10

Ανεξαρτησία

10.1 Ανεξαρτησία για οικογένειες συνόλων και τυχαίες μεταβλητές

Στην παράγραφο αυτή δουλεύουμε σε χώρο πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$.

Δίνουμε καταρχάς τον ορισμό της ανεξαρτησίας για ενδεχόμενα, σύνολα ενδεχομένων, και τυχαίες μεταβλητές.

Ορισμός 10.1. Έστω $(A_i)_{i\in I}$ στοιχεία της $\mathcal{F}$. Τα $(A_i)_{i\in I}$ λέγονται **ανεξάρτητα** αν για κάθε $J \subset I$ πεπερασμένο ισχύει ότι

$$\mathbf{P}\left(\cap_{i\in J} A_i\right) = \prod_{i\in J} \mathbf{P}(A_i). \tag{10.1}$$

Η τομή και το γινόμενο στην τελευταία ισότητα έχουν πεπερασμένο πλήθος όρων.

Ορισμός 10.2. Έστω $(\mathcal{F}_i)_{i\in I}$ οικογένεια συνόλων έτσι ώστε $\mathcal{F}_i \subset \mathcal{F}$ για κάθε $i \in I$. Τα $(\mathcal{F}_i)_{i\in I}$ λέγονται **ανεξάρτητα** αν για κάθε $J \subset I$ πεπερασμένο και $A_i \in \mathcal{F}_i$ για κάθε $i \in J$ ισχύει η (10.1).

Ορισμός 10.3. Έστω $\{(E_i, \mathcal{E}_i) : i \in I\}$ μετρήσιμοι χώροι και $(X_i)_{i\in I}$ οικογένεια τυχαίων μεταβλητών με $X_i : \Omega \to E_i$ για κάθε $i \in I$. Οι $(X_i)_{i\in I}$ λέγονται **ανεξάρτητες** αν οι αντίστοιχες σ-άλγεβρες $(\sigma(X_i))_{i\in I}$, που είναι υποσύνολα της $\mathcal{F}$, είναι ανεξάρτητες.

Παρατήρηση 10.4. Ο Ορισμός 10.3, σύμφωνα με τον Ορισμό 10.2, απαιτεί

$$\mathbf{P}(X_{i_1} \in A_{i_1}, X_{i_2} \in A_{i_2}, \ldots, X_{i_n} \in A_{i_n}) = \mathbf{P}(X_{i_1} \in A_{i_1})\,\mathbf{P}(X_{i_2} \in A_{i_2}) \cdots \mathbf{P}(X_{i_n} \in A_{i_n}) \tag{10.2}$$

για κάθε $n \geq 2$, κάθε επιλογή δεικτών $i_1, i_2, \ldots, i_n \in I$, και κάθε $A_{i_1} \in \mathcal{E}_{i_1}, \ldots, A_{i_n} \in \mathcal{E}_{i_n}$, αφού κάθε στοιχείο μιας $\sigma(X_i)$ είναι της μορφής $X_i^{-1}(A_i) = \{X_i \in A_i\}$ με $A_i \in \mathcal{E}_i$. Το ενδεχόμενο στο αριστερό μέλος της τελευταίας σχέσης είναι συντομογραφία του ενδεχομένου $X_{i_1}^{-1}(A_{i_1}) \cap X_{i_2}^{-1}(A_{i_2}) \cap \cdots \cap X_{i_n}^{-1}(A_{i_n})$.

Σύμβαση: Στο εξής, όποτε λέμε ότι κάποιες τυχαίες μεταβλητές είναι ανεξάρτητες, θα εννοείται ότι ορίζονται στον ίδιο χώρο πιθανότητας, δηλαδή έχουν το ίδιο πεδίο ορισμού. Αυτό εξασφαλίζει ότι το ενδεχόμενο $X_{i_1}^{-1}(A_{i_1}) \cap X_{i_2}^{-1}(A_{i_2}) \cap \cdots \cap X_{i_n}^{-1}(A_{i_n})$ στο αριστερό μέλος της (10.2) είναι στοιχείο της $\mathcal{F}$, που είναι το πεδίο ορισμού της $\mathbf{P}$.

Παράδειγμα 10.5 (Δυό ανεξάρτητες τυχαίες μεταβλητές)**.** Θεωρούμε το πείραμα δύο ρίψεων ενός νομίσματος που φέρνει κορώνα με πιθανότητα p. Το σύνηθες μέτρο που μοντελοποιεί το πείραμα είναι τέτοιο ώστε τα αποτελέσματα των δύο ρίψεων να είναι ανεξάρτητες τυχαίες μεταβλητές.

Πιο συγκεκριμένα, ο δειγματικός μας χώρος είναι ο $\Omega = \{K, \Gamma\} \times \{K, \Gamma\}$ και σ-άλγεβρα η $\mathcal{F} = \mathscr{P}(A)$. Για $\omega \in \Omega$ θέτουμε

$$p_\omega = \begin{cases} p(1-p) & \text{αν } \omega = (K, \Gamma) \text{ ή } (\Gamma, K), \\ p^2 & \text{αν } \omega = (K, K), \\ (1-p)^2 & \text{αν } \omega = (\Gamma, \Gamma). \end{cases}$$

Έστω $\mathbf{P}$ το μοναδικό μέτρο πιθανότητας στην $\mathcal{F}$ με $\mathbf{P}(\omega) = p_\omega$. Έστω $E = \{K, \Gamma\}$ και $\mathcal{E} = \mathscr{P}(E)$. Θεωρούμε τις τυχαίες μεταβλητές $X, Y : \Omega \to E$, με $X((x, y)) = x$ και $Y((x, y)) = y$. Τότε, η X είναι η ένδειξη της πρώτης ρίψης και η Y η ένδειξη της δεύτερης.

Ισχυρισμος: Οι X, Y είναι ανεξάρτητες.
Θα δείξουμε ότι για κάθε $A, B \in \mathcal{E}$ ισχύει

$$\mathbf{P}(X \in A, Y \in B) = \mathbf{P}(X \in A)\,\mathbf{P}(Y \in B). \tag{10.3}$$

- Αν $A = \emptyset$ ή $B = \emptyset$, η (10.3) ισχύει.
- Αν $A = \{K, \Gamma\}$, τότε $\{X \in A, Y \in B\} = \{Y \in B\}$, και $\mathbf{P}(X \in A) = 1$. Άρα η (10.3) πάλι ισχύει.
- Αν $B = \{K, \Gamma\}$, η (10.3) αποδεικνύεται όμοια.

Τέλος, μένουν οι περιπτώσεις που τα A, B είναι μονοσύνολα. Για παράδειγμα, αν $A = \{K\}$ και $B = \{\Gamma\}$, έχουμε

$$\mathbf{P}(X = K, Y = \Gamma) = \mathbf{P}(\{(K,\Gamma), (K,K)\} \cap \{(K,\Gamma), (\Gamma,\Gamma)\}) = \mathbf{P}(\{K,\Gamma\}) = p(1-p).$$

Όμως

$$\mathbf{P}(\{(K,\Gamma),(K,K)\}) = p(1-p) + pp = p$$
$$\mathbf{P}(\{(K,\Gamma),(\Gamma,\Gamma)\}) = p(1-p) + (1-p)^2 = 1-p.$$

Πολλαπλασιάζοντας κατά μέλη έχουμε ότι

$$\mathbf{P}(\{(K,\Gamma),(K,K)\})\,\mathbf{P}(\{(K,\Gamma),(\Gamma,\Gamma)\}) = p(1-p)$$

και έτσι η (10.3) ισχύει πάλι.
Όμοια αποδεικνύονται και οι υπόλοιπες περιπτώσεις όπου τα A, B είναι μονοσύνολα.

Το επόμενο θεώρημα διευκολύνει τον έλεγχο ανεξαρτησίας δύο τυχαίων μεταβλητών.

Θεώρημα 10.6. *Έστω $(E, \mathcal{E})$, $(G, \mathcal{G})$ μετρήσιμοι χώροι και $X : \Omega \to E$, $Y : \Omega \to G$ τυχαίες μεταβλητές. Θεωρούμε τη σχέση*

$$\mathbf{P}(X \in A, Y \in B) = \mathbf{P}(X \in A)\,\mathbf{P}(Y \in B). \tag{$*$}$$

Τα εξής είναι ισοδύναμα:

(i) Οι X, Y είναι ανεξάρτητες.

(ii) Η $(*)$ ισχύει για κάθε $A \in \mathcal{E}$ και $B \in \mathcal{G}$.

(iii) Η $(*)$ ισχύει για κάθε $A \in \mathcal{C}$, $B \in \mathcal{D}$, όπου $\mathcal{C}$, $\mathcal{D}$ οικογένειες κλειστές στις πεπερασμένες τομές με $\sigma(\mathcal{C}) = \mathcal{E}$, $\sigma(\mathcal{D}) = \mathcal{G}$.

Σημαντική είναι η ισοδυναμία των (i) και (iii). Δηλαδή αρκεί να ελέγξουμε την $(*)$ για λιγότερα σύνολα από ότι απαιτεί η (ii) ώστε να διαπιστώσουμε την ανεξαρτησία των X, Y.

Απόδειξη. Η (ii) είναι αναδιατύπωση του ορισμού της ανεξαρτησίας και προφανώς συνεπάγεται την (iii). Μένει να δείξουμε ότι η (iii) συνεπάγεται την (ii).
Έστω $A \in \mathcal{C}$ και

$$\mathcal{D}_1(A) := \{B \in \mathcal{G} : \text{ η } (*) \text{ ισχύει για τα } A, B\}.$$

Έχουμε ότι $\mathcal{D} \subset \mathcal{D}_1(A)$ από υπόθεση και η $\mathcal{D}_1(A)$ είναι κλάση Dynkin (Άσκηση 3.1). Άρα $\delta(\mathcal{D}) \subset \mathcal{D}_1(A)$. Επειδή η $\mathcal{D}$ είναι κλειστή στις πεπερασμένες τομές, το θεώρημα μονότονης κλάσης δίνει ότι $\sigma(\mathcal{D}) = \delta(\mathcal{D})$. Άρα $\sigma(\mathcal{D}) \subset \mathcal{D}_1(A)$, δηλαδή η $(*)$ ισχύει για κάθε $A \in \mathcal{C}$ και $B \in \mathcal{G}$. Τώρα για $B \in \mathcal{G}$ θέτουμε

$$\mathcal{D}_2(B) := \{A \in \mathcal{E} : \text{ η } (*) \text{ ισχύει για τα } A, B\}.$$

Όμοια, όπως με το $\mathcal{D}_1(A)$, δείχνουμε ότι $\mathcal{D}_2(B) = \mathcal{E}$, και έτσι αποδείχθηκε η (ii). ∎

Στις στοιχειώδεις πιθανότητες μαθαίνουμε (χωρίς απόδειξη) ότι δύο τυχαίες μεταβλητές X, Y είναι ανεξάρτητες αν και μόνο αν η από κοινού συνάρτηση κατανομή τους, $F_{X,Y}$, γράφεται ως $F_{X,Y}(x, y) = F_X(y)F_Y(y)$ για κάθε $x, y \in \mathbb{R}$. Τώρα είμαστε σε θέση να το αποδείξουμε.

Πόρισμα 10.7. *Έστω $X, Y : \Omega \to \mathbb{R}$ τυχαίες μεταβλητές. Τότε οι X, Y είναι ανεξάρτητες αν και μόνο αν*

$$\mathbf{P}(X \le x, Y \le y) = \mathbf{P}(X \le x)\,\mathbf{P}(Y \le y)$$

για κάθε $x, y \in \mathbb{R}$.

Απόδειξη. Προκύπτει από το Θεώρημα 10.6 αν πάρουμε $C = \mathcal{D} = \{(-\infty, a] : a \in \mathbb{R}\}$. ∎

Πρόταση 10.8. *Έστω X, Y όπως στη διατύπωση του Θεωρήματος 10.6. Τα εξής είναι ισοδύναμα:*

(i) Οι X, Y είναι ανεξάρτητες.

(ii) Οι $f(X), g(Y)$ είναι ανεξάρτητες για κάθε $f : E \to \mathbb{R}$, $g : G \to \mathbb{R}$ μετρήσιμες.

