Κεφάλαιο 1

Κεφάλαιο 2

Κεφάλαιο 3

Κεφάλαιο 4

Κεφάλαιο 5

Κεφάλαιο 6

Νόμοι Μεγάλων Αριθμών

6.1 Σύγκλιση Ακολουθίας Τυχαίων Μεταβλητών

Στην μαθηματική ανάλυση έχουμε, σε γενικές γραμμές, τρία είδη συγκλίσεων μιας ακολουθίας πραγματικών συναρτήσεων $\{f_n,\ n \geq 1\}$ προς μία πραγματική συνάρτηση f. Αυτές είναι:

(α) Η **κατά σημείο** σύγκλιση, δηλαδή

$$\lim_{n\to\infty} f_n(x) = f(x), \quad \text{για κάθε } x \in \mathbb{R},$$

(β) Η **ομοιόμορφη** σύγκλιση, δηλαδή

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} |f_n(x) - f(x)| = 0,$$

(γ) Η σύγκλιση στον L^p (εδώ $p \geq 1$), κατά την οποία,

$$\lim_{n\to\infty} \int |f_n(x) - f(x)|^p dx = 0.$$

Στις πιθανότητες, έχοντας υπ΄ όψιν ότι οι τ.μ. είναι και μαθηματικές συναρτήσεις ($X : \Omega \longrightarrow \mathbb{R}$), θα μπορούσαμε, κατ΄

αναλογία, να μελετάμε αντίστοιχες συγκλίσεις, όπως π.χ. την κατά σημείο σύγκλιση (δηλ. για κάθε $\omega \in \Omega$, $X_n(\omega) \to X(\omega)$), την ομοιόμορφη σύγκλιση ($\sup_{\omega\in\Omega} |X_n(\omega) - X(\omega)| \to 0$), και την L^p σύγκλιση ($\mathbb{E}|X_n - X|^p \to 0$). Αναγνωρίζεται όμως εύκολα ότι στην κατά σημείο, όπως και στην ομοιόμορφη σύγκλιση, δεν λαμβάνουμε υπ΄ όψιν μας την δομή του χώρου Ω, στον οποίο διεξάγεται το πείραμα, και, στην ουσία, οι συγκλίσεις αυτές δεν είναι **στοχαστικές**. Για παράδειγμα, έστω ότι $\Omega = \{K, \mathit{\Gamma}\}$, με $\mathbb{P}(\{K\}) = 0.999$ και $\mathbb{P}(\{\mathit{\Gamma}\}) = 0.001$, $X(K) = 0$, $X(\mathit{\Gamma}) = 1$, και $X_n(K) = 1/n$, $X_n(\mathit{\Gamma}) = (-1)^n n$. Τότε έχουμε

$$X_n(K) \to X(K), \quad X_n(\mathit{\Gamma}) \nrightarrow X(\mathit{\Gamma})$$

(γενικά η $X_n(\mathit{\Gamma})$ δεν συγκλίνει), οπότε δεν ισχύει ότι $X_n \to X$. Όμως τα ω για τα οποία $X_n(\omega) \nrightarrow X(\omega)$ (δηλ. το μονοσύνολο $\{\mathit{\Gamma}\}$) αποτελούν ενδεχόμενο με «μικρή πιθανότητα», $\mathbb{P}(\{\mathit{\Gamma}\}) = 0.001$, οπότε μπορούμε να πούμε ότι «με μεγάλη πιθανότητα, $X_n \to X$».

Αν υποθέσουμε ότι τα ω για τα οποία $X_n(\omega) \nrightarrow X(\omega)$ έχουν πιθανότητα 0, δηλ. $\mathbb{P}(\{\omega : X_n(\omega) \nrightarrow X(\omega)\}) = 0$ ή, ισοδύναμα, $\mathbb{P}(\{\omega : X_n(\omega) \to X(\omega)\}) = 1$, τότε, ουσιαστικά, ισχύει ότι $X_n \to X$ σ.π. (ως προς το μέτρο $\mathbb{P}$), δηλ. $X_n \to X$ με πιθ. 1, και άρα, από «στοχαστικής» απόψεως, $X_n \to X$.

Γενικά στις πιθανότητες μελετάμε τέσσερα είδη συγκλίσεων:

1. **Σύγκλιση κατά πιθανότητα.** Λέμε ότι η ακολουθία των τ.μ. $\{X_n,\ n \geq 1\}$ συγκλίνει κατά πιθανότητα (in probability) στην τ.μ. X, όταν οι τ.μ. X_n και X είναι ορισμένες **στον ίδιο** χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, και ισχύει:

$$\text{Για κάθε } \varepsilon > 0,\ \mathbb{P}(|X_n - X| \geq \varepsilon) \to 0, \quad \text{καθώς } n \to \infty.$$

Συμβολισμός: $X_n \xrightarrow{\mathbf{p}} X$.
Με άλλα λόγια, $X_n \xrightarrow{\mathbf{p}} X$ όταν, για κάθε δοθέν $\varepsilon > 0$ (οσοδήποτε μικρό), η πιθανότητα όπως η X_n διαφέρει από την X περισσότερο από ε τείνει στο 0, καθώς $n \to \infty$. Φυσικά, αφού οι X_n και X είναι τ.μ., το σύνολο $\{\omega : |X_n(\omega) - X(\omega)| \geq \varepsilon\}$ είναι ενδεχόμενο, δηλ. $\{|X_n - X| \geq \varepsilon\} \in \mathcal{A}$.

Ισοδύναμα, μπορούμε να πούμε ότι $X_n \xrightarrow{\mathbf{p}} X$ όταν

$$\text{Για κάθε } \varepsilon > 0,\ \mathbb{P}(|X_n - X| < \varepsilon) \to 1, \quad \text{καθώς } n \to \infty.$$

2. **Σύγκλιση με πιθανότητα 1 (σχεδόν βεβαίως).** Λέμε ότι η ακολουθία τ.μ. $\{X_n,\ n \geq 1\}$ συγκλίνει στην τ.μ. X **με πιθανότητα 1**, ή **σχεδόν βεβαίως** (συμβολισμός: $X_n \xrightarrow{\mathbf{a.s.}} X$, από το a.s. = almost surely = σχεδόν βεβαίως), όταν οι X_n και X είναι ορισμένες **στον ίδιο** χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, και ισχύει ότι

$$\mathbb{P}\left(\lim_n X_n = X\right) = 1,$$

δηλ. $\mathbb{P}(C) = 1$, όπου $C = \{\omega : X_n(\omega) \to X(\omega)\}$.

Ισοδύναμα, $X_n \xrightarrow{\textbf{a.s.}} X$ όταν και μόνο όταν $\mathbb{P}(\lim X_n \neq X) = 0$, δηλ. $\mathbb{P}(C^c) = 0$, όπου $C^c = \Omega \setminus C = \{\omega :$ η $X_n(\omega)$

δεν συγκλίνει, ή η $X_n(\omega)$ συγκλίνει και $\lim X_n(\omega) \neq X(\omega)\}$. Εδώ θεωρούμε ότι η ακολουθία $X_n(\omega)$ συγκλίνει, ακόμα και αν $\lim X_n(\omega) = +\infty$ ή $-\infty$. Η σύγκλιση αυτή ονομάζεται και σχεδόν παντού σύγκλιση (κυρίως σε μαθηματικά κείμενα), ή και **ισχυρή** σύγκλιση (θα δούμε αργότερα γιατί).

Παρατηρούμε ότι έχει νόημα η πιθανότητα $\mathbb{P}(C)$, διότι το σύνολο $C \in \mathcal{A}$, επειδή

$$\begin{aligned} C = &\bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \left(\left\{ \omega : |X_k - X| \leq \frac{1}{m} \right\} \cap \{\omega : |X| < \infty\} \right) \\ &\bigcup \left[\{\omega : X = +\infty\} \cap \left(\bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \{\omega : X_k \geq m\} \right) \right] \\ &\bigcup \left[\{\omega : X = -\infty\} \cap \left(\bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \{\omega : X_k \leq -m\} \right) \right]. \end{aligned}$$

Τα δύο τελευταία σύνολα χρειάζονται όταν η X είναι εκτεταμένη τ.μ., αλλά εμείς στα επόμενα θα θεωρούμε ότι η X είναι συνήθης (πεπερασμένη) τ.μ., δηλ. $\{X = -\infty\} = \{X = +\infty\} = \varnothing$ (ισοδύναμα, $\{|X| < \infty\} = \Omega$), εκτός αν ρητά αναφέρεται το αντίθετο.

3. Σύγκλιση στον L^p. Λέμε ότι η ακολουθία τ.μ. $\{X_n,\ n \geq 1\}$ συγκλίνει στην τ.μ. X στον L^p ($p \geq 1$), όταν οι X_n και X είναι ορισμένες **στον ίδιο** χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, $\mathbb{E}|X|^p < \infty$,

$\mathbb{E}|X_n|^p < \infty$ από κάποιο n και πάνω (δηλ. τελικά για κάθε n), και

$$\mathbb{E}|X_n - X|^p \to 0, \quad \text{καθώς } n \to \infty.$$

Συμβολισμός: $X_n \xrightarrow{\boldsymbol{L^p}} X$.

Οι πιο ενδιαφέρουσες περιπτώσεις προκύπτουν για $p = 1$ και $p = 2$:

$p = 1$: $\quad X_n \xrightarrow{\boldsymbol{L^1}} X \overset{\text{ορ.}}{\Longleftrightarrow} \mathbb{E}|X_n - X| \to 0 \quad$ (σύγκλιση στον L^1).

$p = 2$: $\quad X_n \xrightarrow{\boldsymbol{L^2}} X \overset{\text{ορ.}}{\Longleftrightarrow} \mathbb{E}(X_n - X)^2 \to 0$ (σύγκλιση στον L^2 ή σύγκλιση κατά μέσο τετράγωνο).

4. Σύγκλιση κατά κατανομή (ή κατά νόμο ή ασθενής σύγκλιση). Η ακολουθία των τ.μ. $\{X_n, n \geq 1\}$ (σε αυθαίρετους χώρους πιθανότητας $(\Omega_n, \mathcal{A}_n, \mathbb{P}_n)$) λέμε ότι συγκλίνει κατά κατανομή, ή ασθενώς (convergence in distribution, weak convergence) προς την τ.μ. X (ορισμένη σε κάποιον, αυθαίρετο, χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$), όταν

$$\mathbb{P}_n[X_n \leq x] \to \mathbb{P}[X \leq x], \quad \text{καθώς } n \to \infty,$$

για κάθε x για το οποίο $\mathbb{P}[X = x] = 0$. Συμβολισμός: $X_n \xrightarrow{\mathbf{d}} X$.

Στην πραγματικότητα, η σύγκλιση αυτή είναι σύγκλιση των αντιστοίχων *σ.κ.* F_n προς την *σ.κ.* F της X, στα σημεία συνεχείας της F, και γι΄ αυτό γράφουμε ισοδύναμα

$$F_n \xrightarrow{\mathbf{d}} F \iff F_n(x) \to F(x), \text{ για κάθε σημείο συνεχείας } x \text{ της } F,$$

και δεν χρειάζεται καν να αναφερθούμε στις τ.μ. X_n, X.

Είναι φανερό ότι υπάρχει «ποσοτική διαφορά» μεταξύ των συγκλίσεων $1, 2, 3$, και «ποσοτική» και «ποιοτική» διαφορά μεταξύ των τριών πρώτων συγκλίσεων και της ασθενούς. Η ασθενής σύγκλιση εξαρτάται αποκλειστικά από την συμπεριφορά των επαγόμενων μέτρων στον $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, αφού περιγράφεται **πλήρως** από τις *σ.κ.* F_n και F και, συνεπώς, **δεν εξαρτάται άμεσα** από τις ίδιες τις τ.μ. X_n και X, σε αντίθεση με τις άλλες τρεις συγκλίσεις.

Για να γίνει κατανοητή η σημασία των σημείων συνεχείας ως προς την κατά κατανομή σύγκλιση, δίνουμε το εξής απλό παράδειγμα.

Παράδειγμα 6.1 Έστω $(\Omega, \mathcal{A}, \mathbb{P})$ ένας χώρος πιθανότητας, και a_n μία ακολουθία πραγματικών αριθμών, τέτοια ώστε

$$a_n \to a \in \mathbb{R}, \ a_n < a \text{ για άπειρα } n, \text{ και } a_n > a \text{ για άπειρα } n.$$

[π.χ. $a_n = a + (-1)^n/n$.] Θεωρούμε τις τ.μ. $X_n(\omega) = a_n$ και $X(\omega) = a$ (σταθερές τ.μ.), και παρατηρούμε ότι $X_n(\omega) \to$

$X(\omega)$ για κάθε $\omega \in \Omega$ (και άρα, $X_n \xrightarrow{\mathbf{a.s.}} X$ και, ακόμη περισσότερο, $X_n \to X$ κατά σημείο). Όμως οι σ.κ. των X_n, X είναι οι

$$F_n(x) = \begin{cases} 0, & \text{αν } x < a_n, \\ 1, & \text{αν } x \geq a_n, \end{cases} \quad \text{και} \quad F(x) = \begin{cases} 0, & \text{αν } x < a, \\ 1, & \text{αν } x \geq a. \end{cases}$$

Αν και, προφανώς, ισχύει ότι $F_n(x) \to F(x)$ για $x \neq a$, η ακολουθία $F_n(a)$ δεν συγκλίνει, αφού $0 = \liminf F_n(a) < \limsup F_n(a) = 1$. Φυσικά, αν παίρναμε $a_n > a$ και $a_n \to a$, τότε θα είχαμε $F_n(a) \to 0 \neq F(a) = 1$, οπότε και πάλι δεν ισχύει $F_n \to F$ κατά σημείο. Συνεπώς, ένας «χρήσιμος» ορισμός ασθενούς σύγκλισης δεν θα έπρεπε να απαιτεί σύγκλιση και στα σημεία ασυνέχειας της F, αφού, σύμφωνα με τα παραπάνω, ακόμα και η σχεδόν βέβαιη (ισχυρή) σύγκλιση (ή και η ομοιόμορφη) μπορεί να μην τον ικανοποιεί, δεδομένου ότι οι X_n του παραδείγματος συγκλίνουν **ομοιόμορφα** (ως προς ω) προς την X. $\square$

Οι διαφορετικοί τρόποι σύγκλισης συγκρίνονται στο επόμενο θεώρημα.

Θεώρημα 6.2 Θεωρούμε τις τ.μ. $\{X_n,\ n \geq 1\}$ και X ορισμένες στον χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$. Τότε:

(i) Εάν $X_n \xrightarrow{\mathbf{a.s.}} X$ τότε $X_n \xrightarrow{\mathbf{p}} X$.

(ii) Εάν $X_n \xrightarrow{\boldsymbol{L^p}} X$ (για κάποιο $p \geq 1$) τότε $X_n \xrightarrow{\mathbf{p}} X$.

(iii) Εάν $X_n \xrightarrow{\mathbf{p}} X$ τότε $X_n \xrightarrow{\mathbf{d}} X$.

Απόδειξη: (ii) Έστω $\varepsilon > 0$. Από την ανισότητα Markov,

$$\mathbb{P}(|X_n - X| \geq \varepsilon) = \mathbb{P}(|X_n - X|^p \geq \varepsilon^p) \leq \frac{\mathbb{E}|X_n - X|^p}{\varepsilon^p} \to 0,$$

αφού $X_n \xrightarrow{L^p} X$.

(iii) Έστω $F_{X_n}(x) = \mathbb{P}(X_n \leq x)$, $n = 1, 2, \ldots$, και ας υποθέσουμε ότι το x είναι ένα σημείο συνεχείας της $F_X(x) = \mathbb{P}(X \leq x)$. Για $m = 1, 2, \ldots$ έχουμε

$$\begin{aligned}
F_X\left(x - \tfrac{1}{m}\right) &= \mathbb{P}\big(X \leq x - \tfrac{1}{m}\big) \\
&= \mathbb{P}\big(X \leq x - \tfrac{1}{m}, X_n \leq x\big) + \mathbb{P}\big(X \leq x - \tfrac{1}{m}, X_n > x\big) \\
&\leq \mathbb{P}(X_n \leq x) + \mathbb{P}\big(|X_n - X| \geq \tfrac{1}{m}\big) \\
&= F_{X_n}(x) + \mathbb{P}\big(|X_n - X| \geq \tfrac{1}{m}\big),
\end{aligned}$$

και ομοίως,

$$\begin{aligned}
F_{X_n}(x) &= \mathbb{P}(X_n \leq x) \\
&= \mathbb{P}\big(X_n \leq x, X \leq x + \tfrac{1}{m}\big) + \mathbb{P}\big(X_n \leq x, X > x + \tfrac{1}{m}\big) \\
&\leq \mathbb{P}\big(X \leq x + \tfrac{1}{m}\big) + \mathbb{P}\big(|X_n - X| \geq \tfrac{1}{m}\big) \\
&= F_X\left(x + \tfrac{1}{m}\right) + \mathbb{P}\big(|X_n - X| \geq \tfrac{1}{m}\big).
\end{aligned}$$

Συνδυάζοντας τα προηγούμενα,

$$F_X\left(x - \tfrac{1}{m}\right) - \mathbb{P}\big(|X_n - X| \geq \tfrac{1}{m}\big) \leq F_{X_n}(x),$$

και

$$F_{X_n}(x) \le F_X\left(x+\tfrac{1}{m}\right) + \mathbb{P}\big(|X_n - X| \ge \tfrac{1}{m}\big).$$

Παίρνοντας όρια (του $n \to \infty$) για m σταθερό, έχουμε

$$F_X\left(x-\tfrac{1}{m}\right) \le \liminf_n F_{X_n}(x) \le \limsup_n F_{X_n}(x) \le F_X\left(x+\tfrac{1}{m}\right).$$

Φυσικά, για σταθερό x, τα $a = \liminf F_{X_n}(x)$, $b = \limsup F_{X_n}(x)$ είναι σταθεροί πραγματικοί αριθμοί, με $0 \le a \le b \le 1$. Άρα, για κάθε $m = 1, 2, \ldots$,

$$F_X\left(x-\tfrac{1}{m}\right) \le a \le b \le F_X\left(x+\tfrac{1}{m}\right),$$

και παίρνοντας όρια για $m \to \infty$, προκύπτει ότι $a = b = F_X(x)$ (αφού το x είναι σημείο συνεχείας της F_X), που σημαίνει ότι

$$\lim_n F_{X_n}(x) = F_X(x)$$

για κάθε σημείο συνεχείας x της F_X, δηλ. $X_n \xrightarrow{\mathbf{d}} X$.

(i) Θα αποδείξουμε στο επόμενο λήμμα ότι $X_n \xrightarrow{\mathbf{a.s.}} X$ εάν και μόνο εάν

$$\text{Για κάθε } \varepsilon > 0,\ \mathbb{P}\left(\sup_{k\ge n} |X_k - X| > \varepsilon\right) \to 0,\ \text{ καθώς } n \to \infty.$$

Προφανώς $\sup_{k\ge n} |X_k - X| \ge |X_n - X|$, και συνεπώς,

$$\{|X_n - X| > \varepsilon\} \subset \left\{\sup_{k\ge n} |X_k - X| > \varepsilon\right\},$$

από την οποία προκύπτει η αποδεικτέα. $\square$

Λήμμα 6.3 $X_n \xrightarrow{\text{a.s.}} X$ εάν και μόνο εάν

$$\text{Για κάθε } \varepsilon > 0,\ \mathbb{P}\left(\sup_{k\geq n}|X_k - X| > \varepsilon\right) \to 0,\ \text{ καθώς } n \to \infty,$$

δηλ. αν και μόνο αν

$$Y_n = \sup_{k\geq n}|X_k - X| \xrightarrow{\mathbf{P}} 0.$$

Απόδειξη: Κατ΄ αρχήν σημειώνουμε ότι εδώ, όπως και στο Θεώρημα 6.2, η X είναι τ.μ. **μη εκτεταμένη**, δηλ. $-\infty < X(\omega) < +\infty$ για κάθε $\omega \in \Omega$. Είναι σαφές ότι $X_n(\omega) \to X(\omega)$ αν και μόνο αν για κάθε $m = 1, 2, \ldots$, $|X_n(\omega) - X(\omega)| \leq \frac{1}{m}$ τελικά για κάθε n. Στην «γλώσσα» των ενδεχομένων, αυτό γράφεται ως εξής:

$$\begin{aligned}\{\omega : X_n(\omega) \to X(\omega)\} &= \bigcap_{m=1}^{\infty} \liminf_n \left\{\omega : |X_n - X| \leq \frac{1}{m}\right\} \\ &= \bigcap_{m=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} \left\{\omega : |X_k - X| \leq \frac{1}{m}\right\}.\end{aligned}$$

Έτσι,

$$\{\omega : X_n(\omega) \nrightarrow X(\omega)\} = \{\omega : X_n(\omega) \to X(\omega)\}^c = \bigcup_{m=1}^{\infty}\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} \left\{\omega : |X_k - X| > \frac{1}{m}\right\}.$$

Παρατηρούμε ότι $\bigcup_{k=n}^{\infty}\left\{|X_k - X| > \frac{1}{m}\right\} = \left\{\sup_{k\geq n}|X_k - X| > \frac{1}{m}\right\}$,
και ότι η ακολουθία ενδεχομένων $B_n^m = \bigcup_{k=n}^{\infty}\left\{|X_k - X| > \frac{1}{m}\right\}$
είναι φθίνουσα ως προς n, με

$$\lim_n B_n^m = B^m = \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\left\{|X_k - X| > \frac{1}{m}\right\}.$$

[Παρατηρήστε ότι $B^m = \limsup_n \{|X_n - X| > 1/m\}$.] Επομένως,

$$\{\omega : X_n(\omega) \nrightarrow X(\omega)\} = \bigcup_{m=1}^{\infty} B^m.$$

Αν τώρα $X_n \xrightarrow{\mathbf{a.s.}} X$, τότε έχουμε ότι $\mathbb{P}(X_n \nrightarrow X) = 0$, και συνεπώς,

$$\mathbb{P}(B^m) \leq \mathbb{P}\left(\bigcup_{m=1}^{\infty} B^m\right) = \mathbb{P}(X_n \nrightarrow X) = 0,$$

δηλ. $\mathbb{P}(B^m) = 0$ για κάθε m. Έτσι,

$$\lim_n \mathbb{P}\left(\sup_{k \geq n} |X_k - X| > \frac{1}{m}\right) = \lim_n \mathbb{P}(B_n^m) = \mathbb{P}\left(\lim_n B_n^m\right) = \mathbb{P}(B^m) = 0,$$

που σημαίνει ότι $\mathbb{P}\left(\sup_{k \geq n} |X_k - X| > \varepsilon\right) \to 0$ (του $n \to \infty$) για κάθε $\varepsilon > 0$, διότι για τυχόν $\varepsilon > 0$, μπορούμε να εκλέξουμε $m = [1/\varepsilon] + 1 > 1/\varepsilon$, έτσι ώστε

$$\left\{\sup_{k \geq n} |X_k - X| > \varepsilon\right\} \subset \left\{\sup_{k \geq n} |X_k - X| > \frac{1}{m}\right\} = B_n^m.$$

Αντίστροφα, ας υποθέσουμε ότι $\mathbb{P}\left(\sup_{k \geq n} |X_k - X| > \varepsilon\right) \to 0$ για κάθε $\varepsilon > 0$. Θέτοντας $\varepsilon = \frac{1}{m}$, προκύπτει ότι $\mathbb{P}\left(\sup_{k \geq n} |X_k - X| > \frac{1}{m}\right) \to 0$, δηλ. $\mathbb{P}(B_n^m) \to 0$ (καθώς $n \to \infty$). Αυτό σημαίνει ότι $\mathbb{P}(B^m) = \mathbb{P}\left(\lim_n B_n^m\right) = \lim_n \mathbb{P}(B_n^m) = 0$, και συνεπώς,

$$\mathbb{P}\left(\bigcup_{m=1}^{\infty} B^m\right) = 0,$$

δηλ. $\mathbb{P}(X_n \nrightarrow X) = 0$ ή, ισοδύναμα, $X_n \xrightarrow{\text{a.s.}} X$. $\square$

Κατ΄ αντιστοιχία, μπορεί να δειχθεί (ανισότητα Lyapounov) ότι εάν $X_n \xrightarrow{L^r} X$ για κάποιο $r \geq 1$, τότε $X_n \xrightarrow{L^p} X$ για κάθε $p \in [1, r]$, ενώ το αντίστροφο δεν ισχύει (Άσκηση 6.1). Θα πρέπει να σημειώσουμε ότι, γενικά, οι αντίστροφες συνεπαγωγές του Θεωρήματος 6.2 δεν ισχύουν. Για να αποδειχθεί αυτό χρειάζονται αντιπαραδείγματα (Άσκηση 6.1). Ισχύουν όμως τα εξής.

Πρόταση 6.4 (i) Αν ο Ω είναι αριθμήσιμος και $\mathcal{A} = \mathcal{P}(\Omega)$ τότε

$$X_n \xrightarrow{\mathbf{p}} X \iff X_n \xrightarrow{\text{a.s.}} X.$$

(ii) $X_n \xrightarrow{\mathbf{p}} X$ όταν και μόνο όταν για κάθε υπακολουθία $\{X_{n_k},\ k \geq 1\}$ της ακολουθίας $\{X_n,\ n \geq 1\}$, υπάρχει υποϋπακολουθία $\left\{X_{n_{k_i}},\ i \geq 1\right\}$ της $\{X_{n_k},\ k \geq 1\}$, τέτοια ώστε $X_{n_{k_i}} \xrightarrow{\text{a.s.}} X$ καθώς $i \to \infty$.
(iii) Αν $X_n \xrightarrow{\mathbf{d}} \alpha$ (πεπερασμένη σταθερά), και οι $X_n,\ X(\omega) \equiv \alpha$ είναι ορισμένες στον ίδιο χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, τότε $X_n \xrightarrow{\mathbf{p}} \alpha$.
[Λόγω του τελευταίου, και επειδή, από τα θεωρήματα ύπαρξης, μπορούμε πάντα να βρούμε έναν χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, και τ.μ. Y_n ισόνομες με τις X_n, και Y ισόνομη με την $X \equiv \alpha$, στον χώρο αυτόν, θεωρούμε τις συγκλίσεις $X_n \xrightarrow{\mathbf{d}} \alpha$ και $X_n \xrightarrow{\mathbf{p}} \alpha$ **ισοδύναμες**, χωρίς να απαιτούμε να ορίζονται οι X_n

και $X \equiv \alpha$ στον ίδιο χώρο. Έτσι, οι συμβολισμοί

$$X_n \xrightarrow{\mathbf{d}} \alpha, \quad X_n \xrightarrow{\mathbf{p}} \alpha, \quad \text{και} \quad F_n \xrightarrow{\mathbf{d}} F_{[\alpha]},$$

όπου $F_{[\alpha]} = I_{[\alpha,+\infty)}$, έχουν **ακριβώς** την ίδια σημασία.]

Απόδειξη: (i) Άσκηση 6.2. (iii) Άσκηση 6.2. (ii) Έστω ότι $X_n \xrightarrow{\mathbf{p}} X$, και ας θεωρήσουμε μία υπακολουθία $\{X_{n_k},\ k \geq 1\} \subset \{X_n,\ n \geq 1\}$. Η $p_{n_k} = \mathbb{P}(|X_{n_k} - X| \geq \varepsilon)$ είναι υπακολουθία της $p_n = \mathbb{P}(|X_n - X| \geq \varepsilon)$ (με $p_n \to 0$, από υπόθεση), και άρα, $p_{n_k} \to 0$ καθώς $k \to \infty$. Συνεπώς, $X_{n_k} \xrightarrow{\mathbf{p}} X$ καθώς $k \to \infty$. Άρα, για κάθε $\varepsilon > 0$, $\lim_{k\to\infty} \mathbb{P}(|X_{n_k} - X| \geq \varepsilon) = 0$. Επομένως, για $\varepsilon = 1/i$ (όπου i σταθερό, $i = 1, 2, \ldots$), μπορούμε να επιλέξουμε έναν αρκετά μεγάλο ακέραιο $k_0 = k_0(i)$, έτσι ώστε

$$\mathbb{P}\left(|X_{n_k} - X| \geq \frac{1}{i}\right) \leq \frac{1}{2^i}, \quad \text{για κάθε } k \geq k_0 = k_0(i).$$

Θέτουμε $k_1 = k_0(1)$, $k_2 = \max\{k_0(2), k_1 + 1\}$, $\ldots$, $k_i = \max\{k_0(i), k_{i-1} + 1\}$, $\ldots$, οπότε $1 \leq k_1 < k_2 < k_3 < \cdots$, οι k_i είναι ακέραιοι (για κάθε i), η $\{X_{n_{k_i}},\ i \geq 1\}$ είναι υπακολουθία της $\{X_{n_k},\ k \geq 1\}$ και, από κατασκευή, ισχύει ότι

$$\mathbb{P}\left(|X_{n_{k_i}} - X| \geq \frac{1}{i}\right) \leq \frac{1}{2^i}, \quad i = 1, 2, \ldots\ .$$

Επομένως,

$$\sum_{i=1}^{\infty} \mathbb{P}\left(|X_{n_{k_i}} - X| \geq \frac{1}{i}\right) \leq \sum_{i=1}^{\infty} \frac{1}{2^i} = 1 < \infty,$$

και από το πρώτο Λήμμα Borel-Cantelli,

$$\mathbb{P}\left(\limsup_i \left\{|X_{n_{k_i}} - X| \geq \frac{1}{i}\right\}\right) = 0.$$

Όμως,

$$\limsup_i \left\{\omega : |X_{n_{k_i}}(\omega) - X(\omega)| \geq \frac{1}{i}\right\} \supset \left\{\omega : X_{n_{k_i}}(\omega) \nrightarrow X(\omega), \quad \text{καθώς } i \to \infty\right\},$$

διότι, προφανώς,

$$\liminf_i \left\{\omega : |X_{n_{k_i}}(\omega) - X(\omega)| < \frac{1}{i}\right\} \subset \left\{\omega : X_{n_{k_i}}(\omega) \to X(\omega) \text{ καθώς } i \to \infty\right\},$$

και τα δύο τελευταία ενδεχόμενα είναι τα συμπληρώματα των προηγουμένων. Επομένως,

$$\mathbb{P}\left(X_{n_{k_i}} \nrightarrow X\right) = 0,$$

και συνεπώς, $X_{n_{k_i}} \xrightarrow{\text{a.s.}} X$ καθώς $i \to \infty$.

