# Συμπληρωματα

## 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 $x \in \mathbb{R}$,

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

και παρατηρούμε ότι $x^+, x^- \geq 0$ 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 $\{1, 2, \ldots, n\}$ για κάποιο $n \in \mathbb{N}$) και $x_i \geq 0$ για κάθε $i \in I$. Τότε θέτουμε[2]

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

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

**Περίπτωση 2**. Αν η τυχαία μεταβλητή $X$ παίρνει μόνο μή θετικές τιμές[3] $\{x_i : i \in I\}$ όπου $I$ είναι ένα αριθμήσιμο σύνολο και $x_i \leq 0$ για κάθε $i \in I$. Τότε θέτουμε πάλι

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

το οποίο ή είναι ένας αριθμός στο $(-\infty, 0]$ ή είναι $-\infty$. Ορίζεται πάντοτε.

**Περίπτωση 3**. 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 $\infty$ τότε θέτουμε

$$E(X) := E(X^+) - E(X^-)$$

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

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

αν τουλάχιστον ένα απο τα δύο ολοκληρώματα είναι $\neq \infty$, και τότε η $E(X)$ είναι πραγματικός αριθμός, ή $-\infty$, ή $\infty$. Αν και τα δύο ολοκληρώματα είναι $\infty$, τότε η $E(X)$ δεν ορίζεται. Βέβαια αν $X \geq 0$ ή $X \leq 0$,

[1] Γράφουμε τότε $X \geq 0$.
[2] Δεν έχει σημασία η σειρά άθροισης γιατί είναι όλοι αριθμοί με ίδιο πρόσημο.
[3] Γράφουμε τότε $X \leq 0$.

τότε η $E(X)$ ορίζεται οπωσδήποτε. Και γενικότερα, όταν για κάποιο $a \in \mathbb{R}$ έχουμε $X(\omega) \geq a$ για κάθε $\omega \in \Omega$ ή $X(\omega) \leq a$ για κάθε $\omega \in \Omega$.

Παραδειγματα

**1**. Θεωρούμε διακριτή τυχαία μεταβλητή με συνάρτηση πιθανότητας

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

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)

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

Έχουμε

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

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

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

Άρα η $E(X)$ δεν ορίζεται.

## 2. Πυκνοτητα συνεχους κατανομης[4]

Συνεχή κατανομή λέμε μία 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

$$F(b) - F(a) = \int_a^b f(t)\, dt$$

για κάθε $a, b \in \mathbb{R}$ με $a < b$. Μια τέτοια $f$ λέγεται *πυκνότητα* για την κατανομή, 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 $f \ge 0, \int_{\mathbb{R}} f(x)\, dx = 1$.

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 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 $t \in \mathbb{R} \setminus \{a_1, a_2, \ldots, a_n\}$,
- $f$ συνεχής στο $\mathbb{R} \setminus \{a_1, a_2, \ldots, a_n\}$.

Τότε η $f$ είναι μιά πυκνότητα που αντιστοιχεί στην 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 $\mathbb{R} \setminus J_1$, ή αν είναι ασυνεχής, να είναι ασυνεχής σε πεπερασμένο σύνολο σημείων $J_2$ του $\mathbb{R} \setminus J_1$. Θέτουμε τότε $J := J_1 \cup J_2$ και $f := F'$ στο $\mathbb{R} \setminus J$. 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 $(-\pi/2, \pi/2)$. Ποιά είναι η κατανομή της $Y$;

**Λύση**

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

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

Άρα

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

[4] Αυτή η παράγραφος είναι χρήσιμη για την πλήρη δικαιολόγηση ενός συγκεκριμένου τύπου ασκήσεων, όπως η εφαρμογή παρακάτω.

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

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

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

Αποδειξη της προτασης

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

$$F(b)-F(a)=\int_a^b f(t)\,dt. \tag{1}$$

**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 $\varepsilon$ (καθότι $f\geq 0$), 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 (1) ισχύει.

**Περίπτωση 3**. Αν τα μόνα σημεία του $J$ του στο $[a,b]$ είναι τα $a,b$. Επιλέγουμε $c\in(a,b)$, εφαρμόζουμε την (1) 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 $a\leq x_1<x_2<\cdots<x_k\leq b$, τότε εφαρμόζουμε την (1) 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 $\mathbb{R}$.