Verjetnost in statistika

Zapiski po predavanjih izr. prof. dr. Jaka Smrekarja
Napisal: Jon Pascal Miklavčič

2024/10/06

Kazalo

Del I

Verjetnost

Poglavje 1

Slučajni vektorji

1.1 Slučajni vektorji

Spomnimo se:

Slučajna spremenljivka na je taka funkcija $X : \Omega \to \mathbb{R}$, na verejetnostnem prostoru $(\Omega, \mathcal{F}, P)$, za katero so množice (praslike):

$$\{\omega \in \Omega : X(\omega) \leq x\}$$

v $\mathcal{F}$, se pravi dogodki za vsak $x \in \mathbb{R}$

Oznaka:

$$\{\omega \in \Omega : X(\omega) \leq x\} \equiv \{\omega \in \Omega \mid X(\omega) \in (-\infty, x]\} \equiv X^{-1}\left((-\infty, x]\right)$$

Posledično je za slučajno spremenljivko X definirana *komulativna porazdelitvena funkcija* $F_X : \mathbb{R} \to [0, 1]$:

$$F_X(x) = P(X \leq x) = P\left(X^{-1}((-\infty, x])\right)$$

Definicija 1. *Slučajni vektor* je taka preslikava $X = (X_1, \ldots, X_n) : \Omega \to \mathbb{R}^n$ na verejetnostnem prostoru $(\Omega, \mathcal{F}, P)$, za katero so množice:

$$\{X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n\} := \{\omega \in \Omega : X_1(\omega) \leq x_1, \ldots, X_n(\omega) \leq x_n\}$$

v $\mathcal{F}$, se pravi dogodki za vse n-terice $x = (x_n, \ldots, x_n) \in \mathbb{R}^n$.

Oznaka:

$$\begin{aligned}\{X_1 \le x_1, X_2 \le x_2, \ldots, X_n \le x_n\} &\equiv \{\omega \in \Omega : X_1(\omega) \le x_1, \ldots, X_n(\omega) \le x_n\} \\ &\equiv \{X_1 \in (-\infty, x_1], \ldots, X_n \in (-\infty, x_n]\} \\ &\equiv \{X \in (-\infty, x_1] \times \ldots \times (-\infty, x_n]\} \\ &\equiv X^{-1}\left((-\infty, x_n] \times \ldots \times (-\infty, x_n]\right)\end{aligned}$$

Definicija 2. *Komulativna porazdelitvena funkcija slučajnega vektorja* je funkcija $F_X : \mathbb{R}^n \to [0,1]$:

$$F_X(x) \equiv F_{(X_1,\ldots,X_n)}(x_1, \ldots, x_n) = P(X_1 \le x_1, \ldots, X_n \le x_n)$$

Lastnosti komulativne porazdelitvene funkcije:

1. $$\begin{aligned}\lim_{x_1 \to -\infty} F_X(x_1, \ldots, x_n) &= \lim_{x_1 \to -\infty} P\Big(X \in (-\infty, x_1] \times \underbrace{\cdots \times (-\infty, x_n]}_{K}\Big) \\ &= \lim_{n \to \infty} P\left(X \in (-\infty, -n] \times K\right) \qquad (\star) \\ &= P\left(\bigcap_{n=1}^{\infty} \{X \in (-\infty, -n] \times K\}\right) \\ &= P\Bigg(X \in \underbrace{\bigcap_{n=1}^{\infty} (-\infty, -n] \times K}_{=\emptyset}\Bigg) \\ &= 0\end{aligned}$$

   Kjer smo v $(\star)$ vrstici uporabili zveznost P oziroma:

   $$\lim_{n \to \infty} P(E_n) = P\left(\bigcup_{n=1}^{\infty} E_n\right) \quad \text{za} \quad E_1 \supseteq E_2 \supseteq E_3 \supseteq \cdots$$

   Ta limita velja, tudi če proti ∞ pošljemo poljuben x_i. Velja še:

   $$\begin{aligned}\lim_{(x_1,\ldots,x_n) \to (\infty,\ldots,\infty)} F_X(x_1, \ldots, x_n) &= \lim_{n \to \infty} P(X \in (-\infty, n] \times \cdots \times (-\infty, n]) \\ &= P\left(\bigcup_{n=1}^{\infty} \{X_1 \le n, \ldots, X_n \le n\}\right) \\ &= P(X \in \mathbb{R}^n) \\ &= P(\Omega) \\ &= 1\end{aligned}$$

2. *Monotnost:* če je $x_i \leq y_i$ za vse $i \in \{1, \ldots, n\}$, potem je:

$$F_X(x) \leq F_X(y)$$

Dokaz. Sledi iz monotonosti verjetnostne preslikave. $\square$

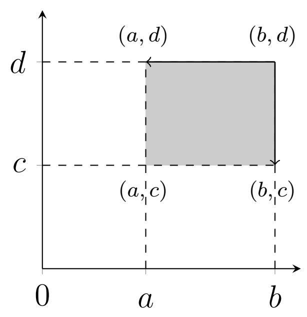
3. *Zveznost z desne:*

$$\lim_{y \downarrow x} F_X(y) = F_X(x)$$

Tukaj $y \downarrow x$ interpretiramo kot $y_i \downarrow x_i$ za vse $i \in \{1, \ldots, n\}$.

Lastnosti 1., 2. in 3. karakterizirajo družino (abstraktnih) komulativnih porazdelitvenih funkcij v primeru slučajne spremenljivke t.j. $n = 1$. V večrazsežnem primeru to ne drži. Poglejmo si $n = 2$. Za $a < b$ in $c < d$ izračunajmo:

$$P((X, Y) \in (a, b] \times (c, d]) = F_{(X,Y)}(b, d) - F_{(X,Y)}(a, d) - F_{(X,Y)}(b, c) + F_{(X,Y)}(a, c)$$



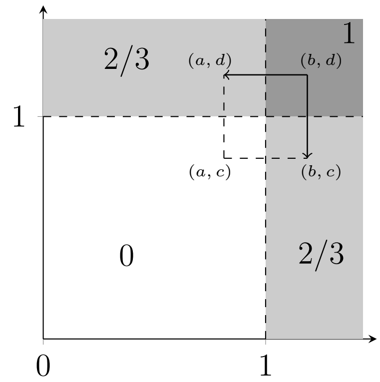
Slika 1.1: $P((X, Y) \in (a, b] \times (c, d])$

Ker je to verjetnost mora veljati:

$$F_{(X,Y)}(b, d) - F_{(X,Y)}(a, d) - F_{(X,Y)}(b, c) + F_{(X,Y)}(a, c) \geq 0 \tag{4}$$

Zgled.

$$F(x, y) = \begin{cases} 1 & ; x \geq 1, y \geq 1 \\ \frac{2}{3} & ; x \geq 1, y \in [0, 1) \\ \frac{2}{3} & ; x \in [0, 1), y \geq 1 \\ 0 & ; \text{sicer} \end{cases}$$



Slika 1.2: Verjetnost $P((X,Y) \in (a,b] \times (c,d])$ za to k.p.f. [1]

Funkcija sicer zadošča lastnostim 1., 2. in 3., ampak za pravokotnik $(a,b] \times (c,d]$, kot na skici velja:

$$F(b,d) - F(a,d) - F(b,c) + F(a,c) = 1 - \frac{2}{3} - \frac{2}{3} + 0 < 0$$

Torej ne more biti komulativna porazdelitvena funkcija. $\Diamond$

Izrek 1. *Če $F : \mathbb{R}^2 \to [0,1]$ zadošča lastnostim 1., 2., 3. in 4. (velja $F(b,d) - F(a,d) - F(b,c) + F(a,c) \geq 0$, za vse četverice $a < b$ in $c < d$), potem je F komulativna porazdelitvena funkcija nekega slučajnega vektorja $(X,Y) : \Omega \to \mathbb{R}^2$.*

Očitna posplošitev velja tudi za $n \geq 3$.

Trditev 1. *Če je $X = (X_1, \ldots, X_n)$ slučajni vektor, so vsi podvektorji tudi slučajni vektorji.*

Dokaz. Na primer, za podvektor $(X_1, \ldots, X_{n-1})$:

$$\{X_1 \leq x_1, \ldots, X_{n-1} \leq x_{n-1}\} = \bigcup_{k=1}^{\infty} \{X_1 \leq x_1, \ldots, X_{n-1} \leq x_{n-1}, X_n \leq k\} \qquad (\star)$$

$\square$

Posebej sledi, da so komponente slučajnih vektorjev, funkcije $X_1, X_2, \ldots, X_n$, slučajne spremenljivke. Iz $(\star)$ je očitno tudi, da komulativne porazdelitvene funkcije podvektorjev dobimo tako, da ustrezne kooridante pošljemo proti ∞:

$$F_{(X_1,\ldots,X_{n-1})}(x_1, \ldots, x_{n-1}) = \lim_{x_n \to \infty} F_{(X_1,\ldots,X_n)}(x_1, \ldots, x_n)$$

Komulativne porazdelitvene funkcije F_{X_i} imenujemo tudi *robne* ali *marginale* komulativne porazdelitvene funckije (glede na $F_{(X_1,\ldots,X_n)}$).

[1] k.p.f. je okrajšava za kumulativno porazdelitveno funkcijo. To se v skripti pojavi še večkrat.

Opomba. Naj bo $X : \Omega \to \mathbb{R}$ s.s.[2] Potem so naslednje množice dogodki:

- $$\begin{aligned}\{X \in (a, b]\} &= \{X \in (-\infty, b]\} \setminus \{X \in (-\infty, a]\} \\ &= \{X \in (-\infty, b]\} \cap \{X \in (-\infty, a]\}^c\end{aligned}$$
- $\{X \in (a, b)\}$, saj je to enako $\bigcup_{n=1}^{\infty} \{X \in (a, c_n]\}$ za $c_n \uparrow b$
- $\{X \in [a, b]\} = \bigcap_{n=1}^{\infty} \left\{X \in \left(a - \frac{1}{n}, b + \frac{1}{n}\right)\right\}$
- $\{X = x\} = \{X \in (x - 1, x]\} \setminus \{X \in (x - 1, x)\}$

Zgornje lahko "kombiniramo" (števne unije, komplementi, preseki). Izkaže se, da obstajajo verjetnosti $P(X \in \mathcal{B})$, kjer je $\mathcal{B} \subseteq \mathbb{R}$ poljubna Borelova množica.

Opomba. Pripomnimo, da podobno velja za slučajni vektor X. Množice:

$$\{X \in \mathcal{B}\} = X^{-1}(\mathcal{B}) \subseteq \Omega$$

so dogodki za Borelove množice $\mathcal{B} \subseteq \mathbb{R}^n$

Zgled (Borelove množice)**.**

1. $\subseteq \mathbb{R}$
   - $(a, b), (-\infty, b), \ldots$ vsi intervali;
   - $\{x\} = (x - 1, x] \setminus (x - 1, x)$ singeltoni;
   - Števne unije Borelovih množic;
   - Števni preseki Borelovih množic;
   - Komplementi Borelovih množic.
2. $\subseteq \mathbb{R}^2$
   - Krogle;
   - Pravokotniki;
   - Enoelementne množice;
   - Unije, preseki, komplemeti kot v $\mathbb{R}$;
   - Če je $G : \mathbb{R}^2 \to \mathbb{R}^2$ zvezno parcialno odvedljiva bijekcija, je $G(\mathcal{B})$ Borelova za vsako Borelovo množico $\mathcal{B}$.

$\Diamond$

[2]s.s. je okrajšava za slučajno spremenljivko. To se v skripti pojavi še večkrat.

1.2 Zvezni slučajni vektorji

Definicija 3. Slučajni vektor $X : \Omega \to \mathbb{R}^n$ ima zvezno gostoto, če obstaja taka "zvezna" funkcija $f_X : \mathbb{R}^n \to [0, \infty)$, da zanjo velja:

$$P(X \in \mathcal{B}) = \int_{\mathcal{B}} f_X(x_1, \ldots, x_n)\, dx_1 \cdots dx_n = \int_{\mathcal{B}} f_X(x)\, dx$$

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

Mislimo si, da gre za posplošeni Reimannov integral. Zvezna funkcija $f_X : \mathbb{R}^n \to [0, \infty)$ je taka, pri kateri je množica točk nezveznosti zanemarljiva za n-terni integral. Če ima slučajni vektor X "zvezno" gostoto, pravimo, da je zvezen. V tem primeru je $F_X : \mathbb{R}^n \to [0,1]$ *zvezna* funkcija.

