

## Κεφάλαιο 8

# Στοιχεία θεωρίας πιθανοτήτων

Η θεωρία πιθανοτήτων αναπτύσσεται σε ένα χώρο μέτρου  $(X, \mathcal{M}, \mu)$  όπου  $\mu(X) = 1$ . Ένας τέτοιος χώρος μέτρου ονομάζεται χώρος μέτρου πιθανότητας και το μέτρο  $\mu$  μέτρο πιθανότητας. Παρόλα αυτά δεν είναι σωστό να συμπεράνει κανείς ότι η θεωρία πιθανοτήτων είναι υποκλάδος της θεωρίας μέτρου, αφού οι κεντρικές ιδέες και οι τεχνικές είναι ξεχωριστές και ιδιαίτερες από αυτές της θεωρίας μέτρου.

### 8.1 Βασικές έννοιες

Η θεωρία πιθανοτήτων έχει τη δική της ορολογία η οποία αναπτύχθηκε πριν τη θεωρία μέτρου. Για αυτό θα χρησιμοποιούμε τη δική της ορολογία.

Θεωρία Μέτρου	Θεωρία πιθανοτήτων
Χώρος μέτρου με ολικό μέτρο 1. Συμβολισμός: $(X, \mathcal{M}, \mu)$	Δειγματικός χώρος ή χώρος πιθανότητας. Συμβολισμός: $(\Omega, \mathcal{B}, \mathbb{P})$
$(\sigma)$ -άλγεβρα	$(\sigma)$ -δακτύλιος
Μετρήσιμο σύνολο	Ενδεχόμενο
Μετρήσιμη πραγματική συνάρτηση. Τυπικό σύμβολο: $f$	Τυχαία μεταβλητή. Τυπικό σύμβολο $X$
Ολοκλήρωμα της $f$ . Σύμβολο: $\int f d\mu$	Μέση τιμή, μαθηματική ελπίδα της $X$ . Σύμβολο $\mathbb{E}(X)$
$L_p$ -συνάρτηση	Τυχαία μεταβλητή με πεπερασμένη $p$ -ροπή (moment)
Σύγκλιση ως προς το μέτρο	Σύγκλιση ως προς πιθανότητα
Σχεδόν παντού (σπ)	Σχεδόν σίγουρα (σσ)
Borel μέτρο πιθανότητας στο $\mathbb{R}$	Κατανομή
Μετασχηματισμός Fourier (μέτρου στο $\mathbb{R}$ )	Χαρακτηριστική συνάρτηση (κατανομής)
Χαρακτηριστική συνάρτηση	Δείκτρια συνάρτηση

Στη θεωρία πιθανοτήτων συνήθως δεν γράφουμε τα ορίσματα των συναρτήσεων (που τώρα λέγονται τυχαίες μεταβλητές). Έτσι όταν γράφουμε  $\{X > a\}$  εννοούμε το ενδεχόμενο (δηλαδή το μετρήσιμο σύνολο)  $\{\omega \in \Omega : X(\omega) > a\}$ . Έτσι το  $\mathbb{P}\{X > a\}$  είναι η πιθανότητα η τυχαία μεταβλητή  $X$  να είναι μεγαλύτερη από  $a$ , δηλαδή το μέτρο του συνόλου  $\{\omega \in \Omega : X(\omega) > a\}$ . Γενικώς θα χρησιμοποιούμε τη γλώσσα των πιθανοτήτων εκτός από το ότι θα προτιμάμε να λέμε ότι μια τυχαία μεταβλητή είναι  $L_p$ .

**Ορισμός 8.1.1** Αν η  $X$  είναι  $L_2$  τυχαία μεταβλητή, τότε είναι και  $L_1$  τυχαία μεταβλητή και η μέση τιμή της  $\mathbb{E}(X)$ , δηλαδή η ποσότητα

$$\int_{\Omega} X(\omega) d\mathbb{P}(\omega),$$

είναι καλά ορισμένη. Η  $L_2$  νόρμα της τυχαίας μεταβλητής  $X - \mathbb{E}(X)$  ονομάζεται τυπική απόκλιση της  $X$  (δηλαδή τυπική απόκλιση της  $X$  από τη μέση τιμή της), και το τετράγωνο αυτής της νόρμας ονομάζεται διακύμανση της  $X$ . Γράφουμε

$$\sigma(X) = (\mathbb{E}|X - \mathbb{E}(X)|^2)^{1/2}$$

και

$$\text{Var}(X) = \sigma^2(X) = \mathbb{E}|X - \mathbb{E}(X)|^2 = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

αντίστοιχα.

Θα χρειαστούμε την ακόλουθη κατασκευή. Αν έχουμε ένα χώρο πιθανότητας  $(\Omega, \mathcal{B}, \mathbb{P})$ , ένα μετρήσιμο χώρο  $(\Omega', \mathcal{B}')$  και μια απεικόνιση  $\varphi : \Omega \rightarrow \Omega'$  τότε ορίζεται μια πιθανότητα  $\mathbb{P}_{\varphi}$  στον χώρο  $(\Omega', \mathcal{B}')$ , το «push forward-μέτρο», θέτοντας

$$\mathbb{P}_{\varphi}(E) = \mathbb{P}(\varphi^{-1}(E)).$$

Το ότι το  $\mathbb{P}_{\varphi}$  είναι πράγματι μέτρο πιθανότητας προκύπτει άμεσα από το ότι η  $\varphi^{-1}$  αντιμετατίθεται με ενώσεις και τομές.

**Πρόταση 8.1.2** Με τον παραπάνω συμβολισμό, αν η  $f : \Omega' \rightarrow \mathbb{R}$  είναι μια μετρήσιμη συνάρτηση τότε  $\int_{\Omega'} f d\mathbb{P}_{\varphi} = \int_{\Omega} (f \circ \varphi) d\mathbb{P}$  εφόσον και τα δύο ολοκληρώματα υπάρχουν.

*Απόδειξη:* Αν η  $f$  είναι η χαρακτηριστική συνάρτηση  $\chi_E$  για  $E \in \mathcal{B}'$  το ζητούμενο είναι απλώς ο ορισμός του  $\mathbb{P}_{\varphi}$ , αφού  $\chi_E \circ \varphi = \chi_{\varphi^{-1}(E)}$ . Για το γενικό αποτέλεσμα χρησιμοποιούμε γραμμικούς συνδυασμούς και παίρνουμε όρια.  $\square$

Για κάθε τυχαία μεταβλητή  $X$  στον  $\Omega$  το μέτρο  $\mathbb{P}_X$  είναι ένα Borel μέτρο πιθανότητας το οποίο ονομάζεται κατανομή της  $X$ , και η συνάρτηση

$$F(t) = \mathbb{P}_X((-\infty, t]) = \mathbb{P}(X \leq t) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \leq t\})$$

(η οποία καθορίζει το μέτρο πιθανότητας  $\mathbb{P}_X$  σύμφωνα με το Θεώρημα 2.2.2) ονομάζεται συνάρτηση κατανομής της  $X$ .

