Αναμενόμενη Τιμή και Διακύμανση Τυχαίας Μεταβλητής

Περιεχόμενα

Έστω ο χώρος πιθανότητας $(\Omega, \mathcal{F}, P)$ και η τυχαία μεταβλητή $X : \Omega \to \mathbb{R}$. Η αναμενόμενη τιμή $E(X)$ της τ.μ. X ορίζεται ως εξής

$$E(X) = \int_\Omega X(\omega) P(d\omega)$$

ή

$$E(X) = \int_\Omega X(\omega) P(d\omega) = \int_R x P_X(dx) = \int_R x dP_X,$$

Σημ.: Η αναμενόμενη τιμή της πραγματικής τ.μ.$X : \Omega \to \mathbb{R}$ υπάρχει στον $\mathbb{R}$ αν και μόνο αν $E(|X|) < \infty$.

Βασικές Ιδιότητες

Έστω δύο τυχαίες μεταβλητές X, Y και $c \in \mathbb{R}$, τότε ισχύουν οι παρακάτω ιδιότητες

α) $E(c) = c$

β) $E(cX) = c \cdot E(X)$

γ) $E(X+Y) = E(X) + E(Y)$

δ) Αν $X \leq Y$ τότε $E(X) \leq E(Y)$

ε) $E(\alpha X + \beta Y) = E(\alpha X) + E(\beta Y) = \alpha \cdot E(X) + \beta \cdot E(Y)$

Συντελεστής Ασυμμετρίας

Έστω η τ.μ. X για την οποία ισχύει $E(|X|)^3 < \infty$ και έχει μέση τιμή μ και τυπική απόκλιση σ. Ο συντελεστής ασυμμετρίας γ_3 ορίζεται από τη σχέση

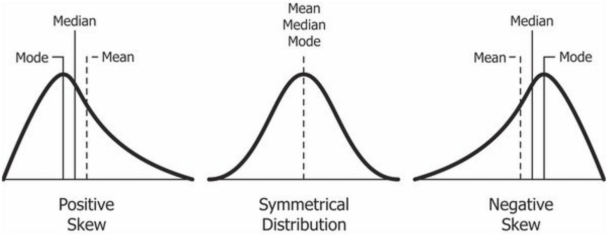
$$\gamma_3 = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right]$$

Οι κατανομές γενικά διακρίνονται σε:

- συμμετρικές ($\gamma_3 = 0$)
- θετικά ασύμμετρες ($\gamma_3 > 0$)
- αρνητικά ασύμμετρες ($\gamma_3 < 0$)

Επίσης είναι

$$\gamma_3 = \frac{1}{\sigma^3} E(X - \mu)^3 = \frac{\mu_3}{\sigma^3}$$


Median
Mode
Mean
Positive Skew
Mean
Median
Mode
Symmetrical Distribution
Median
Mean
Mode
Negative Skew

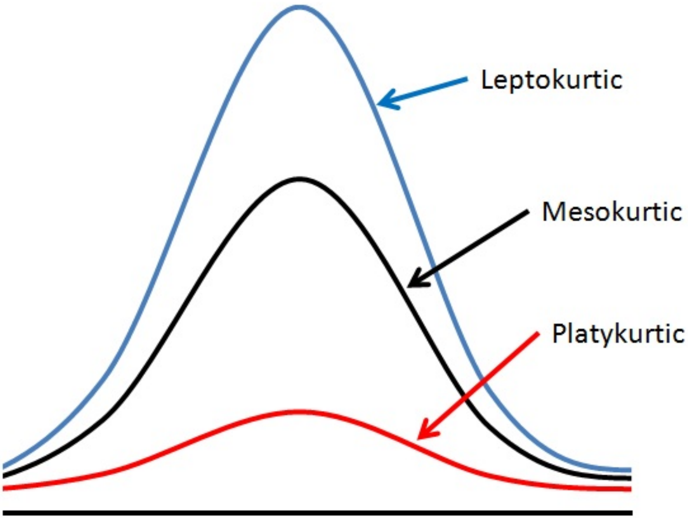
Συντελεστής Κύρτωσης

Έστω η τ.μ. Χ για την οποία ισχύει $E(|X|)^4 < \infty$ και έχει μέση τιμή μ και τυπική απόκλιση σ. Ο συντελεστής κύρτωσης υπολογίζεται από την παρακάτω σχέση

$$\gamma_4 = E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right]$$

Η κύρτωση χαρακτηρίζει την αιχμηρότητα της καμπύλης μιας κατανομής.
Με βάση την κύρτωση, οι κατανομές διακρίνονται σε:

- λεπτόκυρτες ($\gamma_4 > 3$)
- μεσόκυρτες ($\gamma_4 = 3$)
- πλατύκυρτες ($\gamma_4 < 3$)


Leptokurtic
Mesokurtic
Platykurtic

Αναμενόμενη Τιμή Διακριτής Τυχαίας Μεταβλητής

Έστω η διακριτή τ.μ. $X : \Omega \to \mathbb{R}, R_X = \{x_1, x_2, x_3, \ldots\}$ με συνάρτηση πιθανότητας $f_X(x), x \in \mathbb{R}$ και ότι το ολοκλήρωμα $\int_{\mathbb{R}} x dP_X$ υπάρχει στον $\mathbb{R}$.

Τότε η αναμενόμενη τιμή της τ.μ. X είναι

$$E(X) = \sum_{x \in R_X} x f_X(x) = \sum_{x} x P(X = x).$$

Αναμενόμενη Τιμή Συνεχούς Τυχαίας Μεταβλητής

Έστω η συνεχής τ.μ. X και ότι υπάρχει η αναμενόμενη τιμή της X στο $\mathbb{R}$, τότε

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$$

Θεώρημα

Αν ισχύει $|x| < c, \forall\, x \in R_x$ για κάποιο $c \in (0, +\infty)$, τότε η αναμενόμενη τιμή $E(X)$ υπάρχει.

Το θεώρημα δείχνει ότι η αναμενόμενη τιμή υπάρχει για οποιαδήποτε τυχαία μεταβλητή της οποίας τα αποτελέσματα είναι φραγμένα σε απόλυτη τιμή, επειδή το αντίστοιχο άθροισμα ή ολοκλήρωμα συγκλίνει.

Έλεγχος ύπαρξης αναμενόμενης τιμής- ποια είναι τα βήματα;

Αναμενόμενη Τιμή Συνάρτησης Τυχαίας Μεταβλητής

Έστω ο χώρος πιθανότητας $\Omega, \mathcal{F}, P$ και η τ.μ. $X : \Omega \to \mathbb{R}$ ορισμένη σε αυτό το χώρο. Έστω P_X η κατανομή της X και ότι η $g : \mathbb{R} \to \mathbb{R}$ είναι μια Borel μετρήσιμη συνάρτηση. Η μέση τιμή της συνάρτησης $g(X)$ ορίζεται ως εξής

$$E[g(X)] = \begin{cases} \sum_j g(x_j)P(X = x_j) & \text{X διακριτή τ.μ.} \\ \\ \int_{\mathbb{R}} g(x) f_X(x) dx & \text{X συνεχής τ.μ.} \end{cases}$$

Η αναμενόμενη τιμή $E[g(X)]$ υπάρχει στον $\mathbb{R}$ (δηλ. η $g(X)$ ειναι P-ολοκληρώσιμη) αν και μόνο αν ισχύει $E[|g(X)|] < \infty$.

- Αν η τ.μ. X είναι διακριτή τότε

$$E[g(X)] = \sum_j g(x_j) P(X = x_j)$$

- Αν η τ.μ. X είναι συνεχής τότε

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$$

Κριτήρια της P-ολοκληρωσιμότητας της $g(X)$

- Αν η τ.μ. X είναι διακριτή και η g συνεχής

$$E[g(X)] = \sum_j |g(x_j)| P(X = x_j) < \infty$$

- Αν η τ.μ. X είναι συνεχής και η g συνεχής

$$E[g(X)] = \int_{-\infty}^{\infty} |g(x)| f_X(x) dx < \infty$$

Ανισότητα Jensen

Συγκρίνει την αναμενόμενη τιμή μιας συνάρτησης μιας τ.μ. με τη συνάρτηση της αναμενόμενης τιμής της τ.μ. για μια συγκεκριμένη οικογένεια συναρτήσεων g.