Zgled. $X \sim \mathrm{U}(a,b)$ (enakomerna zvezna porazdelitev na (a,b)). Vse točke so "enako verjetne". Natančneje za $a \leq c < d \leq b$:

$$P\left(X \in (c,d)\right) = \int_c^d f_{\mathrm{U}(a,b)}(x)\, dx, \quad \text{kjer je} \quad f_{\mathrm{U}(a,b)}(x) = \frac{1}{b-a} \cdot \mathbb{1}_{(a,b)}(x) = \begin{cases} \frac{1}{b-a} & ; x \in (a,b) \\ 0 & ; x \notin (a,b) \end{cases}$$

Slika 1.3: $f_X(x)$ za enakomerno porazdelitev

$\Diamond$

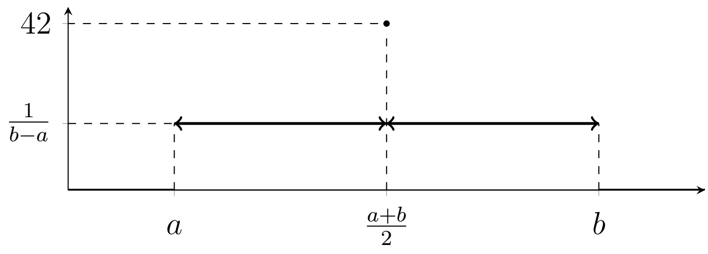
Definicija 4. Za abstraktno množico M in $A \subseteq M$ je $\mathbb{1}_A : M \to \{0,1\}$ *indikator množice* A funkcija:

$$\mathbb{1}_A(m) = \begin{cases} 1 & ; m \in A \\ 0 & ; m \notin A \end{cases}$$

Opomba. Velja: $P(X \in A) = E\left(\mathbb{1}_A(X)\right)$

Zgled. V primeru $X \sim \mathrm{U}(a,b)$ je gostota X tudi:

$$f_X(x) = \begin{cases} \frac{1}{b-a} & ; x \in (a,b) \setminus \left\{\frac{a+b}{2}\right\} \\ 42 & ; x = \frac{a+b}{2} \\ 0 & ; x \notin (a,b) \end{cases}$$



Slika 1.4: $f_X(x)$ za enakomerno porazdelitev

$\Diamond$

Zgled. $D = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$ odprt enotski krog. Pravimo, da $(X,Y) \sim \mathrm{U}(D)$, če ima gostoto:

$$f_{(X,Y)}(x,y) = \frac{1}{\pi} \cdot \mathbb{1}_D(x,y)$$

Množica točk nezveznosti zgornje gostote je enotska krožnica. To je zanemarljiva množica za $\int_{\mathbb{R}^2}$. $\Diamond$

Oglejmo si dvorazsežen primer ($n = 2$):

Za $(X,Y) : \Omega \to \mathbb{R}^2$ z gostoto $f_{(X,Y)} : \mathbb{R}^2 \to [0,\infty)$ velja:

$$\begin{aligned} F_{(X,Y)}(x,y) &= P((X,Y) \in (-\infty,x] \times (-\infty,y]) \\ &= \int_{(-\infty,x]\times(-\infty,y]} f_{(X,Y)}(u,v)\, dudv \\ &= \int_{-\infty}^{x} \int_{-\infty}^{y} f_{(X,Y)}(u,v)\, dvdu \qquad (\star) \\ &= \int_{-\infty}^{y} \int_{-\infty}^{x} f_{(X,Y)}(u,v)\, dudv \end{aligned}$$

Kjer smo v $(\star)$ vrstici uporabili Fubinijev izrek (integral mora biti absolutno konvergenten).

Robni komulativni porazdelitveni funkciji se glasita:

$$
\begin{aligned}
F_X(x) &= \lim_{y\to\infty} F_{(X,Y)}(x,y) \\
&= \lim_{y\to\infty} \int_{-\infty}^{y}\int_{-\infty}^{x} f_{(X,Y)}(u,v)\,dudv \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{x} f_{(X,Y)}(u,v)\,dudv \\
&= \int_{-\infty}^{x}\int_{-\infty}^{\infty} f_{(X,Y)}(u,v)\,dvdu
\end{aligned}
$$

$$
\begin{aligned}
F_Y(y) &= \lim_{x\to\infty} F_{(X,Y)}(x,y) \\
&= \lim_{x\to\infty} \int_{-\infty}^{x}\int_{-\infty}^{y} f_{(X,Y)}(u,v)\,dudv \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{y} f_{(X,Y)}(u,v)\,dudv \\
&= \int_{-\infty}^{y}\int_{-\infty}^{\infty} f_{(X,Y)}(u,v)\,dvdu
\end{aligned}
$$

Od tod sledita robni gostoti:

$$f_X(x) = F_X'(x) = \int_{-\infty}^{\infty} f_{(X,Y)}(x,v)\,dv$$

$$f_Y(y) = F_Y'(y) = \int_{-\infty}^{\infty} f_{(X,Y)}(u,y)\,du$$

V točkah odevljivosti k.p.f. F, oz. za skoraj vse x oz. y.

1.3 Dvofazna normalna porazdelitev

Naj bo $\mu = \begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix} \in \mathbb{R}^2$ in naj bo $\Sigma = \begin{bmatrix}\sigma_1^2 & \sigma_{12}\\ \sigma_{12} & \sigma_2^2\end{bmatrix}$, kjer je $\sigma_1, \sigma_2 \in (0,\infty)$ in $\sigma_{12} \in \mathbb{R}$, simetična pozitivno definitna 2×2 matirka ($\iff$ $\det(\Sigma) > 0$). Tedaj ima slučajni vektor (X,Y) dvorazsežno normalno porazdelitev s parametroma μ in Σ, če ima gostoto:

$$f_{(X,Y)}(x,y) = \frac{1}{2\pi}\frac{1}{\sqrt{\det\Sigma}} \cdot \exp\left(-\frac{1}{2}\left\langle \Sigma^{-1}\left(\begin{bmatrix}x\\ y\end{bmatrix} - \begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix}\right), \begin{bmatrix}x\\ y\end{bmatrix} - \begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix}\right\rangle\right)$$

Gostota $f_{X,Y}$ je zvezna povsod na $\mathbb{R}^2$.

Opomba. Pišemo $(X,Y) \sim \mathrm{N}(\mu,\Sigma) = N\left(\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix}, \begin{bmatrix}\sigma_1^2 & \sigma_{12}\\ \sigma_{12} & \sigma_2^2\end{bmatrix}\right)$.

Seveda je:

$$\Sigma^{-1} = \frac{1}{{\sigma_1}^2{\sigma_2}^2 - \sigma_{12}^2}\begin{bmatrix} \sigma_2^2 & -\sigma_{12} \\ -\sigma_{12} & \sigma_1^2 \end{bmatrix}$$

Tako dobimo: $\left\langle \Sigma^{-1}\left(\begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}\right), \begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}\right\rangle = (\star)$ prepišemo v:

$$\begin{aligned}
\left\langle \Sigma^{-1}\begin{bmatrix} x-\mu_1 \\ y-\mu_2 \end{bmatrix}, \begin{bmatrix} x-\mu_1 \\ y-\mu_2 \end{bmatrix}\right\rangle &= \frac{1}{\sigma_1^2\sigma_2^2 - \sigma_{12}^2}\left\langle \begin{bmatrix} \sigma_2^2 & -\sigma_{12} \\ -\sigma_{12} & \sigma_1^2 \end{bmatrix}\begin{bmatrix} x-\mu_1 \\ y-\mu_2 \end{bmatrix}, \begin{bmatrix} x-\mu_1 \\ y-\mu_2 \end{bmatrix}\right\rangle \\
&= \frac{1}{\sigma_1^2\sigma_2^2 - \sigma_{12}^2}\left\langle \begin{bmatrix} \sigma_2^2\left(x-\mu_1\right) - \sigma_{12}\left(y-\mu_2\right) \\ \sigma_1^2\left(y-\mu_2\right) - \sigma_{12}\left(x\mu_1\right) \end{bmatrix}, \begin{bmatrix} x-\mu_1 \\ y-\mu_2 \end{bmatrix}\right\rangle \\
&= \frac{\sigma_2^2\left(x-\mu_1\right)^2 - 2\sigma_{12}\left(x-\mu_1\right)\left(y-\mu_2\right) + \sigma_1^2\left(y-\mu_2\right)^2}{\sigma_1^2\sigma_2^2 - \sigma_{12}^2}
\end{aligned}$$

Vpeljemo $\rho = \frac{\sigma_{12}}{\sigma_1\sigma_2}$. Ker je $\det\Sigma = {\sigma_1}^2{\sigma_2}^2\left(1-\rho^2\right)$ je $\det\Sigma > 0 \iff \rho^2 < 1$, oz. $\rho \in (-1,1)$. S parametrom ρ lahko nadomestimo σ_{12}. Dobimo:

$$\left\langle \Sigma^{-1}\begin{bmatrix} x-\mu_1 \\ y-\mu_2 \end{bmatrix}, \begin{bmatrix} x-\mu_1 \\ y-\mu_2 \end{bmatrix}\right\rangle = \tfrac{1}{1-\rho^2}\left(\left(\tfrac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\tfrac{x-\mu_1}{\sigma_1}\right)\left(\tfrac{y-\mu_2}{\sigma_2}\right) + \left(\tfrac{y-\mu_2}{\sigma_2}\right)^2\right)$$

Ekvivalentno lahko torej normalno porazdelitev parametriziramo s parametri:

$$\mu_1 \in \mathbb{R}, \mu_2 \in \mathbb{R}, \sigma_1 \in (0,\infty), \sigma_2 \in (0,\infty), \rho \in (-1,1)$$

Dobimo:

$$f_{(X,Y)}(x,y) = \tfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left(-\tfrac{1}{2(1-\rho^2)}\left(\left(\tfrac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\tfrac{x-\mu_1}{\sigma_1}\right)\left(\tfrac{y-\mu_2}{\sigma_2}\right) + \left(\tfrac{y-\mu_2}{\sigma_2}\right)^2\right)\right)$$

Vidimo, da so nivojnice funkcije $f_{(X,Y)}$ krivulje, ki so elipse s središčem v (μ_1,μ_2).

Oglejmo si poseben primer $\mu_1 = \mu_2 = 0, \sigma_1 = \sigma_2 = 1$ (ekvivalentno, obravnavamo gostoto transformacije $U = \frac{X-\mu_1}{\sigma_1}, V = \frac{Y-\mu_2}{\sigma_2}$). Nivojnice se krivulje:

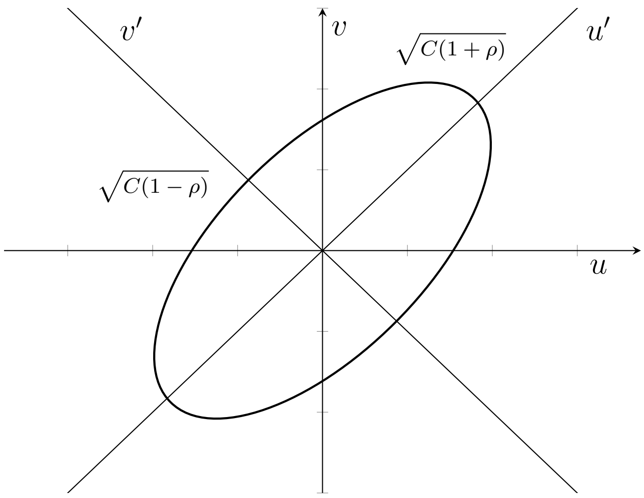
$$u^2 - 2\rho uv + v^2 = C\left(1-\rho^2\right)$$
$$\left\langle \begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix}, \begin{bmatrix} u \\ v \end{bmatrix}\right\rangle = C\left(1-\rho^2\right)$$

To je kvadratna forma oblike $q_A(x) = \langle Ax, x\rangle$. Izračunajmo lastni vrednosti matrike $A = \begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}$:

$$\begin{aligned}\det(A - \lambda I) &= \begin{vmatrix} 1-\lambda & -\rho \\ -\rho & 1-\lambda \end{vmatrix} \\ &= (1-\lambda)^2 - \rho^2 \\ &= (1-\lambda-\rho)(1-\lambda+\rho)\end{aligned}$$

Lastni vrednosti sta $1 \pm \rho$.

$$\lambda_1 = 1-\rho : \begin{bmatrix} \rho & -\rho \\ -\rho & \rho \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies \text{lastni vektor je} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$
$$\lambda_1 = 1+\rho : \begin{bmatrix} -\rho & -\rho \\ -\rho & -\rho \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies \text{lastni vektor je} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$



Slika 1.5: Nivojnica N $\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right)$

V koordinatah $\begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{2} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$ dobimo enačbo:

$$(1-\rho)u'^2 + (1+\rho)v'^2 = C\left(1-\rho^2\right) \iff \frac{u'^2}{C(1+\rho)} + \frac{v'^2}{C(1-\rho)} = 1$$

Naj bo $(X, Y) \sim N\left(\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix}, \begin{bmatrix}\sigma_1^2 & \sigma_{12}\\ \sigma_{12} & \sigma_2^2\end{bmatrix}\right)$. Izkaže se, da za robni porazdelitvi velja:

$$X \sim N\left(\mu_1, {\sigma_1}^2\right) \quad \text{in} \quad Y \sim N\left(\mu_2, \sigma_2^2\right)$$

Nadalje se izkaže, da je $\sigma_{12} = K(X, Y)$. Posledično je ρ t.i. Pearsonov korelacijski koeficient.

