

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
5	POGOJNE PORAZDELITVE in	
	POGOJNA PRIČAKOVANA VREDNOST	13
6	MOMENTI in	
	MOMENTNO-RODOVNA FUNKCIJA	15
7	LIMITNI IZREKI	18
8	TOČKOVNO OCENJEVANJE	21
8.1	Metode za konstrukcijo cenilk	22
9	INTERVALSKO OCENJEVANJE	24
10	PREIZKUŠANJE DOMNEV	26

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:

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

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

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

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

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

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 $\iff \rho(X, Y) = 0$.

5 POGOJNE PORAZDELITVE in POGOJNA PRIČAKOVANA VREDNOST

Definicija 5.1. Naj bo $\mathcal{B} \subseteq \mathbb{R}$ Borelova množica. Pogojna verjetnost, da Y zavzame vrednost v \mathcal{B} pri pogoju $X = x$ je

$$\begin{aligned}\mathbb{P}(Y \in \mathcal{B} \mid X = x) &= \lim_{h \searrow 0} \mathbb{P}(Y \in \mathcal{B} \mid X \in (x - h, x + h)) = \\ &= \lim_{h \searrow 0} \frac{\mathbb{P}(Y \in \mathcal{B}, X \in (x - h, x + h))}{\mathbb{P}(X \in (x - h, x + h))},\end{aligned}$$

če ta limita obstaja.

Trditev 5.1. Naj ima (X, Y) “zvezno” gostoto $f_{(X,Y)}$ in naj bo $f_X(x)$ zvezna v x . Tedaj je

$$\mathbb{P}(Y \in \mathcal{B} \mid X = x) = \int_{\mathcal{B}} \frac{f_{(X,Y)}(x, y)}{f_X(x)} dy.$$

Opomba. Če sta X, Y neodvisni, je

$$f_{(Y|X)}(y \mid x) = f_Y(y).$$

Brez privzetka zveznosti je

$$\mathbb{P}(Y \in \mathcal{B} \mid X = x) = \mathbb{P}(Y \in \mathcal{B}).$$

Trditev 5.2. Velja tako imenovani zakon popolne verjetnosti:

$$\mathbb{P}(Y \in \mathcal{B}) = \int_{-\infty}^{\infty} \mathbb{P}(Y \in \mathcal{B} \mid X = x) f_X(x) dx.$$

Definicija 5.2. Pogojna pričakovana vrednost Y , pogojna na $X = x$, je pričakovana vrednost pogojne porazdelitve $(Y \mid X = x)$. V primeru, ko ima (X, Y) “zvezno” gostoto, to pomeni

$$\mathbb{E}(Y \mid X = x) = \int_{-\infty}^{\infty} y \cdot f_{(Y|X)}(y \mid x) dy.$$

Trditev 5.3. Zakon *popolne pričakovane vrednosti*:

$$\mathbb{E}(Y) = \int_{-\infty}^{\infty} \mathbb{E}(Y \mid X = x) f_X(x) dx.$$

Komentar. Velja tudi:

$$\begin{aligned} \mathbb{E}(g(Y) \mid X = x) &= \int_{-\infty}^{\infty} g(y) f_{(Y \mid X)}(y \mid x) dy \\ \text{in} \quad \mathbb{E}(g(Y)) &= \int_{-\infty}^{\infty} \mathbb{E}(g(Y) \mid X = x) f_X(x) dx. \end{aligned}$$

Definicija 5.3. Pogojna pričakovana vrednost $\mathbb{E}(Y \mid X)$ je slučajna spremenljivka, definirana¹ kot:

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

Če pišemo $g(x) = \mathbb{E}(Y \mid X = x)$, je $\mathbb{E}(Y \mid X) = g(X)$ *transformacija* slučajne spremenljivke X .

Sledi:

$$\begin{aligned} \mathbb{E}(\mathbb{E}(Y \mid X)) &= \mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(y) f_X(x) dx = \\ &= \int_{-\infty}^{\infty} \mathbb{E}(Y \mid X = x) f_X(x) dx = \mathbb{E}(Y). \end{aligned}$$

Pravimo, da je $\mathbb{E}(Y \mid X)$ najboljši približek za Y , če poznamo X .

¹Na istem verjetnostnem prostoru.