Αντιστρόφως, ας υποθέσουμε, αντιθέτως, ότι $X_n \overset{\mathbf{p}}{\nrightarrow} X$. Τότε, για κάποιο $\varepsilon > 0$, υπάρχει υπακολουθία $\{X_{n_k},\ k \geq 1\}$ της $\{X_n,\ n \geq 1\}$, τέτοια ώστε

$$\mathbb{P}\left(|X_{n_k} - X| > \varepsilon\right) \geq \varepsilon, \quad \text{για } k = 1, 2, \ldots .$$

Είναι σαφές ότι οποιαδήποτε υποϋπακολουθία $\{X_{n_{k_i}},\ i \geq 1\}$ της $\{X_{n_k},\ k \geq 1\}$ θα ικανοποιεί την ίδια σχέση, δηλ.

$$\mathbb{P}\left(|X_{n_{k_i}} - X| > \varepsilon\right) \geq \varepsilon, \quad \text{για } i = 1, 2, \ldots,$$

οπότε $X_{n_{k_i}} \overset{\mathbf{p}}{\nrightarrow} X$, και άρα, $X_{n_{k_i}} \overset{\text{a.s.}}{\nrightarrow} X$. □

Είναι προφανές ότι εάν $X_n \xrightarrow{\mathbf{a.s.}} X$ και, ταυτόχρονα, $X_n \xrightarrow{\mathbf{a.s.}} Y$, τότε οι τ.μ. X και Y είναι ίσες με πιθ. 1, επειδή

$$\{\omega : X_n(\omega) \to X(\omega)\} \cap \{\omega : X_n(\omega) \to Y(\omega)\} \subset \{\omega : X(\omega) = Y(\omega)\}.$$

Αυτό δείχνει την **μοναδικότητα του ορίου της ισχυρής σύγκλισης**. Άρα, οι σχέσεις $X_n \xrightarrow{\mathbf{a.s.}} X$ και $X_n \xrightarrow{\mathbf{a.s.}} Y$ συνεπάγονται την $X = Y$ με πιθ. 1, $\mathbb{P}(X = Y) = 1$, σχέση την οποία δηλώνουμε, πολλές φορές, και ως $X \stackrel{\mathbf{a.s.}}{=} Y$. Από το παραπάνω λήμμα έπεται άμεσα και το επόμενο πόρισμα, το οποίο δείχνει την μοναδικότητα των ορίων και στις άλλες περιπτώσεις, εκτός της ασθενούς σύγκλισης (για την οποία βλ. Άσκηση 6.4).

Πόρισμα 6.5 (μοναδικότητα των ορίων της L^p σύγκλισης, της σύγκλισης κατά πιθανότητα, και της σχεδόν βέβαιης σύγκλισης) Εάν $X_n \xrightarrow{\mathbf{p}} X$ και, ταυτόχρονα, $X_n \xrightarrow{\mathbf{p}} Y$, τότε $X \stackrel{\mathbf{a.s.}}{=} Y$.

[Το συμπέρασμα ισχύει και όταν μία (ή και οι δύο) από τις παραπάνω (κατά πιθανότητα) συγκλίσεις αντικατασταθεί με την σχεδόν βέβαιη (ισχυρή) αντίστοιχη, ή την L^p-αντίστοιχη, για κάποιο $p \geq 1$.]

Απόδειξη: Έστω ότι $X_n \xrightarrow{\mathbf{p}} X$ και, ταυτόχρονα, $X_n \xrightarrow{\mathbf{p}} Y$. Αφού $X_n \xrightarrow{\mathbf{p}} X$ τότε, λόγω της Πρότασης 6.4(ii), η $\{X_n,\ n \geq 1\}$ θα έχει υπακολουθία $\{X_{n_k},\ k \geq 1\}$, τέτοια ώστε $X_{n_k} \xrightarrow{\mathbf{a.s.}} X$ (του $k \to \infty$). Τότε όμως, αφού $X_n \xrightarrow{\mathbf{p}} Y$, και η X_{n_k} είναι υπακολουθία της X_n, προκύπτει, από την Πρόταση 6.4(ii) και πάλι, ότι η $\{X_{n_k},\ k \geq 1\}$ θα έχει υπακολουθία $\{X_{n_{k_i}},\ i \geq 1\}$, τέτοια ώστε $X_{n_{k_i}} \xrightarrow{\text{a.s.}} Y$ (του $i \to \infty$). Άρα, $X_{n_{k_i}} \xrightarrow{\text{a.s.}} X$ (ως υπακολουθία της $X_{n_k} \xrightarrow{\text{a.s.}} X$) και, ταυτόχρονα, $X_{n_{k_i}} \xrightarrow{\text{a.s.}} Y$, που αποδεικνύει το ζητούμενο. [Οι άλλες συγκλίσεις (ισχυρή και L^p) συνεπάγονται την κατά πιθανότητα σύγκλιση, λόγω του Θεωρήματος 6.2.] $\square$

Παρατήρηση 6.6 Σημειώνουμε ότι για τυχούσα ακολουθία *σ.κ.* $\{F_n,\ n \geq 1\}$, με $F_n \xrightarrow{\text{d}} F$ (όπου η F είναι *σ.κ.*), υπάρχει χώρος πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, στον οποίον μπορούμε να κατασκευάσουμε τ.μ. Y_n και Y, τέτοιες ώστε $Y_n(\omega) \to Y(\omega)$ για κάθε $\omega \in \Omega$, $F_{Y_n} = F_n$, και $F_Y = F$. [Αυτό είναι το περίφημο Θεώρημα του Skorohod, βλ. Κεφ. 7, Θεώρημα 7.15.] Μάλιστα, η κατασκευή αυτή μπορεί να γίνει στον χώρο $((0,1), \mathcal{B}((0,1)), \lambda)$, όπου, ορίζοντας $\widetilde{Y}_n(\omega) = \inf\{x : F_n(x) \geq \omega\}$, $\widetilde{Y}(\omega) = \inf\{x : F(x) \geq \omega\}$, $0 < \omega < 1$, αποδεικνύεται ότι $\widetilde{Y}_n(\omega) \to \widetilde{Y}(\omega)$, εκτός ίσως από ένα (το πολύ) αριθμήσιμο σύνολο, και συγκεκριμένα, το ενδεχόμενο $A = \{\omega : \widetilde{Y}_n(\omega) \nrightarrow \widetilde{Y}(\omega)\} \subset (0,1)$. Θέτοντας τότε $Y_n = \widetilde{Y}_n I_{A^c}$, $Y = \widetilde{Y} I_{A^c}$, έχουμε ότι $Y_n \stackrel{\text{a.s.}}{=} \widetilde{Y}_n$, $Y \stackrel{\text{a.s.}}{=} \widetilde{Y}$, και $Y_n(\omega) \to Y(\omega)$ για κάθε $\omega \in \Omega = (0,1)$. Προφανώς, αφού ισχύει ότι $Y_n \stackrel{\text{a.s.}}{=} \widetilde{Y}_n$, έπεται ότι $F_{Y_n} = F_{\widetilde{Y}_n} = F_n$, και ομοίως, $F_Y = F_{\widetilde{Y}} = F$, επειδή $Y \stackrel{\text{a.s.}}{=} \widetilde{Y}$.

Το Θεώρημα (εμφύτευσης) του Skorohod μας δίνει τη δυνατότητα να αναγάγουμε την κατά κατανομή σύγκλιση σε ισχυρή (σχεδόν βέβαιη) σύγκλιση, και έτσι, να απλοποιήσουμε πολλές από τις υποθέσεις που πρέπει να ικανοποιούνται όταν εφαρμόζουμε γνωστά θεωρήματα. Για παράδειγμα, εάν οι $\{X_n,\ n \geq 1\}$ είναι ομοιόμορφα ολοκληρώσιμες, και εάν $X_n \xrightarrow{\mathbf{d}} X$, τότε μπορούμε να συμπεράνουμε ότι και

$$\mathbb{E}(X_n) \to \mathbb{E}(X), \quad (\text{και ότι } \ \mathbb{E}|X| < \infty)$$

χωρίς να χρειάζεται οι X_n και X να είναι ορισμένες στον ίδιο χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, και ούτε, βέβαια, να απαιτείται η ισχυρή σύγκλιση $X_n \xrightarrow{\mathbf{a.s.}} X$ (αλλά μόνο η ασθενής σύγκλιση $X_n \xrightarrow{\mathbf{d}} X$, δηλ. $F_{X_n} \xrightarrow{\mathbf{d}} F_X$). Αυτό οφείλεται στο γεγονός ότι η συνθήκη για ομοιόμορφη ολοκληρωσιμότητα,

$$\sup_n \mathbb{E}[|X_n| I(|X_n| \geq \alpha)] \to 0, \quad \text{καθώς } \alpha \to +\infty,$$

εξαρτάται από τις X_n μόνο μέσω των αντιστοίχων *σ.κ.* F_n, αφού,

$$\mathbb{E}[|X_n| I(|X_n| \geq \alpha)] = \int_{|x| \geq \alpha} |x| dF_n(x).$$

Συνεπώς, οι εμφυτευμένες τ.μ. Y_n, Y, του Θεωρήματος Skorohod, θα είναι έτσι ώστε $Y_n \xrightarrow{\mathbf{a.s.}} Y$ (αφού $Y_n \to Y$ κατά σημείο), $Y_n \overset{\mathbf{d}}{=} X_n$ (δηλ. οι Y_n και X_n είναι ισόνομες), και $Y \overset{\mathbf{d}}{=} X$, οπότε $\mathbb{E}(X_n) = \mathbb{E}(Y_n)$, $\mathbb{E}(X) = \mathbb{E}(Y)$, και

$$\mathbb{E}[|X_n| I(|X_n| \geq \alpha)] = \mathbb{E}[|Y_n| I(|Y_n| \geq \alpha)].$$

Άρα, οι Y_n είναι ομοιόμορφα ολοκληρώσιμες, και $Y_n \xrightarrow{\textbf{a.s.}} Y$, που σημαίνει (από το κλασικό Θεώρημα Ομοιόμορφης Ολοκληρωσιμότητας – Θεώρημα ;;) ότι $\mathbb{E}|Y| < \infty$, και $\mathbb{E}(Y_n) \to \mathbb{E}(Y)$. Τότε όμως, λόγω ισονομίας, $\mathbb{E}(X_n) = \mathbb{E}(Y_n) \to \mathbb{E}(Y) = \mathbb{E}(X)$ (και $\mathbb{E}|X| = \mathbb{E}|Y| < \infty$), δηλαδή $\mathbb{E}(X_n) \to \mathbb{E}(X)$. [Φυσικά, δεν μπορούμε να συμπεράνουμε την $\mathbb{E}|X_n - X| \to 0$, αν και ισχύει ότι $\mathbb{E}|Y_n - Y| \to 0$. (γιατί;)]

Ανάλογες εφαρμογές μπορούν να γίνουν και στα Θεωρήματα Μονότονης και Κυριαρχημένης Σύγκλισης. Φυσικά, για να είναι δυνατή η εφαρμογή των θεωρημάτων αυτών στην περίπτωση που, απλώς, $X_n \xrightarrow{\textbf{d}} X$, θα πρέπει να γίνουν κατάλληλες τροποποιήσεις όσον αφορά στην διάταξη τυχαίων μεταβλητών, αφού δεν απαιτείται τώρα οι X_n (και X) να είναι ορισμένες στον ίδιο χώρο πιθανότητας. Συγκεκριμένα, η «σχεδόν βέβαιη» διάταξη, $X_1 \le X_2$ με πιθ. 1, θα πρέπει να αντικατασταθεί από την **στοχαστική διάταξη**, $F_{X_1}(x) \ge F_{X_2}(x)$, για κάθε $x \in \mathbb{R}$.

Για παράδειγμα, εάν $F_{X_n} \xrightarrow{\textbf{d}} F_X$, $F_{|X_n|} \ge F_{|Y|}$ για κάθε n, και η Y είναι ολοκληρώσιμη, δηλ. $\mathbb{E}|Y| \left(= \int_{\mathbb{R}} |y| dF_Y(y)\right) < \infty$, τότε ισχύει ότι $\mathbb{E}|X| < \infty$ (δηλ. και η X είναι ολοκληρώσιμη), και μάλιστα, $\mathbb{E}(X_n) \to \mathbb{E}(X)$. [Θεώρημα Κυριαρχημένης Σύγκλισης, όπου η συνθήκη $X_n \xrightarrow{\textbf{a.s.}} X$ αντικαταστάθηκε από την $X_n \xrightarrow{\textbf{d}} X$, η δε συνθήκη $|X_n| \le |Y|$ για κάθε n, αντικαταστάθηκε από την στοχαστική της ανάλογη, $F_{|X_n|} \ge F_{|Y|}$ για κάθε n.]

Η στοχαστική διάταξη, $F_{X_1}(x) \geq F_{X_2}(x)$ για κάθε $x \in \mathbb{R}$, δηλώνεται συνήθως στην βιβλιογραφία ως $X_1 \leq_{\mathrm{st}} X_2$, και ισοδυναμεί με την $F_{X_1}^{-1}(t) \leq F_{X_2}^{-1}(t)$ για κάθε $t \in (0,1)$, όπου $F^{-1}(t) = \inf\{x : F(x) \geq t\}$, $0 < t < 1$, είναι η γενικευμένη αντίστροφη τυχούσας συνάρτησης κατανομής F – βλ. Θεώρημα ;; και Άσκηση ;;.2. Έτσι, όταν $X_1 \leq_{\mathrm{st}} X_2$, τότε οι αντίστοιχες εμφυτευμένες τ.μ. Y_1, Y_2, του Θεωρήματος Skorohod, θα ικανοποιούν την $Y_1 \leq Y_2$, με πιθανότητα 1. Για περισσότερες λεπτομέρειες βλ. Κεφ. 7, Ορισμός 7.16, Προτάσεις 7.17, 7.19, 7.24, και Θεωρήματα 7.20, 7.21, 7.23.

6.2 Ασθενής Νόμος Μεγάλων Αριθμών

Αναμφίβολα, οι Νόμοι των Μεγάλων Αριθμών (Laws of Large Numbers) κατέχουν κεντρική θέση στην Θεωρία των Πιθανοτήτων. Εκτός από τις πολλές και ενδιαφέρουσες εφαρμογές τους, περιγράφουν, με την αναγκαία μαθηματική αυστηρότητα, την έννοια της **Στατιστικής Ισορροπίας**, αυτοπροσδιορίζοντας ακόμα και την έννοια της ίδιας της πιθανότητας. Στην παράγραφο αυτή θα πάρουμε μία πρώτη ιδέα για την σπουδαιότητά τους, μελετώντας την πιο απλή περίπτωση (πεπερασμένες διασπορές) του ασθενούς νόμου.

Γενικά, λέμε ότι η ακολουθία τ.μ. $\{X_n,\ n \geq 1\}$ ικανοποιεί έναν **νόμο μεγάλων αριθμών**, όταν υπάρχουν ακολουθίες $a_n \in \mathbb{R}$, και $b_n > 0$, τέτοιες ώστε

$$\frac{X_1 + \ldots + X_n - a_n}{b_n} \xrightarrow{\mathbf{p}} 0, \quad \text{ή} \quad \frac{X_1 + \ldots + X_n - a_n}{b_n} \xrightarrow{\mathbf{a.s.}} 0.$$

Φυσικά, η δεύτερη σχέση είναι ισχυρότερη από την πρώτη (η οποία, στην ουσία, είναι ισοδύναμη με την κατά κατανομή σύγκλιση), και γι΄ αυτό μιλάμε για **ισχυρό νόμο μεγάλων αριθμών** (σχεδόν βέβαιη/ισχυρή σύγκλιση – strong law of large numbers – SLLN), και για **ασθενή νόμο** (κατά πιθανότητα/κατά κατανομή/ασθενής σύγκλιση – weak law of large numbers – WLLN).

Είναι τώρα πολύ εύκολο να δειχθεί το εξής θεώρημα.

Θεώρημα 6.7 (ασθενής νόμος μεγάλων αριθμών) Έστω $\{X_n, n \geq 1\}$ μία ανεξάρτητη ακολουθία τ.μ., με $\mathbb{E}(X_n^2) < \infty$, $\mathbb{E}(X_n) = \mu_n$ $\mathrm{Var}(X_n) = \sigma_n^2$ (για κάθε n), και ας υποθέσουμε ότι

$$\frac{1}{n^2} \sum_{j=1}^{n} \sigma_j^2 \to 0, \quad \text{καθώς } n \to \infty.$$

Τότε, για τα μερικά αθροίσματα $S_n = X_1 + \ldots + X_n$, ισχύει ότι

$$\frac{S_n}{n} - \frac{\mu_1 + \ldots + \mu_n}{n} \xrightarrow{\mathbf{p}} 0, \quad \text{καθώς } n \to \infty.$$

Απόδειξη: Έστω $\varepsilon > 0$. Αφού, λόγω γραμμικότητας της μέσης τιμής, ισχύει ότι $\mathbb{E}(S_n) = \mu_1 + \ldots + \mu_n$, η ανισότητα

Chebychev (Πόρισμα ;;) δίνει

$$\mathbb{P}\left(\left|\frac{S_n - \mathbb{E}(S_n)}{n}\right| \geq \varepsilon\right) = \mathbb{P}(|S_n - \mathbb{E}(S_n)| \geq n\varepsilon) \leq \frac{\mathrm{Var}(S_n)}{n^2\varepsilon^2} = \frac{1}{\varepsilon^2}\frac{1}{n^2}\sum_{j=1}^{n}\sigma_j^2 \to 0,$$

επειδή $\mathrm{Var}(S_n) = \sigma_1^2 + \ldots + \sigma_n^2$, διότι οι X_n είναι ανεξάρτητες (βλ. Εφαρμογή ;;). $\square$

Η ποσότητα $S_n/n = (X_1 + \ldots + X_n)/n$, η οποία είναι κεντρικής σημασίας για τις Πιθανότητες και την Στατιστική, παριστάνει τον μέσο όρο των $X_1, X_2, \ldots, X_n$, και, επομένως, ονομάζεται **δειγματικός μέσος** (sample mean), συμβολίζεται δε με $\overline{X}_n$ (ή και $\overline{X}$, όταν δεν είναι αναγκαίο να δηλωθεί ρητά ο αριθμός, n, των προσθετέων).

Άμεσα προκύπτει το εξής πόρισμα.

Πόρισμα 6.8 (ασθενής νόμος μεγάλων αριθμών) Εάν οι X_n είναι ανεξάρτητες και ισόνομες, με[1] $\mathbb{E}(X_1) = \mu$, $\mathrm{Var}(X_1) = \sigma^2 < \infty$, τότε

$$\overline{X}_n \overset{\text{ορ.}}{=\!=} \frac{1}{n}\sum_{j=1}^{n} X_j \xrightarrow{\mathbf{P}} \mu.$$

Απόδειξη: Λόγω ισονομίας είναι $\mathbb{E}(X_n) = \mu_n = \mu$ και $\mathrm{Var}(X_n) =$

[1] Είναι πάγια τακτική στις πιθανότητες η μέση τιμή μιας τ.μ. να συμβολίζεται με μ, και η διασπορά της με σ^2. Όπως ίσως έγινε αντιληπτό, αποφύγαμε μέχρι τώρα αυτόν τον συμβολισμό, επειδή με το γράμμα μ, γενικά, δηλώναμε κάποιο μέτρο σε αντίστοιχο χώρο μέτρου, το δε σ πολλές φορές παρίστανε κάποια σ-άλγεβρα. Στο κεφάλαιο αυτό θα χρησιμοποιούμε τον συνήθη συμβολισμό, δηλ. το μ θα δηλώνει μέση τιμή και το σ^2 διασπορά, αφού περιοριζόμαστε αυστηρά σε μέτρα πιθανότητας, και έτσι δεν υπάρχει κίνδυνος σύγχυσης.

$\sigma_n^2 = \sigma^2$ για κάθε n, οπότε

$$\frac{1}{n^2}\sum_{j=1}^{n}\sigma_j^2 = \frac{\sigma^2}{n} \to 0,$$

και, με άμεση εφαρμογή του Θεωρήματος 6.7, παίρνουμε την

$$\overline{X}_n - \mu \xrightarrow{\mathbf{p}} 0.$$

Τώρα, η σχέση $\overline{X}_n \xrightarrow{\mathbf{p}} \mu$, που ισοδυναμεί με την $\overline{X}_n \xrightarrow{\mathbf{d}} \mu$, έπεται άμεσα αν γράψουμε $\overline{X}_n = (\overline{X}_n - \mu) + \mu$, και εφαρμόσουμε την παρακάτω χρήσιμη και απλή πρόταση, γνωστή ως Θεώρημα Slutsky. $\square$

Πρόταση 6.9 (Θεώρημα Slutsky) Έστω ότι $X_n \xrightarrow{\mathbf{d}} X$ και $Y_n \xrightarrow{\mathbf{d}} c$, όπου $c \in \mathbb{R}$, σταθερά. [Οι X_n, Y_n, X μπορεί να έχουν (ή να μην έχουν) οποιαδήποτε σχέση μεταξύ τους.][2] Τότε ισχύουν τα εξής:

(i) $X_n + Y_n \xrightarrow{\mathbf{d}} X + c$.

(ii) $X_n Y_n \xrightarrow{\mathbf{d}} cX$.

(iii) Αν $c \neq 0$, τότε $X_n/Y_n \xrightarrow{\mathbf{d}} X/c$.

Απόδειξη: Άσκηση 6.3. $\square$

Στο Πόρισμα 6.8 υποθέσαμε πεπερασμένη διασπορά για την X_1. Αυτό δεν είναι αναγκαίο διότι, όπως θα δούμε στην συνέχεια, υπό τις ίδιες συνθήκες (και ακόμα ασθενέστερες) ισχύει

[2] φυσικά εννοείται ότι κάθε ζεύγος (X_n, Y_n) είναι ορισμένο στον ίδιο χώρο πιθανότητας $(\Omega_n, \mathcal{A}_n, \mathbb{P}_n)$, ώστε να έχουν νόημα οι πράξεις $X_n(\omega) + Y_n(\omega)$, $X_n(\omega)Y_n(\omega)$, και $X_n(\omega)/Y_n(\omega)$.

Ο ισχυρός νόμος μεγάλων αριθμών,

$$\overline{X}_n \xrightarrow{\mathbf{a.s.}} \mu.$$

Για την ακρίβεια, η μόνη υπόθεση που απαιτείται ώστε να έχουμε ισχυρή σύγκλιση, είναι η ύπαρξη της μέσης τιμής, ότι δηλ. $\mathbb{E}(X_1) = \mu \in \mathbb{R}$. Ο Khinchine έδωσε μία απόδειξη του ασθενούς νόμου ($\overline{X}_n \xrightarrow{\mathbf{P}} \mu$), χωρίς την υπόθεση πεπερασμένης διασποράς, χρησιμοποιώντας περικοπή τ.μ. (τεχνική την οποία θα χρησιμοποιήσουμε και παρακάτω, όταν θα αποδείξουμε τον ισχυρό νόμο). Ο πιο εύκολος τρόπος για να δειχθεί ο ασθενής νόμος, σε αυτό το γενικότερο πλαίσιο, είναι με χρήση χαρακτηριστικών συναρτήσεων (βλ. Κεφ. ;;, Θεώρημα ;;).

Παράδειγμα 6.10 Ας υποθέσουμε ότι ένα πείραμα $(\Omega, \mathcal{A}, \mathbb{P})$ εκτελείται «άπειρες φορές», και μας ενδιαφέρει να προσδιορίσουμε την $\mathbb{P}(A)$, για δοθέν ενδεχόμενο $A \in \mathcal{A}$. Θεωρώντας τον χώρο (βλ. Κεφ. ;;)

$$(\Omega^\infty, \mathcal{A}^\infty, \mathbb{P}^\infty),$$

έχουμε ότι

$$\begin{aligned}
\mathbb{P}^\infty(\overbrace{\Omega \times \cdots \times \Omega}^{n-1} \times A \times \Omega \times \cdots) &= \mathbb{P}^n(\Omega \times \cdots \times \Omega \times A) \\
&= (\mathbb{P} \times \cdots \times \mathbb{P})(\Omega \times \cdots \times \Omega \times A) \\
&= \mathbb{P}(\Omega) \cdots \mathbb{P}(\Omega)\mathbb{P}(A) \\
&= \mathbb{P}(A).
\end{aligned}$$

Άρα, οι τ.μ.

$$X_n(\omega) = X_n(\omega_1, \omega_2, \ldots) = \begin{cases} 1, & \text{αν } \omega_n \in A, \\ 0, & \text{αν } \omega_n \notin A, \end{cases}$$

είναι ανεξάρτητες και ισόνομες, και μάλιστα, είναι δείκτριες τ.μ., αφού $X_n = I_{A_n}$, όπου

$$A_n = \underbrace{\Omega \times \cdots \times \Omega}_{n-1} \times A \times \Omega \times \cdots \in \mathcal{A}^\infty.$$

Προφανώς, $\mathbb{E}(X_n) = \mathbb{P}(A)$, και $\mathrm{Var}(X_n) = \mathbb{P}(A)(1 - \mathbb{P}(A))$, και, βέβαια, η

$$S_n = X_1 + \ldots + X_n$$

παριστάνει το πλήθος εμφανίσεων του A στις πρώτες n ανεξάρτητες δοκιμές, οπότε[3] $S_n \sim \mathbf{b}(n, \mathbb{P}(A))$, με $\mathbb{P}(S_n = k) = \binom{n}{k} (\mathbb{P}(A))^k (1 - \mathbb{P}(A))^{n-k}$, $k = 0, 1, \ldots, n$. Φυσικά, λόγω του ασθενούς νόμου, ισχύει ότι

$$\overline{X}_n = \frac{S_n}{n} \xrightarrow{\mathbf{p}} \mathbb{P}(A). \text{ Δηλαδή: Για κάθε } \varepsilon > 0,\ \mathbb{P}\left(\left|\overline{X}_n - \mathbb{P}(A)\right| \geq \varepsilon\right) \to 0.$$

Με απλά λόγια, καθώς $n \to \infty$, η πιθανότητα όπως ο δειγματικός μέσος, $\overline{X}_n$ (που παριστάνει την σχετική συχνότητα εμφανίσεως του A στις n πρώτες δοκιμές), διαφέρει από την $\mathbb{P}(A)$ περισσότερο από $\varepsilon > 0$, τείνει στο 0 (για κάθε $\varepsilon > 0$), δηλ. η πιθανότητα αυτή γίνεται, τελικά, αμελητέα. Αυτό επιβεβαιώνει την έννοια της στατιστικής ισορροπίας, μέσω της

[3] Ο συμβολισμός $X \sim F$ δηλώνει ότι «η τ.μ. X έχει συνάρτηση κατανομής F». Επίσης, $\mathbf{b}(n, p)$ είναι η συνάρτηση κατανομής της διωνυμικής, με παραμέτρους n και p, η οποία περιγράφει τον αριθμό των επιτυχιών σε n ανεξάρτητες δοκιμές Bernoulli(p), καθεμιά με πιθανότητα επιτυχίας p.

οποίας ο von Mises επιχείρησε να ορίσει την $\mathbb{P}(A)$, βλ. Κεφ. 1, αλλά η σημαντική διαφορά είναι ότι, τώρα, η στατιστική ισορροπία είναι **θεώρημα**, που προκύπτει, απλώς, από τα αξιώματα της πιθανοθεωρίας. Επιπλέον, παρατηρούμε ότι **δεν εξασφαλίζεται** η μαθηματική (κατά σημείο) σύγκλιση, $\overline{X}_n(\omega) \to \mathbb{P}(A)$, για κάθε ω (και αυτό δεν συμβαίνει ούτε καν στον ισχυρό νόμο των μεγάλων αριθμών). Το καλύτερο που μπορούμε να ελπίζουμε ότι θα ισχύει (και το οποίο, πράγματι, ισχύει) είναι ότι $\overline{X}_n(\omega) \xrightarrow{\textbf{a.s.}} \mathbb{P}(A)$ (δηλ. σχεδόν βέβαιη σύγκλιση). $\square$

Μία πολύ όμορφη εφαρμογή (στην μαθηματική ανάλυση) είναι η κατασκευή των **πολυωνύμων Bernstein**, βαθμού n, που αντιστοιχούν στην f (με πεδίο ορισμού το διάστημα $[0, 1]$):

$$B_n(x; f) = \sum_{k=0}^{n} f(\tfrac{k}{n}) \binom{n}{k} x^k (1-x)^{n-k}, \quad 0 \leq x \leq 1.$$

Τότε έχουμε το εξής πόρισμα.