Opomba. Pripomnimo, da sledi, da porazdelitev vektorja (X, Y) ni določena z robnima porazdelitvama.

Zgled. Oglejmo si primer $(X, Y) \sim N\left(\begin{bmatrix}0\\ 0\end{bmatrix}, \begin{bmatrix}1 & \rho\\ \rho & 1\end{bmatrix}\right)$. Tedaj se gostota (X, Y) glasi:

$$f_{(X,Y)}(x, y) = \frac{1}{2\pi} \cdot e^{-\frac{1}{2(1-\rho^2)}\left(x^2-2\rho xy+y^2\right)}$$

$\rho = 0 \implies$ standardna dvorazsežna normalna porazdelitev. $\diamondsuit$

Posplošitev:

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

$$f_X(x) = (2\pi)^{-\frac{n}{2}} (\det \Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\left\langle \Sigma^{-1}(x-\mu), x-\mu\right\rangle\right)$$

Poglavje 2

Neodvisnost

Definicija 5. Slučajne spremenljivke $X_1, \ldots, X_n$, so neodvisne, če velja:

$$F_X\left(x_1, \ldots, x_n\right) = F_{X_1}\left(x_1\right) \cdot F_{X_2}\left(x_2\right) \cdots F_{X_n}\left(x_n\right) \tag{$\star$}$$

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

Enakost $(\star)$ lahko prepišemo v:

$$P\left(X \in (-\infty, x_1] \times \ldots \times (-\infty, x_n]\right) = P\left(X_1 \in (-\infty, x_1]\right) \cdots P\left(X_n \in (-\infty, x_n]\right)$$

Izkaže se, da so komponente vektorja X neodvisne natanko tedaj, ko velja:

$$P\left(X_1 \in B_1, X_2 \in B_2, \ldots, X_n \in B_n\right) = P\left(X_1 \in B_1\right) \cdot P\left(X_2 \in B_2\right) \cdots P\left(X_n \in B_n\right)$$

za vse n-terice Borelovih množic $B_i \subseteq \mathbb{R}$.

Torej za slučajni spremenljivki X, Y velja, da sta neodvisni natanko tedaj, ko:

$$P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B)$$

za vse intervale $A, B \subseteq \mathbb{R}$

Spomnimo se, da so dogodki $E_1, E_2, \ldots, E_n$ neodvisni:

$$P\left(E_{i_1} \cap \cdots \cap E_{i_k}\right) = P\left(E_{i_1}\right) \cdots P\left(E_{i_k}\right)$$

za $\forall k \in \{2, 3, \ldots, n\}$ in vsako k-terico $i_1 < i_2 < \ldots < i_k$.

Opomba. Če so $X_1, \ldots, X_n$ neodvisne, so tudi paroma neodvisne, t.j. neodvisni sta X_i in X_j za vsak par $i \neq j$. Obratno ne velja v splošnem.

Trditev 2. *Naj ima slučajni vektor* (X, Y) *"zvezno" gostoto* $f_{(X,Y)} : \mathbb{R}^2 \to [0, \infty)$. X *in* Y *sta neodvisni* $\iff$

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

za skoraj vse realne pare $(x, y) \in \mathbb{R}^2$.

Dokaz. Definicija neodvisnosti pravi:

$$P((X, Y) \in (-\infty, x] \times (-\infty, y]) = P(X \in (-\infty, x])P(Y \in (-\infty, y])$$

Oziroma:

$$\int_{(-\infty,x]\times(-\infty,y]} f_{(X,Y)}(u, v)\, dudv = \int_{(-\infty,x]} f_X(u)\, du \int_{(-\infty,y]} f_Y(v)\, dv$$
$$\int_{-\infty}^{x} \int_{-\infty}^{y} f_{(X,Y)}(u, v)\, dvdu = \int_{-\infty}^{x} f_X(u)\, du \int_{-\infty}^{y} f_Y(v)\, dv$$

($\Longrightarrow$) Uporabimo osnovni izrek analize in odvajamo najprej po x:

$$\int_{-\infty}^{y} f_{(X,Y)}(x, y)\, dv = f_X(x) \int_{-\infty}^{y} f_Y(v)\, dv$$

Odvajati smemo skoraj v vseh točka $x \in \mathbb{R}$, saj gostote niso vedno povsod zvezne. Odvjamo še po y:

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

Velja za skoraj vse x in y.[3]

($\Longleftarrow$) Če veja $f_{(X,Y)}(u, v) = f_X(u) f_Y(v)$ za skoraj vse u, v veja tudi:

$$\int_{(-\infty,x]\times(-\infty,y]} f_{(X,Y)}(u, v)\, dudv = \int_{(-\infty,x]\times(-\infty,y]} f_X(u) \cdot f_Y(v)\, dudv$$

za vse[4] x in y. Leva starn je $F_{(X,Y)}(x, y)$, na desni strani pa dvojni integral spremeninmo v dvakratnega:

$$\begin{aligned}
\int_{(-\infty,x]\times(-\infty,y]} f_X(u) \cdot f_Y(v)\, dudv &= \int_{(-\infty,x]} f_X(u) \int_{(-\infty,y]} f_Y(v)\, dvdu \\
&= \int_{(-\infty,x]} f_X(u)\, du \int_{(-\infty,y]} f_Y(v)\, dv \\
&= F_X(x) \cdot F_Y(y)
\end{aligned}$$

□

[3]Enakost za skoraj vse x in y pomeni enakost za vse integrale po Borelovih množicah.

[4]Če sta funkciji skoraj povsod enaki je integral funkcij enak povsod.

Posledica 1. *Če je* $(X,Y) \sim N\left(\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix}, \begin{bmatrix}\sigma_1^2 & \sigma_{12}\\ \sigma_{12} & \sigma_2^2\end{bmatrix}\right)$ *potem sta* X *in* Y *neodvisni* $\iff \sigma_{12} \iff \rho = 0$.

Dokaz. ($\Longrightarrow$) Spomnimo se, da ima (X,Y) gostoto:

$$f_{(X,Y)}(x,y) = \frac{1}{2\pi}\cdot\frac{1}{\sqrt{\det\Sigma}}\cdot e^{-\frac{1}{2}\left\langle \Sigma^{-1}(\vec{x}-\vec{\mu}),\vec{x}-\vec{\mu}\right\rangle}$$

Če za f_X f_Y izberemo "standardne" funkcije zaradi njihove zveznosti, dobimo neodvisnost natanko tedaj, ko velja:

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

za vse pare x,y, kjer sta:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2} \qquad \text{in} \qquad f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{1}{2}\left(\frac{y-\mu_2}{\sigma_2}\right)^2}$$

To pomeni:

$$\frac{1}{2\pi}\cdot\frac{1}{\sqrt{\det\Sigma}}\cdot e^{-\frac{1}{2}\left\langle \Sigma^{-1}(\vec{x}-\vec{\mu}),\vec{x}-\vec{\mu}\right\rangle} = \frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2}\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{1}{2}\left(\frac{y-\mu_2}{\sigma_2}\right)^2}$$

Če privzamemo veljavnost zgornjega razcepa in vstavimo $(x,y) = (\mu_1,\mu_2)$ dobimo:

$$\frac{1}{2\pi}\frac{1}{\sigma_1\sigma_2\sqrt{1-\rho^2}} = \frac{1}{\sqrt{2\pi}\sigma_1}\frac{1}{\sqrt{2\pi}\sigma_2} \implies 1-\rho^2 = 1 \implies \rho = 0$$

($\Longleftarrow$) Očitno. Vstavimo $\rho = 0$. □

Opomba. Za $\mu \in \mathbb{R}^n$ in simetrično pozitivno definitno matriko $\Sigma \in \mathbb{R}^{n\times n}$ je slučajni vektor $X : \Omega \to \mathbb{R}^n$ porazdeljen normalno s parametroma μ in Σ ($X \sim N(\mu,\Sigma)$), če ima gostoto:

$$f_X(x) = (2\pi)^{-\frac{n}{2}}(\det\Sigma)^{-\frac{1}{2}}e^{-\frac{1}{2}\left\langle\Sigma^{-1}(x-\mu),x-\mu\right\rangle}$$

Izkaže se, da sledi:

$$X_i \sim N(\mu_i,\Sigma_{ii}) \quad \text{in da je} \quad \Sigma_{ij} = K(X_i,X_j) \quad \text{za} \quad i \neq j.$$

Dalje sledi, da so $X_1,\ldots,X_n$ neodvisne $\iff$ Σ diagonalna. Posledično so komponente večrazsežno normalno porazdeljenega slučajnega vektorja neodvisne natanko tedaj, ko so paroma neodvisne.

Posledica 2. *Naj ima slučajni vektor (X,Y) "zvezno" gostoto $f_{X,Y} : \mathbb{R}^2 \to [0,\infty)$. X in Y neodvisni $\iff$*

$$f_{(X,Y)}(x,y) = \phi(x) \cdot \psi(y)$$

za skoraj vse pare x in y in neki nenegativni integrabilni funkciji ϕ in ψ.

Dokaz. ($\implies$) Že vemo.

($\impliedby$) Iz razcepa $f_{(X,Y)}(x,y) = \phi(x) \cdot \psi(y)$ sledi:

$$F_{(X,Y)}(x,y) = \int_{-\infty}^{x} \phi(u)\,du \int_{-\infty}^{y} \psi(v)\,dv$$

Zato:

$$\lim_{y\to\infty} F_{(X,Y)}(x,y) = \lim_{y\to\infty} \int_{-\infty}^{x} \phi(u)\,du \int_{-\infty}^{y} \psi(v)\,dv$$
$$F_X(x) = \int_{-\infty}^{x} \phi(u)\,du \int_{-\infty}^{\infty} \psi(v)\,dv$$

Posledično je po osnovnem izreku analize:

$$f_X(x) = \phi(x) \cdot \int_{-\infty}^{\infty} \psi(v)\,dv$$

Simetrično je:

$$f_Y(y) = \int_{-\infty}^{\infty} \phi(v)\,dv \cdot \psi(y)$$

Ker je $\int_{\mathbb{R}^2} f_{(X,Y)}(u,v)\,dudv = 1$ sledi:

$$\int_{-\infty}^{\infty} \phi(u)\,du \int_{-\infty}^{\infty} \psi(v)\,dv = 1$$

Sledi:

$$\phi(x) \cdot \psi(y) = \phi(x) \int_{-\infty}^{\infty} \psi(v)\,dv \int_{-\infty}^{\infty} \phi(u)\,du\,\psi(y) = f_X(x) \cdot f_Y(y)$$

□

Poglavje 3

Funkcije slučajnih spremenljivk in slučajnih vektorjev

3.1 Slučajne spremenljivke

Naj bo X slučajna spremenljivka in $g : \mathbb{R} \to \mathbb{R}$ zvezna bijekcija. Potem je $Y = g \circ X : \Omega \to \mathbb{R}$ tudi slučajna spremenljivka, saj je:

$$\begin{aligned}
\{Y \leq y\} &\equiv \{\omega \in \Omega : g(X(\omega)) \leq y\} \\
&= \{\omega \in \Omega : g(X(\omega)) \in (-\infty, y]\} \\
&= \{\omega \in \Omega : X(\omega) \in g^{-1}\left((-\infty, y]\right)\} \\
&= \{X \in \underbrace{g^{-1}((-\infty, y])}_{\text{Borelova množica}}\} \in \mathcal{F}
\end{aligned}$$

Torej je $\{Y \leq y\}$ dogodek. Pravimo, da je Y *funkcija slučajne spremenljivke* X. Vprašanje: kako je porazdelitveno funkcijo za Y ponazoriti s pomočjo g in porazdelitvene funkcije za X?

