PRAVDEPODOBNOSŤ a ŠTATISTIKA

Náhodný vektor a jeho charakteristiky

doc. RNDr. Štefan Peško, CSc.

Katedra matematických metód a operačnej analýzy, FRI ŽU

21. mája 2019

Náhodný vektor

Nech je daný pravdepodobnostný priestor $(\Omega, \mathcal{B}, \mathcal{P})$ a nech $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ sú náhodné premenné definované na tomto priestore. Potom *n* rozmerným náhodným vektorom rozumieme usporiadanú *n*-ticu $\mathbb{X} = (\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n)$.

Poznámka:
Obmedzíme sa len na štúdium vzťahov medzi dvoma náhodnými premennými t.j. budeme sa zaoberať dvojrozmerným náhodným vektorom s tým, že získané poznatky sa dajú ;-) jednoducho zovšeobecniť na viacrozmerný náhodný vektor.

Dohoda:
Ďalej pod pojmom náhodný vektor budeme v tejto prednáške rozumieť dvojrozmerný náhodný vektor.

Združená distribučná funkcia

Združenou distribučnou funkciou náhodného vektora $(\mathbf{X}, \mathbf{Y})$ nazývame reálnu funkciu $F : \Re^2 \to \Re$ definovanú rovnosťou

$$F_{\mathbf{X},\mathbf{Y}}(x, y) = \mathcal{P}([\mathbf{X} < x] \cap [\mathbf{Y} < y]). \tag{1}$$

Dohoda:

Pre $\mathcal{P}([\mathbf{X} < x] \cap [\mathbf{Y} < y])$ budeme používať skrátený zápis $\mathcal{P}(\mathbf{X} < x, \mathbf{Y} < y)$.

Poznámka:

Zápis (1) znamená, že hodnota distribučnej funkcie v bode (x, y) sa rovná pp., že náhodná premenná $\mathbf{X}$ je menšia než x a súčasne náhodná premenná $\mathbf{Y}$ je menšia než y.

Tvrdenie 13

Distribučná funkcia $F_{\mathbf{X},\mathbf{Y}}$ náhodného vektora $(\mathbf{X}, \mathbf{Y})$ je funkcia

a) kde $\forall (x, y) \in \Re^2$ platí

$$0 \leq F_{\mathbf{X},\mathbf{Y}}(x, y) \leq 1. \tag{2}$$

b) neklesajúca v oboch premenných t.j. $\forall x_1, x_2, y_1, y_2 \in \Re$

$$F_{\mathbf{X},\mathbf{Y}}(x_1, y_1) \leq F_{\mathbf{X},\mathbf{Y}}(x_2, y_2) \text{ ak } x_1 < x_2, y_1 < y_2. \tag{3}$$

c) zľava spojitá v oboch premenných t.j. $\forall (a, b) \in \Re^2$ platí

$$\lim_{\substack{x \to a^- \\ y \to b^-}} F_{\mathbf{X},\mathbf{Y}}(x, y) = F_{\mathbf{X},\mathbf{Y}}(a, b). \tag{4}$$

d)

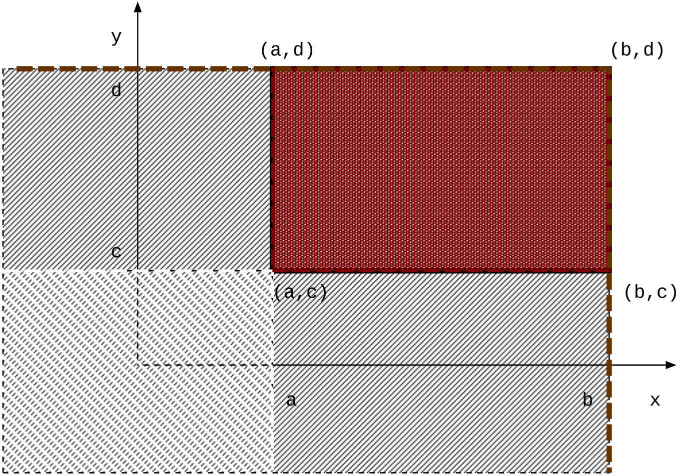
$$\lim_{\substack{x \to -\infty \\ y \to -\infty}} F_{\mathbf{X},\mathbf{Y}}(x, y) = 0, \quad \lim_{\substack{x \to \infty \\ y \to \infty}} F_{\mathbf{X},\mathbf{Y}}(x, y) = 1. \tag{5}$$