Απόδειξη. $(i) \Rightarrow (ii)$ Έστω $A, B \in \mathscr{B}(\mathbb{R})$. Τότε,

$$\begin{aligned}\mathbf{P}(f(X) \in A, g(Y) \in B) &= \mathbf{P}(X \in f^{-1}(A), Y \in g^{-1}(B)) = \mathbf{P}(X \in f^{-1}(A))\,\mathbf{P}(Y \in g^{-1}(B))\\ &= \mathbf{P}(f(X) \in A)\,\mathbf{P}(g(Y) \in B).\end{aligned}$$

Η δεύτερη ισότητα προκύπτει από την ανεξαρτησία των X, Y. Το ζητούμενο δείχθηκε.
$(ii) \Rightarrow (i)$ Έστω $\Gamma \in \mathcal{E}, \Delta \in \mathcal{G}$. Το (ii) για τις $f = \mathbf{1}_C$ και $g = \mathbf{1}_D$ δίνει

$$\mathbf{P}(\{f(X) \in \{1\}\} \cap \{g(X) \in \{1\}\}) = \mathbf{P}(\{f(X) \in \{1\}\})\,\mathbf{P}(\{g(X) \in \{1\}\})$$

δηλαδή

$$\mathbf{P}(X \in \Gamma, Y \in D) = \mathbf{P}(X \in \Gamma)\,\mathbf{P}(Y \in D),$$

που είναι το ζητούμενο. ∎

Θεώρημα 10.9. *Έστω $X, Y : \Omega \to [-\infty, \infty]$ ανεξάρτητες τυχαίες μεταβλητές μη αρνητικές ή με $\mathbf{E}\,|X|, \mathbf{E}\,|Y| < \infty$. Τότε*

$$\mathbf{E}(XY) = \mathbf{E}(X)\,\mathbf{E}(Y).$$

Απόδειξη. Το σχέδιο της απόδειξης είναι να δείξουμε κάτι φαινομενικά ισχυρότερο. Δηλαδή το εξής. Ισχυρισμος:

$$\mathbf{E}\{f(X)g(Y)\} = \mathbf{E}(f(X))\,\mathbf{E}(g(Y)) \tag{10.4}$$

για κάθε $f : [-\infty, \infty] \to [-\infty, \infty]$, $g : [-\infty, \infty] \to [-\infty, \infty]$ μετρήσιμες που είναι μη αρνητικές ή ικανοποιούν $\mathbf{E}\,|f(X)|, \mathbf{E}\,|g(Y)| < \infty$.
Θα δείξουμε την (10.4) σταδιακά με τον γνωστό τρόπο (Τυπική Μηχανή).

Βήμα 1. Αν $f = \mathbf{1}_A$ και $g = \mathbf{1}_B$, όπου $A, B \in \mathscr{B}([-\infty, \infty])$, έχουμε ότι

$$\begin{aligned}\mathbf{E}(f(X)g(Y)) &= \mathbf{E}(\mathbf{1}_{X\in A}\mathbf{1}_{Y\in B}) = \mathbf{E}(\mathbf{1}_{\{X\in A\}\cap\{Y\in B\}}) = \mathbf{P}(X \in A, Y \in B)\\ &= \mathbf{P}(X \in A)\,\mathbf{P}(Y \in B) = \mathbf{E}(\mathbf{1}_{X\in A})\,\mathbf{E}(\mathbf{1}_{Y\in B}).\end{aligned}$$

Η τέταρτη ισότητα ισχύει γιατί οι X, Y είναι ανεξάρτητες. Και έτσι προκύπτει η (10.4) για τις συγκεκριμένες f, g.

Βήμα 2. Αν $f, g \ge$ απλές μετρήσιμες, έστω

$$f = \sum_{i=1}^{m} a_i \mathbf{1}_{A_i}, \quad g = \sum_{j=1}^{n} b_j \mathbf{1}_{B_j},$$

σε κανονική μορφή, όπου $A_i, B_j \in \mathscr{B}([-\infty, \infty])$, τότε

$$\mathbf{E}(f(X)g(Y)) = \mathbf{E}\Big(\sum_{i=1}^{m}\sum_{j=1}^{n} a_i b_j \mathbf{1}_{A_i}(X)\mathbf{1}_{B_j}(Y)\Big).$$

Και λόγω γραμμικότητας και του προηγούμενου βήματος, η τελευταία μέση τιμή ισούται με

$$\sum_{i=1}^{m}\sum_{j=1}^{n} a_i b_j\, \mathbf{E}(\mathbf{1}_{A_i}(X))\, \mathbf{E}(\mathbf{1}_{B_j}(Y)) = \mathbf{E}(f(X))\, \mathbf{E}(g(Y)).$$

Βήμα 3. Αν $f, g \geq 0$ μετρήσιμες, τότε υπάρχουν αύξουσες ακολουθίες $(r_n)_{n\in\mathbb{N}}$ και $(s_n)_{n\in\mathbb{N}}$ μη αρνητικών απλών συναρτήσεων έτσι ώστε

$$\lim_{n\to\infty} r_n = f, \quad \lim_{n\to\infty} s_n = g.$$

Συνεπώς, από τα προηγούμενα έχουμε ότι

$$\mathbf{E}(r_n(X)s_n(Y)) = \mathbf{E}(r_n(X))\, \mathbf{E}(s_n(Y)) \quad \text{για κάθε } n \in \mathbb{N}.$$

Επειδή οι ακολουθίες $(r_n)_{n\in\mathbb{N}}$ και $(s_n)_{n\in\mathbb{N}}$ είναι αύξουσες, για $n \to \infty$, από το θεώρημα μονότονης σύγκλισης έχουμε ότι

$$\mathbf{E}(f(X)g(Y)) = \mathbf{E}(f(X))\, \mathbf{E}(g(Y)).$$

Δηλαδή η (10.4) ισχύει για τις f, g.

Βήμα 4. Αν οι f, g είναι μετρήσιμες με $\mathbf{E}\,|f(X)|, \mathbf{E}\,|g(Y)| < \infty$, έχουμε

$$\begin{aligned}
\mathbf{E}\{f(X)g(Y)\} &= \mathbf{E}\{(f^+(X) - f^-(X))(g^+(Y) - g^-(Y))\} \\
&= \mathbf{E}\{f^+(X)g^+(Y)\} - \mathbf{E}\{f^+(X)g^-(Y)\} - \mathbf{E}\{f^-(X)g^+(Y)\} + \mathbf{E}\{f^-(X)g^-(Y)\} \\
&= \mathbf{E}(f^+(X))\, \mathbf{E}(g^+(Y)) - \mathbf{E}(f^+(X))\, \mathbf{E}(g^-(Y)) - \mathbf{E}(f^-(X))\, \mathbf{E}(g^+(Y)) + \mathbf{E}(f^-(X))\, \mathbf{E}(g^-(Y)) \\
&= \mathbf{E}(f(X))\, \mathbf{E}(g(Y)).
\end{aligned}$$

Στην τρίτη ισότητα χρησιμοποιήσαμε το ότι η (10.4) ισχύει για μη αρνητικές μετρήσιμες. Επίσης στον τελευταίο υπολογισμό δεν εμφανίζεται πουθενά κάποια απροσδιόριστη μορφή $\infty - \infty$ γιατί οι f, g ικανοποιούν $\mathbf{E}\,|f(X)|, \mathbf{E}\,|g(Y)| < \infty$. Έτσι ο ισχυρισμός αποδείχθηκε.

Επιστρέφουμε τώρα στην απόδειξη του θεωρήματος. Αν $X, Y \geq 0$, εφαρμόζουμε τον ισχυρισμό για

$$f(x) = g(x) = \begin{cases} x & \text{αν } x \in [0, \infty], \\ 0 & \text{αν } x \in [-\infty, 0). \end{cases}$$

Στην περίπτωση που $\mathbf{E}\,|X|, \mathbf{E}\,|Y| < \infty$, εφαρμόζουμε τον ισχυρισμό για $f(x) = g(x) = x$ για κάθε $x \in [-\infty, \infty]$. ■

Παρατήρηση 10.10. Ανάλογα με τα Θεωρήματα 10.6, 10.9 και την Πρόταση 10.8 ισχύουν αν αντί δύο έχουμε περισσότερες ανεξάρτητες τυχαίες μεταβλητές, έστω $X_1, X_2, \dots, X_n$. Για παράδειγμα

$$\mathbf{E}\{f_1(X_1)f_2(X_2)\cdots f_n(X_n)\} = \mathbf{E}\{f_1(X_1)\}\, \mathbf{E}\{f_2(X_2)\} \cdots \mathbf{E}\{f_n(X_n)\}$$

με τις $f_1, f_2, \dots, f_n$ μετρήσιμες και μη αρνητικές ή με $\mathbf{E}\,|f_k(X_k)| < \infty$ για κάθε k. Η απόδειξη των αντίστοιχων αυτών ισχυρισμών γίνεται με επαγωγή.

Υπενθυμίζουμε εδώ ότι για $(X_k)_{1\leq k\leq n}$ τυχαίες μεταβλητές στον ίδιο χώρο πιθανότητας, με πραγματικές τιμές και με $\mathbf{E}(X_k^2) < \infty$ για κάθε $k \in \{1, 2, \dots, n\}$, ισχύει

$$\mathrm{Var}(X_1 + X_2 + \cdots + X_n) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n) + 2\sum_{1\leq i<j\leq n} \mathrm{Cov}(X_i, X_j). \tag{10.5}$$

Η απόδειξη γίνεται όπως ακριβώς την έχουμε δει στις στοιχειώδεις πιθανότητες. Όταν οι $(X_k)_{1\leq k\leq n}$ είναι ανεξάρτητες, το προηγούμενο θεώρημα δίνει ότι όλες οι συνδιακυμάνσεις είναι 0, οπότε

$$\mathrm{Var}(X_1 + X_2 + \cdots + X_n) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n) \tag{10.6}$$

10.2 Ανεξαρτησία και ομαδοποίηση

Αν οι τυχαίες μεταβλητές X, Y, Z, V, W είναι ανεξάρτητες μεταξύ τους, περιμένουμε και οι $X+Y, Z^2W, |V|$ να είναι ανεξάρτητες εφόσον χρησιμοποιούν διαφορετικά ανεξάρτητα συστατικά. Θα διατυπώσουμε ένα θεώρημα που δίνει αποτελέσματα αυτής της μορφής. Προηγουμένως, δίνουμε το αντίστοιχο αποτέλεσμα για σ-άλγεβρες. Η απόξειξή του δίνεται στο Παράρτημα Β′.

Θεώρημα 10.11. *Έστω $(\mathcal{F}_i)_{i\in I}$ ανεξάρτητη οικογένεια σ-αλγεβρών και $\{I_j : j \in J\}$ διαμέριση [1] του συνόλου δεικτών I. Για κάθε $j \in J$ θεωρούμε τη σ-άλγεβρα*

$$\mathcal{G}_j := \sigma\left(\cup_{i\in I_j}\mathcal{F}_i\right).$$

Οι $(\mathcal{G}_j)_{j\in J}$ είναι ανεξάρτητες.

Για το επόμενο θεώρημα, σε πρώτη ανάγνωση καλό είναι να υποθέσει κανείς ότι τα σύνολα I, I_j είναι πεπερασμένα και επομένως οι συναρτήσεις f_j ορίζονται σε χώρους της μορφής $\mathbb{R}^d$. Θεωρούμε ότι ο $\mathbb{R}^{I_j}$ είναι εφοδιασμένος με τη σ-άλγεβρα γινόμενο $\otimes_{i\in I_j}\mathscr{B}(\mathbb{R})$ (όλοι οι όροι του γινομένου είναι ίδιοι). Η απόξειξη και αυτού του θεωρήματος δίνεται στο Παράρτημα Β′.

Θεώρημα 10.12. *Έστω $(X_i)_{i\in I}$ ανεξάρτητες τυχαίες μεταβλητές, $\{I_j : j \in J\}$ διαμέριση του συνόλου δεικτών I, και για κάθε $j \in J$, μετρήσιμη συνάρτηση $f_j : \mathbb{R}^{I_j} \to \mathbb{R}$. Για κάθε $j \in J$ θεωρούμε τη συνάρτηση $Y_j := f_j((X_i)_{i\in I_j}) : \Omega \to \mathbb{R}$. Οι $(Y_j)_{j\in J}$ είναι ανεξάρτητες τυχαίες μεταβλητές.*

Για παράδειγμα, αν οι τυχαίες μεταβλητές $(X_n)_{n\geq 1}$ είναι ανεξάρτητες, τότε και οι $(Y_n)_{n\geq 1}$ με $Y_n := \sum_{i=2^n+1}^{2^{n+1}} X_i$ για κάθε $n \geq 1$ είναι ανεξάρτητες. Όμοια είναι ανεξάρτητα και τα σύνολα $\{X_{2n} + X_{2n+1} > 0\}, n \geq 1$.

10.3 Ανεξαρτησία=Μέτρο γινόμενο

Πρόταση 10.13. *Έστω $X := (X_1, X_2, \ldots, X_n)$ τυχαία μεταβλητή με τιμές στον $\mathbb{R}^n$. Τότε οι $X_1, X_2, \ldots, X_n$ είναι ανεξάρτητες αν και μόνο αν*

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

Απόδειξη. $\Rightarrow$ Η τιμή του μέτρου $\mathbf{P}^X$ σε έναν μετρήσιμο κύλινδρο $A := A_1 \times A_2 \times \cdots \times A_n$ του $\mathbb{R}^n$ είναι

$$\begin{aligned}\mathbf{P}^X(A) &= \mathbf{P}(X_1 \in A_1, \cdots, X_n \in A_n) = \mathbf{P}(X_1 \in A_1)\,\mathbf{P}(X_2 \in A_2)\cdots\mathbf{P}(X_n \in A_n)\\ &= \mathbf{P}^{X_1}(A_1)\,\mathbf{P}^{X_2}(A_2)\cdots\mathbf{P}^{X_n}(A_n).\end{aligned}$$

Στη δεύτερη ισότητα χρησιμοποιήσαμε την ανεξαρτησία των $X_1, \ldots, X_n$. Όμως το $\mathbf{P}^{X_1} \otimes \mathbf{P}^{X_2} \otimes \ldots \otimes \mathbf{P}^{X_n}$ είναι το μοναδικό μέτρο που παίρνει την τιμή $\mathbf{P}^{X_1}(A_1)\,\mathbf{P}^{X_2}(A_2)\cdots\mathbf{P}^{X_n}(A_n)$ στο A. Η ζητούμενη ισότητα έπεται.

