Συνεχείς τυχαίες μεταβλητές

Πιθανότητες και Στατιστική

Τμήμα Πληροφορικής, Πανεπιστήμιο Πειραιώς

2023-2024

Συνεχείς τυχαίες μεταβλητές

Μια ΤΜ X ονομάζεται **συνεχής** όταν το σύνολο τιμών S αυτής είναι διάστημα ή ένωση διαστημάτων.
Η μη αρνητική ολοκληρώσιμη συνάρτηση f/S για την οποία ισχύει

$$\int_S f(x)dx = 1 \quad \text{και} \quad P(X \in A) = \int_A f(x)dx, \text{ για κάθε διάστημα } A \subseteq S$$

ονομάζεται **συνάρτηση πυκνότητας πιθανότητας** (PDF) της ΤΜ X.
Συνήθως επεκτείνουμε την f σε όλο το $\mathbb{R}$, θέτοντας $f(x) = 0$ για κάθε $x \in \mathbb{R} \setminus S$. Τότε, η αθροιστική συνάρτηση κατανομής πιθανότητας (CDF) F ισούται με

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt, \qquad \text{για κάθε } x \in \mathbb{R},$$

οπότε, σύμφωνα με το πρώτο Θεμελιώδες Θεώρημα του Ολοκληρωτικού Λογισμού, η F είναι συνεχής και επιπλέον είναι παραγωγίσιμη σε κάθε x όπου η f είναι συνεχής, με

$$f(x) = F'(x), \qquad \text{για κάθε } x \text{ όπου } f \text{ συνεχής.}$$

Συνεχείς τυχαίες μεταβλητές

Βασικές ιδιότητες: Για κάθε $a, b \in \mathbb{R}$, με $a \leq b$, είναι

- $P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = \int_a^b f(x)dx.$
- Για κάθε $a \in \mathbb{R}$, η πιθανότητα η X να ισούται ακριβώς με a είναι 0, δηλαδή
$$P(X = a) = \int_a^a f(x)dx = 0$$
- $P(X > a) = P(X \geq a) = P(a \leq X < +\infty) = \int_a^{\infty} f(x)dx$
- $P(a \leq X \leq b) = F(b) - F(a)$

Μέση τιμή και διακύμανση

Η μέση τιμή μ και η διακύμανση σ^2 της συνεχούς ΤΜ X δίδονται από τους τύπους

$$\mu = E(X) = \int_S xf(x)dx = \int_{-\infty}^{\infty} xf(x)dx$$

$$\sigma^2 = V(X) = E((X - E(X))^2) = E((X - \mu)^2) = \int_S (x - \mu)^2 f(x)dx$$

Γενικότερα, για οποιαδήποτε συνάρτηση $g : S \to \mathbb{R}$, είναι

$$E(g(X)) = \int_S g(x)f(x)dx = \int_{-\infty}^{\infty} g(x)f(x)dx$$

Εύκολα αποδεικνύονται οι ακόλουθες ιδιότητες, όπως και στην περίπτωση διακριτής ΤΜ:

1. $E(aX + bY) = aE(X) + bE(Y)$
2. $V(aX + b) = a^2 V(X)$.
3. $V(X) = E(X^2) - (E(X))^2$.

Ανάλογα αποδεικνύεται και η ακόλουθη πρόταση:

Πρόταση

Αν οι συνεχείς τ. μ. $X_1, X_2, \ldots, X_n$ είναι ανεξάρτητες και $g_1, g_2, \ldots, g_n$ οποιεσδήποτε συναρτήσεις, όπου $g_i : S_{X_i} \to \mathbb{R}$, τότε:

1. Οι $g_1(X_1), g_2(X_2), \ldots, g_n(X_n)$ είναι ανεξάρτητες
2. $E(g_1(X_1)g_2(X_2)\cdots g_n(X_n)) = E(g_1(X_1))E(g_2(X_2))\cdots E(g_n(X_n))$
3. $V(g_1(X_1) + g_2(X_2) + \cdots + g_n(X_n)) =$ $V(g_1(X_1)) + V(g_2(X_2)) + \cdots + V(g_n(X_n))$

Παραδείγματα

Παράδειγμα

Να βρεθεί η σταθερά c ώστε η συνάρτηση $f(x) = \begin{cases} cx^2, & x \in [0,1] \\ 0, & x \notin [0,1] \end{cases}$ να είναι PDF κάποιας ΤΜ X. Στη συνέχεια, να βρεθεί ο τύπος της CDF, η πιθανότητα $P(X \leq 2/3)$, η μέση τιμή και η διακύμανση της X.

