Συμπληρωματα θεωριας

1. Ορισμος της μεσης τιμης

Η διαφορά του ορισμού της μέσης τιμής που δίνουμε παρακάτω απο αυτόν του βιβλίου είναι το ότι επιτρέπει η μέση τιμή να παίρνει τις τιμές $-\infty, \infty$.
Καταρχάς, ορίζουμε για $x \in \mathbb{R}$,

$$x^+ = \begin{cases} x & \text{αν } x \geqslant 0 \\ 0 & \text{αν } x < 0 \end{cases} \qquad x^- = \begin{cases} 0 & \text{αν } x \geqslant 0 \\ |x| & \text{αν } x < 0 \end{cases}$$

και παρατηρούμε ότι $x^+, x^- \geqslant 0$ και $|x| = x^- + x^+$, $x = x^+ - x^-$.

Μεση τιμη για διακριτες τυχαιες μεταβλητες

Περίπτωση 1. Αν η τυχαία μεταβλητή X παίρνει μόνο μή αρνητικές τιμές[1] $\{x_i : i \in I\}$ όπου I είναι ένα αριθμήσιμο σύνολο (δηλαδή ισοπληθικό με το $\mathbb{N}$ ή με το $\{1, 2, \dots, n\}$ για κάποιο $n \in \mathbb{N}^+$) και $x_i \geqslant 0$ για κάθε $i \in I$. Τότε θέτουμε[2]

$$E(X) := \sum_{i \in I} x_i P(X = x_i), \tag{1}$$

το οποίο ή είναι ένας αριθμός στο $[0, \infty)$ ή είναι ∞. Ορίζεται πάντοτε.

Περίπτωση 2. Αν η X παίρνει τιμές στο $\mathbb{R}$, έστω $\{x_i : i \in I\}$ όπου I είναι ένα αριθμήσιμο σύνολο. Τότε οι μεταβλητές X^+, X^- είναι διακριτές με μη αρνητικές τιμές. Προφανώς (με βάση την Περίπτωση 1)

$$E(X^+) = \sum_{i \in I : x_i > 0} x_i P(X = x_i), \;\; E(X^-) = \sum_{i \in I : x_i < 0} |x_i| P(X = x_i).$$

- Αν $E(X^+) = E(X^-) = \infty$ τότε η $E(X)$ δεν ορίζεται.
- Αν τουλάχιστον μία απο τις $E(X^+), E(X^-)$ είναι διαφορετική από το ∞ τότε θέτουμε

  $$E(X) := E(X^+) - E(X^-), \tag{2}$$

  το οποίο μπορεί να είναι ή ένας πραγματικός αριθμός, ή $-\infty$, ή ∞. Και προφανώς

  $$E(X) = \sum_{i \in I} x_i P(X = x_i).$$

Μεση τιμη για συνεχεις τυχαιες μεταβλητες

Έστω X συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας f. Ακολουθούμε ανάλογη διαδικασία όπως και πρίν. Θέτουμε

$$E(X) := \int_{-\infty}^{\infty} t f(t)\, dt \tag{3}$$

όπου το τελευταίο ολοκλήρωμα ορίζεται να ισούται με

$$\int_0^{\infty} t f(t)\, dt - \int_{-\infty}^{0} |t| f(t)\, dt$$

αν τουλάχιστον ένα απο τα δύο ολοκληρώματα είναι $\neq \infty$, και τότε η $E(X)$ είναι πραγματικός αριθμός, ή $-\infty$, ή ∞. Αν και τα δύο ολοκληρώματα είναι ∞, τότε το γενικευμένο ολοκλήρωμα στην (3), όπως και η $E(X)$, δεν ορίζεται. Βέβαια αν $X \geqslant 0$ ή $X \leqslant 0$, τότε η $E(X)$ ορίζεται οπωσδήποτε.

---

[1] Γράφουμε τότε $X \geqslant 0$.

[2] Δεν έχει σημασία η σειρά άθροισης γιατί είναι όλοι αριθμοί με ίδιο πρόσημο.

Και γενικότερα, όταν για κάποιο $a \in \mathbb{R}$ έχουμε $X(\omega) \geqslant a$ για κάθε $\omega \in \Omega$ ή $X(\omega) \leqslant a$ για κάθε $\omega \in \Omega$.