Πόρισμα 6.11 Εάν η f είναι συνεχής στο διάστημα $[0,1]$, τότε[4]

$$\max_{x\in[0,1]} |f(x) - B_n(x;f)| \to 0, \quad \text{καθώς } n \to \infty.$$

Απόδειξη: Η «ιδέα» είναι κρυμμένη στην εξής παρατήρηση. Εάν οι $X_1, X_2, \ldots$ είναι ανεξάρτητες και ισόνομες Bernoulli(x) τ.μ. $(0 \le x \le 1)$, τότε η $S_n = X_1 + \ldots + X_n \sim \mathbf{b}(n,x)$, οπότε

$$\mathbb{E}\left[f\left(\frac{S_n}{n}\right)\right] = \sum_{k=0}^{n} f(\tfrac{k}{n})\binom{n}{k} x^k (1-x)^{n-k},$$

που είναι το $B_n(x;f)$. Όμως, $S_n/n \xrightarrow{\mathbf{p}} x$, από τον ασθενή νόμο των μεγάλων αριθμών, και άρα, «περιμένουμε» ότι και

$$\mathbb{E}\left[f\left(\frac{S_n}{n}\right)\right] \to \mathbb{E}[f(x)] = f(x), \ \text{ομοιόμορφα ως προς } x, \ \text{καθώς } n \to \infty.$$

[4]Επειδή η f είναι συνεχής, το $B_n(x;f)$ είναι πολυώνυμο (άρα συνεχής συνάρτηση), και το $[0,1]$ είναι συμπαγές, έπεται ότι η συνεχής συνάρτηση $g_n(x) = |f(x) - B_n(x;f)|$, $x \in [0,1]$, μεγιστοποιείται για κάποιο (ή κάποια) $x_n \in [0,1]$, και η μέγιστη τιμή, $g_n(x_n)$, είναι ένας, καλά ορισμένος, πραγματικός αριθμός ($g_n(x_n) < \infty$). Το Πόρισμα 6.11 μας εξασφαλίζει την **ομοιόμορφη** σύγκλιση, μιας ακολουθίας πολυωνύμων, προς την συνεχή f (στο $[0,1]$), και μάλιστα, κατασκευάζει ακριβώς τα πολυώνυμα αυτά (προσεγγιστικά πολυώνυμα), για δοθείσα f. Φυσικά, το Θεώρημα Weierstrass (ότι κάθε συνεχής συνάρτηση προσεγγίζεται ομοιόμορφα από πολυώνυμα στο $[a,b]$), όχι μόνο προκύπτει από το Πόρισμα 6.11, αλλά, επιπλέον, βρίσκεται ακριβής, και σχετικά απλή, κατασκευή τέτοιων πολυωνύμων. Παρατηρήστε ότι το $[0,1]$ μπορεί να μεταφερθεί, με μία απλή γραμμική απεικόνιση, σε οποιοδήποτε συμπαγές διάστημα, οπότε το συμπέρασμα ισχύει για κάθε διάστημα $[a,b]$. Συγκεκριμένα, αν η f είναι συνεχής στο $[a,b]$, τότε, θέτοντας $\widetilde{f}(x) = f(a+(b-a)x)$, $x \in [0,1]$, βλέπουμε ότι η $\widetilde{f}$ είναι συνεχής στο $[0,1]$, οπότε

$$\max_{x\in[0,1]} |\widetilde{f}(x) - B_n(x;\widetilde{f})| \to 0, \quad \text{καθώς } n \to \infty,$$

(από το συμπέρασμα για το διάστημα $[0,1]$), που σημαίνει ότι και

$$\max_{y\in[a,b]} \left| f(y) - B_n\left(\frac{y-a}{b-a};\widetilde{f}\right)\right| \to 0, \quad \text{καθώς } n \to \infty,$$

και το προσεγγιστικό πολυώνυμο, βαθμού n, για την f, στο $[a,b]$, λαμβάνει την μορφή

$$P_n(y) = \frac{1}{(b-a)^n}\sum_{k=0}^{n} f(a+\tfrac{k}{n}(b-a))\binom{n}{k}(y-a)^k(b-y)^{n-k}, \quad \text{για } a \le y \le b.$$

Οι λεπτομέρειες είναι οι εξής: Αφού η f είναι συνεχής στο $[0,1]$, θα είναι ομοιόμορφα φραγμένη (δηλ. $|f(x)| \leq M < \infty$ για κάποιο M, και για κάθε $x \in [0,1]$), και ομοιόμορφα συνεχής (δηλ. για $\varepsilon > 0$, $|f(x) - f(y)| \leq \varepsilon$, όταν $|y - x| \leq \delta = \delta(\varepsilon)$, και $x, y \in [0,1]$). Έστω λοιπόν $\varepsilon > 0$ (αυθαίρετα μικρό), και έστω $\delta = \delta(\varepsilon) > 0$ το δ που απαιτείται έτσι ώστε $|f(x) - f(y)| \leq \varepsilon$ όταν $|x - y| \leq \delta$. Έστω επίσης M, $0 \leq$

$M < \infty$, ένα φράγμα της $|f|$. Τότε έχουμε:

$$\begin{aligned}
|f(x) - B_n(x;f)| &= \left| f(x) - \sum_{k=0}^{n} f(\tfrac{k}{n}) \binom{n}{k} x^k (1-x)^{n-k} \right| \\
&= \left| \sum_{k=0}^{n} \left(f(x) - f(\tfrac{k}{n}) \right) \binom{n}{k} x^k (1-x)^{n-k} \right| \\
&\le \sum_{k=0}^{n} \left| f(x) - f(\tfrac{k}{n}) \right| \binom{n}{k} x^k (1-x)^{n-k} \\
&= \left(\sum_{k: \left|\frac{k}{n} - x\right| \le \delta} + \sum_{k: \left|\frac{k}{n} - x\right| > \delta} \right) \left| f(x) - f(\tfrac{k}{n}) \right| \binom{n}{k} x^k (1-x)^{n-k} \\
&\le \sum_{k: \left|\frac{k}{n} - x\right| \le \delta} \varepsilon \binom{n}{k} x^k (1-x)^{n-k} \\
&\qquad + \sum_{k: \left|\frac{k}{n} - x\right| > \delta} \{ |f(x)| + |f(\tfrac{k}{n})| \} \binom{n}{k} x^k (1-x)^{n-k} \\
&\le \sum_{k=0}^{n} \varepsilon \binom{n}{k} x^k (1-x)^{n-k} + \sum_{k: \left|\frac{k}{n} - x\right| > \delta} 2M \binom{n}{k} x^k (1-x)^{n-k} \\
&= \varepsilon + 2M \, \mathbb{P} \left[\left| \frac{S_n}{n} - \mathbb{E}\left(\frac{S_n}{n} \right) \right| > \delta \right] \qquad \text{(ανισότητα Chebychev)} \\
&\le \varepsilon + 2M \frac{\mathrm{Var}(S_n/n)}{\delta^2} = \varepsilon + \frac{2M}{\delta^2} \, \frac{x(1-x)}{n} \\
&\le \varepsilon + \frac{M}{2n\delta^2},
\end{aligned}$$

όπου, στον παραπάνω υπολογισμό, δικαιολογείται η χρήση της ανισότητας Chebychev, επειδή

$$\mathbb{E}\left(\frac{S_n}{n}\right) = x, \quad \text{και} \quad \mathrm{Var}\left(\frac{S_n}{n}\right) = \frac{x(1-x)}{n} \quad \left(\leq \frac{1}{4n}\right).$$

Επομένως,

$$a_n = \max_{x\in[0,1]} |f(x) - B_n(x; f)| \leq \varepsilon + \frac{M}{2n\delta^2} \to \varepsilon, \ \text{καθώς}\ n \to \infty,$$

που σημαίνει ότι $\limsup a_n \leq \varepsilon$ (φυσικά, $a_n \geq 0$). Επειδή το $\varepsilon > 0$ είναι τυχόν, έπεται ότι $\limsup a_n = 0$, δηλαδή $a_n \to 0$. $\square$

Στην πραγματικότητα, η απόδειξη στηρίχθηκε μόνο στην ανισότητα Chebychev (και όχι στον ασθενή νόμο των μεγάλων αριθμών), αλλά η ιδέα κρύβεται στον ασθενή νόμο. Άλλωστε, και ο ίδιος ο ασθενής νόμος είναι άμεσο πόρισμα της ανισότητας Chebychev, όταν η διασπορά είναι πεπερασμένη.

6.3 Ισχυρός Νόμος Μεγάλων Αριθμών για Ανεξάρτητη Ακολουθία

Στην συνέχεια θα επικεντρωθούμε στην απόδειξη του ισχυρού νόμου των μεγάλων αριθμών (και χωρίς την υπόθεση πεπερασμένων διασπορών). Για τον σκοπό αυτό χρειαζόμαστε μία ανισότητα ισχυρότερη από την Chebychev.

Πρόταση 6.12 (μεγιστική ανισότητα Kolmogorov) Εάν οι τ.μ. $X_1, X_2, \ldots, X_n$ είναι ανεξάρτητες, και $\mathbb{E}|X_j| < \infty$ για $j = 1, 2, \ldots, n$, τότε για κάθε $a > 0$,

$$\mathbb{P}\left[\max_{1\le j\le n} |S_j - \mathbb{E}(S_j)| \ge a\right] \le \frac{1}{a^2}\mathrm{Var}(S_n),$$

όπου $S_j = X_1 + \ldots + X_j$, $j = 1, \ldots, n$.

Απόδειξη: Αν $\mathrm{Var}(S_n) = \infty$, δεν έχουμε τίποτα να δείξουμε. Ας υποθέσουμε λοιπόν ότι $\mathrm{Var}(S_n) < \infty$, που σημαίνει ότι $\mathbb{E}(X_1 + \ldots + X_n)^2 < \infty$, οπότε και[5]

$$\mathbb{E}(X_1^2 + \ldots + X_n^2) < \infty.$$

Επομένως, $\mathbb{E}\left(X_j^2\right) < \infty$ για $j = 1, 2, \ldots, n$, δηλ. $\mathrm{Var}(X_j) < \infty$ για $j = 1, 2, \ldots, n$.

Χωρίς βλάβη της γενικότητας υποθέτουμε[6] ότι $\mathbb{E}(X_j) = 0$ για κάθε j. Τότε έχουμε ότι

$$\mathrm{Var}(S_n) = \mathbb{E}(S_n^2) = \int_\Omega S_n^2 d\mathbb{P}.$$

---

[5] Αυτό αποδεικνύεται ως εξής: Η X_iX_j $(i < j)$ είναι ολοκληρώσιμη, επειδή οι X_i και X_j είναι ανεξάρτητες, και $\mathbb{E}|X_i| < \infty$, $\mathbb{E}|X_j| < \infty$, οπότε και $\mathbb{E}|X_iX_j| = \mathbb{E}|X_j|\mathbb{E}|X_j| < \infty$, από το Θεώρημα ;;. Αφού $(X_1 + \ldots + X_n)^2 = X_1^2 + \ldots + X_n^2 + 2\sum_{i<j} X_iX_j \ge X_1^2 + \ldots + X_n^2 - 2\sum_{i<j} |X_iX_j|$, δηλ.

$$X_1^2 + \ldots + X_n^2 \le (X_1 + \ldots + X_n)^2 + 2\sum_{i<j} |X_iX_j|,$$

έπεται ότι και η $X_1^2 + \ldots + X_n^2$ είναι ολοκληρώσιμη.

[6] Αν αυτό δεν συμβαίνει, τότε θέτουμε $\widetilde{X}_j = X_j - \mathbb{E}(X_j)$, οπότε $\widetilde{S}_j = \widetilde{X}_1 + \ldots + \widetilde{X}_j = S_j - \mathbb{E}(S_j)$, και άρα, αν αποδείξουμε ότι η πρόταση ισχύει όταν $\mathbb{E}(X_j) = 0$, $j = 1, 2, \ldots, n$, τότε η πρόταση θα ισχύει και για τις τ.μ. $\widetilde{X}_j$, δηλ.

$$\mathbb{P}\left[\max_{1\le j\le n} |S_j - \mathbb{E}(S_j)| \ge a\right] = \mathbb{P}\left[\max_{1\le j\le n} |\widetilde{S}_j| \ge a\right] \le \frac{1}{a^2}\mathrm{Var}(\widetilde{S}_n) = \frac{1}{a^2}\mathrm{Var}(S_n),$$

αφού $\mathrm{Var}(\widetilde{S}_n) = \mathrm{Var}(S_n)$.

Έστω[7] $A_k = \{|S_1| < a, |S_2| < a, \ldots, |S_{k-1}| < a, |S_k| \geq a\}$, $k = 1, 2, \ldots, n$, οπότε $A = \{\max_{1\leq k\leq n} |S_k| \geq a\} = \bigcup_{k=1}^n A_k$, με $A_iA_j = \varnothing$ για $i \neq j$. Προφανώς,

$$\begin{aligned}\mathrm{Var}\,(S_n) &= \int_\Omega S_n^2 d\mathbb{P} \geq \int_A S_n^2 d\mathbb{P} \overset{A_k \text{ ξένα}}{=\!=\!=} \sum_{k=1}^n \int_{A_k} S_n^2 d\mathbb{P} \\ &= \sum_{k=1}^n \int_{A_k} (S_k + S_n - S_k)^2 d\mathbb{P} \geq \sum_{k=1}^n \int_{A_k} [S_k^2 + 2S_k(S_n - S_k)] d\mathbb{P}\,.\end{aligned}$$

[Η τελευταία ανισότητα προέρχεται από την $(a+b)^2 \geq a^2 + 2ab$.] Εφαρμόζοντας την ανισότητα Cauchy-Schwarz, και την $(a - b)^2 \leq 2a^2 + 2b^2$, έχουμε

$$\begin{aligned}\int_\Omega |S_k||S_n - S_k| d\mathbb{P} &= \mathbb{E}(|S_k||S_n - S_k|) \\ &\leq \sqrt{\mathbb{E}\,(S_k^2)\,\mathbb{E}\,(S_n - S_k)^2} \\ &\leq \sqrt{\mathbb{E}\,(S_k^2)\,[2\,\mathbb{E}\,(S_n^2) + 2\,\mathbb{E}\,(S_k^2)]},\end{aligned}$$

και συνεπώς,

$$\mathbb{E}(|S_k||S_n - S_k|) < \infty,$$

διότι[8] $\mathbb{E}(S_j^2) = \mathrm{Var}\,(S_j) = \mathrm{Var}\,(X_1) + \ldots + \mathrm{Var}\,(X_j) = \mathbb{E}(X_1^2) + \ldots + \mathbb{E}(X_j^2) < \infty$. Επομένως, από την γραμμικότητα

[7] εδώ $A_1 = \{|S_1| \geq a\} = \{|X_1| \geq a\}$

[8] ή, αν προτιμάτε, χρησιμοποιήστε την ανισότητα $S_j^2 \leq j(X_1^2 + \ldots + X_j^2) \leq j(X_1^2 + \ldots + X_n^2)$

του ολοκληρώματος, προκύπτει ότι

$$\begin{aligned}\int_{A_k}[S_k^2+2S_k(S_n-S_k)]d\mathbb{P} &= \int_{A_k} S_k^2 d\mathbb{P} + 2\int_{A_k} S_k(S_n-S_k)d\mathbb{P}\\ &= \int_{A_k} S_k^2 d\mathbb{P} + 2\int_{\Omega} I_{A_k}S_k(S_n-S_k)d\mathbb{P}\,.\end{aligned}$$

Παρατηρούμε τώρα ότι η $S_k = X_1+\ldots+X_k$ είναι $\sigma(X_1,\ldots,X_k)$–μετρήσιμη, και η I_{A_k} επίσης. Άρα, η $I_{A_k}S_k$ είναι $\sigma(X_1,\ldots,X_k)$–μετρήσιμη (ως συνάρτηση των $X_1,\ldots X_k$), ενώ η $S_n-S_k = X_{k+1}+\ldots+X_n$ είναι $\sigma(X_{k+1},\ldots,X_n)$–μετρήσιμη (ως συνάρτηση των υπολοίπων X_j – για $k=n$ είναι $\{\varnothing,\Omega\}$–μετρήσιμη, αφού είναι σταθερή, $S_n-S_n\equiv 0$). Άρα, οι $I_{A_k}S_k$ και S_n-S_k είναι ανεξάρτητες τ.μ. [Διότι οι $X_1,X_2,\ldots,X_n$ είναι ανεξάρτητες τ.μ. και, συνεπώς, οι $\sigma(X_1,\ldots,X_k)$ και $\sigma(X_{k+1},\ldots,X_n)$ είναι ανεξάρτητες κλάσεις ενδεχομένων (ισχύει και για $k=n$).] Επομένως, αφού είναι και ολοκληρώσιμες, προκύπτει ότι

$$2\int_{\Omega} I_{A_k}S_k(S_n-S_k)d\mathbb{P} = 2\,\mathbb{E}(I_{A_k}S_k(S_n-S_k)) = 2\,\mathbb{E}(I_{A_k}S_k)\,\mathbb{E}(S_n-S_k) = 0,$$

διότι $\mathbb{E}(S_j) = 0$ για $j = 1, \ldots, n$. Συνδυάζοντας τα παραπάνω, έχουμε

$$\begin{aligned}
\mathrm{Var}(S_n) &\geq \sum_{k=1}^{n} \int_{A_k} S_k^2 d\mathbb{P} = \sum_{k=1}^{n} \int_{\Omega} I_{A_k} S_k^2 d\mathbb{P} = \sum_{k=1}^{n} \mathbb{E}(I_{A_k} S_k^2) \\
&\geq \sum_{k=1}^{n} \mathbb{E}(a^2 I_{A_k}) && (\text{διότι } I_{A_k} S_k^2 \geq a^2 I_{A_k}) \\
&= a^2 \sum_{k=1}^{n} \mathbb{P}(A_k) \\
&= a^2 \mathbb{P}(A), && (\text{αφού } A_i A_j = \varnothing, \quad \bigcup_{k=1}^{n} A_k = A)
\end{aligned}$$

δηλαδή

$$a^2 \mathbb{P}\left[\max_{1 \leq k \leq n} |S_k| \geq a\right] \leq \mathrm{Var}(S_n),$$

η οποία αποδεικνύει το ζητούμενο. $\square$

Είναι προφανές ότι

$$\{\omega \in \Omega : |S_n - \mathbb{E}(S_n)| \geq a\} \subset \left\{\omega \in \Omega : \max_{1 \leq k \leq n} |S_k - \mathbb{E}(S_k)| \geq a\right\}, \text{ για κάθε } a > 0,$$

και, κατά συνέπεια,

$$\mathbb{P}[|S_n - \mathbb{E}(S_n)| \geq a] \leq \mathbb{P}\left[\max_{1 \leq k \leq n} |S_k - \mathbb{E}(S_k)| \geq a\right].$$

Επομένως, η ανισότητα του Kolmogorov είναι πολύ ισχυρότερη από αυτήν του Chebychev, διότι και οι δύο ανισότητες δίνουν το ίδιο άνω φράγμα, $\mathrm{Var}(S_n)/a^2$, για τις δύο παραπάνω πιθανότητες.

Στα επόμενα θα χρειαστούμε επίσης το παρακάτω λήμμα.

Λήμμα 6.13 Έστω x_n μία ακολουθία πραγματικών αριθμών. Τότε

(i) Εάν $x_n \to x \in \mathbb{R}$ τότε και $(x_1 + \ldots + x_n)/n \to x$.

(ii) Εάν η σειρά $\sum_{k=1}^{\infty} x_k/k$ συγκλίνει σε κάποιον πραγματικό αριθμό, τότε[9]

$$\frac{x_1 + \ldots + x_n}{n} \to 0.$$

Απόδειξη: (i) Αφού $x_n \to x$, τότε για τυχόν $\varepsilon > 0$, $|x_n - x| \leq \varepsilon$, όταν $n \geq n_0 = n_0(\varepsilon)$. Άρα, για $n > n_0$,

$$\begin{aligned}
\left|\frac{x_1 + \ldots + x_n}{n} - x\right| &= \left|\frac{(x_1 - x) + \ldots + (x_n - x)}{n}\right| \\
&\leq \left|\frac{(x_1 - x) + \ldots + (x_{n_0} - x)}{n}\right| \\
&\quad + \frac{n - n_0}{n}\left|\frac{(x_{n_0+1} - x) + \ldots + (x_n - x)}{n - n_0}\right| \\
&\leq \frac{c}{n} + \left(1 - \frac{n_0}{n}\right)\frac{|x_{n_0+1} - x| + \ldots + |x_n - x|}{n - n_0} \\
&\leq \frac{c}{n} + \left(1 - \frac{n_0}{n}\right)\varepsilon \to \varepsilon, \quad \text{καθώς } n \to \infty.
\end{aligned}$$

[Εδώ $c = |(x_1 - x) + \ldots + (x_{n_0} - x)|$, ανεξάρτητο του n.] Άρα,

$$\limsup_{n\to\infty}\left|\frac{x_1 + \ldots + x_n}{n} - x\right| \leq \varepsilon,$$

και επειδή το $\varepsilon > 0$ είναι τυχόν, $(x_1 + \ldots + x_n)/n \to x$.

[9] Το αποτέλεσμα αυτό ισχύει γενικότερα, και είναι γνωστό ως **Λήμμα του Kronecker**: Αν $0 < b_n \nearrow \infty$, και η σειρά $\sum_{k=1}^{\infty} x_k/b_k$ συγκλίνει σε πραγματικό αριθμό, τότε $(x_1 + \ldots + x_n)/b_n \to 0$.

(ii) Θέτουμε $s = \sum_{k=1}^{\infty} x_k/k$ (από την υπόθεση, $s \in \mathbb{R}$), και $s_n = \sum_{k=1}^{n} x_k/k$, οπότε $s_n \to s$. Όμως, $s_n - s_{n-1} = x_n/n$ (όπου $s_0 = 0$), οπότε $x_n = n(s_n - s_{n-1})$, και έτσι,

$$\begin{aligned}
\frac{x_1 + \ldots + x_{n+1}}{n+1} &= \frac{1}{n+1}\sum_{k=1}^{n+1} k(s_k - s_{k-1}) \\
&= \frac{1}{n+1}(s_1 + 2(s_2 - s_1) + \ldots + n(s_n - s_{n-1}) + (n+1)(s_{n+1} - s_n)) \\
&= \frac{1}{n+1}(-s_1 - s_2 - \ldots - s_n + (n+1)s_{n+1}) \\
&= -\frac{n}{n+1}\frac{s_1 + \ldots + s_n}{n} + s_{n+1}.
\end{aligned}$$

Από την υπόθεση, $s_n \to s$ (και άρα, $s_{n+1} \to s$), και λόγω του (i), $(s_1 + \ldots + s_n)/n \to s$, που σημαίνει ότι

$$\frac{x_1 + \ldots + x_{n+1}}{n+1} \to -1 \cdot s + s = 0,$$

δηλαδή $(x_1 + \ldots + x_n)/n \to 0$. $\square$

Επίσης θα χρειαστούμε το εξής λήμμα.

Λήμμα 6.14 Για τυχούσα ακολουθία πραγματικών αριθμών, έστω x_n, θέτουμε $s_n = x_1 + \ldots + x_n$, $t_n = \sup_{k \geq n} |s_k - s_n|$, και $t = \inf_n t_n$. Τότε, αν $t = 0$, η σειρά $\sum_{k=1}^{\infty} x_k$ συγκλίνει σε πραγματικό αριθμό (δηλ. $s_n \to s \in \mathbb{R}$).

Απόδειξη: Έστω $\varepsilon > 0$. Αφού ισχύει ότι $\inf_n t_n = 0$, θα υπάρχει κάποιο n_0, τέτοιο ώστε $t_{n_0} < \varepsilon/2$, δηλαδή

$$\sup_{k \geq n_0} |s_k - s_{n_0}| < \varepsilon/2.$$

Έστω m, n με $m, n \geq n_0$. Τότε,

$$|s_m - s_n| = |s_m - s_{n_0} + s_{n_0} - s_n| \leq |s_m - s_{n_0}| + |s_n - s_{n_0}| < \varepsilon$$

(αφού $m \geq n_0, n \geq n_0$), και άρα, από το κριτήριο σύγκλισης του Cauchy, η ακολουθία s_n συγκλίνει στο $\mathbb{R}$. □

Έχουμε τώρα την εξής πρόταση.

Πρόταση 6.15 Εάν η $\{X_n,\ n \geq 1\}$ είναι τυχούσα ακολουθία ανεξαρτήτων τ.μ. στον χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, με $\mathrm{Var}(X_n) < \infty$, και $\mathbb{E}(X_n) = 0$ $(n = 1, 2, \ldots)$, τέτοια ώστε η σειρά των διασπορών συγκλίνει,

$$\sum_{n=1}^{\infty} \mathrm{Var}(X_n) < \infty,$$

τότε υπάρχει μη εκτεταμένη τ.μ. X ($|X(\omega)| < \infty$ για κάθε $\omega \in \Omega$), τέτοια ώστε

$$X = \sum_{n=1}^{\infty} X_n, \quad \text{με πιθανότητα 1.}$$

[Δηλαδή, $\sum\limits_{k=1}^{n} X_k \xrightarrow{\mathbf{a.s.}} X$.]

Απόδειξη: Θέτουμε $S_n = X_1 + \ldots + X_n$,

$$T_n = \sup_{k \geq n} |S_k - S_n|, \quad \text{και} \quad T = \inf_n T_n,$$

οπότε οι S_n, T_n, T είναι τ.μ. στον $(\Omega, \mathcal{A}, \mathbb{P})$, στον οποίο είναι ορισμένες οι X_n (δεν αποκλείουμε την περίπτωση οι T_n και T να είναι εκτεταμένες). Από το Λήμμα 6.14, αν $\omega \in \{T = 0\}$ (δηλ.

$T(\omega) = 0$), τότε η $S_n(\omega)$ συγκλίνει σε πραγματικό αριθμό, έστω $X(\omega)$. Θέτουμε λοιπόν

$$X(\omega) = \begin{cases} \lim_n S_n(\omega), & \text{αν } T(\omega) = 0, \\ 0, & \text{αν } T(\omega) > 0, \end{cases}$$

και παρατηρούμε ότι η X είναι (μη εκτεταμένη) τ.μ., και μάλιστα,

$$X(\omega) = \sum_{n=1}^{\infty} X_n(\omega), \quad \text{όταν } T(\omega) = 0.$$

Αν δείξουμε ότι $\mathbb{P}(T = 0) = 1$, τότε θα έχουμε ότι και

$$\mathbb{P}(S_n \to X) \geq \mathbb{P}(T = 0) = 1,$$

δηλαδή,

$$S_n \xrightarrow{\textbf{a.s.}} X.$$

Αρκεί λοιπόν να δειχθεί ότι $\mathbb{P}(T > 0) = 0$ (αφού η T είναι μη αρνητική και, πιθανώς, εκτεταμένη), και επομένως, αρκεί να δειχθεί ότι

$$\mathbb{P}\left(T > \frac{1}{m}\right) = 0, \quad \text{για κάθε } m = 1, 2, \dots,$$

αφού $\{T > 0\} = \bigcup_{m=1}^{\infty} \{T > \frac{1}{m}\}$. Από τον ορισμό της $T = \inf_n T_n$, έχουμε ότι $T \leq T_n$, και άρα,

$$\left\{T > \frac{1}{m}\right\} \subset \left\{T_n > \frac{1}{m}\right\}, \quad \text{για } n = 1, 2, \dots,$$

επειδή η σχέση $T(\omega) > \frac{1}{m}$ συνεπάγεται την $T_n(\omega) \geq T(\omega) > \frac{1}{m}$. Όμως,

$$\begin{aligned}
\left\{T_n > \frac{1}{m}\right\} &= \left\{\sup_{k\geq n}|S_k - S_n| > \frac{1}{m}\right\} = \bigcup_{k=n}^{\infty}\left\{|S_k - S_n| > \frac{1}{m}\right\} \\
&= \bigcup_{k=n+1}^{\infty}\left\{|S_k - S_n| > \frac{1}{m}\right\} \qquad (\text{διότι } \left\{|S_n - S_n| > \tfrac{1}{m}\right\} = \varnothing) \\
&= \lim_{N\to\infty}\bigcup_{k=n+1}^{N}\left\{|S_k - S_n| > \frac{1}{m}\right\} \\
&= \lim_{N\to\infty}\left\{\max_{n+1\leq k\leq N}|S_k - S_n| > \frac{1}{m}\right\} \\
&\subset \lim_{N\to\infty}\left\{\max_{n+1\leq k\leq N}|S_k - S_n| \geq \frac{1}{m}\right\} \\
&= \lim_{N\to\infty}\left\{\max_{1\leq j\leq N-n}|X_{n+1} + \ldots + X_{n+j}| \geq \frac{1}{m}\right\}.
\end{aligned}$$

Συνεπώς,

$$\begin{aligned}
\mathbb{P}\left(T_n > \frac{1}{m}\right) &\leq \mathbb{P}\left[\lim_{N\to\infty}\left\{\max_{1\leq j\leq N-n}|X_{n+1} + \ldots + X_{n+j}| \geq \frac{1}{m}\right\}\right] \\
&= \lim_{N\to\infty}\mathbb{P}\left[\max_{1\leq j\leq N-n}|X_{n+1} + \ldots + X_{n+j}| \geq \frac{1}{m}\right] \\
&\leq \limsup_{N\to\infty} m^2\,\mathrm{Var}\,(X_{n+1} + \ldots + X_N) \qquad (\text{από την Πρόταση 6.12}) \\
&= m^2 \limsup_{N\to\infty}\sum_{j=n+1}^{N}\mathrm{Var}\,(X_j) \\
&= m^2\sum_{j=n+1}^{\infty}\mathrm{Var}\,(X_j).
\end{aligned}$$

Επομένως, ισχύει η ανισότητα

$$\mathbb{P}\left(T_n > \frac{1}{m}\right) \leq m^2\sum_{j=n+1}^{\infty}\mathrm{Var}\,(X_j),$$

και επειδή $\sum_{j=1}^{\infty} \mathrm{Var}(X_j) < \infty$ (από υπόθεση), έπεται ότι

$$\limsup_{n\to\infty} \mathbb{P}\left(T_n > \frac{1}{m}\right) \le m^2 \limsup_{n\to\infty} \sum_{j=n+1}^{\infty} \mathrm{Var}(X_j) = 0,$$

δηλαδή $\lim_{n\to\infty} \mathbb{P}\left(T_n > \frac{1}{m}\right) = 0$ (για κάθε $m = 1, 2, \ldots$). Από την ανισότητα

$$\mathbb{P}\left(T > \frac{1}{m}\right) \le \mathbb{P}\left(T_n > \frac{1}{m}\right), \quad n = 1, 2, \ldots,$$

η οποία ισχύει επειδή $T \le T_n$, προκύπτει αμέσως ότι (παίρνοντας όρια για $n \to \infty$)

$$\mathbb{P}\left(T > \frac{1}{m}\right) \le \lim_{n\to\infty} \mathbb{P}\left(T_n > \frac{1}{m}\right) = 0,$$

δηλαδή, $\mathbb{P}\left(T > \frac{1}{m}\right) = 0$. Αυτό αποδεικνύει την πρόταση. $\square$

Πόρισμα 6.16 Έστω $\{X_n, n \ge 1\}$ μία ανεξάρτητη ακολουθία τ.μ., με $\mathrm{Var}(X_n) < \infty$ και $\mathbb{E}(X_n) = \mu_n$ για κάθε n. [Φυσικά, $\mu_n \in \mathbb{R}$ για κάθε n.] Εάν $\sum_{n=1}^{\infty} \mathrm{Var}(X_n) < \infty$, τότε η σειρά

$$\sum_{n=1}^{\infty} (X_n - \mu_n)$$

συγκλίνει, σχεδόν βεβαίως, προς κάποια (μη εκτεταμένη) τ.μ. X.