Γενικότερα, αν  $X_j : \Omega \rightarrow \mathbb{R}$  για  $j = 1, 2, \dots, n$  ορίζουμε το τυχαίο διάνυσμα  $X = (X_1, X_2, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n$  και το μέτρο  $\mathbb{P}_X$  στον  $\mathbb{R}^n$  ονομάζεται από κοινού κατανομή των  $X_1, X_2, \dots, X_n$ . Η κατανομή ή η από κοινού κατανομή για περισσότερες από μία τυχαίες μεταβλητές είναι χρήσιμη γιατί με τη βοήθειά της αποδεικνύονται διάφορες ιδιότητες. Έτσι αν  $X, Y : \Omega \rightarrow \mathbb{R}$  τυχαίες μεταβλητές, από την Πρόταση 8.1.2,

- αν  $\varphi = X : \Omega \rightarrow \mathbb{R} =: \Omega'$  και  $f(t) = t : \Omega' = \mathbb{R} \rightarrow \mathbb{R}$ , τότε

$$\mathbb{E}(X) = \int_{\Omega} X d\mathbb{P}(\omega) = \int_{\mathbb{R}} t d\mathbb{P}_X(t).$$

- αν  $\varphi = X : \Omega \rightarrow \mathbb{R} =: \Omega'$  και  $f(t) = t - \mathbb{E}(X) : \Omega' = \mathbb{R} \rightarrow \mathbb{R}$ , τότε

$$\text{Var}(X) = \int_{\Omega} |X - \mathbb{E}(X)|^2 d\mathbb{P}(\omega) = \int_{\mathbb{R}} |t - \mathbb{E}(X)|^2 d\mathbb{P}_X(t).$$

- αν  $\varphi = (X, Y) : \Omega \rightarrow \mathbb{R}^2 =: \Omega'$  και  $f(t) = t + s : \Omega' = \mathbb{R}^2 \rightarrow \mathbb{R}$ , τότε

$$\mathbb{E}(X + Y) = \int_{\Omega} (X + Y) d\mathbb{P}(\omega) = \int_{\mathbb{R}^2} (t + s) d\mathbb{P}_{(X,Y)}(t, s).$$

Δηλαδή η γνώση της από κοινού κατανομής μάς επιτρέπει τον υπολογισμό των παραμέτρων των τυχαίων μεταβλητών ως ολοκληρώματα σε πραγματικούς χώρους.

**Ορισμός 8.1.3** Λέμε ότι μια οικογένεια τυχαίων μεταβλητών  $\{X_a\}$  είναι *ισοκατανεμημένη* αν όλες οι  $X_a$  έχουν την ίδια κατανομή, δηλαδή αν ισχύει  $\mathbb{P}_{X_a} = \mathbb{P}_{X_b}$  για κάθε  $a, b$ .

Στη Θεωρία Πιθανοτήτων μια από τις σημαντικότερες έννοιες είναι αυτή της (στοχαστικής) ανεξαρτησίας. Ας υποθέσουμε ότι το ενδεχόμενο  $E$  στον χώρο πιθανότητας  $(\Omega, \mathcal{B}, \mathbb{P})$  έχει θετική πιθανότητα:  $\mathbb{P}(E) > 0$ . Εύκολα βλέπει κανείς ότι η ποσότητα  $\mathbb{P}_E(F) = \mathbb{P}(E \cap F)/\mathbb{P}(E)$  ορίζει ένα μέτρο πιθανότητας το οποίο ονομάζεται *πιθανότητα δοθέντος του  $E$*  ή *πιθανότητα υπό συνθήκη* (τη συνθήκη της βεβαιότητας του  $E$ ). Πρόκειται δηλαδή για την πιθανότητα να συμβεί το ενδεχόμενο  $F$  δοθέντος ότι το ενδεχόμενο  $E$  συμβαίνει ήδη. Αν αυτή η πιθανότητα είναι ίση με την  $\mathbb{P}(F)$  τότε λέμε ότι το ενδεχόμενο  $F$  είναι *ανεξάρτητο του  $E$* . Ένας άλλος τρόπος να περιγράψουμε την ανεξαρτησία του  $F$  από το  $E$  είναι να πούμε ότι «το ποσοστό του μέτρου που το  $F$  καταλαμβάνει μέσα στο  $E$  είναι ίσο με το ποσοστό του μέτρου που καταλαμβάνει το  $F$  μέσα στον  $\Omega$  (που έχει μέτρο ίσο με 1)». Αυτή η συνθήκη είναι ισοδύναμη με το να ισχύει

$$\mathbb{P}(E \cap F) = \mathbb{P}(E)\mathbb{P}(F).$$

Η σχέση αυτή είναι συμμετρική ως προς  $E$  και  $F$ , και όταν είναι αληθής συνεπάγεται ότι και το  $E$  είναι ανεξάρτητο του  $F$ . Επίσης έχει νόημα ακόμα και όταν  $\mathbb{P}(E) = 0$ .

Γενικότερα, μια συλλογή ενδεχομένων  $(E_a)_{a \in A}$  λέγεται *ανεξάρτητη* αν

$$\mathbb{P}(E_{a_1} \cap \dots \cap E_{a_n}) = \prod_{j=1}^n \mathbb{P}(E_{a_j})$$

για οποιαδήποτε διακεκριμένα  $a_1, \dots, a_n \in A$  και για κάθε  $n \in \mathbb{N}$ . Παρατηρούμε εδώ ότι για να είναι τα  $(E_a)_{a \in A}$  ανεξάρτητα δεν αρκεί να είναι ανά δύο ανεξάρτητα (δείτε Άσκηση ι). Ορίζουμε επίσης ότι μια οικογένεια τυχαίων μεταβλητών  $(X_a)_{a \in A}$  στον χώρο πιθανότητας  $(\Omega, \mathcal{B}, \mathbb{P})$  λέγεται *ανεξάρτητη* αν τα ενδεχόμενα  $\{X_a \in B_a\} = X_a^{-1}(B_a)$  είναι ανεξάρτητα για οποιαδήποτε Borel υποσύνολα  $B_a$

του  $\mathbb{R}$ . Αυτή τη συνθήκη μπορούμε να τη γράψουμε και ως εξής: γράφοντας  $X_j$  για την  $X_{a_j}$  οι  $(X_a)_{a \in A}$  είναι ανεξάρτητες αν

$$(8.1) \quad \prod_{j=1}^n \mathbb{P}(X_j^{-1}(B_j)) = \mathbb{P}(X_1^{-1}(B_1) \cap \dots \cap X_n^{-1}(B_n)).$$

Παρατηρούμε τώρα ότι  $\omega \in \bigcap_j X_j^{-1}(B_j)$  αν και μόνο αν  $X_j(\omega) \in B_j$  για κάθε  $j = 1, \dots, n$ , αν και μόνο αν  $(X_1(\omega), \dots, X_n(\omega)) \in B_1 \times \dots \times B_n$ , δηλαδή  $\omega \in (X_1, \dots, X_n)^{-1}(B_1 \times \dots \times B_n)$ . Άρα η (8.1) γράφεται

$$(8.2) \quad \prod_{j=1}^n \mathbb{P}(X_j^{-1}(B_j)) = \mathbb{P}((X_1, \dots, X_n)^{-1}(B_1 \times \dots \times B_n)) \\ = \mathbb{P}_{(X_1, \dots, X_n)}(B_1 \times \dots \times B_n).$$

Αλλά

$$\prod_{j=1}^n \mathbb{P}(X_j^{-1}(B_j)) = \prod_{j=1}^n \mathbb{P}_{X_j}(B_j) = \left( \prod_{j=1}^n \mathbb{P}_{X_j} \right) (B_1 \times \dots \times B_n).$$

Αφού τα  $B_j$  είναι οποιαδήποτε σύνολα Borel, αυτές οι ποσότητες είναι ίσες αν και μόνο αν

$$\mathbb{P}_{(X_1, \dots, X_n)} = \prod_{j=1}^n \mathbb{P}_{X_j}.$$

Δηλαδή, οι τυχαίες μεταβλητές  $(X_a)_{a \in A}$  είναι ανεξάρτητες αν και μόνο αν η από κοινού κατανομή οποιουδήποτε πεπερασμένου πλήθους από αυτές ισούται με το γινόμενο των κατανομών τους.