Παραδειγματα

1. Θεωρούμε διακριτή τυχαία μεταβλητή με συνάρτηση πιθανότητας

$$f(x) = \begin{cases} \frac{3}{\pi^2}\frac{1}{x^2} & \text{αν } x \in \mathbb{Z}\backslash\{0\}, \\ 0 & \text{διαφορετικά.} \end{cases}$$

Προφανώς

$$E(X^+) = \sum_{k=1}^{\infty} k\frac{3}{\pi^2}\frac{1}{k^2} = \infty = \frac{3}{\pi^2}\sum_{k=1}^{\infty}\frac{1}{k} = \infty,$$
$$E(X^-) = \sum_{k=-1}^{-\infty} |k|\frac{3}{\pi^2}\frac{1}{k^2} = \frac{3}{\pi^2}\sum_{k=1}^{\infty}\frac{1}{k} = \infty.$$

Οπότε η $E(X)$ δεν ορίζεται.

2. Θεωρούμε διακριτή τυχαία μεταβλητή με συνάρτηση πιθανότητας

$$f(x) = \begin{cases} \frac{3}{2\pi^2}\frac{1}{x^2} & \text{αν } x \in \mathbb{Z},\ x < 0, \\ \frac{1}{4} & \text{αν } x = 0, \\ \frac{1}{2^{x+1}} & \text{αν } x \in \mathbb{Z},\ x > 0. \end{cases}$$

Τώρα όπως πριν έχουμε $E(X^-) = \sum_{k=1}^{\infty} |k|\frac{3}{2\pi^3}\frac{1}{k^2} = \infty$, και

$$E(X^+) = \sum_{k=1}^{\infty} k\frac{1}{2^{k+1}} < \infty$$

απο το κριτήριο του λόγου, για παράδειγμα. Μάλιστα το τελευταίο άθροισμα είναι 1, αλλά μας ενδιαφέρει απλώς το ότι είναι πεπερασμένο. Έπεται ότι $E(X) = 1 - \infty = -\infty$.

3. Θεωρούμε συνεχή τυχαία μεταβλητή X με πυκνότητα (ελέγξτε ότι είναι όντως πυκνότητα)

$$f(x) = \begin{cases} \frac{1}{|x|^3} & \text{αν } x \leqslant -1, \\ \frac{1}{2x^2} & \text{αν } x \geqslant 1. \end{cases}$$

Έχουμε

$$\int_{-\infty}^{0} |x| f(x)\,dx = ... = \int_{1}^{\infty} \frac{1}{x^2}\,dx = [-1/x]_1^{\infty} = 1,$$
$$\int_{0}^{\infty} x f(x)\,dx = \int_{1}^{\infty} \frac{1}{2x}\,dx = \infty.$$

Άρα $E(X) = \infty$.

4. Θεωρούμε συνεχή τυχαία μεταβλητή X με πυκνότητα

$$f(x) = \frac{1}{\pi}\frac{1}{(x^2+1)} \quad \text{για κάθε } x \in \mathbb{R}.$$

Έχουμε

$$\int_{-\infty}^{0} |x| f(x)\,dx = \int_{0}^{\infty} x f(x)\,dx = \frac{1}{\pi}\int_{0}^{\infty} \frac{x}{x^2+1}\,dx \geqslant \frac{1}{\pi}\int_{1}^{\infty}\frac{1}{2x}\,dx = \infty.$$

Άρα η $E(X)$ δεν ορίζεται.

2. Πυκνότητα συνεχούς τυχαίας μεταβλητής

Αυτή η παράγραφος είναι χρήσιμη για την πλήρη δικαιολόγηση λύσης ασκήσεων στις οποίες αναζητούμε την πυκνότητα μιας τ.μ. $g(X)$ γνωρίζοντας την πυκνότητα, f_X, της X [δες την εφαρμογή παρακάτω].

Συνεχή τυχαία μεταβλητή λέμε οποιαδήποτε τ.μ. της οποίας η συνάρτηση κατανομής $F : \mathbb{R} \to [0,1]$ είναι συνεχής συνάρτηση. Αυτές που καλούμε εμείς στην τάξη συνεχείς είναι αυτές που σωστότερα λέγονται *απολύτως συνεχείς*. Δηλαδή αυτές για τις οποίες υπάρχει $f : \mathbb{R} \to [0,\infty)$ με

$$F(b) - F(a) = \int_a^b f(t)\,dt \tag{4}$$

για κάθε $a, b \in \mathbb{R}$ με $a < b$. Μια τέτοια f λέγεται *πυκνότητα* για την X (και για την F), και αν υπάρχει δεν είναι μοναδική. Αλλάζοντας την σε πεπερασμένα το πλήθος σημεία εξακολουθεί να ικανοποιεί την παραπάνω ισότητα. Αποδεικνύεται ότι πυκνότητες είναι ακριβώς οι (Lebesgue μετρήσιμες) συναρτήσεις που ικανοποιούν $f \geqslant 0, \int_{\mathbb{R}} f(x)\,dx = 1$.

