

Τυχαίες Μεταβλητές (τ.μ.)

Τυχαία Μεταβλητή (τ.μ.): συνάρτηση $X(.)$ με πεδίο ορισμού τον δειγματικό χώρο Ω και πεδίο τιμών ένα σύνολο πραγματικών αριθμών.

$$X(.) : \Omega \rightarrow D \subseteq \mathbb{R}$$

- Διακριτές τ.μ.
- Συνεχείς τ.μ.

Η πιθανοτική συμπεριφορά μιας τ.μ., δηλ. οι πιθανότητες που αντιστοιχούν στις διάφορες τιμές ή σε διαστήματα τιμών της τ.μ., δίνουν την κατανομή πιθανότητας (ή απλώς κατανομή) της τ.μ. X

Διακριτές τ.μ.

Συνάρτηση Πιθανότητας (σ.π.) της X :

$$f(x) = P(X = x) = P(\omega \in \Omega : X(\omega) = x) \quad \forall x \in \mathbb{R}$$

με πεδίο ορισμού το σύνολο των πραγματικών αριθμών και πεδίο τιμών στο διάστημα $[0,1]$.

Ιδιότητες

1. $f(x) \geq 0 \quad \forall x \in \mathbb{R}$

2. $\sum_x f(x) = 1 \quad \forall x \in \mathbb{R}$

Διακριτές τ.μ. (συνέχεια)

Αθροιστική συνάρτηση κατανομής (α.σ.κ.) της X :

$$F(x) = P(X \leq x) \quad \forall x \in \mathbb{R}$$

με πεδίο ορισμού το σύνολο των πραγματικών αριθμών και πεδίο τιμών στο διάστημα $[0,1]$.

Ιδιότητες

1. Μη αρνητική και μη φθίνουσα

$$2. \lim_{x \rightarrow x_0^+} F(x) = F(x_0) \quad x_0 \in \mathbb{R}$$

$$3. \lim_{x \rightarrow \infty} F(x) = 1 \quad \text{και} \quad \lim_{x \rightarrow -\infty} F(x) = 0$$

$$4. F(x) = \sum_{k \leq x} f(k) \quad x \in \mathbb{R}$$

$$5. f(x) = F(x) - F(x^-) \quad \forall x \in \mathbb{R} \quad \text{όταν} \quad F(x^-) = \lim_{x \rightarrow x^-} F(x)$$

Παράδειγμα 1: Κατανομή Bernulli

Έστω ένα πείραμα με δύο πιθανά αποτελέσματα: επιτυχία (1) ή αποτυχία (0). Εάν η πιθανότητα επιτυχίας είναι p , η τ.μ. X που μετράει τον αριθμό των επιτυχιών ακολουθεί *κατανομή Bernulli*.

$$f(x) = \begin{cases} p & \text{όταν } x = 1 \\ q = 1 - p & \text{όταν } x = 0 \end{cases}$$

όπου $0 < p < 1$.

$$F(x) = \begin{cases} 0 & \text{όταν } x < 0 \\ q & \text{όταν } 0 \leq x < 1 \\ 1 & \text{όταν } x \geq 1 \end{cases}$$

Παράδειγμα 2: Διωνυμική Κατανομή

Έστω πείραμα με n ανεξάρτητες επαναλήψεις δοκιμών Bernoulli, όταν η πιθανότητα επιτυχίας είναι p ($0 < p < 1$).
Η τ.μ. X που μετράει τον αριθμό των επιτυχιών στις n επαναλήψεις ακολουθεί *διωνυμική κατανομή* $\mathcal{B}(n, p)$.

$$f(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & x = 0, 1, 2, \dots, n \\ 0 & \text{αλλού} \end{cases}$$

Η α.σ.κ. είναι:

$$F(x) = P(X \leq x) = \begin{cases} 0 & x < 0 \\ \sum_{k=0}^{\lfloor x \rfloor} \binom{n}{k} p^k (1-p)^{n-k} & 0 \leq x < n \\ 1 & x \geq n \end{cases}$$

Παράδειγμα 3: Γεωμετρική Κατανομή

Έστω πείραμα με ανεξάρτητες επαναλήψεις δοκιμών Bernulli, όταν η πιθανότητα επιτυχίας είναι p ($0 < p < 1$).
Η τ.μ. X που μετράει τον αριθμό των επαναλήψεων που απαιτούνται μέχρι την εμφάνιση της πρώτης επιτυχίας ακολουθεί γεωμετρική κατανομή $\mathcal{G}(p)$.