**Πρόταση 8.1.4** Έστω ότι οι  $\{X_{nj} : 1 \leq j \leq J(n), 1 \leq n \leq N\}$  είναι ανεξάρτητες τυχαίες μεταβλητές και έστω ότι οι  $f_n : \mathbb{R}^{J(n)} \rightarrow \mathbb{R}$  είναι Borel μετρήσιμες συναρτήσεις για  $1 \leq n \leq N$ . Τότε οι τυχαίες μεταβλητές  $Y_n = f_n(X_{n1}, X_{n2}, \dots, X_{nJ(n)})$  για  $1 \leq n \leq N$  είναι ανεξάρτητες.

*Απόδειξη:* Για  $n = 1, \dots, N$  θέτουμε  $X_n = (X_{n1}, X_{n2}, \dots, X_{nJ(n)})$ . Αν  $B_1, \dots, B_N$  Borel υποσύνολα του  $\mathbb{R}$  πρέπει να δείξουμε ότι τα ενδεχόμενα

$$Y_n^{-1}(B_n) = (f_n \circ X_n)^{-1}(B_n) = X_n^{-1}(f_n^{-1}(B_n))$$

είναι ανεξάρτητα. Όπως και (8.1) και (8.2) έχουμε ότι

$$(Y_1, \dots, Y_N)^{-1}(B_1 \times \dots \times B_N) = \bigcap_{n=1}^N Y_n^{-1}(B_n) \\ = \bigcap_{n=1}^N X_n^{-1}(f_n^{-1}(B_n)) \\ = (X_1, \dots, X_N)^{-1}(f_1^{-1}(B_1) \times \dots \times f_N^{-1}(B_N)).$$

Οπότε

$$\mathbb{P}_{(Y_1, \dots, Y_N)}(B_1 \times \dots \times B_N) = \mathbb{P}((Y_1, \dots, Y_N)^{-1}(B_1 \times \dots \times B_N)) \\ = \mathbb{P}((X_1, \dots, X_N)^{-1}(f_1^{-1}(B_1) \times \dots \times f_N^{-1}(B_N))) \\ = \mathbb{P}_{(X_{11}, \dots, X_{1J(1)}, X_{21}, \dots, X_{2J(2)}, \dots, X_{N1}, \dots, X_{NJ(N)})}(f_1^{-1}(B_1) \times \dots \times f_N^{-1}(B_N)),$$

και η τελευταία λόγω της ανεξαρτησίας των  $X_{nj}$  ισούται με

$$\begin{aligned} &= \left( \prod_{n=1}^N \prod_{j=1}^{J(n)} \mathbb{P}_{X_{nj}} \right) (f_1^{-1}(B_1) \times \dots \times f_N^{-1}(B_N)) \\ &= \int \chi_{f_1^{-1}(B_1) \times \dots \times f_N^{-1}(B_N)}(x_{11}, \dots, x_{1J(1)}, \dots, x_{N1}, \dots, x_{NJ(N)}) d \left( \prod_{n=1}^N \prod_{j=1}^{J(n)} \mathbb{P}_{X_{nj}} \right) \\ &= \int \chi_{f_1^{-1}(B_1)}(x_{11}, \dots, x_{1J(1)}) \cdots \chi_{f_N^{-1}(B_N)}(x_{N1}, \dots, x_{NJ(N)}) d \left( \prod_{n=1}^N \prod_{j=1}^{J(n)} \mathbb{P}_{X_{nj}} \right). \end{aligned}$$

Από το Θεώρημα Fubini η τελευταία παράσταση είναι ίση με

$$\begin{aligned} \prod_{n=1}^N \int \chi_{f_n^{-1}(B_n)}(x_{n1}, \dots, x_{nJ(n)}) d \left( \prod_{j=1}^{J(n)} \mathbb{P}_{X_{nj}} \right) &= \prod_{n=1}^N \left( \prod_{j=1}^{J(n)} \mathbb{P}_{X_{nj}}(f_n^{-1}(B_n)) \right) \\ &= \prod_{n=1}^N \mathbb{P}_{X_n}(f_n^{-1}(B_n)) = \prod_{n=1}^N \mathbb{P}(X_n^{-1}(f_n^{-1}(B_n))) \\ &= \prod_{n=1}^N \mathbb{P}(Y_n^{-1}(B_n)). \end{aligned}$$

Δηλαδή  $\mathbb{P}_{(Y_1, \dots, Y_N)}(B_1 \times \dots \times B_N) = \prod_{n=1}^N \mathbb{P}_{Y_n}(B_n)$ , άρα οι  $Y_n$  είναι ανεξάρτητες.  $\square$

**Ορισμός 8.1.5 (συνέλιξη μέτρων)** Αν τα  $\mu_j, j = 1, 2, \dots, n$  είναι μέτρα Borel στο  $\mathbb{R}$  ορίζουμε το μέτρο συνέλιξη  $\mu_1 * \mu_2 * \dots * \mu_n$  να δίνεται από τον τύπο

$$\mu_1 * \dots * \mu_n(E) = \int \dots \int \chi_E(t_1 + \dots + t_n) d\mu_1(t_1) \dots d\mu_n(t_n).$$

**Πρόταση 8.1.6** Αν οι  $X_1, \dots, X_n$  είναι ανεξάρτητες τυχαίες μεταβλητές, τότε

$$\mathbb{P}_{X_1 + \dots + X_n} = \mathbb{P}_{X_1} * \dots * \mathbb{P}_{X_n}.$$

*Απόδειξη:* Θέτουμε  $T(t_1, \dots, t_n) = \sum_{j=1}^n t_j$ . Φανερά  $X_1 + \dots + X_n = T \circ (X_1, \dots, X_n)$ . Άρα

$$\begin{aligned} \mathbb{P}_{X_1 + \dots + X_n}(B) &= \mathbb{P}_{T \circ (X_1, \dots, X_n)}(B) = \mathbb{P}((T \circ (X_1, \dots, X_n))^{-1}(B)) \\ &= \mathbb{P}((X_1, \dots, X_n)^{-1}(T^{-1}(B))) = \mathbb{P}_{(X_1, \dots, X_n)}(T^{-1}(B)) \\ &= \prod_{j=1}^n \mathbb{P}_{X_j}(T^{-1}(B)) \\ &= \int \dots \int \chi_{T^{-1}(B)}(t_1, \dots, t_n) d\mathbb{P}_{X_1} \dots d\mathbb{P}_{X_n} \end{aligned}$$

και επειδή  $(t_1, \dots, t_n) \in T^{-1}(B)$  αν και μόνο αν  $\sum_{j=1}^n t_j \in B$

$$\begin{aligned} &= \int \dots \int \chi_B(t_1 + \dots + t_n) d\mathbb{P}_{X_1} \dots d\mathbb{P}_{X_n} \\ &= \mathbb{P}_{X_1} * \dots * \mathbb{P}_{X_n}. \end{aligned}$$