$\Leftarrow$ Ελέγχουμε τη σχέση (10.2). Έστω $A_1, A_2, \ldots, A_n \in \mathscr{B}(\mathbb{R})$. Τότε

$$\begin{aligned}\mathbf{P}(X_1 \in A_1, \cdots, X_n \in A_n) &= \mathbf{P}^{X_1} \otimes \mathbf{P}^{X_2} \otimes \ldots \otimes \mathbf{P}^{X_n}(A_1 \times A_2 \times \cdots \times A_n)\\ &= \mathbf{P}^{X_1}(A_1)\,\mathbf{P}^{X_2}(A_2)\cdots\mathbf{P}^{X_n}(A_n)\\ &= \mathbf{P}(X_1 \in A_1)\,\mathbf{P}(X_2 \in A_2)\cdots\mathbf{P}(X_n \in A_n)\end{aligned}$$

Στην πρώτη ισότητα χρησιμοποιήσαμε την υπόθεση και στη δεύτερη τον ορισμό του μέτρου γινομένου. ■

[1] Δηλαδή τα $I_j (j \in J)$, είναι μη κενά, ξένα ανα δύο, και έχουν ένωση το I.

10.4 Κατασκευή ανεξάρτητων τυχαίων μεταβλητών με δεδομένη κατανομή

Υπάρχουν ανεξάρτητες τυχαίες μεταβλητές; Πώς μπορούμε να κατασκευάσουμε τέτοιες και βέβαια τον χώρο πιθανότητας στον οποίο αυτές ορίζονται; Τη λύση σε αυτά τα προβλήματα δίνουν οι χώροι γινόμενο, τους οποίους είδαμε στην Παράγραφο 10.3. Σε αυτή την παράγραφο θα τους χρησιμοποιήσουμε για να κατασκευάσουμε ανεξάρτητες τυχαίες μεταβλητές με επιθυμητές ιδιότητες.

Υπενθυμίζουμε ότι αν $(E,\mathcal{E})$ είναι μετρήσιμος χώρος και $X:\Omega\to E$ τυχαία μεταβλητή, κατανομή της X λέμε το μέτρο πιθανότητας $\mathbf{P}^X$ που ορίζεται στον $(E,\mathcal{E})$ με $\mathbf{P}^X(A)=\mathbf{P}(X^{-1}(A))$ για κάθε $A\in\mathcal{E}$.

Έστω I σύνολο δεικτών και $(E_i,\mathcal{E}_i,\mathbf{P}_i)_{i\in I}$ οικογένεια χώρων πιθανότητας. Θέλουμε να κατασκευάσουμε έναν χώρο πιθανότητας $(\Omega,\mathcal{F},\mathbf{P})$ και τυχαίες μεταβλητές $X_i:\Omega\to E_i$ έτσι ώστε

(α) Η κατανομή της X_i να είναι η $\mathbf{P}_i$ για κάθε $i\in I$.

(β) Οι $(X_i)_{i\in I}$ να είναι οικογένεια ανεξάρτητων τυχαίων μεταβλητών.

Το να κάνουμε ένα από τα (α) ή (β) είναι εύκολο. Αυτό που είναι μη τετριμμένο είναι να κάνουμε και τα δύο μαζί. Και το κατορθώνουμε με χρήση του χώρου γινομένου ως εξής:

Έστω $\Omega=\prod_{i\in I}E_i, \mathcal{F}=\otimes_{i\in I}\mathcal{E}_i$, και $\mathbf{P}$ το μέτρο γινόμενο των $\mathbf{P}_i$, $i\in I$. Για κάθε $r\in I$ ορίζουμε $X_r:\Omega\to E_r$ ως

$$X_r((\omega_i)_{i\in I})=\omega_r,$$

δηλαδή η X_r είναι η προβολή στη r-συντεταγμένη.

Εύκολα βλέπουμε ότι η X_r είναι τυχαία μεταβλητή γιατί για $A\in E_r$ έχουμε

$$X_r^{-1}(A)=\prod_{i\in I}A_i,\ \text{ με } A_i=\begin{cases}\Omega & \text{αν } i\neq r,\\ A & \text{αν } i=r.\end{cases}$$

Άρα $X_r^{-1}(A)\in\mathcal{F}$ ως μετρήσιμος κύλινδρος.

Θα δείξουμε τώρα ότι πράγματι αυτή η κατασκευή ικανοποιεί τις απαιτήσεις του προβλήματος.

Πρόταση 10.14. *Έστω $(\Omega,\mathcal{F},\mathbf{P})$, $(E_i,\mathcal{E}_i,\mathbf{P}_i)_{i\in I}$ όπως πριν και $X_r:\Omega\to E_r$, $r\in I$, η r-προβολή. Τότε*

(i) Η X_r έχει κατανομή $\mathbf{P}_r$.

(ii) Οι $(X_r)_{r\in I}$ είναι ανεξάρτητες.

Απόδειξη. (*i*) Έστω $A\in E_r$. Έχουμε,

$$\mathbf{P}^{X_r}(A)=\mathbf{P}(X_r^{-1}(A))=\mathbf{P}\Big(\prod_{i\in I}A_i\Big)=\prod_{i\in I}\mathbf{P}_i(A_i)=\mathbf{P}_r(A),$$

γράφοντας το $X_r^{-1}(A)$ ως μετρήσιμο κύλινδρο όπως προηγουμένως. Άρα, $\mathbf{P}^{X_r}=\mathbf{P}_r$.

(*ii*) Έστω $J=\{j_1,j_2,\dots,j_n\}\subset I$, και $B_{j_1}\in\mathcal{E}_{j_1},B_{j_2}\in\mathcal{E}_{j_2},\dots,B_{j_n}\in\mathcal{E}_{j_n}$. Τότε

$$\mathbf{P}(\{X_{j_1}\in B_{j_1}\}\cap\{X_{j_2}\in B_{j_2}\}\cap\cdots\cap\{X_{j_n}\in B_{j_n}\})=\mathbf{P}\Big(\prod_{i\in I}A_i\Big),$$

όπου

$$A_i=\begin{cases}\Omega_i & \text{αν } i\in I\backslash J,\\ B_i, & \text{αν } i\in J.\end{cases}$$

Από τον ορισμό του $\mathbf{P}$ έχουμε ότι

$$\begin{aligned}\mathbf{P}\Big(\prod_{i\in I}A_i\Big)&=\prod_{i\in I}\mathbf{P}(A_i)=\prod_{i\in J}\mathbf{P}(A_i)=\mathbf{P}_{j_1}(B_{j_1})\mathbf{P}_{j_2}(B_{j_2})\dots\mathbf{P}_{j_n}(B_{j_n})\\&=\mathbf{P}(X_{j_1}\in B_{j_1})\mathbf{P}(X_{j_2}\in B_{j_2})\cdots\mathbf{P}(X_{j_n}\in B_{j_n}).\end{aligned}$$

Συνεπώς οι $\{X_r:r\in I\}$ είναι ανεξάρτητες. ∎

Μια συνέπεια της Πρότασης 10.14 είναι ότι για δεδομένη κατανομή $\mathbf{Q}$ και σύνολο I υπάρχει σύνολο ανεξάρτητων και ισόνομων τυχαίων μεταβλητών $(X_i)_{i\in I}$ που καθεμία έχει κατανομή $\mathbf{Q}$.

Ασκήσεις

Στις ασκήσεις πιο κάτω οι τυχαίες μεταβλητές που εμφανίζονται παίρνουν τιμές στο $\mathbb{R}$.

10.1 Έστω $\mathcal{F}_1, \mathcal{F}_2 \subset \mathcal{F}$ σ-άλγεβρες στον $(\Omega, \mathcal{F}, \mathbf{P})$. Αν για κάθε $A \in \mathcal{F}_1$ ισχύει ότι $\mathbf{P}(A) = 0$ ή $\mathbf{P}(A) = 1$, να δείξετε ότι η $\mathcal{F}_1$ είναι ανεξάρτητη της $\mathcal{F}_2$.

10.2 Έστω X, Y ανεξάρτητες τυχαίες μεταβλητές ώστε $\mathbf{P}(X = Y) = 1$. Να δειχθεί ότι υπάρχει $c \in \mathbb{R}$ ώστε $\mathbf{P}(X = c) = 1$, δηλαδή με πιθανότητα 1 η X είναι σταθερή τυχαία μεταβλητή.

10.3 Έστω X, Y ανεξάρτητες ισόνομες τυχαίες μεταβλητές με πεπερασμένη διασπορά σ^2. Να δειχθεί ότι $\mathbf{E}\{(X - Y)^2/2\} = \sigma^2$.

10.4 Έστω X, Y τυχαίες μεταβλητές με πεπερασμένη δεύτερη ροπή (δηλαδή $\mathbf{E}(X^2), \mathbf{E}(Y^2) < \infty$) ώστε οι $X, X - Y$ να είναι ανεξάρτητες αλλά και οι $Y, X - Y$ να είναι ανεξάρτητες. Να δειχθεί ότι υπάρχει $c \in \mathbb{R}$ ώστε $\mathbf{P}(X - Y = c) = 1$. [Υποδ.: $Y = X - (X - Y)$].

10.5 Έστω X, Y ανεξάρτητες τυχαίες μεταβλητές με $\mathbf{E}\,|X + Y| < \infty$. Να δειχθεί ότι $\mathbf{E}\,|X| < \infty$ και $\mathbf{E}\,|Y| < \infty$.

10.6* Έστω X τυχαία μεταβλητή και $f, g : \mathbb{R} \to \mathbb{R}$ αύξουσες συναρτήσεις. Υποθέτουμε ότι οι μέσες τιμές $\mathbf{E}\{f(X)g(X)\}$, $\mathbf{E}\{f(X)\}, \mathbf{E}\{g(X)\}$ ορίζονται και είναι πραγματικοί αριθμοί. Να δειχθεί ότι

$$\mathbf{E}\{f(X)g(X)\} \ge \mathbf{E}\{f(X)\}\,\mathbf{E}\{g(X)\}.$$

Αυτή είναι ειδική περίπτωση της ανισότητας FKG (Fortuin-Kasteleyn-Ginibre. 1971).

10.7 Έστω $X_1, X_2, \ldots, X_n$ ανεξάρτητες τυχαίες μεταβλητές με $X_i \sim \exp(a_i)$ για κάθε $i = 1, 2, \ldots, n$ όπου $a_1, a_2, \ldots, a_n \in (0, \infty)$ είναι σταθερές. Ποια η κατανομή της $m := \min\{X_1, X_2, \ldots, X_n\}$;

10.8 Έστω $(X_n)_{n\in\mathbb{N}}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών έτσι ώστε $\mathbf{E}(X_1) = 0, \mathbf{E}(X_1^2) = 1$ και υπάρχει σταθερά $M > 0$ ώστε $|X_i| \le M$ στο Ω για κάθε $i \in \mathbb{N}$. Θέτουμε $S_n = \sum_{k=1}^n X_k$ για κάθε $n \ge 1$. Να δείξετε ότι

(α) $\mathbf{E}(S_n^2) = n$.

(β) $\mathbf{E}(S_n^4) = 3n^2 - 2n$.

(γ) $\frac{S_n}{n} \xrightarrow{\mathbf{P}} 0$ καθώς $n \to \infty$.

[Υποδειξη για το (γ): Χρησιμοποιήστε την ανισότητα Markov και το (α).]

10.9 Έστω $(X_n)_{n\ge1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών καθεμία με την ομοιόμορφη κατανομή στο $(0, 1)$. Για κάθε $n \ge 1$ θεωρούμε τις τυχαίες μεταβλητές

$$m_n = \min\{X_1, X_2, \ldots, X_n\},$$
$$M_n = \max\{X_1, X_2, \ldots, X_n\}.$$

Να δειχθεί ότι $m_n \to 0$ και $M_n \to 1$ κατά πιθανότητα καθώς $n \to \infty$.

10.10 Για κάθε πίνακα $A \in \mathbb{C}^{n\times n}$, η ορίζουσα του και η permanent του ορίζονται από τους τύπους

$$\det(A) := \sum_{\pi\in S_n} \operatorname{sgn}(\pi) a_{1,\pi(1)} a_{2,\pi(2)} \cdots a_{n,\pi(n)}, \tag{10.7}$$

$$\operatorname{per}(A) := \sum_{\pi\in S_n} a_{1,\pi(1)} a_{2,\pi(2)} \cdots a_{n,\pi(n)}. \tag{10.8}$$

S_n είναι το σύνολο των μεταθέσεων στο $\{1, 2, \ldots, n\}$ και $\operatorname{sgn}(\pi)$ είναι το πρόσημο της μετάθεσης π (1 για άρτια μετάθεση και -1 για περιττή).

