Folytonos valószínűségi változók

Az X egy folytonos valószínűségi változó, ha a lehetséges kimenetelek halmaza egy intervallum vagy egy félegyenes vagy minden valós szám. Ekkor ezt a valószínűségi változót vagy az eloszlásfüggvényével vagy az ún. sűrűségfüggvényével lehet leírni.

Ha X egy (folytonos) v.v., akkor az ő eloszlásfüggvénye:

$$F(x) = P(X < x)$$

Erre teljesül, hogy

1. $F(x)$ monoton nő
2. $\lim_{x\to\infty} F(x) = 1$, $\lim_{x\to-\infty} F(x) = 0$

Folytonos X v.v. esetén egy konkrét x érték valószínűsége 0, azaz $P(X = x) = 0$.

Annak a valószínűségét, hogy az X v.v. az $[a, b]$ intervallumba esik, a következőképpen számíthatjuk ki:

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

$$F(b) - F(a)$$

Példa: Legyen az X v.v. eloszlásfüggvénye

$$F(x) = \begin{cases} 0, & \text{ha} \quad x \leq 1 \\ 1 - \frac{1}{x^3}, & \text{ha} \quad x > 1 \end{cases}$$

Például

$$P(2 < X < 3) = F(3) - F(2) = (1 - \frac{1}{3^3}) - (1 - \frac{1}{2^3}) = \frac{1}{2^3} - \frac{1}{3^3}$$

Az $f(x) = F'(x)$ függvényt az X v.v. <u>sűrűségfüggvény</u>ének hívjuk.

Példa: Az előző eloszlásfüggvénnyel bíró v.v. sűrűségfüggvénye:

$$f(x) = \begin{cases} 0, & \text{ha} \quad x \leq 1 \\ \frac{3}{x^4}, & \text{ha} \quad x > 1 \end{cases}$$

Az $f(x)$ sűrűségfüggvény tulajdonságai:

1. $f(x) \geq 0$, minden $x \in \mathbb{R}$ esetén
2. $\displaystyle\int_{-\infty}^{\infty} f(x)dx = 1$

Ha az X v.v. sűrségfüggvénye adott, akkor ebből eloszlásfüggvényt a következőképpen számolunk:

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

Egy intervallumba esés valószínűsége ekkor

$$P(a < X < b) = F(b) - F(a) = \int_{a}^{b} f(x)dx$$

Az X folytonos v.v. <u>p-edrendű kvantilise</u> (vagy fraktilise) azon x valós szám, amelyre

$$F(x) = p$$

Tipikusan $p = 0{,}01$, 0,05 0,95, 0,99.

Példa: Számítsuk ki a fenti X v.v. esetén a p=0,05 kvantilist!

Megoldás: Olyan x-et keresünk, melyre

$$F(x) = 0,05$$

Ekkor

$$0,05 = 1 - \frac{1}{x^3} \Rightarrow x = 1,017$$

Az X folytonos v.v. várható értékét az átlagértéket használva szintén értelmezhetjük. Ekkor ezt a következő integrállal számolhatjuk:

$$\int_{-\infty}^{\infty} xf(x)dx$$

ha ez az improprius integrál létezik. Példa: A fenti X v.v. esetén

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_{1}^{\infty} x\frac{3}{x^4}dx = \left[3\frac{x^{-2}}{-2}\right]_1^{\infty} =$$

$$\left[\frac{-3}{2x^2}\right]_1^{\infty} = \frac{3}{2}$$

A szórás értelmezéséhez a szórásnégyzetet vezetjük be ismét:

$$D^2(X) = E(X^2) - (E(X))^2,$$

ahol

$$E(X^2) \int_{-\infty}^{\infty} x^2 f(x) dx,$$

ha ez létezik.

Példa: A fenti X szórása a következő:

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx = \int_{1}^{\infty} x^2 \frac{3}{x^4} dx = \left[3 \frac{x^{-1}}{-1} \right]_1^{\infty} =$$

$$\left[\frac{-3}{x} \right]_1^{\infty} = 3.$$

Innen

$$D^2(X) = E(X^2) - (E(X))^2 = 3 - 1,5^2 = 0,75 \Rightarrow D(X) = \sqrt{0,75}$$

Nevezetes folytonos eloszlások

Egyenletes eloszlás

Az X folytonos v.v. egyenletes az $[a, b]$ intervallumon, ha

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{ha} \quad a < x < b \\ 0, & \text{egyébként} \end{cases}$$

Ekkor az eloszlásfüggvény:

$$F(x) = \begin{cases} 0, & \text{ha} \quad x \leq a \\ \frac{x-a}{b-a}, & \text{ha} \quad a < x < b \\ 1, & \text{ha} \quad x > b \end{cases}$$

