mathrm{d}\,\mathbf{P}^{X_1}\otimes\mathbf{P}^{X_2}\otimes\ldots\otimes\mathbf{P}^{X_n}(x_1, x_2, \dots, x_n)\\
&= \mathbf{P}^{X_1}\otimes\mathbf{P}^{\widehat{X_2}\otimes}\ldots\otimes\mathbf{P}^{X_n}(u).
\end{aligned}$$

Συνεπώς, από το θεώρημα μοναδικότητας έχουμε ότι

$$\mathbf{P}^X = \mathbf{P}^{X_1} \otimes \mathbf{P}^{X_2} \otimes \ldots \otimes \mathbf{P}^{X_n},$$

το οποίο από την Πρόταση 10.13 ισοδυναμεί με το ζητούμενο. ■

Έστω X, Y τυχαίες μεταβλητές με τιμές στον ίδιο χώρο. Αν οι X και Y έχουν την ίδια κατανομή, δηλαδή $\mathbf{P}^X = \mathbf{P}^Y$, θα γράφουμε[1] $X \overset{d}{=} Y$.

Ορισμός 13.12. Έστω X τυχαία μεταβλητή με τιμές στο $\mathbb{R}$. Λέμε ότι η X έχει συμμετρική κατανομή αν $X \overset{d}{=} -X$, δηλαδή, για κάθε $A \in \mathscr{B}(\mathbb{R})$ ισχύει $\mathbf{P}(X \in A) = \mathbf{P}(-X \in A)$.

Παράδειγμα 13.13. Έστω X τυχαία μεταβλητή με τιμές στο $\mathbb{R}$ και πυκνότητα f άρτια συνάρτηση. Τότε η X έχει συμμετρική κατανομή. Πράγματι,

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

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

Παράδειγμα 13.14. (i) Έστω $X_1, X_2, \ldots, X_n$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές τέτοιες ώστε $X_1 \sim \mathbf{Bernoulli}(p)$. Αν $Y = \sum_{j=1}^{n} X_j$, τότε $Y \sim \mathbf{Bin}(n, p)$. Πράγματι, η χαρακτηριστική συνάρτηση καθεμίας από τις X_j ισούται με $\phi_{X_1}(t) = e^{it}p + 1 - p$ και από την Πρόταση 13.4(iii) έπεται ότι

$$\phi_Y(t) = \phi_{X_1}(t)^n = (e^{it}p + 1 - p)^n,$$

που είναι η χαρακτηριστική συνάρτηση της $\mathbf{Bin}(n, p)$. Το θεώρημα μοναδικότητας (Πόρισμα 13.10) δίνει ότι $Y \sim \mathbf{Bin}(n, p)$.

(ii) Έστω X, Y ανεξάρτητες τυχαίες μεταβλητές τέτοιες ώστε $X \sim \mathbf{Poisson}(\lambda)$ και $Y \sim \mathbf{Poisson}(\mu)$. Τότε για τη $Z \sim X + Y$ έχουμε ότι $Z \sim \mathbf{Poisson}(\lambda + \mu)$. 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 $\mathbf{Poisson}(\lambda + \mu)$, και από το θεώρημα μοναδικότητας η Z έχει κατανομή $\mathbf{Poisson}(\lambda + \mu)$.

(iii) Έστω $X \sim \mathbf{Bin}(n, p)$ και $Y \sim \mathbf{Bin}(m, p)$, ανεξάρτητες τυχαίες μεταβλητές. Τότε η $Z = X + Y$ έχει κατανομή $\mathbf{Bin}(n + m, p)$. Αυτό προκύπτει εύκολα με τη χρήση χαρακτηριστικών συναρτήσεων ή με χρήση του (i), αναπαριστώντας τις X, Y ως άθροισμα ανεξάρτητων τυχαίων μεταβλητών κατανομής $\mathbf{Bernoulli}(p)$.

(iv) Έστω $X \sim \mathbf{N}(\mu, \sigma^2)$ και $Y \sim \mathbf{N}(\nu, \tau^2)$ ανεξάρτητες τυχαίες μεταβλητές. Τότε, η $Z = X + Y$ έχει κατανομή $\mathbf{N}(\mu + \nu, \sigma^2 + \tau^2)$. 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 $\mathbf{N}(\mu + \nu, \sigma^2 + \tau^2)$ και από το θεώρημα μοναδικότητας προκύπτει το ζητούμενο.

[1] Το d από το αρχικό της λέξης distribution (κατανομή).

(v) Έστω $X \sim \mathbf{Gamma}(a_1, \lambda)$ και $Y \sim \mathbf{Gamma}(a_2, \lambda)$ ανεξάρτητες τυχαίες μεταβλητές. Τότε, δουλεύοντας όμοια με τα προηγούμενα, δείχνουμε ότι η $Z = X + Y$ έχει κατανομή $\mathbf{Gamma}(a_1 + a_2, \lambda)$.

Μια συνέπεια αυτού του αποτελέσματος είναι η εξής αναπαράσταση της κατανομής χι τετράγωνο με p βαθμούς ελευθερίας ($p \in \mathbb{N}^+$). Αν οι $Y_1, Y_2, \ldots, Y_p$ είναι ανεξάρτητες τυχαίες μεταβλητές με κατανομή $\mathbf{N}(0,1)$, τότε η

$$X := Y_1^2 + Y_2^2 + \ldots + Y_p^2$$

έχει κατανομή χ_p^2. Υπενθυμίζουμε ότι χ_p^2 είναι η κατανομή $\mathbf{Gamma}(\frac{p}{2}, \frac{1}{2})$.

Για την απόδειξη αυτού του ισχυρισμού, δείχνουμε με τη γνωστή τεχνική από τις στοιχειώδεις πιθανότητες ότι αν $Y \sim \mathbf{N}(0,1)$, τότε η $Y^2 \sim \mathbf{Gamma}(\frac{1}{2}, \frac{1}{2})$ (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.

13.5 Ροπογεννήτριες

Για μια τυχαία μεταβλητή X με τιμές στο $[-\infty, \infty]$ η ροπογεννήτριά της είναι η συνάρτηση $M_X : \mathbb{R} \to [0, \infty]$ 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 $t \le 0$ (αντίστοιχα, για κάθε $t \ge 0$), και μάλιστα $M_X(t) \le 1$ για εκείνα τα t. Πάντοτε $M_X(0) = 1$, ενώ το δεδομένο $M_X(t) < \infty$ για κάποιο $t \ne 0$ έχει χρήσιμες συνέπειες. Καταγράφουμε δύο από αυτές στο επόμενο λήμμα (δες επίσης την Άσκηση 6.7).

Λήμμα 13.15. *(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 $e^{tX} \le e^{\varepsilon X} + 1$ (παίρνουμε τις περιπτώσεις $X(\omega) \ge 0$ και $X(\omega) < 0$), άρα $M_X(t) < \infty$. Έπειτα, πάλι παίρνοντας περιπτώσεις, έχουμε $0 \le \varepsilon^k (X^+)^k \le k!\, e^{\varepsilon X}$ 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 7.9), 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 τ.μ., δεν την κωδικοποιεί.

Μελετούμε τώρα την περίπτωση που το D_X περιέχει ένα ανοιχτο διάστημα $(-\varepsilon, \varepsilon)$ γύρω από το 0.

Πρόταση 13.16. *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

$$M_X(t) = \sum_{k=0}^{\infty} \frac{\mathbf{E}(X^k)}{k!} t^k \tag{13.5}$$

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

$$\mathbf{E}\left(\sum_{k=0}^{\infty} \frac{|t^k X^k|}{k!}\right) = \mathbf{E}(e^{|tX|}) < \mathbf{E}(e^{-tX}) + \mathbf{E}(e^{tX}) < \infty.$$

(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 «απόδειξή» του. Στην $M_X(t) = \mathbf{E}(e^{tX})$ παραγωγίζουμε k φορές και παίρνουμε

$$M_X^{(k)}(t) = \mathbf{E}(X^k e^{tX}). \tag{13.6}$$

Δηλαδή περνάμε την παράγωγο μέσα από τη μέση τιμή. Το ότι αυτό είναι σωστό αποδεικνύεται με χρήση του θεωρήματος κυριαρχημένης σύγκλισης, αλλά το παραλείπουμε. Έπειτα θέτουμε $t = 0$ στην (13.6).

Αντιπαραβάλλουμε τη χαρακτηριστική συνάρτηση με τη ροπογεννήτρια μιας τυχαίας μεταβλητής. Γράφουμε (+) στα προτερήματα και (−) στα ελαττώματα.