Έστω $A = (a_{i,j})_{1\le i,j\le n} \in \mathbb{C}^n$ δεδομένος πίνακας. Θεωρούμε ανεξάρτητες τυχαίες μεταβλητές $\{u_{i,j} : 1 \le i, j \le n\}$ με τιμές στο $\mathbb{R}$ ώστε καθεμία έχει μέση τιμή 0 και διασπορά 1. Ορίζουμε τον πίνακα $B = (u_{i,j}\sqrt{a_{i,j}})_{1\le i,j\le n}$ όπου $\sqrt{a_{i,j}}$ είναι μία τετραγωνική ρίζα του $a_{i,j}$. Να δειχθεί ότι

$$\mathbf{E}(\det(B^2)) = \operatorname{per}(A).$$

11

Τα λήμματα Borel-Cantelli και ο νόμος 0-1 του Kolmogorov

Τα βασικότερα εργαλεία για την απόδειξη θεωρημάτων που αφορούν τη σχεδόν βέβαιη σύγκλιση είναι δύο απλά αποτελέσματα, τα δύο λήμματα Borel-Cantelli, τα οποία θα δούμε σε αυτό το κεφάλαιο. Επίσης θα δούμε τον νόμο 0-1 του Kolmogorov ο οποίος είναι πολύ χρήσιμος όταν θέλουμε να δείξουμε ότι μια ιδιότητα ισχύει με πιθανότητα 1. Σύμφωνα με αυτόν τον νόμο, αν η ιδιότητα έχει μια συγκεκριμένη μορφή, τότε αναγκαστικά έχει πιθανότητα 0 ή 1. Επομένως για να δείξουμε ότι ισχύει με πιθανότητα 1, αρκεί να δείξουμε ότι ισχύει με θετική πιθανότητα. Και το τελευταίο πολλές φορές είναι σημαντικά ευκολότερο.

11.1 Τα λήμματα Borel-Cantelli

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

$$\limsup_{n\geq 1} A_n = \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k = \{\omega \in \Omega : \omega \text{ ανήκει σε άπειρα } A_n\}.$$

Πρόταση 11.1 (Πρώτο λήμμα Borel-Cantelli). *Έστω $(A_n)_{n\geq 1}$ ακολουθία ενδεχομένων στον $(\Omega, \mathcal{F}, \mathbf{P})$. Άν $\sum_{n=1}^{\infty} \mathbf{P}(A_n) < \infty$, τότε*

$$\mathbf{P}(\limsup_{n\geq 1} A_n) = 0.$$

Απόδειξη. Θεωρούμε την τυχαία μεταβλητη $X = \sum_{n=1}^{\infty} \mathbf{1}_{A_n}$. Τότε

$$\limsup_{n\geq 1} A_n = \{\omega \in \Omega : \omega \text{ ανήκει σε άπειρα } A_n\} = \{\omega \in \Omega : X(\omega) = \infty\} = \{X = \infty\},$$

και

$$\mathbf{E}(X) = \sum_{n=1}^{\infty} \mathbf{E}(\mathbf{1}_{A_n}) = \sum_{n=1}^{\infty} \mathbf{P}(A_k) < \infty.$$

Επειδή $\mathbf{E}(X) < \infty$ και $X \geq 0$, η Πρόταση 6.14(iii) δίνει ότι $\mathbf{P}(X = \infty) = 0$, δηλαδή $\mathbf{P}(\limsup_{n\geq 1} A_n) = 0$. ■

Λήμμα 11.2. *Αν $A_1, A_2, \ldots, A_n$, $n \geq 2$, ανεξάρτητα ενδεχόμενα στον $(\Omega, \mathcal{F}, \mathbf{P})$, τότε και τα $A_1^c, A_2^c, \ldots, A_n^c$ είναι ανεξάρτητα.*

Απόδειξη. Δείχνουμε πρώτα τον εξής ισχυρισμό.

Ισχυρισμός: Τα $A_1, A_2, \ldots, A_{n-1}, A_n^c$ είναι ανεξάρτητα.

Πράγματι, για $k \leq n$, έστω δείκτες $1 \leq i_1 < i_2 < \ldots < i_k \leq n$.

- Αν $i_k < n$, τότε τα $A_{i_1}, A_{i_2}, \ldots, A_{i_k}$ είναι ανεξάρτητα από υπόθεση. Επομένως

$$\mathbf{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_{k-1}} \cap A_{i_k}) = \mathbf{P}(A_{i_1})\,\mathbf{P}(A_{i_2}) \cdots \mathbf{P}(A_{i_k}) \tag{11.1}$$

- Αν $i_k = n$, τότε

$$\begin{aligned}\mathbf{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_{k-1}} \cap A_n^c) &= \mathbf{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_{k-1}}) - \mathbf{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_{k-1}} \cap A_n)\\ &= \mathbf{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_{k-1}}) - \mathbf{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_{k-1}})\,\mathbf{P}(A_n)\\ &= \mathbf{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_{k-1}})(1 - \mathbf{P}(A_n))\\ &= \mathbf{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_{k-1}})\,\mathbf{P}(A_n^c)\\ &= \mathbf{P}(A_{i_1})\,\mathbf{P}(A_{i_2})\cdots\mathbf{P}(A_n^c).\end{aligned}$$

Στη δεύτερη και στην τελευταία ισότητα χρησιμοποιήσαμε την ανεξαρτησία των $A_1, A_2, \ldots, A_{i_k}$.

Άρα και στις δύο περιπτώσεις η πιθανότητα της τομής ισούται με το γινομενο των πιθανοτήτων και ο ισχυρισμός αποδείχθηκε.
Χρησιμοποιώντας τον ισχυρισμό, δείχνουμε επαγωγικά ότι τα $A_1^c, A_2^c, \ldots, A_n^c$ είναι ανεξάρτητα. ■

Το δεύτερο λήμμα Borel-Cantelli αφορά την περίπτωση που η σειρά $\sum_{n=1}^{\infty} \mathbf{P}(A_n)$ απειρίζεται. Όμως τώρα υποθέτουμε επιπλέον ότι τα $(A_n)_{n\geq 1}$ είναι ανεξάρτητα. Η ακριβής διατύπωση έχει ως εξής.

Πρόταση 11.3 (Δεύτερο λήμμα Borel-Cantelli)**.** *Έστω $(A_n)_{n\geq 1}$ ακολουθία ανεξάρτητων ενδεχομένων στον $(\Omega, \mathcal{F}, \mathbf{P})$. Άν $\sum_{n=1}^{\infty} \mathbf{P}(A_n) = \infty$, τότε*

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

Απόδειξη. Θα δείξουμε ότι $\mathbf{P}(\{\limsup_{n\geq 1} A_n\}^c) = 0$. Έχουμε ότι

$$\mathbf{P}(\{\limsup_{n\geq 1} A_n\}^c) = \mathbf{P}(\cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k^c) = \lim_{n\to\infty} \mathbf{P}(\cap_{k=n}^{\infty} A_k^c)$$

εφόσον η ακολουθία $(B_n)_{n\geq 1}$ με $B_n = \cap_{k=n}^{\infty} A_k^c$ για κάθε $n \geq 1$ είναι αύξουσα. Για δεδομένο $n \geq 1$ έχουμε

$$\begin{aligned}\mathbf{P}(\cap_{k=n}^{\infty} A_k^c) &= \lim_{m\to\infty} \mathbf{P}(\cap_{k=n}^{m} A_k^c) = \lim_{m\to\infty} \prod_{k=n}^{m} \mathbf{P}(A_k^c)\\ &= \lim_{m\to\infty} \prod_{k=n}^{m} (1 - \mathbf{P}(A_n)) \leq \lim_{m\to\infty} \prod_{k=n}^{m} e^{-\mathbf{P}(A_k)}\\ &= \lim_{m\to\infty} e^{-\sum_{k=n}^{m} \mathbf{P}(A_k)} = e^{-\sum_{k=n}^{\infty} \mathbf{P}(A_k)} = e^{-\infty} = 0.\end{aligned}$$

Στη δεύτερη ισότητα χρησιμοποιήσαμε το Λήμμα 11.2, ενώ η ανισότητα προκύπτει από την $1 + x \leq e^x$ για κάθε $x \in \mathbb{R}$. Άρα $\mathbf{P}((\limsup_{n\geq 1} A_n)^c) = 0$ και το ζητούμενο αποδείχθηκε. ■

Αν παραλείψουμε την υπόθεση της ανεξαρτησίας από τη διατύπωση του δεύτερου λήμματος Borel-Cantelli, τότε το συμπέρασμά του ενδέχεται να μην ισχύει. Για παράδειγμα, αν πάρουμε λ το μέτρο Lebesgue στον $\Omega := [0, 1]$ (που είναι μέτρο πιθανότητας) και $A_n = (0, 1/n)$ για κάθε $n \geq 1$, τότε $\lambda(A_n) = 1/n$, και η σειρά $\sum_{n=1}^{\infty} \lambda(A_n) = \infty$, ενώ $\limsup_{n\geq 1} A_n = \emptyset$. Κανένας αριθμός δεν ανήκει σε άπειρα από τα A_n. Βέβαια τα $(A_n)_{n\geq 1}$ δεν είναι ανεξάρτητα.

Παράδειγμα 11.4. Θεωρούμε το πείραμα ρίψης ενός νομίσματος άπειρες (αριθμήσιμες) φορές που φέρνει «κορώνα» (K) με πιθανότητα $p \in (0, 1)$. Θα υπολογίσουμε την πιθανότητα $\mathbf{P}$(K έρχεται απειρες φορές).

Ο χώρος πιθανότητας του πειράματος είναι ο χώρος γινόμενο $(\Omega, \mathcal{F}, \mathbf{P})$ των $(\Omega_n, \mathcal{F}_n, \mathbf{P}_n)_{n\geq 1}$, όπου, για κάθε $n \geq 1$, $\Omega_n = \{K, \Gamma\}$ (Γ = το ενδεχόμενο «γράμματα») , $\mathcal{F}_n = \mathscr{P}(\Omega_n)$ και $\mathbf{P}_n = \mathbf{P}^{(p)}$ ($\mathbf{P}^{(p)}$ το μέτρο πιθανότητας με $\mathbf{P}^{(p)}(\{K\}) = p$). Για $n \geq 1$, θεωρούμε το ενδεχόμενο A_n = {έρχεται K στη n ρίψη} και την τυχαία μεταβλητή $X_n : \Omega \to \{K, \Gamma\}$ με $X_n(\omega) = \omega_n$ = το αποτέλεσμα της n ρίψης.

Γνωρίζουμε ότι οι $(X_n)_{n\in\mathbb{N}}$ είναι ανεξάρτητες (Πρόταση 10.8) και, εφόσον $A_n = X_n^{-1}(\{K\})$, έχουμε ότι τα $(A_n)_{n\geq 1}$ είναι ανεξάρτητα. Επιπλέον, $\mathbf{P}(A_n) = p$, άρα

$$\sum_{n=1}^{\infty} \mathbf{P}(A_n) = \infty.$$

Από το 2ο λήμμα Borel-Cantelli έχουμε ότι $\mathbf{P}(\limsup_{n\geq 1} A_n) = 1$, δηλαδή $\mathbf{P}(K$ έρχεται άπειρες φορές$) = 1$.

Παρατήρηση 11.5. Συνήθως για το $\limsup A_n$ χρησιμοποιούμε τον συμβολισμό

$$\{A_n \text{ συμβαίνει άπειρες φορές}\},$$

και το σκεπτικό του είναι το εξής. Ο χώρος πιθανότητας $(\Omega, \mathcal{F}, \mathbf{P})$ μοντελοποιεί ένα πείραμα, και μια πραγματοποίηση του πειράματος είναι ένα $\omega \in \Omega$ (στο προηγούμενο παράδειγμα η πραγματοποίηση είναι ένα $\omega \in \{K, \Gamma\}^{\mathbb{N}^+}$). Πολλές φορές παρ' όλ' αυτά, αντιλαμβανόμαστε ότι το πείραμα γίνεται σε πολλά στάδια (π.χ. ρίχνουμε ένα νόμισμα άπειρες φορές, τη μία μετά την άλλη και όχι μονομιάς). Ορίζουμε A_n να είναι ένα σύνολο που αφορά το στάδιο n, και αυτό πραγματοποιείται αν $\omega \in A_n$. Τότε το να συμβεί το A_n άπειρες φορές (δηλαδή για άπειρα n) σημαίνει ακριβώς ότι η πραγματοποίηση ω ανήκει σε άπειρα από τα $\{A_n : n \geq 1\}$.