Ερώτημα: Δεδομένης μιας συνεχούς συνάρτησης κατανομής F, πώς αποδεικνύουμε ότι αυτή προέρχεται από πυκνότητα, και πώς βρίσκουμε μια πυκνότητα f;

Αν η F προέρχεται απο πυκνότητα, τότε θα έχουμε $f(x) = F'(x)$ σε πάρα πολλά σημεία (σε όλα εκτός απο ένα σύνολο μέτρου μηδέν). Αυτό είναι ένα σημείο εκκίνησης.

Για το μάθημα, θα μας αρκέσει η εξής πρόταση.

Πρόταση. Έστω συνεχής συνάρτηση κατανομής F για την οποία υπάρχει ένα πεπερασμένο σύνολο $J = \{a_1, a_2, \ldots, a_n\} \subset \mathbb{R}$ και συνάρτηση $f : \mathbb{R} \to [0,\infty)$ έτσι ώστε

- $F'(t) = f(t)$ για κάθε $t \in \mathbb{R} \backslash \{a_1, a_2, \ldots, a_n\}$,
- f συνεχής στο $\mathbb{R} \backslash \{a_1, a_2, \ldots, a_n\}$.

Τότε η f είναι μιά πυκνότητα που αντιστοιχεί στη συνάρτηση κατανομής F.

Συμπέρασμα: Δεδομένης μιας συνεχούς F, υπολογίζουμε την F', και έστω ότι αυτή υπάρχει παντού εκτός από ένα πεπερασμένο σύνολο J_1. Το ευνοϊκό σενάριο είναι η F' να είναι συνεχής στο $\mathbb{R} \backslash J_1$, ή αν είναι ασυνεχής, να είναι ασυνεχής σε πεπερασμένο σύνολο σημείων J_2 του $\mathbb{R} \backslash J_1$. Θέτουμε τότε $J := J_1 \cup J_2$ και $f := F'$ στο $\mathbb{R} \backslash J$. Επεκτείνουμε την f αυθαίρετα στο J δίνοντάς της κάποιες μη αρνητικές τιμές. Για παράδειγμα, $f = 0$ στο J. Τότε για τις F, f εφαρμόζεται η πιο πάνω πρόταση και παίρνουμε ότι η f είναι μια πυκνότητα για την F.

Εφαρμογη

Έστω X συνεχής τυχαία μεταβλητή με πυκνότητα

$$f_X(x) = \frac{1}{\pi} \frac{1}{(1 + x^2)} \text{ για κάθε } x \in \mathbb{R}.$$

Θέτουμε $Y = \tan^{-1} X$, όπου $\tan^{-1}$ είναι η αντίστροφη του περιορισμού της $\tan$ στο $(-\pi/2, \pi/2)$. Ποια είναι η κατανομή της Y;

Λύση

Για $t \in \mathbb{R}$ έχουμε $F_Y(t) := P(Y \leqslant t)$, το οποίο ισούται με 0 για $t \leqslant -\pi/2$ και με 1 για $t \geqslant \pi/2$ επειδή η Y παίρνει τιμές στο $(-\pi/2, \pi/2)$. Έπειτα, για $t \in (-\pi/2, \pi/2)$,

$$F_Y(t) = P(\tan^{-1} X \leqslant t) = P(X \leqslant \tan t) = F_X(\tan t).$$

Άρα

$$F_Y(t) = \begin{cases} 0 & \text{αν } t \leqslant -\pi/2, \\ F_X(\tan t) & \text{αν } t \in (-\pi/2, \pi/2), \\ 1 & \text{αν } t \leqslant -\pi/2. \end{cases}$$