Απόδειξη: Προφανής από την Πρόταση 6.15, διότι οι τ.μ. $\widetilde{X}_n = X_n - \mu_n$ ικανοποιούν τις υποθέσεις της, αφού $\mathrm{Var}(\widetilde{X}_n) =$

$\mathrm{Var}(X_n)$, και $\mathbb{E}(\widetilde{X}_n) = 0$ για κάθε n. $\square$

Είμαστε τώρα σε θέση να διατυπώσουμε και να αποδείξουμε τον πρώτο ισχυρό νόμο των μεγάλων αριθμών.

Θεώρημα 6.17 (ισχυρός νόμος μεγάλων αριθμών (του Kolmogorov) για ανεξάρτητη ακολουθία) Έστω $\{X_n,\ n \geq 1\}$ μία ανεξάρτητη ακολουθία τ.μ., για την οποία ισχύει ότι $\mathbb{E}(X_n) = 0$ και $\mathrm{Var}(X_n) < \infty$ για κάθε n. Αν

$$\sum_{n=1}^{\infty} \frac{\mathrm{Var}(X_n)}{n^2} < \infty$$

τότε

$$\overline{X}_n \stackrel{\text{ορ.}}{=\!=} \frac{X_1 + \ldots + X_n}{n} \xrightarrow{\text{a.s.}} 0.$$

Απόδειξη: Θέτουμε $Y_n = X_n/n$, οπότε η Y_n είναι ανεξάρτητη ακολουθία τ.μ. με $\mathrm{Var}(Y_n) = \mathrm{Var}(X_n)/n^2$, και άρα, $\sum\limits_{n=1}^{\infty} \mathrm{Var}(Y_n) < \infty$. Επίσης, $\mathbb{E}(Y_n) = 0$ για κάθε n. Από την Πρόταση 6.15, υπάρχει (μη εκτεταμένη) τ.μ. X, τέτοια ώστε

$$\sum_{k=1}^{n} \frac{X_k(\omega)}{k} \to X(\omega), \quad \text{καθώς } n \to \infty,$$

εκτός, ίσως, από κάποια $\omega \in \Omega$, τα οποία ανήκουν σε ένα ενδεχόμενο A, με $\mathbb{P}(A) = 0$. [Φυσικά, η $X(\omega)$ είναι πεπερασμένη για $\omega \in A^c$.] Από το Λήμμα 6.13(ii), αν $\omega \in A^c$ τότε

$$\frac{X_1(\omega) + \ldots + X_n(\omega)}{n} \to 0, \quad \text{καθώς } n \to \infty,$$

που σημαίνει ότι
$$\left\{\omega : \frac{S_n(\omega)}{n} \to 0\right\} \supset A^c.$$
[Εδώ, ως συνήθως, $S_n(\omega) = X_1(\omega) + \ldots + X_n(\omega)$.] Αφού $\mathbb{P}(A) = 0$, έχουμε, προφανώς, ότι
$$\mathbb{P}(A^c) = 1,$$
και συνεπώς,
$$\mathbb{P}\left(\frac{S_n}{n} \to 0\right) = 1,$$
δηλαδή
$$\overline{X}_n = \frac{S_n}{n} \xrightarrow{\textbf{a.s.}} 0. \quad \square$$

Πόρισμα 6.18 (ισχυρός νόμος μεγάλων αριθμών (του Kolmogorov) για ανεξάρτητη ακολουθία) Έστω $\{X_n,\ n \geq 1\}$ μία ανεξάρτητη ακολουθία τ.μ., με $\mathrm{Var}(X_n) < \infty$ και $\mathbb{E}(X_n) = \mu_n$ για κάθε n. Εάν $\sum_{n=1}^{\infty} \mathrm{Var}(X_n)/n^2 < \infty$, τότε
$$\frac{X_1 + \ldots + X_n}{n} - \frac{\mu_1 + \ldots + \mu_n}{n} \xrightarrow{\textbf{a.s.}} 0,$$
δηλαδή,
$$\overline{X}_n - \overline{\mu}_n \xrightarrow{\textbf{a.s.}} 0,$$
όπου
$$\overline{\mu}_n = \frac{\mu_1 + \ldots + \mu_n}{n}, \quad n = 1, 2, \ldots\ .$$
Αν, επιπλέον, ισχύει ότι[10]
$$\frac{\mu_1 + \ldots + \mu_n}{n} \to \mu \in \mathbb{R}, \quad \text{καθώς } n \to \infty,$$

[10] δηλαδή, αν ισχύει ότι $\overline{\mu}_n \to \mu \in \mathbb{R}$. Αυτό συμβαίνει όταν, για παράδειγμα, $\mu_n = \mu$ για κάθε n, ή, γενικότερα, όταν $\mathbb{E}(X_n) = \mu_n \to \mu \in \mathbb{R}$, καθώς $n \to \infty$, βλ. Λήμμα 6.13(i).

τότε
$$\overline{X}_n = \frac{X_1 + \ldots + X_n}{n} \xrightarrow{\mathbf{a.s.}} \mu.$$
Απόδειξη: Οι τ.μ. $\widetilde{X_n} = X_n - \mathbb{E}(X_n) = X_n - \mu_n$ ικανοποιούν τις υποθέσεις του Θεωρήματος 6.17, άρα ισχύει ότι
$$\frac{X_1 - \mu_1 + \ldots + X_n - \mu_n}{n} \xrightarrow{\mathbf{a.s.}} 0,$$
δηλαδή,
$$\overline{X}_n - \overline{\mu}_n = \frac{X_1 + \ldots + X_n}{n} - \frac{\mu_1 + \ldots + \mu_n}{n} \xrightarrow{\mathbf{a.s.}} 0.$$
Αν, επιπροσθέτως, $(\mu_1 + \ldots + \mu_n)/n \to \mu \in \mathbb{R}$, καθώς $n \to \infty$, τότε έχουμε ότι $\overline{\mu}_n \to \mu$ (άρα και $\overline{\mu}_n \xrightarrow{\mathbf{a.s.}} \mu$), $\overline{X}_n - \overline{\mu}_n \xrightarrow{\mathbf{a.s.}} 0$, και τελικά,
$$\overline{X}_n = \left(\overline{X}_n - \mu_n\right) + \mu_n \xrightarrow{\mathbf{a.s.}} 0 + \mu = \mu.$$
[Διότι, αν $X_n \xrightarrow{\mathbf{a.s.}} X$ και $Y_n \xrightarrow{\mathbf{a.s.}} c$ (όπου $c \in \mathbb{R}$, σταθερά), τότε $X_n + Y_n \xrightarrow{\mathbf{a.s.}} X + c$ (πρβλ. Άσκηση 6.3(γ)), και το εφαρμόσαμε αυτό για $X_n = \overline{X}_n - \overline{\mu}_n \xrightarrow{\mathbf{a.s.}} 0$, και $Y_n(\omega) = \overline{\mu}_n$ (για κάθε $\omega \in \Omega) \to \mu$.] $\square$

Αξίζει να σημειωθεί ότι ο ισχυρός νόμος του Θεωρήματος 6.17 (ή του Πορίσματος 6.18) **δεν καλύπτει πλήρως** τον ασθενή νόμο του Θεωρήματος 6.7. Πράγματι, ο ισχυρός νόμος εξάγει μεν ισχυρότερο συμπέρασμα,[11] αλλά το Πόρισμα 6.18 χρησιμοποιεί την ελαφρώς **ισχυρότερη υπόθεση**, $\sum_{n=1}^{\infty} \mathrm{Var}(X_n)/n^2 < \infty$, αντί

[11] ότι, δηλ., ισχύει η ισχυρή σύγκλιση $\overline{X}_n - \overline{\mu}_n \xrightarrow{\mathbf{a.s.}} 0$, αντί της $\overline{X}_n - \overline{\mu}_n \xrightarrow{\mathbf{P}} 0$, δηλ., αντί της ασθενούς σύγκλισης $\overline{X}_n - \overline{\mu}_n \xrightarrow{\mathbf{d}} 0$

της αντίστοιχης του Θεωρήματος 6.7, $(1/n^2)\sum_{k=1}^{n} \mathrm{Var}(X_k) \to 0$. Το γεγονός ότι η σχέση $\sum_{n=1}^{\infty} \mathrm{Var}(X_n)/n^2 < \infty$ συνεπάγεται την $(1/n^2)\sum_{k=1}^{n} \mathrm{Var}(X_k) \to 0$ προκύπτει από το Λήμμα 6.13(ii), αφού η σύγκλιση της σειράς

$$\sum_{n=1}^{\infty} \frac{\frac{1}{n}\mathrm{Var}(X_n)}{n},$$

μας εξασφαλίζει ότι

$$\frac{1}{n}\left(\mathrm{Var}(X_1) + \frac{1}{2}\mathrm{Var}(X_2) + \ldots + \frac{1}{n}\mathrm{Var}(X_n)\right) \to 0,$$

και συνεπώς,

$$\begin{aligned}\frac{1}{n^2}\left(\mathrm{Var}(X_1) + \ldots + \mathrm{Var}(X_n)\right) &= \frac{1}{n}\left(\frac{1}{n}\mathrm{Var}(X_1) + \ldots + \frac{1}{n}\mathrm{Var}(X_n)\right)\\ &\le \frac{1}{n}\left(\mathrm{Var}(X_1) + \frac{1}{2}\mathrm{Var}(X_2) + \ldots + \frac{1}{n}\mathrm{Var}(X_n)\right) \to 0.\end{aligned}$$

Φυσικά, για να αποδειχθεί το Πόρισμα 6.18 χρησιμοποιήθηκε η πολύ ισχυρή ανισότητα του Kolmogorov, καθώς και κάποιες, σχετικά πολύπλοκες, τεχνικές, ενώ η απόδειξη του Θεωρήματος 6.17 είναι σχεδόν προφανής.

Συγκρίνοντας τους δύο νόμους από την άποψη των εφαρμογών, παρατηρούμε τα εξής: Όταν οι $\{X_n,\ n \ge 1\}$ είναι ανεξάρτητες και ισόνομες με πεπερασμένη διασπορά σ^2 και μέσο μ, το Θεώρημα 6.7 μας εξασφαλίζει ότι

$$\overline{X}_n \xrightarrow{\mathbf{p}} \mu,$$

ενώ το Πόρισμα 6.18 (αφού $\sum_{n=1}^{\infty} \sigma^2/n^2 = \sigma^2\pi^2/6 < \infty$) συνεπάγεται την

$$\overline{X}_n \xrightarrow{\textbf{a.s.}} \mu,$$

που σημαίνει ότι ισχύει το εξής (βλ. Λήμμα 6.3):

$$\text{Για κάθε } \varepsilon > 0, \ \lim_{n\to\infty} \mathbb{P}\left(\sup_{k\geq n} \left|\overline{X}_k - \mu\right| \geq \varepsilon\right) = 0.$$

Άρα, η πιθανότητα να υπάρξει **έστω και ένας** δειγματικός μέσος $\overline{X}_k$, $k \geq n$, που να διαφέρει από τον θεωρητικό μέσο μ περισσότερο από ε, γίνεται αμελητέα από κάποιο $n = n(\varepsilon)$ και πάνω. Δηλαδή, για αρκούντως μεγάλο n, **όλοι οι δειγματικοί μέσοι** $\overline{X}_k$, $k \geq n$, θα βρίσκονται στο διάστημα $(\mu - \varepsilon, \mu + \varepsilon)$, με πολύ μεγάλη πιθανότητα. Ο ασθενής νόμος μας διαβεβαιώνει ότι αυτό συμβαίνει για τον $\overline{X}_n$, για αρκούντως μεγάλο n, αλλά δεν μας εξασφαλίζει ότι αυτή η κατάσταση θα συνεχίζεται, εσαεί, και για όλους τους επόμενους (άπειρους το πλήθος) δειγματικούς μέσους.

6.4 Ισχυρός Νόμος για Ανεξάρτητη και Ισόνομη Ακολουθία

Στην περίπτωση κατά την οποία οι X_n είναι και ισόνομες (εκτός από ανεξάρτητες), και $\mathbb{E}|X_n| < \infty$, $\mathbb{E}(X_n) = \mu$, τότε ο ισχυρός νόμος ισχύει χωρίς την υπόθεση πεπερασμένης διασποράς για τις X_n. Για να αποδείξουμε αυτό το πολύ σπουδαίο αποτέλεσμα του Kolmogorov, θα χρησιμοποιήσουμε μερικά ακόμη βοηθητικά αποτελέσματα.

Λήμμα 6.19 Για μία τ.μ. X με $\{|X| = +\infty\} = \varnothing$,

$$\sum_{n=1}^{\infty} \mathbb{P}(|X| \geq n) \leq \mathbb{E}|X| \leq 1 + \sum_{n=1}^{\infty} \mathbb{P}(|X| \geq n).$$

Απόδειξη: Αν και προκύπτει άμεσα από την έκφραση (βλ. Άσκηση ;;.8)

$$\mathbb{E}|X| = \int_0^{\infty} \mathbb{P}(|X| \geq x)dx = \int_0^{\infty} \mathbb{P}(|X| > x)dx,$$

θα δώσουμε και μία άλλη απόδειξη. Προφανώς,

$$\sum_{n=0}^{\infty} nI_{A_n} \leq |X| \leq \sum_{n=0}^{\infty}(n+1)I_{A_n},$$

όπου $A_n = \{\omega : n \leq |X(\omega)| < (n+1)\}$. Άρα, ισχύει η ανισότητα

$$\mathbb{E}\left(\sum_{n=0}^{\infty} nI_{A_n}\right) \leq \mathbb{E}|X| \leq \mathbb{E}\left(\sum_{n=0}^{\infty}(n+1)I_{A_n}\right),$$

και από τα Θεωρήματα Beppo Levi και Tonelli,

$$\begin{aligned}
\mathbb{E}\left(\sum_{n=0}^{\infty} nI_{A_n}\right) \overset{\text{B. Levi}}{=\!=\!=} \sum_{n=1}^{\infty} \mathbb{E}(nI_{A_n}) \quad &= \quad \sum_{n=1}^{\infty} n\,\mathbb{P}(n \leq |X| < n+1) \\
= \quad \sum_{n=1}^{\infty} \mathbb{P}(n \leq |X| < n+1) \left(\sum_{k=1}^{n} 1\right) & \\
\overset{\text{Tonelli}}{=\!=\!=} \sum_{k=1}^{\infty}\sum_{n=k}^{\infty} \mathbb{P}(n \leq |X| < n+1) \quad &= \quad \sum_{k=1}^{\infty} \mathbb{P}(|X| \geq k).
\end{aligned}$$

Ομοίως, αφού $\sum_{n=0}^{\infty} I_{A_n}(\omega) \equiv 1$ για κάθε $\omega \in \Omega$, το άνω φράγμα για την $\mathbb{E}|X|$ ισούται με

$$\mathbb{E}\left(\sum_{n=0}^{\infty} nI_{A_n} + \sum_{n=0}^{\infty} I_{A_n}\right) = \mathbb{E}\left(1 + \sum_{n=0}^{\infty} nI_{A_n}\right) = 1 + \mathbb{E}\left(\sum_{n=0}^{\infty} nI_{A_n}\right)$$
$$= 1 + \sum_{k=1}^{\infty} \mathbb{P}(|X| \geq k). \quad \square$$

Λήμμα 6.20 Για κάθε $k = 1, 2, \ldots$, ισχύει η ανισότητα $\displaystyle\sum_{n=k}^{\infty} \frac{1}{n^2} \leq \frac{2}{k}$.

Απόδειξη: Αφού $\dfrac{1}{n^2} \leq \dfrac{1}{(x-1)^2}$ για $n \leq x \leq n+1$, έχουμε (για $n = 2, 3, \ldots$)

$$\int_n^{n+1} \frac{1}{n^2} dx \leq \int_n^{n+1} \frac{1}{(x-1)^2} dx = \int_{n-1}^{n} \frac{1}{x^2} dx,$$

και συνεπώς,

$$\sum_{n=k}^{\infty} \frac{1}{n^2} = \frac{1}{k^2} + \sum_{n=k+1}^{\infty} \frac{1}{n^2} \leq \frac{1}{k} + \sum_{n=k+1}^{\infty} \int_{[n-1,n)} \frac{1}{x^2} dx = \frac{1}{k} + \int_{[k,+\infty)} \frac{1}{x^2} dx = \frac{2}{k}. \quad \square$$

Είμαστε τώρα σε θέση να διατυπώσουμε και να αποδείξουμε το εξής σημαντικότατο θεώρημα.

Θεώρημα 6.21 (ισχυρός νόμος μεγάλων αριθμών για ανεξάρτητη και ι-σόνομη ακολουθία, Kolmogorov) Εάν η ακολουθία (εκτεταμένων) τ.μ. $\{X_n,\ n \geq 1\}$ είναι ανεξάρτητη και ισόνομη, με $\mathbb{E}|X_n| <$

∞ και $\mathbb{E}(X_n) = \mu$, τότε

$$\overline{X}_n = \frac{X_1 + \ldots + X_n}{n} \xrightarrow{\textbf{a.s.}} \mu.$$

Απόδειξη: Χωρίς βλάβη της γενικότητας μπορούμε να υποθέσουμε ότι όλες οι τ.μ. X_n δεν είναι εκτεταμένες, δηλ. $\{|X_n| = +\infty\} = \varnothing$ για κάθε n. Πράγματι, θέτοντας $A_n = \{\omega : |X_n(\omega)| = +\infty\}$, έχουμε ότι $\mathbb{P}(A_n) = 0$, επειδή $\mathbb{E}|X_n| < \infty$. Έτσι, $\mathbb{P}(A) = 0$, όπου $A = \bigcup_{n=1}^{\infty} A_n$, οπότε θέτοντας $\widetilde{X}_n = X_n I_{A^c}$, έχουμε ότι η ακολουθία $\{\widetilde{X}_n,\ n \geq 1\}$ είναι ανεξάρτητη[12] και ισόνομη, και $\widetilde{X}_n = X_n$ για κάθε n, με πιθ. 1 (δηλ. $\widetilde{X}_n \overset{\textbf{a.s.}}{=} X_n$). Επομένως, $\mathbb{E}|\widetilde{X}_n| = \mathbb{E}|X_n| < \infty$, και $\mathbb{E}(\widetilde{X}_n) = \mathbb{E}(X_n) = \mu$. Αν λοιπόν δεχθούμε ότι το θεώρημα ισχύει για μη εκτεταμένες τ.μ., τότε θα έχουμε ότι

$$\frac{\widetilde{X}_1 + \ldots + \widetilde{X}_n}{n} \xrightarrow{\textbf{a.s.}} \mu,$$

δηλ. θα υπάρχει ένα ενδεχόμενο B, με $\mathbb{P}(B) = 1$, τέτοιο ώστε

$$\frac{\widetilde{X}_1(\omega) + \ldots + \widetilde{X}_n(\omega)}{n} \to \mu, \quad \text{για κάθε } \omega \in B.$$

Τότε όμως, για κάθε $\omega \in BA^c = B \setminus A$, θα έχουμε ότι

$$\overline{X}_n(\omega) = \frac{X_1(\omega) + \ldots + X_n(\omega)}{n} = \frac{\widetilde{X}_1(\omega) + \ldots + \widetilde{X}_n(\omega)}{n} \to \mu,$$

και προφανώς, $\mathbb{P}(BA^c) = 1$. Έτσι, θα προκύψει ότι $\mathbb{P}(\overline{X}_n \to \mu) \geq \mathbb{P}(BA^c) = 1$, δηλαδή ότι και $\overline{X}_n \xrightarrow{\textbf{a.s.}} \mu$. Στην συνέχεια

[12] Γιατί είναι ανεξάρτητη;

λοιπόν, και χωρίς βλάβη της γενικότητας, υποθέτουμε ότι όλες οι X_n δεν είναι εκτεταμένες.

Θα αποδείξουμε πρώτα το θεώρημα στην ειδική περίπτωση που $\mu = 0$, δηλ. υποθέτουμε ότι $\mathbb{E}(X_n) = \mu = 0$ για κάθε n (το γενικό συμπέρασμα για οποιοδήποτε μ προκύπτει εύκολα από την περίπτωση $\mu = 0$ – βλ. παρακάτω). Θεωρούμε την ακολουθία τ.μ.

$$Y_n = X_n I(|X_n| \le n) = \begin{cases} X_n, & \text{αν } -n \le X_n \le n, \\ 0, & \text{αν } |X_n| > n. \end{cases}$$

Η ακολουθία Y_n είναι προφανώς ανεξάρτητη (αφού η Y_n είναι συνάρτηση της X_n, και άρα, $\sigma(X_n)$–μετρήσιμη), και, επειδή όλες οι τ.μ. Y_n είναι φραγμένες, έπεται ότι έχουν και πεπερασμένη διασπορά. Η απόδειξη χωρίζεται σε 3 μέρη:

Μέρος 1. Αρχικά, στόχος μας είναι να δείξουμε ότι η σειρά $\sum_{n=1}^{\infty} \mathrm{Var}(Y_n)/n^2$ συγκλίνει, έτσι ώστε να μπορέσουμε να εφαρμόσουμε το Πόρισμα 6.18 στην ανεξάρτητη ακολουθία τ.μ. $\{Y_n,\ n \ge 1\}$. Έστω λοιπόν F η (κοινή) σ.κ. των τ.μ. X_n. Προφανώς,

$$\mathbb{E}(Y_n) = \mathbb{E}[X_n I(|X_n| \le n)] = \int_{\mathbb{R}} x I(|x| \le n) dF(x) = \mathbb{E}[X_1 I(|X_1| \le n)],$$
$$\mathbb{E}(Y_n^2) = \mathbb{E}[X_n^2 I(|X_n| \le n)] = \int_{\mathbb{R}} x^2 I(|x| \le n) dF(x) = \mathbb{E}[X_1^2 I(|X_1| \le n)].$$

[Φυσικά, $\mathbb{E}(Y_n^2) \le n^2 < \infty$.] Επομένως,

$$\begin{aligned}\operatorname{Var}(Y_n) &\leq \mathbb{E}(Y_n^2) = \mathbb{E}[X_n^2 I(|X_n| \leq n)] = \mathbb{E}[X_1^2 I(|X_1| \leq n)] \\ &= \sum_{k=1}^{n} \mathbb{E}[X_1^2 I(k-1 < |X_1| \leq k)] \\ &\leq \sum_{k=1}^{n} k\,\mathbb{E}[|X_1| I(k-1 < |X_1| \leq k)],\end{aligned}$$

όπου στις δύο τελευταίες σχέσεις χρησιμοποιήθηκε η ταυτότητα

$$X_1^2(\omega) I(|X_1(\omega)| \leq n) = \sum_{k=1}^{n} X_1^2(\omega) I(k-1 < |X_1(\omega)| \leq k), \quad \omega \in \Omega,$$

η οποία ισχύει και όταν $X_1(\omega) = 0$, και η προφανής ανισότητα

$$X_1^2(\omega) I(k-1 < |X_1(\omega)| \leq k) \leq k |X_1(\omega)| I(k-1 < |X_1(\omega)| \leq k).$$

Συνεπώς, χρησιμοποιώντας την προηγούμενη ανισότητα για την $\operatorname{Var}(Y_n)$, έχουμε

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{1}{n^2} \operatorname{Var}(Y_n) \quad &\leq \quad \sum_{n=1}^{\infty} \frac{1}{n^2} \sum_{k=1}^{n} k\,\mathbb{E}[|X_1| I(k-1 < |X_1| \leq k)] \\ &\overset{\text{Tonelli}}{=\!=} \sum_{k=1}^{\infty} k\,\mathbb{E}[|X_1| I(k-1 < |X_1| \leq k)] \left(\sum_{n=k}^{\infty} \frac{1}{n^2} \right) \\ &\overset{\text{Λήμμα 6.20}}{\leq} 2 \sum_{k=1}^{\infty} \mathbb{E}[|X_1| I(k-1 < |X_1| \leq k)] \\ &\overset{\text{Beppo Levi}}{=\!=} 2\,\mathbb{E}\left[|X_1| \sum_{k=1}^{\infty} I(k-1 < |X_1| \leq k) \right] \\ &= \quad 2\,\mathbb{E}|X_1| < \infty.\end{aligned}$$

[Η τελευταία ισότητα ισχύει επειδή $|X_1| = |X_1| \sum_{k=1}^{\infty} I(k-1 < |X_1| \le k)$, αφού η X_1 δεν είναι εκτεταμένη.] Άρα, $\sum_{n=1}^{\infty} \mathrm{Var}(Y_n)/n^2 < \infty$, που σημαίνει ότι (Πόρισμα 6.18)

$$\frac{Y_1 + \ldots + Y_n}{n} - \frac{\mathbb{E}(Y_1) + \ldots + \mathbb{E}(Y_n)}{n} \xrightarrow{\text{a.s.}} 0.$$

Μέρος 2. Τώρα είναι εύκολο να δείξουμε ότι ισχύει και η σχέση $(Y_1 + \ldots + Y_n)/n = \overline{Y}_n \xrightarrow{\text{a.s.}} 0$. Αυτή θα προκύψει άμεσα, και πάλι από το Πόρισμα 6.18, αν αποδείξουμε ότι $\mathbb{E}(Y_n) \to 0$. Η τελευταία σχέση όμως έπεται άμεσα από το Θεώρημα Κυριαρχημένης Σύγκλισης. Πράγματι, έχουμε $\mathbb{E}(Y_n) = \mathbb{E}[X_n I(|X_n| \le n)] = \mathbb{E}[X_1 I(|X_1| \le n)]$, δηλ. $\mathbb{E}(Y_n) = \mathbb{E}(Z_n)$, όπου $Z_n = X_1 I(|X_1| \le n)$. Προφανώς, ισχύει ότι $Z_n = X_1 I(|X_1| \le n) \to X_1$ (για κάθε ω), και $|Z_n| = |X_1 I(|X_1| \le n)| \le |X_1|$, με $\mathbb{E}|X_1| < \infty$. Άρα, $\mathbb{E}(Y_n) = \mathbb{E}(Z_n) \to \mathbb{E}(X_1) = 0$ (Θεώρημα Κυριαρχημένης Σύγκλισης – βλ. Ιδιότητα (M7) της §5.6 του Κεφ. ;;), που σημαίνει ότι (βλ. και Πόρισμα 6.18)

$$\overline{Y}_n = \frac{Y_1 + \ldots + Y_n}{n} \xrightarrow{\text{a.s.}} 0.$$

Μέρος 3. Στο τρίτο και τελευταίο στάδιο της απόδειξης θα δείξουμε ότι τα ενδεχόμενα $\{\overline{Y}_n \xrightarrow{\text{a.s.}} 0\}$ και $\{\overline{X}_n \xrightarrow{\text{a.s.}} 0\}$ είναι ισοδύναμα, δηλ. διαφέρουν μόνο κατά ένα ενδεχόμενο που έχει πιθανότητα 0. Έστω λοιπόν $A = \left\{\omega : \overline{Y}_n(\omega) \to 0\right\}$, οπότε $\mathbb{P}(A) = 1$ από τα προηγούμενα, και $B = \{\omega : Y_n(\omega) \neq$

$X_n(\omega)$ για άπειρα $n\} = \limsup\{X_n \neq Y_n\}$. Θα δείξουμε ότι για κάθε $\omega \in A \setminus B$,

$$\overline{X}_n(\omega) = \frac{X_1(\omega) + \ldots + X_n(\omega)}{n} \to 0, \quad \text{καθώς } n \to \infty.$$

Πράγματι, $\omega \in A$ σημαίνει ότι $\overline{Y}_n(\omega) \to 0$, ενώ $\omega \notin B$ σημαίνει ότι $X_n(\omega) = Y_n(\omega)$ από ένα n_0 και πάνω. Άρα, αν $\omega \in A \setminus B$, τότε υπάρχει ένας δείκτης $n_0 = n_0(\omega)$, τέτοιος ώστε για κάθε $n > n_0$, $|X_n(\omega) - Y_n(\omega)| = 0$. Συνεπώς,

$$\begin{aligned}
|\overline{X}_n(\omega)| &\leq |\overline{Y}_n(\omega)| + |\overline{X}_n(\omega) - \overline{Y}_n(\omega)| \\
&= |\overline{Y}_n(\omega)| + \left| \frac{X_1(\omega) + \ldots + X_n(\omega)}{n} - \frac{Y_1(\omega) + \ldots + Y_n(\omega)}{n} \right| \\
&= |\overline{Y}_n(\omega)| + \frac{1}{n} \left| \sum_{k=1}^{n} (X_k(\omega) - Y_k(\omega)) \right| \\
&\leq |\overline{Y}_n(\omega)| + \frac{1}{n} \sum_{k=1}^{n} |X_k(\omega) - Y_k(\omega)| \qquad (\text{και, επειδή } \omega \notin B) \\
&= |\overline{Y}_n(\omega)| + \frac{1}{n} \sum_{k=1}^{n_0} |X_k(\omega) - Y_k(\omega)| \to 0, \quad \text{καθώς } n \to \infty,
\end{aligned}$$

διότι $\overline{Y}_n(\omega) \to 0$ (επειδή $\omega \in A$), ενώ και ο δεύτερος προσθετέος τείνει στο 0, επειδή ο $|X_1(\omega) - Y_1(\omega)| + \ldots + |X_{n_0}(\omega) - Y_{n_0}(\omega)|$ είναι ένας σταθερός πραγματικός αριθμός, ανεξάρτητος του n. Αυτό σημαίνει ότι

$$\left\{ \omega : \frac{X_1(\omega) + \ldots + X_n(\omega)}{n} \to 0 \right\} \supset A \setminus B.$$

Αφού λοιπόν έχουμε δείξει ότι $\mathbb{P}(A) = 1$, αρκεί να δείξουμε ακόμα ότι και $\mathbb{P}(B) = 0$, διότι θα προκύψει τότε ότι $\mathbb{P}(A \smallsetminus B) = 1$, και από τα προηγούμενα,

$$\mathbb{P}\left(\frac{X_1(\omega) + \ldots + X_n(\omega)}{n} \to 0\right) \geq \mathbb{P}(A \smallsetminus B) = 1,$$

δηλαδή,

$$\frac{X_1 + \ldots + X_n}{n} \xrightarrow{\textbf{a.s.}} 0.$$

Είναι όμως $B = \limsup\{X_n \neq Y_n\} = \limsup B_n$, όπου $B_n = \{\omega : X_n(\omega) \neq Y_n(\omega)\}$, και προφανώς,

$$\begin{aligned}
\sum_{n=1}^{\infty} \mathbb{P}(B_n) &= \sum_{n=1}^{\infty} \mathbb{P}(X_n \neq Y_n) \\
&= \sum_{n=1}^{\infty} \mathbb{P}(|X_n| > n) \qquad (\text{διότι } Y_n = X_n \text{ αν και μόνο αν } |X_n| \leq n) \\
&= \sum_{n=1}^{\infty} \mathbb{P}(|X_1| > n) \qquad (\text{λόγω ισονομίας}) \\
&\leq \sum_{n=1}^{\infty} \mathbb{P}(|X_1| \geq n) \overset{\text{Λήμμα 6.19}}{\leq} \mathbb{E}|X_1| < \infty,
\end{aligned}$$

δηλ. $\sum_{n=1}^{\infty} \mathbb{P}(B_n) < \infty$. Αφού η σειρά $\sum_{n=1}^{\infty} \mathbb{P}(B_n)$ συγκλίνει, έπεται, από το πρώτο Λήμμα Borel-Cantelli, ότι $\mathbb{P}(\limsup B_n) = 0$, δηλαδή $\mathbb{P}(B) = 0$, που αποδεικνύει το θεώρημα στην περίπτωση που $\mu = 0$.