Παράδειγμα 11.6. Έστω $(X_n)_{n\geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με $X_1 \sim Exp(1)$, δηλαδή με πυκνότητα $f(x) = e^{-x}\mathbf{1}_{x>0}$. Θα δείξουμε ότι

$$\overline{\lim_{n\to\infty}} \frac{X_n}{\log n} = 1 \text{ με πιθανότητα } 1.$$

[Μεταφράζοντας τη σημασία του $\overline{\lim}$, αυτό σημαίνει ότι για κάθε $\varepsilon > 0$ ισχύει $X_n/\log n > 1+\varepsilon$ για *πεπερασμένα* το πλήθος n, ενώ ισχύει $X_n/\log n > 1-\varepsilon$ για *άπειρα* το πλήθος n. Άμεσα βλέπουμε τη συνάφεια των λημμάτων Borel-Cantelli με το ερώτημα.]
Για κάθε $n \geq 1$ και $r > 0$, θέτουμε $A_n^{(r)} = \{X_n \geq r\log n\}$. Τότε,

$$\mathbf{P}(A_n^{(r)}) = \mathbf{P}(X_n \geq r\log n) = e^{-r\log n} = \frac{1}{n^r}.$$

Έστω $r > 1$. Τότε

$$\sum_{n=1}^{\infty} \mathbf{P}(A_n) < \infty,$$

και από το πρώτο λήμμα Borel-Cantelli έχουμε ότι $\mathbf{P}(\limsup_{n\geq 1} A_n^{(r)}) = 0$, άρα

$$\mathbf{P}\left(\overline{\lim_{n\to\infty}} \frac{X_n}{\log n} \geq r\right) = 0.$$

Συνεπώς, θέτοντας $C_r = \left\{\overline{\lim_{n\to\infty}} \frac{X_n}{\log n} \geq r\right\}$, έχουμε ότι

$$\left\{\overline{\lim_{n\to\infty}} \frac{X_n}{\log n} > 1\right\} = \cup_{k=1}^{\infty} C_{1+\frac{1}{k}},$$

άρα $\mathbf{P}\left(\overline{\lim_{n\to\infty}} \frac{X_n}{\log n} > 1\right) = 0$ εφόσον $\mathbf{P}(C_{1+\frac{1}{k}}) = 0$ για κάθε $k \geq 1$. Επομένως,

$$\mathbf{P}\left(\overline{\lim_{n\to\infty}} \frac{X_n}{\log n} \leq 1\right) = 1. \tag{11.2}$$

Έστω $r = 1$. Τότε, εφόσον τα $(A_n)_{n\geq 1}$ είναι ανεξάρτητα (οι $(X_n)_{n\geq 1}$ είναι ανεξάρτητες) και

$$\sum_{n=1}^{\infty} \mathbf{P}(A_n^{(1)}) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty,$$

από το 2ο λήμμα Borel-Cantelli, έχουμε ότι $\mathbf{P}(\limsup_{n\geq 1} A_n^{(1)}) = 1$. Συνεπώς, για $\omega \in \limsup_{n\geq 1} A_n^{(1)}$ ισχύει ότι $X_n(\omega) \geq \log n$ για άπειρα $n \geq 1$. Δηλαδή,

$$\overline{\lim_{n\to\infty}} \frac{X_n(\omega)}{\log n} \geq 1,$$

από το οποίο προκύπτει ότι $\mathbf{P}\left(\overline{\lim_{n\to\infty}} \frac{X_n}{\log n} \geq 1\right) = 1$. Η τελευταία ισότητα μαζί με την (11.2) δίνουν το ζητούμενο.

Παρατήρηση 11.7. Στο προηγούμενο παράδειγμα, ορίσαμε κάποια σύνολα (τα $A_n^{(r)}$ και C_r) και μιλήσαμε για τις πιθανότητες τους. Τυπικά θα έπρεπε προηγουμένως να δείξουμε ότι είναι στοιχεία της $\mathcal{F}$, δηλαδή είναι μετρήσιμα σύνολα. Δεν το κάναμε, ούτε θα το κάνουμε στο εξής για τα σύνολα που θα ορίζουμε. Όλα θα είναι μετρήσιμα, και η τυπική δικαιολόγηση αφήνεται στον αναγνώστη.

11.2 Ο νόμος 0-1 του Kolmogorov*

Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας, $((E_n, \mathcal{E}_n))_{n\geq 1}$ μετρήσιμοι χώροι και $X_n : \Omega \to E_n$, $n \geq 1$, τυχαίες μεταβλητές. Για $n \geq 1$ θέτουμε

$$\mathscr{C}_n := \sigma(\{X_k : k \geq n+1\}),$$

τη σ-άλγεβρα που παράγεται από τις $X_{n+1}, X_{n+2}, \ldots$

Ορισμός 11.8. Η **τελική σ-άλγεβρα** που παράγεται από τις $(X_n)_{n\geq 1}$ ορίζεται ως

$$\mathscr{C}_\infty = \cap_{n=1}^{\infty} \mathscr{C}_n.$$

Τι σημαίνει πρακτικά για ένα ενδεχόμενο A να ανήκει στην $\mathscr{C}_\infty$; Σημαίνει ότι για κάθε $n \geq 1$ η πραγματοποίση ή όχι του A δεν εξαρτάται από την τιμή που παίρνουν οι πρώτες n από τις X_i. Δηλαδή οποιοδήποτε δεδομένο πεπερασμένο πλήθος από τις X_i δεν επηρεάζει την πραγματοποίηση του A. Αυτή η ασαφής περιγραφή θα γίνει ξεκάθαρη στο επόμενο παράδειγμα.

Παράδειγμα 11.9. Έστω $(X_n)_{n\geq 1}$ όπως προηγουμένως με τιμές στο $\mathbb{R}$. Θεωρούμε τα σύνολα

$$A = \{\omega : \overline{\lim}_{k\to\infty} X_k(\omega) \geq 1\}, \quad B = \left\{\omega : \overline{\lim}_{n\to\infty} \tfrac{1}{n} \textstyle\sum_{k=1}^{n} X_k(\omega) \geq 0\right\},$$

$$\Gamma = \{\omega : \inf_{n\geq 1} X_n(\omega) \leq 0\}, \qquad \Delta = \left\{\omega : \sum_{n=1}^{\infty} X_n(\omega) \leq 10\right\}.$$

Τα A, B ανήκουν στην $\mathscr{C}_\infty$, ενώ τα Γ, Δ δεν ανήκουν σε αυτήν αναγκαστικά.

Πράγματι, όσον αφορά τα A, B, έχουμε ότι για κάθε $m \geq 1$ ισχύει

$$A = \{\omega : \overline{\lim_{k\to\infty}} X_k(\omega) \geq 1\} = \{\omega : \overline{\lim_{k\to\infty}} X_{m+1+k}(\omega) \geq 1\} \in \mathscr{C}_m$$

και

$$B = \left\{\omega : \overline{\lim_{n\to\infty}} \frac{1}{n} \sum_{k=1}^{n} X_k(\omega) \geq 0\right\} = \left\{\omega : \overline{\lim_{n\to\infty}} \left(\frac{1}{n} \sum_{k=1}^{m} X_k(\omega) + \frac{1}{n} \sum_{k=m+1}^{n} X_k(\omega)\right) \geq 0\right\}$$
$$= \left\{\omega : \overline{\lim_{n\to\infty}} \frac{1}{n} \sum_{k=m+1}^{n} X_k(\omega) \geq 0\right\} \in \mathscr{C}_m$$

εφόσον (προσοχή, το άθροισμα δεν εξαρτάται από το n)

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{m} X_k(\omega) = 0.$$

Άρα $A, B \in \mathscr{C}_\infty$.

Για τα Γ και Δ αρκεί να παρατηρήσουμε ότι για συγκεκριμένες επιλογές των τυχαίων μεταβλητών $(X_n)_{n\geq 1}$, τα σύνολα Γ και Δ εξαρτώνται από την τιμή του X_1. Για παράδειγμα, στο Γ, αν οι $(X_n)_{n\geq 1}$ είναι ανεξάρτητες και $X_1 \sim N(0,1)$, $X_k \sim U(1,3)$ (ομοιόμορφη στο $(1,3)$) για κάθε $k \geq 2$, τότε $\Gamma = \{X_1 \leq 0\}$, το οποίο έχει πιθανότητα 1/2. Αν υποθέσουμε ότι $\Gamma \in \mathscr{C}_\infty$, τότε $\Gamma \in \sigma(\{X_k : k \geq 2\})$, και αυτή η σ-άλγεβρα είναι ανεξάρτητη από την $\sigma(X_1)$, η οποία περιέχει το Γ. Άρα το Γ είναι ανεξάρτητο από το εαυτό του, δηλαδή $\mathbf{P}(\Gamma \cap \Gamma) = \mathbf{P}(\Gamma)\,\mathbf{P}(\Gamma)$, το οποίο δίνει ότι $\mathbf{P}(\Gamma) \in \{0,1\}$. Άτοπο.

Πριν κάνουμε τις παραπάνω αποδείξεις για τα A, B, βλέπουμε ότι για δεδομένο ω (δηλαδή για μία πραγματοποίηση του πειράματος), το αν $\omega \in A$, δηλαδή το αν το A πραγματοποιήθηκε, δεν εξαρτάται από τις πρώτες τιμές της $(X_n(\omega))_{n\geq 1}$. Για το A, η τιμή του $\overline{\lim}$ μένει η ίδια αν αλλάξουμε, π.χ., τους πρώτους 1000 όρους της ακολουθίας. Το ίδιο συμβαίνει και με το B. Αυτή η παρατήρηση μας πείθει ότι $A, B \in \mathscr{C}_\infty$ και τη χρησιμοποιούμε στην τυπική απόδειξη.

Το βασικό αποτέλεσμα αυτής της παραγράφου αφορά την τελική σ-άλγεβρα ακολουθίας *ανεξάρτητων* τυχαίων μεταβλητών.

Θεώρημα 11.10 (Νόμος 0-1 του Kolmogorov)**.** *Έστω $(X_n)_{n\geq 1}$ ανεξάρτητες τυχαίες μεταβλητές στον $(\Omega, \mathcal{F}, \mathbf{P})$ και $\mathscr{C}_\infty$ η τελική σ-άλγεβρά τους. Αν $C \in \mathscr{C}_\infty$, τότε $\mathbf{P}(C) = 0$ ή 1.*

Η απόξειξη του θεωρήματος δίνεται στο Παράρτημα Β΄.

Το θεώρημα χρησιμοποιείται συνήθως για να δείξουμε ότι ένα γεγονός έχει πιθανότητα 1. Δείχνουμε ότι ανήκει στην τελική σ-άλγεβρα και έχει θετική πιθανότητα. Μια τέτοια χρήση γίνεται στην Άσκηση 11.16. Το να δείξει κανείς ότι το γεγονός του ερωτήματος (β) της άσκησης έχει θετική πιθανότητα είναι πολύ απλό, ενώ το να δείξει ότι έχει πιθανότητα 1 είναι αρκετά περίπλοκο αν δεν χρησιμοποιήσουμε τον νόμο 0-1 του Kolmogorov.

Άμεση συνέπεια του νόμου 0-1 είναι το ακόλουθο πόρισμα.

Πόρισμα 11.11. *Έστω $(X_n)_{n\in\mathbb{N}}$ ανεξάρτητες τυχαίες μεταβλητές στον $(\Omega, \mathcal{F}, \mathbf{P})$ και $X : \Omega \to [-\infty, \infty]$ τυχαία μεταβλητή $\mathscr{C}_\infty$-μετρήσιμη. Τότε η X είναι σταθερή με πιθανοτητα 1.*

Απόδειξη. Εφόσον η X είναι $\mathscr{C}_\infty$-μετρήσιμη, τα σύνολα $\{X = -\infty\}, \{X = \infty\}$ είναι στοιχεία της $\mathscr{C}_\infty$, και επομένως έχουν πιθανότητα 0 ή 1. Άν κάποιο από αυτά έχει πιθανότητα 1, δείχθηκε το ζητούμενο. Διαφορετικά, έχουμε ότι η X παίρνει τιμές στο $\mathbb{R}$. Σε αυτή την περίπτωση, για τη συνάρτηση κατανομής της, F, το Θεώρημα 11.10 δίνει ότι για κάθε $x \in \mathbb{R}$ ισχύει

$$F(x) = \mathbf{P}(X \leq x) = \mathbf{P}(X^{-1}((-\infty, x])) \in \{0,1\}. \tag{11.3}$$

Ξέρουμε όμως ότι η $\mathbf{F}$ είναι αύξουσα, δεξιά συνεχής και $\mathbf{F}(-\infty) = 0$, $\mathbf{F}(\infty) = 1$. Αυτές οι ιδιότητες μαζί με την (11.3) συνεπάγονται ότι υπάρχει $c \in \mathbb{R}$ τέτοιο ώστε

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

Συνεπώς $\mathbf{P}(X = c) = \mathbf{F}(c) - \mathbf{F}(c-) = 1$, δηλαδή η X ισούται με τη σταθερά c με πιθανότητα 1. ∎

Παρατήρηση 11.12. Στην απόδειξη του Πορίσματος 11.11, από το ότι η X είναι $\mathscr{C}_\infty$-μετρήσιμη χρησιμοποιήσαμε μόνο ότι όλα τα σύνολα της $\mathscr{C}_\infty$ έχουν πιθανότητα 0 ή 1. Έτσι, το ίδιο επιχείρημα δίνει ότι αν η X είναι τυχαία μεταβλητή στον $(\Omega, \mathcal{F}, \mathbf{P})$ με τιμές στο $[-\infty, \infty]$ και $\mathcal{A} \subset \mathcal{F}$ σ-άλγεβρα τέτοια ώστε η X να είναι $\mathcal{A}/\mathscr{B}([-\infty, \infty])$-μετρήσιμη και $\mathbf{P}(A) \in \{0,1\}$ για κάθε $A \in \mathcal{A}$, τότε η X είναι σταθερή με πιθανότητα 1.