Η χαρακτηριστική συνάρτηση:

(i) Είναι πάντοτε πεπερασμένος αριθμός. (+)

(ii) Ο υπολογισμός της ενδέχεται να εμπλέκει ολοκληρώματα μιγαδικών συναρτήσεων. (−)

(iii) Χαρακτηρίζει την κατανομή της X. Δύο τ.μ. με ίδια χαρακτηριστική συνάρτηση έχουν την ίδια κατανομή (Πόρισμα 13.10). (+)

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 τ.μ. 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. (+)

13.6 Χαρακτηριστικές συναρτήσεις μέσω ροπογεννητριών*

Έχοντας υπολογίσει κανείς την ροπογεννήτρια M_X της X είναι δελεαστικό να προσπαθήσει να υπολογίσει τη χαρακτηριστική συνάρτηση ως

$$\phi_X(t) = \mathbf{E}(e^{itX}) \doteq M_X(it).$$

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 $t \to it$ δίνει $e^{t^2/2}$, που είναι λάθος. Τι καλύτερο έχει ο τύπος $e^{t^2/2}$ από τον $e^{|t|^2/2}$;

Πρόταση 13.17. *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 $h : \{z \in \mathbb{C} : |Re(z)| < \varepsilon\} \to \mathbb{C}$ 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 $A_\varepsilon := \{z \in \mathbb{C} : |\mathrm{Re}(z)| < \varepsilon\}$. Θέτουμε $g : A_\varepsilon \to \mathbb{C}$ με $g(z) := \mathbf{E}(e^{zX})$ για κάθε $z \in A_\varepsilon$.

Ισχυρισμος: Η g είναι καλά ορισμένη[2] 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 $|z| < \varepsilon - |\mathrm{Re}(z_0)|$ ισχύει

$$g(z_0 + z) = \mathbf{E}(e^{z_0 X} e^{zX}) = \mathbf{E}\left\{ e^{z_0 X} \sum_{k=0}^{\infty} \frac{(zX)^k}{k!} \right\} = \sum_{k=0}^{\infty} \frac{\mathbf{E}\left\{X^k e^{z_0 X}\right\}}{k!} z^k. \tag{13.7}$$

Χρειάζεται δικαιολόγηση μόνο η τελευταία ισότητα. Δηλαδή η αλλαγή σειράς μέσης τιμής και άθροισης. Αυτό έπεται από το θεώρημα Fubini αφού

$$\mathbf{E}\left\{ \sum_{k=0}^{\infty} \left| e^{z_0 X} \frac{(zX)^k}{k!} \right| \right\} = \mathbf{E}\left\{ e^{\mathrm{Re}(z_0)X + |zX|} \right\} \le \mathbf{E}\left(e^{|X|(|z| + |\mathrm{Re}(z_0)|)} \right) < \infty.$$

Το ότι η τελευταία ποσότητα είναι πεπερασμένη έπεται από το ότι $|z| + |\mathrm{Re}(z_0)| < \varepsilon$ και την υπόθεση (i). Εδώ λοιπον είναι κρίσιμη η υπόθεση ότι η M_X είναι πεπερασμένη στο $(-\varepsilon, \varepsilon)$. Επίσης συμπεραίνουμε ότι στο δεξί μέλος της (13.7) 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 $\varepsilon - |\mathrm{Re}(z_0)| > 0$, 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 13.18. (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 $h : \mathbb{C} \to \mathbb{C}$ με $h(z) = e^{z^2/2}$ για κάθε $z \in \mathbb{C}$ είναι

[2] Το καλά ορισμένη σημαίνει ότι η μέση τιμή μπορεί να οριστεί και είναι στοιχείο του $\mathbb{C}$. Δεν εμφανίζεται κάποια μορφή $\infty - \infty$.

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 13.17 εφαρμόζεται και δίνει ότι $\phi_X(t) = h(it) = e^{-t^2/2}$ για κάθε $t \in \mathbb{R}$.

(ii) Μια $X \sim \mathbf{Exp}(\lambda)$ ($\lambda > 0$) 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 $h : \mathbb{C}\setminus\{\lambda\} \to \mathbb{C}$ με $h(z) = \lambda/(\lambda - z)$ είναι αναλυτική στο πεδίο ορισμού της [το οποίο περιέχει μια λωρίδα της μορφής $\{z : |\mathrm{Re}(z)| < \varepsilon\}$ με $\varepsilon > 0$, π.χ. με $\varepsilon = \lambda$.] και συμφωνεί με τη M_X στο $(-\infty, \lambda)$. Άρα η Πρόταση 13.17 εφαρμόζεται και δίνει ότι $\phi_X(t) = h(it) = \lambda/(\lambda - it)$ για κάθε $t \in \mathbb{R}$.

Συνέπεια της απόδειξης της Πρότασης 13.17 είναι το εξής θεώρημα μοναδικότητας για ροπογεννήτριες. Το διατυπώνουμε μόνο για τυχαίες μεταβλητές με τιμές στο $\mathbb{R}$.

Θεώρημα 13.19 (Θεώρημα μοναδικότητας για ροπογεννήτριες)**.** *Έστω X, Y τυχαίες μεταβλητές με τιμές στον $\mathbb{R}$. Αν*

(i) υπάρχει $\varepsilon > 0$ ώστε 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 13.17 δίνουν ότι οι συναρτήσεις $g(z) = \mathbf{E}(e^{zX}), h(z) = \mathbf{E}(e^{zY})$ είναι αναλυτικές στο $A_\varepsilon := \{z \in \mathbb{C} : |\mathrm{Re}(z)| < \varepsilon\}$. 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 13.10 (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 13.10) 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).

13.7 Άθροισμα ανεξάρτητων τυχαίων μεταβλητών

Ορισμός 13.20. 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 $A \in \mathscr{B}(\mathbb{R})$.

Παρατήρηση 13.21. Εύκολα βλέπουμε ότι η συνέλιξη είναι συμμετρική, δηλαδή $\mu * \nu = \nu * \mu$. Επίσης, ισχύει ότι

$$\mu * \nu(A) = \int \int \mathbf{1}_{(A-y)}(x)\, d\mu(x)\, d\nu(y) = \int \mu(A - y)\, d\nu(y) = \int \nu(A - x)\, d\mu(x).$$

Θεώρημα 13.22. *Έστω X, Y ανεξάρτητες τυχαίες 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 10.13). Για $A \in \mathscr{B}(\mathbb{R})$ έχουμε

$$\begin{aligned}\mathbf{P}^{X+Y}(A) &= \mathbf{P}(X+Y \in A) = \mathbf{E}\{\mathbf{1}_A(X+Y)\}\\ &= \int \mathbf{1}_A(x+y)\,\mathrm{d}(\mathbf{P}^X \otimes \mathbf{P}^Y)(x,y) = \int\int \mathbf{1}_A(x+y)\,\mathrm{d}\,\mathbf{P}^X(x)\,\mathrm{d}\,\mathbf{P}^Y(y)\\ &= \mathbf{P}^X * \mathbf{P}^Y(A).\end{aligned}$$

Στην πρώτη ισότητα της δεύτερης γραμμής χρησιμοποιήσαμε την Πρόταση 7.2 για τη συνάρτηση $g : \mathbb{R}^2 \to \mathbb{R}$ με $g(x, y) = \mathbf{1}_A(x + y)$ και την τυχαία μεταβλητή (X, Y). ■

Θεώρημα 13.23. *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

$$f_Z(z) = \int f_X(z-y)\,\mathrm{d}\,\mathbf{P}^Y(y)$$

για κάθε $z \in \mathbb{R}$.

(ii) Αν οι X, Y έχουν αντίστοιχα πυκνότητες f_X, f_Y, τότε η

$$f_Z(z) = \int f_X(z-y) f_Y(y)\,\mathrm{d}y = \int f_X(x) f_Y(z-x)\,\mathrm{d}x$$

για κάθε $z \in \mathbb{R}$ είναι πυκνότητα της Z.

Απόδειξη. (i) Έστω $A \in \mathscr{B}(\mathbb{R})$. Τότε, από το Θεώρημα 13.22,

$$\begin{aligned}\mathbf{P}(Z \in A) &= \int \mathbf{P}^X(A-y)\,\mathrm{d}\,\mathbf{P}^Y(y) = \int\int_{A-y} f_X(x)\,\mathrm{d}x\,\mathrm{d}\,\mathbf{P}^Y(y)\\ &\overset{x=z-y}{=} \int\int_A f_X(z-y)\,\mathrm{d}z\,\mathrm{d}\,\mathbf{P}^Y(y) = \int_A\int f_X(z-y)\,\mathrm{d}\,\mathbf{P}^Y(y)\,\mathrm{d}z,\end{aligned}$$

άρα η $f_Z(z) = \displaystyle\int f_X(z-y)\,\mathrm{d}\,\mathbf{P}^Y(y)$ είναι πυκνότητα της Z.