Vzemimo, da je $g : (a, b) \to (c, d)$ strogo naraščujoča zvezna bijekcija, z zalogo vrednosti (a, b) za $-\infty \leq a < b \leq \infty$, torej $P(X \in (a, b)) = 1$. Vzemimo, da je f strogo naraščujoča, zvezna funkcija, z zalogo vrednosti (a, b). Za $y \in (a, b)$ je:

$$\begin{aligned}
F_Y(y) = P(g \circ X \leq y) &= P\left(X \leq g^{-1}(y)\right) \\
&= F_X\left(g^{-1}(y)\right)
\end{aligned}$$

Za $y \leq a$ je $F_Y(y) = 0 = P(Y \leq y)$ (nemogoč dogodek), za $y \geq b$ pa $F_Y(y) = 1$.

Poglejmo si najprej diskreten primer. Če je X diskretna slučajna spremenljivka z vrednostmi $\{x_i\}_{i \in I}$ in je g funkcija, ki preslika množico $\{x_i\}_{i \in I}$ na množico $\{y_i\}_{i \in J}$, je $f \circ X$ diskretna slučajna spremenljivka z verjetnostno funkcijo:

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

Zgled.

$$X \sim \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}, \quad g(x) = x^2$$

Potem:

$$X^2 \sim \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

◊

Poglejmo si zdaj še zvezen primer. Če je X slučajna spremenljivka z zvezno gostoto f_X, ki je od 0 različna na interavlu (a, b) in $g : (a, b) \to (c, d)$ zvezna bijekcija, potem je tudi Y zvezno porazdeljena slučajna spremenljivka in velja:

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

Če je f odvedljiva, sledi:

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

Podobno ravnamo v primeru, ko je f strogo padajoča funkcija:

$$\begin{aligned} F_{g(X)}(y) &= P(g(X) \leq y) \\ &= P\left(X \geq g^{-1}(y)\right) \\ &= 1 - P\left(X < g^{-1}(y)\right) \\ &= 1 - F_X\left(g^{-1}(y)-\right) \end{aligned}$$

Nadaljujemo podobno kot prej. Za splošno odvedljivo bijekcijo $f : (a, b) \to (c, d)$ potem velja transformacijska formula:

$$f_{g(X)}(y) = \frac{f_X\left(g^{-1}(y)\right)}{\left|g'\left(g^{-1}(y)\right)\right|} = p_X\left(g^{-1}(y)\right) \cdot \left(g^{-1}(y)\right)', \quad y \in (c, d)$$

Če pa funkcija ni povsod naraščujoča oz. padajoča, potem njeno definicijsko območje razdelimo na intervale monotnosti in obravnavamo vsakega posebaj.

Zgled. Naj bo $X \sim \mathrm{N}(0,1)$, $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ in $\alpha, \beta \in \mathbb{R}, \alpha \neq 0$. Kakšna je porazdelitev $\alpha X + \beta$?

Gre za $g(X)$, kjer je $g(x) = \alpha x + \beta$. To je zvezno odvedljiva bijekcija $g : \mathbb{R} \to \mathbb{R}$, za katero $g^{-1}(z) = \frac{z-\beta}{\alpha}$ in $g'(x) = \alpha$. Transformacijska formula nam da:

$$f_{aX+b}(z) = \frac{1}{|\alpha|} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{z-\beta}{\alpha}\right)^2}$$

Torej $\alpha X + \beta \sim \mathrm{N}\left(\beta, \alpha^2\right)$ ♢

Zgled. Oglejmo si primer nemonotone funkcije g. Naj bo $g : \mathbb{R} \to [0, \infty), g(x) = x^2$.

Ideja: $P\left(X^2 \in (c,d)\right)$ bomo izpeljali, kot $\int_c^d \ldots$ in s tem prepozanali gostoto slučajne spremenljivke X^2.

Dovolj je obravnavati primer $0 < c < d < \infty$:

$$\begin{aligned}
P\Big(X^2 \in (c,d)\Big) &= P\Big(c < X^2 < d\Big) \\
&= P(\sqrt{c} < |X| < \sqrt{d}) \\
&= P(\sqrt{c} < X < \sqrt{d}) + P(-\sqrt{d} < X < -\sqrt{c}) \\
&= \underbrace{\int_{\sqrt{c}}^{\sqrt{d}} f_X(x)\, dx}_{z=x^2,\, x=\sqrt{z}} + \underbrace{\int_{-\sqrt{d}}^{\sqrt{c}} f_X(x)\, dx}_{z=x^2,\, x=-\sqrt{z}} \\
&= \int_c^d f_X(\sqrt{z}) \frac{1}{2\sqrt{z}}\, dz + \int_d^c f_X(-\sqrt{z}) \frac{-1}{2\sqrt{z}}\, dz \\
&= \int_c^d \frac{f_X(\sqrt{z}) + f_X(-\sqrt{z})}{2\sqrt{z}}\, dz
\end{aligned}$$

Torej:

$$f_{X^2}(z) = \frac{f_X(\sqrt{z}) + f_X(-\sqrt{z})}{2\sqrt{z}}$$

♢

Zgled. Naj bo $X \sim \mathrm{N}(0,1)$ in $Y = X^2$ (mali g je torej kvadriranje). Za $y < 0$ je $F_Y(y) = 0 = P(Y \leq y)$ nemogoč dogodek, saj $Y \geq 0$. Za $y \geq 0$ uporabimo zgornjo formulo in dobimo:

$$f_{X^2}(y) = f_Y(y) = \frac{f_X(\sqrt{y}) + f_X(-\sqrt{y})}{2\sqrt{y}} = \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} + \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}}}{2\sqrt{y}} = \frac{1}{\sqrt{2\pi}} y^{-\frac{1}{2}} \cdot e^{-\frac{y}{2}}$$

Torej je:

$$X^2 \sim \mathrm{Gama}\left(\frac{1}{2}, \frac{1}{2}\right) = \chi_1^2$$

♢

3.2 Slučajni vektorji

Trditev 3. *Naj bo* $X : \Omega \to \mathbb{R}^n$ *slučajni vektor in* $h : \mathbb{R}^n \to \mathbb{R}$ *zvezna funkcija. Tedaj je* $h(X) = h(X_1, X_2, \ldots, X_n)$ *slučajna spremenljivka.*

Dokaz.

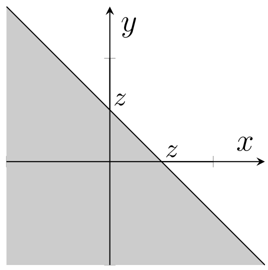
$$\begin{aligned}\{h(X_1, \ldots, X_n) \le z\} &= \{h(X) \in (-\infty, z]\} \\ &= \left\{X \in h^{-1}((-\infty, z])\right\} \\ &= \left\{\omega \in \Omega \mid X(\omega) \in h^{-1}((-\infty, z])\right\}\end{aligned}$$

Tu je $h^{-1}((-\infty, z]) = \{x \in \mathbb{R}^n \mid h(x) \in (-\infty, z]\}$ praslika. Ker je h zvezna in je $(-\infty, z]$ zaprta v $\mathbb{R}$ je $h^{-1}((-\infty, z])$ zaprta in zato Borelova. □

Zgled. Naj ima (X, Y) "zvezno" gostoto $f_{(X,Y)}$. Zanima nas porazdelitev $X + Y$ (tu interpretiramo $X + Y = h(X, Y)$ za $h : \mathbb{R}^2 \to \mathbb{R}$, $h(x, y) = x + y$)[5]. Izračunajmo k.p.f. slučajne spremenljivke $X + Y$:

$$\begin{aligned}P(X + Y \le z) &= P\left((X, Y) \in h^{-1}((-\infty, z])\right) \\ &= P((X, Y) \in \{(x, y) \mid x + y \le z\}) \\ &= \int_{x+y\le z} f_{(X,Y)}(x, y)\, dxdy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{z-y} f_{(X,Y)}(x, y)\, dxdy \qquad (\star) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{z} f_{(X,Y)}(u - y, y)\, dudy \\ &= \int_{-\infty}^{z} \int_{-\infty}^{\infty} f_{(X,Y)}(u - y, y)\, dydu\end{aligned}$$

Kjer smo v $(\star)$ vrstici vpeljali novo spremenljivko $u = x + y$.



Slika 3.1: $\{(x, y) \mid x + y \le z\}$

[5] To ni bijekcija; preslikava ni injektivna, saj za $(2, 1)$ in $(1, 2)$ dobimo isti rezultat, je pa surjektivna. Zato praslika vrne vse pare (x, y), za katere velja, da so enaki neki vrednosti $h(x, y)$. Recimo $h(x, y) = z$ in $z = 3$. Praslika potem vrne vse pare (x, y), da je $x + y = 3$. V našem primeru vrne vse pare (x, y), ki so $\le z$

Sledi:

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_{(X,Y)}(z-y, y)\, dy$$

Če sta X in Y neodvisni dobimo:

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(z-y) \cdot f_Y(y)\, dy$$

Slednji pravimo *konvolucija* funkcij f_X in f_Y. $\Diamond$

Trditev 4. *Naj bosta X in Y neodvisni slučajni spremenljivki. $X \sim \chi^2(m)$ in $Y \sim \chi^2(n)$. Potem je:*

$$X + Y \sim \chi^2(m+n)$$

Dokaz. Vzemimo $X \sim \chi^2(m)$ in $Y \sim \chi^2(n)$. Izračunajmo $X+Y$. Ta je nenegativna (saj je χ^2 negativna) in zato $f_{X+Y}(z) = 0$ za $z \leq 0$. Za $z > 0$ pa je:

$$\begin{aligned}
f_{X+Y}(z) &= \frac{1}{2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right) \cdot 2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} \int_0^z x^{\frac{m}{2}-1} e^{-\frac{x}{2}} (z-x)^{\frac{n}{2}-1} e^{-\frac{z-x}{2}}\, dx \\
&= \frac{1}{2^{\frac{m+n}{2}} \cdot \Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} e^{-\frac{z}{2}} \int_0^z x^{\frac{m}{2}-1}(z-x)^{\frac{n}{2}-1}\, dx \qquad (\star) \\
&= \frac{1}{2^{\frac{m+n}{2}} \cdot \Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} e^{-\frac{z}{2}} z^{\frac{m+n}{2}-1} \underbrace{\int_0^1 t^{\frac{m}{2}-1} \cdot (1-t)^{\frac{n}{2}-1}\, dt}_{\beta\left(\frac{m}{2},\frac{n}{2}\right) = \frac{\Gamma\left(\frac{m}{2}\right)\cdot\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{m}{2}+\frac{n}{2}\right)}} \\
&= \frac{1}{2^{\frac{m+n}{2}}\Gamma\left(\frac{m+n}{2}\right)} z^{\frac{m+n}{2}-1} e^{-\frac{z}{2}}
\end{aligned}$$

Kjer smo v $(\star)$ vrstici vpeljali novo spremenljivko $tz = x$. Torej je:

$$X + Y \sim \chi^2(m+n)$$

$\square$

Posledica 3. *Naj bodo neodvisne slučajne spremenljivke $X_1, X_2, \ldots, X_m \sim \mathrm{N}(0,1)$ porazdeljene standardno normalno. Potem je $Y = X_1^2 + X_2^2 + \ldots + X_n^2 \sim \chi^2(n)$.*

Dokaz. Vemo že, da je $X_i^2 \sim \chi^2(1)$. Posledica sledi takoj iz trdive, ter trditve da so kvadrati neodvisnih slučajnih spremenljivk med seboj neodvisni. $\square$

V splošnem porazdelitev slučajne spremenljivke $h(X_1, \ldots, X_n)$ najpreprosteje dobimo z dopolnituijo funkcije $h : \mathbb{R}^n \to \mathbb{R}$ do zvezno diferenciabilne preslikave $g : \mathbb{R}^n \to \mathbb{R}^n$ (na primer $h(x,y) = x+y$ lahko dopolnimo do $g(x,y) = (x+y, y)$) in izračunamo porazdelitev transformacije $g(X)$.