Tvrdenie 14

Nech $(\mathbf{X}, \mathbf{Y})$ je náhodný vektor určený distribučnou funkciu $F_{\mathbf{X},\mathbf{Y}}$ a $(a, b), (c, d)$ intervaly. Potom

$$
\begin{aligned}
(i)\quad & \mathcal{P}(a < \mathbf{X} \leq b, \mathbf{Y} \leq d) = F_{\mathbf{X},\mathbf{Y}}(b, d) - F_{\mathbf{X},\mathbf{Y}}(a, d),\\
(ii)\quad & \mathcal{P}(\mathbf{X} \leq b, c < \mathbf{Y} \leq d) = F_{\mathbf{X},\mathbf{Y}}(b, d) - F_{\mathbf{X},\mathbf{Y}}(b, c),\\
(iii)\quad & \mathcal{P}(a < \mathbf{X} \leq b, c < \mathbf{Y} \leq d) =\\
& \qquad = F_{\mathbf{X},\mathbf{Y}}(b, d) - F_{\mathbf{X},\mathbf{Y}}(a, d) + F_{\mathbf{X},\mathbf{Y}}(a, c) - F_{\mathbf{X},\mathbf{Y}}(b, c).
\end{aligned}
$$

$\mathcal{P}(a < \mathbf{X} \leq b, c < \mathbf{Y} \leq d)$



Združená pravdepodobnostná funkcia

Združenou pravdepodobnostnou funkciou náhodného vektora $(\mathbf{X}, \mathbf{Y})$ diskrétnych náhodných premenných nazývame reálnu funkciu $p_{\mathbf{X},\mathbf{Y}} : \Re^2 \to \Re$ definovanú rovnosťou

$$p_{\mathbf{X},\mathbf{Y}}(x_i, y_j) = \mathcal{P}(\mathbf{X} = x_i, \mathbf{Y} = y_j), \tag{6}$$

pre nanajvýš počítateľný počet hodnôt x_i resp. y_j náhodnej premennej $\mathbf{X}$ resp. $\mathbf{Y}$ pre ktorú platí

$$\sum_i \sum_j \mathcal{P}(\mathbf{X} = x_i, \mathbf{Y} = y_j) = 1. \tag{7}$$

Poznámka: V prípade konečného počtom hodnôt sa často funkcia $p_{\mathbf{X},\mathbf{Y}}(x, y)$ reprezentuje tabuľkou

$x; y$	y_1	y_2	y_3	$\dots$	y_n
x_1	$p(x_1, y_1)$	$p(x_1, y_2)$	$p(x_1, y_3)$	$\dots$	$p(x_1, y_n)$
x_2	$p(x_2, y_1)$	$p(x_2, y_2)$	$p(x_2, y_3)$	$\dots$	$p(x_2, y_n)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\dots$	$\vdots$
x_m	$p(x_m, y_1)$	$p(x_m, y_2)$	$p(x_m, y_3)$	$\dots$	$p(x_m, y_n)$

Príklad 6.1

Budeme tri krát hádzať spravodlivou mincou. Nech premenná $\mathbf{X}$ udáva počet pokusov pred prvým hodom znaku a premenná $\mathbf{Y}$ počet po sebe idúcich znakov. Zostavte tabuľku $p_{\mathbf{X},\mathbf{Y}}$.