Megmutatható, hogy

$$E(X) = \frac{a+b}{2}, \quad D(X) = \frac{b-a}{\sqrt{12}}$$

Példák:

1. Legyen *RND* a $[0,1]$ intervallumban egyenletes eloszlású v.v. Ekkor

$$f(x) = \begin{cases} 1, & \text{ha} \qquad 0 < x < 1 \\ 0, & \text{egyébként} \end{cases}$$

$$F(x) = \begin{cases} 0, & \text{ha} \quad x \leq 0 \\ x, & \text{ha} \quad 0 < x < 1 \\ 1, & \text{ha} \quad x > 1 \end{cases}$$

$$E(RND) = \frac{1}{2}, \quad D(RND) = \frac{1}{\sqrt{12}}$$

2. Hogyan kaphatun a $[2,5]$ intervallumon egyenletes eloszlású v.v.-t az *RND*-ből?

$$2 + 3RND$$

Exponenciális eloszlás

Az X folytonos v.v. λ paraméterű exponenciális eloszlást követ, ha

$$F(x) = \begin{cases} 0, & \text{ha} \quad x < 0 \\ 1 - e^{-\lambda x}, & \text{ha} \quad x \geq 0 \end{cases}$$

és ekkor

$$f(x) = \begin{cases} 0, & \text{ha} \quad x < 0 \\ \lambda e^{-\lambda x}, & \text{ha} \quad x \geq 0 \end{cases}$$

Megmutatható, hogy

$$E(X) = \frac{1}{\lambda}, \quad D(X) = \frac{1}{\lambda}$$

Az exponenciális eloszlású X v.v. rendelkezik az ún. örökifjú tulajdonsággal: tetszőleges $s, t > 0$ érték esetén

$$P(X > s + t | X > s) = P(X > t)$$

Példák exponenciális eloszlású v.v.-ra:

1. Radioaktív anyag bomlásának ideje
2. Menzai poharak élettartam
3. Egy műszaki készülék élettartama

Példa: Egy radiaktív anyag felezési ideje 20 év.

1. 50 év alatt az anyag hány százaléka bomlik el?
2. Mennyi idő alatt bomlik el az anyag 75%-a?

Megoldás: Jelölje X egy véletlenül kiválasztott radioaktív atom elbomlásának idejét. Ez exponenciális eloszlást követ. Először meghatározzuk λ-t. Tudjuk:

$$0,5 = P(X < 20) = 1 - e^{-\lambda 20} \to \lambda = -\frac{1}{20} \ln \frac{1}{2} = 0,0346$$

Így az első kérdésre a válasz:

$$P(X < 50) = 1 - e^{-0,0346*50} = 0,822$$

A második kérdéshez olyan x-et kell találni, hogy

$$0,75 = P(X < x) = 1 - e^{-0,0346x},$$

innen $x = 40$

Normális eloszlás

Az X folytonos v.v. standard normális eloszlású, ha sűrűségfüggvénye:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}$$

Ekkor ennek eloszlásfüggvénye:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

Megmutatható, hogy ekkor

$$E(X) = 0 \text{ és } D(X) = 1$$

Az eloszlásfüggvény értékei táblázatból határozhatók meg. A táblázat csak $x > 0$ értékeket tartalmaz, negatívakra a

$$\Phi(-x) = 1 - \Phi(x)$$

összefüggéssel számolunk.

Például:

$$\Phi(2) = 0,9772,$$

$$\Phi(-1) = 1 - \Phi(1) = 1 - 0,8413 = 0,1587$$

Az általános (m, σ) paraméterű normális eloszlású X v.v. sűrűségfüggvénye:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

Az eloszlásfüggvényt a standard normális eloszlás eloszlásfüggvényéből határozhatjuk meg:

$$F(x) = P(X < x) = \Phi(\frac{x-m}{\sigma})$$

Ekkor

$$E(X) = m, \text{ és } D(X) = \sigma$$

Legyen X egy (m, σ) paraméterű normális eloszlású v.v. Határozza meg a

1. $P(X < m + \sigma)$
2. $P(X > m - 3\sigma)$
3. $P(m - 2\sigma < X < m + 2\sigma)$ valószínűségeket!