Trditev 5. *Naj bo $X : \Omega \to \mathbb{R}^n$ slučajni vektor z "zvezno" gostoto $f_X : \mathbb{R}^n \to [0, \infty)$. Dalje naj bo $g : \mathbb{R}^n \to \mathbb{R}^n$ (ali $g : D \to E$ za primerni množici $D, E \subseteq \mathbb{R}^n$, kjer $P(X \in D) = 1$) zvezno diferenciabilna bijekcija. Tedaj ima slučajni vektor $g(X) : \Omega \to \mathbb{R}^n$ gostoto:*

$$f_{g(X)}(z) = f\left(g^{-1}(z)\right) |\det J_{g^{-1}}(z)| = \frac{f\left(g^{-1}(z)\right)}{|\det J_g\left(g^{-1}(z)\right)|}$$

za vse $z \in \mathbb{R}^n$ (oziroma $z \in E$).

Dokaz. Rdeča nit: $P(g(X) \in X)$ želimo izraziti kot $\int_C \ldots$, kjer bo $\ldots$ gostota:

$$\begin{aligned} P(g(X) \in C) &= P\left(X \in g^{-1}(C)\right) \\ &= \int_{g^{-1}(C)} f_X(x)\, dx \qquad (\star) \\ &= \int_C f_X\left(g^{-1}(z)\right) |\det J_{g^{-1}}(z)|\; dz \end{aligned}$$

Kjer smo v $(\star)$ vrstici vpeljali novo spremenljivko $x = g^{-1}(z)$, $dx = |\det J_{g^{-1}}(z)| dz$. □

Zapišimo dvorazsežen primer v "tradicionalnih" oznakah:

Transformacijo $g(X, Y)$ zapišemo kot:

$$(U, V) = (U(X, Y), V(X, Y))$$

Njen inverz $g^{-1}(U, V)$ pa kot:

$$(X, Y) = (X(U, V), Y(U, V))$$

Tedaj:

$$|\det J_{g^{-1}}(u, v)| = \left| \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}_{(u,v)} \right|$$

in

$$|\det J_g(g^{-1}(u, v))| = \left| \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}_{(x(u,v), y(u,v))} \right|$$

Transformacijska formula se tedaj glasi:

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

Upoštevaje $g \circ g^{-1} = \mathrm{id}$ je posledično $J_g\left(g^{-1}(u,v)\right) \cdot J_{g^{-1}}(u,v) = I$. Zgornje lahko prepišemo v:

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

Zgled.

1. Oglejmo si transformacijo:

$$U = X + Y, V = Y \implies Y = V, X = U - V$$

   To pomeni $g(x,y) = (x+y, y)$ in $g^{-1}(u,v) = (u-v, v)$.

$$\det J_g(x,y) = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1$$

   Torej:

$$f_{(X+Y,Y)}(u,v) = \frac{f_{(X,Y)}(u-v,v)}{1}$$

   Gostota slučajne spremenljivke $X+Y$ je robna gostota oziroma:

$$f_{X+Y}(u) = \int_{-\infty}^{\infty} f_{(X,Y)}(u-v,v)\,dv$$

2. Oglejmo si še transformacijo:

$$U = \frac{X}{Y}, V = Y \implies X = UV, Y = V$$

   Pripomnimo, da je transformacija $g(x,y) = \left(\frac{x}{y}, y\right)$ definirana le na $\{Y \neq 0\} = \mathbb{R} \times \mathbb{R} \backslash \{0\}$. Razumemo lahko $g : \mathbb{R} \times \mathbb{R} \backslash \{0\} \to \mathbb{R} \times \mathbb{R} \backslash \{0\}$. Za "zvezen" slučajni vektor (X,Y) seveda velja $P(Y=0) = 0$ oziroma $P((X,Y) \in \mathbb{R} \times \mathbb{R} \backslash \{0\}) = 1$. Po transformacijski formuli velja:

$$f_{\left(\frac{X}{Y},Y\right)}(u,v) = \frac{f_{(X,Y)}(uv,v)}{\left| \begin{vmatrix} \frac{1}{y} & * \\ 0 & 1 \end{vmatrix}_{y=y(u,v)} \right|} = |v| f_{(X,Y)}(uv,v)$$

   in:

$$f_{\frac{X}{Y}}(u) = \int_{-\infty}^{\infty} |v| f_{(X,Y)}(uv,v)\,dv$$

$\Diamond$

Trditev 6. *Naj za* $X : \Omega \to \mathbb{R}^n$ *velja* $X \sim \mathrm{N}(\mu, \Sigma)$ $(\mu \in \mathbb{R}^n, \Sigma \in \mathbb{R}^{n\times n}$ *s.p.d.) in naj bo* $A : \mathbb{R}^n \to \mathbb{R}^n$ *obrljiva matrika. Dalje naj bo* $\nu \in \mathbb{R}^n$. *Tedaj:*

$$AX + \nu \sim \mathrm{N}\left(A\mu + \nu, A\Sigma A^\top\right)$$

Dokaz. Imamo $g(x) = Ax + \nu$ in $g^{-1}(z) = A^{-1}(z - \nu)$. g je "linearna" (afina) preslikava $\implies J_g(x) \equiv A$. Transformacijska formula pove:

$$\begin{aligned} f_{AX+\nu}(z) &= \frac{1}{|\det A|} f_X\left(A^{-1}(z-\nu)\right) \\ &= \frac{1}{\sqrt{|\det A|^2}} (2\pi)^{-\frac{n}{2}} (\det \Sigma)^{-\frac{1}{2}} e^{-\frac{1}{2}\left\langle \Sigma^{-1}\left(A^{-1}(z-\nu)-\mu\right), A^{-1}(z-\nu)-\mu\right\rangle} && (\star) \\ &= (2\pi)^{-\frac{n}{2}} \left(\det\left(A\Sigma A^\top\right)\right)^{-\frac{1}{2}} e^{-\frac{1}{2}\left\langle \left(A^\top\right)^{-1}\Sigma^{-1}A^{-1}(z-\nu-A\mu), z-\nu-A\mu\right\rangle} \end{aligned}$$

V vrstici $(\star)$ velja:

$$\begin{aligned} |\det A| &= \sqrt{|\det A|^2} = \sqrt{\mathrm{del}\, A \cdot \det A^\top} = \sqrt{\det\left(AA^\top\right)} \\ A^{-1}(z-\nu) - \mu &= A^{-1}\left((z-\nu) - A\mu\right) \end{aligned}$$

□

Posledica 4. *Naj bo* $X \sim N(\mu, \Sigma)$. *Razcepimo* $\Sigma = LL^\top$ *(npr. Cholesky). Sledi, da je* $X \sim LZ + \mu$, *kjer je* $Z \sim N\left(0, I_{n\times n}\right)$ *n-razsežna standardna normalna porazdelitev. To je klučnega pomena za simulacijo vzorčenja.*[6]

[6]To pomeni, da lahko do katere koli normalne porazdelitve pridemo tako, da naredimo afino preslikavo s spodnje trikotno matriko in nekim vektorjem na standarni normalni porazdelitvi.

Poglavje 4

Matematično upanje oz. pričakovna vrednost

Definicija 6. Naj bo $X : \Omega \to \mathbb{R}$ slučajna spremenljivka z "zvezno" gostoto f_X. Tedaj je *pričakovana vrednost* slučajne spremenljivke X definirana kot:

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

za zvezne porazdelitve, če integral absolutno konvergira, t.j. $\int_{-\infty}^{\infty} |x| f_X(x)\,dx < \infty$ in:

$$E(X) = \sum_i x_i P\left(X = x_i\right)$$

za diskretne porazdelitve, če vrsta absolutno konvergira, t.j. $\sum_i |x_i| P\left(X = x_i\right) < \infty$.

Če je $X : \Omega \to \mathbb{R}^n$ slučajni vektor, definiramo pričakovano vrednost "po komponentah":

$$E\left(\begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}\right) = \begin{bmatrix} E(X_1) \\ \vdots \\ E(X_n) \end{bmatrix}$$

Podobno velja za slučajno matriko.

Zgled.

- Če je X porazdeljena Cauchyjevo $E(X)$ ne obstaja. Če sta X in Y neodvisni in porazdeljeni standardno normalno, t.j. $X, Y \sim \mathrm{N}(0,1)$ je $\frac{X}{Y}$ porazdeljena Cauchyjevo.

- $X \sim \mathrm{N}\left(\mu, \sigma^2\right) \implies E(X) = \mu$

- $X \sim \mathrm{U}(a,b) \implies E(X) = \frac{a+b}{2}$

$\Diamond$

Trditev 7. *Naj bo X slučajna spremenljivka z "zvezno" gostoto f_X in naj bo $g : \mathbb{R} \to \mathbb{R}$ (ali $g : (a,b) \to (c,d)$) zvezna funkcija. Tedaj je:*

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

če obstaja.

Dokaz. Za primer, ko je $X : \Omega \to (a,b)$ (t.j. $P(X \in (a,b)) = 1$) in je $g : (a,b) \to (c,d)$ zvezno diferenciabilna bijekcija. Vemo, da v tem primeru velja:

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

Po definiciji je tedaj:

$$\begin{aligned} E(g(X)) &= \int_{(c,d)} f_X\left(g^{-1}(z)\right) \left|\left(g^{-1}\right)'(z)\right| dz \qquad (\star) \\ &= \int_{(a,b)} g(x) f_X(x)\, dx \\ &= \int_a^b g(x) f_X(x)\, dx \qquad \text{za} \quad a < b \end{aligned}$$

Kjer smo v $(\star)$ vrstici vpeljali novo spremenljivko $x = g^{-1}(z)$. $\square$

Zgled. Posebej izračunajmo $E(|X|)$ za "zvezno" slučajno spremenljivko X. Najprej izračunajmo njeno komulativno porazdelitveno funkcijo:

$$\begin{aligned} & F_{|X|}(z) = P(|X| \le z) = P(-z \le X \le z) = F_X(z) - F_X(-z) \\ \implies & f_{|X|}(z) = f_X(z) + f_X(-z) \qquad \text{za} \quad z > 0 \end{aligned}$$

Sledi:

$$\begin{aligned} E(|X|) &= \int_0^{\infty} z\left(f_X(z) + f_X(-z)\right) dz \\ &= \int_0^{\infty} x f_X(x)\, dx + \int_0^{-\infty} (-x) f_X(x)\, (-dx) \\ &= \int_{-\infty}^{\infty} |x| f_X(x)\, dx \end{aligned}$$

$\Diamond$

Posledica 5. *Za $p \neq 0$ velja:*

$$E\left(|X|^p\right) = \int_0^{\infty} z^p \left(f_X(z) + f_X(-z)\right) dz = \int_{-\infty}^{\infty} |x|^p f_X(x)\, dx$$

Dokaz. V prvi enakosti smo uporabili trditev in $F_{|X|}(z^p) = P(|X|^p \le z^p)$ v drugi enakosti pa račun zgoraj. $\square$

Trditev 8. *Naj ima slučajni vektor (X,Y) "zvezno" gostoto $f_{(X,Y)}$ in naj bo $h : \mathbb{R}^2 \to \mathbb{R}$ (ali $D \to \mathbb{R}$ za $P((X,Y) \in D) = 1$) zvezna funkcija. Tedaj je:*

$$E(h(X,Y)) = \int_{\mathbb{R}^2} h(x,y) f_{(X,Y)}(x,y)\, dxdy$$

Dokaz. Za primer, ko je h mogoče dopolniti do zveno diferenciabilne bijekcije $g : D \to E$ kot $g(x,y) = (h(x,y), k(x,y))$, kjer sta D in E odprti množici v $\mathbb{R}^2$. Tedaj je:

$$f_{h(X,Y)}(u) = \int_{-\infty}^{\infty} f_{g(X,Y)}(u,v)\, dv = \int_{-\infty}^{\infty} f_{(X,Y)}\left(g^{-1}(u,v)\right) |\det J_{g^{-1}}(u,v)|\, dv$$

in

$$\begin{aligned} E(h(X,Y)) &= \int_{-\infty}^{\infty} u f_{h(X,Y)}(u)\, du \\ &= \int_{\mathbb{R}^2} u f_{(X,Y)}\left(g^{-1}(u,v)\right) |\det J_{g^{-1}}(u,v)|\, dudv \qquad (\star) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y) f_{(X,Y)}(x,y)\, dx\, dy \end{aligned}$$

Kjer smo v $(\star)$ vrstici naredili substitucijo $(x,y) = g^{-1}(u,v)$, $(u,v) = g(x,y)$ in $dxdy = |\det J_{g^{-1}}(u,v)|\, dudv$. $\square$

Zgled.

