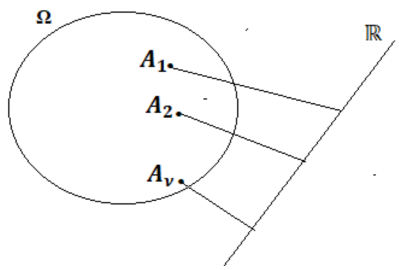
Τυχαίες μεταβλητές

Τυχαία μεταβλητή (τ.μ.) λέγεται κάθε συνάρτηση που απεικονίζει το σύνολο των δυνατών αποτελεσμάτων ενός πειράματος τύχης στο σύνολο των πραγματικών αριθμών.



Οι τυχαίες μεταβλητές συμβολίζονται με τα κεφαλαία γράμματα X, Y, Z,…., ενώ οι τιμές που παίρνουν με τα μικρά γράμματα $x_1, x_2, \dots, x_\nu$ ή $y_1, y_2, \dots, y_\nu$ ή $z_1, z_2, \dots, z_k$. Διακρίνονται σε **ποιοτικές** όταν εκφράζουν ποιοτικά χαρακτηριστικά ενός πληθυσμού (π.χ. το χρώμα των ανθέων ενός φυτού, το φύλο, η εθνικότητα, το επάγγελμα ενός ατόμου, κ.λ.π.) και **ποσοτικές** όταν μπορούν να μετρηθούν (π.χ. το ύψος ενός φυτού, ο αριθμός των σπόρων, το βάρος ενός ζώου, κ.λ.π.).

Επίσης μια τυχαία μεταβλητή μπορεί να είναι **διακριτή (ή απαριθμητή)** εάν το σύνολο των τιμών της είναι πεπερασμένο ή απείρως αριθμήσιμο ή **συνεχής** εάν παίρνει τιμές σε ένα υποσύνολο των πραγματικών αριθμών, δηλαδή σε ένα διάστημα (α, β) με $-\infty \leq \alpha < \beta \leq \infty$. Το πλήθος των βακτηριδίων στη δειγματοληπτική πλάκα, ο αριθμός των γεννήσεων ή των θανάτων που συμβαίνουν σε μια κτηνοτροφική μονάδα, ο αριθμός των ημερών βροχής σε έναν μήνα ή σε ένα έτος σε μια συγκεκριμένη περιοχή είναι διακριτές τ.μ., ενώ ο χρόνος λειτουργίας ενός λαμπτήρα, το ύψος, το βάρος, η θερμοκρασία είναι συνεχείς τ.μ.

Αν η τυχαία μεταβλητή X είναι **διακριτή**, τότε η συνάρτηση με πεδίο ορισμού τις τιμές που παίρνει η τ.μ. X και πεδίο τιμών τις πιθανότητες των τιμών αυτών λέγεται **συνάρτηση πιθανότητας** της τ.μ. X, συμβολίζεται με $p(x) = P(X = x)$ και έχει τις εξής ιδιότητες:

i) $\boldsymbol{p(x) \geq 0}$, **για κάθε** $\boldsymbol{x}$**,**

ii) $\boldsymbol{\sum_x p(x) = 1}$

Αν η τυχαία μεταβλητή X είναι **συνεχής**, τότε υπάρχει μια μη αρνητική συνάρτηση $f(x)$, τέτοια ώστε:

$$P(X \in A) = \int_A f(x)dx$$

για κάθε διάστημα A του πεδίου τιμών της X , τότε η $f(x)$ λέγεται **συνάρτηση πυκνότητας πιθανότητας** και έχει τις εξής ιδιότητες:

i) $\boldsymbol{f(x) \geq 0}$, **για κάθε** $\boldsymbol{x \in \mathcal{R}}$

ii) $\boldsymbol{\int_{-\infty}^{\infty} f(x)dx = 1}$

Εάν X μία τυχαία μεταβλητή, τότε η **μέση ή αναμενόμενη τιμή** της τ.μ. X συμβολίζεται με $\boldsymbol{E(X)}$ ή $\boldsymbol{EX}$ ή $\boldsymbol{\mu}$ και ορίζεται από τη σχέση:

$$E(X) = \begin{cases} \sum_x x \cdot p(x) & \text{εάν η τ.μ.} X \text{ είναι διακριτή} \\ \int_{-\infty}^{\infty} x \cdot f(x)dx & \text{εάν η τ.μ. } X \text{ είναι συνεχής} \end{cases}$$

Γενικότερα **μέση ή αναμενόμενη τιμή** της $g(X)$ ορίζεται η:

$$Eg(X) = \begin{cases} \sum_x g(x) \cdot p(x) & \text{εάν η τ.μ.} X \text{ είναι διακριτή} \\ \int_{-\infty}^{\infty} g(x) \cdot f(x)dx & \text{εάν η τ.μ. } X \text{ είναι συνεχής} \end{cases}$$

