

Μέση Τιμή

Έστω X τ.μ. και $f_X(x)$ η σ.π. ή σ.π.π. της X

Μέση ή αναμενόμενη τιμή της X είναι ο αριθμός:

$$E(X) = \mu_X = \sum xf_X(x) \quad \text{αν η } X \text{ είναι διακριτή, και}$$

$$E(X) = \mu_X = \int_{-\infty}^{\infty} xf_X(x)dx \quad \text{αν η } X \text{ είναι συνεχής.}$$

Παραδείγματα:

(1) Έστω X ο αριθμός που φέρνει το ζάρι. Η μέση τιμή της X είναι:

$$E(X) = 1\frac{1}{6} + 2\frac{1}{6} + 3\frac{1}{6} + 4\frac{1}{6} + 5\frac{1}{6} + 6\frac{1}{6} = \frac{21}{6} = 3.5$$

(2) Έστω X τ.μ. με σ.π.π $f(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & \text{αλλού} \end{cases}$

Η αναμενόμενη τιμή της X είναι:

$$E(X) = \int_{-\infty}^{+\infty} xf_X(x)dx = \int_0^1 x2xdx = \int_0^1 2x^2dx = \frac{2x^3}{3}\bigg|_0^1 = \frac{2}{3}$$

Μέση τιμή κατανομής Bernulli

Η τ.μ X έχει σ.π. $f(x) = P(X = x) = \begin{cases} p & x = 1 \\ q & x = 0 \\ 0 & \text{αλλού} \end{cases}$ όπου $q = 1 - p$
και μέση τιμή: $E(X) = 1(p) + 0(q) = p$

Μέση τιμή διωνυμικής κατανομής

Η τ.μ X έχει σ.π. $f(x) = \binom{n}{x} p^x (1-p)^{n-x}$ όταν $x = 0, 1, 2, \dots, n$

και μέση τιμή: $E(X) = \sum_{x=0}^n x \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=1}^n x \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} =$
 $= \sum_{x=1}^n \frac{n!}{(x-1)!(n-x)!} p^x (1-p)^{n-x} \stackrel{(y=x-1)}{=} np \sum_{y=0}^{n-1} \frac{(n-1)!}{(y)!(n-1-y)!} p^y (1-p)^{n-1-y} =$
 $= np \sum_{y=0}^{n-1} \binom{n-1}{y} p^y (1-p)^{n-1-y} = np(p+1-p)^{n-1} = np$

Μέση τιμή κατανομής Poisson

Η τ.μ X έχει σ.π. $f(x) = P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$ όταν $x = 0, 1, 2, 3, \dots$ και $\lambda > 0$

και μέση τιμή:
$$E(X) = \sum_{x=0}^{\infty} x \frac{e^{-\lambda} \lambda^x}{x!} = \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} \stackrel{(y=x-1)}{=} \\ = \lambda e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^y}{(y)!} = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

Μέση τιμή γεωμετρικής κατανομής

Η τ.μ X έχει σ.π. $f(x) = P(X = x) = (1-p)^{x-1} p$ όταν $x = 1, 2, \dots$

και μέση τιμή:
$$E(X) = \sum_{x=1}^{\infty} x(1-p)^{x-1} p = p \sum_{x=1}^{\infty} x(1-p)^{x-1} = p \frac{1}{(1-(1-p))^2} = \frac{1}{p}$$

Μέση τιμή εκθετικής κατανομής

Η τ.μ X έχει σ.π. $f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & \text{αλλού} \end{cases}$

και μέση τιμή: $E(X) = \int_{-\infty}^{+\infty} x f_X(x) dx = \int_0^{\infty} x \lambda e^{-\lambda x} dx =$
 $= -x e^{-\lambda x} \Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda x} dx = -\frac{1}{\lambda} e^{-\lambda x} \Big|_0^{\infty} = \frac{1}{\lambda}$

Μέση τιμή ομοιόμορφης κατανομής

Η τ.μ X έχει σ.π. $f_X(x) = \begin{cases} \frac{1}{\beta - \alpha} & \alpha < x < \beta \\ 0 & \text{αλλού} \end{cases}$

και μέση τιμή: $E(X) = \int_{-\infty}^{+\infty} x f_X(x) dx = \int_a^b x \frac{1}{b-a} dx = \frac{1}{b-a} \frac{x^2}{2} \Big|_a^b = \frac{(b^2 - a^2)}{2(b-a)} = \frac{(b+a)}{2}$

Μέση Τιμή (συνέχεια)

Για ορισμένες τ.μ. η μέση τιμή δεν υπάρχει διότι απειρίζεται.