Στην γενική περίπτωση που $\mathbb{E}(X_n) = \mu$, θεωρούμε τις τ.μ. $\widetilde{X}_n = X_n - \mu$, οπότε οι $\widetilde{X}_n$ είναι ανεξάρτητες και ισόνομες, με $\mathbb{E}(\widetilde{X}_n) = 0$. Άρα, από αυτά που έχουμε μέχρι στιγμής αποδείξει,

$$\frac{\widetilde{X}_1 + \ldots + \widetilde{X}_n}{n} \xrightarrow{\text{a.s.}} 0, \quad \text{δηλαδή} \quad \frac{X_1 + \ldots + X_n}{n} - \mu \xrightarrow{\text{a.s.}} 0,$$

που σημαίνει ότι (Άσκηση 6.3(γ))

$$\frac{X_1 + \ldots + X_n}{n} \xrightarrow{\text{a.s.}} \mu. \quad \square$$

6.5 Αντίστροφος και Εκτεταμένος Ισχυρός Νόμος

Είναι ενδιαφέρον το γεγονός ότι για μία ανεξάρτητη και ισόνομη ακολουθία, οι δειγματικοί μέσοι συγκλίνουν, με πιθ. 1 (ισχυρά), προς κάποια πραγματική σταθερά **όταν και μόνο όταν υπάρχει η μέση τιμή**, και μάλιστα, η σταθερά αυτή είναι ακριβώς η **μέση τιμή**. Επομένως, η ύπαρξη της μέσης τιμής είναι **ακριβώς** η συνθήκη που χρειαζόμαστε για την σχεδόν βέβαιη σύγκλιση των δειγματικών μέσων. Ουσιαστικά, αυτό μας παρέχει μία ακόμη βαθύτερη αιτιολόγηση του γεγονότος ότι αποδεχθήκαμε τον ορισμό της μέσης τιμής ως ολοκλήρωμα. Σχετικά, έχουμε το επόμενο αποτέλεσμα.

Θεώρημα 6.22 (αντίστροφο του ισχυρού νόμου) Έστω $\{X_n,\ n \geq 1\}$ μία ανεξάρτητη και ισόνομη ακολουθία τ.μ., για την οποία ισχύει ότι

$$\overline{X}_n = \frac{X_1 + \ldots + X_n}{n} \xrightarrow{\text{a.s.}} \mu,$$

όπου $\mu \in \mathbb{R}$ κάποια σταθερά. Τότε $\mathbb{E}|X_1| < \infty$, και $\mathbb{E}(X_1) = \mu$.

Απόδειξη: Έστω $S_n = X_1 + \ldots + X_n$. Προφανώς, για $n \geq 2$ ισχύει ότι

$$\frac{X_n}{n} = \frac{S_n - S_{n-1}}{n} = \frac{S_n}{n} - \frac{n-1}{n}\frac{S_{n-1}}{n-1},$$

και αφού

$$\frac{S_n}{n} \xrightarrow{\text{a.s.}} \mu, \text{ και } -\frac{n-1}{n}\frac{S_{n-1}}{n-1} \xrightarrow{\text{a.s.}} -\mu, \quad (\text{βλ. Άσκηση } 6.3(\text{γ}))$$

συμπεραίνουμε ότι

$$\frac{X_n}{n} \xrightarrow{\text{a.s.}} 0, \quad \text{δηλ.} \quad \mathbb{P}\left(\frac{X_n}{n} \nrightarrow 0\right) = 0.$$

Αυτό σημαίνει ότι $\mathbb{P}[\limsup\{|X_n| \geq n\}] = 0$. Πράγματι, αν εκλέξουμε ένα $\omega \in \limsup\{\omega : |X_n(\omega)| \geq n\}$, τότε θα ισχύει ότι $|X_n(\omega)| \geq n$ για άπειρα n, και συνεπώς, για το συγκεκριμένο αυτό ω θα ισχύει ότι $X_n(\omega)/n \nrightarrow 0$. Με άλλα λόγια, ισχύει η ανισότητα $\limsup\{\omega : |X_n(\omega)| \geq n\} \subset \{\omega : X_n(\omega)/n \nrightarrow 0\}$, και το συμπέρασμα προκύπτει άμεσα, αφού το ενδεχόμενο $\{\omega : X_n(\omega)/n \nrightarrow 0\}$ έχει πιθανότητα 0. Όμως, η ακολουθία ενδεχομένων $B_n = \{\omega : |X_n(\omega)| \geq n\}$ είναι ανεξάρτητη, διότι $B_n \in \sigma(X_n)$, και οι $\sigma(X_n)$ είναι ανεξάρτητες. Αν ήταν $\sum_{n=1}^{\infty} \mathbb{P}(B_n) = +\infty$, τότε θα προέκυπτε, από το δεύτερο Λήμμα Borel-Cantelli, ότι $\mathbb{P}(\limsup B_n) = 1$ (άτοπο). Συνεπώς, $\sum_{n=1}^{\infty} \mathbb{P}(B_n) < +\infty$, δηλ. $\sum_{n=1}^{\infty} \mathbb{P}(|X_n| \geq n) < +\infty$, και, λόγω ισονομίας, $\sum_{n=1}^{\infty} \mathbb{P}(|X_1| \geq n) < +\infty$. Η τελευταία ανισότητα δείχνει ότι (βλ. Λήμμα 6.19)

$$\mathbb{E}|X_1| < +\infty.$$

Επομένως, $\mathbb{E}(X_1) = c \in \mathbb{R}$, και ο ισχυρός νόμος (Θεώρημα 6.21) συνεπάγεται την σχέση

$$\frac{X_1 + \ldots + X_n}{n} \xrightarrow{\textbf{a.s.}} c.$$

Όμως υποθέσαμε ότι

$$\frac{X_1 + \ldots + X_n}{n} \xrightarrow{\textbf{a.s.}} \mu,$$

και συνεπώς, $\mu = c$, και $\mathbb{E}(X_1) = \mu$. [Το όριο της σχεδόν βέβαιης σύγκλισης είναι μοναδικό (με πιθ. 1), και άρα πρέπει $\mu = c$, με πιθ. 1, όμως τα μ, c είναι σταθερές, άρα $\mu = c$ υποχρεωτικά.] $\square$

Κατ΄ αναλογίαν με το Θεώρημα 6.21, ισχύει ο ισχυρός νόμος και για $\mu = \pm\infty$:

Θεώρημα 6.23 (εκτεταμένος ισχυρός νόμος) Αν η ακολουθία $\{X_n,\ n \geq 1\}$ είναι ανεξάρτητη και ισόνομη, με $\mathbb{E}(X_1) = +\infty$ (δηλ. $\mathbb{E}(X_1^+) = +\infty$, $\mathbb{E}(X_1^-) < +\infty$), τότε

$$\frac{X_1 + \ldots + X_n}{n} \xrightarrow{\textbf{a.s.}} +\infty.$$

Ομοίως, αν $\mathbb{E}(X_1) = -\infty$, τότε

$$\frac{X_1 + \ldots + X_n}{n} \xrightarrow{\textbf{a.s.}} -\infty.$$

Σημειώνεται ότι όταν γράφουμε $Y_n \xrightarrow{\textbf{a.s.}} +\infty$, αυτό σημαίνει ότι $\mathbb{P}(C) = 1$, όπου

$$C = \{\omega : Y_n(\omega) \to +\infty\} = \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \{\omega : Y_k(\omega) \geq m\} = \bigcap_{m=1}^{\infty} \liminf_n \{Y_n \geq m\},$$

και ανάλογη σημασία έχει και η δήλωση $Y_n \xrightarrow{\textbf{a.s.}} -\infty$.

Απόδειξη: Υποθέτουμε ότι $\mathbb{E}(X_1) = +\infty$, δηλαδή $\mathbb{E}(X_1^+) = +\infty$, $\mathbb{E}(X_1^-) < +\infty$, σταθεροποιούμε έναν ακέραιο $M \in \{1, 2, \ldots\}$, και θέτουμε $Y_{n,M} = \min\{X_n, M\}$. Προφανώς, $Y_{n,M}^- = X_n^-$, $Y_{n,M}^+ = \min\{X_n^+, M\} \leq M$, οπότε $\mathbb{E}|Y_{n,M}| < \infty$, και

$$\begin{aligned}
\mathbb{E}(Y_{n,M}) &= \mathbb{E}(Y_{n,M}^+) - \mathbb{E}(Y_{n,M}^-) = \mathbb{E}(\min\{X_n^+, M\}) - \mathbb{E}(X_n^-) \\
&= \mathbb{E}(\min\{X_n^+, M\} - X_n^-) = \mathbb{E}(\min\{X_n, M\}) = \mathbb{E}(\min\{X_1, M\}).
\end{aligned}$$

Άρα, λόγω του Θεωρήματος 6.21, ισχύει ότι

$$\frac{Y_{1,M} + \ldots + Y_{n,M}}{n} \xrightarrow{\textbf{a.s.}} \mu(M), \quad \text{καθώς } n \to \infty,$$

όπου $\mu(M) = \mathbb{E}(Y_{1,M}) = \mathbb{E}(\min\{X_1, M\}) \in \mathbb{R}$, επειδή οι τ.μ. $Y_{n,M}$, $n = 1, 2, \ldots$, είναι ανεξάρτητες (ως $\sigma(X_n)$–μετρήσιμες), ισόνομες (αφού οι X_n είναι ισόνομες), και με πεπερασμένη μέση τιμή, $\mu(M)$, όπως παραπάνω. Από το Θεώρημα Μονότονης Σύγκλισης προκύπτει ότι $\mathbb{E}(Y_{1,M}^+) \to +\infty$, καθώς $M \to \infty$, επειδή $0 \leq Y_{1,M}^+ = \min\{X_1^+, M\} \nearrow X_1^+$, καθώς $M \to \infty$, και $\mathbb{E}(X_1^+) = +\infty$. Επομένως, ισχύει ότι και $\mu(M) = \mathbb{E}(Y_{1,M}) = \mathbb{E}(Y_{1,M}^+) - \mathbb{E}(Y_{1,M}^-) = \mathbb{E}(Y_{1,M}^+) - \mathbb{E}(X_1^-) \to +\infty$, καθώς $M \to \infty$. Θέτουμε

$$A_M = \left\{ \omega : \frac{Y_{1,M}(\omega) + \ldots + Y_{n,M}(\omega)}{n} \to \mu(M), \quad \text{καθώς } n \to \infty \right\},$$

οπότε $\mathbb{P}(A_M) = 1$ για $M = 1, 2, \ldots$, σύμφωνα με τα προηγούμενα. Αν διαλέξουμε ένα $\omega \in \bigcap_{M=1}^{\infty} A_M$, τότε, αφού $X_n(\omega) \geq Y_{n,M}(\omega)$ για κάθε n, ω και M, έπεται ότι

$$\frac{X_1(\omega) + \ldots + X_n(\omega)}{n} \geq \frac{Y_{1,M}(\omega) + \ldots + Y_{n,M}(\omega)}{n} \to \mu(M), \ \text{καθώς } n \to \infty,$$

οπότε, για το συγκεκριμένο αυτό ω,

$$\liminf_{n\to\infty} \frac{X_1(\omega) + \ldots + X_n(\omega)}{n} \geq \mu(M).$$

Όμως, αφού το συγκεκριμένο αυτό ω ανήκει σε όλα τα ενδεχόμενα A_M, έπεται ότι το

$$\liminf_{n\to\infty} \frac{X_1(\omega) + \ldots + X_n(\omega)}{n} = \alpha(\omega)$$

θα είναι ένας εκτεταμένος πραγματικός αριθμός ($\alpha(\omega) \in [-\infty, +\infty]$), με την ιδιότητα

$$\alpha(\omega) \geq \mu(M), \quad \text{για κάθε } M = 1, 2, \ldots,$$

και έτσι, $\alpha(\omega) = +\infty$, διότι το $\alpha(\omega)$ είναι σταθερό ως προς M, ενώ το $\mu(M)$ τείνει στο $+\infty$, καθώς $M \to +\infty$. Αυτό σημαίνει ότι

$$\bigcap_{M=1}^{\infty} A_M \subset \left\{ \omega : \frac{X_1(\omega) + \ldots + X_n(\omega)}{n} \to +\infty \right\},$$

και αφού $\mathbb{P}(A_M) = 1$ για κάθε $M = 1, 2, \ldots$, έπεται ότι και $\mathbb{P}\left(\bigcap_{M=1}^{\infty} A_M\right) = 1$, επίσης. Άρα,

$$\mathbb{P}\left(\frac{X_1 + \ldots + X_n}{n} \to +\infty\right) = 1.$$

Η περίπτωση $\mathbb{E}(X_1) = -\infty$ ανάγεται στην προηγούμενη, θεωρώντας την ακολουθία $\{-X_n,\ n \geq 1\}$. $\square$

Αξίζει να σημειώσουμε ότι το αντίστροφο του Θεωρήματος 6.23 **δεν ισχύει**. Αν, για παράδειγμα, για την ανεξάρτητη και ισόνομη ακολουθία τ.μ. $\{X_n,\ n \geq 1\}$, οι δειγματικοί μέσοι συγκλίνουν σχεδόν βέβαια προς το $+\infty$, τότε δεν μπορούμε να συμπεράνουμε ότι και $\mathbb{E}(X_1) = +\infty$. Φυσικά, από τα Θεωρήματα 6.21 και 6.23 προκύπτει ότι, σε αυτήν την περίπτωση, αποκλείεται να ισχύει $\mathbb{E}|X_1| < +\infty$, ή $\mathbb{E}(X_1) = -\infty$, χωρίς όμως να μπορεί να αποκλειστεί και η περίπτωση $\mathbb{E}(X_1^+) = \mathbb{E}(X_1^-) = +\infty$, κατά την οποία η μέση τιμή **δεν ορίζεται**. Τέτοια παραδείγματα έχουν βρεθεί από τους C. Derman και H. Robbins[13] (βλ. Chung, 1974, σελ. 131).

6.6 Ισχυρός Νόμος του Etemadi

Η απόδειξη του Θεωρήματος 6.21 χρησιμοποιεί **ουσιωδώς** την ανεξαρτησία των X_n, διότι στηρίζεται στο Θεώρημα 6.17, το

[13]Derman, C.; Robbins, H. (1955). The strong law of large numbers when the first moment does not exist. *Proc. Nat. Acad. Sci. U.S.A.*, vol. 41, pp. 586–587.

οποίο, με την σειρά του, κάνει χρήση της Πρότασης 6.12, δηλ. της μεγιστικής ανισότητας του Kolmogorov. Το 1981, ο N. Etemadi[14] παρουσίασε μία καινούρια απόδειξη, η οποία, εκτός των άλλων, είναι και ισχυρότερη από αυτήν του Kolmogorov (δηλ. του Θεωρήματος 6.21), διότι το συμπέρασμα ισχύει όταν οι ισόνομες τ.μ. X_n είναι, απλώς, **ανεξάρτητες ανά δύο**, χωρίς να απαιτείται η πλήρης ανεξαρτησία για όλη την ακολουθία $\{X_n,\ n \geq 1\}$. Στην παράγραφο αυτή θα παρουσιάσουμε το αποτέλεσμα και την απόδειξη του Etemadi.

Θεώρημα 6.24 (ισχυρός νόμος του Etemadi) Αν η ακολουθία τ.μ. $\{X_n,\ n \geq 1\}$ αποτελείται από ισόνομες και ανεξάρτητες ανά δύο τ.μ., και αν $\mathbb{E}|X_1| < +\infty$ και $\mathbb{E}(X_1) = \mu$, τότε

$$\overline{X}_n = \frac{X_1 + \ldots + X_n}{n} \xrightarrow{\textbf{a.s.}} \mu.$$

Απόδειξη: Κατ΄ αρχήν παρατηρούμε ότι είναι αρκετό να αποδείξουμε το θεώρημα για μη αρνητικές τ.μ. Πράγματι, χωρίζοντας σε θετικό και αρνητικό μέρος, $X_n = X_n^+ - X_n^-$, και υποθέτοντας ότι ισχύει το θεώρημα για μη αρνητικές, ισόνομες, και ανεξάρτητες ανά δύο τ.μ., με πεπερασμένη μέση τιμή, θα προκύψει

[14]Etemadi, N. (1981). An elementary proof of the strong law of large numbers. *Z. Wahrsch. Verw. Gebiete*, vol. 55, no. 1, pp. 119–122.

ότι[15]

$$\frac{X_1^+ + \ldots + X_n^+}{n} \xrightarrow{\textbf{a.s.}} \mathbb{E}(X_1^+), \quad \frac{X_1^- + \ldots + X_n^-}{n} \xrightarrow{\textbf{a.s.}} \mathbb{E}(X_1^-),$$

οπότε και

$$\frac{X_1 + \ldots + X_n}{n} = \frac{X_1^+ + \ldots + X_n^+}{n} - \frac{X_1^- + \ldots + X_n^-}{n} \xrightarrow{\textbf{a.s.}} \mathbb{E}(X_1^+) - \mathbb{E}(X_1^-) = \mu.$$

Υποθέτουμε λοιπόν ότι $X_n \geq 0$, ότι οι X_n είναι ισόνομες και ανεξάρτητες ανά δύο, και ότι $\mathbb{E}(X_1) = \mu < \infty$. Θέτουμε $Y_k = X_k I(X_k \leq k)$, $S_n = \sum_{k=1}^n X_k$, και $\widetilde{S}_n = \sum_{k=1}^n Y_k$. Επίσης, θέτουμε $u_n = [a^n] \to \infty$, όπου $a > 1$ είναι κάποιος (αυθαίρετος) πραγματικός αριθμός, και $[x]$ δηλώνει το ακέραιο μέρος του x. Για τυχόν $\varepsilon > 0$, θα δείξουμε πρώτα ότι

$$S(\varepsilon) = \sum_{n=1}^{\infty} \mathbb{P}\left[\frac{\left|\widetilde{S}_{u_n} - \mathbb{E}(\widetilde{S}_{u_n})\right|}{u_n} > \varepsilon\right] < \infty. \tag{6.1}$$

Πράγματι, οι τ.μ. Y_k είναι φραγμένες και ανεξάρτητες ανά δύο, ενώ οι X_k είναι ισόνομες, και έτσι έχουμε

$$\begin{aligned} \mathrm{Var}(\widetilde{S}_n) &= \sum_{k=1}^n \mathrm{Var}(Y_k) \leq \sum_{k=1}^n \mathbb{E}(Y_k^2) = \sum_{k=1}^n \mathbb{E}[X_k^2 I(X_k \leq k)] \\ &= \sum_{k=1}^n \mathbb{E}[X_1^2 I(X_1 \leq k)] \leq n\mathbb{E}[X_1^2 I(X_1 \leq n)], \end{aligned}$$

[15] Παρατηρήστε ότι και οι δύο ακολουθίες $\{X_n^+,\ n \geq 1\}$ και $\{X_n^-,\ n \geq 1\}$, αποτελούνται από ισόνομες, και ανεξάρτητες ανά δύο τ.μ., με πεπερασμένη μέση τιμή.

οπότε, από την ανισότητα Chebychev και το Θεώρημα Beppo Levi,

$$S(\varepsilon)=\sum_{n=1}^{\infty}\mathbb{P}\left[|\widetilde{S}_{u_n}-\mathbb{E}(\widetilde{S}_{u_n})|>\varepsilon u_n\right]\leq\sum_{n=1}^{\infty}\frac{\mathrm{Var}(\widetilde{S}_{u_n})}{\varepsilon^2u_n^2}$$
$$\leq\frac{1}{\varepsilon^2}\sum_{n=1}^{\infty}\frac{1}{u_n^2}u_n\,\mathbb{E}[X_1^2I(X_1\leq u_n)]=\frac{1}{\varepsilon^2}\mathbb{E}\left[X_1^2\sum_{n=1}^{\infty}\frac{I(X_1\leq u_n)}{u_n}\right].$$

Θεωρούμε την συνάρτηση g, με τύπο

$$g(x)=\sum_{n=1}^{\infty}\frac{I(x\leq u_n)}{u_n},\quad x\geq 0,$$

οπότε η παραπάνω ανισότητα γράφεται ως εξής:

$$S(\varepsilon)\leq\frac{1}{\varepsilon^2}\mathbb{E}[X_1^2g(X_1)].\tag{6.2}$$

Είναι σαφές ότι για τυχόν $x>0$,

$$g(x)=\sum_{n:u_n\geq x}\frac{1}{u_n}=\sum_{n\geq N(x)}\frac{1}{u_n}=\sum_{n\geq N(x)}\frac{1}{[a^n]},$$

όπου $N(x)=\min\{n:u_n\geq x\}$. Χρησιμοποιώντας την ανισότητα $1/[y]\leq 2/y$, που ισχύει για κάθε $y\geq 1$, έχουμε (για τυχόν $x>0$)

$$g(x)=\sum_{n\geq N(x)}\frac{1}{[a^n]}\leq\sum_{n\geq N(x)}\frac{2}{a^n}=\frac{2}{a^{N(x)}}\sum_{n=0}^{\infty}\frac{1}{a^n}=\frac{2a}{a-1}\frac{1}{a^{N(x)}}=\frac{K(a)}{a^{N(x)}},$$

όπου θέσαμε $K(a)=2a/(a-1)$. Όμως, από τον ορισμό του ακεραίου $N(x)$, έχουμε $u_{N(x)}\geq x$, δηλαδή $[a^{N(x)}]\geq x$, και επειδή $y\geq[y]$ για κάθε y, έπεται ότι $a^{N(x)}\geq x$, και τελικά,

$$g(x) \leq \frac{K(a)}{a^{N(x)}} \leq \frac{K(a)}{x},$$

οπότε ισχύει ότι $x^2 g(x) \leq K(a)x$, για κάθε $x \geq 0$. Άρα, $X_1^2 g(X_1) \leq K(a)X_1$, και από την (6.2) παίρνουμε την (6.1), επειδή $\mathbb{E}(X_1) < \infty$.

Από την (6.1) και το πρώτο Λήμμα Borel–Cantelli προκύπτει ότι

$$\mathbb{P}\left[\frac{\left|\widetilde{S}_{u_n} - \mathbb{E}(\widetilde{S}_{u_n})\right|}{u_n} > \varepsilon \text{ για άπειρα } n\right] = \mathbb{P}\left[\limsup_n \left\{\frac{\left|\widetilde{S}_{u_n} - \mathbb{E}(\widetilde{S}_{u_n})\right|}{u_n} > \varepsilon\right\}\right] = 0,$$

και επειδή το $\varepsilon > 0$ είναι τυχόν, συμπεραίνουμε ότι

$$\frac{\widetilde{S}_{u_n} - \mathbb{E}(\widetilde{S}_{u_n})}{u_n} \xrightarrow{\textbf{a.s.}} 0. \tag{6.3}$$

Τώρα παρατηρούμε ότι $X_1 I(X_1 \leq n) \nearrow X_1$, και συνεπώς, από το Θεώρημα Μονότονης Σύγκλισης, $\mathbb{E}(Y_n) = \mathbb{E}[X_1 I(X_1 \leq n)] \to \mathbb{E}(X_1) = \mu$. Επομένως (βλ. Λήμμα 6.13(i)), $\mathbb{E}(\widetilde{S}_n)/n = (\mathbb{E}(Y_1) + \ldots + \mathbb{E}(Y_n))/n \to \mu$, και άρα, $\mathbb{E}(\widetilde{S}_{u_n})/u_n \to \mu$, αφού η $u_n = [a^n]$ είναι ακολουθία ακεραίων που τείνει στο άπειρο. Έτσι, από την (6.3) παίρνουμε την

$$\frac{\widetilde{S}_{u_n}}{u_n} \xrightarrow{\textbf{a.s.}} \mu. \tag{6.4}$$

Τώρα, όπως στην απόδειξη του Θεωρήματος 6.21, έχουμε (βλ. Λήμμα 6.19)

$$\sum_{n=1}^{\infty} \mathbb{P}(X_n \neq Y_n) = \sum_{n=1}^{\infty} \mathbb{P}(X_n > n) = \sum_{n=1}^{\infty} \mathbb{P}(X_1 > n) \leq \mu < +\infty,$$

οπότε, και πάλι από το πρώτο Λήμμα Borel-Cantelli,

$$\mathbb{P}(X_n \neq Y_n \text{ για άπειρα } n) = 0,$$

και έτσι,

$$\frac{S_n - \widetilde{S}_n}{n} \xrightarrow{\text{a.s.}} 0, \quad \text{οπότε και} \quad \frac{S_{u_n} - \widetilde{S}_{u_n}}{u_n} \xrightarrow{\text{a.s.}} 0. \tag{6.5}$$

Προφανώς, οι σχέσεις (6.4) και (6.5) συνεπάγονται την

$$\frac{S_{u_n}}{u_n} \xrightarrow{\text{a.s.}} \mu, \quad \text{καθώς } n \to \infty, \tag{6.6}$$

όπου $u_n = [a^n]$, και η (6.6) ισχύει για κάθε $a > 1$. Παρατηρούμε τώρα ότι για τυχόντα φυσικό αριθμό $k \geq a$, υπάρχουν ακέραιοι της μορφής u_n και u_{n+1}, τέτοιοι ώστε $u_n \leq k \leq u_{n+1}$. Επειδή οι τ.μ. X_n είναι μη αρνητικές, έπεται ότι $0 \leq S_{u_n} \leq S_k \leq S_{u_{n+1}}$, και αφού $0 < 1/u_{n+1} \leq 1/k \leq 1/u_n$, συμπεραίνουμε την ανισότητα

$$\left(\frac{u_n}{u_{n+1}}\right)\frac{S_{u_n}}{u_n} \leq \frac{S_k}{k} \leq \left(\frac{u_{n+1}}{u_n}\right)\frac{S_{u_{n+1}}}{u_{n+1}}. \tag{6.7}$$

Είναι τώρα εύκολο να διαπιστώσουμε ότι $u_{n+1}/u_n = [a^{n+1}]/[a^n] \to a$, και έτσι, οι (6.6) και (6.7) δείχνουν ότι, με πιθανότητα 1,

$$\frac{1}{a}\mu \leq \liminf_{k\to\infty} \frac{S_k}{k} \leq \limsup_{k\to\infty} \frac{S_k}{k} \leq a\mu. \tag{6.8}$$

Επειδή το $a > 1$ είναι αυθαίρετο, μπορούμε π.χ. να εκλέξουμε $a = 1 + 1/m$, οπότε η (6.8) μας εξασφαλίζει ότι για κάθε $m = 1, 2, \ldots$, τα ενδεχόμενα

$$A_m = \left\{\omega : \frac{m}{m+1}\mu \leq \liminf \frac{S_n(\omega)}{n} \leq \limsup_n \frac{S_n(\omega)}{n} \leq \frac{m+1}{m}\mu\right\}$$

έχουν πιθανότητα 1. Είναι σαφές ότι και το ενδεχόμενο $A = \bigcap_{m=1}^{\infty} A_m$ έχει πιθανότητα 1, και ότι

$$A = \left\{ \omega : \frac{S_n(\omega)}{n} \to \mu \right\},$$

πράγμα που αποδεικνύει το θεώρημα. $\square$

Ασκήσεις Κεφ. 6:

6.1. (α) Αποδείξτε ότι αν $X_n \xrightarrow{L^r} X$ για κάποιο $r \geq 1$, τότε $X_n \xrightarrow{L^p} X$ για κάθε p, με $1 \leq p \leq r$.

(β) Βρείτε (αντι)παράδειγμα για κάθε μία από τις (i)–(v), ορίζοντας ακολουθία τ.μ. $\{X_n, n \geq 1\}$, και τ.μ. X (σε κάποιον χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$), έτσι ώστε:

(i) $X_n \xrightarrow{d} X$, αλλά $X_n \overset{p}{\nrightarrow} X$.

(ii) $X_n \xrightarrow{p} X$, αλλά $X_n \overset{a.s.}{\nrightarrow} X$, και $X_n \overset{L^p}{\nrightarrow} X$ για κανένα $p \geq 1$.

(iii) $X_n \xrightarrow{a.s.} X$, αλλά $X_n \overset{L^p}{\nrightarrow} X$ για κανένα $p \geq 1$.

(iv) Για δοθέν $p \geq 1$, $X_n \xrightarrow{L^p} X$, αλλά $X_n \overset{a.s.}{\nrightarrow} X$.

(v) Για δοθέντα p, r, με $1 \leq p < r$, $X_n \xrightarrow{L^p} X$, αλλά $X_n \overset{L^r}{\nrightarrow} X$.

6.2. Δείξτε ότι:

(i) Αν ο Ω είναι αριθμήσιμος και $\mathcal{A} = \mathcal{P}(\Omega)$, τότε $X_n \xrightarrow{p} X \iff X_n \xrightarrow{a.s.} X$.