Εάν X μία τυχαία μεταβλητή, τότε η **διασπορά ή διακύμανση** της τ.μ. X συμβολίζεται με $\boldsymbol{Var(X)}$ ή $\boldsymbol{V(X)}$ ή $\boldsymbol{\sigma^2}$ και ορίζεται από τη σχέση:

$$V(X) = E(X-\mu)^2 = \begin{cases} \sum_x (X-\mu)^2 \cdot p(x) & \text{εάν η τ.μ.} X \text{ είναι διακριτή} \\ \int_{-\infty}^{\infty} (X-\mu)^2 \cdot f(x)dx & \text{εάν η τ.μ. } X \text{ είναι συνεχής} \end{cases}$$

Ιδιότητα (για τον υπολογισμό) της διασποράς:

$$\boldsymbol{V(X) = E(X^2) - [E(X)]^2}$$

Η **τυπική απόκλιση** της τ.μ. X συμβολίζεται με $\sigma(X)$ και ορίζεται από τη σχέση:

$$\sigma(X) = \sqrt{V(X)}$$

Η **μέση ή αναμενόμενη τιμή** μιας τ.μ. *X* είναι κατά κάποιο τρόπο το κέντρο της πιθανότητας της *X* (αντίστοιχο του κέντρου βάρους) και δίνει το σημείο εκείνο γύρω από το οποίο παίρνει τιμές η τ.μ. *X*. Υπολογίζεται εύκολα, είναι εύχρηστη, αλλά έχει το μειονέκτημα να επηρεάζεται από ακραίες τιμές. Η **διασπορά ή διακύμανση** είναι ένας δείκτης που μας λέει πόσο συγκεντρωμένες είναι οι τιμές της τ.μ. *X* γύρω από τη μέση τιμή. Αν η διασπορά είναι μικρή, τότε οι τιμές της *X* κυμαίνονται γύρω τη μέση τιμή, ενώ αν η διασπορά είναι μεγάλη τότε υπάρχουν μεγάλες αποκλίσεις από τη μέση τιμή.

<u>Παράδειγμα</u>: Σε ένα δοχείο υπάρχουν 4 άσπρες και 3 κόκκινες σφαίρες. Παίρνουμε χωρίς επανάθεση μία-μία τρεις σφαίρες. i) Να βρεθεί η συνάρτηση πιθανότητας της τ.μ. *X*, που ορίζεται ως εξής: $X = \{$ ο αριθμός των κόκκινων σφαιρών που επιλέγονται$\}$ ii) Να βρεθεί η μέση τιμή και η τυπική απόκλιση της τ.μ. *X*.

i) Η τ.μ. *X* παίρνει τις τιμές $X = 0, 1, 2, 3$ με τις αντίστοιχες πιθανότητες:

$$P(X=0) = \frac{\text{ευνοικές περιπτώσεις}}{\text{δυνατές περιπτώσεις}} = \frac{\binom{4}{3}\binom{3}{0}}{\binom{7}{3}} = \frac{\frac{4!}{3!\,1!}\cdot\frac{3!}{0!\,3!}}{\frac{7!}{3!\,4!}} = \frac{4}{35}$$

$$P(X=1) = \frac{\binom{4}{2}\binom{3}{1}}{\binom{7}{3}} = \frac{\frac{4!}{2!\,2!}\cdot\frac{3!}{1!\,2!}}{\frac{7!}{3!\,4!}} = \frac{6\cdot 3}{35} = \frac{18}{35}$$

$$P(X=2) = \frac{\binom{4}{1}\binom{3}{2}}{\binom{7}{3}} = \frac{\frac{4!}{1!\,3!}\cdot\frac{3!}{2!\,1!}}{\frac{7!}{3!\,4!}} = \frac{4\cdot 3}{35} = \frac{12}{35}$$

$$P(X=3) = \frac{\binom{4}{0}\binom{3}{3}}{\binom{7}{3}} = \frac{\frac{4!}{0!\,4!}\cdot\frac{3!}{3!\,0!}}{\frac{7!}{3!\,4!}} = \frac{1\cdot 1}{35} = \frac{1}{35}$$

Επομένως η συνάρτηση πιθανότητας της διακριτής τ.μ. *X* είναι:

x	0	1	2	3
$p(x)$	$4/35$	$18/35$	$12/35$	$1/35$

ii) Μέση τιμή της τ.μ. X:

$$E(X) = \sum_x x \cdot p(x) = 0 \cdot\ 4/35 + 1 \cdot 18/35 + 2 \cdot 12/35 + 3 \cdot 1/35 = 45/35 = 1.29$$