6 MOMENTI in MOMENTNO-RODOVNA FUNKCIJA

Definicija 6.1. Naj bo X slučajna spremenljivka in naj bosta $k \in \mathbb{N}$ in $a \in \mathbb{R}$. Moment slučajne spremenljivke X glede na a je

$$m_k(a) = \mathbb{E}((X - a)^k),$$

če obstaja, torej če $\mathbb{E}(|X - a|^k) < \infty$.

V zveznem primeru je

$$m_k(a) = \int_{-\infty}^{\infty} (x - a)^k f(x) dx,$$

če integral absolutno konvergira.

Pravimo:

- $m_k(0) \dots k$ -ti *začetni* moment
- $m_k(\mathbb{E}(X)) \dots k$ -ti *centralni* moment

Trditev 6.1. Če obstaja $m_n(a)$, potem obstajajo $m_k(a)$ za $k < n$.

Trditev 6.2. Če obstaja $m_n(a)$ za neki $a \in \mathbb{R}$, obstaja $m_n(b)$ za $b \in \mathbb{R}$.

Posledica. Če m_n obstaja, je

$$m_n(b) = \sum_{k=0}^n \binom{n}{k} (a - b)^{n-k} m_k(a).$$

Izrek 6.1. Naj obstajajo vsi momenti $m_k = \mathbb{E}(X^k)$ in naj vrsta

$$\sum_{k=0}^{\infty} \frac{1}{k!} t^k \mathbb{E}(X^k)$$

absolutno konvergira za neki $t > 0$. Potem je porazdelitev slučajne spremenljivke X enolično določena z momenti. Če ima Y enako lastnost in velja

$$\mathbb{E}(X^k) = \mathbb{E}(Y^k) \quad \forall k,$$

sta X in Y enako porazdeljeni.

Komentar. To pomeni, da je komulativna porazdelitvena funkcija $F_X : \mathbb{R} \rightarrow [0, 1]$ določena s *števnim* naborom števil.

Definicija 6.2. *Momentno-rodovna funkcija* slučajne spremenljivke X je funkcija

$$M_X(t) = \mathbb{E}(e^{tX}),$$

ki je definirana za tista realna števila $t \in \mathbb{R}$, za katera pričakovana vrednost obstaja. Vedno je

$$M_X(0) = \mathbb{E}(1) = 1.$$

Za $t \neq 0$ je $y(x) = e^{tx}$ zvezno odvedljiva bijekcija $\mathbb{R} \rightarrow (0, \infty)$ in za slučajno spremenljivko X z gostoto f_X velja

$$M_X(t) = \sum_{k=0}^{\infty} \frac{1}{k!} t^k \mathbb{E}(X^k).$$

To pomeni, da je M_X *analitična*; porazdelitev slučajne spremenljivke X je *enolično* določena z momenti.

Trditev 6.3.

$$M_{aX+b}(t) = e^{tb} M_X(at)$$

Posledica.

$$M_{\mathcal{N}(\mu, \sigma^2)}(t) = M_{\sigma \cdot \mathcal{N}(0,1) + \mu}(t) = e^{t\mu} \cdot e^{\frac{(\sigma t)^2}{2}}$$

Trditev 6.4. Če sta slučajni spremenljivki X in Y *neodvisni*, velja

$$M_{X+Y}(t) = M_X(t) \cdot M_Y(t).$$

Lema 1. Če sta slučajni spremenljivki X in Y *neodvisni* in $f, g : \mathbb{R} \rightarrow \mathbb{R}$ (ali $f : (a, b) \rightarrow \mathbb{R}$, $g : (c, d) \rightarrow \mathbb{R}$) *zvezni* funkciji, potem sta $f(X)$ in $g(Y)$ neodvisni:

$$\begin{aligned} \mathbb{P}(f(X) \in \mathcal{B}, g(Y) \in \mathcal{C}) &= \mathbb{P}(X \in f^{-1}(\mathcal{B}), Y \in g^{-1}(\mathcal{C})) = \\ &= \mathbb{P}(X \in f^{-1}(\mathcal{B})) \cdot \mathbb{P}(Y \in g^{-1}(\mathcal{C})), \end{aligned}$$

kjer smo pri zadnjem enačaju upoštevali *neodvisnost*.

7 LIMITNI IZREKI

Izrek 7.1 (Krepki zakon velikih števil). Verjetnost tistih vzorcev $s = (\omega_1, \omega_2, \omega_3, \dots) \in S$, za katere je

$$\lim_{n \rightarrow \infty} \frac{X(\omega_1) + \dots + X(\omega_n)}{n} = \lim_{n \rightarrow \infty} \frac{X_1(s) + \dots + X_n(s)}{n} = \mathbb{E}(x),$$

je enaka 1.