(ii) Αν οι $\{X_n,\ n \geq 1\}$ και X είναι ορισμένες στον ίδιο χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, τότε $X_n \xrightarrow{\mathbf{d}} \alpha \iff X_n \xrightarrow{\mathbf{p}} \alpha$ (όπου $\alpha \in \mathbb{R}$, σταθερά).

6.3. (α) Δείξτε το Θεώρημα του Slutsky, ότι δηλ. αν $X_n \xrightarrow{\mathbf{d}} X$ και $Y_n \xrightarrow{\mathbf{d}} c$ $(\in \mathbb{R})$, τότε $X_n + Y_n \xrightarrow{\mathbf{d}} X + c$, $X_nY_n \xrightarrow{\mathbf{d}} cX$, και $X_n/Y_n \xrightarrow{\mathbf{d}} X/c$ (εφ' όσον $c \neq 0$).

(β) Έστω g μία συνεχής συνάρτηση $(g : \mathbb{R} \longrightarrow \mathbb{R})$. Δείξτε ότι αν $X_n \xrightarrow{\mathbf{d}} X$ τότε $g(X_n) \xrightarrow{\mathbf{d}} g(X)$, ότι αν $X_n \xrightarrow{\mathbf{p}} X$ τότε $g(X_n) \xrightarrow{\mathbf{p}} g(X)$, και ότι αν $X_n \xrightarrow{\mathbf{a.s.}} X$ τότε $g(X_n) \xrightarrow{\mathbf{a.s.}} g(X)$. Αν η g είναι συνεχής στο $c \in \mathbb{R}$, και $X_n \xrightarrow{\mathbf{d}} c$, τότε $g(X_n) \xrightarrow{\mathbf{d}} g(c)$.

(γ) Αν $X_n \xrightarrow{\mathbf{a.s.}} X$ και $Y_n \xrightarrow{\mathbf{a.s.}} Y$, και το $X + Y$ ορίζεται για κάθε ω (δηλ. δεν είναι της μορφής $(+\infty) + (-\infty)$, ή $(-\infty) + (+\infty)$), τότε $X_n + Y_n \xrightarrow{\mathbf{a.s.}} X + Y$. Αν $Y(\omega) \equiv c \in \mathbb{R}$, τότε $X_nY_n \xrightarrow{\mathbf{a.s.}} cX$.

6.4. Αποδείξτε ότι το όριο της ασθενούς σύγκλισης είναι μοναδικό (με την έννοια ότι αν $X_n \xrightarrow{\mathbf{d}} X$ και $X_n \xrightarrow{\mathbf{d}} Y$, τότε $F_X = F_Y$, δηλ. οι X και Y έχουν την ίδια κατανομή).

6.5. Αποδείξτε την εξής γενίκευση του δευτέρου Λήμματος Borel-Cantelli: Θεωρούμε μία ακολουθία ενδεχομένων $\{A_n,\ n \geq 1\}$, για την οποία ισχύει ότι $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = +\infty$. Αν, επιπροσθέτως, ισχύει ότι

$$\liminf_{n\to\infty} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} \mathbb{P}(A_i A_j)}{\left[\sum_{k=1}^{n} \mathbb{P}(A_k)\right]^2} \leq 1,$$

τότε $\mathbb{P}(\limsup A_n) = 1$.

Ειδικότερα, αν $\mathbb{P}(A_iA_j) = \mathbb{P}(A_i)\mathbb{P}(A_j)$ για κάθε $i \neq j$, και $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = +\infty$, τότε $\mathbb{P}(\limsup A_n) = 1$.
[Επομένως, για να ισχύει το δεύτερο Λήμμα Borel-Cantelli, είναι αρκετό τα ενδεχόμενα $\{A_n,\ n \geq 1\}$ να είναι **ανά ζεύγη ανεξάρτητα.**]

[**Υπόδειξη:** Έστω $X_n = I_{A_n}$, και $S_n = \sum_{j=1}^{n} X_j$. Τότε $\mathbb{E}(S_n) = \sum_{j=1}^{n} \mathbb{P}(A_j) \to +\infty$, και $\mathrm{Var}(S_n) = (\vartheta_n - 1)[\mathbb{E}(S_n)]^2$, όπου ϑ_n το πηλίκο του οποίου λαμβάνουμε το κατώτερο όριο. Επίσης,

$$\mathbb{P}(S_n \leq x) \leq \frac{(\vartheta_n - 1)[\mathbb{E}(S_n)]^2}{[\mathbb{E}(S_n) - x]^2},$$

διότι $\mathbb{P}(S_n \leq x) \leq \mathbb{P}[|S_n - \mathbb{E}(S_n)| \geq \mathbb{E}(S_n) - x] \leq \mathrm{Var}(S_n)/[\mathbb{E}(S_n) - x]^2$, για $0 \leq x < \mathbb{E}(S_n)$. Άρα, αφού $\liminf \vartheta_n \leq 1$, έπεται ότι

$$\liminf_{n\to\infty} \mathbb{P}(S_n \leq x) = 0, \quad \text{για κάθε } x \in \mathbb{R}.$$

Όμως, $\mathbb{P}(\sup_{k\geq 1} S_k \leq x) \leq \mathbb{P}(S_n \leq x)$, $n = 1, 2, \ldots$, και συνεπώς, λαμβάνοντας κατώτερο όριο (του $n \to \infty$), έπεται ότι $\mathbb{P}(\sup_{k\geq 1} S_k \leq x) = 0$, για κάθε $x \in \mathbb{R}$. Τελικά, $\mathbb{P}(\limsup A_n) = \mathbb{P}(\sup_{k\geq 1} S_k = +\infty) = 1$.]

6.6. Αποδείξτε το αντίστροφο του ισχυρού νόμου του Etemadi, δηλ. ότι αν οι X_n είναι ισόνομες, και ανεξάρτητες ανά δύο, και αν

$$\overline{X}_n = \frac{X_1 + \ldots + X_n}{n} \xrightarrow{\text{a.s.}} \mu,$$

όπου $\mu \in \mathbb{R}$ κάποια σταθερά, τότε $\mathbb{E}|X_1| < \infty$, και $\mathbb{E}(X_1) = \mu$.

6.7. Δείξτε ότι $\mathbb{E}(X^2) < +\infty$ αν και μόνο αν $\sum_{n=1}^{\infty} n\mathbb{P}(|X| \geq n) < +\infty$.

6.8. (Θεώρημα Cantelli) Εάν η ακολουθία τ.μ. $\{X_n,\ n \geq 1\}$ είναι ανεξάρτητη, και $\mathbb{E}(X_n^4) < +\infty$ για κάθε n, και $\mathbb{E}|X_n - \mathbb{E}(X_n)|^4 \leq c < +\infty$ για $n = 1, 2, \ldots$, τότε

$$\frac{X_1 + \ldots + X_n}{n} - \frac{\mathbb{E}(X_1) + \ldots + \mathbb{E}(X_n)}{n} \xrightarrow{\text{a.s.}} 0.$$

[**Υπόδειξη:** Υποθέστε ότι $\mathbb{E}(X_n) = 0$ για κάθε n. Αν $S_n = X_1 + \ldots + X_n$, τότε

$$\mathbb{E}(S_n^4) = \sum_{i=1}^{n} \mathbb{E}(X_i^4) + 6 \sum_{1 \leq i < j \leq n} \mathbb{E}(X_i^2)\mathbb{E}(X_j^2) \leq (n + 3n(n-1))c < 3n^2 c,$$

οπότε $\sum_{n=1}^{\infty} \mathbb{E}[(S_n/n)^4] \leq 3c \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$. Εφαρμόστε την ανισότητα Chebychev για να δείξετε ότι για κάθε $\varepsilon > 0$, $\sum_{n=1}^{\infty} \mathbb{P}(\frac{|S_n|}{n} \geq \varepsilon) < +\infty$. Χρησιμοποιήστε τώρα το Λήμμα 6.3 και το πρώτο Λήμμα Borel-Cantelli.]

6.9. (Θεώρημα Borel) Ένας αριθμός $x \in (0,1)$ λέγεται **φυσιολογικός**, όταν το δυαδικό του ανάπτυγμα, $x = \sum_{n=1}^{\infty} x_n/2^n$ (όπου $x_n \in \{0,1\}$), περιέχει «το ίδιο πλήθος» από «0» και «1», δηλ. όταν ισχύει ότι

$$\lim_{n\to\infty} \frac{x_1 + \ldots + x_n}{n} = \frac{1}{2}.$$

Δείξτε ότι το μέτρο Lebesgue των φυσιολογικών αριθμών του $(0,1)$ είναι 1.

[**Υπόδειξη:** Θεωρήστε τις συναρτήσεις Rademacher R_n (Άσκηση ;;.5), στον χώρο $((0,1), \mathcal{B}((0,1)), \lambda)$, και χρησιμοποιήστε τον ισχυρό νόμο των μεγάλων αριθμών.]

6.10. (ταυτότητα Wald) Έστω $\{X_n,\ n \geq 1\}$ μία ανεξάρτητη και ισόνομη ακολουθία τ.μ. στον $(\Omega, \mathcal{A}, \mathbb{P})$, και N μία τ.μ., με $N(\omega) \in \{1,2,\ldots\}$ για κάθε $\omega \in \Omega$, και $\{\omega : N(\omega) = k\} \in \sigma(X_1,\ldots,X_k)$, για $k = 1,2,\ldots$. Θέτουμε $S = X_1 + \ldots + X_N$, δηλ. $S(\omega) = X_1(\omega) + X_2(\omega) + \ldots + X_{N(\omega)}(\omega)$. Αν $\mathbb{E}(N) < \infty$, και $\mathbb{E}|X_1| < \infty$, τότε,

$$\mathbb{E}(S) = \mathbb{E}(X_1)\mathbb{E}(N).$$

[**Υπόδειξη:** «Τυπικά»:

$$\begin{aligned}
\mathbb{E}(S) &= \mathbb{E}\left[\sum_{n=1}^{\infty}(X_1+\ldots+X_n)I(N=n)\right] = \mathbb{E}\left[\sum_{n=1}^{\infty} I(N=n)\sum_{k=1}^{n} X_k\right] \\
&= \sum_{n=1}^{\infty}\sum_{k=1}^{n} \mathbb{E}[X_k I(N=n)] = \sum_{k=1}^{\infty} \mathbb{E}\left[X_k \sum_{n=k}^{\infty} I(N=n)\right] \\
&= \sum_{k=1}^{\infty} \mathbb{E}[X_k I(N\geq k)] = \sum_{k=1}^{\infty} \mathbb{E}(X_k)\mathbb{P}(N\geq k) \\
&= \sum_{k=1}^{\infty} \mathbb{E}(X_1)\mathbb{P}(N\geq k) = \mathbb{E}(X_1)\mathbb{E}(N).
\end{aligned}$$

Γιατί είναι σωστά όλα αυτά;]

6.11. Δύο ακολουθίες τ.μ. $\{X_n,\ n\geq 1\}$ και $\{Y_n,\ n\geq 1\}$ λέγονται **ασυμπτωτικά ισοδύναμες** όταν $\sum_{n=1}^{\infty} \mathbb{P}(X_n \neq Y_n) < \infty$. Δείξτε ότι

$$\mathbb{P}\left(\sum_{n=1}^{\infty} X_n \text{ συγκλίνει}\right) = \mathbb{P}\left(\sum_{n=1}^{\infty} Y_n \text{ συγκλίνει}\right),$$

και ότι για τυχόν $\alpha\in\mathbb{R}$,

$$\mathbb{P}\left(\frac{X_1+\ldots+X_n}{n}\to\alpha\right) = \mathbb{P}\left(\frac{Y_1+\ldots+Y_n}{n}\to\alpha\right),$$

όταν οι $\{X_n,\ n\geq 1\}$ και $\{Y_n,\ n\geq 1\}$ είναι ασυμπτωτικά ισοδύναμες.

[**Υπόδειξη:** Λήμμα Borel-Cantelli.]

Κεφάλαιο 7

Ανακεφαλαίωση Συγκλίσεων – Θεώρημα Skorohod

Το παρόν κεφάλαιο έχει επαναληπτικό χαρακτήρα. Συνοψίζονται τα πιο βασικά αποτελέσματα που έχουν συζητηθεί μέχρι τώρα, σχετικά με την σύγκλιση ακολουθίας τυχαίων μεταβλητών (τ.μ.). Περιέχεται, επίσης, η απόδειξη κάποιων σημαντικών προτάσεων, οι οποίες έχουν αποδειχθεί και στο Κεφ. 6. Αυτό έγινε σκόπιμα, ώστε το μέρος 2 (Κεφ. 6–;;) να αποτελεί αυτόνομη, προς διδασκαλία, ενότητα, χωρίς, κατά το δυνατόν, πολλές παραπομπές στα προηγούμενα, αν και υποτίθεται ότι ο αναγνώστης γνωρίζει τα βασικά αποτελέσματα της Θεωρίας Μέτρου και Ολοκλήρωσης, που συζητήθηκαν στα Κεφ. ;;–;;, καθώς και τους νόμους των μεγάλων αριθμών του Κεφ. 6.

Ουσιαστικά, τα καινούριο αποτέλεσμα που παρουσιάζεται εδώ είναι το Θεώρημα εμφύτευσης του Skorohod, καθώς και οι εφαρμογές του σε άλλα (κλασικά) θεωρήματα. Η χρησιμότητα του Θεωρήματος Skorohod οφείλεται στο γεγονός ότι, στις Εφαρμοσμένες Πιθανότητες και την Στατιστική, σπανίως έχουμε πληροφορία για τον δειγματικό χώρο. Αντίθετα, έχουμε πληροφορία για τις κατανομές των τυχαίων μεταβλητών και την στοχαστική σχέση μεταξύ τους (ανεξαρτησία κ.λπ.), και το ενδιαφέρον επικεντρώνεται στην κατά κατανομή σύγκλιση των υπό μελέτη στατιστικών συναρτήσεων, και εν γένει, στην συμπεριφορά των αντιστοίχων συναρτήσεων κατανομής (*σ.κ.*), δηλ. του επαγόμενου μέτρου πιθανότητας στον $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Έτσι, τα κεντρικά θεωρήματα της Θεωρίας Ολοκλήρωσης (Μονότονη και Κυριαρχημένη Σύγκλιση, Θεώρημα Ομοιόμορφης Ολοκληρωσιμότητας κ.λπ.) δεν μπορούν να εφαρμοστούν άμεσα, ακριβώς επειδή οι τ.μ. που υπεισέρχονται σε αυτά είναι γνωστές μόνο μέσω των αντιστοίχων *σ.κ.*

Το κατάλληλο αποτέλεσμα που επιτρέπει την αναγωγή της ασθενούς σύγκλισης σε ισχυρή, εμφυτεύοντας τις υπό μελέτη τ.μ. σε κοινό χώρο πιθανότητας, είναι ακριβώς το Θεώρημα Skorohod.

7.1 Σύγκλιση Ακολουθίας Τυχαίων Μεταβλητών

Υπενθυμίζονται οι βασικές μορφές στοχαστικής σύγκλισης που χρησιμοποιούνται στις πιθανότητες.

Σύγκλιση με πιθανότητα 1 (σχεδόν βέβαιη ή ισχυρή σύγκλιση).

Λέμε ότι η ακολουθία τ.μ. $\{X_n,\ n \geq 1\}$ συγκλίνει προς την τ.μ. X, με πιθανότητα 1 (almost surely – συμβολισμός: $X_n \xrightarrow{\text{a.s.}} X$), όταν οι X_n και X είναι ορισμένες **στον ίδιο** χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, και

$$\mathbb{P}(\lim X_n = X) = 1,$$

δηλ. $\mathbb{P}(C) = 1$, όπου $C = \{\omega :\ X_n(\omega) \to X(\omega)\}$.

Σύγκλιση κατά πιθανότητα.

Λέμε ότι η ακολουθία των τ.μ. $\{X_n,\ n \geq 1\}$ συγκλίνει, κατά πιθανότητα, (in probability) προς την τ.μ. X, όταν οι τ.μ. X_n και X είναι ορισμένες **στον ίδιο** χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, και ισχύει:

$$\text{Για κάθε } \varepsilon > 0,\ \ \mathbb{P}(|X_n - X| \geq \varepsilon) \to 0,\ \text{ καθώς } n \to \infty.$$

Συμβολισμός: $X_n \xrightarrow{\mathbf{p}} X$.

Σύγκλιση στον L^p.

Λέμε ότι η ακολουθία τ.μ. $\{X_n,\ n \geq 1\}$ συγκλίνει προς την τ.μ. X κατά L^p $(p \geq 1)$, όταν οι X_n και X είναι ορισμένες **στον ίδιο** χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, και ισχύουν οι σχέσεις $\mathbb{E}|X|^p < \infty$, $\mathbb{E}|X_n|^p < \infty$ (από κάποιο n και πάνω), και

$$\mathbb{E}|X_n - X|^p \to 0 \ \text{ καθώς } n \to \infty.$$

Συμβολισμός: $X_n \xrightarrow{\boldsymbol{L^p}} X$.

Ιδιαίτερο ενδιαφέρον παρουσιάζουν οι περιπτώσεις $p = 1$ και

$p = 2$:

$p = 1$: $X_n \xrightarrow{L^1} X \overset{\text{ορ.}}{\Longleftrightarrow} \mathbb{E}|X_n - X| \to 0$ (σύγκλιση στον L^1).

$p = 2$: $X_n \xrightarrow{L^2} X \overset{\text{ορ.}}{\Longleftrightarrow} \mathbb{E}(X_n - X)^2 \to 0$ (σύγκλιση στον L^2 ή κατά μέσο τετράγωνο).

Σύγκλιση κατά κατανομή (ασθενής σύγκλιση).

Η ακολουθία των τ.μ. $\{X_n,\ n \geq 1\}$ λέμε ότι συγκλίνει κατά κατανομή (in distribution) προς την τ.μ. X (συμβολισμός: $X_n \xrightarrow{\mathbf{d}} X$), όταν[1]

$$\mathbb{P}[X_n \leq x] \to \mathbb{P}[X \leq x],\ \text{καθώς } n \to \infty,$$

για κάθε x για το οποίο $\mathbb{P}[X = x] = 0$.

Παρατήρηση 7.1 Η σύγκλιση κατά κατανομή **δεν απαιτεί** οι τ.μ. $\{X_n,\ n \geq 1\}$ και X να είναι ορισμένες στον ίδιο χώρο πιθανότητας. Η σύγκλιση αυτή είναι συνέπεια της συμπεριφοράς των σ.κ. F_n και F (των τ.μ. X_n και X), **και όχι** των ιδίων των X_n και X. Ο συμβολισμός της σύγκλισης μέσω των σ.κ. λαμβάνει την μορφή $F_n \xrightarrow{\mathbf{d}} F$, και ισοδυναμεί με την δήλωση:

$$F_n(x) \to F(x),\ \text{για κάθε σημείο συνεχείας } x \text{ της } F.$$

Θα γράφουμε $X_n \xrightarrow{\mathbf{d}} X$, ή $F_n \xrightarrow{\mathbf{d}} F$, ανάλογα με το αν θέλουμε να δώσουμε έμφαση στις τ.μ., ή τις σ.κ. τους. Πάντως, οποιοσδήποτε συμβολισμός θα σημαίνει ακριβώς το ίδιο πράγμα.

Θεώρημα 7.2 (σχέση τρόπων σύγκλισης) Έστω $\{X_n,\ n \geq 1\}$ και X τ.μ., ορισμένες στον χώρο ίδιο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$. Τότε:

---

[1] ορθότερα, θα έπρεπε να γράφαμε $\mathbb{P}_n[X_n \leq x] \to \mathbb{P}[X \leq x]$, αφού τα μέτρα πιθανότητας $\mathbb{P}_n$ επιτρέπεται να μεταβάλλονται με το n, αλλά προτιμάμε να χρησιμοποιήσουμε τον συνήθη συμβολισμό.

(i) Εάν $X_n \xrightarrow{\textbf{a.s.}} X$, τότε $X_n \xrightarrow{\textbf{p}} X$.

(ii) Εάν $X_n \xrightarrow{\boldsymbol{L}^{\mathbf{p}}} X$, τότε $X_n \xrightarrow{\textbf{p}} X$.

(iii) Εάν $X_n \xrightarrow{\boldsymbol{L}^{\mathbf{p}}} X$, τότε $X_n \xrightarrow{\boldsymbol{L}^{\mathbf{r}}} X$, για $1 \leq r \leq p$.

(iv) Εάν $X_n \xrightarrow{\textbf{p}} X$, τότε $X_n \xrightarrow{\textbf{d}} X$.

Από το θεώρημα αυτό γίνεται φανερό γιατί ονομάζουμε «ασθενή» την κατά κατανομή σύγκλιση.

Οι αποδείξεις του Θεωρήματος 7.2 έχουν δοθεί στο Κεφ. 6. Το (i), το οποίο δεν είναι τελείως προφανές, προκύπτει από την παρακάτω πρόταση. Για λόγους πληρότητας θα δοθεί και εδώ η απόδειξη της.

Πρόταση 7.3 $X_n \xrightarrow{\textbf{a.s.}} X$ αν και μόνο αν για κάθε $\varepsilon > 0$,

$$\mathbb{P}\left(\sup_{k \geq n} |X_k - X| > \varepsilon\right) \to 0, \quad \text{καθώς } n \to \infty.$$

Απόδειξη: Έστω $\varepsilon > 0$. Θέτουμε

$$\begin{aligned}
A^{\varepsilon} &= \{\omega \in \Omega : |X_n(\omega) - X(\omega)| > \varepsilon \text{ για άπειρες τιμές του } n\} \\
&= \limsup_n \{|X_n - X| > \varepsilon\} \\
&= \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \{|X_k - X| > \varepsilon\} \\
&= \bigcap_{n=1}^{\infty} \left\{\sup_{k \geq n} |X_k - X| > \varepsilon\right\} \\
&= \lim_n \left\{\sup_{k \geq n} |X_k - X| > \varepsilon\right\}.
\end{aligned}$$

Ας υποθέσουμε ότι $X_n \xrightarrow{\textbf{a.s.}} X$. Τότε $\mathbb{P}(\lim X_n = X) = 1$, και άρα, $\mathbb{P}(A^\varepsilon) = 0$, διότι $A^\varepsilon \subset \{\omega \in \Omega : X_n(\omega) \nrightarrow X(\omega)\}$. Συνεπώς,

$$\mathbb{P}\bigg(\lim_n \Big\{\sup_{k\geq n} |X_k - X| > \varepsilon\Big\}\bigg) = \lim_n \mathbb{P}\bigg(\sup_{k\geq n} |X_k - X| > \varepsilon\bigg) = 0,$$

διότι $A_n \to A \Rightarrow \mathbb{P}(A_n) \to \mathbb{P}(A)$, και το εφαρμόσαμε αυτό για την φθίνουσα ακολουθία ενδεχομένων

$$\begin{aligned} A_n &= \Big\{\sup_{k\geq n} |X_k - X| > \varepsilon\Big\} \\ &= \{\omega \in \Omega : \text{υπάρχει } k \geq n \text{ για το οποίο } |X_k(\omega) - X(\omega)| > \varepsilon\} \\ &= \bigcup_{k=n}^{\infty} \{|X_k - X| > \varepsilon\} \to A^\varepsilon. \end{aligned}$$

Αντίστροφα, εάν για κάθε $\varepsilon > 0$,

$$\mathbb{P}\bigg(\sup_{k\geq n} |X_k - X| > \varepsilon\bigg) \to 0,$$

δηλαδή

$$\lim_n \mathbb{P}\bigg(\sup_{k\geq n} |X_k - X| > \varepsilon\bigg) = 0,$$

τότε

$$\mathbb{P}\bigg(\lim_n \Big\{\sup_{k\geq n} |X_k - X| > \varepsilon\Big\}\bigg) = 0,$$

δηλαδή $\mathbb{P}(A^\varepsilon) = 0$. Τότε όμως, επειδή

$$\{\omega \in \Omega : X_n(\omega) \nrightarrow X(\omega)\} = \bigcup_{m=1}^{\infty} A^{1/m}, \qquad (\text{γιατί;})$$

έπεται ότι

$$\mathbb{P}(\lim X_n \neq X) = \mathbb{P}\bigg(\bigcup_{m=1}^{\infty} A^{1/m}\bigg) \leq \sum_{m=1}^{\infty} \mathbb{P}(A^{1/m}) = 0,$$

και άρα, $\mathbb{P}(\lim X_n = X) = 1$, όπως έπρεπε να δειχθεί. $\square$

Πόρισμα 7.4 Εάν $X_n \xrightarrow{\text{a.s.}} X$, τότε $X_n \xrightarrow{\text{p}} X$.

Απόδειξη: Προφανώς $\{|X_n - X| > \varepsilon\} \subset \left\{\sup_{k\geq n} |X_k - X| > \varepsilon\right\}$.
$\square$

Μερικές φορές είναι πολύ χρήσιμο, ιδιαίτερα σε κάποιες αποδείξεις, να έχουμε ισχυρή σύγκλιση στην θέση της σύγκλισης κατά πιθανότητα. Για παράδειγμα, αν η $g : \mathbb{R} \longrightarrow \mathbb{R}$ είναι μία συνεχής συνάρτηση, και γνωρίζουμε ότι $X_n \xrightarrow{\text{a.s.}} X$, τότε μπορούμε εύκολα να συμπεράνουμε ότι και $g(X_n) \xrightarrow{\text{a.s.}} g(X)$, επειδή $g(x_n) \to g(x)$ για κάθε ακολουθία $x_n \to x$ (λόγω συνεχείας της g), και άρα, τα ω για τα οποία ισχύει ότι $X_n(\omega) \to X(\omega)$, ικανοποιούν και την $g(X_n(\omega)) \to g(X(\omega))$, που σημαίνει ότι $\{\omega \in \Omega : g(X_n(\omega)) \to g(X(\omega))\} \supset \{\omega \in \Omega : X_n(\omega) \to X(\omega)\}$, και έτσι, $\mathbb{P}[g(X_n) \to g(X)] \geq \mathbb{P}[X_n \to X] = 1$. Το επόμενο λήμμα μας δείχνει πως η κατά πιθανότητα σύγκλιση ισοδυναμεί με ισχυρή σύγκλιση μέσω κατάλληλης υπακολουθίας. Η απόδειξή του επίσης θα περιληφθεί και εδώ, για λόγους πληρότητας.

Λήμμα 7.5 Η ακολουθία τ.μ. $\{X_n,\ n \geq 1\}$, συγκλίνει κατά πιθανότητα στην τ.μ. X, εάν και μόνο εάν, κάθε υπακολουθία $\{n_k,\ k = 1, 2, \ldots\}$, περιέχει μία υποϋπακολουθία $\{n_{k(i)},\ i =$

$1, 2, \ldots\}$, τέτοια ώστε

$$X_{n_{k(i)}} \xrightarrow{\mathbf{a.s.}} X, \quad \text{καθώς } i \to \infty.$$

Απόδειξη: Ας υποθέσουμε ότι $X_n \xrightarrow{\mathbf{p}} X$. Θέτουμε

$$p_{n,m} = \mathbb{P}\left(|X_n - X| > \frac{1}{m}\right), \quad n, m = 1, 2, \ldots .$$

Από την υπόθεση, $p_{n,m} \to 0$ καθώς $n \to \infty$, για κάθε σταθερό m. Έστω $\{n_k,\ k = 1, 2 \ldots\}$ αυθαίρετη υπακολουθία. Προφανώς,

$$p_{n_k,m} \to 0, \quad \text{καθώς } k \to \infty. \qquad (\text{για κάθε σταθερό } m)$$

Παρατηρούμε ότι μπορούμε να εκλέξουμε μία υποϋπακολουθία $\{n_{k(i)}\ i \geq 1\}$ της $\{n_k,\ k \geq 1\}$, τέτοια ώστε $p_{n_{k(i)},i} \leq 2^{-i}$, για $i = 1, 2, \ldots$. Πράγματι, για $i = 1$ εκλέγουμε k_0, τέτοιο ώστε $p_{n_k,1} \leq 1/2$ για $k \geq k_0$ (αφού $p_{n_k,1} \to 0$ καθώς $k \to \infty$), και θέτουμε $k(1) = k_0$. Στην συνέχεια, αν έχουμε εκλέξει $k(1) < k(2) < \cdots < k(i-1)$, και αφού $p_{n_k,i} \to 0$ καθώς $k \to \infty$, μπορούμε να βρούμε k_0, τέτοιο ώστε

$$p_{n_k,i} < \frac{1}{2^i}, \quad \text{όταν } k \geq k_0,$$

και τότε ορίζουμε $k(i) = \max\{k_0, k(i-1)+1\}$. Προφανώς, η ακολουθία $\{k(i),\ i \geq 1\}$ είναι υπακολουθία της $\{k, k = 1, 2, \ldots\}$, και επομένως,

$$\{n_{k(i)},\ i \geq 1\} \subset \{n_k,\ k \geq 1\}.$$

Από κατασκευή, $p_{n_{k(i)},i} < 2^{-i}$ για κάθε i, δηλαδή,

$$\mathbb{P}\left(|X_{n_{k(i)}} - X| > \frac{1}{i}\right) \leq \frac{1}{2^i}, \quad i = 1, 2, \ldots,$$

και επομένως,

$$\sum_{i=1}^{\infty} \mathbb{P}\left(|X_{n_{k(i)}} - X| > \frac{1}{i}\right) \leq \sum_{i=1}^{\infty} \frac{1}{2^i} = 1 < \infty.$$

Κατά συνέπεια, το πρώτο Λήμμα Borel-Cantelli συνεπάγεται την σχέση

$$\mathbb{P}\left(\limsup_i \left\{|X_{n_{k(i)}} - X| > \frac{1}{i}\right\}\right) = 0. \tag{7.1}$$

Έστω

$$C = \left\{\omega \in \Omega : X_{n_{k(i)}}(\omega) \to X(\omega) \text{ καθώς } i \to \infty\right\},$$

$$A = \left\{\omega \in \Omega : |X_{n_{k(i)}}(\omega) - X(\omega)| > \frac{1}{i} \text{ για άπειρες τιμές του } i\right\}.$$

Τότε, $A^c = \left\{\omega \in \Omega : |X_{n_{k(i)}}(\omega) - X(\omega)| \leq \frac{1}{i} \text{ τελικά για όλα τα } i\right\}$. Προφανώς, $A^c \subset C$, και $\mathbb{P}(A) = 0$, από την (7.1). Άρα, $\mathbb{P}(C) = 1$, δηλαδή,

$$X_{n_{k(i)}} \xrightarrow{\textbf{a.s.}} X, \quad \text{καθώς } i \to \infty.$$

Ας υποθέσουμε τώρα ότι $X_n \overset{\mathbf{p}}{\nrightarrow} X$. Τότε, για κάποιο $\varepsilon > 0$, υπάρχει υπακολουθία $\{n_k,\ k \geq 1\}$, τέτοια ώστε

$$\mathbb{P}\left(|X_{n_k} - X| \geq \varepsilon\right) > \varepsilon, \quad \text{για } k = 1, 2, \ldots \,. \qquad \text{(γιατί;)}$$

Είναι φανερό ότι δεν μπορεί να βρεθεί υπακολουθία $\{n_{k(i)},\ i \geq 1\}$ της $\{n_k,\ k \geq 1\}$, τέτοια ώστε $X_{n_{k(i)}} \xrightarrow{\mathbf{p}} X$, καθώς $i \to \infty$, επειδή $\mathbb{P}(|X_{n_{k(i)}} - X| \geq \varepsilon) > \varepsilon$, για κάθε i. Άρα, $X_{n_{k(i)}} \overset{\mathbf{p}}{\nrightarrow} X$, και συνεπώς, $X_{n_{k(i)}} \overset{\mathbf{a.s.}}{\nrightarrow} X$.