Označme H-hlava a Z-znak. Môže nastat 8 rôznych výsledkov hodu - náhodných udalostí: HHH,ZHH,HZH,ZZH,HHZ,ZHZ,HZZ,ZZZ. Ak je výsledok hodu HHH, potom $X = 0$, $Y = 0$; ak HHZ potom $X = 2$, $Y = 1$, atď. Náhodné veličiny môžu nadobúdať hodnoty $X \in \{0, 1, 2\}$ a $Y \in \{0, 1, 2, 3\}$ ako vidieť z tejto tabuľky:

$x; y$	0	1	2	3
0	HHH	ZHH,ZHZ	ZZH	ZZZ
1	-	HZH	HZZ	-
2	-	HHZ	-	-

Tabuľka 1: Výsledky hodov priaznivých hodnotám $(\mathbf{X}, \mathbf{Y})$

Jedná sa o hody spravodlivou mincou a tak $\mathcal{P}(H) = \mathcal{P}(Z) = 0.5$
Výsledky hodov sú nezávislé, $\mathcal{P}(HHH) = \cdots = \mathcal{P}(ZZZ) = 0.5^3$.
Náhodné udalosti ZHH a ZHZ sú disjunktné, preto

$$\mathcal{P}(ZHH \cup ZHZ) = \mathcal{P}(ZHH) + \mathcal{P}(ZHZ) = 0.25.$$

$x; y$	0	1	2	3
0	0.125	0.25	0.125	0.125
1	0	0.125	0.125	0
2	0.	0.125	0	0

Tabuľka 2: Tabuľka združených pravdepodobností vektora $(\mathbf{X}, \mathbf{Y})$

Ľahko overíme, že podobne ako pre náhodnú premennú aj pre náhodný vektor platí

$$0 \le p_{\mathbf{X},\mathbf{Y}}(x, y) \le 1, \ \sum_{x=0}^{2} \sum_{y=0}^{3} p_{\mathbf{X},\mathbf{Y}}(x, y) = 1.$$

Združená distribučná funkcia diskrétneho náhodného vektora

Združenou distribučnou funkciou náhodného vektora $(\mathbf{X}, \mathbf{Y})$ diskrétnych náhodných premenných rozumieme reálnu funkciu $F_{\mathbf{X},\mathbf{Y}} : \Re^2 \to \Re$ definovanú rovnosťou

$$F_{\mathbf{X},\mathbf{Y}}(x, y) = \sum_{x_i < x} \sum_{y_j < y} p_{\mathbf{X},\mathbf{Y}}(x_i, y_j), \tag{8}$$

kde x_i resp. y_j sú prípustné hodnoty náhodnej premennej $\mathbf{X}$ resp. $\mathbf{Y}$.

Príklad 6.1 – pokračovanie

Zostavte tabuľku $F_{\mathbf{X},\mathbf{Y}}$ z tabuľky $p_{\mathbf{X},\mathbf{Y}}$ náhodného vektora $(\mathbf{X}, \mathbf{Y})$.

Z odvodenej tabuľky $p_{\mathbf{X},\mathbf{Y}}$

$x; y$	0	1	2	3
0	0.125	0.250	0.125	0.125
1	0	0.125	0.125	0
2	0	0.125	0	0

Tabuľka 3: Tabuľka združených pravdepodobností $p_{\mathbf{X},\mathbf{Y}}$

vypočítame podľa vzťahu (8) tabuľku $F_{\mathbf{X},\mathbf{Y}}$

$x; y$	$(-\infty, 0\rangle$	$(0, 1\rangle$	$(1, 2\rangle$	$(2, 3\rangle$	$(3, \infty)$
$(-\infty, 0\rangle$	0	0	0	0	0
$(0, 1\rangle$	0	0.125	0.375	0.5	0.625
$(1, 2\rangle$	0	0.125	0.5	0.750	0.875
$(2, \infty)$	0	0.125	0.625	0.875	1

Tabuľka 4: Tabuľka združenej distribučnej funkcie $F_{\mathbf{X},\mathbf{Y}}$

Združená hustota pravdepodobnosti a združená distribučná funkcia spojitého náhodného vektora