Trditev 7.1 (Markova neenakost). Naj bo X slučajna spremenljivka s pričakovano vrednostjo in $a > 0$. Tedaj je

$$\mathbb{P}(|X| > a) \leq \frac{1}{|a|} \mathbb{E}(|X|).$$

Posledica (Čebiševa neenakost). Naj bo $\mathbb{E}(X^2) < \infty$ in bo $\varepsilon > 0$. Tedaj je

$$\mathbb{P}(|X - \mathbb{E}(X)| > \varepsilon) = \frac{\text{Var}(X)}{\varepsilon^2}.$$

Izrek 7.2 (Šibki zakon velikih števil Markova). Naj bodo X_1, X_2, \dots neodvisne in enako porazdeljene slučajne spremenljivke z varianco $\sigma^2 < \infty$ in pričakovano vrednostjo μ . Tedaj $\forall \varepsilon > 0$ velja

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\left|\frac{X_1 + \dots + X_n}{n} - \mu\right| > \varepsilon\right) = 0.$$

Definicija 7.1. Naj bodo Y, Y_1, Y_2, Y_3, \dots slučajne spremenljivke, definirane na skupnem verjetnostnem prostoru. Pravimo, da zaporedje $\{Y_n\}_n$ verjetnostno konvergira k Y , pišemo $Y_n \xrightarrow[n \rightarrow \infty]{p} Y$, če $\forall \varepsilon > 0$ velja

$$\lim_{n \rightarrow \infty} \mathbb{P}(|Y_n - Y| > \varepsilon) = 0.$$

Komentar. Če so X_1, X_2, \dots neodvisne enako porazdeljene slučajne spremenljivke s pričakovano vrednostjo μ in disperzijo σ^2 , potem

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow[n \rightarrow \infty]{p} \mu.$$

$\forall \varepsilon > 0 :$

$$\lim_{n \rightarrow \infty} \mathbb{P}(|\bar{X} - \mu| > \varepsilon) = 0.$$

Definicija 7.2. Zaporedje $\{Y_n\}_n$ konvergira *skoraj gotovo* k Y , $Y_n \xrightarrow[n \rightarrow \infty]{s.g.} Y$, če je

$$\mathbb{P}(\lim_{n \rightarrow \infty} Y_n = Y) = \mathbb{P}(\{s \mid \exists \lim_{n \rightarrow \infty} Y_n(s) = Y(s)\}) = 1.$$

Izrek 7.3 (Krepki zakon velikih števil Kolmogorova). Naj bodo X_1, X_2, \dots neodvisne in enako porazdeljene slučajne spremenljivke s pričakovano vrednostjo μ . Tedaj velja

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow[n \rightarrow \infty]{s.g.} \mu.$$

Komentar. • obstoj variance ni potreben

- za $X : \Omega \rightarrow \mathbb{R}$ smo konstruirali $X_i : \Omega^{\mathbb{N}} \rightarrow \mathbb{R}$

Trditev 7.2. Iz *skoraj gotove* konvergence sledi konvergenca v *verjetnosti*.

Trditev 7.3. Naj bo $g : \mathbb{R} \rightarrow \mathbb{R}$ zvezna funkcija:

- (1) Če velja $Y_n \xrightarrow[n \rightarrow \infty]{s.g.} Y$, potem $g(Y_n) \xrightarrow[n \rightarrow \infty]{s.g.} g(Y)$.
- (2) Če velja $Y_n \xrightarrow[n \rightarrow \infty]{p} Y$, potem $g(Y_n) \xrightarrow[n \rightarrow \infty]{p} g(Y)$.

Izrek 7.4 (Centralni limitni izrek). Naj bo X_1, X_2, \dots zaporedje neodvisnih in enako porazdeljenih slučajnih spremenljivk s pričakovano vrednostjo μ in varianco $\sigma^2 < \infty$. Tedaj $\forall x \in \mathbb{R}$:

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{\frac{X_1 + \dots + X_n}{n} - \mu}{\frac{\sigma}{\sqrt{n}}} \leq x \right) = \Phi(x) = \mathbb{P}(\mathcal{N}(0, 1) \leq x).$$

Komentar. $\frac{\frac{X_1 + \dots + X_n}{n} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$ je standardizacija od \bar{X} . Izjava centralnega limitnega izreka pravi, da komulativne porazdelitvene funkcije standardiziranih vzorčnih povprečij po točkah konvergirajo k komulativni porazdelitveni funkciji Φ . Za realni števili $a < b$ sledi