$$f(x) = \begin{cases} p(1-p)^{x-1} & x = 1, 2, \dots \\ 0 & \text{αλλού} \end{cases}$$

Η α.σ.κ. είναι:

$$F(x) = \begin{cases} 0 & x < 1 \\ 1 - (1-p)^{[x]} & x \geq 1 \end{cases}$$

Παράδειγμα 4:

Αρνητική διωνυμική κατανομή

Έστω πείραμα με ανεξάρτητες επαναλήψεις δοκιμών Bernoulli, όταν η πιθανότητα επιτυχίας είναι p ($0 < p < 1$). Η τ.μ. X που μετράει τον αριθμό των επαναλήψεων που απαιτούνται μέχρι την r επιτυχία ακολουθεί αρνητική διωνυμική κατανομή $\mathcal{NB}(r, p)$.

$$f(x) = \begin{cases} \binom{x-1}{r-1} p^r (1-p)^{x-r} & x = r, r+1, \dots \\ 0 & \text{αλλού} \end{cases}$$

Παράδειγμα 5: Κατανομή Poisson

Μια τ.μ. έχει την *κατανομή Poisson* $\mathcal{P}(\lambda)$ όταν η συνάρτηση πιθανότητας της X δίνεται από τον τύπο:

$$f(x) = P(X = x) = \begin{cases} \frac{e^{-\lambda} \lambda^x}{x!} & x = 0, 1, 2, 3, \dots \\ 0 & \text{αλλού} \end{cases} \quad \text{όπου } \lambda > 0$$

Η τ.μ. X μετράει την εμφάνιση τυχαίων περιστατικών σε καθορισμένο χρονικό διάστημα π.χ. εισερχόμενες κλήσεις ανά ώρα σε κάποιο τηλεφωνικό κέντρο.

Η α.σ.κ.

$$F(x) = \begin{cases} 0 & x < 0 \\ e^{-\lambda} \sum_{k=0}^{[x]} \frac{\lambda^k}{k!} & x \geq 0 \end{cases}$$

Συμπεράσματα για διακριτές τ.μ.

Έστω X διακριτή τ.μ. και η α.σ.κ $F(x)$. Τότε ισχύει:

$$P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$$

Επίσης μπορούμε να γράψουμε:

$$F(t) = \sum_{x=-\infty}^{\lfloor t \rfloor} f(x)$$

όπου $\lfloor t \rfloor$ παριστάνει το ακέραιο μέρος του t .

Συνεχείς Τ.μ.

Συνάρτηση πυκνότητας πιθανότητας (σ.π.π.) της X :
συνάρτηση με πεδίο ορισμού το σύνολο των
πραγματικών αριθμών και πεδίο τιμών στο διάστημα
 $[0,1]$, τέτοια ώστε:

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

Ιδιότητες

$$1. \quad f(x) \geq 0 \quad \forall x \in \mathbb{R}$$

$$2. \quad \int_{-\infty}^{\infty} f(x)dx = 1$$

Για μια συνεχή τ.μ. ισχύει: $P(X = a) = \int_a^a f(x)dx = 0$ και
 $P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b)$

Συνεχείς τ.μ. (συνέχεια)

Αθροιστική συνάρτηση κατανομής (α.σ.κ.) της X :

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(u) du \quad \forall x \in \mathbb{R}$$

με πεδίο ορισμού το σύνολο των πραγματικών αριθμών
και πεδίο τιμών στο διάστημα $[0,1]$.

Ιδιότητες

1. μη αρνητική και μη φθίνουσα
2. συνεχής συνάρτηση στο $\mathbb{R}$
3. $\lim_{x \rightarrow \infty} F(x) = 1$ και $\lim_{x \rightarrow -\infty} F(x) = 0$
4. $f(x) = \frac{d}{dx} F(x) \quad \forall x \in \mathbb{R}$

Ισχύει: $P(a < X \leq b) = F(b) - F(a)$ και $P(X > x) = 1 - F(x)$

Παράδειγμα 1:

Ομοιόμορφη Κατανομή (Uniform)

$X \sim \mathcal{U}(\alpha, \beta)$ με $\alpha < \beta$, όταν:

$$f_X(x) = \begin{cases} \frac{1}{\beta - \alpha} & \alpha < x < \beta \\ 0 & \text{αλλού} \end{cases}$$

Αθροιστική Συνάρτηση κατανομής της X :

$$F_X(x) = \begin{cases} 0 & x < \alpha \\ \frac{x - \alpha}{\beta - \alpha} & \alpha \leq x \leq \beta \\ 1 & x > \beta \end{cases}$$

Παράδειγμα 2:

Εκθετική Κατανομή (Exponential)

$X \sim \mathcal{E}(\lambda)$ όταν:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & \text{αλλού} \end{cases} \quad \text{όπου } \lambda \text{ θετική σταθερά}$$

Αθροιστική Συνάρτηση κατανομής της X : $F_X(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & \text{αλλού} \end{cases}$

Χρησιμοποιείται για να εκφράσει:

- το χρόνο μεταξύ αφίξεων διαφορετικών πελατών σε κατάστημα
- το χρόνο συνομιλίας σε μια τηλεφωνική επικοινωνία
- το χρόνος ζωής ηλεκτρονικών κομματιών σε μία συσκευή
-

Διδιάστατες τ.μ.