Združenou hustotou pravdepodobnosti náhodného vektora $(\mathbf{X}, \mathbf{Y})$ spojitých náhodných premenných nazývame nezápornú, integrovateľnú reálnu funkciu $f_{\mathbf{X},\mathbf{Y}} : \Re^2 \to \Re$ pre ktorú platí

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\mathbf{X},\mathbf{Y}}(x, y) \mathrm{d}x \mathrm{d}y = 1. \tag{9}$$

Združenou distribučnou funkciou náhodného vektora $(\mathbf{X}, \mathbf{Y})$ spojitých náhodných premenných rozumieme reálnu funkciu $F_{\mathbf{X},\mathbf{Y}} : \Re^2 \to \Re$ definovanú rovnosťou

$$F_{\mathbf{X},\mathbf{Y}}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{\mathbf{X},\mathbf{Y}}(u, v) \mathrm{d}u \mathrm{d}v, \tag{10}$$

ak existuje hustota pravdepodobnosti $f_{\mathbf{X},\mathbf{Y}}(x, y)$.

Platí vzťah $f_{\mathbf{X},\mathbf{Y}}(x, y) = \frac{\partial^2 F_{\mathbf{X},\mathbf{Y}}(x,y)}{\partial x \partial y}$ for $\forall x \in \Re, \forall y \in \Re$.

Príklad 6.2

Majme náhodný vektor $(\mathbf{X}, \mathbf{Y})$ so združenou funkciou hustoty

$$f_{\mathbf{X},\mathbf{Y}}(x,y) = \begin{cases} \alpha(x+y) & \text{ak } 0 \le x \le 1, 0 \le y \le 1, \\ 0 & \text{ináč.} \end{cases}$$

Určte konštantu α a zostrojte distribučnú funkciu $F_{\mathbf{X},\mathbf{Y}}$.

Z podmienky (9) a z nezáporných hodnôt hustoty dostaneme

$$1 = \int_0^1 \int_0^1 \alpha(x+y)\mathrm{d}x\mathrm{d}y = \alpha \int_0^1 \Big[\int_0^1 (x+y)\mathrm{d}y\Big]\mathrm{d}x$$
$$= \alpha \int_0^1 \Big[xy + \frac{y^2}{2}\Big]_{y=0}^{y=1} \mathrm{d}x = \alpha \int_0^1 \Big(x + \frac{1}{2}\Big)\mathrm{d}x =$$
$$= \alpha\Big[\frac{x^2}{2} + \frac{x}{2}\Big]_0^1 = \alpha.$$

Aby bola funkcia $f_{\mathbf{X},\mathbf{Y}}(x,y)$ združenou hustotou pp., musí $\alpha = 1$.

Z definície distribučnej funkcie (10) dostaneme pre premenné $0 \le x \le 1$, $0 \le y \le 1$

$$
\begin{aligned}
F_{\mathbf{X},\mathbf{Y}}(x,y) &= \int_0^x \int_0^y (u+v)\mathrm{d}u\mathrm{d}v = \int_0^x \Big[\int_0^y (u+v)\mathrm{d}v\Big]\mathrm{d}u \\
&= \int_0^x \Big[uv + \frac{v^2}{2}\Big]_{v=0}^{v=y}\mathrm{d}u == \int_0^x \Big(uy + \frac{y^2}{2}\Big)\mathrm{d}u \\
&= \Big[\frac{u^2y}{2} + \frac{y^2u}{2}\Big]_0^x = \frac{xy}{2}(x+y).
\end{aligned}
$$

a tak

$$
F_{\mathbf{X},\mathbf{Y}}(x,y) = \begin{cases}
0 & \text{ak } x \le 0 \text{ alebo } y \le 0, \\
\frac{xy}{2}(x+y) & \text{ak } 0 < x \le 1 \text{ a } 0 < y \le 1, \\
\frac{x}{2}(x+1) & \text{ak } 0 < x \le 1 \text{ a } y \ge 1, \\
\frac{y}{2}(1+y) & \text{ak } 0 < y \le 1 \text{ a } x \ge 1, \\
1 & \text{ak } x > 1 \text{ a } y > 1,
\end{cases}
$$