- Poglejmo si $h(x,y) = x + y$:

$$\begin{aligned} E(X+Y) &= \int_{\mathbb{R}^2} (x+y) f_{(X,Y)}(x,y)\, dxdy \\ &= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f_{(X,Y)}(x,y)\, dy\, dx + \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f_{(X,Y)}(x,y)\, dx\, dy \\ &= \int_{-\infty}^{\infty} x f_X(x)\, dx + \int_{-\infty}^{\infty} y f_Y(y)\, dy \\ &= E(X) + E(Y) \end{aligned}$$

  Dokazali smo *aditivnost* pričakovane vrednosti.

- Poglejmo si še $h(x,y) = xy$:

$$E(XY) = \int_{\mathbb{R}^2} xy f_{(X,Y)}(x,y)\, dxdy$$

  Če integral absolutno konvergira.

$\Diamond$

Posledica 6. *Če sta X in Y neodvisni slučajni spremenljivki, za kateri obstajata $E(X)$ in $E(Y)$, obstaja tudi $E(XY)$ in velja:*

$$E(XY) = E(X) \cdot E(Y)$$

Dokaz.

$$\begin{aligned} E(|XY|) &= \iint_{\mathbb{R}^2} |xy| f_{(X,Y)}(x,y)\, dxdy \\ &= \iint_{\mathbb{R}^2} |x||y| f_X(x) f_Y(y)\, dxdy \\ &= \int_{\mathbb{R}} |x| f_X(x)\, dx \int_{\mathbb{R}} |y| \cdot f_Y(y)\, dy < \infty \end{aligned}$$

Pokazali smo obstoj. Enak račun pokaže $E(XY) = E(X) \cdot E(Y)$. $\square$

Posledica 7. *Če je $X \leq Y$ (z verjetnostjo 1) velja $E(X) \leq E(Y)$, če obstaja $E(Y)$.*

Dokaz. Iz aditivnosti sledi:

$$E(Y) = E(X) + E(Y - X)$$

Ker je $Y - X \geq 0$ z verjetnostjo 1, je $E(Y - X) \geq 0$. $\square$

Izrek 2. *Če obstajata $E\left(X^2\right)$ in $E\left(Y^2\right)$, potem obstaja tudi $E(|XY|)$ in velja:*

$$E(|XY|) \leq \sqrt{E\left(X^2\right) E\left(Y^2\right)}$$

Enakost velja $\iff$

$$Y = \pm\sqrt{\frac{E\left(Y^2\right)}{E\left(X^2\right)}} X \qquad \text{z verjetnostjo } 1.$$

Dokaz. Za poljubni nenegativni števili u in v velja:

$$uv \leq \frac{1}{2}\left(u^2 + v^2\right) \qquad \left(\iff (u - v)^2 \geq 0\right)$$

Torej za nanegativni slučajni spremenljivki U in V velja:

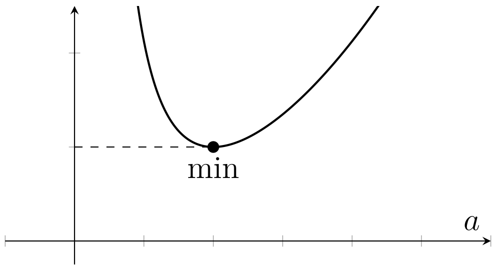
$$UV \leq \frac{1}{2}\left(U^2 + V^2\right)$$

kjer velja enačaj, le za $\omega \in \Omega$, v katerih je $U(\omega) = V(\omega)$. Če vstavimo $U = a|X|$ in $V = \frac{1}{a}|Y|$ za $a > 0$, dobimo neenakost:

$$|XY| \leq \frac{1}{2}\left(a^2 X^2 + \frac{1}{a^2} Y^2\right)$$

Iz $E\left(X^2\right) < \infty$ in $E\left(Y^2\right) < \infty$ ter monotnosti sledi $E(\mid XY \mid) < \infty$, torej smo dokazali obstoj in:

$$E(|XY|) \leq \frac{1}{2}\left(a^2 E\left(X^2\right) + \frac{1}{a^2}E\left(Y^2\right)\right)$$



Slika 4.1: Graf $a^2 + \frac{1}{a^2}$

Minimum desne strani (po $a \in (0, \infty)$) je dosežen pri $a^2 = \sqrt{\frac{E(Y^2)}{E(X^2)}}$ in znaša:

$$\frac{1}{2}\left(\sqrt{\frac{E\left(Y^2\right)}{E\left(X^2\right)}}E\left(X^2\right) + \sqrt{\frac{E\left(X^2\right)}{E\left(Y^2\right)}}E\left(Y^2\right)\right) = \sqrt{E\left(X^2\right)E\left(Y^2\right)}$$

torej je:

$$E(|XY|) \leq \sqrt{E\left(X^2\right)E\left(Y^2\right)}$$

Enačaj velja, če je $U = V$, torej $a|X| = \frac{1}{a}|Y|$:

$$|Y| = a^2|X| = \sqrt{\frac{E\left(Y^2\right)}{E\left(X^2\right)}}|X|$$

$\square$

Posledica 8. *Če je $E\left(X^2\right) < \infty$, je $E(|X|) < \infty$ (vzamemo $Y \equiv 1$) in $E(|X|) \leq \sqrt{E\left(X^2\right)}$.*

Poglavje 5

Disperzija, kovarianca in variančno-kovariančne matrike

5.1 Disperzija in standardni odklon

Definicija 7. Naj obstaja $E\left(X^2\right)$. *Disperzija* oz. *varianca* je definirana kot:

$$D(X) \equiv \mathrm{Var}(X) := E\left((X - E(X))^2\right)$$

$D(X)$ je torej pričakovano kvadratično odstopanje od $E(X)$. Velja tudi:

$$\begin{aligned} E\left((X - E(X))^2\right) &= E\left(X^2 - 2E(X)X + (E(X))^2\right) \\ &= E\left(X^2\right) - 2E(X)E(X) + (E(X))^2 \\ &= E\left(X^2\right) - (E(X))^2 \end{aligned}$$

Zato je:

$$D(X) = E\left(X^2\right) - (E(X))^2$$

Če ima X "zvezno" gostoto f_X, iz trditve o $E(g(X))$ sledi:

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

Definicija 8. *Standardna deviacija* oz. *standardni odklon* je definirana kot:

$$\sigma(X) := \sqrt{D(X)}$$

Opomba. Standardni odklon ima iste "enote" kot X. Intuitivnejša mera za odstopanje vrednost X od $E(X)$ bi bila $E(|X - E(X)|)$. Razlog za to, da favoriziramo $\sigma(X)$ oz. $D(X)$ je ta, da je $D(X)$ porojena s skalarnim produktom.

Zgled.

- $X \sim \mathrm{U}\left(-\frac{1}{2}, \frac{1}{2}\right)$: $f_X(x) = \mathbb{1}_{\left(-\frac{1}{2}, \frac{1}{2}\right)}(x)$ in $E(X) = 0$. Zato je:

$$D(X) = E\left(X^2\right) = \int_{\mathbb{R}} x^2 \cdot \mathbb{1}_{\left(-\frac{1}{2}, \frac{1}{2}\right)}(x)\, dx = \int_{-\frac{1}{2}}^{\frac{1}{2}} x^2\, dx = \frac{1}{12}$$

- $X \sim \mathrm{N}(0,1)$: Velja $E(X) = 0$, torej:

$$\begin{aligned} D(X) = E\left(X^2\right) &= \int_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx \\ &= \sqrt{\frac{2}{\pi}} \int_0^{\infty} x^2 e^{-\frac{x^2}{2}}\, dx \qquad (\star) \\ &= \sqrt{\frac{2}{\pi}} \int_0^{\infty} \sqrt{2t} \cdot e^{-t}\, dt \\ &= \frac{2}{\sqrt{\pi}} \int_0^{\infty} t^{\frac{3}{2}-1} e^{-t}\, dt \\ &= \frac{2}{\sqrt{\pi}} \Gamma\left(\frac{3}{2}\right) \\ &= \frac{2}{\sqrt{\pi}} \frac{1}{2} \Gamma\left(\frac{1}{2}\right) \\ &= 1 \end{aligned}$$

Kjer smo v $(\star)$ vrstici vpeljali novo spremenljivko $t = \frac{x^2}{2}$.

$\Diamond$

Lastnosti $D(X)$:

1. $D(X) \geq 0$ in $D(X) = 0 \iff P(X = E(X)) = 1$, t.j.: $X : \begin{pmatrix} E(X) \\ 1 \end{pmatrix}$.

2. $D(X)$ je minimum funkcije $a \mapsto E\left((X-a)^2\right)$

   Dokaz.

$$\begin{aligned} E\left((X-a)^2\right) &= E((X - E(X) + E(X) - a)^2) \\ &= D(X) + 2(E(X) - a)\underbrace{E(X - E(X))}_{=0} + \underbrace{(E(X) - a)^2}_{\text{minimalno } 0} \end{aligned}$$

$\square$

3. $D(aX+b) = a^2 D(X)$ za $a, b \in \mathbb{R}$.

   Dokaz.
   $$D(aX+b) = E\left((aX - aE(X))^2\right) = a^2 E\left(X - E(X)\right)^2$$
   □

4. $D(X+Y) = D(X) + D(Y)$, če sta X in Y neodvisni:

   Dokaz.
   $$\begin{aligned} D(X+Y) &= E\left((X+Y)^2\right) - E(X+Y)^2 \\ &= D(X) + D(Y) + 2\underbrace{(E(XY) - E(X)\cdot E(Y))}_{=0} \end{aligned}$$
   □

Zgled.

- $X \sim \mathrm{U}(a,b) \implies \frac{X-a}{b-a} - \frac{1}{2} \sim \mathrm{U}\left(-\frac{1}{2}, \frac{1}{2}\right)$. Vemo, da je $D\left(\frac{X-a}{b-a} - \frac{1}{2}\right) = \frac{1}{12}$, torej:
  $$D\left(\frac{X-a}{b-a} - \frac{1}{2}\right) = \frac{1}{(b-a)^2} \cdot D(X) \implies D(X) = \frac{(b-a)^2}{12}$$
- $X \sim \mathrm{N}\left(\mu, \sigma^2\right) \Rightarrow \frac{X-\mu}{\sigma} \sim \mathrm{N}(0,1)$. Vemo, da je $D\left(\frac{X-\mu}{\sigma}\right) = 1$, torej:
  $$D\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2} \cdot D(X) \implies D(X) = \sigma^2$$

◊

Opomba. Končno vemo, da je σ^2 v $\mathrm{N}\left(\mu, \sigma^2\right)$ disperzija.

Aproksimacija pričakovne vrednosti z disperzijo:

Zanima nas $E(g(X))$ za slučajno spremenljivko X in realno funkcijo g. Če je g dovoljkrat odvedljiva, jo lahko aproksimiramo:
$$g(x) \approx g(\mu) + (x-\mu) \cdot g'(\mu) + \frac{1}{2} \cdot (x-\mu)^2 \cdot g''(\mu)$$
Vstavimo $\mu = E(X)$ in $X(w)$ namesto x, uporabimo pričakovano vrednost in dobimo:
$$E(g(X)) \approx g(E(X)) + \frac{1}{2} \cdot D(X) \cdot g''(E(X))$$
V praski je ta približek tipično dober, še boljše lastnosti pa imajo analogni približki $E(h(X,Y))$.

5.2 Kovarianca

Definicija 9. *Kovarianca* slučajnih spremenljivk X in Y je:

$$K(X,Y) \equiv \mathrm{Cov}(X,Y) = E((X-E(X))(Y-E(Y)))$$

če obstaja.

Če sta X in Y neodvisni in imata pričakovano vrednost, je $K(X,Y)=0$.

Izrek 3 (Cauchy-Schwartzova neenakost)**.** *Če obstajata $E\left(X^2\right)$ in $E\left(Y^2\right)$, potem obstaja $K(X,Y)$ in velja:*

$$|K(X,Y)| \leq \sqrt{D(X)D(Y)}$$

Enakost velja $\iff$

$$\frac{X-E(X)}{\sigma(X)} = \frac{Y-E(Y)}{\sigma(Y)} \quad \textit{oz.} \quad \frac{X-E(X)}{\sigma(X)} = -\frac{Y-E(Y)}{\sigma(Y)} \quad \textit{z verjetnostjo 1.}$$

V tem primeru pravimo, da sta X in Y skoraj gotovo linearno povezana.

Dokaz. Izrek 2 uporabimo za $X-E(X)$ in $Y-E(Y)$ v vlogah X in Y. □

Pravimo, da $K(X,Y)$ meri jakost linearne povezanosti slučajnih spremenljivk X in Y. Iz formule za $E(h(X,Y))$ v primeru vektorja z "zvezno" gostoto $f_{(X,Y)}$ sledi:

$$K(X,Y) = \int_{\mathbb{R}^2} (x-E(X))(y-E(Y)) f_{(X,Y)}(x,y)\,dxdy$$

če integral absolutno konvergira.