Έστω ότι σε ένα πείραμα ενδιαφερόμαστε για δύο τυχαίες μεταβλητές, X και Y . Οι τ.μ. μπορούν να είναι διακριτές ή συνεχείς.

(α) Διακριτές διδιάστατες τ.μ.

Κοινή συνάρτηση πιθανότητας (κ.σ.π.) των X και Y είναι μία συνάρτηση $f(x,y)$ με πεδίο ορισμού το $\mathbb{R}^2$ και πεδίο τιμών στο $[0,1]$, όταν

$$f(x,y) = P(X = x, Y = y) \quad \forall (x,y) \in \mathbb{R}^2$$

Ισχύει:

$$1. \quad f(x,y) \geq 0 \quad \forall (x,y) \in \mathbb{R}^2$$

$$2. \quad \sum_x \sum_y f(x,y) = 1$$

Διδιάστατες τ.μ. (συνέχεια)

Εάν ενδιαφερόμαστε για την πιθανοτική συμπεριφορά της X μόνο ή της Y μόνο, ορίζουμε τις *περιθώριες συναρτήσεις πιθανότητας*:

$$f_X(x) = \sum_y f(x, y) \quad x \in \mathbb{R} \quad (\text{περιθώρια σ.π. της } X)$$

$$f_Y(y) = \sum_x f(x, y) \quad y \in \mathbb{R} \quad (\text{περιθώρια σ.π. της } Y)$$

Ορίζουμε την *κοινή (αθροιστική) συνάρτηση κατανομής (κ.α.σ.κ.)* ως τη συνάρτηση με τιμή

$$F(x, y) = P(X \leq x, Y \leq y) = \sum_{k \leq x} \sum_{l \leq y} f(k, l)$$

Διδιάστατες τ.μ. (συνέχεια)

(β) Συνεχείς διδιάστατες τ.μ.

Κοινή συνάρτηση πυκνότητας πιθανότητας (κ.σ.π.π.) των X και Y είναι μία συνάρτηση $f(x,y)$ με πεδίο ορισμού το $\mathbb{R}^2$ και πεδίο τιμών στο $[0,1]$ που ικανοποιεί τις ιδιότητες:

$$1. \quad f(x,y) \geq 0 \quad \forall (x,y) \in \mathbb{R}^2$$

$$2. \quad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) dx dy = 1$$

και ισχύει:

$$P[(X,Y) \in E] = \iint_{(x,y) \in E} f(x,y) dx dy \quad \text{όπου} \quad E \subset \mathbb{R}^2$$

Π.χ.

$$P(a < X < b, c < Y < d) = \int_c^d \int_a^b f(x,y) dx dy$$

Διδιάστατες τ.μ. (συνέχεια)

Εάν ενδιαφερόμαστε για την πιθανοτική συμπεριφορά της X μόνο ή της Y μόνο, τότε ορίζουμε τις *περιθώριες συναρτήσεις πυκνότητας πιθανότητας*:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy \quad x \in \mathbb{R} \quad (\text{περιθώρια σ.π.π. της } X)$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx \quad y \in \mathbb{R} \quad (\text{περιθώρια σ.π.π. της } Y)$$

Ορίζουμε την *κοινή (αθροιστική) συνάρτηση κατανομής (κ.α.σ.κ.)* ως τη συνάρτηση με τιμή

$$F(x, y) = P(X \leq x, Y \leq y) = \int_{-\infty}^y \int_{-\infty}^x f(u, v) du dv$$

Διδιάστατες τ.μ. (συνέχεια)

Ιδιότητες κοινής αθροιστικής συνάρτησης κατανομής

1. μη αρνητική και μη φθίνουσα
2. συνεχής στο $\mathbb{R}^2$
3. $\lim_{\substack{x \rightarrow \infty \\ y \rightarrow \infty}} F(x, y) = 1$
4. $\lim_{\substack{x \rightarrow -\infty \\ y \rightarrow -\infty}} F(x, y) = \lim_{x \rightarrow -\infty} F(x, y) = \lim_{y \rightarrow -\infty} F(x, y) = 0$
5. $\lim_{x \rightarrow \infty} F(x, y) = F_Y(y)$ και $\lim_{y \rightarrow \infty} F(x, y) = F_X(x)$
6. $f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}$