Marginálne distribučné funkcie

Ak je $F_{\mathbf{X},\mathbf{Y}}$ združená distribučná funkcia náhodného vektora $(\mathbf{X},\mathbf{Y})$, potom marginálnou distribučnou funkciou náhodnej premennej $\mathbf{X}$ resp. $\mathbf{Y}$ rozumieme distribučné funkcie $F_{\mathbf{X}}$ resp. $F_{\mathbf{Y}}$ určené vzťahmi

$$F_{\mathbf{X}}(x) = \mathcal{P}(\mathbf{X} < x) = \lim_{y\to\infty} F_{\mathbf{X},\mathbf{Y}}(x,y),\ x \in \Re, \tag{11}$$

$$F_{\mathbf{Y}}(y) = \mathcal{P}(\mathbf{Y} < y) = \lim_{x\to\infty} F_{\mathbf{X},\mathbf{Y}}(x,y),\ y \in \Re. \tag{12}$$

Marginálne rozdelenie diskrétneho náhodného vektora

Ak je $p_{\mathbf{X},\mathbf{Y}}(x,y), x \in H_{\mathbf{X}}, y \in H_{\mathbf{Y}}$ združená pravdepodobnostná funkcia náhodného vektora $(\mathbf{X},\mathbf{Y})$, potom marginálnou pravdepodobnostná funkciou náhodnej premennej $\mathbf{X}$ resp. $\mathbf{Y}$ rozumieme pravdepodobnostné funkcie $p_{\mathbf{X}}$ resp. $p_{\mathbf{Y}}$ určené vzťahmi

$$p_{\mathbf{X}}(x) = \mathcal{P}(\mathbf{X} = x) = \sum_{y \in H_{\mathbf{Y}}} p_{\mathbf{X},\mathbf{Y}}(x,y),\ x \in H_{\mathbf{X}}, \tag{13}$$

$$p_{\mathbf{Y}}(y) = \mathcal{P}(\mathbf{Y} = y) = \sum_{x \in H_{\mathbf{X}}} p_{\mathbf{X},\mathbf{Y}}(x,y),\ y \in H_{\mathbf{Y}}. \tag{14}$$

Poznámka:

Keď zadáme združenú pravdepodobnostnú funkciu $p_{\mathbf{X},\mathbf{Y}}$ tabuľkou, potom riadkovým súčtom zodpovedajú pp. $p_{\mathbf{X}}$ a stĺpcovým súčtom zodpovedajú pp. $p_{\mathbf{Y}}$.

Príklad 6.1 – pokračovanie

Doplníme tabuľku $p_{\mathbf{X},\mathbf{Y}}$ o riadkové a stĺpcové súčty pp. a dostaneme

$x; y$	0	1	2	3	$p_{\mathbf{X}}$
0	0.125	0.25	0.125	0.125	0.625
1	0	0.125	0.125	0	0.25
2	0	0.125	0	0	0.125
$p_{\mathbf{Y}}$	0.125	0.5	0.25	0.125	1

Tabuľka 5: Marginálne pravdepodobnosti $p_{\mathbf{X}}, p_{\mathbf{Y}}$ v tabuľke $p_{\mathbf{X},\mathbf{Y}}$

Marginálne distribučné funkcie $F_{\mathbf{X}}, F_{\mathbf{Y}}$

$$F_{\mathbf{X}}(x) = \begin{cases} 0 & \text{ak } x \leq 0, \\ 0.625 & \text{ak } 0 < x \leq 1, \\ 0.875 & \text{ak } 1 < x \leq 2, \\ 1 & \text{ak } x > 2. \end{cases} \quad F_{\mathbf{Y}}(y) = \begin{cases} 0 & \text{ak } y \leq 0, \\ 0.125 & \text{ak } 0 < y \leq 1, \\ 0.625 & \text{ak } 1 < y \leq 2, \\ 0.875 & \text{ak } 2 < y \leq 3, \\ 1 & \text{ak } x > 3. \end{cases}$$

