

Verjetnost in statistika - definicije, trditve in izreki

Oskar Vavtar
po predavanjih profesorja Jaka Smrekarja

2020/21

Kazalo

1	SLUČAJNI VEKTORJI	3
2	NEODVISNOST	5
3	FUNKCIJE in TRANSFORMACIJE	
	slučajnih spremenljivk in vektorjev	6
3.1	Slučajne spremenljivke	6
3.2	Slučajni vektorji	7
4	PRIČAKOVANA VREDNOST	
	zveznih slučajnih spremenljivk in vektorjev	8
4.1	Pričakovana vrednost funkcij slučajnih spremenljivk in vektorjev	8
4.2	Disperzija, kovarianca in variančno-kovariančno matrika . . .	10

1 SLUČAJNI VEKTORJI

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

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

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

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

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

(2) MONOTONOST:

Če je $x_i \leq y_i$ za $\forall i \in \{1, \dots, n\}$, je

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

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 spremenljivk ($n = 1$). V večrazsežnem prostoru to ne drži.

Izrek 1.1. Če je $F : \mathbb{R}^2 \rightarrow [0, 1]$ 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 $(X, Y) : \Omega \rightarrow \mathbb{R}^2$.

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

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

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

Posplošitev. Pravimo, da ima vektor $\vec{X} = (X_1, \dots, X_n)$ *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, \dots, X_n slučajnega vektorja $\vec{X} = (X_1, \dots, X_n)$ so *neodvisne*, če velja

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

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

Opomba. Enakost lahko prepišemo v

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

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 slučajnih spremenljivk in vektorjev

3.1 Slučajne spremenljivke

Trditev 3.1 (Diskretni primer). Če je X *diskretna slučajna spremenljivka* z vrednostmi $\{x_i \mid i \in I\}$ in je g funkcija, ki preslika množico $\{x_i \mid i \in I\}$ na množico $\{y_j \mid j \in J\}$, je $g \circ X = g(X)$ *diskretna slučajna spremenljivka* z verjetnostno funkcijo

$$\mathbb{P}(g(X) = y_j) = \sum_{i: g(x_i)=y_j} \mathbb{P}(X = x_i)$$

Trditev 3.2 (Zvezni primer). Naj ima slučajna spremenljivka X 'zvezno' gostoto f_X , ki je različna od 0 natanko na intervalu (a, b) , kjer $-\infty \leq a \leq b \leq \infty$.

Naj bo $g : (a, b) \rightarrow (c, d)$ *zvezna bijekcija*. Zanima nas funkcija $g(X)$ slučajne spremenljivke X .

Velja tudi

$$\mathbb{P}(X \in (a, b)) = \int_a^b f_X(x) dx = \int_{\mathbb{R}} f_X(x) dx = 1.$$

Velja $\{g(X) \leq z\} = \{X \leq g^{-1}(z)\}$. Ker je X slučajna spremenljivka, so $\{X \leq g^{-1}(z)\}$ dogodki za $z \in (c, d)$. Sledi, da je $g(X)$ slučajna spremenljivka in velja:

$$F_{g(X)}(z) = \mathbb{P}(g(X) \leq z) = \begin{cases} 1 & ; \quad z \geq d \\ F_X(g^{-1}(z)) & ; \quad z \in (c, d) \\ 0 & ; \quad z \leq c \end{cases}$$

Če je g *odvedljiva*, sledi:

$$f_{g(X)}(z) = \frac{d}{dz} F_{g(X)}(z) = \begin{cases} \frac{f_X(g^{-1}(z))}{g'(g^{-1}(z))} & ; \quad z \in (c, d) \\ 0 & ; \quad \text{sicer} \end{cases}$$

Za splošno *odvedljivo bijekcijo* $g : (a, b) \rightarrow (c, d)$ velja t.i. *transformacijska formula*:

$$f_{g(X)}(z) = \frac{f_X(g^{-1}(z))}{|g'(g^{-1}(z))|}, \quad \text{za } z \in (c, d).$$

3.2 Slučajni vektorji

Trditev 3.3. Naj bo $\vec{X} : \Omega \rightarrow \mathbb{R}^n$ slučajni vektor in naj bo $h : \mathbb{R}^n \rightarrow \mathbb{R}$ zvezna funkcija. Tedaj je $h(\vec{X}) = h(X_1, \dots, X_n)$ slučajna spremenljivka.

Trditev 3.4. Naj bo $\vec{X} : \Omega \rightarrow \mathbb{R}^n$ slučajni vektor z gostoto $f_{\vec{X}}$. Dalje naj bo $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ (ali $g : D \rightarrow E$ za primerni množici $D, E \subset \mathbb{R}^n$, kjer $\mathbb{P}(\vec{X} \in D) = 1$) zvezno diferenciablena bijekcija. Tedaj ima slučajni vektor $g(\vec{X})$ gostoto

$$f_{g(\vec{X})}(\vec{z}) = f_{\vec{X}}(g^{-1}(\vec{z})) \cdot |\det Jg^{-1}(\vec{z})|$$

v točkah $\vec{z} \in \mathbb{R}^n$ (oz. $\vec{z} \in E$).

