Ενότητα 9

Βασικές Συνεχείς Κατανομές

1. Ομοιόμορφη Κατανομή

Ορισμός

Έστω X μια συνεχής τυχαία μεταβλητή με πυκνότητα την

$$f(x)=\begin{cases}\dfrac{1}{\beta-a}, & a\le x\le \beta\\ 0, & x<a \text{ ή } x>\beta.\end{cases}$$

Η κατανομή

της τ.μ. X συμβολίζεται με $U(a,\beta)$ και καλείται ομοιόμορφη ή ορθογώνια στο διάστημα $[a,\beta]$. Τα σημεία a και β είναι παράμετροι της κατανομής.

Θεώρημα

Έστω ότι η τυχαία μεταβλητή X έχει την ομοιόμορφη κατανομή $U(a, \beta)$. Τότε η μέση τιμή και η διασπορά αυτής δίνονται από τις

$$\mu = E(X) = \frac{a+\beta}{2}, \quad \sigma^2 = V(X) = \frac{(\beta - a)^2}{12}.$$

Απόδειξη

Η μέση τιμή της τ.μ. X, σύμφωνα με τον ορισμό, είναι

$$\mu = E(X) = \frac{1}{\beta - a}\int_a^\beta x dx = \left[\frac{x^2}{2(\beta - a)}\right]_a^\beta = \frac{\beta^2 - a^2}{2(\beta - a)}$$

και επειδή $(\beta^2 - a^2) = (\beta - a)(\beta + a)$,

$$\mu = E(X) = \frac{a+\beta}{2}.$$

Επίσης είναι

$$E(X^2) = \frac{1}{\beta - a}\int_a^\beta x^2 dx = \left[\frac{x^3}{3(\beta - a)}\right]_a^\beta = \frac{\beta^3 - a^3}{3(\beta - a)}$$

και επειδή $\beta^3 - a^3 = (\beta - a)(\beta^2 + a\beta + a^2)$,

$$E(X^2) = \frac{a^2 + a\beta + \beta^2}{3}.$$

Η διασπορά της τ.μ. X είναι τότε

$$\begin{aligned}\sigma^2 = V(X) = E(X^2) - [E(X)]^2 &= \frac{a^2 + a\beta + \beta^2}{3} - \frac{a^2 + 2a\beta + \beta^2}{4}\\ &= \frac{a^2 - 2a\beta + \beta^2}{12} = \frac{(\beta - a)^2}{12}.\end{aligned}$$

Παράδειγμα

Έστω ότι ο συρμός φθάνει σε συγκεκριμένο σταθμό του μετρό κάθε 10 λεπτά, αρχίζοντας τα δρομολόγιά του στις 5 π.μ. Αν ένας επιβάτης φθάνει στο σταθμό σε χρόνο ο οποίος κατανέμεται ομοιόμορφα στο διάστημα 7:20 ως 7:40, να υπολογισθούν οι πιθανότητες να περιμένει το συρμό (α) το πολύ 4 λεπτά και (β) τουλάχιστον 7 λεπτά.

Λύση

Έστω X ο χρόνος άφιξης του επιβάτη στο σταθμό, μετρούμενος σε λεπτά με αρχή τη χρονική στιγμή 7:20. Τότε η τ.μ. X έχει την ομοιόμορφη κατανομή στο διάστημα $[0, 20]$ και έτσι

$$F(x) = \begin{cases} 0, & x < 0 \\ \dfrac{x}{20}, & 0 \leq x < 20 \\ 1, & x \geq 20. \end{cases}$$

(α) Το ενδεχόμενο A ο επιβάτης να περιμένει το πολύ 4 λεπτά είναι ισοδύναμο με το ενδεχόμενο να φθάσει στο σταθμό στο διάστημα 7:26 ως 7:30 ή στο διάστημα 7:36 ως 7:40. Επομένως

$$P(A)=P(6<X\leq 10)+P(16<X\leq 20)=\{F(10)-F(6)\}+\{F(20)-F(16)\}=\frac{2}{5}.$$