(ii) Έστω $z \in \mathbb{R}$. Τότε, από το (i) και την Πρόταση 7.8, έχουμε

$$f_Z(z) = \int f_X(z-y)\,\mathrm{d}\,\mathbf{P}^Y(y) = \int f_X(z-y) f_Y(y)\,\mathrm{d}y.$$

Η δεύτερη ισότητα στην εκφώνηση προκύπτει με μια απλή αλλαγή μεταβλητής. ■

Το προηγούμενο θεώρημα συμπληρώνει την τεχνική προσδιορισμού κατανομής αθροίσματος που είδαμε στο Παράδειγμα 13.14. Το θεώρημα είναι χρήσιμο όταν η χαρακτηριστική συνάρτηση του αθροίσματος δεν είναι κάποια από τις γνωστές χαρακτηριστικές συναρτήσεις. Μια τέτοια περίπτωση περιγράφεται στην Άσκηση 13.12.

Ασκήσεις

13.1 Να δειχθεί 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 ότι για την τυχαία μεταβλητή X με τιμές στο $\mathbb{R}$ και ένα $t_0 \neq 0$ ισχύει $|\phi_X(t_0)| = 1$. Να δειχθεί 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 $a = 0$.

13.3 Έστω X τυχαία μεταβλητή με τιμές στο $\mathbb{R}$ ώστε $\mathbf{E}|X| < \infty$. Να δειχθεί ότι η ϕ_X είναι παραγωγίσιμη στο $\mathbb{R}$ και $\phi'_X(0) = i\,\mathbf{E}(X)$.
[Υπόδειξη: Θεώρημα κυριαρχημένης σύγκλισης.]

13.4 Έστω $X \sim \Gamma(a, \lambda)$. Δηλαδή η X έχει πυκνότητα

$$f_X(x) = \frac{\lambda^a}{\Gamma(a)}\, x^{a-1} e^{-\lambda x}\, \mathbf{1}_{x \geq 0}.$$

Να δειχθεί ότι η X έχει χαρακτηριστική συνάρτηση

$$\phi_X(t) = \frac{1}{\left(1 - \frac{it}{\lambda}\right)^a}$$

για κάθε $t \in \mathbb{R}$.

13.5 Έστω $a > 0$ και X τυχαία μεταβλητή με πυκνότητα

$$f(x) = \frac{a}{2|x|^{a+1}}\, \mathbf{1}_{|x| \geq 1}.$$

Να δειχθεί ότι $M_X(t) = \infty$ για κάθε $t \neq 0$.

13.6 Έστω X τυχαία μεταβλητή με κατανομή $N(0, 1)$ και $Y := e^X$. Να δειχθεί ότι

(α) $\mathbf{E}(Y^k) < \infty$ για κάθε $k \in \mathbb{N}$.

(β) $M_Y(t) < \infty$ αν και μόνο αν $t \leq 0$.

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 Λέμε ότι η τυχαία μεταβλητή X έχει κατανομή **Cauchy** αν έχει πυκνότητα $f(x) = \frac{1}{\pi(1+x^2)}$, $x \in \mathbb{R}$. Θεωρούμε γνωστό για αυτή την άσκηση ότι $\phi_X(u) = e^{-|u|}$ για κάθε $u \in \mathbb{R}$. Να δείξετε ότι:

(α) Αν $X, Y \sim$ **Cauchy** ανεξάρτητες, τότε $\frac{X+Y}{2} \sim$ **Cauchy**.

(β) Αν $X_1, X_2, \ldots, X_n \sim$ **Cauchy** ανεξάρτητες, τότε $\frac{X_1+X_2+\ldots+X_n}{n} \sim$ **Cauchy**.

13.10 Έστω X τυχαία μεταβλητή όπως στην Άσκηση 13.5 με $a \in (0, 2)$. Να δειχθεί ότι υπάρχει σταθερά $C(a) \in (0, \infty)$ ώστε

$$\lim_{t \to 0} \frac{1 - \phi_X(t)}{t^a} = C(a). \tag{13.8}$$

13.11* Έστω X τυχαία μεταβλητή με πυκνότητα $f(x) = |x|^{-3}\mathbf{1}_{|x| \geq 1}$. Να δειχθεί ότι για κάθε $t \in [-1, 1]$ ισχύει

$$|\phi_X(t) - 1 - t^2 \log|t|| \leq 3t^2. \tag{13.9}$$

13.12 Έστω $X, Y \sim \mathbf{U}(0, 1)$ ανεξάρτητες τυχαίες ματαβλητές. Να δείξετε ότι η $Z = X + Y$ έχει πυκνότητα

$$f_Z(z) = \begin{cases} z & \text{αν } z \in (0, 1], \\ 2 - z & \text{αν } z \in (1, 2), \\ 2 & \text{αν } z \in \mathbb{R} \setminus (0, 2). \end{cases}$$

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 στο $\mathbb{R}$. Λέμε ότι η $(\mu_n)_{n\geq 1}$ **συγκλίνει ασθενώς** στο μ αν

$$\mu_n((-\infty, x]) \to \mu((-\infty, x])$$

καθώς $n \to \infty$ 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 συγκίνει 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, άρα είναι πυκνό. Και επειδή οι συναρτήσεις κατανομής είναι δεξιά συνεχείς, έπεται ότι ισούνται σε όλο το $\mathbb{R}$. Το Θεώρημα 4.10 δίνει ότι $\mu = \nu$.

Ορισμός 14.2. Έστω $(X_n)_{n\geq 1}, X$ τυχαίες μεταβλητες με τιμές στο $\mathbb{R}$. Λέμε ότι η $(X_n)_{n\geq 1}$ *συγκλίνει κατά κατανομή* στη X και γράφουμε[1]

$$X_n \Rightarrow X \text{ ή } X_n \xrightarrow{\mathrm{d}} X \text{ ή } X_n \xrightarrow{\mathscr{L}} X,$$

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*

$$F_{X_n}(x) \to F_X(x)$$

καθώς $n \to \infty$ 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 4.1). Δηλαδή είναι λίγα. Σε οποιοδήποτε διάστημα θετικού μήκους μπορούμε να βρούμε σημείο συνέχειας της F.

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 $X_n(\omega) - X(\omega)$.

[1] d από το distribution, και $\mathscr{L}$ από το law.

(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 L^p, κατά πιθανότητα). Το Θεώρημα 14.12 πιο κάτω αφορά αυτό το ερώτημα.

Παράδειγμα 14.5. Έστω $(p_n)_{n\geq 1}$ ακολουθία στο $(0,1)$ έτσι ώστε $p_n \to 0$ και $(X_n)_{n\geq 1}$ ακολουθία τυχαίων μεταβλητών έτσι ώστε $X_n \sim$ **Γεωμετρική**(p_n) για κάθε $n \geq 1$. Θα δείξουμε ότι

$$p_n X_n \Rightarrow X,$$

όπου X τυχαία μεταβλητή τέτοια ώστε $X \sim \mathbf{Exp}(1)$.

Η F_X είναι συνεχής στο $\mathbb{R}$. Συγκεκριμένα, έχουμε ότι

$$F_X(x) = \int_{-\infty}^{x} e^{-t}\mathbf{1}_{t>0}\,\mathrm{d}t = \begin{cases} 0 & \text{αν } x \leq 0, \\ 1 - e^{-x} & \text{αν } x > 0. \end{cases}$$

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 $x > 0$,

$$\begin{aligned} F_{p_n X_n}(x) &= \mathbf{P}\left(X_n \leq \frac{x}{p_n}\right) = 1 - \mathbf{P}\left(X_n > \frac{x}{p_n}\right) \\ &= 1 - \mathbf{P}\left(X_n > \left[\frac{x}{p_n}\right]\right) = 1 - (1-p_n)^{\left[\frac{x}{p_n}\right]}, \end{aligned}$$

εφόσον για $Y \sim$ **Γεωμετρική**(p) ισχύει $\mathbf{P}(Y \geq k) = (1-p)^{k-1}$. Τώρα παρατηρούμε ότι

$$(1-p_n)^{\left[\frac{x}{p_n}\right]} = e^{\left[\frac{x}{p_n}\right]p_n \frac{\log(1-p_n)}{p_n}} \to e^{-x}$$

για $n \to \infty$. Άρα, $F_{p_n X_n}(x) \to 1 - e^{-x}$. 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 πιό πυκνό τρόπο καθώς $n \to \infty$. Κατά μία 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 $W_n := \max\{X_1, X_2, \ldots, X_n\}$ και $Z_n := W_n - \log n$. Θα δείξουμε ότι

$$W_n - \log n \Rightarrow Z \tag{14.1}$$

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

$$\begin{aligned} F_{W_n - \log n}(t) &:= \mathbf{P}(W_n - \log n \leq t) = \mathbf{P}(\max\{X_1, X_2, \ldots, X_n\} \leq t + \log n) \\ &= \mathbf{P}(X_1 \leq t + \log n, \ldots, X_n \leq t + \log n) \\ &= \mathbf{P}(X_1 \leq t + \log n)^n = (1 - e^{-t-\log n})^n = \left(1 - \frac{e^{-t}}{n}\right)^n. \end{aligned}$$

Στην τέταρτη ισότητα χρησιμοποιούμε ότι οι X_i είναι ανεξάρτητες και ισόνομες. Στην πέμπτη ι- σότητα ότι $t + \log n > 0$ και ότι η X_1 ακολουθεί την εκθετική κατανομή με παράμετρο 1. Άρα $\lim_{n\to\infty} F_{W_n - \log n}(t) = e^{-e^{-t}} = F_Z(t)$ για κάθε $t \in \mathbb{R}$, που αποδεικνύει το ζητούμενο.

