Verjetnost in statistika - definicije, trditve in izreki

Oskar Vavtar
po predavanjih profesorja Jaka Smrekarja

2020/21

Kazalo

1 SLUČAJNI VEKTORJI

Definicija 1.1 (Komulativna porazdelitvena funkcija)**.** Slučajni vektor je taka funkcija/preslikava $\vec{X} = (X_1, \dots, X_n) : \Omega \to \mathbb{R}^n$, kjer je Ω verjetnostni prostor, za katero so množice

$$\begin{aligned}
\{X_1 \leq x_1, \dots, X_n \leq x_n\} &= \{X_1 \in (-\infty, x_1], \dots, X_n \in (-\infty, x_n]\} \\
&= \{\vec{X} \in (-\infty, x_1] \times \dots \times (-\infty, x_n]\} \\
&= \vec{X}^{-1}\left((-\infty, x_1] \times \dots \times (-\infty, x_n]\right)
\end{aligned}$$

dogodki za vse *realne* n-terice $(x_1, \dots, x_n) \in \mathbb{R}^n$.

Komulativna porazdelitvena funkcija slučajnega vektorja $\vec{X}$ je funkcija $F_X : \mathbb{R}^n \to [0, 1]$ s predpisom

$$F_{\vec{X}}(x_1, \dots, x_n) = \mathbb{P}(X_1 \leq x_1, \dots, X_n \leq x_n).$$

Trditev 1.1 (Lastnosti KPF)**.**

(1)

$$\begin{aligned}
\lim_{x_i \to -\infty} F_{\vec{X}}(x_1, \dots, x_n) &= 0 \\
\lim_{\substack{x_1 \to \infty \\ \dots \\ x_n \to \infty}} F_{\vec{X}}(x_1, \dots, x_n) &= 1
\end{aligned}$$

(2) Monotonost:
Če je $x_i \leq y_i$ za $\forall i \in \{1, \dots, n\}$, je

$$F_{\vec{X}}(x_1, \dots, x_n) = F_{\vec{X}}(\vec{x}) \leq F_{\vec{X}}(y).$$

To sledi iz monotonosti $\mathbb{P}$.

(3) Zveznost z desne:

$$\lim_{\vec{y} \searrow \vec{x}} F_{\vec{X}(\vec{y})} = F_{\vec{X}}(\vec{x})$$

Tu $\vec{y} \searrow \vec{x}$ interpretiramo kot $\vec{y_i} \searrow \vec{x_i}$ za $\forall i$.

Opomba. Lastnosti (1), (2) in (3) karakterizirajo družino abstraktnih komulativnih porazdelitvenih funkcij v primeru slučajnih spremeljivk ($n = 1$). V večrazsežnem prostoru to ne drži.

Izrek 1.1. Če je $F : \mathbb{R}^2 \to [0, 1]$ zadošča (1), (2), (3) in (4):

$$F(b, d) - F(a, d) - F(b, c) + F(a, c) \;\geq\; 0$$

za vse četverice $a < b$ in $c < d$, je F *komulativna porazdelitvena funkcija* nekega slučajnega vektorja $(X, Y) : \Omega \to \mathbb{R}^2$.

Definicija 1.2 ('Zvezni' slučajni vektorj)**.** Slučajni vektor $\vec{X} : \Omega \to \mathbb{R}^n$ ima (*zvezno*) gostoto, če obstaja taka *zvezna* funkcija $f_{\vec{X}} : \mathbb{R}^n \to [0, \infty)$, da zanjo velja

$$\mathbb{P}(\vec{x} \in \mathcal{B}) \;=\; \int_{\mathcal{B}} f_{\vec{X}}(x_1, \ldots, x_n)\; dx_1 \ldots dx_n,$$

za vsako Borelovo množico $\mathcal{B} \subset \mathbb{R}^n$.

Posplošitev. Pravimo, da ima vektor $\vec{X} = (X_1, \ldots, X_n)$ n-razsežno normalno porazdelitev s parametrom $\vec{\mu} \in \mathbb{R}$ in $\Sigma \in \mathbb{R}^{n \times n}$ (*simetrična* in *pozitivno definitna*), če ima gostoto:

$$f_{\vec{X}}(\vec{x}) \;=\; (2\pi)^{-\frac{n}{2}} (\det \Sigma)^{-1} e^{-\frac{1}{2} \langle\; \Sigma^{-1}(\vec{x}-\vec{\mu}) \;,\; (\vec{x}-\vec{\mu}) \;\rangle}$$

2 NEODVISNOST

Definicija 2.1. Komponente $X_1, \ldots, X_n$ slučajnega vektorja $\vec{X} = (X_1, \ldots, X_n)$ so *neodvisne*, če velja

$$F_{\vec{X}}(x_1, \ldots, x_n) \;=\; F_{X_1}(x_1) \cdot \ldots \cdot F_{X_n}(x_n)$$

za vse n-terice $(x_1, \ldots, x_n) \in \mathbb{R}^n$.

Opomba. Enakost lahko prepišemo v

$$\mathbb{P}\left(\vec{X} \in (-\infty, x_1] \times \ldots \times (-\infty, x_n]\right) \;=\; \mathbb{P}\left(X_1 \in (-\infty, x_1]\right) \cdot \ldots \cdot \mathbb{P}\left(X_n \in (-\infty, x_n]\right)$$

Trditev 2.1. Naj ima (X, Y) *'zvezno'* gostoto $f(X, Y)$ in naj bosta f_X in f_Y robni gostoti. Tedaj sta X in Y *neodvisni* natanko takrat, ko

$$f_{(X,Y)}(x, y) \;=\; f_X(x) \cdot f_Y(y)$$

za skoraj vse pare x in y.

Trditev 2.2 (Posledica prejšnje trditve)**.** Naj ima (X, Y) *'zvezno'* gostoto $f_{(X,Y)}(x, y)$. Tedaj sta X in Y neodvisni natanko takrat, ko velja

$$f_{(X,Y)}(x, y) \;=\; \Phi(x)\Psi(y)$$

za skoraj vse x in y, za neki *nenegativni* integralski funkciji.

3 FUNKCIJE in TRANSFORMACIJE vektorjev slučajnih spremenljivk