Επιστρέφουμε στο Παράδειγμα 11.6. Εκεί η $Z := \overline{\lim}_{n\to\infty}(X_n/\log n)$ είναι μια $\mathscr{C}_\infty$-μετρήσιμη τυχαία μεταβλητή με τιμές στο $[0,\infty]$ και οι $(X_n)_{n\geq 1}$ είναι ανεξάρτητες. Το Πόρισμα 11.11 εφαρμόζεται. Άρα εκ των προτέρων ξέρουμε ότι η Z είναι σταθερή με πιθανότητα 1.

Ασκήσεις

11.1* Έστω $(A_n)_{n\geq 1}$ ακολουθία ανεξάρτητων ενδεχομένων με $\mathbf{P}(A_n) < 1$ για κάθε $n \geq 1$ και $\mathbf{P}(\cup_{n=1}^{\infty} A_n) = 1$. Να δείξετε ότι $\sum_{n=1}^{\infty} \mathbf{P}(A_n) = \infty$.

11.2 Έστω $(X_n)_{n\geq 1}$ ακολουθία τυχαίων μεταβλητών με $\mathbf{P}(X_n \neq 0) = \frac{1}{n^2}$ για κάθε $n \geq 1$. Να δείξετε ότι με πιθανότητα 1 για κάθε $\omega \in \Omega$, υπάρχει $n(\omega) \in \mathbb{N}$ ώστε $X_n(\omega) = 0$ για κάθε $n \geq n(\omega)$. (Συνεπώς $X_n \to 0$ με πιθανότητα 1.)

11.3 Στην Άσκηση 10.9, να δειχθεί ότι επίσης $m_n \to 0$ και $M_n \to 1$ σχεδόν βέβαια καθώς $n \to \infty$.

11.4 Έστω $(X_n)_{n\geq 1}$ ακολουθία ανεξάρτητων τυχαίων μεταβλητών ώστε

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

Να δειχθεί ότι

(α) $X_n \to 0$ κατά πιθανότητα καθώς $n \to \infty$,

(β) αλλά δεν ισχύει $X_n \to 0$ σχεδόν βεβαίως. Μάλιστα $\mathbf{P}(\lim_{n\to\infty} X_n = 0) = 0$.

11.5 Έστω $(X_n)_{n\geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με τιμές στο $[0,\infty)$ ώστε $\mathbf{P}(X_1 > 0) > 0$. Να δειχθεί ότι με πιθανότητα 1 ισχύει

$$\sum_{n=1}^{\infty} X_n = \infty.$$

11.6 Έστω $(X_n)_{n\geq 1}$ ακολουθία τυχαίων μεταβλητών με τιμές στο $\mathbb{R}$. Να δείξετε ότι υπάρχει ακολουθία $(a_n)_{n\geq 1}$ πραγματικών θετικών αριθμών τέτοια ώστε $\mathbf{P}\Big(\lim_{n\to\infty} \frac{X_n}{a_n} = 0\Big) = 1$.

11.7 Έστω $(X_n)_{n\geq 1}$ ακολουθία ανεξάρτητων τυχαίων μεταβλητών με τιμές στο $\mathbb{R}$. Να δείξετε ότι για την τυχαία μεταβλητή $X^* = \sup_{n\geq 1} X_n$ ισχύει ότι $\mathbf{P}(X^* < \infty) = 1$ αν και μόνο αν υπάρχει $M \in \mathbb{R}$ έτσι ώστε $\sum_{n=1}^{\infty} \mathbf{P}(X_n > M) < \infty$.

11.8* Έστω $(X_n)_{n\geq 1}$ ακολουθία ανεξάρτητων τυχαίων μεταβλητών με $X_n \sim N(0,1)$ για κάθε $n \geq 1$. Να δείξετε ότι με πιθανότητα 1 ισχύει

$$\overline{\lim_{n\to\infty}} \frac{X_n}{\sqrt{2\log n}} = 1. \tag{11.4}$$

11.9 Έστω $(X_n)_{n\geq 1}$ ακολουθία ανεξάρτητων τυχαίων μεταβλητών με $X_n \sim Exp(1)$ για κάθε $n \geq 1$. Αν $M_n := \max\{X_1, X_2, \ldots, X_n\}$ για κάθε $n \geq 1$, να δείξετε ότι με πιθανότητα 1 ισχύει

$$\lim_{n\to\infty} \frac{M_n}{\log n} = 1. \tag{11.5}$$

11.10 Έστω $(X_n)_{n\geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με $\mathbf{P}(X_1 = K) = \mathbf{P}(X_1 = \Gamma) = 1/2$. Οι $(X_n)_{n\geq 1}$ καταγράφουν τα αποτελέσματα ακολουθίας ρίψεων ενός αμερόληπτου νομίσματος. Για κάθε $n \geq 1$ θέτουμε

$$L_n := \max\{m \geq 1 : X_n = X_{n+1} = \ldots = X_{n+m-1}\}.$$

Για παράδειγμα, αν οι ρίψεις δώσουν το αποτέλεσμα $(K, K, \Gamma, \Gamma, K, \Gamma, \Gamma, \Gamma, K, \ldots)$, τότε $L_1(\omega) = 2$, $L_5(\omega) = 1$, και $L_6(\omega) = 3$. Δείξτε ότι

(α) $\overline{\lim_{n\to\infty}} \frac{L_n}{\log_2 n} \leq 1$, με πιθανότητα 1.

(β) $\underline{\lim_{n\to\infty}} L_n = 1$, με πιθανότητα 1.

11.11* Έστω $(X_n)_{n\geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών. Να δείξετε ότι τα ακόλουθα είναι ισοδύναμα:

(α) $\lim_{n\to\infty} \dfrac{X_n}{n} = 0$.

(β) $\mathbf{E}\,|X_1| < \infty$.

11.12 Έστω $(X_n)_{n\geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με τιμές στο $[0,\infty)$, και $c > 1$. Να δείξετε ότι τα ακόλουθα είναι ισοδύναμα:

(α) $\mathbf{E}\{(\log X_1)^+\} < \infty$.

(β) $\sum_{n=1}^{\infty} \frac{X_n}{c^n} < \infty$ με πιθανότητα 1.

11.13* Έστω $(X_n)_{n\geq 1}$ ακολουθία ισόνομων και ανεξάρτητων τυχαίων μεταβλητών με τιμές στο $\mathbb{R}$ ώστε η κατανομή της X_1 να μην είναι συγκεντρωμένη σε ένα σημείο (δηλαδή δεν υπάρχει $c \in \mathbb{R}$ με $\mathbf{P}(X = c) = 1$). Να δειχθεί ότι

$$\mathbf{P}(\lim_{n\to\infty} X_n \text{ υπάρχει }) = 0.$$

11.14 Έστω $(X_n)_{n\geq 1}$ θετικές τυχαίες μεταβλητές. Να δειχθεί ότι με πιθανότητα 1 ισχύει

$$\overline{\lim}\, \frac{1}{n} \log X_n \leq \overline{\lim}\, \frac{1}{n} \log \mathbf{E}\, X_n. \tag{11.6}$$

11.15 Έστω $(X_i)_{i\geq 1}$ ακολουθία τυχαίων μεταβλητών με τιμές στο $\mathbb{R}$ και $\mathcal{C}_\infty := \cap_{n=1}^{\infty} \sigma(X_n, X_{n+1}, ...)$ η τελική σ-άλγεβρα.
(α) Υποθέτουμε ότι οι X_i έχουν θετικές τιμές. Ποιες από τις παρακάτω τυχαίες μεταβλητές είναι $\mathcal{C}_\infty$ μετρήσιμες;

(i) $\underline{\lim}_{n\to\infty} \dfrac{X_n}{n}$, (ii) $\underline{\lim}_{n\to\infty} \dfrac{X_1 + X_2 + \cdots + X_n}{n}$, (iii) $\overline{\lim}_{n\to\infty} n(X_1 + X_2 + \cdots + X_n)$,
(iv) $\displaystyle\sum_{k=1}^{\infty} \frac{X_k}{2^k}$, (v) $\overline{\lim}_{n\to\infty} (X_n + X_{n+1})$.

(β) Ποια από τα παρακάτω σύνολα είναι στοιχεία της $\mathcal{C}_\infty$;

(i) $\left\{\displaystyle\sum_{n=1}^{\infty} |X_n| < \infty\right\}$, (ii) $\{X_1 + X_2 + \cdots + X_n = 0$ για άπειρα $n\}$,
(iii) $\left\{\overline{\lim}_{n\to\infty} n|X_1 + X_2 + \cdots + X_n| \leq 1\right\}$, (iv) $\left\{\displaystyle\sum_{n=1}^{\infty} X_n \text{ συγκλίνει σε πραγματικό άριθμό}\right\}$,
(v) $\left\{\displaystyle\sum_{k=n}^{2n} X_k > 0 \text{ για άπειρα } n\right\}$.

[Σχόλιο: Καταρχάς, τα ερωτήματα να απαντηθούν διαισθητικά. Έπειτα, σχετικά με το (β), για τα σύνολα τα οποία είναι στοιχεία της $\mathcal{C}_\infty$ να αποδειχθεί αυτό τυπικά. Για τα υπόλοιπα, να μην αποδειχθεί τίποτε. Για εκείνα δεν ισχυριζόμαστε ότι πάντοτε δεν είναι στοιχεία της $\mathcal{C}_\infty$. Εξαρτάται από τη συγκεκριμένη επιλογή της ακολουθίας $(X_i)_{i\geq 1}$. Παρόμοιο σχόλιο ισχύει για το μέρος (α) της άσκησης.]

11.16 Έστω $(X_n)_{n\geq 1}$ ακολουθία ανεξάρτητων τυχαίων μεταβλητών, καθεμία με κατανομή την τυπική κανονική $N(0,1)$. Για κάθε $n \geq 1$ θέτουμε $S_n := X_1 + X_2 + \cdots + X_n$.
(α) Για κάθε $A > 0$ και $n \geq 1$ να δειχθεί ότι

$$\mathbf{P}\left(\frac{S_n}{\sqrt{n}} \geq A\right) = 1 - \Phi(A) > 0,$$

όπου Φ είναι η συνάρτηση κατανομής της $N(0,1)$. [Εδώ μπορείτε να χρησιμοποιήσετε πράγματα για κανονικές τυχαίες μεταβλητές και αθροίσματα τους από τις στοιχειώδεις πιθανότητες.]

(β) Για κάθε $A > 0$, με πιθανότητα 1 ισχύει

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

(γ) Με πιθανότητα 1 ισχύει

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

11.17 Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας και $\{A_i : i \in I\}$ στοιχεία της $\mathcal{F}$.

(α) Θεωρούμε τις τυχαίες μεταβλητές $X_i := \mathbf{1}_{A_i}, i \in I$. Να δειχθεί ότι οι $\{X_i : i \in I\}$ είναι ανεξάρτητες αν και μόνο αν τα $\{A_i : i \in I\}$ είναι ανεξάρτητα.

(β) Υποθέτουμε ότι $I = \mathbb{N}^+$. Αν τα $\{A_n : n \in \mathbb{N}^+\}$ είναι ανεξάρτητα, τότε τα σύνολα $\liminf_{n\geq 1} A_n, \limsup_{n\geq 1} A_n$ έχουν πιθανότητα 0 ή 1.

11.18 Έστω $(X_n)_{n\geq 1}$ ακολουθία ανεξάρτητων τυχαίων μεταβλητών με τιμές στο $\mathbb{R}$. Θεωρούμε τη δυναμοσειρά

$$f(z) = \sum_{n=0}^{\infty} X_n z^n.$$

(α) Να δειχθεί ότι η ακτίνα σύγκλισης R της f είναι μετρήσιμη ως προς την τελική σ-άλγεβρα των $(X_n)_{n\geq 1}$ και άρα είναι σταθερή με πιθανότητα 1.

(β) Αν υποθέσουμε επιπλέον ότι καθεμία από τις $(X_n)_{n\geq 1}$ έχει κατανομή $N(0,1)$, τότε με πιθανότητα 1 ισχύει $R = 1$.

11.19 Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας, και $\{A_n : n \geq 1\}$ στοιχεία της $\mathcal{F}$ τα οποία είναι ανα δύο ανεξάρτητα (δηλαδή $\mathbf{P}(A_i \cap A_j) = \mathbf{P}(A_i)\,\mathbf{P}(A_j)$ για κάθε $i, j \geq 1, i \neq j$). Θέτουμε

$$S_n := \sum_{i=1}^{n} \mathbf{1}_{A_i}$$

και

$$s_n := \sum_{i=1}^{n} \mathbf{P}(A_i).$$

(α) Να δειχθεί ότι

$$\mathbf{E}(S_n^2) = s_n + s_n^2 - \sum_{i=1}^{n} \mathbf{P}(A_i)^2 \leq s_n + s_n^2.$$

(β) Για κάθε $\varepsilon > 0$, να δειχθεί ότι

$$\mathbf{P}(S_n \geq \varepsilon s_n) \geq (1-\varepsilon)^2 \frac{s_n^2}{s_n^2 + s_n}.$$

[Υπόδειξη: Άσκηση 6.6.]

(γ)* Αν επιπλέον ισχύει ότι $\sum_{i=1}^{\infty} \mathbf{P}(A_i) = \infty$, να δειχθεί ότι $\mathbf{P}(\limsup_n A_n) = 1$.

[Η άσκηση αυτή γενικεύει το 2ο λήμμα Borel-Cantelli κατά το ότι υποθέτουμε τα $\{A_n : n \geq 1\}$ ανά δύο ανεξάρτητα και όχι απαραίτητα πλήρως ανεξάρτητα.]