$\square$

**Πρόταση 8.1.7** Έστω ότι οι  $X_1, \dots, X_n$  είναι ανεξάρτητες τυχαίες μεταβλητές. Αν  $X_j \in L_1$  για κάθε  $j$ , τότε  $\prod_{j=1}^n X_j \in L_1$  και

$$\mathbb{E} \left( \prod_{j=1}^n X_j \right) = \prod_{j=1}^n \mathbb{E}(X_j).$$

Απόδειξη: Για  $f(t_1, \dots, t_n) = \prod_{j=1}^n |t_j|$  ισχύει  $\prod_{j=1}^n |X_j| = f(X_1, \dots, X_n)$ . Άρα από την Πρόταση 8.1.2 ισχύει

$$\int f d\mathbb{P}_{(X_1, \dots, X_n)} = \int f \circ (X_1, \dots, X_n) d\mathbb{P} = \int \prod_{j=1}^n |X_j| d\mathbb{P} = \mathbb{E} \left( \prod_{j=1}^n |X_j| \right).$$

Άρα χρησιμοποιώντας πρώτα την ανεξαρτησία των  $X_j$  και στη συνέχεια το Θεώρημα Fubini παίρνουμε

$$\begin{aligned} \mathbb{E} \left( \prod_{j=1}^n |X_j| \right) &= \int f d\mathbb{P}_{(X_1, \dots, X_n)} \\ &= \int f d \left( \prod_{j=1}^n \mathbb{P}_{X_j} \right) = \prod_{j=1}^n \int |t_j| d\mathbb{P}_{X_j}(t_j) \\ &= \prod_{j=1}^n \mathbb{E}|X_j|. \end{aligned}$$

Συνεπώς  $\prod_{j=1}^n |X_j| \in L^1$  και άρα  $\prod_{j=1}^n X_j \in L^1$ . Το ίδιο επιχείρημα χωρίς τις απόλυτες τιμές δίνει το ζητούμενο.  $\square$

**Πόρισμα 8.1.8** Έστω ότι οι  $X_1, \dots, X_n$  είναι ανεξάρτητες τυχαίες μεταβλητές. Αν  $X_j \in L_2$  για κάθε  $j$ , τότε  $\sigma^2(X_1 + \dots + X_n) = \sum_{j=1}^n \sigma^2(X_j)$ .

Απόδειξη: Θέτοντας  $Y_j = X_j - \mathbb{E}X_j$  αυτές είναι ανεξάρτητες από την Πρόταση 8.1.4. Έτσι, επειδή  $\mathbb{E}Y_j = 0$  συμπεραίνουμε ότι  $\mathbb{E}(Y_i Y_j) = \mathbb{E}Y_i \mathbb{E}Y_j = 0$  για κάθε  $i \neq j$ . Άρα

$$\begin{aligned} \sigma^2(X_1 + \dots + X_n) &= \mathbb{E}(|(X_1 + \dots + X_n) - \mathbb{E}(X_1 + \dots + X_n)|^2) \\ &= \mathbb{E}(((X_1 - \mathbb{E}X_1) + \dots + (X_n - \mathbb{E}X_n))^2) \\ &= \mathbb{E}((Y_1 + \dots + Y_n)^2) = \sum_{j=1}^n \mathbb{E}Y_j^2 + 2 \sum_{j < i} \mathbb{E}(Y_i Y_j) = \sum_{j=1}^n \mathbb{E}Y_j^2 \\ &= \sum_{j=1}^n \sigma^2(X_j). \end{aligned}$$

$\square$

Η ανεξαρτησία είναι μια πολύ απαιτητική ιδιότητα. Για παράδειγμα, αν οι  $X$  και  $Y$  είναι ανεξάρτητες τυχαίες μεταβλητές στο  $\mathbb{R}$ , με  $\mathbb{E}(X) = 0$  και  $f$  μια Borel μετρήσιμη πραγματική συνάρτηση ώστε η  $f \circ Y$  να είναι ολοκληρώσιμη, τότε από την Πρόταση 8.1.4 οι  $X$  και  $f \circ Y$  είναι ανεξάρτητες, οπότε

$$\mathbb{E}(X(f \circ Y)) = \mathbb{E}(X)\mathbb{E}(f \circ Y) = 0.$$

Δηλαδή η  $X$  είναι ορθογώνια σε κάθε συνάρτηση της  $Y$ . Αυτός είναι και ο λόγος που είναι δύσκολο να κατασκευάσει κανείς μια ακολουθία ανεξάρτητων τυχαίων μεταβλητών ακόμα και στο  $[0, 1]$  με το μέτρο Lebesgue χρησιμοποιώντας οικείες συναρτήσεις από τον απειροστικό λογισμό.

**Παράδειγμα 8.1.9 (συναρτήσεις Rademacher)** Για κάθε  $n \in \mathbb{N}$  ορίζουμε τη  $n$ -στη συνάρτηση Rademacher  $r_n(x) = \text{sign} \sin(\pi 2^n x)$ . Για παράδειγμα,

$$r_1(x) = \begin{cases} 1 & \text{αν } x \in (0, \frac{1}{2}) \\ -1 & \text{αν } x \in (\frac{1}{2}, 1) \end{cases} \quad r_2(x) = \begin{cases} 1 & \text{αν } x \in (0, \frac{1}{4}) \cup (\frac{1}{2}, \frac{3}{4}) \\ -1 & \text{αν } x \in (\frac{1}{4}, \frac{1}{2}) \cup (\frac{3}{4}, 1) \end{cases}$$

$$r_3(x) = \begin{cases} 1 & \text{αν } x \in (0, \frac{1}{8}) \cup (\frac{2}{8}, \frac{3}{8}) \cup (\frac{4}{8}, \frac{5}{8}) \cup (\frac{6}{8}, \frac{7}{8}) \\ -1 & \text{αν } x \in (\frac{1}{8}, \frac{2}{8}) \cup (\frac{3}{8}, \frac{4}{8}) \cup (\frac{5}{8}, \frac{6}{8}) \cup (\frac{7}{8}, 1) \end{cases} \quad \text{κλπ.}$$

Οι συναρτήσεις αυτές είναι ανεξάρτητες. Πράγματι, ελέγχουμε εύκολα ότι  $\mathbb{P}(r_k = 1) = \mathbb{P}(r_k = -1) = 1/2$  για κάθε  $k \in \mathbb{N}$ . Επίσης για κάθε  $k_1 < \dots < k_n \in \mathbb{N}$  και για κάθε πρόσημο  $\theta_j \in \{-1, 1\}$  ισχύει

$$\mathbb{P}(r_{k_1} = \theta_1 \text{ και } r_{k_2} = \theta_2 \text{ και } \dots r_{k_n} = \theta_n) = \frac{1}{2^n}.$$

Πράγματι, για  $n = 2$  σε κάθε διάστημα στο οποίο η  $r_{k_1}$  είναι ίση με  $\theta_1$  η  $r_{k_2}$  είναι ίση με  $\theta_2$  στο μισό από αυτό. Δηλαδή

$$\mathbb{P}(r_{k_1} = \theta_1 \text{ και } r_{k_2} = \theta_2) = \frac{1}{4}$$

και συνεχίζουμε επαγωγικά.