Διδιάστατες τ.μ. (συνέχεια)

Έστω X και Y τυχαίες μεταβλητές με κ.α.σ.κ. $F(x, y)$.
Οι τ.μ. X και Y λέγονται *ανεξάρτητες* αν και μόνο αν:

$$F(x, y) = F_X(x) \cdot F_Y(y) \quad \forall (x, y) \in \mathbb{R}^2$$

όπου $F_X(x)$ και $F_Y(y)$ οι περιθώριες α.σ.κ. των X και Y .

Ισοδύναμα, για τις τ.μ. X και Y ισχύει ότι είναι
ανεξάρτητες αν και μόνο αν:

$$f(x, y) = f_X(x) \cdot f_Y(y) \quad \forall (x, y) \in \mathbb{R}^2$$

όπου $f(x, y)$, $f_X(x)$ και $f_Y(y)$ είναι σ.π.π. (για συνεχείς τ.μ.) και σ.π (για διακριτές τ.μ.)

Συναρτήσεις τυχαίων μεταβλητών

Έστω X τυχαία μεταβλητή με γνωστή κατανομή.

Εάν $Y = g(X)$ είναι συνάρτηση της X τότε και η Y είναι τυχαία μεταβλητή. Ζητάμε την πιθανοτική συμπεριφορά της Y . Όταν οι X και Y είναι διακριτές τ.μ. το πρόβλημα είναι απλό.

Στην περίπτωση που η X είναι συνεχής και η συνάρτηση $y = g(x)$ είναι (γνησίως) μονότονη και παραγωγίσιμη στην περιοχή των δυνατών τιμών της X , υπάρχει η αντίστροφη συνάρτηση $x = g^{-1}(y)$ και ισχύει:

$$f_Y(y) = f_X(g^{-1}(y)) \cdot \left| \frac{dg^{-1}(y)}{dy} \right| = f_X(x) \cdot \left| \frac{dx}{dy} \right| \quad \forall y$$

Αυτό είναι γνωστό σαν **Θεώρημα αλλαγής μεταβλητής**

Συναρτήσεις διδιάστατων τ.μ.

Έστω X και Y συνεχείς τ.μ. με κ.σ.π.π. $f_{X,Y}(x,y)$.

Εάν $Z = g_1(X,Y)$ και $W = g_2(X,Y)$ είναι συναρτήσεις των X και Y , ο μετασχηματισμός $(x,y) \rightarrow (z,w)$ είναι 1-1 και οι g_1, g_2 είναι παραγωγίσιμες ως προς x και ως προς y στην περιοχή για την οποία $f_{X,Y}(x,y) \neq 0$, τότε υπάρχει ο αντίστροφος μετασχηματισμός που δίνεται από τις συναρτήσεις $x = h_1(z,w)$ και $y = h_2(z,w)$ και η κ.σ.π.π. των Z και W είναι:

$$\begin{aligned} f_{Z,W}(z,w) &= f_{X,Y}(h_1(z,w), h_2(z,w)) \cdot \left| \frac{\partial(h_1(z,w), h_2(z,w))}{\partial(z,w)} \right| = \\ &= f_{X,Y}(x,y) \cdot \left| \frac{\partial(x,y)}{\partial(z,w)} \right| \quad \forall (z,w) \end{aligned}$$

(Θεώρημα αλλαγής μεταβλητών για διδιάστατες μεταβλητές)

Υπό συνθήκη κατανομές

Έστω X και Y δύο τ.μ. με $f_{X,Y}(x,y)$ η κ.σ.π. ή η κ.σ.π.π τους, αν είναι διακριτές ή συνεχείς, αντίστοιχα.

Η υπό συνθήκη κατανομή της Y όταν η X πάρει μια συγκεκριμένη τιμή x δίνεται από την υπό συνθήκη σ.π.

$$f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)} \text{ ή αλλιώς } P(Y=y | X=x) = \frac{P(X=x, Y=y)}{P(X=x)} \quad \forall y \in \mathbb{R}$$

όταν οι X και Y είναι διακριτές τ.μ. ή από την υπό συνθήκη σ.π.π.

$$f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)} \quad \forall y \in \mathbb{R}$$

όταν οι X και Y είναι συνεχείς τ.μ.