Επειδή $F_X(-\infty) = 0$ και $F_X(\infty) = 1$, έπεται ότι η F_Y είναι συνεχής. Επίσης, για $t \in (-\pi/2, \pi/2)$,

$$F_Y'(t) = F_X'(\tan t)(\tan t)' = \frac{1}{\pi(1+\tan^2 t)}(1+\tan^2 t) = \frac{1}{\pi}.$$

Και βέβαια $F_Y'(t) = 0$ για $|t| > \pi/2$. Επομένως η F_Y' υπάρχει και είναι συνεχής στο συμπλήρωμα του πεπερασμένου συνόλου $\{-\pi/2, \pi/2\}$. Με βαση την παραπάνω πρόταση, η Y είναι συνεχής με πυκνότητα, και μια πυκνότητα για την Y είναι η

$$f_Y(t) := \begin{cases} 0 & \text{αν } |t| \geqslant \pi/2, \\ \frac{1}{\pi} & \text{αν } t \in (-\pi/2, \pi/2). \end{cases}$$

Άρα η Y έχει την ομοιόμορφη κατανομή στο $(-\pi/2, \pi/2)$.

Αποδειξη της προτασης

Η απόδειξη είναι άμεση συνέπεια του δεύτερου θεμελιώδους θεωρήματος του ολοκληρωτικού λογισμού και του ορισμού του γενικευμένου ολοκληρώματος. Πάει ως εξής.
Θέλουμε να δείξουμε ότι για κάθε $a < b$ ισχύει

$$F(b) - F(a) = \int_a^b f(t)\,dt. \tag{5}$$

Αν $J = \varnothing$, τότε η (5) έπεται από το δεύτερο θεμελιώδες θεώρημα του ολοκληρωτικού λογισμού ($f = F'$ και η F' είναι συνεχής στο $[a, b]$, και άρα ολοκληρώσιμη). Αν $J \neq \varnothing$, διακρίνουμε τις εξής περιπτώσεις.
Περίπτωση 1. Αν το μόνο σημείο του J στο $[a, b]$ είναι το a, τότε επειδή η F είναι συνεχής έχουμε

$$F(b) - F(a) = \lim_{\varepsilon \to 0^+} \{F(b) - F(a + \varepsilon)\}.$$

Επειδή η F' είναι συνεχής στο $[a + \varepsilon, b]$, το δεύτερο θεμελιώδες θεώρημα του ολοκληρωτικού λογισμού δίνει

$$F(b) - F(a + \varepsilon) = \int_{a+\varepsilon}^b F'(t)\,dt = \int_{a+\varepsilon}^b f(t)\,dt.$$

Το όριο του τελευταίου ολοκληρώματος για $\varepsilon \to 0^+$ υπάρχει γιατί είναι φθίνουσα συνάρτηση του ε (αφού $f \geqslant 0$), και είναι εξ ορισμού η τιμή του $\int_a^b f(t)\,dt$. Για παράδειγμα, είναι δυνατόν να έχουμε $f(a+) = \infty$ (όπως συμβαίνει με την $1/\sqrt{x}$ στο $a = 0$) ή το δεξιό όριο της f στο a να μην υπάρχει.

Περίπτωση 2. Αν το μόνο σημείο του J στο $[a, b]$ είναι το b. Δουλεύουμε όμοια όπως στην προηγούμενη περίπτωση. Πάλι η (5) ισχύει.

Περίπτωση 3. Αν τα μόνα σημεία του J του στο $[a, b]$ είναι τα a, b. Επιλέγουμε $c \in (a, b)$, εφαρμόζουμε την (5) στα διαστήματα $[a, c]$, $[c, b]$ (το οποίο ξέρουμε από τις προηγούμενες περιπτώσεις), και προσθέτουμε κατά μέλη.

Περίπτωση 4. Αν τα σημεία του J ανάμεσα στα a, b είναι τα $x_1, x_2, \ldots, x_k$ με $a \leqslant x_1 < x_2 < \cdots < x_k \leqslant b$, τότε εφαρμόζουμε την (5) στα διαστήματα $[a, x_1], [x_1, x_2], \ldots, [x_k, b]$, και προσθέτουμε τις ισότητες που προκύπτουν κατά μέλη. ■

Είναι προφανές απο την απόδειξη της πρότασης ότι η υπόθεση «J πεπερασμένο» μπορεί να αντικατασταθεί απο την

$$J \cap [a,b] \text{ πεπερασμένο για κάθε πεπερασμένο υποδιάστημα } [a,b] \text{ του } \mathbb{R},$$

δηλαδή το J δεν έχει σημεία συσσώρευσης στο $\mathbb{R}$.