Για τον υπολογισμό της τυπικής απόκλισης θα πρέπει πρώτα να υπολογίσουμε τη διασπορά της X. Θα την υπολογίσουμε χρησιμοποιώντας την ιδιότητα:

$V(X) = E(X^2) - [E(X)]^2$ όμως:

$$E(X^2) = \sum_x x^2 \cdot p(x) = 0^2 \cdot\ 4/35 + 1^2 \cdot 18/35 + 2^2 \cdot 12/35 + 3^2 \cdot 1/35 = 75/35$$

Επομένως:

$$V(X) = E(X^2) - [E(X)]^2 = \frac{75}{35} - \left(\frac{45}{35}\right)^2 = 0.49$$

$\sigma(X) = \sqrt{V(X)} = \sqrt{0.49} = 0.7$ η τυπική απόκλιση της X.

Λύσεις ασκήσεων από το **φυλλάδιο 5 – Ασκήσεις σε Τυχαίες Μεταβλητές και Κατανομές**

1. Μια διακριτή τυχαία μεταβλητή X έχει συνάρτηση πιθανότητας

x	0	1	2	3	4
$p(x)$	$1/16$	$4/16$	$6/16$	c	$1/16$

Να βρεθούν η σταθερά c, η πιθανότητα $P(0 \leq X < 3)$, η δεσμευμένη πιθανότητα $P(X \geq 3/1 < X < 4)$, η μέση τιμή $E(X)$ και η διασπορά $V(X)$ της X.

Σύμφωνα με την ιδιότητα της συνάρτησης πιθανότητας έχουμε:

$$\sum_x p(x) = 1 \Rightarrow p(0) + p(1) + \cdots + p(4) = 1 \Rightarrow 1/16 + 4/16 + 6/16 + c + 1/16 = 1$$

$$\Rightarrow 12/16 + c = 1 \Rightarrow \boldsymbol{c = 4/16} \text{ και } \boldsymbol{p(3) = 4/16}$$

$$P(0 \leq X < 3) = p(0) + p(1) + p(2) = 1/16 + 4/16 + 6/16 = 11/16$$

Για τον υπολογισμό της δεσμευμένης πιθανότητας που ακολουθεί υπενθυμίζουμε ότι:

$$P(A/B) = \frac{P(AB)}{P(B)}$$

Επομένως:

$$P(X \geq 3/1 < X < 4) = \frac{P((X \geq 3) \cap (1 < X < 4))}{P(1 < X < 4)} = \frac{p(3)}{p(2) + p(3)} = \frac{4/16}{6/16 + 4/16} = \frac{2}{5}$$

καθώς: $X \geq 3 = \{3,\ 4\}$, $1 < X < 4 = \{2,\ 3\}$, επομένως $(X \geq 3) \cap (1 < X < 4) = \{3\}$

$$E(X) = \sum_x x \cdot p(x) = 0 \cdot p(0) + 1 \cdot p(1) + \cdots + 4 \cdot p(4) = 0 \cdot 1/16 + 1 \cdot 4/16 + \cdots + 4 \cdot 1/16$$
$$= 2$$

$V(X) = E(X - \mu)^2 = \sum_x (x - \mu)^2 \cdot p(x)$ ή $V(X) = E(X^2) - [E(X)^2]$

Αλλά

$$E(X^2) = \sum_x x^2 \cdot p(x) = 0^2 \cdot p(0) + 1^2 \cdot p(1) + \cdots + 4^2 \cdot p(4) = 5$$

Τότε $V(X) = E(X^2) - [E(X)^2] = 5 - 2^2 = 1$

2. Μια διακριτή τυχαία μεταβλητή X έχει συνάρτηση πιθανότητας

x	10	20	30	40	50
$p(x)$	$c/8$	$c/2$	$3c/4$	$c/2$	$c/8$

Αφού υπολογιστεί η σταθερά c, να βρεθούν οι πιθανότητες $P(X = 40)$,
$P(X \geq 30/X < 40)$, καθώς επίσης και η μέση τιμή $E(X)$ και η διασπορά $V(X)$ της X.