Λύση

Πρέπει

$$\int_{-\infty}^{+\infty} f(x)dx = 1 \Leftrightarrow c\int_0^1 x^2 dx = 1 \Leftrightarrow c\left[\frac{x^3}{3}\right]_0^1 = 1 \Leftrightarrow c\frac{1}{3} = 1 \Leftrightarrow c = 3.$$

Άρα,

$$f(x) = \begin{cases} 3x^2, & x \in [0,1] \\ 0, & x \notin [0,1] \end{cases}$$

Παραδείγματα

Λύση (συνέχεια)

Επομένως, $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} 0, & x \in (-\infty, 0) \\ x^3 & x \in [0, 1] \\ 1 & x \in (1, +\infty) \end{cases}$

Άρα, $P(X \leq 2/3) = F(2/3) = (2/3)^3 = 8/27$.

Επίσης, $E(X) = \int_{-\infty}^{+\infty} xf(x)dx = \int_0^1 x3x^2 dx = \left[\frac{3x^4}{4}\right]_0^1 = \frac{3}{4}$.

Τέλος,

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)dx = \int_0^1 3x^4 dx = \left[\frac{3x^5}{5}\right]_0^1 = \frac{3}{5},$$

οπότε

$$V(X) = E(X^2) - E(X)^2 = \frac{3}{5} - \frac{9}{16} = \frac{3}{80}.$$

Παράδειγμα

Έστω ότι ρίχνουμε με κλειστά μάτια ένα μαγνητικό βελάκι σε ένα στόχο με σχήμα κυκλικού δίσκου ακτίνα 10 εκατοστών με κέντρο την αρχή των αξόνων. Έστω D η ΤΜ που ισούται με την απόσταση του σημείου που στοχεύσαμε με το βελάκι από την αρχή των αξόνων. Θεωρούμε ότι το βελάκι καταλήγει πάντα εντός στόχου.)

i) Να βρεθεί το σύνολο τιμών της D, η CDF και η PDF της D.

ii) Να βρεθεί η μέση τιμή και η διακύμανση της D.

iii) Να υπολογισθούν οι πιθανότητες $P(4 \leq D \leq 6)$, $P(D \leq 1)$.

Λύση

Έστω D η ΤΜ που αντιστοιχεί στην απόσταση από το κέντρο, με $S_D = [0, r]$, όπου $r = 10$. Θεωρώντας όλα τα σημεία ισοπίθανα, θα είναι

$$F_D(x) = P(D \leq x) = \frac{\pi x^2}{\pi r^2} = \frac{x^2}{r^2},$$

Λύση (συνέχεια)

οπότε (αφού η F είναι παραγωγίσιμη) $f_D(x) = F'_D(x) = \frac{2x}{r^2}$.
Κατόπιν τούτου, είναι

$$E(D) = \int_0^r x f_D(x) dx = \int_0^r \frac{2x^2}{r^2} dx = \frac{2}{3r^2}\left[x^3\right]_0^r = \frac{2r}{3}$$

$$E(D^2) = \int_0^r x^2 f_D(x) dx = \int_0^r \frac{2x^3}{r^2} dx = \frac{2}{4r^2}\left[x^4\right]_0^r = \frac{r^2}{2}$$

$$V(D) = E(D^2) - (E(D))^2 = \frac{r^2}{2} - \frac{4r^2}{9} = \frac{r^2}{18}$$

Τέλος, είναι
$P(4 \leq D \leq 6) = F_D(6) - F_D(4) = (6^2 - 4^2)/r^2 = 20/r^2 = 0.2$ και
$P(D \leq 1) = F_D(1) = 1/r^2 = 0.01$.

Άσκηση

Μια συνεχής ΤΜ X, με τιμές στο $\mathbb{R}$, έχει PDF

$$f(x) = \begin{cases} ax^2 e^{-kx} & x \geq 0, \\ 0 & x < 0, \end{cases} \qquad \text{όπου } k > 0.$$

i) Να βρεθεί η σταθερά a.

ii) Να βρεθεί η CDF της X.

iii) Να βρεθεί η πιθανότητα $P(0 \leq X \leq 1/k)$.

Λύση

i) Πρέπει $\int_{-\infty}^{+\infty} f(x)dx = 1$, δηλαδή $\int_0^{+\infty} ax^2 e^{-kx} dx = 1$. Θέτουμε

$$I(x) = \int_0^x t^2 e^{-kt} dt \qquad \text{και} \quad I = \lim_{x \to +\infty} I(x) = \int_0^{+\infty} t^2 e^{-kt} dt$$

Λύση (συνέχεια)