Marginálne rozdelenie spojitého náhodného vektora

Ak je $f_{\mathbf{X},\mathbf{Y}}(x,y)$ združená hustota pravdepodobnosti náhodného vektora $(\mathbf{X},\mathbf{Y})$, potom marginálnou distribučná funkciou náhodnej premennej $\mathbf{X}$ resp. $\mathbf{Y}$ rozumieme distribučné funkcie $F_{\mathbf{X}}$ resp. $F_{\mathbf{Y}}$ určené vzťahmi

$$F_{\mathbf{X}}(x) = \mathcal{P}(\mathbf{X} < x) = \int_{-\infty}^{x}\int_{-\infty}^{\infty} f_{\mathbf{X},\mathbf{Y}}(u,v)\mathrm{d}u\mathrm{d}v, \qquad (15)$$

$$F_{\mathbf{Y}}(y) = \mathcal{P}(\mathbf{Y} < y) = \int_{-\infty}^{\infty}\int_{-\infty}^{y} f_{\mathbf{X},\mathbf{Y}}(u,v)\mathrm{d}u\mathrm{d}v. \qquad (16)$$

Marginálna hustota pravdepodobnosti náhodnej premennej $\mathbf{X}$ resp. $\mathbf{Y}$ je definovaná takto

$$f_{\mathbf{X}}(x) = \int_{-\infty}^{\infty} f_{\mathbf{X},\mathbf{Y}}(x,y)\mathrm{d}y, \ \ x \in \Re, \qquad (17)$$

$$f_{\mathbf{Y}}(y) = \int_{-\infty}^{\infty} f_{\mathbf{X},\mathbf{Y}}(x,y)\mathrm{d}x, \ \ y \in \Re. \qquad (18)$$

Príklad 6.2 – pokračovanie

Máme náhodný vektor $(\mathbf{X}, \mathbf{Y})$ so združenou funkciou hustoty

$$f_{\mathbf{X},\mathbf{Y}}(x,y) = \begin{cases} x+y & \text{ak } 0 \le x \le 1, 0 \le y \le 1, \\ 0 & \text{ináč.} \end{cases}$$

Určte marginálne distribučné funkcie a hustoty pravdepodobnosti.

V intervale $\langle 0, 1 \rangle$ máme podľa vzťahov (17) a (18)

$$f_{\mathbf{X}}(x) = \int_{-\infty}^{\infty} f_{\mathbf{X},\mathbf{Y}}(x,y)\mathrm{d}y = \int_0^1 (x+y)\mathrm{d}y = \left[xy + \frac{y^2}{2}\right]_{y=0}^{y=1} = x + \frac{1}{2},$$

$$f_{\mathbf{Y}}(y) = \int_{-\infty}^{\infty} f_{\mathbf{X},\mathbf{Y}}(x,y)\mathrm{d}x = y + \frac{1}{2}.$$

a tak

$$f_{\mathbf{X}}(x) = \begin{cases} x+\frac{1}{2} & \text{ak } 0 < x \le 1, \\ 0 & \text{ináč.} \end{cases} \qquad f_{\mathbf{Y}}(y) = \begin{cases} y+\frac{1}{2} & \text{ak } 0 < y \le 1, \\ 0 & \text{ináč.} \end{cases}$$

Podľa vzťahov (15) a (16)

$$F_{\mathbf{X}}(x) = \int_{-\infty}^{x} f_{\mathbf{X}}(t)\mathrm{d}t = \int_{0}^{x} \left(t + \frac{1}{2}\right)\mathrm{d}t = \left[\frac{t^2}{2} + \frac{t}{2}\right]_{t=0}^{t=x} = \frac{x}{2}(x+1),$$

$$F_{\mathbf{Y}}(y) = \int_{-\infty}^{y} f_{\mathbf{X},\mathbf{Y}}(x,y)\mathrm{d}x = \frac{y}{2}(y+1).$$

a tak

$$F_{\mathbf{X}}(x) = \begin{cases} 0 & \text{ak } x \leq 0, \\ \frac{x}{2}(x+1) & \text{ak } 0 < x \leq 1, \\ 1 & \text{ak } x > 1. \end{cases} \quad F_{\mathbf{Y}}(y) = \begin{cases} 0 & \text{ak } y \leq 0, \\ \frac{y}{2}(y+1) & \text{ak } 0 < y \leq 1, \\ 1 & \text{ak } x > 1. \end{cases}$$