Πιο προχωρημενα αποτελεσματα

Επιστρέφουμε στο ερώτημα που τέθηκε στην αρχή αυτή της παραγράφου. Δηλαδή, αν δίνεται η συνάρτηση κατανομής, F, μιας τυχαίας μεταβλητής, μπορούμε να αποφανθούμε αν αυτή η τυχαία μεταβλητή έχει πυκνότητα; Η ικανή και αναγκαία συνθήκη είναι η F να είναι **απόλυτα συνεχής** (δες Θεώρημα 14.22 στο [1]). Δηλαδή για κάθε $\varepsilon > 0$ υπάρχει $\delta > 0$ ώστε αν το $n \in \mathbb{N}^+$ και $\{(a_i, b_i) : i = 1, 2, \ldots, n\}$ είναι ξένα ανά δύο διαστήματα με $\sum_{i=1}^{n}(b_i - a_i) < \delta$, τότε ισχύει

$$\sum_{i=1}^{n} |F(b_i) - F(a_i)| < \varepsilon. \tag{6}$$

Αυτό ισχύει, για παράδειγμα, αν η F είναι Lipschitz.

Ικανές συνθήκες ώστε να είναι η F (και γενικότερα μια αύξουσα F) απόλυτα συνεχής δίνουν τα ακόλουθα αποτελέσματα. Το δεύτερο είναι γενίκευση του πρώτου και έχει πιο δύσκολη απόδειξη.

Πρόταση 1. Αν η F είναι συνεχής (στο $\mathbb{R}$) και υπάρχει σύνολο $A \subset \mathbb{R}$ χωρίς σημείο συσσώρευσης στο $\mathbb{R}$ ώστε η F' να υπάρχει και να είναι πεπερασμένη για κάθε $x \in \mathbb{R}\backslash A$, τότε η F είναι απόλυτα συνεχής.

Πρόταση 2. Αν η F είναι συνεχής (στο $\mathbb{R}$) και υπάρχει σύνολο $A \subset \mathbb{R}$ αριθμήσιμο ώστε η F' να υπάρχει και να είναι πεπερασμένη για κάθε $x \in \mathbb{R}\backslash A$, τότε η F είναι απόλυτα συνεχής.

Και στις δύο περιπτώσεις μια πυκνότητα είναι η

$$f(x) := \begin{cases} F'(x) & \text{αν } x \in \mathbb{R}\backslash A, \\ 0 & \text{αν } x \in A. \end{cases}$$

Η απόδειξη της Πρότασης 1 έπεται από το Θεώρημα 7.19 και το Πόρισμα 14.9 στο [1] ενώ της 2 από το Πόρισμα 14.9 και την Άσκηση 14-15 στο [1]. Στην ξένη βιβλιογραφία, το θέμα των απολύτως συνεχών συναρτήσεων καλύπτεται, για παράδειγμα, στο Κεφάλαιο 7 του [3] και στην Παράγραφο 18 του [2].

[1] Γ. Κουμουλλής, Σ. Νεγρεπόντης. Θεωρία Μέτρου. Εκδόσεις Συμμετρία.
[2] E. Hewitt-K. Stromberg. Real and abstract analysis. Springer-Verlag, 1969.
[3] R. Kannan-C.K. Krueger. Advanced analysis on the real line. Springer-Verlag, 1996.

3. Κατασκευη τ.μ. ομοιομορφης στο $(0,1)$ με χρηση ενος νομισματος

[Η παράγραφος αυτή δεν χρειάζεται για την εξέταση του μαθήματος. Ωστόσο, βοηθάει στην κατανόηση της συμπεριφοράς των συνεχών τυχαίων μεταβλητών.]

Θεωρούμε ακολουθία ρίψεων ενός αμερόληπτου νομίσματος, και θέτουμε

$$E_i = \begin{cases} 0 & \text{αν η } i \text{ ρίψη έφερε γράμματα,} \\ 1 & \text{αν η } i \text{ ρίψη έφερε κορώνα.} \end{cases}$$

Οι E_i είναι ανεξάρτητες μεταξύ τους, και καθεμία έχει κατανομή Bernoulli με πιθανότητα επιτυχίας $p = 1/2$. Θέτουμε

$$X := \sum_{k=1}^{\infty} \frac{E_k}{2^k}. \tag{7}$$

Δηλαδή ο X γράφεται στο δυαδικό σύστημα ως $X = 0.E_1E_2E_3...$ Αυτός είναι ένας τυχαίος αριθμός στο $[0,1]$. Θα δείξουμε ότι ακολουθεί την ομοιόμορφη κατανομή στο $(0,1)$.