Megoldás

$$P(X < m+\sigma) = F(m+\sigma) = \Phi(\frac{(m+\sigma) - m}{\sigma}) = \Phi(1) = 0,8413$$

$$P(X > m - 3\sigma) = 1 - P(X < m - 3\sigma) = 1 - F(m - 3\sigma) =$$

$$1 - \Phi(\frac{(m - 3\sigma) - m}{\sigma}) = 1 - \Phi(-3) = 1 - (1 - \Phi(3)) =$$

$$= \Phi(3) = 0,9986$$

$$P(m-2\sigma < X < m+2\sigma) = F(m+2\sigma) - F(m-2\sigma) =$$

$$\Phi(\frac{(m+2\sigma)-m}{\sigma}) - \Phi(\frac{(m-2\sigma)-m}{\sigma}) =$$

$$\Phi(2) - \Phi(-2) = \Phi(2) - (1-\Phi(2)) = 2\Phi(2) - 1 =$$

$$2 * 0,9772 - 1 = 0,9544$$

Példa: Határozza meg az (m, σ) paraméterű normális eloszlású X v.v. p=0,05 kvantilisét!

Megoldás: Olyan x-et keresünk, amelyre

$$0,05 = P(X < x) = \Phi(\frac{x-m}{\sigma})$$

Ehhez keresünk egy a pozitív valós számot, hogy $\Phi(-a) = 0,05$:

$$0,05 = \Phi(-a) = 1 - \Phi(a) \Rightarrow \Phi(a) = 0,95 \Rightarrow a = 1,64$$

Így

$$\Phi(\frac{x-m}{\sigma}) = \Phi(-1,64) \Rightarrow \frac{x-m}{\sigma} = -1,64$$

$$x = m - 1,64\sigma$$

Miért olyan fontos a normális eloszlás?

A v.v. általában normális eloszlású lesz, ha az sok véletlen jelenség összegeként áll elő.

Példák:

1. Egy árúház bevétele egy átlagos hétköznap
2. Egy lakoház gázfogyasztása egy átlagos téli napon
3. 1000 pénzérmét feldobva a fejek száma
4. Az 1kg-os cukor tényleges súlya

Moivre-Laplace tétel

Ha X egy n és p paraméterű binomiális eloszlású v.v., ahol n nagy (tipikusan $n \geq 30$) és p nem pici, akkor ennek eloszlásfüggvénye jól közleíthető az $(np, \sqrt{np(1-p)})$ paraméterű normális eloszlású v.v. eloszlásfüggvényével.

Példa: 720-szor dobunk egy kockával. Határozza meg annak a valószínűségét, hogy a 6-osok száma legfeljebb 110! Adjon becslést is a valószínűségre!

Megoldás: Legyen X az a v.v., amelyik azt adja meg, hogy a 720 dobás közül hány 6-os van. Ez egy $n = 720$, $p = \frac{1}{6}$ paraméterű binomiális eloszlású v.v., ezért

$$P(\text{Legfeljebb 110 db 6-os}) = P(X \leq 110) =$$

$$\sum_{k=0}^{110} P(X = k) = \sum_{k=0}^{110} \binom{720}{k} \left(\frac{1}{6}\right)^k \left(1 - \frac{1}{6}\right)^{720-k}$$

Ezt tudjuk az $m = np = 120$ és $\sigma = \sqrt{720 * \frac{1}{6} * \frac{5}{6}} = 10$ paraméterű normális eloszlással közelíteni. Ezért

$$P(X \leq 110) \approx \Phi(\frac{110 - 120}{10}) = \Phi(-1) = 1 - \Phi(1) =$$

$$1 - 0,8413 = 0,1587.$$

A Moivre-Laplace tétel egy általánosítása a Centrális határeloszlás tétel:

Centrális határeloszlás tétel Legyenek $X_1, X_2, \ldots, X_n$ független azonos eloszlású v.v.-k. Legyen $m = E(X_i)$ és $\sigma = D(X_i)$. Ha n elég nagy, akkor az

$$X = X_1 + X_2 + \cdots + X_n$$

v.v. eloszlásfüggvénye jól közelíthető $(nm, \sqrt{n}\sigma)$ paraméterű normális eloszlású v.v. eloszlásfüggvényével.

Normális eloszlásból származtatott eloszlások

1. Az X lognormális eloszlású v.v., ha $\ln X$ v.v. normális eloszlású.

   Lognormális eloszlás általában aprítási , osztódási, mállási folyamatoknál lép fel, mint a végtermék súlyának, köbtartalmának vagy valamilyen méretnek az eloszlása.

2. Legyenek $X_1, X_2, \ldots, X_n$ független standard normális eloszlású v.v.-k. Ekkor

$$X = X_1^2 + X_2^2 + \cdots + X_n^2$$

   v.v-t n-szabadságfokú χ^2-eloszlásúnak mondjuk. Ennek az eloszlásnak a statisztikában van központi szerepe.