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \in (a, b] \right) = \mathbb{P} \left(\bar{X} \in \underbrace{\left(\mu + a \frac{\sigma}{\sqrt{n}}, \mu + b \frac{\sigma}{\sqrt{n}} \right]}_c \right).$$

Torej:

$$\mathbb{P}(\bar{X} \in (c, d]) \approx \Phi \left(\frac{d - \mu}{\frac{\sigma}{\sqrt{n}}} \right) - \Phi \left(\frac{c - \mu}{\frac{\sigma}{\sqrt{n}}} \right) = F_{\mathcal{N}(\mu, \frac{\sigma^2}{n})}(d) - F_{\mathcal{N}(\mu, \frac{\sigma^2}{n})}(c)$$

oziroma

$$\mathbb{P}(\bar{X} \in (c, d]) \approx \mathbb{P} \left(\mathcal{N} \left(\mu, \frac{\sigma^2}{n} \right) \in (c, d] \right) \quad \text{za velike } n.$$

Rečemo tudi

$$\bar{X} \sim \mathcal{N} \left(\mu, \frac{\sigma^2}{n} \right).$$

Izrek 7.5 (Izrek o zveznosti). Če za slučajne spremenljivke Y_1, Y_2, \dots , ki imajo momentno-rodovne funkcije na nekem intervalu $(-\delta, \delta)$, $\delta > 0$, velja

$$\lim_{n \rightarrow \infty} M_{Y_n}(t) = e^{\frac{t^2}{2}} = M_{\mathcal{N}(0,1)}(t) \quad \forall t \in (-\delta, \delta),$$

potem $\forall x \in \mathbb{R}$ velja

$$\lim_{n \rightarrow \infty} \mathbb{P}(Y_n \leq x) = \Phi(x).$$

8 TOČKOVNO OCENJEVANJE

Definicija 8.1. Naj bo c realnoštevilska “karakteristika” proučevane porazdelitve². Cenilka za c je funkcija slučajnega vzorca $T = T(X_1, \dots, X_n)$, s katero ocenjujemo c . Cenilka je določena s funkcijo $T : \mathbb{R}^n \rightarrow \mathbb{R}$; za $T(X_1, \dots, X_n) = \bar{X}$ je

$$T(x_1, \dots, x_n) = \frac{1}{n}(x_1 + \dots + x_n).$$

Cenilka T karakteristike c je *nepristranska*, če velja

$$\mathbb{E}(T(X_1, \dots, X_n)) = c(X)$$

za vsako dopustno porazdelitev. Natančneje: za vsako dopustno porazdelitev (k.p.f.) $F \in \mathcal{F}$ in vsak vzorec $X_1, \dots, X_n \stackrel{NEP}{\sim} F$ velja

$$\mathbb{E}(T(X_1, \dots, X_n)) = c(F).$$

Definicija 8.2. Naj bosta U in V *nepristranski* cenilki za $c(X)$ v modelu \mathcal{F} . Tedaj ima U enakomerno manjšo varianco od V , če velja

$$\text{Var}(U(X_1, \dots, X_n)) \leq \text{Var}(V(X_1, \dots, X_n))$$

$\forall F \in \mathcal{F}$ in vsak vzorec $X_1, \dots, X_n \stackrel{NEP}{\sim} F$.

Definicija 8.3. Zaporedje cenilk T_n za karakteristike c , je *dosledno* (angl. *consistent*), če $\forall F \in \mathcal{F}$ in vsako neskončno zaporedje $X_1, X_2, \dots \stackrel{NEP}{\sim} F$ zaporedje $T_n(X_1, \dots, X_n)$ konvergira h konstanti $c(F)$ verjetnostno.

Komentar. Če X_1, X_2, \dots generiramo kot *neodvisne* replikacije slučajne spremenljivke $X : \Omega \rightarrow \mathbb{R}$, lahko X_i razumemo kot funkcijo na prostoru neskončnih vzorcev

$$S = \Omega^{\mathbb{N}} = \{\text{neskončna zaporedja iz } \Omega\}.$$

²npr. $c = \mathbb{E}(X)$, $c = \text{Var}(X)$, $c = \mathbb{P}(X \geq 2)$, ...