Παρατηρούμε τώρα ότι για κάθε σύνολο Borel  $A$  του  $\mathbb{R}$  ισχύει  $r_k^{-1}(A) = r_k^{-1}(A \cap \{-1, 1\})$ , και συνεπώς ο έλεγχος της ανεξαρτησίας αρκεί να γίνει για σύνολα  $A$  της μορφής  $\emptyset, \{-1\}, \{1\}$  και  $\{-1, 1\}$ . Αλλά κάθε  $r_k$  δίνει τιμές με βεβαιότητα στο  $\{-1, 1\}$  και ποτέ στο  $\emptyset$ . Συνεπώς αρκεί ο έλεγχος να γίνει για σύνολα  $A$  της μορφής  $\{-1\}$  και  $\{1\}$ . Έτσι, αρκεί να δείξουμε ότι

$$\mathbb{P}(r_{k_1} = \theta_1 \text{ και } r_{k_2} = \theta_2 \text{ και } \dots r_{k_n} = \theta_n) = \prod_{j=1}^n P(r_{k_j} = \theta_j),$$

για οποιαδήποτε πρόσημα  $\theta_j$ . Αλλά αυτό είναι προφανές, αφού από πριν έχουμε

$$\mathbb{P}(r_{k_1} = \theta_1 \text{ και } r_{k_2} = \theta_2 \text{ και } \dots r_{k_n} = \theta_n) = \frac{1}{2^n} = \prod_{j=1}^n P(r_{k_j} = \theta_j).$$

Ο απλούστερος τρόπος για να δοθούν περισσότερα παραδείγματα ανεξαρτησίας είναι σε χώρους γινόμενο. Ας υποθέσουμε ότι  $\Omega = \Omega_1 \times \dots \times \Omega_n$ ,  $B = B_1 \otimes \dots \otimes B_n$  και  $\mathbb{P} = \mathbb{P}_1 \times \dots \times \mathbb{P}_n$ . Οποιοσδήποτε τυχαίες μεταβλητές  $X_1, \dots, X_n$  στο  $\Omega$  όπου η  $X_j$  εξαρτάται μόνο από τη μεταβλητή της  $j$  συντεταγμένης είναι ανεξάρτητες. Πράγματι, γράφουμε κάθε  $X_j$  ως σύνθεση μιας  $f_j$  και της προβολής στην  $j$  συντεταγμένη, δηλαδή  $X_j = f_j \circ \text{proj}_j$ , οπότε ισχύει

$$(8.3) \quad (X_1, \dots, X_n)^{-1}(B_1 \times \dots \times B_n) = f_1^{-1}(B_1) \times \dots \times f_n^{-1}(B_n),$$

και άρα

$$\begin{aligned}
 \mathbb{P}_{(X_1, \dots, X_n)}(B_1 \times \dots \times B_n) &= \mathbb{P}(f_1^{-1}(B_1) \times \dots \times f_n^{-1}(B_n)) \\
 &= \prod_{j=1}^n \mathbb{P}_j(f_j^{-1}(B_j)) = \prod_{j=1}^n \mathbb{P}(\text{proj}_j^{-1}(f_j^{-1}(B_j))) \\
 &= \prod_{j=1}^n \mathbb{P}(X_j^{-1}(B_j)) = \prod_{j=1}^n (\mathbb{P}_{X_j}(B_j)) \\
 (8.4) \qquad &= \left( \prod_{j=1}^n \mathbb{P}_{X_j} \right) (B_1 \times \dots \times B_n).
 \end{aligned}$$

Οι ίδιες ιδέες μπορούν να λειτουργήσουν και για άπειρο πλήθος μεταβλητών παρόλο που η απόδειξη είναι πολύ πιο απαιτητική.

Σε ένα τοπολογικό χώρο ορίζεται βεβαίως η έννοια της Borel  $\sigma$ -άλγεβρας  $B_X$  ως η  $\sigma$ -άλγεβρα που παράγεται από τα ανοικτά σύνολα του χώρου. Ένα Borel μέτρο  $\mu$  στον  $X$  λέγεται *εξωτερικά κανονικό* στο σύνολο  $E$  αν

$$\mu(E) = \inf\{\mu(U) : U \supseteq E, U \text{ ανοικτό}\},$$

και *εσωτερικά κανονικό* στο  $E$  αν

$$\mu(E) = \sup\{\mu(F) : F \subseteq E, F \text{ συμπαγές}\}.$$

Αν το  $\mu$  είναι εξωτερικά και εσωτερικά κανονικό σε κάθε σύνολο Borel τότε ονομάζεται *κανονικό μέτρο*. Το να είναι ένα μέτρο κανονικό είναι αρκετά απαιτητική συνθήκη. Μια πιο «χαλαρή» απαίτηση είναι ένα μέτρο να είναι «μέτρο Radon»: λέμε ότι το  $\mu$  είναι μέτρο Radon στον  $X$  αν είναι μέτρο Borel στον  $X$ , πεπερασμένο στα συμπαγή σύνολα, εξωτερικά κανονικό στα σύνολα Borel, και εσωτερικά κανονικό στα ανοικτά. Θα δείξουμε αργότερα ότι τα Radon μέτρα είναι και εσωτερικά κανονικά εφόσον είναι  $\sigma$ -πεπερασμένα.

Ένας τοπολογικός χώρος  $X$  λέγεται *χώρος Hausdorff* αν για κάθε  $x, y \in X$  υπάρχουν ανοικτά ξένα σύνολα  $U, V$  ώστε  $x \in U$  και  $y \in V$ .

Το ακόλουθο θεώρημα θα αποδειχθεί μετά τη μελέτη των μέτρων Radon. Προς το παρόν μας δίνει απάντηση στο αν υπάρχει (άπειρη) ακολουθία ανεξάρτητων τυχαίων μεταβλητών.

**Θεώρημα 8.1.10** Έστω ότι οι  $\Omega_a$  για  $a \in A$  είναι μια οικογένεια συμπαγών χώρων Hausdorff, και για κάθε  $a$  έστω ότι το  $\mathbb{P}_a$  είναι ένα Radon μέτρο πιθανότητας στον  $\Omega_a$ . Τότε υπάρχει μοναδικό μέτρο Radon  $\mathbb{P}$  στον  $\Omega := \prod_{a \in A} \Omega_a$  ώστε για οποιαδήποτε διακεκριμένα  $a_1, \dots, a_n \in A$  η εικόνα του  $\mathbb{P}$  στον χώρο  $\prod_{j=1}^n \Omega_{a_j}$  με τη συνήθη προβολή είναι το γινόμενο των  $\mathbb{P}_{a_1}, \dots, \mathbb{P}_{a_n}$ .

Παρατηρούμε εδώ ότι η απαίτηση να είναι οι χώροι  $\Omega_a$  συμπαγείς δεν μας εμποδίζει να εφαρμόσουμε το παραπάνω θεώρημα σε μη συμπαγείς χώρους, όπως στο  $\mathbb{R}$ , αφού μπορούμε να εφαρμόσουμε το θεώρημα στη συμπαγοποίηση (κατά Alexandroff) του χώρου, θέτοντας  $\mathbb{P}_a(\{\infty\}) = 0$ . Έτσι με βάση το παραπάνω θεώρημα και τη συζήτηση που έγινε στις σχέσεις (8.3) και (8.4) μπορούμε να έχουμε οσοδήποτε ανεξάρτητες τυχαίες μεταβλητές με προκαθορισμένη κατανομή στο  $\mathbb{R}$ .