Lastnosti $K(X,Y)$:

1. Če obstajajo $E(X)$, $E(Y)$ in $E(XY)$, obstaja tudi $K(X,Y)$ in velja:

   $$K(X,Y) = E(XY) - E(X)E(Y)$$

   Dokaz. Iz $(X-E(X))(Y-E(X)) = XY - E(X)\cdot Y - X\cdot E(Y) + E(X)E(Y)$ vidimo, da $K(X,Y)$ obstaja. Podoben račun pokaže zgornjo enakost. □

2. K je simetrična bilinearna funkcija.

Dokaz. Simetričnost je očitna. Linearnost je dovolj preveriti v eni spremenljivki:

$$\begin{aligned}
K\left(aX_1+bX_2,Y\right) &= E\left(\left(aX_1+bX_2-aE\left(X_1\right)-bE\left(X_2\right)\right)\left(Y-E(Y)\right)\right)\\
&= aE\left(\left(X_1-E\left(X_1\right)\right)\left(Y-E(Y)\right)\right)+\\
&\qquad\qquad\qquad + bE\left(\left(X_2-E\left(X_2\right)\right)\left(Y-E(Y)\right)\right)\\
&= aK\left(X_1,Y\right)+bK\left(X_2,Y\right)
\end{aligned}$$

□

3. Če obstajata $E\left(X^2\right)$ in $E\left(Y^2\right)$ velja:

$$D(X+Y)=D(X)+2K(X,Y)+D(Y)$$

4. Če obstajajo $E\left(X_1^2\right),\ldots,E\left(X_n^2\right)$ velja:

$$\begin{aligned}
D\left(X_1+\cdots+X_n\right) &= \sum_{i=1}^{n} D\left(X_i\right)+2\sum_{1\le i<j\le n} K\left(X_i,X_j\right)\\
&= \sum_{1\le i,j\le n} K\left(X_i,X_j\right)
\end{aligned}$$

Opomba. Velja $K(X,X)=D(X)$.

5.3 Variančno-kovariančne matrike

Definicija 10. Če je $X=(X_1,\ldots,X_n)$ slučajni vektor, je njegova *variančno-kovariančna matrika* matrika $\mathrm{Var}(X)\in\mathbb{R}^{n\times n}$ z elementi:

$$[\mathrm{Var}(X)]_{i,j}=K\left(X_i,X_j\right)$$

To je simetrična matrika z diagonalnimi elementi $D\left(X_1\right),\ldots,D\left(X_n\right)$.

Zgled. Oglejmo si $\begin{bmatrix}X\\Y\end{bmatrix}\sim \mathrm{N}\left(\begin{bmatrix}\mu_1\\\mu_2\end{bmatrix},\begin{bmatrix}\sigma_1^2 & \sigma_{12}\\ \sigma_{12} & \sigma_2^2\end{bmatrix}\right)$. Želimo izračunati $K(X,Y)$. Naj bo $A=\begin{bmatrix}1 & b\\ 0 & 1\end{bmatrix}$ za primerno število b. Vemo, da je:

$$A\begin{bmatrix}X\\Y\end{bmatrix}=\begin{bmatrix}X+bY\\Y\end{bmatrix}\sim \mathrm{N}\left(\begin{bmatrix}\mu_1+b\mu_2\\\mu_2\end{bmatrix},A\Sigma A^\top\right)$$

Tu je:

$$A\Sigma A^\top = + \begin{bmatrix} \sigma_1^2 + 2b\sigma_{12} + b^2\sigma_2^2 & \sigma_{12} + b\sigma_2^2 \\ \sigma_{12} + b\sigma_2^2 & \sigma_2^2 \end{bmatrix}$$

Naj bo $b = -\frac{\sigma_{12}}{\sigma_2^2}$. Dobimo:

$$A\Sigma A^\top = \begin{bmatrix} \sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^2} & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$$

Iz razdelka o neodvisnosti sledi (za ta b), da sta $X + bY$ in Y neodvisni. Posledično je $K(X + bY, Y) = 0$, po drugi strani pa je enaka:

$$K(X, Y) - \frac{\sigma_{12}}{\sigma_2^2} K(Y, Y) = K(X, Y) - \sigma_{12} \implies K(X, Y) = \sigma_{12}$$

$\Diamond$

Domača naloga: Prepričaj se, da iz zgornje izpeljave sledi:

$$\begin{bmatrix} X \\ Y \end{bmatrix} \sim \mathrm{N}\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} \right) \implies Y \sim \mathrm{N}\left(\mu_2, \sigma_2^2\right)$$

5.4 Korelacijski koeficient

Definicija 11. Če obstajata $E\left(X^2\right)$ in $E\left(Y^2\right)$, lahko definiramo *Pearsonov korelacijski koeficient*:

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

Opazimo:

$$\rho(X, Y) = \frac{E((X - E(X))(Y - E(Y)))}{\sigma(X)\sigma(Y)} = E\left(\frac{X - E(X)}{\sigma(X)} \cdot \frac{Y - E(Y)}{\sigma(Y)} \right)$$

Zgled. Za dvorazsežno normalno porazdeljen slučajni vektor (X, Y) je $\rho(X, Y) = \frac{\sigma_{12}}{\sigma_1\sigma_2}$ natanko parameter "ρ" iz alternativne parametrizacije porazdelitve. $\Diamond$

Enakosti $\rho(X, Y) = 1$ in $\rho(X, Y) = -1$ (oz. $|\rho(X, Y)| = 1$) nastopita natanko tedaj, ko sta X in Y skoraj gotovo v linearni zvezi. Torej:

$$\rho = 1 \iff \frac{X - E(X)}{\sigma(X)} = \frac{Y - E(Y)}{\sigma(Y)} \quad \text{z verjetnostjo 1}$$

$$\rho = -1 \iff \frac{X - E(X)}{\sigma(X)} = -\frac{Y - E(Y)}{\sigma(Y)} \quad \text{z verjetnostjo 1}$$

Definicija 12. Slučajni spremenljivki X in Y sta *nekorelirani*, če je $E(XY) = E(X) \cdot E(Y)$.

Opomba.

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

Poglavje 6

Pogojne porazdelitve in pogojno matematično upanje

Uvod:

Proučujemo slučajni vektor (X, Y):

$$X(\text{človek}) = \text{DKT}(\text{človek}) \qquad \text{je diastolični krvni tlak;}$$
$$X(\text{človek}) = \text{SKT}(\text{človek}) \qquad \text{je sistolični krvni tlak.}$$

Zanima nas:

(i) Izmerimo si DKT, izkaže se, da je enak 80 mm Hg. Zanima nas verjetnost, da je SKT ≤ 120 mm Hg.

(ii) Zanima nas pričakovana vrednost SKT, pri pogoju, da je DKT $= 80$ mm Hg.

Vprašanje (i) bomo interpretirali kot $P(Y \leq 120 \mid X = 80)$. To *ni* elementarna pogojna verjetnost, ker je ulomek:

$$\frac{P(Y \leq 120 \text{ in } X = 80)}{P(X = 80)} = \frac{0}{0}$$

nedoločen zaradi zveznosti X. Kako bi potem pri zvezni slučajni spremenljivki X definirali $P(Y \in \mathcal{B} \mid X = x)$? Ker je $\{x\} = \bigcap_{h>0}(x - h, x + h)$ postopamo takole:

Definicija 13. Naj bo $\mathcal{B} \subseteq \mathbb{R}$ Borelova množica (npr. interval). Pogojna verjetnost, da Y zavzame vrednost v $\mathcal{B}$, pri pogoju $X = x$, je enaka:

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

če ta limita obstaja. Potreben pogoj za obstoj je, da so $P(X \in (x-h, x+h)) > 0$ za vsa dovolj majhna števila h. To velja (na primer), če je f_X zvezna v x in $f_X(x) > 0$.

Opomba. Če je X diskretna slučajna spremenljivka z vrednostmi $x_1, \ldots, x_n$ velja:

$$\{X \in (x_i - h, x_i + h)\} = \{X = x_i\}$$

za vsa dovolj majhna števila h. Sledi, da je:

$$\lim_{h \downarrow 0} P(Y \in \mathcal{B} \mid X \in (x-h, x+h))$$

v resnici elementarna pogojna verjetnost.

Trditev 9. *Naj ima (X, Y) "zvezno" gostoto $f_{(X,Y)}$ in velja:*

1. *$f_X(x) > 0$, ter f_X zvezna v točki x;*
2. *funkcija $t \mapsto \int_{\mathbb{R}} f_{(X,Y)}(t, y)\, dy$ je zvezna v x.*

Tedaj obstaja zgornja limita in je enaka:

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

Dokaz.

$$\frac{P(Y \in \mathcal{B} \text{ in } X \in (x-h, x+h))}{P(X \in (x-h, x+h))} = \frac{\iint_{(x-h,x+h)\times\mathcal{B}} f_{(X,Y)}(t, y)\, dtdy \cdot \frac{1}{2h}}{\int_{(x-h,x+h)} f_X(t)\, dt \cdot \frac{1}{2h}}$$

Imenovalec:

$$\begin{aligned}
\frac{1}{2h} \int_{x-h}^{x+h} f_X(t)\, dt &= \frac{1}{2h} \left(\int_{-\infty}^{x+h} f_X(t)\, dt - \int_{-\infty}^{x-h} f_X(t)\, dt \right) \\
&= \frac{1}{2h} \left(F_X(x+h) - F_X(x-h) \right) \\
&= \frac{1}{2} \left(\frac{F_X(x+h) - F_X(x)}{h} + \frac{F_X(x) - F_X(x-h)}{h} \right)
\end{aligned}$$

Iz predpostavke 1. sledi, da ta izraz konvergira k $f_X(x)$.

Števec:

$$\frac{1}{2h} \cdot \int_{x-h}^{x+h} \left(\int_{\mathcal{B}} f_{(X,Y)}(t, y)\, dy \right) dt$$

Po izreku o srednji vrednosti in predpostavki 2. sledi, da ta izraz konvergira k integralu $\int_{\mathcal{B}} f_{(X,Y)}(x, y)\, dy$. □

Naj bo $f_X(x) > 0$. Ker je $y \mapsto \frac{f_{(X,Y)}(x,y)}{f_X(x)}$ nenegativna funkcija in velja:

$$\int_{\mathbb{R}} \frac{f_{(X,Y)}(x,y)}{f_X(x)}\, dy = 1$$

je ta funkcija gostota neke porazdelitve. Tej porazdelitvi pravimo *pogojna porazdelitev* slučajne spremenljivke Y pri pogoju $X = x$. Sestavljajo jo pogojne verjetnosti $P(Y \in \mathcal{B} \mid X = x)$ za Borelove množice $\mathcal{B} \subseteq \mathbb{R}$.

Če ima (X, Y) "zvezno" gostoto $f_{(X,Y)}$, je verjetnost, da X zavzame tak x, ki zadošča predpostavkam trditve, enaka 1. Oziroma, za:

$$D = \{x \in \mathbb{R} \mid f_X(x) > 0 \text{ in } f_X \text{ je zvezna pri } x\}$$

velja:

$$P(X \in D) = \int_D f_X(x)\, dx = \int_{\mathbb{R}} f_X(x)\, dx = 1$$

Zaradi popolnosti (za teoretične namene) definiramo t.i. *pogojno gostoto*:

$$f_{Y|X}(y \mid x) = \begin{cases} \frac{f_{(X,Y)}(x,y)}{f_X(x)} & ; f_X \text{ zvezna in pozitivna v } x \\ \phi_0(y) & ; \text{ sicer} \end{cases}$$

kjer je ϕ_0 gostota neke fiksne porazdelitve na $\mathbb{R}$. Pri tej definiciji je za $\forall x \in \mathbb{R}$ funkcija $y \mapsto f_{(Y|X)}(y \mid x)$ gostota neke porazdelitve. Končno redefiniramo:

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

Opomba. Če sta X in Y neodviski, je $f_{Y|X}(y \mid x) = f_Y(y)$ za skoraj vse x. Brez privzetka zveznosti je $P(Y \in \mathcal{B} \mid X = x) = P(Y \in \mathcal{B})$.

Trditev 10. *Velja t.i. zakon popolne verjetnosti:*

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

za zvezno porazdelitve in:

$$P(Y \in \mathcal{B}) = \sum_i P(Y \in \mathcal{B} \mid X = x_i) \cdot P(X = x_i)$$

za diskretno porazdelitev in vsako Borelovo množico $\mathcal{B} \subseteq \mathbb{R}$.