Za dvorazsežne vektorje, kjer je $\vec{z} = (u, v)$ in $(U, V) = g(X, Y)$ ter posledično $(X, Y) = g^{-1}(U, V)$ se transformacijska formula glasi

$$f_{(U,V)}(u, v) = f_{(X,Y)}(g^{-1}(u, v)) \cdot |\det Jg^{-1}(u, v)|$$

oziroma

$$f_{(U,V)}(u, v) = f_{(X,Y)}(x(u, v), y(u, v)) \cdot \left\| \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} \right\|.$$

Upošteva se $g \circ g^{-1} = \text{id}$ in posledično

$$Jg(g^{-1}(u, v)) \cdot Jg^{-1}(u, v) = I$$

zgornje prepišemo v

$$f_{(U,V)}(u, v) = \frac{f_{(X,Y)}(g^{-1}(u, v))}{|\det Jg(g^{-1}(u, v))|} = \frac{f_{(X,Y)}(x(u, v), y(u, v))}{\left\| \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} \right\|_{(x(u,v), y(u,v))}}.$$

Trditev 3.5. Naj za $\vec{X} : \Omega \rightarrow \mathbb{R}^n$ velja $\vec{X} \sim N(\vec{\mu}, \Sigma)$ in naj bo $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ obrnljiva matrika. Dalje naj bo $\vec{v} \in \mathbb{R}^n$. Tedaj

$$A\vec{x} + \vec{v} \sim N(A\vec{\mu} + \vec{v}, A\Sigma A^T).$$

4 PRIČAKOVANA VREDNOST zveznih slučajnih spremenljivk in vektorjev

Definicija 4.1. Naj bo X slučajna spremenljivka z 'zvezno' gostoto $f_X : \mathbb{R} \rightarrow [0, \infty)$. *Pričakovana vrednost* slučajne spremenljivke X je

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x \cdot f_X dx,$$

če integral absolutno konvergira.

Za slučajni vektor $\vec{X} = (X_1, \dots, X_n)$ definiramo pričakovano vrednost po komponentah

$$\mathbb{E}(\vec{X}) = (\mathbb{E}(X_1), \dots, \mathbb{E}(X_n)) \in \mathbb{R}^n.$$

Podobno definiramo tudi pričakovano vrednost slučajnih matrik.

4.1 Pričakovana vrednost funkcij slučajnih spremenljivk in vektorjev

Trditev 4.1. Naj bo $g : \mathbb{R} \rightarrow \mathbb{R}$ (ali $g : (a, b) \rightarrow (c, d)$) zvezna funkcija. Tedaj je

$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x) dx,$$

če obstaja.

Posledica. Za $p \neq 0$ velja:

$$\mathbb{E}(|x|^p) = \int_0^{\infty} |z|^p (f_X(z) + f_X(-z)) dz = \dots = \int_{-\infty}^{\infty} |x|^p f_X(x) dx.$$

Trditev 4.2. Naj bo $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ (oziroma $h : D \rightarrow \mathbb{R}$ v primeru, da $\mathbb{P}((X, Y) \in D) = 1$) *zvezna* funkcija in (X, Y) slučajni vektor z gostoto $f_{(X, Y)}$. Tedaj velja

$$\mathbb{E}(h(X, Y)) = \int_{\mathbb{R}^2 \text{ (ali } D)} h(x, y) f_{(X, Y)}(x, y) dx dy,$$

če obstaja.

Posledica.

$$\mathbb{E}(XY) = \int_{\mathbb{R}} x \cdot y \cdot f_{(X, Y)}(x, y) dx dy,$$

če integral absolutno konvergira.

Opomba. Če sta X in Y *neodvisni* slučajni spremenljivki, za kateri obstajata $\mathbb{E}(X)$ in $\mathbb{E}(Y)$, potem obstaja $\mathbb{E}(XY)$ in je

$$\mathbb{E}(XY) = \mathbb{E}(X) \cdot \mathbb{E}(Y).$$

Posledica (Aditivnosti pričakovane vrednosti). Če je $X \leq Y$ (z verjetnostjo 1), velja

$$\mathbb{E}(X) \leq \mathbb{E}(Y)$$

če $\mathbb{E}(|Y|) < \infty$.