Για παράδειγμα, ας υποθέσουμε ότι θέλουμε να βρούμε τυχαίες ανεξάρτητες μεταβλητές  $X_a$  από ένα χώρο πιθανότητας  $\Omega$  στο  $\mathbb{R}$  που να έχουν όλες



κατανομή την γκαουσιανή  $\gamma$ , όπου  $\gamma(A) = \sqrt{2\pi}^{-1} \int_A e^{-t^2/2} dt$ . Θέτουμε  $\Omega_a = \mathbb{R}^*$ , όπου  $\mathbb{R}^*$  η κατά Alexandroff συμπαγοποίηση του  $\mathbb{R}$  με το γκαουσιανό μέτρο  $\gamma$  θέτοντας  $\gamma(\{\infty\}) = 0$ . Σύμφωνα με το παραπάνω θεώρημα το σύνολο  $\Omega = \prod_a \Omega_a$  εφοδιάζεται με (μοναδικό) μέτρο πιθανότητας  $\mathbb{P}$  με τις ιδιότητες που περιγράφει το θεώρημα. Ορίζουμε  $X_a : \Omega \rightarrow \mathbb{R}$  θέτοντας  $X_a((\omega_i)_{i \in A}) = \omega_a$  για κάθε  $(\omega_i)_{i \in A} \in \Omega$ . Αφού το  $\mathbb{P}$  σύμφωνα με το παραπάνω θεώρημα, μετά την προβολή στο  $\prod_{j=1}^n \Omega_{a_j}$  ταυτίζεται με το  $\prod_{j=1}^n \mathbb{P}_{a_j}$  σύμφωνα με τη συζήτηση που μας οδήγησε στη σχέση (8.4) οι  $X_a$  είναι ανεξάρτητες ( $X_a = f_a \circ \text{proj}_a$  με  $f_a(x) = x$ ). Τέλος είναι όλες τους γκαουσιανές, αφού

$$\mathbb{P}_{X_j}(E \subseteq \mathbb{R}^*) = \mathbb{P}(X_j^{-1}(E)) = \mathbb{P}(\text{proj}_j^{-1}(E) \subseteq \Omega) = \mathbb{P}_j(E) = \gamma(E),$$

όπου στην προτελευταία ισότητα χρησιμοποιήσαμε το τελευταίο θεώρημα με  $n = 1$  και  $a_1 = j$ .

## 8.2 Ο νόμος των μεγάλων Αριθμών

**Θεώρημα 8.2.1 (Ασθενής νόμος των μεγάλων αριθμών)** Έστω ότι η  $X_n$  είναι μια ακολουθία ανεξάρτητων τυχαίων μεταβλητών στον  $L_2$  με μέσες τιμές  $\mu_n$  και διακυμάνσεις  $\sigma_n^2$ . Αν  $\lim_{n \rightarrow \infty} n^{-2} \sum_{j=1}^n \sigma_j^2 = 0$  τότε

$$\mathbb{P}\left(\left|\frac{1}{n} \sum_{j=1}^n (X_j - \mu_j)\right| > \epsilon\right) \rightarrow 0$$

καθώς  $n \rightarrow \infty$ .

Απόδειξη: Παρατηρούμε ότι  $\mathbb{E}(n^{-1} \sum_{j=1}^n (X_j - \mu_j)) = 0$  και λόγω της ανεξαρτησίας,

$$\sigma^2\left(\frac{1}{n} \sum_{j=1}^n (X_j - \mu_j)\right) = \frac{1}{n^2} \sum_{j=1}^n \sigma_j^2.$$

Άρα

$$\frac{1}{n^2} \sum_{j=1}^n \sigma_j^2 = \mathbb{E}\left(\frac{1}{n} \sum_{j=1}^n (X_j - \mu_j)\right)^2 \geq \int_{\left\{\left|n^{-1} \sum_{j=1}^n (X_j - \mu_j)\right| > \epsilon\right\}} \epsilon^2,$$

άρα

$$\mathbb{P}\left(\left|\frac{1}{n} \sum_{j=1}^n (X_j - \mu_j)\right| > \epsilon\right) \leq \frac{1}{n^2 \epsilon^2} \sum_{j=1}^n \sigma_j^2 \rightarrow 0.$$

□

**Λήμμα 8.2.2 (Borel-Cantelli)** Έστω ότι τα  $A_n$  είναι μια ακολουθία ενδεχομένων. Αν  $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$  τότε  $\mathbb{P}(\limsup A_n) = 0$ . Αν από την άλλη μεριά τα  $A_n$  είναι ανεξάρτητα και  $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$  τότε  $\mathbb{P}(\limsup A_n) = 1$ .

Απόδειξη: Επειδή  $\limsup A_n = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n$  θα έχουμε

$$\mathbb{P}(\limsup A_n) \leq \mathbb{P}\left(\bigcup_{n=k}^{\infty} A_n\right) \leq \sum_{n=k}^{\infty} \mathbb{P}(A_n) \rightarrow 0$$

καθώς  $k \rightarrow \infty$ .

Επειδή  $(\limsup A_n)^c = \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} A_n^c$  αρκεί να δείξουμε ότι  $\mathbb{P}(\bigcap_{n=k}^{\infty} A_n^c) = 0$  για κάθε  $k \in \mathbb{N}$ . Αλλά τα  $A_n$  είναι ανεξάρτητα από όπου εύκολα προκύπτει ότι και τα  $A_n^c$  είναι ανεξάρτητα (Άσκηση ι). Άρα για κάθε  $N \in \mathbb{N}$  ισχύει

$$\begin{aligned} \mathbb{P}\left(\bigcap_{n=k}^{\infty} A_n^c\right) &= \prod_{n=k}^{\infty} (1 - \mathbb{P}(A_n)) \\ &\leq \prod_{n=k}^{\infty} e^{-\mathbb{P}(A_n)} = e^{-\sum_{n=k}^{\infty} \mathbb{P}(A_n)} \rightarrow 0, \end{aligned}$$

καθώς  $N \rightarrow \infty$ , ολοκληρώνοντας την απόδειξη.  $\square$

**Πρόταση 8.2.3 (Ανισότητα Kolmogorov)** Έστω ότι οι  $X_1, \dots, X_n$  είναι ανεξάρτητες τυχαίες μεταβλητές μέσης τιμής 0 και με διακυμάνσεις  $\sigma_1^2, \dots, \sigma_n^2$ . Αν  $S_k = X_1 + \dots + X_k$  τότε για κάθε  $\epsilon > 0$  ισχύει

$$\mathbb{P}\left(\max_{1 \leq k \leq n} |S_k| \geq \epsilon\right) \leq \frac{1}{\epsilon^2} \sum_{k=1}^n \sigma_k^2.$$

Απόδειξη: Θεωρούμε τα σύνολα

$$A_k = \{\omega : |S_j| < \epsilon \text{ για } j = 1, \dots, k-1 \text{ και } |S_k| \geq \epsilon\}.$$

Φανερά αυτά είναι μεταξύ τους ξένα και

$$\bigcup_{k=1}^n A_k = \{\omega : \max |S_k| \geq \epsilon\}.$$

Άρα

$$\mathbb{P}\{\max_k |S_k| \geq \epsilon\} = \sum_{k=1}^n \mathbb{P}(A_k) \leq \frac{1}{\epsilon^2} \sum_{k=1}^n \mathbb{E}(\chi_{A_k} S_k^2),$$