Αλλά πρώτα ας πειραματιστούμε λίγο. Ποιός είναι ο ρόλος καθενός E_k; Αν το E_1 βγεί 0, τότε ο X θα ανήκει στο διάστημα $[0,1/2]$. Γιατί ακόμα και όλα τα υπόλοιπα ψηφία να βγούν 1, ο X θα είναι το πολύ

$$\sum_{n=2}^{\infty} \frac{1}{2^n} = \frac{1}{2}.$$

Αν το E_1 βγεί 1, τότε ο X θα ανήκει στο διάστημα $[1/2,1]$. Γιατί θα είναι τουλάχιστον 1/2, και το πολύ

$$\sum_{n=1}^{\infty} \frac{1}{2^n} = 1.$$

Δηλαδή το E_1 αποφασίζει σε ποιό από τα δύο μισά του $[0,1]$ θα ανήκει ο X. Τώρα για τον ρόλο του E_2. Μόλις ξέρουμε το E_1, χωρίζουμε στην μέση το διάστημα που αυτό διάλεξε. Ας υποθέσουμε ότι $E_1 = 1$, οπότε το διάστημα που διάλεξε είναι το $[1/2,1]$, και τα δύο μισά είναι τα $[1/2,3/4]$, $[3/4,1]$. Το E_2 αποφασίζει σε ποιό από αυτά τα δύο διαστήματα θα ανήκει ο X. $E_2 = 0$ σημαίνει επιλογή του αριστερού διαστήματος (δηλ. του $[1/2,3/4]$), ενώ $E_2 = 1$ σημαίνει επιλογή του δεξιού διαστήματος (δηλ. του $[3/4,1]$). Παρόμοιος είναι ο ρόλος και των επόμενων ψηφίων.

Μετά από αυτή την παρατήρηση, ας υπολογίσουμε την πιθανότητα ο X να πέσει στο διάστημα

$$I := \left(\frac{1}{4}, \frac{1}{4} + \frac{1}{8}\right).$$

Αυτό συμβαίνει ακριβώς όταν $E_1 = 0, E_2 = 1, E_3 = 0$. Δηλαδή πρέπει πρώτα να διαλέξουμε το $[0,1/2]$, έπειτα το δεξιό μισό του, δηλαδή το $[1/4,1/2]$, και τέλος το αριστερό μισό του τελευταίου διαστήματος, που είναι το $[1/4,1/4+1/8]$. Η πιθανότητα αυτού του ενδεχομένου είναι

$$P(E_1 = 0, E_2 = 1, E_3 = 0) = P(E_1 = 0)P(E_2 = 0)P(E_3 = 0) = 1/8.$$

Δηλαδή, $P(X \in I) =$ μήκος του I. Αν δείξουμε το ίδιο αποτέλεσμα για όλα τα υποδιαστήματα του $[0,1]$, τότε θα έχουμε δείξει ότι η X ακολουθεί την ομοιόμορφη κατανομή στο $(0,1)$.

Έστω

$$\mathcal{D} := \left\{ \sum_{i=1}^{k} \frac{\varepsilon_i}{2^i} : k \geqslant 1, \varepsilon_i \in \{0,1\} \text{ για κάθε } i = 1,2,\ldots,k, \varepsilon_k = 1 \right\}.$$

Δηλαδή οι αριθμοί στο $(0,1)$ που έχουν δυαδικό ανάπτυγμα πεπερασμένο. Αυτοί αποτελούν γνήσιο υποσύνολο των ρητών του $(0, 1)$ [είναι οι ρητοί οι οποίοι αν γραφούν σε ανάγωγη μορφή έχουν παρανομαστή δύναμη του 2], αλλά αποτελούν ένα πυκνό σύνολο στο $[0,1]$.