Αυτό που μας λέει η σύγκλιση (14.1) 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 τ.μ.] Συνεχίζοντας από την Παρατήρηση 7.4, στο προηγούμενο παράδειγμα, αφού ολες οι $X_1, X_2, \ldots, X_n$ έχουν την ίδια κατανομή με τη X_1, γιατί στην ποσότητα

$$\mathbf{P}(X_1 \leq t + \log n, \ldots, X_n \leq t + \log n)$$

δεν αντικαθιστούμε όλες τις τ.μ. 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 7.2 γιατί, για να είναι η χρήση της σωστή, θα έπρεπε τα δύο *διανύσματα* $(X_1, X_2, \ldots, X_n)$ και $(X_1, X_1, \ldots, X_1)$ 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*

$$\int f(x)\, d\mu_n(x) \to \int f(x)\, d\mu(x) \tag{14.2}$$

για κάθε $f : \mathbb{R} \to \mathbb{R}$ συνεχή και φραγμένη.

Το ίδιο θεώρημα, με όρους τυχαίων μεταβλητών, γράφεται ως εξής.

Θεώρημα 14.9. *Έστω $(X_n)_{n\geq 1}$ τυχαίες μεταβλητές σε χώρους πιθανότητας $(\Omega_n, \mathcal{F}_n, \mathbf{P}_n)$, $n \geq 1$ και X τυχαία μεταβλητή σε χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$ με τιμές στο $\mathbb{R}$. Τότε $X_n \Rightarrow X$ αν και μόνο αν*

$$\mathbf{E}(f(X_n)) \to \mathbf{E}(f(X)) \tag{14.3}$$

για κάθε $f : \mathbb{R} \to \mathbb{R}$ συνεχή και φραγμένη.

Η μέση τιμή στο αριστερό μέλος της (14.3) 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 $\{1, 2, \ldots, n\}$. Θα δείξουμε ότι

$$\frac{1}{n} X_n \Rightarrow U \tag{14.4}$$

καθώς $n \to \infty$, όπου U είναι τυχαία μεταβλητή που ακολουθεί την ομοιόμορφη κατανομή στο $(0, 1)$.

Για $f : \mathbb{R} \to \mathbb{R}$ συνεχή και φραγμένη και $n \to \infty$ έχουμε

$$\mathbf{E}\left\{f\left(\frac{X_n}{n}\right)\right\} = \sum_{k=1}^{n} \mathbf{P}(X_n = k) f\left(\frac{k}{n}\right) = \frac{1}{n} \sum_{k=1}^{n} f\left(\frac{k}{n}\right) \to \int_0^1 f(x)\, dx = \mathbf{E}(f(U))$$

Η σύγκλιση ισχύει γιατί έχουμε ένα άθροισμα Riemann για την f στο $[0, 1]$. Η τελευταία ισότητα ισχύει γιατί η U είχει πυκνότητα $1_{(0,1)}$.

Η σύγκλιση (14.4) είναι αναμενόμενη αφού η κατανομή της X_n/n δίνει μάζα $1/n$ σε καθένα από τα σημεία $\{1/n, 2/n, \ldots, n/n\}$. Η μάζα ισομοιράζεται και τελικά, καθώς $n \to \infty$, κάθε υποδιάστημα του $[0, 1]$ παίρνει μάζα ανάλογη προς το μέγεθός του.

Δίνουμε ακόμη ένα χαρακτηρισμό της σύγκλισης κατά κατανομή. Αυτός είναι χρήσιμος στις εφαρμογές.

Θεώρημα 14.11. *Έστω $(X_n)_{n\geq 1}$ τυχαίες μεταβλητές σε χώρους πιθανότητας $(\Omega_n, \mathcal{F}_n, \mathbf{P}_n)$, $n \geq 1$ και X τυχαία μεταβλητή σε χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$ με τιμές στον $\mathbb{R}$. Τότε τα εξής είναι ισοδύναμα.*

(i) $X_n \Rightarrow X$

(ii) Για κάθε $A \in \mathscr{B}(\mathbb{R})$ με $\mathbf{P}(X \in \partial A) = 0$ ισχύει

$$\lim_{n\to\infty} \mathbf{P}_n(X_n \in A) = \mathbf{P}(X \in A).$$

Απόδειξη. (ii) ⇒ (i). Αυτή η κατεύθυνση είναι εύκολη. Αν το x_0 είναι σημείο συνέχειας της F, τότε το σύνολο $A = (-\infty, x_0]$ έχει $\partial A = \{x_0\}$ και $\mathbf{P}(X \in \{x_0\}) = F(x_0) - F(x_0-) = 0$. Άρα

$$F_{X_n}(x_0) = \mathbf{P}(X_n \in A) \to \mathbf{P}(X \in A) = F(x_0)$$

καθώς $n \to \infty$.

(i) ⇒ (ii). Δίνεται στο Παράρτημα Β′. ■

Το επόμενο θεώρημα δείχνει ότι η σύγκλιση κατά κατανομή είναι η ασθενέστερη μορφή σύγκλισης τυχαίων μεταβλητών από όσες έχουμε δει ως τώρα.

Θεώρημα 14.12. *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 $f : \mathbb{R} \to \mathbb{R}$ συνεχής και φραγμένη, $\varepsilon > 0$, και υπακολουθία $(X_{n_k})_{k\in\mathbb{N}}$ της $(X_n)_{n\geq 1}$ έτσι ώστε:

$$|\mathbf{E}(f(X_{n_k})) - \mathbf{E}(f(X))| \geq \varepsilon. \tag{14.5}$$

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 $X_{n_{k_r}} \xrightarrow{\sigma.\beta.} X$. Αφού η f είναι συνεχής, από το Θεώρημα 8.6 έχουμε ότι $f(X_{n_{k_r}}) \xrightarrow{\sigma.\beta.} f(X)$. 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,

$$\mathbf{E}(f(X_{n_{k_r}})) \to \mathbf{E}(f(X)),$$

το οποίο συγκρούεται με την (14.5). 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 14.13. *Έστω $(X_n)_{n\geq 1}$ ακολουθία τυχαίων μεταβλητών ορισμένες 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

$$\begin{aligned}\mathbf{P}(|X_n - C| > \varepsilon) &= \mathbf{P}(X_n > C + \varepsilon) + \mathbf{P}(X_n < C - \varepsilon)\\ &= 1 - F_{X_n}(C + \varepsilon) + \mathbf{P}(X_n < C - \varepsilon)\\ &\leq 1 - F_{X_n}(C + \varepsilon) + F_n(C - \varepsilon).\end{aligned}$$

Τα $C - \varepsilon$, $C + \varepsilon$ είναι σημεία συνέχειας της F_C ($F_C(x) = \mathbf{1}_{[C,\infty)}(x)$), άρα από το Θεώρημα 14.3,

$$\lim_{n\to\infty} F_{X_n}(C - \varepsilon) = F_C(C - \varepsilon) = 0$$

και

$$\lim_{n\to\infty} F_{X_n}(C + \varepsilon) = F_C(C + \varepsilon) = 1.$$

Συνεπώς $\lim_{n\to\infty} \mathbf{P}(|X_n - C| > \varepsilon) = 0$. ■

14.2 Σφιχτότητα και συμπάγεια

Ορισμός 14.14. Μια οικογένεια $\{\mu_i : i \in I\}$ μέτρων πιθανότητας στο $\mathbb{R}$ λέγεται **σφιχτή** αν για κάθε $\varepsilon > 0$ υπάρχει $M > 0$ έτσι ώστε

$$\mu_i(\mathbb{R}\setminus[-M, M]) < \varepsilon$$

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:

$$\lim_{M\to\infty} \sup_{i\in I} \mu_i(\mathbb{R}\setminus[-M, M]) = 0.$$

Παρατήρηση 14.15. Αν μ 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).

Ορισμός 14.16. 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.