αφού  $\int_{A_k} S_k^2 \geq \epsilon^2 \mathbb{P}(A_k)$ . Από την άλλη,

$$\begin{aligned} \mathbb{E}(S_n^2) &\geq \sum_{k=1}^n \mathbb{E}(\chi_{A_k} S_n^2) = \sum_{k=1}^n \mathbb{E}\left((\chi_{A_k} (S_k^2 + 2S_k(S_n - S_k) + (S_n - S_k)^2))\right) \\ &\geq \sum_{k=1}^n \mathbb{E}(\chi_{A_k} S_k^2) + 2 \sum_{k=1}^n \mathbb{E}(\chi_{A_k} S_k(S_n - S_k)). \end{aligned}$$

Αν τώρα δείξουμε ότι  $\mathbb{E}(\chi_{A_k} S_k(S_n - S_k)) = 0$  τότε θα έχουμε δείξει ότι

$$\mathbb{P}\{\max_{1 \leq k \leq n} |S_k| \geq \epsilon\} \leq \frac{1}{\epsilon^2} \mathbb{E}(S_n^2) = \frac{1}{\epsilon^2} \sum_{k=1}^n \sigma_k^2,$$

ολοκληρώνοντας την απόδειξη. Αλλά η μεταβλητή  $\chi_{A_k} S_k$  είναι συνάρτηση των  $X_1, \dots, X_k$  ενώ η  $S_n - S_k$  είναι συνάρτηση των  $X_{k+1}, \dots, X_n$ , οπότε είναι ανεξάρτητες, και συνεπώς ισχύει

$$\mathbb{E}(\chi_{A_k} S_k(S_n - S_k)) = \mathbb{E}(\chi_{A_k} S_k) \mathbb{E}(S_n - S_k) = \mathbb{E}(\chi_{A_k} S_k) \cdot 0 = 0.$$

$\square$

**Θεώρημα 8.2.4 (Ισχυρός νόμος των μεγάλων αριθμών του Kolmogorov)** Έστω ότι η  $X_n$  είναι μια ακολουθία ανεξάρτητων τυχαίων μεταβλητών στον  $L_2$  με μέσες τιμές  $\mu_n$  και διακυμάνσεις  $\sigma_n^2$ . Αν  $\sum_{n=1}^{\infty} n^{-2} \sigma_n^2 < \infty$  τότε η ακολουθία

$$\frac{1}{n} \sum_{j=1}^n (X_j - \mu_j)$$

συγκλίνει στο 0 σχεδόν σίγουρα για  $n \rightarrow \infty$ .

Απόδειξη: Θέτουμε  $S_n = \sum_{k=1}^n (X_k - \mu_k)$ , και για  $\epsilon > 0$  θέτουμε

$$A_k = \left\{ \omega : \frac{1}{n} |S_n| \geq \epsilon \text{ για κάποιο } 2^{k-1} \leq n < 2^k \right\}.$$

Στο  $A_k$  ισχύει  $|S_n| \geq \epsilon 2^{k-1}$  για κάποιο  $n < 2^k$ . Από την ανισότητα Kolmogorov

$$\begin{aligned} \mathbb{P}(A_k) &\leq \mathbb{P}\left\{ \max_{2^{k-1} \leq n < 2^k} |S_n| \geq \epsilon n \right\} \\ &\leq \frac{1}{\epsilon^2 2^{k-1}} \sum_{n=2^{k-1}}^{2^k-1} \sigma_n^2 \leq \frac{1}{\epsilon^2 4^{k-1}} \sum_{n=1}^{2^k} \sigma_n^2. \end{aligned}$$

Άρα

$$\begin{aligned} \sum_{k=1}^{\infty} \mathbb{P}(A_k) &\leq \frac{4}{\epsilon^2} \sum_{k=1}^{\infty} \sum_{n=1}^{2^k} \frac{1}{2^{2k}} \sigma_n^2 \\ &= \frac{4}{\epsilon^2} \sum_{n=1}^{\infty} \left( \sum_{k \geq \log_2 n} \frac{1}{2^{2k}} \right) \sigma_n^2 \leq \frac{8}{\epsilon^2} \sum_{n=1}^{\infty} \frac{1}{n^2} \sigma_n^2, \end{aligned}$$

όπου η αλλαγή της σειράς άθροισης γίνεται εύκολα χρησιμοποιώντας χαρακτηριστικές συναρτήσεις (και το θεώρημα Tonelli):

$$\sum_{k=1}^{\infty} \sum_{n=1}^{2^k} \frac{1}{2^{2k}} \sigma_n^2 = \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} \chi_{\{1, \dots, 2^k\}}(n) \frac{1}{2^{2k}} \sigma_n^2 \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \chi_{\{1, \dots, 2^k\}}(n) \frac{1}{2^{2k}} \sigma_n^2,$$

και παρατηρώντας ότι  $\chi_{\{1, \dots, 2^k\}}(n) = 1$  αν και μόνο αν  $k \geq \log_2 n$ .

Από το Λήμμα Borel-Cantelli συμπεραίνουμε ότι  $\mathbb{P}(\limsup A_k) = 0$ , δηλαδή σχεδόν σίγουρα  $|S_n|/n \leq \epsilon$ . Άρα  $\limsup_n |S_n|/n \leq \epsilon$  σχεδόν σίγουρα. Επιλέγουμε τέλος  $\epsilon = 1/k$  και άρα το σύνολο  $B_k = \{\omega : \limsup_n |S_n|/n \geq 1/k\}$  έχει πιθανότητα μηδέν, οπότε  $\mathbb{P}(\cup_k B_k) = 0$ , δηλαδή  $\limsup_n |S_n|/n = 0$  σ.σ.  $\square$

**Θεώρημα 8.2.5 (Ισχυρός νόμος των μεγάλων αριθμών του Khinchine)** Έστω ότι η  $X_n$  είναι μια ακολουθία ανεξάρτητων τυχαίων μεταβλητών στον  $L_1$  με την ίδια μέση τιμή  $\mu$ . Τότε η ακολουθία

$$\frac{1}{n} \sum_{j=1}^n X_j$$

συγκλίνει στο  $\mu$  σχεδόν σίγουρα για  $n \rightarrow \infty$ .

Απόδειξη: Αντικαθιστώντας τις  $X_j$  με τις  $X_j - \mu$  μπορούμε να υποθέσουμε χωρίς βλάβη της γενικότητας ότι  $\mu = 0$ . Έστω ότι η  $\nu$  είναι η κοινή κατανομή των  $X_j$ . Έτσι (από την Πρόταση 8.1.2 για  $f(t) = |t|$  και  $\phi = X_j$ ) έχουμε  $\int |t| d\nu(t) = \mathbb{E}|X_j| < \infty$  και  $\int t d\nu(t) = \mathbb{E}(X_j) = 0$ . Θέτουμε

$$Y_j = \begin{cases} X_j & \text{αν } |X_j| \leq j \\ 0 & \text{αλλιώς.} \end{cases}$$

Έτσι έχουμε

$$\begin{aligned} \sum_{j=1}^{\infty} \mathbb{P}\{Y_j \neq X_j\} &= \sum_{j=1}^{\infty} \mathbb{P}\{|X_j| > j\} = \sum_{j=1}^{\infty} \mathbb{P}\{X_j^{-1}\{t : |t| > j\}\} \\ &= \sum_{j=1}^{\infty} \mathbb{P}_{X_j}\{t : |t| > j\} = \sum_{j=1}^{\infty} \nu\{t : |t| > j\} \\ &= \sum_{j=1}^{\infty} j \cdot \nu\{t : j < |t| \leq j+1\} \\ &\leq \int_0^{\infty} \nu\{t : |t| > s\} ds = \int |t| d\nu(t) < \infty. \end{aligned}$$

Άρα από το Λήμμα Borel-Cantelli

$$\mathbb{P}(\limsup\{\omega : Y_j(\omega) \neq X_j(\omega)\}) = 0.$$

Δηλαδή

$$P(\limsup\{\omega : Y_j(\omega) = X_j(\omega) \text{ για μεγάλα } j\}) = 1.$$

Συνεπώς αρκεί να δείξουμε ότι  $n^{-1} \sum_{j=1}^n Y_j \rightarrow 0$  σ.σ.