Izrek 4.1. Naj za slučajni spremenljivki X in Y velja $\mathbb{E}(X^2) < \infty$ in $\mathbb{E}(Y^2) < \infty$. Tedaj obstaja $\mathbb{E}(XY)$ in velja

$$\mathbb{E}(|XY|) \leq \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}.$$

$$\text{Enakost nastopi} \iff \frac{Y}{\sqrt{\mathbb{E}(Y^2)}} = \pm \frac{X}{\sqrt{\mathbb{E}(X^2)}} \text{ skoraj gotovo.}$$

Posledica. Če je $\mathbb{E}(X^2) < \infty$, je $\mathbb{E}(|X|) < \infty$ in

$$\mathbb{E}(|X|) \leq \sqrt{\mathbb{E}(X^2)}.$$

4.2 Disperzija, kovarianca in variančno-kovariančno matrika

Definicija 4.2. *Disperzija (razpršenost, varianca)* slučajne spremenljivke X je

$$\mathbb{E}((X - \mathbb{E}(X))^2),$$

če ta pričakovana vrednost obstaja. Ta obstaja, če obstaja $\mathbb{E}(X^2)$ in tedaj velja

$$\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2.$$

Definicija 4.3. *Standardni odklon* slučajne spremenljivke X definiramo kot

$$\sigma(X) = \sigma_X = \sqrt{\text{Var}(X)}.$$

Pripomnimo, da ima $\sigma(X)$ iste enote kot X in je zato $\sigma(X)$ 'količina' ki je neposredno primerljiva z X .

Trditev 4.3 (Lastnosti variance). Naj za slučajno spremenljivko X obstaja $\mathbb{E}(X^2)$ (torej obstaja tudi $\text{Var}(X)$):

- $\text{Var}(X) \geq 0$ in $\text{Var}(X) = 0 \iff X \equiv \mathbb{E}(X)$ skoraj gotovo,
- $\text{Var}(X)$ je minimum funkcije $a \mapsto \mathbb{E}((X - a)^2)$,
- $\text{Var}(aX + b) = a^2 \text{Var}(X)$,
- $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$, če sta X in Y neodvisni.

Komentar. Vemo, da je σ^2 v $\mathcal{N}(\mu, \sigma^2)$ disperzija.

Opomba. Za *zvezno* slučajno spremenljivko X z gostoto f_X je

$$\text{Var}(X) = \int_{-\infty}^{\infty} (x - \mathbb{E}(X))^2 f_X(x) dx.$$

Definicija 4.4. Kovarianca slučajnih spremenljivk X in Y je

$$Cov(X, Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))),$$

če obstaja. $Cov(X, Y)$ obstaja, če obstajajo $\mathbb{E}(X)$, $\mathbb{E}(Y)$, $\mathbb{E}(XY)$. Če je tako, velja

$$Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

Če sta X in Y neodvisni in imata pričakovano vrednost, potem

$$Cov(X, Y) = 0.$$

Trditev 4.4 (Lastnosti kovariance).

- $Cov(X, Y)$ obstaja, če obstajata $Var(X)$ in $Var(Y)$ ter velja *Cauchy-Schwarzova neenakost*

$$|Cov(X, Y)| \leq \sqrt{Var(X)Var(Y)} = \sigma(X)\sigma(Y),$$

enakost nastopi $\iff \frac{Y - \mathbb{E}(Y)}{\sigma(Y)} = \pm \frac{X - \mathbb{E}(X)}{\sigma(X)}$ skoraj gotovo. V tem primeru pravimo, da sta X in Y skoraj gotovo linearno povezani.

- $Cov(X, X) = Var(X)$
- Kovarianca je *simetrična bilinearna* funkcija. Dovolj je preveriti bilinearnost v eni spremenljivki.
- Če ima (X, Y) *zvezno* gostoto $f_{X,Y}$, potem je

$$Cov(X, Y) = \iint_{\mathbb{R}^2} (x - \mathbb{E}(X))(y - \mathbb{E}(Y))f_{(X,Y)}(x, y) dx dy.$$

- Če imata X in Y disperzijo, jo ima tudi $X + Y$ in velja

$$Var(X + Y) = Var(X) + 2Cov(X, Y) + Var(Y).$$

Posplošitev na disperzijo vrste $X_1 + \dots + X_n$:

$$Var(X_1 + \dots + X_n) = \sum_{i=1}^n Var(X_i) + 2 \sum_{1 \leq i < j \leq n} Cov(X_i, X_j)$$

Definicija 4.5. *Variančno-kovariančna matrika* slučajnega vektorja $\vec{X} = (X_1, \dots, X_n)$ je matrika $Var(\vec{X}) \in \mathbb{R}^{n \times n}$ z elementi

$$[Var(\vec{X})]_{ij} = Cov(X_i, X_j).$$

To je *simetrična* matrika z diagonalnimi elementi $Var(X_1), \dots, Var(X_n)$.

Definicija 4.6. *Pearsonov korelacijski koeficient* slučajnih spremenljivk X in Y z disperzijo je

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sigma(X)\sigma(Y)} \in [-1, 1].$$

Definicija 4.7. Slučajni spremenljivki X in Y sta *nekorelirani* natanko takrat, ko

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y).$$

Vidimo:

- Če sta X in Y *neodvisni* slučajni spremenljivki s pričakovano vrednostjo, sta *nekorelirani*.
- Obrat v splošnem ne drži, drži pa v primeru dvorazsežne normalno porazdeljenega slučajnega vektorja.
- Če imata X in Y disperzijo, sta nekorelirani $\iff \rho(X, Y) = 0$.