Παράδειγμα: Έστω τ.μ. X με σ.π.

$$f_X(x) = \begin{cases} \frac{k}{x^2} & x = 1, 2, 3, 4, \dots \\ 0 & \text{αλλού} \end{cases}$$

Το k θα πάρει μια τιμή έτσι ώστε $\sum_{x=1}^{\infty} \frac{k}{x^2} = 1$.

Η σειρά $\sum_{x=1}^{\infty} \frac{1}{x^2}$ συγκλίνει (αρμονική) και το k μπορεί να οριστεί.

Η μέση τιμή της τ.μ. είναι: $E(X) = \sum_{x=1}^{\infty} x \frac{k}{x^2} = \sum_{x=1}^{\infty} \frac{k}{x}$.

Η σειρά $\sum_{x=1}^{\infty} \frac{1}{x}$ δεν συγκλίνει (αρμονική) και η μέση τιμή δεν ορίζεται.

Θεώρημα αφηρημένου στατιστικού

Έστω X μια τ.μ. και $f_X(x)$ η σ.π. ή σ.π.π. της X . Εάν $Y = g(X)$ μια συνάρτηση της τ.μ., τότε και η Y είναι τ.μ. και η αναμενόμενη της τιμή υπολογίζεται σύμφωνα με τον ορισμό:

$$E(Y) = \mu_Y = \sum_y y f_Y(y) \quad \text{ή} \quad E(Y) = \mu_Y = \int_{-\infty}^{\infty} y f_Y(y) dy$$

Όμως εάν η σ.π. ή σ.π.π της Y δεν είναι γνωστή μπορεί να υπολογιστεί και με τη βοήθεια της $f_X(x)$, ως εξής:

$$E(Y) = E[g(X)] = \sum_x g(x) f_X(x) \quad \text{αν η } X \text{ είναι διακριτή, και}$$
$$E(Y) = E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx \quad \text{αν η } X \text{ είναι συνεχής.}$$

Ιδιότητες Μέσης Τιμής

1. $E(\alpha X + \beta) = \alpha E(X) + \beta$ (γραμμική ιδιότητα)

2. $E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$ (γραμμική ιδιότητα)

3. Αν οι τ.μ. X και Y είναι ανεξάρτητες τότε ισχύει:

$$E(XY) = E(X) E(Y)$$

Διασπορά

Έστω X τ.μ. και $f_X(x)$ η σ.π. ή σ.π.π. της X

Διασπορά ή διακύμανση μιας τ.μ. X είναι ο θετικός αριθμός

$$Var(X) = E[(X - \mu)^2]$$

και είναι μέτρο μεταβλητότητας.

Σύμφωνα με το θεώρημα του αφηρημένου στατιστικού:

$$Var(X) = \sigma_X^2 = \sum (x - \mu_X)^2 f_X(x) \quad \text{αν η } X \text{ είναι διακριτή, και}$$

$$Var(X) = \sigma_X^2 = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx \quad \text{αν η } X \text{ είναι συνεχής.}$$

$$\text{Τυπική απόκλιση της } X: \sigma_X = \sqrt{Var(X)} = \sqrt{E[(X - \mu)^2]}$$

(**Σημ.** Η τυπική απόκλιση μετριέται με τη μονάδα μέτρησης της X .)

Διασπορά (συνέχεια)

Παράδειγμα:

(1) Έστω τ.μ. X με σ.π.

$f_X(3)=0.2, f_X(4)=0.6, f_X(5)=0.2$, και $f_X(x)=0$ οπουδήποτε αλλού.

Αναμενόμενη τιμή της X : $E(X) = 3(0.2) + 4(0.6) + 5(0.2) = 4$

Διασπορά της X : $\text{Var}(X) = (3-4)^2(0.2) + (4-4)^2(0.6) + (5-4)^2(0.2) = 0.4$

Έστω δεύτερη τ.μ. Y με σ.π.