Για τις  $Y_j$  έχουμε

$$\sigma^2(Y_n) = (Y_n^2) - (\mathbb{E}Y_n)^2 \leq \mathbb{E}(Y_n^2) = \int_{|t| \leq n} t^2 d\nu(t).$$

Άρα

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{1}{n^2} \sigma^2(Y_n) &\leq \sum_{n=1}^{\infty} \sum_{j=1}^n \frac{1}{n^2} \int_{j-1 < |t| \leq j} t^2 d\nu(t) \\ &\leq \sum_{n=1}^{\infty} \sum_{j=1}^n \frac{j}{n^2} \int_{j-1 < |t| \leq j} |t| d\nu(t) \\ &\leq \sum_{j=1}^{\infty} \sum_{n=j}^{\infty} \frac{j}{n^2} \int_{j-1 < |t| \leq j} |t| d\nu(t). \end{aligned}$$

Επειδή τώρα  $\sum_{n=j}^{\infty} 1/n^2 \leq 2/j$  συγκρίνοντας με το  $\int_j^{\infty} x^{-2} dx$ , καταλήγουμε στην

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \sigma^2(Y_n) \leq 2 \int_{-\infty}^{\infty} |t| d\nu(t) < \infty.$$

Άρα από τον ισχυρό νόμο των μεγάλων αριθμών για  $\mu_j = \mathbb{E}(Y_j)$  ισχύει

$$\frac{1}{n} \sum_{j=1}^n Y_j - \frac{1}{n} \sum_{j=1}^n \mu_j \rightarrow 0$$

σχεδόν σίγουρα. Αλλά από το Θεώρημα Κυριαρχημένης Σύγκλισης

$$\mu_j = \int_{|t| \leq j} t dv(t) \rightarrow \int_{-\infty}^{\infty} t dv(t) = 0,$$

οπότε  $n^{-1} \sum_{j=1}^n \mu_j \rightarrow 0$ , ολοκληρώνοντας την απόδειξη.  $\square$

### Ασκήσεις

- (i) Αποδείξτε ότι αν τα ενδεχόμενα  $A_1, \dots, A_n$  είναι ανεξάρτητα τότε είναι ανεξάρτητα και τα  $A_1^c, \dots, A_n^c$ , αποδεικνύοντας πρώτα ότι τα  $A_1, \dots, A_{n-1}, A_n^c$  είναι ανεξάρτητα, και συνεχίζοντας με επαγωγή.

### 8.3 Το κεντρικό οριακό θεώρημα

Ένας τοπολογικός χώρος  $X$  λέγεται τοπικά συμπαγής αν κάθε σημείο  $x \in X$  έχει μια συμπαγή περιοχή. Δηλαδή, υπάρχει συμπαγές  $K \subseteq X$  ώστε  $x \in K^\circ$ .

Λέμε ότι μια συνάρτηση  $f : X \rightarrow \mathbb{R}$  μηδενίζεται στο άπειρο όταν για κάθε  $\epsilon > 0$  το σύνολο  $\{x : |f(x)| \geq \epsilon\}$  είναι συμπαγές υποσύνολο του  $X$ .

Θέτουμε

$$C_0(X) = \{f \in C(X) : f \text{ μηδενίζεται στο άπειρο}\},$$

όπου  $C(X)$  ο χώρος των συνεχών συναρτήσεων από τον τοπολογικό χώρο  $X$  στο  $\mathbb{R}$ .

Μπορεί να αποδείξει κανείς ότι αν ο  $X$  είναι τοπικά συμπαγής τότε το  $C_0(X)$  είναι η κλειστή θήκη του

$$C_c(X) = \{f \in C(X) : \text{supp}(f) \text{ συμπαγές}\},$$

στον χώρο  $(C(X), \|\cdot\|_\infty)$ , όπου  $\text{supp}(f) = \{x \in X : f(x) \neq 0\}$ . Επίσης ένα πολύ γνωστό θεώρημα από την συναρτησιακή ανάλυση, το Θεώρημα αναπαράστασης του Riesz, λέει ότι αν ο  $X$  είναι τοπικά συμπαγής τοπολογικός χώρος ο δυϊκός χώρος του χώρου Banach  $(C_0(X), \|\cdot\|_\infty)$  είναι ισομετρικά ισομορφος με τον χώρο των Radon προσημασμένων μέτρων  $(M(X), \|\cdot\|)$ , όπου  $\|\mu\| := |\mu|(X) = \mu^+(X) + \mu^-(X)$  η ολική κύμανση του μέτρου  $\mu$ . Στον χώρο αυτό ορίζουμε την ασθενή άστρο (weak\*) τοπολογία ορίζοντας για μια ακολουθία  $\mu_n \in M(X)$  να συγκλίνει στο  $\mu \in M(X)$  αν και μόνο αν

$$\int f d\mu_n \rightarrow \int f d\mu$$

για κάθε  $f \in C_0(X)$ . Λέμε τότε ότι η  $\mu_n$  συγκλίνει ασθενώς-\* στο  $\mu$  και γράφουμε  $\mu_n \xrightarrow{w^*} \mu$ . Θυμίζουμε ότι η σύγκλιση ορίζει τοπολογία, αφού η σύγκλιση μπορεί να περιγράψει τα κλειστά σύνολα (όσα περιέχουν τα κατά τη σύγκλιση αυτή

οριακά τους σημεία) και συνεπώς ορίζει και τα ανοικτά ως συμπληρώματα των κλειστών.

Το παρακάτω θεώρημα, παρόλο που είναι κεντρικό στη θεωρία πιθανοτήτων, δεν θα το αποδείξουμε γιατί απαιτεί γνώσεις από τον μετασχηματισμό Fourier, που είναι έξω από τους σκοπούς του παρόντος κειμένου.

**Θεώρημα 8.3.1** Έστω ότι η  $X_n$  είναι μια ακολουθία ανεξάρτητων τυχαίων μεταβλητών στον  $L_2$  με κοινή μέση τιμή  $\mu$  και διακύμανση  $\sigma^2$ . Τότε, καθώς  $n \rightarrow \infty$ , η κατανομή της

$$\frac{1}{\sigma\sqrt{n}} \sum_{j=1}^n (X_j - \mu)$$

συγκλίνει ασθενώς-\* στην τυπική κανονική κατανομή  $\gamma_1$ , και ειδικότερα

$$\mathbb{P} \left( \frac{1}{\sigma\sqrt{n}} \sum_{j=1}^n (X_j - \mu) \leq a \right) \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^a e^{-t^2/2} dt.$$

## 8.4 Ασκήσεις

- (i) Έστω ότι  $\Omega = \{1, 2, 3, 4\}$  και  $\mathbb{P}(\{n\}) = 1/4$  για κάθε  $n \in \Omega$ . Βρείτε τρία ενδεχόμενα που να είναι ανεξάρτητα ανά δύο αλλά να μην είναι ανεξάρτητα.
- (ii)