8.1 Metode za konstrukcijo cenilk

Metoda 1 (Metoda momentov). Zanima nas karakteristika $c(X)$ proučevane slučajne spremenljivke $X : \Omega \rightarrow \mathbb{R}$ v nekem modelu. Če je $c(X)$ mogoče izraziti z momenti kot $c(X) = g(m_1, m_2, \dots, m_r)$, potem $c(X)$ ocenjujemo s cenilko $\hat{c} = g(\hat{m}_1, \hat{m}_2, \dots, \hat{m}_r)$. Če je g zvezna, je cenilka \hat{c} dosledna. V praksi imamo najpogostejše parametrični model z vektorskim parametrom $\vartheta = (\vartheta_1, \dots, \vartheta_r)$, kjer izrazimo

$$\begin{aligned} m_1 &= m_1(\vartheta_1, \dots, \vartheta_r) \\ &\vdots \\ m_r &= m_r(\vartheta_1, \dots, \vartheta_r) \end{aligned}$$

Privzemimo, da znamo zgornji sistem razrešiti na ϑ z vektorsko funkcijo $(\vartheta_1, \dots, \vartheta_r) = (g_1(m_1, \dots, m_r), \dots, g_r(m_1, \dots, m_r))$. Potem za ϑ vzamemo cenilko

$$\hat{\vartheta} = (g_1(\hat{m}_1, \dots, \hat{m}_r), \dots, g_r(\hat{m}_1, \dots, \hat{m}_r)).$$

Že vemo: če je $g = (g_1, \dots, g_r)$ zvezna, je $\hat{\vartheta}$ dosledna.

Metoda 2 (Metoda največjega verjetja). Privzemimo parametrični model s prostorom parametrov $\Theta \subset \mathbb{R}^r$ in splošnim parametrom $\vartheta = (\vartheta_1, \dots, \vartheta_r) \in \Theta$. Dalje privzemimo, da imajo dopustne porazdelitve verjetnostne funkcije (v diskretnem modelu) oz. gostote (v zveznem modelu) oblike

$$f(x; \vartheta) = f(x; \vartheta_1, \dots, \vartheta_r).$$

Definicija 8.4. Funkcija verjetja za vzorec velikosti n je $L : \mathbb{R}^n \times \Theta \rightarrow [0, \infty)$,

$$L(x_1, \dots, x_n; \vartheta) = f(x_1; \vartheta) \cdot \dots \cdot f(x_n; \vartheta).$$

Kot funkcija vzorca $\vec{x} = (x_1, \dots, x_n)$ je L gostota vektorja (x_1, \dots, x_n) . V teoriji verjetja L pri danem \vec{x} gledamo kot funkcijo parametra ϑ . Ocenimo za ϑ pri danem \vec{x} po metodi najmanjših kvadratov je tak $\hat{\vartheta} \in \Theta$, pri katerem ima $L(\vec{x}; \vartheta)$ maksimum, torej velja:

$$L(\vec{x}; \hat{\vartheta}) = \max_{\vartheta \in \Theta} L(\vec{x}; \vartheta).$$

Oceno $\hat{\vartheta}$ analitično tipično izračunamo z odvajanjem. Zaradi preprostosti raje odvajamo logaritemsko funkcijo verjetja:

$$\ln(L(\vec{x}; \vartheta)) = \sum_{i=1}^n \ln f(x_i; \vartheta).$$

Stacionarne točke funkcije $\ln L$ imenujemo rešitve *enačb verjetja* (EV)

$$\frac{\partial}{\partial \vartheta_j}(\ln L) = 0, \quad 1 \leq j \leq n.$$

Definicija 8.5 (Cenilka največjega verjetja). Če za vse možne (v diskretnem primeru) ali skoraj vse (v zveznem primeru) realizacije \vec{x} slučajnega vektorja \vec{X} obstaja ocena $\hat{\vartheta}(\vec{x})$ za ϑ po metodi največjega verjetja, to je

$$\max_{\vartheta \in \Theta} L(\vec{x}; \vartheta) = L(\vec{x}; \hat{\vartheta}(\vec{x})),$$

potem funkciji $\hat{\vartheta} : \mathbb{R}^n \rightarrow \bar{\Theta}$ pravimo *cenilka največjega verjetja* (CNV) za ϑ .

Opomba. Imamo parametrični model s parametričnim prostorom Θ .