$f_X(1)=0.4, f_X(2)=0.1, f_X(6)=0.3, f_X(8)=0.2$ και $f_X(x)=0$ οπουδήποτε αλλού.

Αναμενόμενη τιμή της Y : $E(X) = 1(0.4) + 2(0.1) + 6(0.3) + 8(0.2) = 4$

Διασπορά της Y : $\text{Var}(X) = (1-4)^2(0.4) + (2-4)^2(0.1) + (6-4)^2(0.3) + (8-4)^2(0.2) = 8.4$

Ροπές

1. Ροπή γύρω από την αρχή (μηδέν): $\mu'_r = E(X^r)$

Ισχύει ότι: $\mu'_0 = 1$, $\mu'_1 = E(X)$ και $\mu'_2 = E(X^2)$.

2. Κεντρική ροπή ή ροπή γύρω από τη μέση τιμή:

$$\mu_r = E[(X - \mu)^r]$$

Ισχύει ότι: $\mu_0 = 1$, $\mu_1 = 0$ και $\mu_2 = E[(X - \mu)^2] = Var(X)$.

Σχέσεις μεταξύ ροπών

$$\mu_k = E[(X - \mu)^k] = E\left[\sum_{r=0}^k \binom{k}{r} X^r (-\mu)^{k-r}\right] = \sum_{r=0}^k \binom{k}{r} (-\mu)^{k-r} E(X^r) = \sum_{r=0}^k \binom{k}{r} (-\mu)^{k-r} \mu'_r$$

$$\mu'_k = E[X^k] = E[(X - \mu + \mu)^k] = E\left[\sum_{r=0}^k \binom{k}{r} (X - \mu)^r \mu^{k-r}\right] = \sum_{r=0}^k \binom{k}{r} \mu^{k-r} \mu_r$$

Ιδιότητες Διασποράς

1. Από τις σχέσεις μεταξύ ροπών ισχύει ότι:

$$\begin{aligned} \text{Var}(X) &= E[(X - \mu)^2] = E[X^2 - 2X\mu + \mu^2] \\ &= E(X^2) - 2E(X)\mu + \mu^2 = E(X^2) - [E(X)]^2 \end{aligned}$$

2. $\text{Var}(aX + b) = a^2 \text{Var}(X) \quad \forall a, b \in \mathbb{R}$

3. Αν οι X και Y είναι ανεξάρτητες τ.μ. τότε

$$\text{Var}(aX + bY) = a^2 \text{Var}(X) + b^2 \text{Var}(Y) \quad \forall a, b \in \mathbb{R}$$

Διασπορά κατανομής Bernulli

Η τ.μ X έχει σ.π. $f_X(1) = p$, $f_X(0) = 1 - p$ και $f_X(x) = 0$ αλλού

Μέση τιμή: $E(X) = p$ και $E(X^2) = 1(p) + 0(1 - p) = p$

Συνεπώς: $\text{Var}(X) = p - p^2 = p(1 - p) = pq$

Διασπορά διωνυμικής κατανομής

$X \sim \mathcal{B}(n, p)$ όταν $f(x) = \binom{n}{x} p^x (1 - p)^{n-x}$ για $x = 0, 1, 2, \dots, n$

Μέση τιμή: $E(X) = np$ και $E(X^2) = n(n-1)p^2 + np$, οπότε

$$\text{Var}(X) = n(n-1)p^2 + np - n^2p^2 = np(1 - p) = npq$$

Υπολογισμός διασποράς διωνυμικής

$$\begin{aligned} E(X(X-1)) &= \sum_{x=0}^n x(x-1) \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=2}^n \frac{n!}{(x-2)!(n-x)!} p^x (1-p)^{n-x} = \\ &= n(n-1)p^2 \sum_{x=2}^n \frac{(n-2)!}{(x-2)!(n-x)!} p^{x-2} (1-p)^{n-x} \stackrel{(y=x-2)}{=} \\ &= n(n-1)p^2 \sum_{y=0}^{n-2} \frac{(n-2)!}{(y)!(n-2-y)!} p^y (1-p)^{n-2-y} = n(n-1)p^2 (p+1-p)^{n-2} = n(n-1)p^2 \end{aligned}$$