Dokaz. Desna stran je enaka:

$$\begin{aligned}
\int_{-\infty}^{\infty}\left(\int_{\mathcal{B}} f_{(Y|X)}(y \mid x)\,dy\right) f_X(x)\,dx &= \int_{-\infty}^{\infty}\int_{\mathcal{B}} \underbrace{f_{(Y|X)}(y \mid x) f_X(x)}_{=f_{(X,Y)}(x,y) \text{ skoraj povsod}}\,dy\,dx \\
&= \iint_{\mathbb{R}\times\mathcal{B}} f_{(X,Y)}(x,y)\,dxdy \\
&= P((X,Y)\in\mathbb{R}) \\
&= P(Y\in\mathcal{B})
\end{aligned}$$

□

Zgled. Naj bo $\begin{bmatrix} X \\ Y \end{bmatrix} \sim \mathrm{N}\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}\right)$. Izračunajmo $f_{(Y|X)}(y,x) = \frac{f_{(X,Y)}(x,y)}{f_X(x)}$. Zapišemo:

$$f_{(X,Y)}(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left(-\frac{1}{2(1-\rho^2)}\left(\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right)\right)$$

$$\begin{aligned}
f_X(x) &= \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma_1}\exp\left(-\frac{1}{2\sigma_1^2}(x-\mu_1)^2\right) \\
&= \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma_1}\exp\left(-\frac{1}{2(1-\rho^2)}\left(\left(1-\rho^2\right)\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right)\right)
\end{aligned}$$

Sledi:

$$\frac{f_{(X,Y)}(x,y)}{f_X(x)} = \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma_2\sqrt{1-\rho^2}}\exp\left(-\frac{1}{2(1-\rho^2)}\left(\underbrace{\rho^2\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\frac{x-\mu_1}{\sigma_1}\frac{y-\mu_2}{\sigma_2} + \left(\frac{y-\mu_2}{\sigma_2}\right)^2}_{=\left(\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right)^2}\right)\right)$$

Prepoznamo gostoto porazdelitve $(Y \mid X = x) \sim \mathrm{N}\left(\mu_2 + \rho\sigma_2\frac{x-\mu_1}{\sigma_1}, \sigma_2^2\left(1-\rho^2\right)\right)$. ◊

Definicija 14. *Pogojna pričakovana vrednost slučajne spremenljivke Y, pogojeno na $X = x$ je pričakovana vrednost pogojne porazdelitve $(Y \mid X = x)$. V primeru, ko ima (X,Y) "zvezno" gostoto, to pomeni:*

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

Opomba. Velja tudi:

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

Izrek 4. *Velja zakon popolne pričakovane vrednosti:*

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

Dokaz. Desna stran se glasi:

$$\begin{aligned}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} g(y) f_{(Y|X)}(y \mid x)\, dy\right) f_X(x)\, dx &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(y) f_{(X,Y)}(x,y)\, dx\, dy \\ &= \int_{-\infty}^{\infty} g(y) f_Y(y)\, dy \\ &= E(g(Y))\end{aligned}$$

□

Definicija 15. *Pogojna pričakovana vrednost* $E(Y \mid X)$ je *slučajna spremenljivka*, definirana (na istem verjetnostnem prostoru) kot:

$$E(Y \mid X)(\omega) = E(Y \mid X = X(\omega))$$

Opomba. Velja tudi:

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

Če pišemo $g(x) = E(Y \mid X = x)$, je $E(Y \mid X) = g(X)$ transformacija slučajne spremenljivke X. Sledi:

Posledica 9. *Za slučajno spremenljivko $E(g(Y) \mid X)$ velja:*

$$E(E(g(Y) \mid X)) = E(g(Y))$$

Dokaz.

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

□

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

Zgled (nadaljevanje)**.** Za $\begin{bmatrix} X \\ Y \end{bmatrix} \sim \mathrm{N}\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}\right)$ kot prej je:

$$(Y \mid X = x) \sim \mathrm{N}\left(\mu_2 + \rho\sigma_2 \frac{x - \mu_1}{\sigma_1}, \sigma_2^2\left(1 - \rho^2\right)\right)$$

Sledi, da je $E(Y \mid X = x) = \mu_2 + \rho\sigma_2^2 \frac{x-\mu_1}{\sigma_1}$ in zato:

$$E(Y \mid X) = \mu_2 + \rho\sigma_2 \frac{X - \mu_1}{\sigma_1}$$

To funkcijo imenujemo *regresijska funkcija*. ◊

Poglavje 7

Momenti in momentno rodovna funkcija

7.1 Momenti

Definicija 16. Naj bo $k \in \mathbb{N}$ in $a \in \mathbb{R}$. *Moment* slučajne spremenljivke X *reda* k *glede na* a je:

$$m_k(a) = E\left((X-a)^k\right)$$

če obstaja, t.j. $E\left(|X-a|^k\right) < \infty$.

Za a običajno vzamemo:

1. $a = 0$: $z_k = m_k(0) = E\left(X^k\right)$; *$k$- ti začetni moment.*
2. $a = E(X)$: $m_k = m_k(E(X)) = E\left((X-E(X))^k\right)$; *$k$-ti centralni moment.*

Očitno je $z_1 = m_1(0) = E(X)$ in $m_2 = D(X)$.

Trditev 11. *Če obstaja $m_n(a)$, potem obstaja tudi $m_k(a)$ za vsak $k < n$.*

Dokaz. (samo zvezen primer)

$$\begin{aligned}
\int_{-\infty}^{\infty} |x-a|^k f(x)\,dx &= \underbrace{\int_{|x-a|\leq 1} |x-a|^n f(x)\,dx}_{\leq 1} + \int_{|x-a|>1} \underbrace{|x-a|^k f(x)}_{\leq |x-a|^n f(x)}\,dx \\
&\leq 1 + E\left(|X-a|^n\right) \\
&< \infty
\end{aligned}$$

saj $m_n(a)$ obstaja. ∎

Posledica 10. *Če obstaja* $E(X^2)$, *obstaja tudi* $D(X)$.

Trditev 12. *Če obstaja moment* $m_n(a)$ *za nek* $a \in \mathbb{R}$, *potem obstajajo tudi* $m_n(b)$ *za vse* $b \in \mathbb{R}$.

Dokaz.

$$\begin{aligned} E(|X-b|^n) &= E(|X-a+a-b|^n) \\ &\leq E((|x-a|+|a-b|)^n) \\ &= \sum_{k=0}^{n} \binom{n}{k} |a-b|^{n-k} E\left(|x-a|^k\right) \\ &< \infty \end{aligned}$$

saj $E\left(|X|^k\right) < \infty$ po prejšnji trditvi. □

Posledica 11. *Če obstaja* $m_n(a)$, *je:*

$$m_n(b) = \sum_{k=0}^{n} \binom{n}{k} (a-b)^{n-k} E\left((X-a)^k\right)$$

Motivacija za študij momentov:

Denimo, da imamo porazdelitev, ki ima vse zaćetne momente $z_k = m_k(0) = E\left(X^k\right)$ (torej obstajajo vsi momenti). Ali lahko iz njih "rekonstruiramo" celotno porazdelitev. Drugače povedano: če poznamo $E\left(X^k\right)$, ali potem poznamo vse verjetnosti $P(X \in \mathcal{B})$?

Sorodno vprašanje: če za slučajni spremenljivki X in Y velja $E\left(X^k\right) = E\left(Y^k\right)$, ali potem velja $P(X \in \mathcal{B}) = P(Y \in \mathcal{B})$ za vse $\mathcal{B}$.

Izrek 5. *Naj bosta* X *in* Y *slučajni spremenljivki z vsemi momenti in naj vrsti:*

$$\sum_{k=0}^{\infty} \frac{1}{k!} E\left(X^k\right) t^k \quad \textit{in} \quad \sum_{k=0}^{\infty} \frac{1}{k!} E\left(Y^k\right) t^k$$

konvergirata za nek $t > 0$. *V tem primeru sta porazdelitvi* X *in* Y *enolično določeni. Če velja* $E\left(X^k\right) = E\left(Y^k\right)$ $\forall k$ *sta* X *in* Y *enako porazdeljeni. To pomeni:*

$$P(X \in \mathcal{B}) = P(Y \in \mathcal{B})$$

za vse Borelove množice $\mathcal{B}$ *in*

$$F_X(x) = F_Y(x)$$

za vsa števila $x \in \mathbb{R}$.

Opomba. To pomeni, da je (pri določenih predpostavkah) komulativna porazdelitvena funkcija $F_X : \mathbb{R} \to [0,1]$ določena s števnim naborom števil.

7.2 Momentno rodovna funkcija

Definicija 17. *Momentno rodovna funkcija* slučajne spremenljivke X je funkcija:

$$t \mapsto M_X(t) = E\left(e^{tX}\right)$$

za tiste $t \in \mathbb{R}$, za katere obstaja matematično upanje, t.j.: $E\left(e^{tX}\right) < \infty$.

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

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x)\, dx$$

kjer integral konvergira. To je *Laplaceova transformacija* funkcije f_X.

Če je X diskretna slučajna spremenljivka z vrednostmi $\{x_i\}_i$ je:

$$M_X(t) = \sum_i e^{tx_i} P\left(X = x_i\right)$$

Izrek 6. *Naj obstaja $\delta > 0$, da je $M_X(t) < \infty$, za vse $t \in (-\delta, \delta)$. Potem je M_X neskončnokrat zvezno odvedljiva in vsi začetni momenti obstajajo:*

$$\begin{aligned} z_k = E\left(X^k\right) &= M_X^{(k)}(0) \quad \textit{za vse} \quad k \in \mathbb{N} \\ M_X(t) &= \sum_{k=0}^{\infty} \frac{z_k}{k!} t^k \quad \textit{za vse} \quad t \in (-\delta, \delta) \end{aligned}$$

To pomeni, da je M_X analitična. Porazdelitev slučajne spremenljivke X je natančno določena z momenti.

Zgled. Naj bo $X \sim \mathrm{N}(0,1)$. Tedaj je:

$$\begin{aligned} M_{\mathrm{N}(0,1)}(t) = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-t)^2} e^{-\frac{t^2}{2}}\, dx \\ &= e^{-\frac{t^2}{2}} \int_{-\infty}^{\infty} \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-t)^2}}_{f_{\mathrm{N}(t,1)}}\, dx \\ &= e^{-\frac{t^2}{2}} \end{aligned}$$

◊

Trditev 13.

$$M_{aX+b}(t) = e^{bt} M_X(at) \quad \textit{za} \quad a \neq 0, b \in \mathbb{R}$$

Dokaz.

$$\begin{aligned} M_{aX+b}(t) &= E\left(e^{t(aX+b)}\right) \\ &= E\left(e^{(at)X}\, e^{bt}\right) \\ &= e^{bt}\, M_X(at) \end{aligned}$$

□

Zgled.

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

◊

Trditev 14. *Če sta X in Y neodvisni slučajni spremenljivki velja:*

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

Lema 1. *Če sta X in Y neodvisni slučajni spremenljivki in sta g in h poljubni zvezni funkciji, sta tudi $g(X)$ in $h(Y)$ neodvisni slučajni spremenljivki.*

Dokaz leme. Za Borelovi množici $\mathcal{C}$ in $\mathcal{D}$:

$$\begin{aligned} P(g(X) \in \mathcal{C} \text{ in } h(Y) \in \mathcal{D}) &= P\left(X \in g^{-1}(\mathcal{C}) \text{ in } Y \in h^{-1}(\mathcal{D})\right) \\ &= P\left(X \in g^{-1}(\mathcal{C})\right) P\left(Y \in h^{-1}(\mathcal{D})\right) \qquad (\star) \\ &= P(g(X) \in \mathcal{C}) P(h(Y) \in \mathcal{D}) \end{aligned}$$

Kjer smo v $(\star)$ vrstici uporabili neodvisnost X in Y. Upoštevali smo še dejstvo, da so praslike Borelovih množic Borelove množice. □

Dokaz trditve.

$$\begin{aligned} M_{X+Y}(t) &= E\left(e^{t(X+Y)}\right) \\ &= E\left(e^{tX}\, e^{tY}\right) \qquad (\star) \\ &= E\left(e^{tX}\right) E\left(e^{tY}\right) \\ &= M_X(t)\, M_Y(t) \end{aligned}$$

Kjer smo v $(\star)$ vrstici upoštevali neodvisnot po lemi 1. □