11.20 Έστω $(\Omega, \mathcal{F}, \mathbf{P})$ χώρος πιθανότητας και $\{A_n : n \geq 1\}$ στοιχεία της $\mathcal{F}$ για τα οποία ισχύει $\sum_{i=1}^{\infty} \mathbf{P}(A_i) = \infty$ και υπάρχει $C \in (0, \infty)$ ώστε

$$\mathbf{P}(A_i \cap A_j) \leq C\ \mathbf{P}(A_i)\,\mathbf{P}(A_j)$$

για κάθε $i, j \geq 1$. Να δειχθεί ότι

$$\mathbf{P}(\limsup_n A_n) \geq 1/C > 0.$$

12

Ο Ισχυρός Νόμος των Μεγάλων Αριθμών

Στο κεφάλαιο αυτό παρουσιάζουμε ένα από τα σημαντικότερα αποτελέσματα της Θεωρίας Πιθανοτήτων, τον ισχυρό νόμο των μεγάλων αριθμών. Η διατύπωση που θα αποδείξουμε δεν είναι η ισχυρότερη μορφή του νόμου, όμως η απόδειξή της είναι ευκολότερη τεχνικά και διατηρεί αρκετά από τα στοιχεία της απόδειξης της ισχυρής μορφής.

12.1 Το θεώρημα

Θεώρημα 12.1 (Ισχυρός Νόμος των Μεγάλων Αριθμών)**.** *Έστω $(X_n)_{n\geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές στον $(\Omega, \mathcal{F}, \mathbf{P})$ με τιμές στο $\mathbb{R}$ έτσι ώστε $\mathbf{E}(X_1^2) < \infty$. Θέτουμε $\mu = \mathbf{E}(X_1)$ και $S_n = \sum_{k=1}^n X_k$ για κάθε $n \geq 1$. Τότε*

$$\lim_{n\to\infty} \frac{S_n}{n} = \mu \text{ με πιθανότητα 1.}$$

Απόδειξη. Έστω $\sigma^2 = \mathrm{Var}(X_1)$.
Πρώτα θα αποδείξουμε το ζητούμενο για $(X_n)_{n\in\mathbb{N}}$ με τιμές στο $[0, \infty]$. Για $n \geq 1$, θέτουμε $Y_n = \frac{S_n}{n} - \mu$. Τότε

$$\mathbf{E}(Y_n) = \frac{1}{n}\mathbf{E}(S_n) - \mu = \frac{1}{n} n\,\mathbf{E}(X_1) - \mu = 0,$$

και

$$\mathbf{E}(Y_n^2) = \mathrm{Var}(Y_n^2) = \mathrm{Var}\left(\frac{S_n}{n}\right) = \frac{1}{n^2}\mathrm{Var}(S_n).$$

Όμως $\mathrm{Var}(S_n) = n\,\mathrm{Var}(X_1)$ εφόσον οι $(X_n)_{n\geq 1}$ είναι ανεξάρτητες. Επομένως,

$$\mathbf{E}(Y_n^2) = \frac{\sigma^2}{n}.$$

Έτσι, για την υπακολουθία $(Y_{n^2})_{n\geq 1}$, έχουμε ότι

$$\mathbf{E}\left\{\sum_{n=1}^{\infty} Y_{n^2}^2\right\} = \sum_{n=1}^{\infty} \frac{\sigma^2}{n^2} < \infty.$$

Άρα, με πιθανότητα 1, $\sum_{n=1}^{\infty} Y_{n^2}^2 < \infty$, συνεπώς $\lim_{n\to\infty} Y_{n^2} = 0$.

Τώρα, από το $\lim_{n\to\infty} Y_{n^2} = 0$, θέλουμε να περάσουμε στο $\lim_{n\to\infty} Y_n = 0$.
Έστω $k \geq 1$. Θέτουμε $n(k) = [\sqrt{k}]$. Τότε

$$\frac{S_{n(k)^2}}{(n(k)+1)^2} \leq \frac{S_k}{k} \leq \frac{S_{(n(k)+1)^2}}{n(k)^2}$$

εφόσον $n(k) \leq \sqrt{k} \leq n(k)+1$ και η S_n είναι άθροισμα θετικών όρων. Όμως, για $n \to \infty$, με πιθανότητα 1,

$$\frac{S_{n(k)^2}}{(n(k)+1)^2} = \frac{S_{n(k)^2}}{n(k)^2}\Big(\frac{n(k)}{n(k)+1}\Big)^2 \to \mu$$

και
$$\frac{S_{(n(k)+1)^2}}{n(k)^2} = \frac{S_{(n(k)+1)^2}}{(n(k)+1)^2}\Big(\frac{n(k)+1}{n(k)}\Big)^2 \to \mu.$$
Άρα, με πιθανότητα 1, έχουμε ότι $\lim_{k\to\infty} \frac{S_k}{k} = \mu$.

Στην περίπτωση που οι $(X_n)_{n\ge 1}$ παίρνουν τιμές στο $[-\infty, \infty]$, έχουμε
$$\frac{S_n}{n} = \frac{X_1 + X_2 + \ldots + X_n}{n} = \frac{X_1^+ + X_2^+ + \ldots + X_n^+}{n} - \frac{X_1^- + X_2^- + \ldots + X_n^-}{n}.$$
Οι $(X_n^+)_{n\in\mathbb{N}}$, όπως και οι $(X_n^-)_{n\in\mathbb{N}}$, είναι ανεξάρτητες και ισόνομες και $\mathbf{E}\left((X_1^+)^2\right) < \infty$, $\mathbf{E}\left((X_1^-)^2\right) < \infty$ εφόσον $(X_1^+)^2 \le X_1^2$, $(X_1^-)^2 \le X_1^2$. Από τα προηγούμενα, έχουμε ότι
$$\lim_{n\to\infty} \frac{X_1^+ + X_2^+ + \ldots + X_n^+}{n} = \mathbf{E}(X_1^+),$$
και
$$\lim_{n\to\infty} \frac{X_1^- + X_2^- + \ldots + X_n^-}{n} = \mathbf{E}(X_1^+)$$
με πιθανότητα 1. Συνεπώς, με πιθανότητα 1,
$$\lim_{n\to\infty} \frac{S_n}{n} = \mathbf{E}(X_1^+) - \mathbf{E}(X_1^-) = \mu.$$
■

Το συμπέρασμα του θεωρήματος ισχύει και αν αντί της $\mathbf{E}(X_1^2) < \infty$ υποθέσουμε ότι $\mathbf{E}|X_1| < \infty$, δηλαδή κάτι λιγότερο. Αυτή είναι η γενική μορφή του νόμου των μεγάλων αριθμών και στο εξής θα τον θεωρούμε δεδομένο με αυτή, την ισχυρότερη μορφή.

Στην Άσκηση 12.3 διατυπώνεται ένα είδος αντίστροφου του νόμου των μεγάλων αριθμών. Δηλαδή, αν ο μέσος όρος ανεξάρτητων και ισόνομων τυχαίων μεταβλητών συγκλίνει με πιθανότητα 1 σε πεπερασμένο αριθμό, τότε αναγκαστικά $\mathbf{E}|X_1| < \infty$. Άρα η υπόθεση $\mathbf{E}|X_1| < \infty$ στο θεώρημα είναι απαραίτητη για την ισχύ του συμπεράσματος.

Επίσης, στην Άσκηση 12.2 δείχνουμε ότι το συμπέρασμα του νόμου των μεγάλων αριθμών ισχύει ακόμα και όταν η μέση τιμή $\mu := \mathbf{E}(X_1)$ είναι $-\infty$ ή ∞. Άρα το συμπέρασμα ισχύει πάντοτε όταν η $\mathbf{E}(X_1)$ μπορεί να οριστεί.

Όταν η $\mathbf{E}(X_1)$ δεν μπορεί να οριστεί, δηλαδή όταν $\mathbf{E}(X_1^+) = \mathbf{E}(X_1^-) = \infty$, τότε ισχύει ακριβώς ένα από τα εξής

(i) $\lim_{n\to\infty} \frac{S_n}{n} = -\infty$ με πιθανότητα 1.

(ii) $\lim_{n\to\infty} \frac{S_n}{n} = \infty$ με πιθανότητα 1.

(iii) $\underline{\lim}_{n\to\infty} \frac{S_n}{n} = -\infty, \overline{\lim}_{n\to\infty} \frac{S_n}{n} = \infty$ με πιθανότητα 1.

Ποιο από τα σενάρια συμβαίνει εξαρτάται από την κατανομή της X_1 και υπάρχει μάλιστα κριτήριο που το καθορίζει αλλά δεν θα το διατυπώσουμε [δες Erickson (1973)].

Σύμβαση: Αν $(X_n)_{n\ge 1}$ είναι ακολουθία τυχαίων μεταβλητών με τιμές στο $\mathbb{R}$, στο εξής θα συμβολίζουμε με S_n το n-οστό μερικό άθροισμά τους.

Πόρισμα 12.2. *(Ο Ασθενής Νόμος των Μεγάλων Αριθμών) Έστω $(X_n)_{n\ge 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με τιμές στο $\mathbb{R}$ έτσι ώστε $\mu := \mathbf{E}(X_1) \in \mathbb{R}$. Τότε*
$$\lim_{n\to\infty} \frac{S_n}{n} = \mu \text{ κατά πιθανότητα.}$$

Απόδειξη. Έπεται από τον ισχυρό νόμο των μεγάλων αριθμών (ισχυρή μορφή) και το ότι η σχεδόν βέβαιη σύγκλιση συνεπάγεται την κατά πιθανότητα [Θεώρημα 8.2(ii))]. ■

Παρατήρηση 12.3 (Η μέθοδος Monte Carlo)**.** Ο νόμος των μεγάλων αριθμών δικαιολογεί την εξής μέθοδο για τον προσεγγιστικό υπολογισμό μέσης τιμής μιας δεδομένης τυχαίας μεταβλητής X. Παράγουμε με κάποιο τρόπο έναν μεγάλο αριθμό ανεξάρτητων πραγματοποιήσεων της τ.μ. X, έστω $x_1, x_2, \ldots, x_n$, και υπολογίζουμε τον μέσο τους όρο $(x_1 + x_2 + \ldots + x_n)/n$. Αυτός ο μέσος όρος είναι η ζητούμενη προσέγγιση.

Με αυτή τη μέθοδο μπορούμε να βρούμε προσεγγίσεις για οποιαδήποτε ποσότητα μπορεί να παρασταθεί ως μέση τιμή κάποιας τ.μ. Η ποσότητα ενδέχεται να εμφανίζεται φυσιολογικά ως μέση τιμή σε κάποιο πρόβλημα πιθανοτήτων και να μην μπορεί να υπολογιστεί με κλειστό τύπο. Επίσης μπορεί να μην έχει καμία σχέση με πιθανότητες. Για παράδειγμα, για τον αριθμό π έχουμε την αναπαράσταση $\pi = 4\,\mathbf{E}(\mathbf{1}_{U^2+V^2<1})$ όπου οι U, V είναι ανεξάρτητες ομοιόμορφες τυχαίες μεταβλητές στο $(-1, 1)$. Η τ.μ. (U, V) είναι ομοιόμορφη στο τετράγωνο $(-1, 1) \times (-1, 1)$ και η μέθοδος Monte Carlo για την προσέγγιση της $\mathbf{E}(\mathbf{1}_{U^2+V^2<1})$ παίρνει πολλές πραγματοποιήσεις της (U, V) και υπολογίζει το ποσοστό αυτών που πέφτουν μέσα στον μοναδιαίο δίσκο.

12.2 Δύο εφαρμογές

Ανανεωτική θεωρία. Έστω $(X_n)_{n\geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με τιμές στο $[0, \infty)$. Θέτουμε

$$T_0 := 0,$$
$$T_n := X_1 + X_2 + \cdots + X_n$$

για κάθε $n \geq 1$. Η ερμηνεία που σκεφτόμαστε για τις ακολουθίες $(X_n)_{n\geq 1}, (T_n)_{n\geq 1}$ είναι η εξής. Για το φωτισμό ενός δωματίου, έχουμε άπειρο πλήθος λαμπών, κάθε μία από τις οποίες έχει τυχαίο χρόνο ζωής. X_n είναι ο χρόνος ζωής της n λάμπας. Τη χρονική στιγμή 0 τοποθετούμε τη λάμπα 1. Μόλις αυτή καεί, τοποθετούμε τη λάμπα 2 και συνεχίζουμε όμοια. T_n είναι ο χρόνος που καίγεται η λάμπα n. Τώρα για $t \geq 0$, θέτουμε

$$N_t := \sup\{n \geq 0 : T_n \leq t\}.$$

Η τυχαία μεταβλητή N_t μετράει πόσες λάμπες κάηκαν κατά το χρονικό διάστημα $[0, t]$. Έστω ότι $\mu := \mathbf{E}(X_1)$. Αφού κάθε λάμπα ζει περίπου χρόνο μ, για το χρονικό διάστημα $[0, t]$ αναμένουμε να χρειάζονται t/μ λάμπες περίπου. Δηλαδή $N_t \approx t/\mu$.