$$E(X^2) = E[X(X-1)] + E(X) = n(n-1)p^2 + np$$

Διασπορά εκθετικής κατανομής

Η τ.μ X έχει σ.π. $f_X(x) = \lambda e^{-\lambda x}$, όταν $x > 0$

Μέση τιμή: $E(X) = \frac{1}{\lambda}$ και

$$\begin{aligned} E(X^2) &= \int_0^{\infty} x^2 \lambda e^{-\lambda x} dx = -x^2 e^{-\lambda x} \Big|_0^{\infty} + \int_0^{\infty} 2x e^{-\lambda x} dx = \\ &= -\frac{2}{\lambda} x e^{-\lambda x} \Big|_0^{\infty} + \frac{2}{\lambda} \int_0^{\infty} e^{-\lambda x} dx = -\frac{2}{\lambda} \frac{1}{\lambda} e^{-\lambda x} \Big|_0^{\infty} = \frac{2}{\lambda^2} \end{aligned}$$

Συμπεπώς $\text{Var}(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$ και $\sigma_X = \frac{1}{\lambda}$

Συνδιασπορά

Έστω ότι έχουμε δύο τυχαίες μεταβλητές X και Y από κοινού κατανομημένες. Ονομάζουμε *συνδιασπορά* ή *συνδιακύμανση* την ποσότητα:

$$\text{Cov}(X, Y) = \sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)]$$

Ισχύει ότι: $\text{Cov}(X, X) = \text{Var}(X)$

Όπως και στη περίπτωση της διασποράς η συνδιασπορά υπολογίζεται ευκολότερα από τον τύπο:

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

Ιδιότητες συνδιασποράς

(1) Αν οι τ.μ. X και Y είναι ανεξάρτητες τότε $\text{Cov}(X, Y) = 0$.

(2) Αν οι τ.μ. X και Y είναι από κοινού κατανεμημένες, τότε

$$\text{Var}(aX \pm bY) = a^2 \text{Var}(X) + b^2 \text{Var}(Y) \pm 2ab \text{Cov}(X, Y) \quad \forall a, b \in \mathbb{R}$$

Συντελεστής συσχέτισης

Έστω ότι έχουμε δύο τυχαίες μεταβλητές X και Y από κοινού κατανεμημένες. Ονομάζουμε *συντελεστή συσχέτισης* το λόγο:

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

Ιδιότητες:

- (1) $|\rho(X, Y)| \leq 1$
- (2) Εάν για δύο τ.μ. X και Y ισχύει $|\rho(X, Y)| = 1$, τότε η Y είναι γραμμική συνάρτηση της X .
- (3) Αν οι τ.μ. X και Y είναι ανεξάρτητες τότε $\rho = 0$.
- (4) Αν $\rho = 0$, τότε X και Y ασυσχέτιστες τ.μ.

Ροπογεννήτρια συνάρτηση

Έστω X τ.μ. και $f_X(x)$ η σ.π. ή σ.π.π. της X

Ροπογεννήτρια της τ.μ. X είναι η συνάρτηση:

$$M_X(t) = E[e^{tx}] \quad t \in \mathbb{R}$$

Σύμφωνα με το θεώρημα του αφηρημένου στατιστικού:

$$M_X(t) = \sum e^{tx} f_X(x) \quad \text{αν η } X \text{ είναι διακριτή, και}$$

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x) dx \quad \text{αν η } X \text{ είναι συνεχής.}$$

Ισχύει ότι:

$$\mu'_r = \left. \frac{d^r M_X(t)}{dt^r} \right|_{t=0}$$

(η σχέση αυτή δικαιολογεί το όνομα της συνάρτησης)

Ροπογεννήτρια συνάρτηση (συνέχεια)

Θεώρημα (μοναδικότητας)

Αν οι ροπογεννήτριες δύο τ.μ. X και Y ταυτίζονται για όλες τις τιμές του t σε διάστημα γύρω από το μηδέν, τότε οι κατανομές των X και Y είναι ίδιες (ισόνομες τ.μ.).