Επειδή $\mathbf{P}^{X_i}(\mathbb{R}\setminus[-M, M]) = \mathbf{P}(|X_i| > M)$, η οικογένεια $\{\mathbf{P}^{X_i} : i \in I\}$ είναι σφιχτή αν και μόνο αν

$$\lim_{M\to\infty} \sup_{i\in I} \mathbf{P}(|X_i| > M) = 0.$$

Παρατήρηση 14.17. Έστω $(X_n)_{n\geq1}$ ακολουθία τυχαίων μεταβλητών έτσι ώστε $X_n \sim \mathbf{Exp}(\frac{1}{n})$. Τότε η $(X_n)_{n\geq1}$ δεν είναι σφιχτή. Πράγματι, για $M > 0$,

$$\sup_{n\geq1} \mathbf{P}(|X_n| > M) = \sup_{n\geq1} e^{-\frac{M}{n}} = 1.$$

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 $n \to \infty$ αυτό το σημείο συγκέντρωσης απομακρύνεται. Δεν μπορούμε να βρούμε ένα φραγμένο σύνολο ώστε *όλες* οι X_n να πέφτουν εκεί με πιθανότητα κοντά στο 1.

Η έννοια της σφιχτότητας στα μέτρα πιθανότητας είναι ανάλογη της έννοιας της σχετικής συμπάγειας σε μετρικό χώρο. Οι όροι μιας συγκλίνουσας ακολουθίας σε μετρικό χώρο ορίζουν ένα σχετικά συμπαγές σύνολο. Το αντίστοιχο εδώ είναι το ακόλουθο αποτέλεσμα.

Πρόταση 14.18. *Έστω $(\mu_n)_{n\geq1}$, 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 $\mu(\{-M\}) = \mu(\{M\}) = 0$. Από την υπόθεση έχουμε

$$\mu_n(\mathbb{R}\setminus[-M, M]) = \mu_n((-\infty, -M]) + 1 - \mu_n((-\infty, M]) \to \mu(\mathbb{R}\setminus[-M, M])(< \varepsilon/2)$$

για $n \to \infty$. 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 $\mu_1, \mu_2, \ldots, \mu_{n_0-1}$ υπάρχει $\tilde{M} > 0$ έτσι ώστε $\mu_i(\mathbb{R}\setminus[-\tilde{M}, \tilde{M}]) < \varepsilon$ για κάθε $i = 1, 2, \ldots, n_0 - 1$ (Άσκηση 14.8).

Έστω $L = \max\{M, \tilde{M}\}$. Τότε $\mu_n(\mathbb{R}\setminus[-L, L]) < \varepsilon$ για κάθε $n \geq 1$. ■

Το ανάλογο του ότι μια ακολουθία σε έναν συμπαγή μετρικό χώρο έχει συγκλίνουσα υπακολουθία 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 14.19 (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 $\mathbb{R}$.*

Ασκήσεις

14.1 Να αποδειχθεί η (14.4) με χρήση του Θεωρήματος 14.3.

14.2 (α) Έστω X τυχαία μεταβλητή με κατανομή την τυπική Cauchy, δηλαδή με πυκνότητα $f(x) = \pi^{-1}(1+x^2)^{-1}$. Να δειχθεί ότι για κάθε $x > 0$ ισχύει

$$\frac{1}{1+x} < \pi\,\mathbf{P}(X > x) < \frac{1}{x}. \tag{14.6}$$

(β) Έστω $(X_n)_{n\geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών, καθεμία με κατανομή της τυπική 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

$$\lim_{n\to\infty} \mathbf{P}(X_n \in A) = \mathbf{P}(X \in A);$$

	Κατανομή της X	Σύνολο A
(i)	$Poisson(2)$	$(2, 32.1) \cup \{100\}$
(ii)	$Poisson(2)$	$\mathbb{Q}$
(iii)	Γεωμετρική$(1/3)$	$(-1.5, 2.8)$
(iv)	$N(0,1)$	$(-2, \pi)$
(v)	$U(0,1)$	$(0, 1/3)\backslash\mathbb{Q}$
(vi)	Bernouli$(2/5)$ στο $\{0,1\}$	$(0, 1/2) \cup (2, 4)$

14.4 (Αυτή η άσκηση δείχνει πώς αντιμετωπίζουμε όρια της μορφής $\lim_{n\to\infty} \mathbf{P}(X_n \in A_n)$ 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 $\lim_{n\to\infty} \mathbf{P}(X_n \leq n) = 1$.
(β) Για κάθε $x \in \mathbb{R}$ να δειχθεί ότι $\overline{\lim}_{n\to\infty} \mathbf{P}(X_n \in [x - n^{-1}, x + n^{-1}]) \leq \mathbf{P}(X = x)$. 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 $X_n Y_n \Rightarrow 0$.

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 $Y_n \to c$ κατά πιθανότητα, τότε $X_n + Y_n \Rightarrow X + c$.
(β) Αν $X_n \Rightarrow X$ και $Y_n \to c$ κατά πιθανότητα, τότε $X_n Y_n \Rightarrow cX$.
Παρατηρήστε ότι το (α) έχει την εξής άμεση συνέπεια
(γ) Αν $X_n \Rightarrow X$ και $Y_n - X_n \to 0$ κατά πιθανότητα, τότε $Y_n \Rightarrow X$.
Για το (β), για αποφυγή μη ουσιαστικών λεπτομερειών, υποθέστε ότι $c > 0$ και ότι οι Y_n παίρνουν τιμές στο $[0, \infty)$.

14.7 Έστω p μέτρο πιθανότητας στο $\mathbb{R}$. Να δείξετε ότι για κάθε $\varepsilon > 0$ υπάρχει $M > 0$ έτσι ώστε $p(\mathbb{R}\backslash[-M, M]) < \varepsilon$.

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 σφιχτή.

14.9 Έστω $(X_n)_{n\geq 1}$ ακολουθία τυχαίων μεταβλητών και $h : [0,\infty) \to [0,\infty)$ αύξουσα με $\lim_{x\to\infty} h(x) = \infty$ έτσι ώστε $\sup_{n\geq 1} \mathbf{E}\{h(|X_n|)\} < \infty$. Να δείξετε ότι η $(X_n)_{n\geq 1}$ είναι σφιχτή.

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 $\phi_{X_n}(t) \to \phi_X(t)$.

Καθένας από τους τρείς αυτούς τρόπους λειτουργεί καλά σε διαφορετικές περιτώσεις. 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 Lévy

Λήμμα 15.1. *Έστω μ μέτρο πιθανότητας στο $\mathbb{R}$ και $\hat{\mu}$ ο μετασχηματισμός Fourier του. Τότε*

$$\mu\left(\left\{x : |x| > \frac{2}{u}\right\}\right) \leq \frac{1}{u}\int_{-u}^{u}(1 - \hat{\mu}(t))\,\mathrm{d}t \tag{15.1}$$

για κάθε $u > 0$.

Η απόξειξή 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 Lévy) Έστω $(\mu_n)_{n\geq 1}$ ακολουθία μέτρων πιθανότητας στο $\mathbb{R}$ και $(\hat{\mu}_n)_{n\geq 1}$ η ακολουθία μετασχηματισμών Fourier τους.*

(i) Αν μ μέτρο πιθανότητας στο $\mathbb{R}$ έτσι ώστε $\mu_n \Rightarrow \mu$, τότε $\lim_{n\to\infty} \hat{\mu}_n(t) = \hat{\mu}(t)$ για κάθε $t \in \mathbb{R}$.

(ii) Αν το $\lim_{n\to\infty} \hat{\mu}_n(t)$ υπάρχει για κάθε $t \in \mathbb{R}$ και η $f(t) = \lim_{n\to\infty} \hat{\mu}_n(t)$ είναι συνεχής στο 0, τότε υπάρχει μέτρο πιθανότητας μ στο $\mathbb{R}$ ώστε $\hat{\mu}(t) = f(t)$ και $\mu_n \Rightarrow \mu$.

Απόδειξη. (i) Από την υπόθεση και το Θεώρημα 14.8, για κάθε συνάρτηση $f : \mathbb{R} \to \mathbb{R}$ συνεχή και φραγμένη έχουμε

$$\int f(x)\,\mathrm{d}\mu_n(x) \to \int f(x)\,\mathrm{d}\mu(x).$$

Εφόσον οι συναρτήσεις $\cos y$, $\sin y$ είναι συνεχείς και φραγμένες, έχουμε

$$\begin{aligned}\hat{\mu}_n(t) = \int e^{itx}\,\mathrm{d}\mu_n(x) &= \int \cos(tx)\,\mathrm{d}\mu_n(x) + i\int \sin(tx)\,\mathrm{d}\mu_n(x)\\ &\to \int \cos(tx)\,\mathrm{d}\mu(x) + i\int \sin(tx)\,\mathrm{d}\mu(x) = \hat{\mu}(t)\end{aligned}$$

για $n \to \infty$.