Έστω ότι η συνάρτηση $g : I \to \mathbb{R}$ είναι:

- **Κυρτή**, τότε ισχύει:

$$g(E(X)) \leq E(g(X)).$$

- **Κοίλη**, τότε ισχύει:

$$g(E(X)) \geq E(g(X)).$$

Για $\alpha \in [0, 1]$ θα έχουμε:

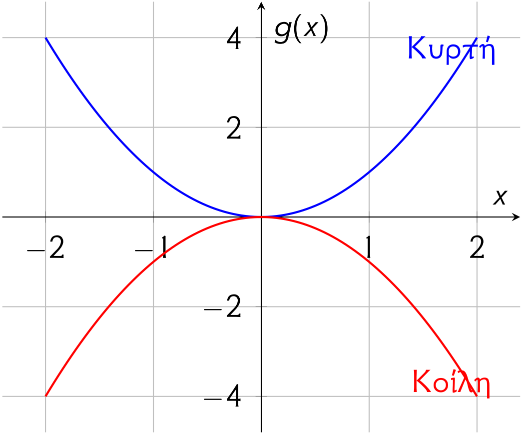
Κυρτή συνάρτηση:

$$g(\alpha x + (1 - \alpha)y) \leq \alpha g(x) + (1 - \alpha)g(y).$$

Κοίλη συνάρτηση:

$$g(\alpha x + (1 - \alpha)y) \geq \alpha g(x) + (1 - \alpha)g(y).$$

Κυρτή και Κοίλη Συνάρτηση



Για να χρησιμοποιήσουμε την ανισότητα του Jensen, πρέπει να προσδιορίσουμε εάν μια συνάρτηση g είναι κυρτή ή κοίλη.

Μια διπλά διαφοροποιήσιμη συνάρτηση $g : I \to \mathbb{R}$ είναι κυρτή αν και μόνο αν $g''(x) \geq 0$ για όλα τα $x \in I$.

Μια διπλά διαφοροποιήσιμη συνάρτηση $g : I \to \mathbb{R}$ είναι κοίλη αν και μόνο αν $g''(x) < 0$ για όλα τα $x \in I$.

Θεώρημα

Έστω μία τυχαία μεταβλητή X και k συναρτήσεις αυτής $g_1(X), g_2(X), \ldots, g_k(X)$. Τότε ισχύει η σχέση:

$$E\left(\sum_{i=1}^{k} g_i(X)\right) = \sum_{i=1}^{k} E(g_i(X)).$$

Απόδειξη

Θέτουμε $g(X) = \sum_{i=1}^{k} g_i(X)$, οπότε:

$$E\left[\sum_{i=1}^{k} g_i(X)\right] = \int_{\mathbb{R}} \sum_{i=1}^{k} g_i(x) P_X(x) dx$$
$$= \sum_{i=1}^{k} \int_{\mathbb{R}} g_i(x) P_X(x) dx.$$

Από τον ορισμό της αναμενόμενης τιμής κάθε $g_i(X)$, προκύπτει:

$$\sum_{i=1}^{k} E[g_i(X)].$$

Άρα

$$E\left[\sum_{i=1}^{k} g_i(X)\right] = \sum_{i=1}^{k} E[g_i(X)].$$

Διακύμανση τυχαίας μεταβλητής

- $$Var(X) = E[X^2] - [EX]^2 = E[(X - \mu_X)^2]$$
- Έστω X τ.μ. και α, β σταθερές
$$Var(\alpha X + \beta) = \alpha^2 Var(X)$$
- X διακριτή τ.μ.
$$Var(X) = \sum_i (x_i - \mu)^2 \cdot P(X = x_i).$$
- X συνεχής τ.μ.
$$Var(X) = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx = \int_{-\infty}^{\infty} x^2 f_X(x) dx - \mu_X^2$$

Αναμενόμενη Τιμή

- Είναι μια θεωρητική ποσότητα που υπολογίζεται από τη συνάρτηση κατανομής μιας τυχαίας μεταβλητής.
- Αντικατοπτρίζει τον «κεντρικό» ή «αναμενόμενο» μέσο όρο των αποτελεσμάτων σε άπειρα επαναλαμβανόμενα δείγματα.

Ορισμός:

- Για διακριτή τυχαία μεταβλητή:

$$E[X] = \sum_i x_i P(X = x_i)$$

- Για συνεχή τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας $f(x)$:

$$E[X] = \int x f(x) dx$$

Μέση Τιμή

- Υπολογίζεται από ένα **συγκεκριμένο δείγμα** δεδομένων και όχι από τη θεωρητική κατανομή.
- Ορίζεται ως ο αριθμητικός μέσος των παρατηρήσεων:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

- Είναι μια εκτίμηση της αναμενόμενης τιμής $E[X]$ της αντίστοιχης τυχαίας μεταβλητής.