Θεώρημα 12.4. *Με πιθανότητα 1 ισχύει*

$$\lim_{t\to\infty} \frac{N_t}{t} = \frac{1}{\mu}.$$

Απόδειξη. Από το νόμο των μεγάλων αριθμών, υπάρχει ένα μετρήσιμο σύνολο $A \subset \Omega$ με $\mathbf{P}(A) = 1$ ώστε για κάθε $\omega \in A$ να ισχύει

$$\lim_{n\to\infty} \frac{T_n}{n} = \mu. \tag{12.1}$$

Από τον ορισμό του N_t έχουμε

$$T_{N_t} \leq t \leq T_{N_t+1},$$

και άρα

$$\frac{T_{N_t}}{N_t} \leq \frac{t}{N_t} \leq \frac{T_{N_t+1}}{N_t} = \frac{T_{N_t+1}}{N_t+1} \frac{N_t+1}{N_t}. \tag{12.2}$$

Ισχυρισμός: Για κάθε $\omega \in \Omega$ ισχύει $N_t \to \infty$.

Πράγματι, επειδή η N_t είναι αύξουσα συνάρτηση του t, αν δεν ισχύει ο ισχυρισμός, τότε η N_t θα ήταν φραγμένη. Δηλαδή θα υπάρχει φυσικός $\ell \geq 1$ ώστε $N_t \leq \ell$ για κάθε $t \geq 0$. Άρα $X_1 + X_2 + \cdots + X_\ell \geq t$ για κάθε $t \geq 0$, το οποίο δεν μπορεί να ισχύει γιατί οι $X_1, X_2, .., X_\ell$ παίρνουν πραγματικές τιμές (και όχι την τιμή ∞).

Για $\omega \in A$, και χρησιμοποιώντας τον ισχυρισμό και την (12.1), έχουμε

$$\lim_{t\to\infty} \frac{T_{N_t}}{N_t} = \lim_{t\to\infty} \frac{T_{N_t+1}}{N_t+1} = \mu.$$

Άρα για $t \to \infty$, η (12.2) δίνει $\lim_{t\to\infty} t/N_t = \mu$ που είναι το ζητούμενο. ∎

Εντροπία. Έστω X διακριτή τυχαία μεταβλητή, S το (αριθμήσιμο) σύνολο τιμών της, και $f(x) = \mathbf{P}(X = x)$ η συνάρτηση πιθανότητας της. **Εντροπία** της X ονομάζουμε τον αριθμό

$$H(X) := -\sum_{x\in S} f(x)\log f(x) \tag{12.3}$$

με τη σύμβαση $0\log 0 = 0$. Ισχύει $H(X) \geq 0$ γιατί $f(x) \in [0, 1]$. Επίσης, $H(X) = -\mathbf{E}\{\log f(X)\}$.

Η εντροπία της X εκφράζει το μέγεθος της αβεβαιότητας που έχουμε για την τιμή που θα πάρει η X αν επιχειρήσουμε να παραγάγουμε μια πραγματοποίηση της. Ας πούμε ότι η X παίρνει τιμές στο $\{x_1, x_2, \ldots, x_k\}$ με αντίστοιχες πιθανότητες $p_j := \mathbf{P}(X = x_j)$. Αν $p_1 = 1$ και $p_2 = \cdots = p_k = 0$, τότε $H(X) = 0$, και βέβαια δεν υπάρχει καμία αβεβαιότητα για την πραγματοποίηση της X, θα είναι x_1. Η εντροπία μεγιστοποιείται όταν $p_1 = p_2 = \cdots = p_k = 1/k$ (Άσκηση 12.8). Τότε όλα τα ενδεχόμενα είναι το ίδιο πιθανά και η αβεβαιότητα μας μέγιστη. Σε περίπτωση άνισων πιθανοτήτων, η αβεβαιότητα είναι μικρότερη γιατί περιμένουμε η X να πάρει μία από τις τιμές που έχουν μεγαλύτερη πιθανότητα.

Τον Ορισμό (12.3) υπαγορεύει ένας υπολογισμός που θα δούμε αμέσως τώρα. Ο ίδιος υπολογισμός δίνει και την εμπειρική σημασία της εντροπίας.

Έστω $(X_k)_{k\geq 1}$ ακολουθία ανεξάρτητων και ισόνομων διακριτών τυχαίων μεταβλητών, καθεμία ισόνομη με τη X. Για $n \geq 1$ σταθερό και δεδομένα $x_1, x_2, \ldots, x_n \in S$, υπολογίζουμε την πιθανότητα οι n πραγματοποιήσεις $(X_1, X_2, \ldots, X_n)$ της X να ταυτιστούν με το διάνυσμα $(x_1, x_2, \ldots, x_n)$. Αυτή ισούται με

$$p_n(x_1, x_2, \ldots, x_n) := \mathbf{P}(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = f(x_1)f(x_2)\cdots f(x_n). \tag{12.4}$$

Για παράδειγμα, όταν η X παίρνει τις τιμές 1 και 0 με πιθανότητες $p := 1/3$ και $2/3$ αντίστοιχα, τότε $f(x) = p^x(1-p)^{1-x}\mathbf{1}_{x\in\{0,1\}}$ και η πιθανότητα εμφάνισης της n-άδας $(x_1, x_2, \ldots, x_n) \in \{0, 1\}^n$ είναι

$$\left(\frac{1}{3}\right)^{x_1+\cdots+x_n}\left(\frac{2}{3}\right)^{n-x_1-\cdots-x_n}$$

Αυτή η πιθανότητα παίρνει τιμές από $(1/3)^n$ ως και $(2/3)^n$. Δεν είναι σταθερή, η τιμή της εξαρτάται από την επιλογή της n-άδας $(x_1, x_2, \ldots, x_n)$. Όμως η πιθανότητα πραγματοποίησης μιας *τυχαίας* n-άδας $(X_1, X_2, \ldots, X_n)$ είναι περίπου η ίδια σχεδόν για κάθε πραγματοποίηση της $(X_1, X_2, \ldots, X_n)$.

Πιο συγκεκριμένα, στη γενική περίπτωση θέλουμε να εκτιμήσουμε την $p_n(X_1, X_2, \ldots, X_n)$. Αυτή η πιθανότητα είναι μια τυχαία μεταβλητή (εξαρτάται από τις $X_1, X_2, \ldots, X_n$), φθίνει εκθετικά με το n και από το νόμο των μεγάλων αριθμών έχουμε

$$\begin{aligned}\frac{1}{n}\log p_n(X_1, X_2, \ldots, X_n) &= \frac{1}{n}\log f(X_1)f(X_2)\cdots f(X_n)\\ &= \frac{\log f(X_1) + \log f(X_2) + \cdots + \log f(X_n)}{n} \to \mathbf{E}(\log f(X_1)) = -H(X)\end{aligned} \tag{12.5}$$

με πιθανότητα 1 καθώς $n \to \infty$. Δηλαδή, με πιθανότητα 1, για μεγάλο n, ισχύει

$$p_n(X_1, X_2, \ldots, X_n) \approx e^{-nH(X)}.$$

Στο παράδειγμα πιο πάνω, $H(X) = -(1/3)\log(1/3) - (2/3)\log(2/3)$ και η $e^{-nH(X)}$ είναι μια πιθανότητα ανάμεσα στις $(1/3)^n, (2/3)^n$.

Ας υποθέσουμε ότι το S είναι πεπερασμένο. Αν ορίσουμε

$$A_n^\varepsilon = \{(x_1, x_2, \dots, x_n) \in S^n : e^{-n(H(X)+\varepsilon)} < p_n(x_1, x_2, \dots, x_n) < e^{-n(H(X)-\varepsilon)}\},$$

τότε $\mathbf{P}(A_n^\varepsilon) \to 1$ λόγω της (12.5). Ακολουθίες στο A_n^ε τις λέμε ε-τυπικές. Αφού όλες τους έχουν πιθανότητα περίπου $e^{-nH(X)}$ και το άθροισμα των πιθανοτήτων τους είναι σχεδόν 1, το A_n^ε έχει πληθικότητα περίπου $e^{nH(X)}$. Στη γενική περίπτωση ισχύει $H(X) < \log|S|$ (Άσκηση 12.8), οπότε το A_n^ε είναι (από πλευράς πληθικότητας) ένα πολύ μικρό κομμάτι του S^n αφού το S^n έχει πληθικότητα $|S|^n = e^{n\log|S|}$. Παρ' όλα αυτά, το A_n^ε συγκεντρώνει σχεδόν όλη την πιθανότητα.

Η έννοια της εντροπίας είναι κεντρικής σημασίας στη Θεωρία Πληροφορίας και έχει πολλές εφαρμογές (δες Cover and Thomas (2014)).

Ασκήσεις

12.1 (Ασθενής Νόμος των Μεγάλων Αριθμών. Ασθενής έκδοση.) Έστω $(X_n)_{n\in\mathbb{N}}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με τιμές στο $\mathbb{R}$ έτσι ώστε $\mathbf{E}(X_1^2) < \infty$. Θέτουμε $\mu = \mathbf{E}(X_1)$. Χωρίς χρήση του ισχυρού νόμου των μεγάλων αριθμών να δειχθεί ότι

$$\lim_{n\to\infty} \mathbf{P}\left(\left|\frac{S_n}{n} - \mu\right| > \epsilon\right) = 0 \text{ για κάθε } \epsilon > 0.$$

Δηλαδή η ακολουθία $\frac{S_n}{n}$ συγκλίνει στο μ κατά πιθανότητα.

12.2* Έστω $(X_n)_{n\in\mathbb{N}}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με τιμές στο $\mathbb{R}$ έτσι ώστε $\mathbf{E}(X_1^+) = \infty$ και $\mathbf{E}(X_1^-) < \infty$. Να δείξετε ότι

$$\lim_{n\to\infty} \frac{S_n}{n} = \infty$$

με πιθανότητα 1.

12.3* (Αντίστροφο του Νόμου των Μεγάλων Αριθμών) Έστω $(X_n)_{n\geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με τιμές στο $\mathbb{R}$ έτσι ώστε $\lim_{n\to\infty} \frac{S_n}{n} = \mu$ σχεδόν βεβαίως με $\mu \in \mathbb{R}$. Να δείξετε ότι $\mathbf{E}|X_1| < \infty$ και $\mathbf{E}(X_1) = \mu$.
[Υπόδειξη: Χρήσιμη είναι η Άσκηση 11.11.]

12.4 Έστω $(X_n)_{n\geq 1}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με $X_1 \sim N(1, 3)$. Να δείξετε ότι

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

με πιθανότητα 1.

12.5 Έστω $(X_n)_{n\in\mathbb{N}}$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με τιμές στο $\mathbb{R}$ έτσι ώστε $\mathbf{E}(|X_1|) < \infty$ και $\mathbf{E}(X_1) > 0$. Να δείξετε ότι

$$\lim_{n\to\infty} S_n = \infty$$

με πιθανότητα 1.

12.6 Έστω $(U_i)_{i\geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών, καθεμία με κατανομή $U(0, 1)$, δηλαδή ομοιόμορφη στο $(0, 1)$. Να δειχθεί ότι

(α) $\lim_{n\to\infty}(U_1U_2\cdots U_n)^{1/n} = e^{-1}$ με πιθανότητα 1.

(β) $\lim_{n\to\infty} U_1U_2\cdots U_n = 0$ με πιθανότητα 1.

(γ) $\lim_{n\to\infty} \frac{U_1^a + \cdots + U_n^a}{n} = \begin{cases} \frac{1}{1+a} & \text{με πιθανότητα 1 αν } a > -1, \\ \infty & \text{με πιθανότητα 1 αν } a \leq -1. \end{cases}$

12.7 Έστω $(X_i)_{i\geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με $\mu = \mathbf{E}(X_1) \in \mathbb{R}$ και $\sigma^2 = V(X_1) < \infty$. Να δειχθεί ότι

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n}(X_k - \mu)^2 = \sigma^2 \text{ με πιθανότητα 1.}$$

12.8 Έστω $k \geq 2, S = \{x_1, \ldots, x_k\}$ σύνολο με k στοιχεία, και X τυχαία μεταβλητή με τιμές στο S. Να δειχθεί ότι $H(X) \leq \log k$ και η ισότητα ισχύει αν και μόνο αν $\mathbf{P}(X = x_j) = 1/k$ για κάθε $1 \leq j \leq k$.
[Υπόδειξη: Εφαρμόστε κατάλληλα την ανισότητα Jensen.]

12.9 Έστω $(X_k)_{k\geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με τιμές στο $S = \{1, 2, \ldots, r\}$ και f η συνάρτηση πιθανότητας της X_1. Έστω επίσης $(Y_k)_{k\geq 1}$ ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με τιμές στο S και g η συνάρτηση πιθανότητας της Y_1. Υποθέτουμε ότι για $k \in S$ ισχύει $f(k) = 0 \Rightarrow g(k) = 0$. Ορίζουμε την p_n όπως στην (12.4). Δείξτε ότι η πιθανότητα $p_n(Y_1, Y_2, \ldots, Y_n)$ φθίνει εκθετικά και υπολογίστε το ρυθμό μείωσης. Αυτή είναι η πιθανότητα στις πρώτες n συντεταγμένες η ακολουθία X να μοιάζει με δείγμα παρμένο από την Y.