(ii) Βημα 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

$$\mu_n\left(\left\{x : |x| > \frac{2}{u}\right\}\right) \leq \frac{1}{u}\int_{-u}^{u}(1 - \hat{\mu}_n(t))\,\mathrm{d}t. \tag{15.2}$$

Από την υπόθεση και το θεώρημα φραγμένης σύγκλισης, η τελευταία ανισότητα δίνει

$$\overline{\lim_{n\to\infty}}\,\mu_n\left(\left\{x : |x| > \frac{2}{u}\right\}\right) \leq \frac{1}{u}\int_{-u}^{u}(1 - f(t))\,\mathrm{d}t. \tag{15.3}$$

Επειδή η f είναι συνεχής στο 0 και $f(0) = 1$ ($\mu_n(0) = 1$, για κάθε $n \geq 1$) το όριο του δεξιού μέλους της ανισότητας για $u \to 0^+$ ισούται με $1 - f(0) = 0$. Άρα υπάρχει $u_0 > 0$ και $n_0 \in \mathbb{N}$ έτσι ώστε

$$\mu_n\left(\left\{x : |x| > \frac{2}{u_0}\right\}\right) < \varepsilon \text{ για κάθε } n \geq n_0. \tag{15.4}$$

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.

Βημα 2: Υπάρχει μέτρο μ ώστε $f(t) = \hat{\mu}(t)$ για κάθε $t \in \mathbb{R}$.

Πράγματι, η $(\mu_n)_{n\geq 1}$ είναι σφιχτή και από το Θεώρημα 14.19 προκύπτει ότι υπάρχει υπακολουθία $(\mu_{n_k})_{n\geq 1}$ τέτοια ώστε να συγκλίνει ασθενώς σε ένα μέτρο πιθανότητας μ ($\mu_{n_k} \Rightarrow \mu$). Λόγω του (i),

$$\lim_{n\to\infty}\hat{\mu}_{n_k}(t) = \hat{\mu}(t), \quad \forall t \in \mathbb{R}.$$

Όμως

$$\lim_{k\to\infty}\hat{\mu}_{n_k}(t) = f(t), \quad \forall t \in \mathbb{R}.$$

Συνεπώς, $f(t) = \hat{\mu}(t)$ για κάθε $t \in \mathbb{R}$.

Βημα 3: Αν μία 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),

$$\lim_{n\to\infty}\hat{\mu}_{\lambda_n}(t) = \hat{\nu}(t), \quad \forall t \in \mathbb{R}.$$

Όμοια με πριν (Βήμα 2), $\hat{\nu}(t) = f(t)$ για κάθε $t \in \mathbb{R}$ και λόγω μοναδικότητας του μετασχηματισμού Fourier, $\nu = \mu$.

Βημα 4:[1] Η $(\mu_n)_{n\geq 1}$ συγκλίνει ασθενώς στο μ.

Έστω ότι αυτό δεν ισχύει. Τότε υπάρχει $h : \mathbb{R} \to \mathbb{R}$ συνεχής και φραγμένη, $\varepsilon > 0$ και $(\lambda_n)_{n\geq 1}$ ακολουθία στο $\mathbb{N}$ έτσι ώστε

$$\left|\int h(x)\,\mathrm{d}\mu_{\lambda_n}(x) - \int h(x)\,\mathrm{d}\mu(x)\right| \geq \varepsilon, \quad \forall n \geq 1. \tag{5}$$

Η $(\mu_{\lambda_n})_{n\geq 1}$ είναι σφιχτή, άρα υπάρχει υπακολουθία της, έστω $(\mu_{\lambda_{r_n}})_{n\geq 1}$, και μέτρο πιθανότητας ν στο $\mathbb{R}$ έτσι ώστε $\mu_{\lambda_{r_n}} \Rightarrow \nu$ (Θεώρημα 14.19). Από τα προηγούμενα (Βήμα 3) προκύπτει ότι $\nu = \mu$, δηλαδή $\mu_{\lambda_{r_n}} \Rightarrow \mu$. Συνεπώς,

$$\lim_{n\to\infty}\int h(x)\,\mathrm{d}\mu_{\lambda_{r_n}}(x) = \int h(x)\,\mathrm{d}\mu(x),$$

το οποίο είναι άτοπο λόγω της (5). ∎

[1] Για την καλύτερη κατανόηση της απόδειξης θεωρείστε την εξής ανάλογη άσκηση απειροστικού λογισμού: Έστω $\ell \in \mathbb{R}$ και $(x_n)_{n\in\mathbf{N}}$ 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 $x_n \to \ell$.

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 $\phi_{X_n}(t) \to \phi_X(t)$ για κάθε $t \in \mathbb{R}$ καθώς $n \to \infty$.*

Απόδειξη. Ισχύει ότι $\phi_{X_n}(t) = \widehat{\mathbf{P}^{X_n}}(t)$ και $\widehat{\mathbf{P}^X}(t) = \widehat{\mathbf{P}^X}(t)$, όπου $\mathbf{P}^{X_n}$, $\mathbf{P}^X$ οι κατανομές των X_n και X αντίστοιχα. Το συμπέρασμα έπεται με εφαρμογή του θεωρήματος συνέχειας του Lévy 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 Lévy. Θα το χρησιμοποιήσουμε πολλές φορές στο εξής και ιδιαίτερα για να αποδείξουμε το κεντρικό οριακό θεώρημα.

15.2 Εφαρμογές

Σε υπολογισμούς ορίων της μορφής $\lim_{n\to\infty} \phi_{X_n}(t)$ χρήσιμο είναι το εξής απλό αποτέλεσμα.

Λήμμα 15.4. *Αν $(c_n)_{n\geq 1}$ είναι ακολουθία στο $\mathbb{C}$ τέτοια ώστε $c_n \to c$, με $c \in \mathbb{C}$, τότε*

$$\Big(1 + \frac{c_n}{n}\Big)^n \to e^c.$$

Η απόδειξή του δίνεται στο Παράρτημα Α′ (Πόρισμα Α′.3).

Παράδειγμα 15.5. Έστω $X_n \sim \mathbf{Bin}(n, \frac{\lambda}{n})$ για κάθε $n \geq 1$ και $X \sim \mathbf{Poisson}(\lambda)$. Τότε

$$\begin{aligned}
\phi_{X_n}(t) = \mathbf{E}(e^{itX_n}) &= \sum_{k=0}^{n} e^{itk}\binom{n}{k}\Big(\frac{\lambda}{n}\Big)^k\Big(1-\frac{\lambda}{n}\Big)^{n-k} \\
&= \sum_{k=0}^{n}\binom{n}{k}\Big(\frac{e^{it}\lambda}{n}\Big)^k\Big(1-\frac{\lambda}{n}\Big)^{n-k} \\
&= \Big(1-\frac{\lambda}{n}+\frac{\lambda e^{it}}{n}\Big)^n \\
&= \Big(1+\frac{\lambda(e^{it}-1)}{n}\Big)^n
\end{aligned}$$

και

$$\lim_{n\to\infty}\Big(1+\frac{\lambda(e^{it}-1)}{n}\Big)^n = e^{\lambda(e^{it}-1)}.$$

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, $\phi_{X_n}(t) \to \phi_X(t)$. Συνεπώς, από το θεώρημα συνέχειας του Lévy προκύπτει ότι $X_n \Rightarrow X$.

Παράδειγμα 15.6. Έστω $X_n \sim \mathbf{Poisson}(n)$ και $Z \sim \mathbf{N}(0,1)$. 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,

$$\lim_{\varepsilon\to 0}\frac{e^{i\varepsilon t}-1-i\varepsilon t}{\varepsilon^2}=\frac{-t^2}{2}.$$

Άρα $\phi_{Z_n}(t)\to\phi_Z(t)=e^{-\frac{t^2}{2}}$ και από το θεώρημα συνέχειας του Lévy 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 $S_n=X_1+\cdots+X_n$ 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 13.11 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

$$C_n(t)=n\left(\frac{t^2}{a_n^2}\log|t|-\frac{t^2}{a_n^2}\log a_n+b(n,t)\right)=n\left(\frac{t^2}{n\log n}\log|t|-\frac{t^2}{2n\log n}\log(n\log n)+b(n,t)\right)\to-\frac{t^2}{2}$$

καθώς $n\to\infty$. Το συμπέρασμα έπεται από το Λήμμα 15.4 και το θεώρημα συνέχειας του Lévy.

Ασκήσεις

15.1 Έστω μ μέτρο πιθανότητας στο $[0,\infty)$. Μετασχηματισμό Laplace του μ λέμε τη συνάρτηση $L:[0,\infty)\to[0,\infty)$ με $L(s)=\int_0^\infty e^{-sx}\,d\mu(x)$ για κάθε $s\geq 0$. Να δειχθεί ότι για κάθε $u>0$ ισχύει

$$\mu\left(\frac{2}{u},\infty\right)\leq\frac{2}{u}\int_0^u(1-L(s))\,ds. \tag{15.5}$$

15.2 (α) Έστω X τυχαία μεταβλητή που ακολουθεί την εκθετική κατανομή με παράμετρο $a>0$. Να υπολογιστεί η χαρακτηριστική συνάρτηση της X.

(β) Έστω Y τυχαία μεταβλητή που ακολουθεί τη γεωμετρική κατανομή με παράμετρο $p\in(0,1]$. Δηλαδή

$$\mathbf{P}(Y=k)=p\,(1-p)^{k-1}$$

για $k=1,2,\ldots$ Να υπολογιστεί η χαρακτηριστική συνάρτηση της Y.