Σύμφωνα με την ιδιότητα της συνάρτησης πιθανότητας έχουμε:

$$\sum_x p(x) = 1 \Rightarrow p(10) + p(20) + \cdots + p(50) = 1 \Rightarrow c/8 + c/2 + 3c/4 + c/2 + c/8 = 1$$
$$\Rightarrow 2c = 1 \Rightarrow c = 1/2$$

Τότε η συνάρτηση πιθανότητας είναι:

x	10	20	30	40	50
$p(x)$	$1/16$	$4/16$	$6/16$	$4/16$	$1/16$

Επομένως $P(X = 40) = 4/16$

$$P(X \geq 30/X < 40) = \frac{P(X \geq 30 \cap X < 40)}{P(X < 40)} = \frac{P(X = 30)}{P(X = 10) + P(X = 20) + P(X = 30)}$$

$$= \frac{6/16}{1/16 + 4/16 + 6/16} = \frac{6/16}{11/16} = \frac{6}{11}$$

$\{X \geq 30 \cap X < 40\} = \{X = 30\}$ καθώς $X \geq 30 = \{30, 40, 50\}$, $X < 40 = \{10, 20, 30\}$

$$E(X) = \sum_x x \cdot p(x) = 10 \cdot p(10) + 20 \cdot p(20) + \cdots + 50 \cdot p(50) =$$

$$= 10 \cdot 1/16 + 20 \cdot 4/16 + 30 \cdot 6/16 + 40 \cdot 4/16 + 50 \cdot 1/16 = 30$$

$V(X) = E(X - \mu)^2 = \sum_x (x - \mu)^2 \cdot p(x)$ ή $V(X) = E(X^2) - [E(X)^2]$

Αλλά

$$E(X^2) = \sum_x x^2 \cdot p(x) = 10^2 \cdot p(10) + 20^2 \cdot p(20) + \cdots + 50^2 \cdot p(50) =$$

$$= 100 \cdot 1/16 + 400 \cdot 4/16 + \cdots + 2500\, 1/16 = 1000$$

Τότε $V(X) = E(X^2) - [E(X)^2] = 1000 - 30^2 = 100$ η διασπορά της X.

και $\sigma(X) = \sqrt{V(X)} = \sqrt{100}$ =10 η τυπική απόκλιση της X.

3. Ο αριθμός των προσκλήσεων ενός κτηνιάτρου από μια κτηνοτροφική μονάδα ανά μήνα είναι διακριτή τυχαία μεταβλητή X με συνάρτηση πιθανότητας:

x	0	1	2	3	4	5
$p(x)$	0.1	0.4	0.2	c	0.1	0.05

α) Βρείτε την πιθανότητα ο κτηνίατρος σε ένα μήνα να προσκληθεί από την κτηνοτροφική μονάδα 3 φορές ακριβώς.
β) Βρείτε την πιθανότητα ο κτηνίατρος σε ένα μήνα να προσκληθεί από την κτηνοτροφική μονάδα το πολύ 3 φορές, όταν είναι γνωστό ότι προσκλήθηκε τουλάχιστον μία φορά.
γ) Βρείτε τη μέση τιμή και τη διασπορά της X .

α)

$$\sum_x p(x) = 1 \Rightarrow p(0) + p(1) + \cdots + p(5) = 1 \Rightarrow 0.1 + 0.4 + 0.2 + c\ + 0.1 + 0.05 = 1$$

$$\Rightarrow 0.85 + c = 1\ \Rightarrow \boldsymbol{c = 0.15}\ \text{και}\ \ \boldsymbol{p(3) = 0.15}$$

β)

$$P(X \leq 3/X \geq 1) = \frac{P(X \leq 3 \cap X \geq 1)}{P(X \geq 1)} = \frac{P(X = 1) + P(X = 2) + P(X = 3)}{1 - P(X = 0)}$$

$$= \frac{0.4 + 0.2 + 0.15}{1 - 0.1} = \frac{0.75}{0.9} = 0.833$$

καθώς $X \leq 3 = \{0, 1, 2, 3\}$, $X \geq 1 = \{1, 2, 3, 4, 5\}$ επομένως $(X \leq 3 \cap X \geq 1) = \{1, 2, 3\}$

$$E(X) = \sum_x x \cdot p(x) = 0 \cdot p(0) + 1 \cdot p(1) + \cdots + 5 \cdot p(5) =$$

$$= 0 \cdot 0.1 + 1 \cdot 0.4 + 2 \cdot 0.2 + 3 \cdot 0.15 + 4 \cdot 0.1 + 5 \cdot 0.05 = 1.9$$

$V(X) = E(X - \mu)^2 = \sum_x (x - \mu)^2 \cdot p(x)$ ή $V(X) = E(X^2) - [E(X)^2]$

Αλλά

$$E(X^2) = \sum_x x^2 \cdot p(x) = 0^2 \cdot p(0) + 1^2 \cdot p(1) + \cdots + 5^2 \cdot p(5) = 0^2 \cdot 0.1 + 1^2 \cdot 0.4$$

$$+ 2^2 \cdot 0.2 + 3^2 \cdot 0.15 + 4^2 \cdot 0.1 + 5^2 \cdot 0.05 = 5.4$$

Τότε $V(X) = E(X^2) - [E(X)^2] = 5.4 - 1.9^2 = 1.79$ η διασπορά της X.