- Množica “možnih” vrednosti se ne spreminja s ϑ .
- Za “skoraj vse vrednosti” se ne spreminja s ϑ : če je $D_{\hat{\vartheta}}$ definicijsko območje, mora veljati

$$\mathbb{P}((X_1, \dots, X_n) \in D_{\hat{\vartheta}}) = 1$$

za vsako dopustno porazdelitev, $\forall \vartheta$ in $\forall X_i \stackrel{NEP}{\sim} \vartheta$.

Izrek 8.1. Naj bodo gostote (verjetnostne funkcije) $f(\vec{x}; \vartheta)$ *dvakrat zvezno parcialno odvedljive* na ϑ_j , naj bo $\{\vec{x} \mid f(\vec{x}; \vartheta) > 0\}$ *neodvisna* od ϑ in naj veljajo še dodatni blagi regularnostni privzetki. Če imajo (EV) enolične rešitve za vse dovolj velike vzorce, ki jih označimo $\hat{\vartheta} = \hat{\vartheta}(x_1, \dots, x_n)$, potem je $\hat{\vartheta}(x_1, \dots, x_n)$ dosledno zaporedje cenilk za ϑ .

9 INTERVALSKO OCENJEVANJE

Definicija 9.1. Naj bo $c = c(X) = c(F_X)$ proučevana karakteristika slučajne spremenljivke X . Naj bo $\alpha \in (0, 1)$ vnaprej podano (“majhno”) in naj bo n velikost vzorca. Interval zaupanja za c stopnje zaupanja $1 - \alpha$ je prireditev $\vec{X} = (X_1, \dots, X_n) \mapsto [L(X_1, \dots, X_n), U(X_1, \dots, X_n)]$ za katero velja:

$$\mathbb{P}([L(X_1, \dots, X_n), U(X_1, \dots, X_n)] \ni c(F)) \geq 1 - \alpha$$

za vsako dopustno porazdelitev F in $\forall X_1, \dots, X_n \stackrel{NEP}{\sim} F$.

Komentar. Interval zaupanja je določen s funkcijama $L, U : \mathbb{R}^n \rightarrow \mathbb{R}$.

Izrek 9.1 (Student, Fisher). Naj bodo $X_1, \dots, X_n \stackrel{NEP}{\sim} \mathcal{N}(\mu, \sigma^2)$. Tedaj sta

$$\bar{X} = \frac{1}{n}(X_1 + \dots + X_n) \quad \text{in} \quad S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

neodvisni slučajni spremenljivki. Velja

$$\frac{n-1}{\sigma^2} \cdot S^2 = \sum_{i=1}^n \left(\frac{X_i - \bar{X}}{\sigma} \right)^2 \sim \chi_{n-1}^2 = \Gamma\left(\frac{n-1}{2}, \frac{1}{2}\right).$$

Definicija 9.2. Naj bosta Z in K neodvisni slučajni spremenljivki in $Z \sim \mathcal{N}(0, 1)$ ter $K \sim \chi_k^2$. Tedaj pravimo, da ima slučajna spremenljivka $\frac{Z}{\sqrt{\frac{K}{k}}}$ Studentovo porazdelitev s k prostorskimi stopnjami. Označimo jo s t_k .

Komentar. Porazdelitev t_k je zaradi neodvisnosti funkcija porazdelitve slučajne spremenljivke Z in slučajne spremenljivke K . Takoj sledi, da je t_k simetrična okrog 0. Očitno sta $\frac{Z}{\sqrt{\frac{K}{k}}}$ in $\frac{-Z}{\sqrt{\frac{K}{k}}}$ enako porazdeljeni.

Posledica. Slučajna spremenljivka

$$t = \frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1}$$

(tu je $S = \sqrt{S^2}$).

10 PREIZKUŠANJE DOMNEV

Definicija 10.1. Statistična domneva (hipoteza) je izjava, da porazdelitev slučajne spremenljivke X pripada neki podmnožici $\mathcal{H} \subseteq \mathcal{F}$.

Definicija 10.2. Preizkus domneve \mathcal{H} proti alternativni \mathcal{A} za vzorec velikosti n je odločitveno pravilo, ki na podlagi realizacije n neodvisnih replikacij proučevane slučajne spremenljivke odloči, ali \mathcal{H} zavrnemo (in s tem sprejmemo \mathcal{A}) ali je ne zavrnemo (oz. jo “sprejmemo”).

Komentar. Verjetnost napake 1. vrste je funkcija, definirana na domnevi \mathcal{H} . Če je \mathcal{H} sestavljena iz ene porazdelitve (“enostavna”), je to v resnici eno število.