και με παραγοντική ολοκλήρωση έχουμε ότι

$$
\begin{aligned}
I(x) &= \int_0^x t^2 \left(\frac{e^{-kt}}{-k}\right)' dt = \left[\frac{t^2 e^{-kt}}{-k}\right]_0^x + \frac{2}{k}\int_0^x te^{-kt}dt \\
&= -\frac{x^2 e^{-kx}}{k} + \frac{2}{k}\int_0^x t\left(\frac{e^{-kt}}{-k}\right)' dt \\
&= -\frac{x^2 e^{-kx}}{k} + \frac{2}{k}\left(\left[\frac{te^{-kt}}{-k}\right]_0^x - \int_0^x \frac{e^{-kt}}{-k}\right) dt \\
&= -\frac{x^2 e^{-kx}}{k} - \frac{2xe^{-kx}}{k^2} + \frac{2}{k^2}\int_0^x e^{-kt}dt \\
&= -\frac{x^2 e^{-kx}}{k} - \frac{2xe^{-kx}}{k^2} + \frac{2}{k^2}\left[\frac{e^{-kt}}{-k}\right]_0^x \\
&= -\frac{x^2 e^{-kx}}{k} - \frac{2xe^{-kx}}{k^2} - \frac{2}{k^3}e^{-kx} + \frac{2}{k^3}.
\end{aligned}
$$

Λύση (συνέχεια)

Παίρνοντας το όριο όταν $x \to +\infty$, βρίσκουμε ότι $I = \lim_{x \to +\infty} I(x) = 2/k^3$.
Εναλλακτικά, χρησιμοποιώντας μετασχηματισμό Laplace, βρίσκουμε απευθείας το ίδιο αποτέλεσμα: $I = \mathcal{L}[x^2](k) = \dfrac{\Gamma(3)}{k^3} = \dfrac{2}{k^3}$.
Επομένως, $a \cdot I = 1 \Leftrightarrow a = k^3/2$ και άρα
$x \in [0, +\infty) \Rightarrow f(x) = \dfrac{k^3}{2} x^2 e^{-kx}$.
ii) Για την CDF $F(x)$ της X, έχουμε ότι $x < 0 \Rightarrow F(x) = 0$ και

$$x \geq 0 \Rightarrow F(x) = \int_0^x f(t)dt = \frac{k^3}{2} I(x) = 1 - \left(1 + kx + \frac{k^2x^2}{2}\right) e^{-kx}.$$

iii) Ισχύει ότι

$$P(0 \leq X \leq 1/k) = F(1/k) - F(0) = 1 - e^{-1}\left(1 + 1 + \frac{1}{2}\right) - 0 = 1 - \frac{5}{2e}.$$

Άσκηση

Έστω X μια συνεχής ΤΜ με PDF $f(x)$ και CDF $F(x)$. Να προσδιορισθούν η CDF $F_Y(x)$ και η PDF $f_Y(x)$ της ΤΜ $Y = |X|$. Αν $f(x) = 1/3$, $x \in [-1, 2]$, να βρεθεί η $f_Y(x)$ της Y.

Λύση

Προφανώς, επειδή $Y \geq 0$, είναι $F_Y(x) = 0$, για κάθε $x \leq 0$, ενώ για $x > 0$ είναι

$$F_Y(x) = P(Y \leq x) = P(|X| \leq x) = P(-x \leq X \leq x) = F(x) - F(-x).$$

Προφανώς, επειδή $Y \geq 0$, είναι $f_Y(x) = 0$, όταν $x < 0$, ενώ σε κάθε $x \geq 0$ όπου f_Y, f συνεχείς, είναι

$$x > 0 \Rightarrow f_Y(x) = F'_Y(x) = F'(x) - (F(-x))' = f(x) + f(-x)$$

Λύση (συνέχεια)

Επομένως, μια PDF για την TM Y είναι η συνάρτηση

$$f_Y(x) = \begin{cases} 0, & x < 0, \\ f(x) + f(-x)), & x \geq 0. \end{cases}$$

Πράγματι, για $x \geq 0$, είναι

$$\begin{aligned} \int_{-\infty}^{x} f_Y(t)dt &= \int_0^x f_Y(t)dt = \int_0^x f(t)dt + \int_0^x f(-t)dt \\ &= \int_0^x f(t)dt - \int_0^{-x} f(t)dt = \int_{-x}^{x} f(t)dt \\ &= F(x) - F(-x) = F_Y(x). \end{aligned}$$

Λύση (συνέχεια)

Αν $f(x) = 1/3$, όταν $x \in [-1, 2]$, τότε έχουμε ότι

$$f_Y(x) = \begin{cases} f(x) + f(-x), & 0 \leq x \leq 2, \\ 0, & \text{αλλιώς}, \end{cases} = \begin{cases} f(x) + f(-x), & 0 \leq x \leq 1, \\ f(x), & 1 \leq x \leq 2, \\ 0, & \text{αλλιώς}, \end{cases}$$

$$= \begin{cases} 2/3, & 0 \leq x \leq 1, \\ 1/3, & 1 \leq x \leq 2, \\ 0, & \text{αλλιώς}. \end{cases}$$