(γ) 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 στην τυχαία μεταβλητή X του ερωτήματος (α) χρησιμοποιώντας

(i) τον χαρακτηρισμό της σύγκλισης κατά κατανομή μέσω συναρτήσεων κατανομής,

(ii) χαρακτηριστικές συναρτήσεις.

15.3 (α) Έστω X τυχαία μεταβλητή που ακολουθεί την κατανομή Poisson με παράμετρο $\lambda > 0$. Να υπολογιστεί η χαρακτηριστική συνάρτηση της X.

(β) Έστω Y τυχαία μεταβλητή που ακολουθεί τη διωνυμική κατανομή με παραμέτρους $n \in \mathbb{N}^+, p \in [0,1]$. Να υπολογιστεί η χαρακτηριστική συνάρτηση της Y.

(γ) Έστω $(X_n)_{n\geq 1}$ ακολουθία τυχαίων μεταβλητών ώστε η X_n να ακολουθεί τη διωνυμική κατανομή με παραμέτους $n, p_n \in (0,1)$. Αν $\lim_{n\to\infty} np_n = \lambda$, να δειχθεί ότι η ακολουθία $(X_n)_{n\geq 1}$ συγκλίνει κατά κατανομή στην τυχαία μεταβλητή X του ερωτήματος (α).

15.4 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 $n \to \infty$.

15.5 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 $n \geq 1$, θέτουμε $S_n := X_1 + X_2 + \cdots + X_n$. Να δειχθεί ότι

$$\frac{S_n}{n} \to a$$

κατά πιθανότητα για $n \to \infty$.

15.6 Έστω $(X_n)_{n\geq 1}$ 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 ασθενώς 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.]

15.7 Έστω $(X_k)_{k\geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με $\mathbf{E}(X_1) = 0, \mathrm{Var}(X_1) = 1$. Για κάθε $n \geq 1$, θέτουμε $S_n := X_1 + X_2 + \cdots + X_n$ και $T_n = S_1 + S_2 + \cdots + S_n$. 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 $n \to \infty$.

15.8 Έστω $(X_n)_{n\geq 1}$ ακολουθία ανεξάρτητων τυχαίων μεταβλητών ώστε για κάθε $k \geq 1$ έχουμε

$$X_k = \begin{cases} -k & \text{με πιθανότητα } 1/(2k^2), \\ k & \text{με πιθανότητα } 1/2k^2, \\ -1 & \text{με πιθανότητα } (1-k^{-2})/2, \\ -1 & \text{με πιθανότητα } (1-k^{-2})/2. \end{cases}$$

Να δειχθεί ότι

(α) $\lim_{n\to\infty} \mathrm{Var}(S_n)/n = 2$,

(β) η ακολουθία $(S_n/\sqrt{n})_{n\geq 1}$ συγκλίνει ασθενώς 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.]

15.9 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 $S_n := X_1 + X_2 + \cdots + X_n$. Να δειχθεί ότι η ακολουθία $(S_n/n^{1/a})_{n\geq 1}$ συγκλίνει ασθενώς 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.

16

Το Κεντρικό Οριακό Θεώρημα

Ο 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 $\sigma^2 := \mathrm{Var}(X_1) \in (0, \infty)$, 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 τ.μ. X με πεπερασμένη πρώτη και δεύτερη ροπή. Χρήσιμη είναι η ανισότητα

$$\left| e^{ix} - 1 - ix - \frac{1}{2}(ix)^2 \right| \le 2x^2, \tag{16.1}$$

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 $|\cos x - 1| \le x^2/2$. Όμοια δείχνουμε ότι $|\sin x - x| \le x^2/2$.

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*

$$\phi_X(t) = 1 + it\mu - \frac{t^2\beta}{2} + \nu(t), \tag{16.2}$$

όπου η ν ικανοποιεί $\lim_{t\to 0} \frac{\nu(t)}{t^2} = 0$.

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

$$1 + itX + \frac{1}{2}(itX)^2,$$

και πάρουμε τη μέση τιμή τους. Η $\nu(t)$ είναι η μέση τιμή της υπόλοιπης δυναμοσειράς.

Απόδειξη. Έστω $A(x) := e^{ix} - 1 - ix - (ix)^2/2$ για κάθε $x \in \mathbb{R}$.

$$|\nu(t)| = \left|\phi_X(t) - 1 - it\mu + \frac{t^2\beta}{2}\right| = \left|\mathbf{E}\left(e^{itX} - 1 - itX - \frac{(itX)^2}{2}\right)\right| \leq \mathbf{E}\,|A(tX)|$$

Για την απόδειξη του λήμματος αρκεί να δείξουμε ότι $\lim_{t\to 0} \mathbf{E}(|A(tX)|/t^2) = 0$. Έστω ακολουθία $(t_n)_{n\in\mathbb{N}}$ στο $\mathbb{R}\backslash\{0\}$ τέτοια ώστε $t_n \to 0$. Τότε

- $A(t_nX)/t_n^2 \to 0$ για $n \to \infty$ λόγω του αναπτύγματος της e^{ix} σε δυναμοσειρά.
- $|A(t_nX)/t_n^2| \leq 2X^2$ για κάθε $n \in \mathbb{N}$ λόγω της (16.1) και $\mathbf{E}(X^2) < \infty$.

Από το θεώρημα κυριαρχημένης σύγκλισης, $\lim_{n\to\infty} \mathbf{E}(|A(t_nX)|/t_n^2) = \mathbf{E}(\lim_{n\to\infty} |A(t_nX)|/t_n^2) = 0$. ■

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 $\mathrm{Var}(X_1) = \sigma^2 \in (0, \infty)$. Θέτουμε $S_n = X_1 + X_2 + \ldots + X_n$ για κάθε $n \geq 1$. Τότε*

$$\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \Rightarrow Z, \quad \text{όπου } Z \sim N(0, 1).$$

Απόδειξη. Θα αποδείξουμε αρχικά την περίπτωση όπου $\mu = 0$, οπότε $\sigma^2 = \mathbf{E}(X^2)$. Θα χρησιμοποιήσουμε το θεώρημα συνέχειας του Lévy (Πόρισμα 15.3). Υπολογίζουμε τη χαρακτηριστική συνάρτηση της $S_n/\sqrt{n\sigma^2}$. Για $t \in \mathbb{R}\backslash\{0\}$,

$$\begin{aligned}
\phi_{\frac{S_n}{\sqrt{n\sigma^2}}}(t) &= \mathbf{E}\left(e^{it\frac{S_n}{\sqrt{n\sigma^2}}}\right) = \mathbf{E}\left(e^{it\frac{X_1}{\sqrt{n\sigma^2}}} \cdots e^{it\frac{X_n}{\sqrt{n\sigma^2}}}\right) \\
&= \mathbf{E}\left(e^{it\frac{X_1}{\sqrt{n\sigma^2}}}\right) \cdots \mathbf{E}\left(e^{it\frac{X_n}{\sqrt{n\sigma^2}}}\right) && \text{(λόγω ανεξαρτησίας)} \\
&= \phi_{X_1}\left(\tfrac{t}{\sqrt{n\sigma^2}}\right) \cdots \phi_{X_n}\left(\tfrac{t}{\sqrt{n\sigma^2}}\right) \\
&= \left(\phi_{X_1}\left(\tfrac{t}{\sqrt{n\sigma^2}}\right)\right)^n && \text{(λόγω ισονομίας).}
\end{aligned}$$

Από το Λήμμα 16.1,

$$\begin{aligned}
\phi_{X_1}\left(\frac{t}{\sqrt{n\sigma^2}}\right) &= 1 + \frac{it}{\sqrt{n\sigma^2}}\mathbf{E}(X_1) - \frac{t^2}{2n\sigma^2}\mathbf{E}\left((X_1)^2\right) + \nu\left(\frac{t}{\sqrt{n\sigma^2}}\right) = 1 - \frac{t^2}{2n} + \nu\left(\frac{t}{\sqrt{n\sigma^2}}\right) \\
&= 1 + \frac{C_n}{n}
\end{aligned}$$

με $C_n = -\frac{t^2}{2} + n\nu\left(\frac{t}{\sqrt{n\sigma^2}}\right) \to -t^2/2$ για $n \to \infty$ λόγω του ότι $\lim_{s\to 0} \nu(s)/s^2 = 0$. Συνεπώς, το Λήμμα 15.4 δίνει

$$\lim_{n\to\infty} \phi_{\frac{S_n}{\sqrt{n\sigma^2}}}(t) = e^{-t^2/2}.$$

Η ίδια σχέση ισχύει προφανώς και για $t = 0$. Μία $Z \sim N(0, 1)$ έχει χαρακτηριστική συνάρτηση $e^{-t^2/2}$. Από το θεώρημα συνέχειας του Lévy, $S_n/\sqrt{n\sigma^2} \Rightarrow Z$.