Αποδείξαμε λοιπόν ότι αν $X_n \overset{\mathbf{p}}{\nrightarrow} X$, τότε υπάρχει υπακολουθία $\{n_k,\ k \geq 1\}$, τέτοια ώστε για κάθε υποϋπακολουθία της $\{n_{k(i)},\ i \geq 1\}$, $X_{n_{k(i)}} \overset{\mathbf{a.s.}}{\nrightarrow} X$, καθώς $i \to \infty$. Αυτό ισοδυναμεί με το ζητούμενο. $\square$

Όπως είδαμε στα προηγούμενα, η ακολουθία τ.μ. $\{X_n,\ n \geq 1\}$, με αντίστοιχες *σ.κ.* $\{F_n,\ n \geq 1\}$, συγκλίνει προς την τ.μ. X, με *σ.κ.* F, όταν $\lim_n F_n(x) = F(x)$, για κάθε x για το οποίο $\mathbb{P}(X = x) = 0$. Ο λόγος που εξαιρούμε τα σημεία ασυνέχειας της F φαίνεται στο παρακάτω παράδειγμα, το οποίο δόθηκε και στο Κεφ. 6.

Παράδειγμα 7.6 Ας υποθέσουμε ότι οι τ.μ. $\{X_n,\ n \geq 1\}$ και η X είναι ορισμένες στον ίδιο χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, και ότι ικανοποιούν τις $\mathbb{P}(X_n = a_n) = 1$ για $n = 1, 2, \ldots$, και $\mathbb{P}(X = a) = 1$. Προφανώς,

$$F_n(x) = \begin{cases} 0, & \text{αν } x < a_n, \\ 1, & \text{αν } x \geq a_n, \end{cases} \quad F(x) = \begin{cases} 0, & \text{αν } x < a, \\ 1, & \text{αν } x \geq a. \end{cases}$$

[Οι X_n και X (αντίστοιχα, οι F_n και F) καλούνται εκφυλισμένες (degenerate) τ.μ. (αντίστοιχα, εκφυλισμένες *σ.κ.*), διότι έχουν διασπορά 0.]

Ας υποθέσουμε ότι $a_n \to a$, και ότι $a_n > a$ για άπειρες τιμές του n, και $a_n < a$ για άπειρες τιμές του n. Τότε, προφανώς, $F_n(x) \to F(x)$ για κάθε $x \neq a$, ενώ εύκολα διαπιστώνεται ότι η ακολουθία $F_n(a)$ δεν συγκλίνει. Εντούτοις, θέτοντας $A_n = \{\omega \in \Omega : X_n(\omega) \neq a_n\}$, $n = 1, 2, \ldots$, $A_0 = \{\omega \in \Omega : X(\omega) \neq a\}$, και $A = \bigcup_{n=0}^{\infty} A_n$, είναι εύκολο να διαπιστωθεί ότι $\mathbb{P}(A) = 0$, και ότι

$$\lim_n X_n(\omega) = X(\omega), \quad \text{για κάθε } \omega \in \Omega \setminus A = A^c.$$

Επομένως, $A^c \subset \{\omega \in \Omega : \lim_n X_n(\omega) = X(\omega)\}$, και συνεπώς,

$$\mathbb{P}(\lim X_n = X) \geq \mathbb{P}(A^c) = 1,$$

δηλ. έχουμε ισχυρή (σχεδόν βέβαιη) σύγκλιση. Αυτό δείχνει ότι ένας «λογικός» ορισμός «ασθενούς» σύγκλισης δεν πρέπει να απαιτεί $F_n(x) \to F(x)$, όταν το x δεν είναι σημείο συνεχείας της F. $\square$

Πρόταση 7.7 Εάν $X_n \xrightarrow{\mathbf{d}} X$, και η X είναι εκφυλισμένη, τότε $X_n \xrightarrow{\mathbf{p}} X$. [Εννοείται ότι οι X_n και X ορίζονται στον ίδιο χώρο πιθανότητας.]

Απόδειξη: Αφού η X είναι εκφυλισμένη, υπάρχει $a \in \mathbb{R}$, τέτοιο ώστε $\mathbb{P}(X = a) = 1$. Έστω F_n και F οι σ.κ. των X_n και X.

Τότε, για τυχόν $\varepsilon > 0$,

$$\begin{aligned}\mathbb{P}(|X_n - X| > \varepsilon) = \mathbb{P}(|X_n - a| > \varepsilon) &= \mathbb{P}(\{X_n > a + \varepsilon\} \cup \{X_n < a - \varepsilon\}) \\ &= \mathbb{P}(X_n > a + \varepsilon) + \mathbb{P}(X_n < a - \varepsilon) \\ &\le \mathbb{P}(X_n > a + \varepsilon) + \mathbb{P}(X_n \le a - \varepsilon) \\ &= 1 - F_n(a + \varepsilon) + F_n(a - \varepsilon) \to 0,\end{aligned}$$

διότι $F_n(a + \varepsilon) \to F(a + \varepsilon) = 1$, $F_n(a - \varepsilon) \to F(a - \varepsilon) = 0$, επειδή τα $a + \varepsilon$ και $a - \varepsilon$ είναι σημεία συνεχείας της F. Αφού η σύγκλιση γίνεται για τυχόν $\varepsilon > 0$, έπεται το ζητούμενο. $\square$

Παρατήρηση 7.8 Στην ειδική περίπτωση που η ακολουθία $\{X_n,\ n \geq 1\}$ συγκλίνει κατά πιθανότητα σε σταθερά, μπορούμε, συμβατικά, να θεωρούμε ότι η κατά πιθανότητα σύγκλιση ικανοποιείται, ακόμα και αν οι X_n ορίζονται σε διαφορετικούς χώρους πιθανότητας $(\Omega_n, \mathcal{A}_n, \mathbb{P}_n)$, λόγω της προηγούμενης πρότασης. [Αρκεί να επιλέξουμε ως X, στην Πρόταση 7.7, την σταθερή τ.μ., $X(\omega) = a$ για κάθε $\omega \in \Omega$ (οπότε η X είναι εκφυλισμένη).]

Πρόταση 7.9 (μοναδικότητα του ορίου της ασθενούς σύγκλισης) Εάν η ακολουθία τ.μ. $\{X_n,\ n \geq 1\}$ είναι τέτοια ώστε

$$X_n \xrightarrow{\mathbf{d}} X \text{ και, ταυτόχρονα, } X_n \xrightarrow{\mathbf{d}} Y,$$

τότε $\mathbb{P}(X \le x) = \mathbb{P}(Y \le x)$ για κάθε $x \in \mathbb{R}$ (δηλ. $X \stackrel{\mathbf{d}}{=} Y$), που σημαίνει ότι οι X και Y έχουν την ίδια σ.κ. – είναι, δηλαδή,

ισόνομες.

Απόδειξη: Έστω F_n, F_X, F_Y, οι σ.κ. των X_n, X, Y. Τότε ισχύουν τα εξής:

$$\begin{aligned} F_n(x) &\to F_X(x), \quad \text{για κάθε } x \text{ για το οποίο } \mathbb{P}(X=x)=0, \\ F_n(x) &\to F_Y(x), \quad \text{για κάθε } x \text{ για το οποίο } \mathbb{P}(Y=x)=0. \end{aligned}$$

Άρα, $F_X(x) = F_Y(x)$, για κάθε x για το οποίο οι F_X, F_Y, είναι συνεχείς και οι δύο. Επειδή όμως η F_X είναι αύξουσα, έχει το πολύ αριθμήσιμο πλήθος ασυνεχειών. Για τον ίδιο λόγο και η F_Y έχει το πολύ αριθμήσιμο πλήθος ασυνεχειών. Αν, λοιπόν, διέφεραν σε κάποιο $x_0 \in \mathbb{R}$, τότε θα διέφεραν και σε ολόκληρο διάστημα, της μορφής $[x_0, x_0+\varepsilon)$, για κάποιο $\varepsilon > 0$ (αφού είναι δεξιά συνεχείς). Αυτό αντίκειται στο γεγονός ότι οι F_X και F_Y διαφέρουν μόνο σε κάποιο αριθμήσιμο σύνολο, και συνεπώς, $F_X = F_Y$. $\square$

Ορισμός 7.10 Ένα σύνολο $A \in \mathcal{B}(\mathbb{R})$ λέγεται F–συνεχές σύνολο (F–continuity set), αναφορικά με την σ.κ. F της τ.μ. X, ή αναφορικά με το επαγόμενο μέτρο πιθανότητας $\mathbb{P}_X$ στον $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, εάν $\mathbb{P}(X \in \partial A) = 0$, όπου ∂A είναι το τοπολογικό σύνορο του A. [Υπενθυμίζεται ότι $\partial A = \text{cl}(A) \setminus \text{int}(A) = \{x \in \mathbb{R} : \text{για κάθε } \varepsilon > 0, \text{ υπάρχει } \alpha \in A, \text{ με } |x-\alpha| < \varepsilon, \text{ και υπάρχει } \beta \in A^c, \text{ με } |x-\beta| < \varepsilon\}$.]

Παρατήρηση 7.11 Εάν οι F_n και F είναι σ.κ., και $F_n \xrightarrow{\mathbf{d}} F$, τότε

$$F_n(x) \to F(x), \quad \text{για κάθε σημείο συνεχείας } x \text{ της } F. \tag{7.2}$$

Ισοδύναμα, αν $I_x = (-\infty, x]$, η (7.2) γράφεται ως εξής:

$$\mathbb{P}(X_n \in I_x) \to \mathbb{P}(X \in I_x), \text{ όταν το } I_x \text{ είναι } F\text{–συνεχές σύνολο}, \tag{7.3}$$

διότι $\partial I_x = \{x\}$, και η F είναι συνεχής στο x αν και μόνο αν $\mathbb{P}(X \in \partial I_x) = 0$.

Η ακριβής σχέση των F–συνεχών συνόλων και της ασθενούς σύγκλισης περιγράφεται στο παρακάτω θεώρημα.[2]

Θεώρημα 7.12 Έστω $\{X_n,\ n \geq 1\}$ μία ακολουθία τ.μ., και X μία τ.μ., με αντίστοιχες σ.κ. $\{F_n,\ n \geq 1\}$, και F. Τότε, τα επόμενα είναι ισοδύναμα:

(i) $X_n \xrightarrow{\mathbf{d}} X$ (δηλ. $F_n \xrightarrow{\mathbf{d}} F$).

(ii) $\mathbb{E}[g(X_n)] \to \mathbb{E}[g(X)]$ για κάθε συνεχή και φραγμένη $g : \mathbb{R} \longrightarrow \mathbb{R}$.

(iii) $\mathbb{P}(X_n \in A) \to \mathbb{P}(X \in A)$ για κάθε F–συνεχές σύνολο $A \in \mathcal{B}(\mathbb{R})$.

Αυτοί οι χαρακτηρισμοί της ασθενούς σύγκλισης χρησιμοποιούνται κυρίως σε γενικεύσεις όπως, για παράδειγμα, όταν

[2] Για την απόδειξη του Θεωρήματος 7.12 βλ. Billingsley (1986), σελ. 344.

ορίζεται η έννοια της ασθενούς σύγκλισης ακολουθίας στοχαστικών ανελίξεων προς κάποια στοχαστική ανέλιξη, κ.λπ.

7.2 Το Θεώρημα του Skorohod

Το επόμενο θεώρημα που θα αποδείξουμε οφείλεται στον Skorohod. Η σπουδαιότητα αυτού του αποτελέσματος έγκειται στο γεγονός ότι μας επιτρέπει να αναγάγουμε/μετασχηματίσουμε την ασθενή σύγκλιση (που γίνεται σε αυθαίρετους χώρους πιθανότητας) σε ισχυρή (και μάλιστα, σε κατά σημείο) σύγκλιση, διατηρώντας την στοχαστική δομή της υπό μελέτη ακολουθίας τ.μ., καθώς και της οριακής τ.μ., στην οποία η εν λόγω ακολουθία συγκλίνει. Κατ΄ ουσίαν, η ιδέα έγκειται στην εμφύτευση των υπό μελέτη τ.μ. σε κατάλληλο χώρο πιθανότητας, στον οποίον μπορούμε να τις ελέγξουμε καλύτερα. Αυτή η εμφύτευση έχει ως άμεση συνέπεια την επέκταση όλων των γνωστών θεωρημάτων σύγκλισης (Μονότονης Σύγκλισης, Κυριαρχημένης Σύγκλισης, και Ομοιόμορφης Ολοκληρωσιμότητας), στα οποία, η μόνη σύγκλιση που απαιτείται είναι η ασθενής (κατά κατανομή) σύγκλιση. Για την απόδειξή του Θεωρήματος του Skorohod θα χρησιμοποιήσουμε το εξής λήμμα, το οποίο είναι και ανεξαρτήτου ενδιαφέροντος (βλ. Θεώρημα ;;).

Λήμμα 7.13 Έστω X τυχούσα τ.μ. με $\sigma.\kappa.$ F. Τότε η τ.μ. Y, που ορίζεται στον χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, με $\Omega = (0,1)$,

$\mathcal{A} = \mathcal{B}((0,1))$ (τα Borel υποσύνολα του $(0,1)$), και $\mathbb{P} = \lambda$ (το μέτρο Lebesgue περιορισμένο στο διάστημα $(0,1)$), ως

$$Y(\omega) = \inf\{t : F(t) \geq \omega\}, \quad \text{για } 0 < \omega < 1,$$

έχει την ίδια σ.κ. F.

Απόδειξη: Έστω $I_\omega = \{t : F(t) \geq \omega\}$, για κάποιο (σταθερό) $\omega \in (0,1)$. Επειδή $F(t) \to 1$ για $t \to +\infty$, και $F(t) \to 0$ για $t \to -\infty$, έπεται ότι το I_ω είναι μη κενό, και γνήσιο υποσύνολο του $\mathbb{R}$. Επίσης, το I_ω είναι διάστημα, διότι αν $t_1 \in I_\omega$ και $t_1 \leq t_2$, έπεται ότι και $t_2 \in I_\omega$. Επομένως, $I_\omega = [b, +\infty)$ ή $I_\omega = (b, +\infty)$, για κάποιο $b \in \mathbb{R}$. Τότε, $b + 1/n \in I_\omega$ για κάθε $n = 1, 2, \ldots$, δηλαδή

$$F\left(b + \frac{1}{n}\right) \geq \omega. \tag{7.4}$$

Η (7.4) δείχνει ότι $F(b) = \lim_n F(b + 1/n) \geq \omega$, και συνεπώς, $b \in I_\omega$. Άρα, $I_\omega = [b, +\infty)$, όπου το άκρο $b = b(\omega) = Y(\omega)$ χαρακτηρίζεται από την ιδιότητα

$$F(b) \geq \omega, \quad \text{και, ταυτόχρονα,} \quad F(b - \varepsilon) < \omega \text{ για κάθε } \varepsilon > 0.$$

Τώρα προφανώς,

$$Y(\omega) \leq x \Rightarrow b \leq x \Rightarrow F(b) \leq F(x) \Rightarrow \omega \leq F(x). \tag{7.5}$$

Όμως και αντίστροφα,

$$\omega \leq F(x) \Rightarrow x \in I_\omega \Rightarrow x \geq b = Y(\omega). \tag{7.6}$$

Από τις (7.5) και (7.6) έπεται ότι για κάθε $x \in \mathbb{R}$, και για κάθε $\omega \in (0,1)$,

$$Y(\omega) \le x \Leftrightarrow \omega \le F(x), \tag{7.7}$$

και συνεπώς, $\{\omega \in (0,1) : Y(\omega) \le x\} = \{\omega \in (0,1) : \omega \le F(x)\}$. Άρα, για κάθε $x \in \mathbb{R}$,

$$\mathbb{P}[Y(\omega) \le x] = \lambda(\{\omega : \omega \le F(x)\}) = \lambda((0, F(x)] \setminus \{1\}) = \lambda((0, F(x)]) = F(x),$$

δηλαδή το ζητούμενο. $\square$

Παρατήρηση 7.14 Προφανώς η $Y(\omega)$ του προηγουμένου λήμματος είναι αύξουσα και αριστερά συνεχής. (γιατί;) Μάλιστα, η συνάρτηση Y μπορεί να θεωρηθεί ως «γενικευμένη αντίστροφη» της σ.κ. F, και γι΄ αυτό συμβολίζεται πολλές φορές με F^{-1}. Ας σημειωθεί ότι το προηγούμενο λήμμα έχει και την ακόλουθη ισοδύναμη «στατιστική» διατύπωση:

Αν U είναι μία τ.μ. με ομοιόμορφη κατανομή στο διάστημα $(0,1)$ [αυτό συνήθως συμβολίζεται με $U \sim U(0,1)$], *τότε για τυχούσα τ.μ. X με σ.κ. F, η τ.μ. $F^{-1}(U)$, όπου $F^{-1}(u) = \inf\{x : F(x) \ge u\}$, $0 < u < 1$, έχει επίσης σ.κ. F, δηλ.*

$$X \stackrel{\mathbf{d}}{=} F^{-1}(U). \tag{7.8}$$

Η σχέση (7.8) περιγράφει, εν συντομία, μία βασική ιδιότητα στην οποία στηρίζεται μεγάλο μέρος της μη παραμετρικής στατιστικής, και καλείται **αντίστροφος μετασχηματισμός πιθανότητας**.

Θεώρημα 7.15 (Skorohod) Θεωρούμε τις τ.μ. $\{X_n,\ n \geq 1\}$ και X, ορισμένες σε αυθαίρετους χώρους πιθανότητας, και τέτοιες ώστε $X_n \xrightarrow{\mathbf{d}} X$. Τότε, υπάρχει χώρος πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, και τ.μ. $\{Y_n,\ n \geq 1\}$ και Y, με τις εξής ιδιότητες:

(i) Οι $\{Y_n,\ n \geq 1\}$ και Y είναι ορισμένες **στον (ίδιο)** χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$.

(ii) $Y_n \stackrel{\mathbf{d}}{=} X_n$, $n \geq 1$, και $Y \stackrel{\mathbf{d}}{=} X$.

(iii) Για κάθε $\omega \in \Omega$, $Y_n(\omega) \to Y(\omega)$. [Δηλ., $Y_n \to Y$ κατά σημείο.]

Μάλιστα, η κατασκευή αυτή μπορεί να γίνει στον χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P}) = ((0,1), \mathcal{B}((0,1)), \lambda)$, όπου λ το μέτρο Lebesgue, περιορισμένο στα Borel υποσύνολα του $(0,1)$. Επίσης, οι $Y_n(\omega)$ και $Y(\omega)$ μπορούν να επιλεγούν έτσι ώστε

$$Y_n(\omega) = \inf\{t : F_n(t) \geq \omega\}, \ \text{και}\ Y(\omega) = \inf\{t : F(t) \geq \omega\}, \ \omega \in (0,1) \setminus N,$$

και $Y_n(\omega) = Y(\omega) = 0$, αν $\omega \in N$, όπου F_n και F οι σ.κ. των X_n και X, αντίστοιχα, και το $N \subset (0,1)$ είναι αριθμήσιμο. [Οπότε το N είναι Borel, με $\lambda(N) = \mathbb{P}(N) = 0$.]

Απόδειξη: Ο χώρος $(\Omega, \mathcal{A}, \mathbb{P})$ είναι αυτός του Λήμματος 7.13. Έστω F_n και F οι σ.κ. των X_n και X, αντίστοιχα. Αρχικά ορίζουμε τις τ.μ. Y_n και Y, ως

$$Y(\omega) = \inf\{t : F(t) \geq \omega\}, \quad Y_n(\omega) = \inf\{t : F_n(t) \geq \omega\}, \quad n = 1, 2, \ldots,$$

για κάθε $\omega \in (0,1)$. Από το Λήμμα 7.13 έχουμε αμέσως ότι οι Y_n και Y έχουν σ.κ. F_n και F, οπότε $X_n \stackrel{\mathbf{d}}{=} Y_n$ και $X \stackrel{\mathbf{d}}{=} Y$. Από την υπόθεση,

$$F_n(x) \to F(x), \quad \text{σε κάθε σημείο συνεχείας } x \text{ της } F.$$

Έστω $\varepsilon > 0$. Σταθεροποιούμε ένα $\omega \in (0,1)$, και εκλέγουμε ένα σημείο συνεχείας x της F, τέτοιο ώστε

$$Y(\omega) - \varepsilon < x < Y(\omega).$$

Λόγω της (7.7) θα είναι τότε $\omega > F(x)$, και επειδή το x είναι σημείο συνεχείας της F, $F_n(x) \to F(x) < \omega$. Άρα, $F_n(x) < \omega$, τελικά για κάθε n, και λόγω της (7.7) και πάλι (εφαρμοζόμενη στην σ.κ. F_n),

$$Y_n(\omega) > x > Y(\omega) - \varepsilon.$$

Επομένως, $\liminf_n Y_n(\omega) \geq Y(\omega)$.

Για τυχόν ω', με $\omega < \omega' < 1$, εκλέγουμε τώρα ένα σημείο συνεχείας y της F, τέτοιο ώστε

$$Y(\omega') \leq y < Y(\omega') + \varepsilon.$$

Λόγω της (7.7) θα είναι $\omega' \leq F(y)$, και άρα, $\omega < F(y)$. Επειδή $F_n(y) \to F(y) > \omega$, έπεται ότι $F_n(y) \geq \omega$, τελικά για κάθε n, και λόγω της (7.7), $Y_n(\omega) \leq y < Y(\omega') + \varepsilon$. Συνεπώς, $\limsup_n Y_n(\omega) \leq Y(\omega')$.

Συνδυάζοντας τα παραπάνω βλέπουμε ότι για κάθε ω, ω', με $0 < \omega < \omega' < 1$, ισχύει ότι

$$Y(\omega) \leq \liminf_n Y_n(\omega) \leq \limsup_n Y_n(\omega) \leq Y(\omega').$$

Επομένως, όταν η Y είναι συνεχής στο ω, προκύπτει ότι (παίρνοντας όρια για $\omega' \searrow \omega$)

$$Y_n(\omega) \to Y(\omega).$$

Επειδή η Y είναι αύξουσα, τα σημεία $\{\omega :$ η Y δεν είναι συνεχής στο $\omega\}$ δημιουργούν ένα το πολύ αριθμήσιμο σύνολο N, και συνεπώς, μέτρου Lebesgue 0. Άρα,

$$\lim_n Y_n(\omega) = Y(\omega), \tag{7.9}$$

εκτός ίσως ενός αριθμήσιμου συνόλου N. Αλλάζοντας τις Y_n και Y (αν χρειάζεται), έτσι ώστε $Y_n(\omega) = Y(\omega) \equiv 0$ για κάθε $\omega \in N$, η (7.9) ικανοποιείται για κάθε $\omega \in (0,1)$. Επιπλέον, οι συναρτήσεις κατανομής των Y_n και Y δεν επηρεάζονται από αυτήν την αλλαγή, διότι, π.χ., ισχύει ότι για κάθε $x \in \mathbb{R}$,

$$\begin{aligned}\mathbb{P}(Y \leq x) &= \lambda(\{\omega \in (0,1) : Y(\omega) \leq x\}) = \lambda(\{Y \leq x\}) \\ &= \lambda(\{Y \leq x\} \cup N) = \lambda(\{Y \leq x\} \setminus N),\end{aligned}$$

αφού $\lambda(N) = 0$. $\square$

7.3 Εφαρμογές του Θεωρήματος Skorohod σε άλλα Κλασικά Θεωρήματα

Για να παρουσιάσουμε την τροποποιημένη γενική μορφή των κλασικών θεωρημάτων Μονότονης και Κυριαρχημένης Σύγκλισης, καθώς και του Θεωρήματος Ομοιόμορφης Ολοκληρωσιμότητας, θα χρειαστούμε την έννοια της **στοχαστικής διάταξης**, η οποία μας επιτρέπει να διατάξουμε τ.μ. που δεν ορίζονται στον ίδιο χώρο πιθανότητας.

Ορισμός 7.16 (στοχαστική διάταξη) Λέμε ότι η τ.μ. X_1 είναι **στοχαστικά μικρότερη** από την X_2 (συμβολισμός: $X_1 \leq_{\mathbf{st}} X_2$, ή $X_2 \geq_{\mathbf{st}} X_1$), όταν οι αντίστοιχες *σ.κ.* F_{X_1} και F_{X_2} ικανοποιούν την (αντίστροφη) ανισότητα:

$$\text{Για κάθε } x \in \mathbb{R}, \;\; F_{X_1}(x) \geq F_{X_2}(x).$$

[Δηλ. όταν ισχύει ότι $\mathbb{P}(X_1 \geq x) \leq \mathbb{P}(X_2 \geq x)$ για κάθε $x \in \mathbb{R}$.]

Είναι σαφές ότι οι δύο σχέσεις $X_1 \leq_{\mathbf{st}} X_2$ και $X_2 \leq_{\mathbf{st}} X_1$ ισοδυναμούν με την $X_1 \stackrel{\mathbf{d}}{=} X_2$, και ότι, γενικότερα, η σχέση $\leq_{\mathbf{st}}$ ορίζει μία σχέση μερικής διάταξης στον χώρο των συναρτήσεων κατανομής και, εξ' επαγωγής, στον χώρο των τ.μ. (στον οποίον, δύο ισόνομες τ.μ. θεωρούνται «ίσες» – αυτό σημαίνει ότι οι τ.μ. που ανήκουν στην ίδια κλάση ισοδυναμίας «ταυτοποιούνται», σύμφωνα με τη σχέση ισοδυναμίας $X_1 \sim X_2 \Leftrightarrow X_1 \stackrel{\mathbf{d}}{=} X_2$). Προφανώς, χρησιμοποιώντας την έννοια της στοχαστικής διάταξης, μπορούμε να εκφράσουμε ανισοτικές σχέσεις **με μόνη γνώση** των αντιστοίχων *σ.κ.*

Δεδομένου ότι η στοχαστική διάταξη επινοήθηκε για να περιγράψει κάποια «γενικευμένη έννοια διατάξεως», οφείλει, προφανώς, να επαληθεύεται από την **συνήθη διάταξη** (αυτήν που ό-λοι γνωρίζουμε).[3] Αυτό είναι, πράγματι, αληθές. Γενικότερα, αν οι X_1 και X_2 είναι ορισμένες στον ίδιο χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, και αν ισχύει η σχέση $X_1 \le X_2$ με πιθ. 1, τότε μπορεί να δειχθεί ότι ισχύει και η σχέση $X_1 \le_{\mathbf{st}} X_2$. Πράγματι, θεωρούμε το ενδεχόμενο $A = \{X_1 > X_2\} \in \mathcal{A}$, το οποίο έχει πιθανότητα 0, από υπόθεση. Τότε είναι σαφές ότι $\mathbb{P}(B) = \mathbb{P}(B \smallsetminus A)$, για κάθε ενδεχόμενο $B \in \mathcal{A}$. Από τον ορισμό του A, προκύπτει ότι για κάθε $x \in \mathbb{R}$, $\{X_1 \le x\} \smallsetminus A \supset \{X_2 \le x\} \smallsetminus A$, και έτσι,

$$\begin{aligned} F_{X_1}(x) &= \mathbb{P}(X_1 \le x) = \mathbb{P}(\{X_1 \le x\} \smallsetminus A) \\ &\ge \mathbb{P}(\{X_2 \le x\} \smallsetminus A) = \mathbb{P}(X_2 \le x) = F_{X_2}(x), \end{aligned}$$

δηλαδή ισχύει ότι και $X_1 \le_{\mathbf{st}} X_2$. [Φυσικά, αν κάτι τέτοιο δεν ίσχυε, η έννοια της στοχαστικής διάταξης θα ήταν θεωρητικού, μόνο, ενδιαφέροντος.] Το σημαντικό είναι ότι ισχύει μία μορφή **αντιστρόφου**, και συγκεκριμένα έχουμε την εξής απλή πρόταση.

Πρόταση 7.17 Εάν $X_1 \le_{\mathbf{st}} X_2$ (και οι X_1, X_2 είναι ορισμένες σε αυθαίρετους χώρους πιθανότητας), τότε οι αντίστοιχες εμφυτευμένες τ.μ. Y_1 και Y_2 του Θεωρήματος 7.15 του Skorohod, στον χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P}) = ((0,1), \mathcal{B}((0,1)), \lambda)$, ικανοποιούν την $Y_1 \le Y_2$, με πιθ. 1.

[3] Δηλ., πρέπει να ικανοποιείται η συνεπαγωγή $X_1 \le X_2 \Rightarrow X_1 \le_{\mathbf{st}} X_2$.