(β) Το ενδεχόμενο B ο επιβάτης να περιμένει τουλάχιστο 7 λεπτά είναι ισοδύναμο με το ενδεχόμενο να φθάσει στο σταθμό στο διάστημα 7:20 ως 7:23 ή 7:30 ως 7:33. Επομένως

$$P(B)=P(0<X\leq 3)+P(10<X\leq 13)=\{F(3)-F(0)\}+\{F(13)-F(10)\}=\frac{3}{10}.$$

Παράδειγμα

Έστω ότι ο χρόνος X σε ώρες που απαιτείται για την επισκευή μιας ηλεκτρονικής συσκευής είναι μια συνεχής τυχαία μεταβλητή με ομοιόμορφη κατανομή στο διάστημα $[1, 3]$. Να υπολογισθούν

(α) ο μέσος χρόνος επισκευής της συσκευής και

(β) το αναμενόμενο κόστος επισκευής της συσκευής, αν το κόστος επισκευής Y σε ευρώ δίνεται από την $Y = 3X^2 + 5$.

Λύση

(α) Η συνάρτηση πυκνότητας της τ.μ. X είναι η

$$f(x) = \frac{1}{2}, \quad 1 \leq x \leq 3$$

και έτσι ο μέσος χρόνος επισκευής της συσκευής είναι

$$E(X) = \frac{1}{2}\int_1^3 x dx = \left[\frac{x^2}{4}\right]_1^3 = 2.$$

(β) Επίσης

$$E(X^2) = \frac{1}{2}\int_1^3 x^2 dx = \left[\frac{x^3}{6}\right]_1^3 = \frac{13}{3}$$

και έτσι το αναμενόμενο κόστος επισκευής της συσκευής είναι

$$E(Y) = E(3X^2 + 5) = 3E(X^2) + 5 = 13 + 5 = 18.$$

2. *Κανονική Κατανομή*

Η σημαντικότερη κατανομή πιθανότητας της Θεωρίας Πιθανοτήτων και της Στατιστικής, με μεγάλο πεδίο εφαρμογών, είναι η **κανονική κατανομή**.

Η κατανομή αυτή χρησιμοποιήθηκε αρχικά από τους De Moivre και Laplace ως προσέγγιση της διωνυμικής κατανομής για μεγάλο αριθμό δοκιμών ν.

Ο Gauss χρησιμοποίησε την κανονική κατανομή ως προσεγγιστική της κατανομής των τυχαίων σφαλμάτων διαφόρων μετρήσεων. Για το λόγο αυτό η κατανομή αυτή αναφέρεται συχνά ως **Γκαουσιανή** κατανομή.

Η ονομασία **κανονική** δόθηκε πιο πρόσφατα από τον Karl Pearson.

Ορισμός

Έστω X μια συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty, \qquad (1)$$

όπου $-\infty < \mu < \infty, 0 < \sigma < \infty$ είναι παράμετροι. Η κατανομή της τ.μ. X συμβολίζεται $N(\mu, \sigma^2)$ και καλείται κανονική κατανομή με παραμέτρους μ και σ^2.

Θεώρημα

Έστω ότι η τυχαία μεταβλητή X ακολουθεί την κανονική κατανομή με συνάρτηση πυκνότητας (1). Τότε η μέση τιμή και η διασπορά αυτής δίνονται από τις

$$E(X) = \mu, \quad V(X) = \sigma^2.$$

Θεώρημα

Έστω ότι η τυχαία μεταβλητή X ακολουθεί τη διωνυμική κατανομή με συνάρτηση πιθανότητας την (4). Τότε η μέση τιμή και η διασπορά της αυτής δίνονται από τις

$$\mu = E(X) = \upsilon p, \quad \sigma^2 = V(X) = \upsilon pq.$$

Σημείωση

Το διάγραμμα της συνάρτησης πυκνότητας της κανονικής κατανομής είναι κωδωνοειδές.
Η μέση τιμή μ προσδιορίζει τη θέση της καμπύλης ως προς τον άξονα των x και η τυπική απόκλιση σ το οξύ ή πεπλατισμένο του σχήματος.