Είναι διαισθητικά προφανές ότι το να φέρνουμε σε μία ακολουθία ρίψεων ενός αμερόληπτου νομίσματος συνεχώς κορώνα έχει πιθανότητα 0. Κάποια στιγμή θα ρθουν γράμματα. Αυτό συμβαίνει και με οποιαδήποτε άλλη συγκεκριμένη ακολουθία, π.χ., την ΚΓΚΓΚΓΚΓΚΓ... Αυτό θα αποδείξουμε τώρα.

Πρόταση 1. *Έστω μία ακολουθία* $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots,)$ *με* $\varepsilon_i \in \{0,1\}$ *για κάθε* i*, δηλαδή μιά άπειρη ακολουθία από 0 και 1. Τότε*

$$P(E_i = \varepsilon_i \text{ για κάθε } i \geqslant 1) = P(E_1 = \varepsilon_1, E_2 = \varepsilon_2, E_3 = \varepsilon_3, ...) = 0$$

Απόδειξη. Για $n \geqslant 1$ σταθερό, η δοσμένη πιθανότητα, έστω q, φράσσεται ως εξής.

$$q = P(E_i = \varepsilon_i \text{ για κάθε } i \geqslant 1) \leqslant P(E_1 = \varepsilon_1, E_2 = \varepsilon_2, \ldots, E_n = \varepsilon_n) = \frac{1}{2^n}.$$

Επειδή αυτό ισχύει για κάθε $n \geqslant 1$, έπεται ότι $q = 0$. ∎

Έστω F η συνάρτηση κατανομή της τυχαίας μεταβλητής X. Δηλαδή $F(x) = P(X \leqslant x)$ για κάθε $x \in \mathbb{R}$. Θα δείξουμε τώρα ότι η F ισούται με την συνάρτηση κατανομής της ομοιόμορφης κατανομής.

Πρόταση 2. *(α) Για κάθε* $c \in [0,1]$ *ισχύει* $P(X = c) = 0$.

(β) Για κάθε $a \in \mathcal{D}$*, ισχύει* $F(a) = a$.

(γ) Για κάθε $x \in (0,1)$*, ισχύει* $F(x) = x$.

Απόδειξη. (α) Αν ο c έχει μοναδική αναπαράσταση στο δυαδικό σύστημα, έστω $c = \sum_{i=1}^{\infty} \frac{c_i}{2^i}$, τότε

$$P(X = c) = P(E_i = c_i \text{ για κάθε } i \geqslant 1) = 0$$

από την προηγούμενη πρόταση.
Υπάρχουν όμως και αριθμοί με δύο αναπαραστάσεις. Αυτοί είναι τα στοιχεία του $\mathcal{D}$. Η μία αναπαράσταση είναι της μορφής $\sum_{i=1}^{k} \frac{\varepsilon_i}{2^i}$, για κάποιο $k \geqslant 1$ και με $\varepsilon_k = 1$, και η άλλη είναι η

$$\sum_{i=1}^{k-1} \frac{\varepsilon_i}{2^i} + \sum_{i=k+1}^{\infty} \frac{1}{2^i}$$

αφού το τελευταίο άθροισμα ισούται με 2^{-k}. Έστω $(c_i)_{i\geqslant 1}, (\hat{c}_i)_{i\geqslant 1}$ οι ακολουθίες ψηφίων των δύο αυτών αναπαραστάσεων. Τότε

$$P(X = c) = P(E_i = c_i \text{ για κάθε } i \geqslant 1) + P(E_i = \hat{c}_i \text{ για κάθε } i \geqslant 1) = 0$$

πάλι λόγω της προηγούμενης πρότασης.

(β) Έστω ότι

$$a = \sum_{i=1}^{k} \frac{\varepsilon_i}{2^i}$$

για κάποιο $k \geqslant 1$ και με $\varepsilon_k = 1$. Ο X θα προκύψει μικρότερος ή ίσος του a αν και μόνο αν υπάρχει δείκτης $r \geqslant 1$ ώστε $\varepsilon_r = 1$, $E_j = \varepsilon_j$ για $1 \leqslant j < r$, και $E_r = 0 < 1 = \varepsilon_r$. Δηλαδή, τα ψηφία τους συμφωνούν αρχικά ώσπου σε κάποια θέση που ο a έχει 1, ο X έχει 0. Γιατί προφανώς αν κάπου που ο a έχει 0 ο X εμφανίσει 1, θα έχουμε αμέσως $X > a$. Και αν ο X έχει 1 σε όλα τα σημεία που ο a έχει 1, τότε κάποια στιγμή μετά το k ψηφίο, ο X θα εμφανίσει 1 (αφού η ακολουθία 0000000... έχει πιθανοτητα 0 λόγω της προηγούμενης πρότασης) και θα βρεθεί πάνω από τον a. Θέτουμε λοιπόν