Απόδειξη: Έστω $F_1 \geq F_2$ οι σ.κ. των X_1, X_2. Είναι σαφές ότι $Y_1(\omega) = \inf\{t : F_1(t) \geq \omega\}$ με πιθ. 1, και ομοίως, $Y_2(\omega) = \inf\{t : F_2(t) \geq \omega\}$ με πιθ. 1. Είναι προφανές ότι το διάστημα, του οποίου λαμβάνουμε το infimum για τον ορισμό της Y_1, περιέχει (είναι μεγαλύτερο από) το αντίστοιχο διάστημα που δίδεται στον ορισμό της Y_2. $\square$

Παρατήρηση 7.18 Είναι προφανές ότι μπορούμε να ορίσουμε στοχαστική διάταξη και για εκφυλισμένες τ.μ. Για παράδειγμα, $X \geq_{\mathbf{st}} 0$ σημαίνει, απλά, ότι $F_X(x) = 0$ για $x < 0$, και δεν δίδει πληροφορία για $x \geq 0$. Πράγματι, η σ.κ. της εκφυλισμένης τ.μ. Y, με $\mathbb{P}(Y = 0) = 1$, είναι η $F_Y(x) = I(x \geq 0)$, $x \in \mathbb{R}$.

Τώρα μπορούμε να δείξουμε εύκολα την εξής πρόταση.

Πρόταση 7.19 (i) Αν $X_1 \leq_{\mathbf{st}} X_2$, και η μέση της X_2 ορίζεται, με $\mathbb{E}(X_2) \in [-\infty, +\infty)$, τότε ορίζεται και η μέση τιμή της X_1, και μάλιστα, $\mathbb{E}(X_1) \leq \mathbb{E}(X_2)$. Ομοίως, αν $X_1 \leq_{\mathbf{st}} X_2$, και η μέση της X_1 ορίζεται, με $\mathbb{E}(X_1) \in (-\infty, +\infty]$, τότε ορίζεται και η μέση τιμή της X_2, και μάλιστα, $\mathbb{E}(X_1) \leq \mathbb{E}(X_2)$.
(ii) Αν $0 \leq_{\mathbf{st}} X_1 \leq_{\mathbf{st}} X_2 \leq_{\mathbf{st}} \cdots$ και $X_n \xrightarrow{\mathbf{d}} X$, τότε $X_n \leq_{\mathbf{st}} X$ για κάθε $n \geq 1$.

Απόδειξη: Τα πάντα προκύπτουν άμεσα από την Πρόταση 7.17 και το Θεώρημα του Skorohod, επειδή οι αντίστοιχες εμφυτευμένες τ.μ. $Y_1,\ Y_2$, για το (i), και $Y, Y_1, Y_2, \ldots$, για το (ii), ικανοποιούν τις $Y_1 \le Y_2$ (με πιθ. 1), και $0 \le Y_1 \le Y_2 \le \cdots \nearrow Y$, με πιθ. 1, και συνεπώς, $Y_n \le Y$ με πιθ. 1. □

Το Θεώρημα Μονότονης Σύγκλισης του Lebesgue, με την βοήθεια του Θεωρήματος 7.15 του Skorohod, λαμβάνει την εξής γενικευμένη μορφή.

Θεώρημα 7.20 (Μονότονης Σύγκλισης) Αν $0 \le_{\mathbf{st}} X_1 \le_{\mathbf{st}} X_2 \le_{\mathbf{st}} \cdots$ και $X_n \xrightarrow{\mathbf{d}} X$, τότε

$$\mathbb{E}(X_n) \to \mathbb{E}(X).$$

Απόδειξη: Θεωρούμε τις εμφυτευμένες τ.μ. Y και $Y_1, Y_2, \ldots$, όπως στην απόδειξη του Θεωρήματος 7.15 του Skorohod. Από το Θεώρημα 7.15 και την Πρόταση 7.19(ii) έχουμε ότι $0 \le Y_1 \le Y_2 \le \cdots \nearrow Y$, με πιθ. 1, οπότε προκύπτει ότι $\mathbb{E}(Y_n) \to \mathbb{E}(Y)$, από το κλασικό Θεώρημα Μονότονης Σύγκλισης (βλ. Θεώρημα ;;, και σχέση (M4) της §5.6 του Κεφ. ;;). Όμως, $X \stackrel{\mathbf{d}}{=} Y$, και $X_n \stackrel{\mathbf{d}}{=} Y_n$ για κάθε n. Άρα, $\mathbb{E}(X) = \mathbb{E}(Y)$, και $\mathbb{E}(X_n) = \mathbb{E}(Y_n)$ για κάθε n, και έτσι, $\mathbb{E}(X_n) = \mathbb{E}(Y_n) \to \mathbb{E}(Y) = \mathbb{E}(X)$. □

Με τον ίδιο ακριβώς τρόπο αποδεικνύεται ένα γενικευμένο Θεώρημα Κυριαρχημένης Σύγκλισης, και ένα γενικευμένο Θεώρημα Ομοιόμορφης Ολοκληρωσιμότητας. Αναφέρουμε απλώς τις εκφωνήσεις τους, αφού η απόδειξή τους γίνεται προφανής, εμφυτεύοντας τις αντίστοιχες τ.μ. σύμφωνα με το Θεώρημα 7.15 και την Πρόταση 7.17.

Θεώρημα 7.21 (Κυριαρχημένης Σύγκλισης) Έστω ότι $X_n \xrightarrow{\mathbf{d}} X$, και ας υποθέσουμε ότι για κάθε n, $|X_n| \leq_{\mathbf{st}} |Y|$ (δηλ. $\mathbb{P}(|X_n| \leq x) \geq \mathbb{P}(|Y| \geq x)$ για κάθε $x \in \mathbb{R}$),[4] όπου η Y είναι κάποια (οποιαδήποτε) ολοκληρώσιμη τ.μ. (δηλ. $\mathbb{E}|Y| < \infty$). Τότε, $\mathbb{E}|X| < \infty$ (και $\mathbb{E}|X_n| < \infty$ για κάθε n), και μάλιστα,

$$\mathbb{E}(X_n) \to \mathbb{E}(X).$$

Στο σημείο αυτό υπενθυμίζεται η έννοια της ομοιόμορφης ολοκληρωσιμότητας (uniform integrability), η οποία έχει δοθεί στο Κεφ. ;;, για τ.μ. που ορίζονται στον ίδιο χώρο πιθανότητας (Ορισμός ;;). Εδώ δίδεται στην πιο γενική της μορφή.

Ορισμός 7.22 (ομοιόμορφη ολοκληρωσιμότητα) Η ακολουθία τ.μ. $\{X_n,\ n \geq 1\}$ (όπου κάθε τυχαία μεταβλητή X_n επιτρέπεται να ορίζεται σε διαφορετικό χώρο πιθανότητας $(\Omega_n, \mathcal{A}_n, \mathbb{P}_n)$) καλείται ομοιόμορφα ολοκληρώσιμη όταν

$$\sup_n \mathbb{E}[|X_n| I(|X_n| \geq \alpha)] \to 0, \quad \text{καθώς } \alpha \to +\infty. \qquad (7.10)$$

Ισοδύναμα χρησιμοποιείται και η διατύπωση «οι X_n είναι ομοιόμορφα ολοκληρώσιμες».

[4] Ακριβέστερα, θα έπρεπε να γράφαμε $\mathbb{P}_n(|X_n| \leq x) \geq \mathbb{P}((|Y| \leq x))$, αφού οι χώροι πιθανότητας επιτρέπεται να μεταβάλλονται με το n, αλλά δεν δημιουργείται σύγχυση από αυτήν την ασυμβατότητα του συμβολισμού. Μάλιστα, ο ίδιος «απλοποιημένος» συμβολισμός θα χρησιμοποιηθεί και στην Πρόταση 7.24(ii), παρακάτω.

Το βασικό θεώρημα που διέπει τις ακολουθίες οι οποίες είναι ομοιόμορφα ολοκληρώσιμες είναι το εξής (πρβλ. Θεώρημα ;; του Κεφ. ;;).

Θεώρημα 7.23 (Ομοιόμορφης Ολοκληρωσιμότητας) Αν η ακολουθία τυχαίων μεταβλητών $\{X_n,\ n \geq 1\}$ είναι ομοιόμορφα ολοκληρώσιμη, και αν $X_n \xrightarrow{\mathbf{d}} X$, τότε

$$\mathbb{E}|X| < \infty, \quad \text{και} \quad \mathbb{E}(X_n) \to \mathbb{E}(X).$$

Επειδή η (7.10) είναι μερικές φορές δύσκολο να δειχθεί απευθείας, είναι χρήσιμες οι παρακάτω δύο απλούστερες ικανές συνθήκες.

Πρόταση 7.24 Εάν οποιαδήποτε από τις παρακάτω συνθήκες ικανοποιείται, τότε η ακολουθία $\{X_n,\ n \geq 1\}$ είναι ομοιόμορφα ολοκληρώσιμη.

(i) Για κάποιο $\varepsilon > 0$, $\sup\limits_n \mathbb{E}|X_n|^{1+\varepsilon} < \infty$.

(ii) Υπάρχει μία τ.μ. Y, με $\mathbb{E}|Y| < \infty$, τέτοια ώστε[5]

$$\mathbb{P}(|X_n| \leq x) \geq \mathbb{P}(|Y| \leq x), \ \text{για}\, n = 1, 2, \ldots, \text{ και για κάθε } x \geq 0.$$

Απόδειξη:(i) Έστω $p = 1 + \varepsilon$, $q = (1+\varepsilon)/\varepsilon$ (οπότε $p \geq 1$, $q \geq 1$, $1/p + 1/q = 1$). Από τις ανισότητες Hölder και Markov έχουμε ότι για κάθε $r > 0$, και $\alpha > 0$,

[5]Παρατηρήστε ότι αυτές είναι ακριβώς οι υποθέσεις του Θεωρήματος Κυριαρχημένης Σύγκλισης, Θεώρημα 7.21.

$$
\begin{aligned}
\mathbb{E}[|X_n|I(|X_n|\geq\alpha)] &\leq (\mathbb{E}|X_n|^p)^{1/p}\,(\mathbb{E}[I(|X_n|\geq\alpha)]^q)^{1/q}\\
&= \left(\mathbb{E}|X_n|^{1+\varepsilon}\right)^{1/(1+\varepsilon)}(\mathbb{E}[I(|X_n|\geq\alpha)])^{1/q}\\
&= \left(\mathbb{E}|X_n|^{1+\varepsilon}\right)^{1/(1+\varepsilon)}(\mathbb{P}[|X_n|\geq\alpha])^{1/q}\\
&\leq \left(\mathbb{E}|X_n|^{1+\varepsilon}\right)^{1/(1+\varepsilon)}\left(\frac{\mathbb{E}|X_n|^r}{\alpha^r}\right)^{1/q}.
\end{aligned}
$$

Εκλέγοντας $r=1+\varepsilon$, παίρνουμε την ανισότητα[6]

$$\mathbb{E}[|X_n|I(|X_n|\geq\alpha)]\leq\frac{1}{\alpha^{\varepsilon}}\mathbb{E}|X_n|^{1+\varepsilon}. \tag{7.11}$$

Επομένως, αν $\sup\limits_n \mathbb{E}|X_n|^{1+\varepsilon}=c<\infty$, τότε

$$\sup_n \mathbb{E}[|X_n|I(|X_n|\geq\alpha)]\leq\frac{c}{\alpha^{\varepsilon}}\to 0,\quad \text{καθώς } \alpha\to+\infty.$$

(ii) Κατ΄ αρχήν σημειώνουμε ότι ισχύει η έκφραση

$$\mathbb{E}(T)=\int_0^{+\infty}\mathbb{P}(T>t)dt, \tag{7.12}$$

για οποιαδήποτε, μη αρνητική, τ.μ. T, ακόμη και αν η T δεν είναι ολοκληρώσιμη (βλ. Άσκηση ;;.8, ή Άσκηση 7.9). Έστω $\alpha>0$. Είναι σαφές ότι η τ.μ. $T=|Y|I(|Y|\geq\alpha)$ είναι μη αρνητική, και ότι για κάθε $t\geq 0$, $\mathbb{P}(T>t)=\min\{\mathbb{P}(|Y|\geq\alpha),\mathbb{P}(|Y|>t)\}$, δηλ., για κάθε t, με $0\leq t<\alpha$, ισχύει η ισότητα $\mathbb{P}(T>t)=\mathbb{P}(|Y|\geq\alpha)$, ενώ για $t\geq\alpha$ ισχύει η $\mathbb{P}(T>t)=\mathbb{P}(|Y|>t)$. Θέτοντας $g(t)=\mathbb{P}(T>t)$, $t\geq 0$, και ολοκληρώνοντας την g στο $[0,+\infty)$, προκύπτει από την (7.12) η έκφραση

[6] Η (7.11) προκύπτει και απευθείας, αν παρατηρήσουμε ότι $\alpha^{\varepsilon}|X_n|I(|X_n|\geq\alpha)\leq|X_n|^{1+\varepsilon}$.

$$\mathbb{E}[|Y|I(|Y| \geq \alpha)] = \int_{\alpha}^{+\infty} \mathbb{P}(|Y| > t)dt + \alpha \mathbb{P}(|Y| \geq \alpha),$$

και, για τον ίδιο λόγο,

$$\mathbb{E}[|X_n|I(|X_n| \geq \alpha)] = \int_{\alpha}^{+\infty} \mathbb{P}(|X_n| > t)dt + \alpha \mathbb{P}(|X_n| \geq \alpha).$$

Παρατηρούμε ότι για κάθε t, $\mathbb{P}(|X_n| > t) \leq \mathbb{P}(|Y| > t)$, από υπόθεση, οπότε ισχύει και η ανισότητα[7] $\mathbb{P}(|X_n| \geq \alpha) \leq \mathbb{P}(|Y| \geq \alpha)$. Επομένως, $\mathbb{E}[|X_n|I(|X_n| \geq \alpha)] \leq \mathbb{E}[|Y|I(|Y| \geq \alpha)]$, και συνεπώς,

$$\sup_n \mathbb{E}[|X_n|I(|X_n| \geq \alpha)] \leq \mathbb{E}[|Y|I(|Y| \geq \alpha)].$$

Αρκεί λοιπόν να αποδείξουμε ότι $\mathbb{E}[|Y|I(|Y| \geq \alpha)] \to 0$, καθώς $\alpha \to +\infty$. Για τον σκοπό αυτό, ας θεωρήσουμε τις μη αρνητικές τ.μ. $Y_n = |Y|I(|Y| \geq n)$, οι οποίες ορίζονται στον ίδιο χώρο πιθανότητας, έστω $(\Omega, \mathcal{A}, \mathbb{P})$, στον οποίο είναι ορισμένη και η τ.μ. Y. Προφανώς, $|Y| = Y_0 \geq Y_1 \geq Y_2 \geq \cdots \geq 0$. Αφού[8] $Y_n \xrightarrow{\textbf{a.s.}} 0$, $|Y_n| \leq |Y|$, και $\mathbb{E}|Y| < \infty$, από το Θεώρημα Κυριαρχημένης Σύγκλισης του Lebesgue έπεται ότι

$$\mathbb{E}(Y_n) \to \mathbb{E}(0) = 0.$$

Συνεπώς, αν $[\alpha]$ παριστάνει το ακέραιο μέρος του α,

$$\mathbb{E}[|Y|I(|Y| \geq \alpha)] \leq \mathbb{E}(Y_{[\alpha]}) \to 0, \quad \text{καθώς } \alpha \to +\infty. \quad \square$$

---

[7]Η ανισότητα αυτή μπορεί να δειχθεί ως εξής: $\mathbb{P}(|X_n| \geq \alpha) = \mathbb{P}\big(\lim_n \{|X_n| > \alpha - \frac{1}{n}\}\big) = \lim_n \mathbb{P}\big(|X_n| > \alpha - \frac{1}{n}\big) \leq \lim_n \mathbb{P}\big(|Y| > \alpha - \frac{1}{n}\big) = \mathbb{P}\big(\lim_n \{|Y| > \alpha - \frac{1}{n}\}\big) = \mathbb{P}(|Y| \geq \alpha)$.

[8]κατ' ακρίβειαν, ισχύει ότι $Y_n(\omega) \to 0$ για κάθε $\omega \in \Omega$ για το οποίο $|Y(\omega)| < \infty$, αλλά το ενδεχόμενο $\{|Y| = \infty\}$ έχει πιθανότητα 0, επειδή η Y υποτέθηκε ολοκληρώσιμη

7.4 Νόμος Μεγάλων Αριθμών

Διάφοροι Νόμοι Μεγάλων αριθμών αποδείχθηκαν στο Κεφ. 6. Συνοψίζοντας, υπενθυμίζονται τα εξής σημαντικά αποτελέσματα, στα οποία θα χρησιμοποιηθεί ο συνήθης συμβολισμός,

$$\overline{X}_n \stackrel{\text{ορ.}}{=\!=} \frac{X_1 + \ldots + X_n}{n}, \tag{7.13}$$

για τον **δειγματικό μέσο** των τ.μ. $X_1, \ldots, X_n$, αλλά και ο συμβολισμός

$$\overline{\mu}_n \stackrel{\text{ορ.}}{=\!=} \frac{\mu_1 + \ldots + \mu_n}{n}, \tag{7.14}$$

για τον **αριθμητικό μέσο** των αριθμών $\mu_1, \ldots, \mu_n$.

Θεώρημα 7.25 (ασθενής νόμος μεγάλων αριθμών) Έστω $\{X_n,\ n \geq 1\}$ μία ακολουθία[9] τ.μ., με $\mathbb{E}(X_n^2) < \infty$, $\mathbb{E}(X_n) = \mu_n$, $\mathrm{Var}(X_n) = \sigma_n^2$, και $\mathrm{Cov}(X_i, X_j) \leq 0$ για $i \neq j$, και ας υποθέσουμε ότι ικανοποιείται η συνθήκη

$$\frac{1}{n^2}\sum_{j=1}^{n} \sigma_j^2 \to 0, \quad \text{καθώς } n \to \infty. \tag{7.15}$$

Τότε ισχύει ότι

$$\overline{X}_n - \overline{\mu}_n \xrightarrow{\mathbf{p}} 0,$$

όπου τα $\overline{X}_n$ και $\overline{\mu}_n$ δίδονται από τις (7.13) και (7.14), αντίστοιχα. Αν, επιπροσθέτως, $\overline{\mu}_n \to \mu \in \mathbb{R}$, καθώς $n \to \infty$, τότε

[9] όχι κατ' ανάγκην ανεξάρτητη

ισχύει και η

$$\overline{X}_n \xrightarrow{\mathbf{p}} \mu.$$

Απόδειξη: Δουλεύουμε ακριβώς όπως και στην απόδειξη του Θεωρήματος 6.7, χρησιμοποιώντας την ανισότητα Chebychev, και παρατηρώντας ότι $\mathrm{Var}(X_1 + \ldots + X_n) \leq \sum_{j=1}^n \sigma_j^2$. $\square$

Θεώρημα 7.26 (ισχυρός νόμος για ανεξάρτητη ακολουθία) Έστω $\{X_n,\ n \geq 1\}$ μία ανεξάρτητη ακολουθία τ.μ., με $\mathbb{E}(X_n^2) < \infty$, $\mathbb{E}(X_n) = \mu_n$, και $\mathrm{Var}(X_n) = \sigma_n^2$. Αν

$$\sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2} < \infty, \tag{7.16}$$

τότε

$$\overline{X}_n - \overline{\mu}_n \xrightarrow{\mathbf{a.s.}} 0,$$

όπου τα $\overline{X}_n$ και $\overline{\mu}_n$ δίδονται από τις (7.13) και (7.14), αντίστοιχα. Αν, επιπροσθέτως, $\overline{\mu}_n \to \mu \in \mathbb{R}$, καθώς $n \to \infty$, τότε

$$\overline{X}_n \xrightarrow{\mathbf{a.s.}} \mu.$$

Θεώρημα 7.27 (ισχυρός νόμος για ανεξάρτητη και ισόνομη ακολουθία) Αν η $\{X_n,\ n \geq 1\}$ είναι μία ανεξάρτητη και ισόνομη ακολουθία τ.μ., και η $\mathbb{E}(X_1)$ ορίζεται, έστω $\mathbb{E}(X_1) = \mu \in [-\infty, +\infty]$, τότε[10]

$$\overline{X}_n = \frac{X_1 + \ldots + X_n}{n} \xrightarrow{\mathbf{a.s.}} \mu.$$

Θεώρημα 7.28 (αντίστροφο του ισχυρού νόμου) Έστω $\{X_n,\ n \geq 1\}$ μία ανεξάρτητη και ισόνομη ακολουθία τ.μ., για την οποία ισχύει ότι

$$\overline{X}_n = \frac{X_1 + \ldots + X_n}{n} \xrightarrow{\mathbf{a.s.}} \mu,$$

όπου $\mu \in \mathbb{R}$ κάποια σταθερά. Τότε $\mathbb{E}|X_1| < \infty$, και $\mathbb{E}(X_1) = \mu$.

Ασκήσεις Κεφ. 7:

7.1. Βρείτε παραδείγματα ακολουθιών τ.μ. $\{X_n,\ n \geq 1\}$, και τ.μ. X, σε κάποιον χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, με τις εξής ιδιότητες:

(i) $X_n \xrightarrow{\mathbf{p}} X$ και $X_n \overset{\mathbf{a.s.}}{\nrightarrow} X$.

(ii) Για δοθέν $p \geq 1$, $X_n \xrightarrow{\mathbf{p}} X$ και $X_n \overset{\boldsymbol{L}^{\mathbf{p}}}{\nrightarrow} X$.

(iii) Για δοθέντα r και p, με $1 \leq r < p$, $X_n \xrightarrow{\boldsymbol{L}^{\mathbf{r}}} X$ και $X_n \overset{\boldsymbol{L}^{\mathbf{p}}}{\nrightarrow} X$.

(iv) $X_n \xrightarrow{\mathbf{d}} X$ και $X_n \overset{\mathbf{p}}{\nrightarrow} X$.

---

[10] Στην περίπτωση που $-\infty < \mu < +\infty$, δηλ. όταν η X_1 είναι ολοκληρώσιμη, είναι αρκετό οι X_n να είναι ισόνομες και **ανά ζεύγη ανεξάρτητες** – βλ. Θεώρημα 6.24, **ισχυρός νόμος του Etemadi**.

7.2. **(απόσταση ολικής κύμανσης (total variation distance))** Αν οι $\{X_n,\ n \geq 1\}$ και X έχουν πυκνότητες f_n και f (αναφορικά με το μέτρο Lebesgue), λέμε ότι η X_n συγκλίνει κατά ολική κύμανση προς την X (συμβολισμός $X_n \xrightarrow{\mathbf{t.v.}} X$) όταν

$$\int_{-\infty}^{+\infty} |f_n(x) - f(x)| dx \to 0.$$

(i) Αποδείξτε ότι αν $X_n \xrightarrow{\mathbf{t.v.}} X$ τότε $X_n \xrightarrow{\mathbf{d}} X$.

(ii) Δώστε παράδειγμα ακολουθίας τ.μ. $\{X_n,\ n \geq 1\}$, και τ.μ. X, με πυκνότητες f_n και f, τέτοιες ώστε $X_n \xrightarrow{\mathbf{d}} X$ και $X_n \overset{\mathbf{t.v.}}{\nrightarrow} X$.

[**Υπόδειξη:** $\int_{-\infty}^{+\infty} |f_n(x) - f(x)| dx = 2 \sup_{A \in \mathcal{B}(\mathbb{R})} |\mathbb{P}(X_n \in A) - \mathbb{P}(X \in A)|$.]

7.3. **(Λήμμα Scheffé)** Αν οι f_n και f είναι πυκνότητες των X_n και X (αναφορικά με το μέτρο Lebesgue), και $f_n \to f$, εκτός ίσως ενός συνόλου N, μέτρου Lebesgue 0, τότε

$$\int_{-\infty}^{+\infty} |f_n(x) - f(x)| dx \to 0.$$

Ισχύει και το αντίστροφο;

7.4. Βρείτε ομοιόμορφα ολοκληρώσιμες τ.μ. X_n, για τις οποίες **δεν υπάρχει** τ.μ. Y, με $\mathbb{E}|Y| < \infty$, τέτοια ώστε για κάθε $t \geq 0$, να ισχύει

$$\mathbb{P}(|X_n| \leq t) \geq \mathbb{P}(|Y| \leq t), \quad \text{για κάθε } n = 1, 2, \ldots .$$

7.5. **(Θεώρημα Pólya)** Αν $F_n \xrightarrow{\mathbf{d}} F$, και η F είναι συνεχής σ.κ., τότε

$$\sup_x |F_n(x) - F(x)| \to 0. \tag{7.17}$$

Βρείτε παράδειγμα σ.κ. F_n και F, για τις οποίες $F_n \xrightarrow{\mathbf{d}} F$ και $\sup_x |F_n(x) - F(x)| \nrightarrow 0$.
[**Υπόδειξη:** Για κάθε $\varepsilon > 0$, υπάρχουν πεπερασμένα το πλήθος σημεία $x_0 = -\infty < x_1 < \cdots < x_n < x_{n+1} = +\infty$, τέτοια ώστε $F(x_{k+1}) - F(x_k) < \varepsilon$, $k = 0, 1, \ldots, n$.]

7.6. **(μετρική Lévy (Lévy distance))** Για δύο σ.κ. F και G, ορίζουμε

$$\mathrm{d}(F,G) = \inf\{\varepsilon \ge 0 : \text{για κάθε } x \in \mathbb{R} \text{ ισχύει ότι } G(x-\varepsilon)-\varepsilon \le F(x) \le G(x+\varepsilon)+\varepsilon\}.$$

Αποδείξτε ότι η d είναι μετρική στον χώρο των σ.κ. $\mathcal{F} = \{F :$ η F είναι συνάρτηση κατανομής$\}$. Δείξτε ότι

$$F_n \xrightarrow{\mathbf{d}} F \iff \mathrm{d}(F_n, F) \to 0.$$

7.7. Αποδείξτε ότι η $W(\omega) = \sup\{x : F(x) \le \omega\}$, $0 < \omega < 1$ (όπου F τυχούσα σ.κ.), μπορεί να χρησιμοποιηθεί στο Λήμμα 7.13, αντί της $Y(\omega)$. Τι διαφορά έχει η Y από την W;

7.8. Αποδείξτε ότι για μία τ.μ. X με σ.κ. F,

$$\mathbb{E}|X| < \infty \text{ αν και μόνο αν } \int_0^1 |F^{-1}(t)|dt < \infty.$$

Επίσης, για αυθαίρετη Borel συνάρτηση g, με $\mathbb{E}|g(X)| < \infty$,

$$\mathbb{E}[g(X)] = \int_0^1 g(F^{-1}(t))dt,$$

όπου στους παραπάνω τύπους μπορούμε να χρησιμοποιήσουμε, ως F^{-1}, τόσο την Y του Λήμματος 7.13, όσο και την W της προηγούμενης άσκησης.

7.9. Δείξτε ότι για τυχούσα τ.μ. X, με σ.κ. F, και πεπερασμένη πρώτη ροπή,

$$\mathbb{E}(X) = \int_0^{+\infty} (1 - F(x))dx - \int_{-\infty}^{0} F(x)dx. \qquad (7.18)$$

Με χρήση της (7.18) βρείτε ανάλογο τύπο για την $\mathbb{E}(X^2)$ (υποθέτοντας ότι είναι πεπερασμένη), και αποδείξτε ότι

$$\mathrm{Var}(X) = 2 \iint_{-\infty<x<y<+\infty} F(x)(1 - F(y))dydx.$$

7.10. Αποδείξτε ότι αν $X_n \xrightarrow{\mathbf{p}} X$ και, ταυτόχρονα, $X_n \xrightarrow{\mathbf{p}} Y$, τότε $\mathbb{P}(X = Y) = 1$. [Δηλ., $X \overset{\mathbf{a.s.}}{=} Y$.]

7.11. Για μία τ.μ. X (ή, για την αντίστοιχη σ.κ. F), οποιοσδήποτε αριθμός $m = m(X)$ $(= m(F))$ ικανοποιεί **ταυτόχρονα** και τις δύο σχέσεις $\mathbb{P}(X \le m) \ge 1/2$, και $\mathbb{P}(X \ge m) \ge 1/2$, ονομάζεται **διάμεσος** της X.

(i) Δείξτε ότι το σύνολο $M = \{m : m$ διάμεσος της $X\}$ είναι μη κενό.

(ii) Βρείτε την μορφή του συνόλου M.

(iii) Αποδείξτε ότι εάν $X_n \xrightarrow{\mathbf{d}} 0$ τότε $m(X_n) \to 0$, όπου $m(X_n)$ είναι οποιοσδήποτε αριθμός του συνόλου $M_n = \{m : m$ διάμεσος της $X_n\}$, δηλαδή $\lim M_n = \{0\}$.

7.12. Μία τ.μ. X (ή η αντίστοιχη σ.κ. F) λέγεται **συμμετρική**, αν οι X και $-X$ έχουν την ίδια σ.κ. Δείξτε ότι η X είναι συμμετρική αν και μόνο αν οι X^+ και X^- έχουν την ίδια σ.κ.

7.13. Αν ο Ω είναι αριθμήσιμος, τότε $X_n \xrightarrow{\mathbf{a.s.}} X$ αν και μόνο αν $X_n \xrightarrow{\mathbf{p}} X$.

7.14. Έστω F μία σ.κ. και

$$F_\alpha(x) = \frac{1}{2\alpha} \int_{x-\alpha}^{x+\alpha} F(t)dt, \quad x \in \mathbb{R},\ \alpha > 0.$$

Δείξτε ότι η F_α είναι συνεχής σ.κ., και εξετάστε εάν μπορούμε να συμπεράνουμε ότι $F_{1/n} \xrightarrow{\mathbf{d}} F$.

7.15. Για τυχούσα συνάρτηση κατανομής F θέτουμε $F^{-1}(t) = \inf\{x : F(x) \geq t\}$, $0 < t < 1$, όπως στο Λήμμα 7.13. Αποδείξτε ότι $F_n \xrightarrow{\mathbf{d}} F$ αν και μόνο αν το σύνολο $\{t \in (0,1) :$

$F_n^{-1}(t) \nrightarrow F^{-1}(t)\}$ είναι μέτρου Lebesgue 0, δηλαδή

$$F_n \xrightarrow{\mathbf{d}} F \quad \text{αν και μόνο αν} \quad F_n^{-1}(t) \to F^{-1}(t) \text{ σχεδόν για κάθε } t \in (0,1).$$

[**Υπόδειξη:** Θεώρημα Skorohod.]