Η ειδική περίπτωση $\mu = 0$, $\sigma = 1$ παρουσιάζει ιδιαίτερο ενδιαφέρον. Σχετικά, η τυποποιημένη τυχαία μεταβλητή $Z = (X - \mu)/\sigma$ έχει συνάρτηση πυκνότητας

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < x < \infty \tag{2}$$

και συνάρτηση κατανομής

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt, \quad -\infty < z < \infty. \tag{3}$$

Η κατανομή $N(0, 1)$ καλείται *τυποποιημένη κανονική κατανομή*.

Η πινακοποίηση της συνάρτησης κατανομής της τυποποιημένης κανονικής κατανομής (3) διευκολύνει τη χρήση της. Τον σχετικό πίνακα συνεπικουρεί και η ακόλουθη ιδιότητά της.

Θεώρημα

Η συνάρτηση κατανομής της τυποποιημένης κανονικής κατανομής ικανοποιεί τη σχέση

$$\Phi(-z) = 1 - \Phi(z), \quad -\infty < z < \infty. \tag{4}$$

Ο υπολογισμός πιθανοτήτων που αφορούν μία κανονική τυχαία μεταβλητή X με μέση τιμή $E(X) = \mu$ και διασπορά $V(X) = \sigma^2$ ανάγεται στον υπολογισμό πιθανοτήτων που αφορούν την τυποποιημένη τυχαία μεταβλητή $Z = (X - \mu)/\sigma$ και γίνεται με τη χρήση των πινάκων της $\Phi(z)$, σύμφωνα με το ακόλουθο πόρισμα.

Πόρισμα

Αν η τυχαία μεταβλητή X ακολουθεί την κανονική κατανομή $N(\mu, \sigma^2)$, τότε

$$P(a \le X \le \beta) = \Phi\left(\frac{\beta-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right), \quad a \le \beta \qquad (5)$$

και ειδικότερα

$$P(X \le \beta) = \Phi\left(\frac{\beta-\mu}{\sigma}\right), \quad P(X \ge a) = 1 - \Phi\left(\frac{a-\mu}{\sigma}\right). \qquad (6)$$

Παράδειγμα

Έστω ότι ο χρόνος εμφάνισης X ενός φωτογραφικού φιλμ ακολουθεί κανονική κατανομή με μέση τιμή $\mu = 30$ min και τυπική απόκλιση $\sigma = 1{,}2$ min.

Να υπολογισθούν η πιθανότητα ο χρόνος εμφάνισης ενός φιλμ να μη υπερβεί τα 28 min

Λύση

Η πιθανότητα ο χρόνος εμφάνισης να μην υπερβεί τα 28 min δίνεται από την

$$P(X \leq 28) = P\left(\frac{X-30}{1{,}2} \leq \frac{28-30}{1{,}2}\right)$$
$$= P(Z \leq -1{,}67) = 1 - \Phi(1{,}67) = 0{,}0475 \cong 5\%.$$

Παράδειγμα

Το ύψος των ανδρών ενός πληθυσμού ακολουθεί την κανονική κατανομή με μέσο $\mu = 175\,\text{cm}$ και τυπική απόκλιση $\sigma = 5\,\text{cm}$. Να υπολογισθούν τα ποσοστά των ανδρών με ύψος

(α) μικρότερο από 175 cm και

(β) μεταξύ 170 cm και 180 cm.

Λύση

(α) Έστω X το ύψος οποιουδήποτε άνδρα του πληθυσμού. Τότε

$$P(X < 175) = P\left(\frac{X-\mu}{\sigma} < \frac{175-175}{5}\right) = P(Z < 0) = \Phi(0) = \frac{1}{2}$$

(β)

$$P(170 < X < 180) = P\left(\frac{170-175}{5} < \frac{X-\mu}{\sigma} < \frac{180-175}{5}\right) = P(-1 < Z < 1)$$
$$= \Phi(1) - \Phi(-1) = 2\Phi(1) - 1 = 2(0{,}8413) - 1 = 0{,}6826.$$