$T := \min\{j \geqslant 1 : E_j < \varepsilon_j\}$, και ονομάζουμε $j_1 < j_2 < \cdots < j_s$ τους δείκτες i που έχουν $\varepsilon_i = 1$. Με βάση την προηγούμενη συζήτηση, έχουμε

$$\begin{aligned} P(X \leqslant a) &= P(T = j_1) + P(T = j_2) + \cdots + P(T = j_s) \\ &= \frac{1}{2^{j_1}} + \frac{1}{2^{j_1}} + \cdots + \frac{1}{2^{j_s}} = a. \end{aligned}$$

Χρησιμοποιήσαμε το ότι $P(T = j) = 1/2^j$. Αυτό ισχύει γιατί

$$P(T = j) = P(E_1 = \varepsilon_1, \cdots, E_{j-1} = \varepsilon_{j-1}, E_j = 0) = \frac{1}{2^j}.$$

(γ) Αυτό προκύπτει από το (β) και την μονοτονία της F. Παίρνουμε $x \in (0, 1)$, και ακολουθίες $(a_n)_{n\geqslant 1}, (b_n)_{n\geqslant 1}$ στοιχείων του $\mathcal{D}$ ώστε $a_n \leqslant x \leqslant b_n$ και $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = x$. Αυτό μπορεί να γίνει γιατί το $\mathcal{D}$ είναι πυκνό στο (0, 1). Η μονοτονία της F και το (β) δίνουν

$$a_n \leqslant x \leqslant b_n \Rightarrow F(a_n) \leqslant F(x) \leqslant F(b_n) \Rightarrow a_n \leqslant F(x) \leqslant b_n.$$

Και για $n \to \infty$, παίρνουμε $F(x) = x$. ■

Μένει να προσδιορίσουμε τις τιμές της F στο $(-\infty, 0] \cup [1, \infty)$. Επειδή $X \in [0, 1]$ πάντοτε, έχουμε αμέσως ότι $F(x) = 0$ για $x < 0$, $F(x) = 1$ για $x \geqslant 1$, και

$$F(0) = P(X \leqslant 0) = P(X = 0) = P(E_i = 0 \text{ για κάθε } i \geqslant 1) = 0$$

με βάση την Πρόταση 1.

Παρατήρηση. Η πιο πάνω αντιστοίχιση μιας ομοιόμορφης τυχαίας μεταβλητής με μια ακολουθία ρίψεων ενός νομίσματος βοηθάει να καταλάβουμε πώς είναι δυνατόν μια τυχαία μεταβλητή να παίρνει κάθε της τιμή με πιθανότητα 0. Η Πρόταση 1 διατυπώνει αυτό το φαινόμενο στην γλώσσα των ρίψεων, και κάνει πιο απτό το φαινόμενο.

Άσκηση. (α) Άν έχουμε στην διάθεσή μας έναν μηχανισμό που να παράγει μια ακολουθία ανεξάρτητων ρίψεων ενος δίκαιου νομίσματος, να κατασκευαστούν δύο ανεξάρτητες τυχαίες μεταβλητές καθεμία με κατανομή την ομοιόμορφη στο $(0, 1)$.

Υποδ.: Με E_i όπως πιο πάνω, θεωρούμε τους τυχαίους αριθμούς X_1, X_2 γραμμένους στο δυαδικό σύστημα ως εξής

$$\begin{aligned} X_1 &:= 0.E_1E_3E_5\ldots, \\ X_2 &:= 0.E_2E_4E_6\ldots \end{aligned}$$

(β) Με τα δεδομένα του προηγούμενου ερωτήματος, να κατασκευαστεί μια ακολουθία $(X_n)_{n\geqslant 1}$ από ανεξάρτητες τυχαίες μεταβλητές καθεμία με κατανομή την ομοιόμορφη στο $(0, 1)$.

(γ) Αν έχουμε στην διάθεση μας μία τυχαία μεταβλητή με κατανομή την ομοιόμορφη στο (0, 1). Να κατασκευαστεί από αυτήν μια ακολουθία $(X_n)_{n\geqslant 1}$ όπως στο προηγούμενο ερώτημα.