Ροπογεννήτρια ομοιόμορφης κατανομής

Η τ.μ X έχει σ.π. $f_X(x) = \frac{1}{\beta - \alpha} \quad \alpha < x < \beta$

Ροπογεννήτρια:

$$\begin{aligned} M_X(t) &= E[e^{tx}] = \int_{-\infty}^{+\infty} e^{tx} f_X(x) dx = \\ &= \frac{1}{\beta - \alpha} \int_{\alpha}^{\beta} e^{tx} dx = \frac{e^{tx}}{t(\beta - \alpha)} \Big|_{\alpha}^{\beta} = \frac{e^{t\beta} - e^{t\alpha}}{t(\beta - \alpha)} \quad t \neq 0 \end{aligned}$$

Ροπογεννήτρια κατανομής Poisson

Η τ.μ X έχει σ.π. $f(x) = P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$ όταν $x = 0, 1, 2, 3, \dots$ και $\lambda > 0$

Ροπογεννήτρια:

$$\begin{aligned} M_X(t) &= E[e^{tx}] = \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda} \lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} = \\ &= e^{-\lambda} e^{\lambda e^t} = e^{\lambda e^t - \lambda} = e^{\lambda(e^t - 1)} \end{aligned}$$

Ανισότητα Chebychev

Έστω X τ.μ. με $E(X) = \mu < \infty$, $Var(X) = \sigma^2 < \infty$

Τότε

$$\forall \varepsilon > 0 \text{ ισχύει } P[|X - \mu| \geq \varepsilon] \leq \frac{\sigma^2}{\varepsilon^2}$$

ή

$$\forall \varepsilon > 0 \text{ ισχύει } P[|X - \mu| < \varepsilon] > 1 - \frac{\sigma^2}{\varepsilon^2}$$

Όταν $\varepsilon = k\sigma$

$$P[|X - \mu| \geq k\sigma] \leq \frac{1}{k^2} \quad \text{ή} \quad P[|X - \mu| < k\sigma] > 1 - \frac{1}{k^2}$$

Άλλα περιγραφικά μέτρα κατανομής

-- παράμετροι θέσης --

Έστω X τ.μ. με $f_X(x)$ η σ.π. ή σ.π.π. της X και $F_X(x)$ η α.σ.κ.
Διάμεσος της X ονομάζεται κάθε αριθμός δ για τον οποίο ισχύει:

$$F(\delta) = P(X \leq \delta) \geq \frac{1}{2} \geq P(X < \delta) = F(\delta^-)$$

Αντίστοιχα, p -ποσοστιαίο σημείο είναι το σημείο x_p για το οποίο ισχύει:

$$F(x_p) = P(X \leq x_p) \geq p \geq P(X < x_p) = F(x_p^-) \quad 0 < p < 1$$

Ενδιαφέρον παρουσιάζουν τα τεταρτημόρια: $x_{0.25}, x_{0.50}, x_{0.75}$

Άλλα περιγραφικά μέτρα κατανομής (συνέχεια)

Επικρατούσα τιμή ή κορυφή ονομάζεται κάθε τοπικό μέγιστο της $f_X(\cdot)$ όταν η X είναι συνεχής.

Αντίστοιχα, όταν η X είναι διακριτή με δυνατές τιμές:

$x_1 < x_2 < \dots < x_k < x_{k+1} < \dots$, μια ενδιάμεση τιμή x_k λέγεται *κορυφή* όταν $f(x_k) > f(x_{k-1})$ και $f(x_k) > f(x_{k+1})$, ενώ μια ακραία τιμή είναι κορυφή όταν είναι μεγαλύτερη από τη γειτονική της.

(μονοκόρυφες, δικόρυφες ή πολυκόρυφες κατανομές)

Άλλα περιγραφικά μέτρα κατανομής (συνέχεια)

-- μέτρα μεταβλητότητας --

- Κεντρική έκταση μιας συνεχούς τ.μ. X : $x_{0.75} - x_{0.25}$
- Ημικεντρική έκταση μιας συνεχούς τ.μ. X : $(x_{0.75} - x_{0.25}) / 2$
- Μέση απόκλιση: $E[|X - \mu|]$
- Διάμεσος απόκλιση: $E[|X - \delta|]$
- Συντελεστής διακύμανσης: $c = \sigma / \mu$ (ποσοστό απόκλισης από την μέση τιμή)
- Συντελεστής λοξότητας: $\alpha_3 = \frac{\mu_3}{\sigma^3} = \frac{E[(X - \mu)^3]}{\sigma^3}$
- Συντελεστής λοξότητας: $\alpha_4 = \frac{\mu_4}{\sigma^4} = \frac{E[(X - \mu)^4]}{\sigma^4}$