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. ■

Ας επιστρέψουμε στο Παράδειγμα 15.7. Αυτό που συμβαίνει εκεί είναι ότι $\mathbf{E}(X_1) = 0$ αλλά $\mathrm{Var}(X_1) = \infty$. Ο νόμος των μεγάλων αριθμών εφαρμόζεται και δίνει ότι $S_n/n \to 0$. Όμως το ότι $\mathrm{Var}(X_1) = \infty$ σημαίνει ότι το άθροισμα S_n έχει μεγαλύτερη μεταβλητότητα από την περίπτωση που $\mathrm{Var}(X_1) < \infty$. 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

$$\mathbf{P}\left(\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \in A\right) \tag{16.3}$$

προσεγγίζεται (για n μεγάλο) από την

$$\mathbf{P}(Z \in A) \tag{16.4}$$

όπου $Z \sim N(0,1)$ και A είναι ένα διάστημα ή μια πεπερασμένη ένωση διαστημάτων (αφού το όριο της πρώτης ποσότητας για $n \to \infty$ 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 $S_n := X_1 + \cdots + X_n$. 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 $\mu = \mathbf{E}(X_1) = 1/2, \sigma^2 := \mathrm{Var}(X_1) = p(1-p) = 1/4$ 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 μας κάνει 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 $X_1, X_2, \ldots, X_n$ και θα χρησιμοποιήσουμε τον μέσο όρο $\bar{X}_n := (X_1 + \cdots + X_n)/n$ 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;

Για την πιθανότητα του γεγονότος που μας ενδιαφέρει έχουμε την εξής προσέγγιση

$$\mathbf{P}\left(\left|\frac{S_n}{n} - \mu\right| \leq \frac{1}{100}\right) = \mathbf{P}\left(\frac{|S_n - n\mu|}{\sqrt{n\sigma^2}} \leq \sqrt{n}\frac{1}{100\,\sigma}\right) \approx \mathbf{P}\left(|Z| \leq \sqrt{n}\frac{1}{100\,\sigma}\right) \tag{16.5}$$

$$= \Phi\left(\sqrt{n}\frac{1}{100\,\sigma}\right) - \Phi\left(-\sqrt{n}\frac{1}{100\,\sigma}\right) = 2\Phi\left(\sqrt{n}\frac{1}{100\,\sigma}\right) - 1. \tag{16.6}$$

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 οποιονδήπτοτε 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

$$\lim_{k\to\infty} \mathbf{P}\left(\frac{|S_k - k\mu|}{\sqrt{k\sigma^2}} \leq \sqrt{n}\frac{1}{100\,\sigma}\right) = \mathbf{P}\left(|Z| \leq \sqrt{n}\frac{1}{100\,\sigma}\right),$$

ήδη για $k = n$ οι δύο αριθμοί

$$\mathbf{P}\left(\frac{|S_n - n\mu|}{\sqrt{n\sigma^2}} \leq \sqrt{n}\frac{1}{100\,\sigma}\right), \mathbf{P}\left(|Z| \leq \sqrt{n}\frac{1}{100\,\sigma}\right) \tag{16.7}$$

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 $A \in \mathscr{B}(\mathbb{R})$ με $\mathbf{P}(Z \in \partial A) = 0$, το ότι

$$\lim_{n\to\infty} \mathbf{P}\left(\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \in A\right) = \mathbf{P}(Z \in A), \tag{16.8}$$

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. Έλεγχο 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

$$\left|\mathbf{P}\left(\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \leq x\right) - \Phi(x)\right| \leq \frac{3\rho}{\sigma^2\sqrt{n}}.$$

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 $\mathbf{P}(S_{100} \geq 125), \mathbf{P}(S_{100} \geq 150)$, καθώς και την προσέγγιση που δίνει το κεντρικό οριακό θεώρημα. Έχουμε $\mathbf{E}(S_{100}) = 100$, $\sqrt{\mathrm{Var}(S_{100})} = 10$. Άρα τα γεγονότα $S_{100} \geq 125, S_{100} \geq 150$ είναι σημαντικές αποκλίσεις από τη μέση τιμή (2.5 και 5 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}

Παρατηρήστε ότι η προσέγγιση για την πιθανότητα $\mathbf{P}(S_{100} \geq 150)$ είναι περίπου το 1/20 της αληθινής τιμής της πιθανότητας.

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 $t \to \infty$ όσο όταν $X_1 \sim N(0, 1)$ ή συντομότερα.

Οι πιθανότητες που μας απασχόλησαν στο προηγούμενο παράδειγμα ήταν ενδεχομένων με πολύ μικρή πιθανοτητα (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 $(X_n)_{n\geq1}$ είναι ανεξάρτητες ισόνομες με $\mathbf{E}(X_1) = 0, \mathrm{Var}(X_1) = 1$ Το κεντρικό οριακό θεώρημα λέει ότι η ακολουθία $R_n := S_n/\sqrt{n}$ συγκλίνει κατά κατανομή σε μια τυχαία μεταβλητή με κατανομή την $N(0, 1)$. 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 $(R_{n_k})_{k\geq1}$ που να συγκλίνει στην W με πιθανότητα 1. Εφαρμόζοντας τα επιχειρήματα της Άσκησης 16.7 (τώρα για την ακολουθία $(R_{n_k})_{k\geq1}$) δείχουμε ότι $\underline{\lim}_{k\to\infty} R_{n_k} = -\infty, \overline{\lim}_{k\to\infty} R_{n_k} = \infty$ με πιθανότητα 1. Άρα η $(R_{n_k})_{k\geq1}$ δεν μπορεί να συγκλίνει με πιθανότητα 1.

Ασκήσεις

Στις ασκήσεις πιο κάτω, αν $(X_i)_{i\geq1}$ είναι ακολουθία ανεξάρτητων τυχαίων μεταβλητών, συμβολίζουμε με S_n το άθροισμα των πρώτων n από αυτές. Δηλαδή $S_n := X_1 + X_2 + \cdots + X_n$ για κάθε $n \geq 1$.

16.1 Έστω $(X_n)_{n\geq1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με $\mathbf{E}(X_1) = \mu \in \mathbb{R}$ και διασπορά $\sigma^2 \in (0, \infty)$. Να δειχθεί ότι

$$\lim_{n\to\infty} \mathbf{P}(S_n > n\mu) = \frac{1}{2}.$$

16.2 Έστω $(X_i)_{i\geq1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με $EX_1 = 2, \mathrm{Var}(X_1) = 1$. Να υπολογιστούν τα όρια

(α) $\lim_{n\to\infty} \mathbf{P}(S_n > 2.1n)$,

(β) $\lim_{n\to\infty} \mathbf{P}(S_n > 2n + \sqrt{n})$,

(γ) $\lim_{n\to\infty} \mathbf{P}(S_n > 10\sqrt{n})$,

(δ) $\lim_{n\to\infty} \mathbf{P}(S_n < 3n)$,

(ε) $\lim_{n\to\infty} \mathbf{P}(S_n > 10^{10})$.

16.3 Έστω $(X_n)_{n\geq1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές έτσι ώστε $X_1 \sim \mathbf{Poisson}(1)$. Να δείξετε ότι

$$\frac{S_n - n}{\sqrt{n}} \Rightarrow Z, \text{ όπου } Z \sim \mathbf{N}(0, 1).$$

16.4 Έστω $p \in (0, 1)$ και $A \in \mathbb{R}$. Να υπολογιστεί το όριο

$$\lim_{n\to\infty} \sum_{k=0}^{np+A\sqrt{n}} \binom{n}{k} p^k (1-p)^{n-k}.$$

16.5 Να δείξετε ότι

$$\lim_{n\to\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = \frac{1}{2}.$$

16.6* Έστω $(X_n)_{n\geq 1}$ ακολουθία ανεξάρτητων ισόνομων τυχαίων μεταβλητών με $\mathbf{E}(X_1) \in \mathbb{R}$ και $\mathrm{Var}(X_1) = 1$. Να υπολογιστεί το όριο

$$\lim_{n\to\infty} \mathbf{P}\left(\left| \frac{X_1 + X_2 + \cdots + X_n - (X_{n+1} + \cdots + X_{2n})}{\sqrt{n}} \right| \leq 1 \right).$$

16.7 Έστω $(X_n)_{n\geq 1}$ ακολουθία ανεξάρτητων ισόνομων τυχαίων μεταβλητών με $\mathbf{E}(X_1) = 0$, $\mathrm{Var}(X_1) = 1$. Να δειχθεί ότι

$$\varlimsup_{n\to\infty} \frac{S_n}{\sqrt{n}} = \infty.$$

[Υπόδειξη: Η στρατηγική της Άσκησης 11.16 λειτουργεί. Απλώς το (α) μέρος χρειάζεται μια μικρή τροποποίηση.]