Pre $0 \leq x \leq 1$ platí (11) a (12)

$$F_{\mathbf{X}}(x) = \lim_{y \to \infty} F_{X,Y}(x,y) = \lim_{y \to 1} \frac{xy}{2}(x+y) = \frac{x}{2}(x+1)$$

a analogicky pre $F_{\mathbf{Y}}(y)$.

Nech je $(\mathbf{X}, \mathbf{Y})$ diskrétny náhodný vektor so združenou pravdepodobnostnou funkciou $p_{\mathbf{X},\mathbf{Y}}$. Podmienené pravdepodobnostné rozdelenie náhodnej premennej $\mathbf{X}$ pri podmienke $\mathbf{Y} = y$ je určené podmienenou pravdepodobnostnou funkciou

$$p_{\mathbf{X},\mathbf{Y}}(x|y) = \mathcal{P}(\mathbf{X} = x|\mathbf{Y} = y) = \frac{p_{\mathbf{X},\mathbf{Y}}(x,y)}{p_{\mathbf{Y}}(y)}, \text{ ak } p_{\mathbf{Y}}(y) > 0. \quad (19)$$

Nech je $(\mathbf{X}, \mathbf{Y})$ je náhodný vektor so združenou hustotou pravdepodobnosti $f_{\mathbf{X},\mathbf{Y}}$. Podmienené pravdepodobnostné rozdelenie náhodnej premennej $\mathbf{X}$ pri podmienke $\mathbf{Y} = y$ je určené podmienenou hustotou pravdepodobnosti

$$f_{\mathbf{X},\mathbf{Y}}(x|y) = \frac{f_{\mathbf{X},\mathbf{Y}}(x,y)}{f_{\mathbf{Y}}(y)}, \text{ ak } f_{\mathbf{Y}}(y) > 0. \quad (20)$$

Príklad 6.1 – pokračovanie

Odvodili sme tabuľku

$x;y$	0	1	2	3
0	0.125	0.25	0.125	0.125
1	0	0.125	0.125	0
2	0	0.125	0	0
$p_\mathbf{Y}$	0.125	0.5	0.25	0.125

Tabuľka 6: Marginálna $p_\mathbf{Y}$ v tabuľke združenej $p_{\mathbf{X},\mathbf{Y}}$

Potom pomocou vzťahu (19) dostaneme

$p_{\mathbf{X},\mathbf{Y}}(x\|y)$	0	1	2	3
0	1	$\frac{1}{2}$	$\frac{1}{2}$	1
1	0	$\frac{1}{4}$	$\frac{1}{2}$	0
2	0	$\frac{1}{4}$	0	0

Tabuľka 7: Tabuľka podmienenej pravdepodobnostnej funkcie $p_{\mathbf{X},\mathbf{Y}}$

Príklad 6.2 – pokračovanie

Pre náhodný vektor $(\mathbf{X}, \mathbf{Y})$ se vypočítali

$$f_{\mathbf{X},\mathbf{Y}}(x,y) = \begin{cases} x+y & \text{ak } 0 \le x \le 1, 0 \le y \le 1, \\ 0 & \text{ináč.} \end{cases}$$

$$f_{\mathbf{X}}(x) = \begin{cases} x+\frac{1}{2} & \text{ak } 0 < x \le 1, \\ 0 & \text{ináč.} \end{cases} \quad f_{\mathbf{Y}}(y) = \begin{cases} y+\frac{1}{2} & \text{ak } 0 < y \le 1, \\ 0 & \text{ináč.} \end{cases}$$

Podľa vzťahu (20) dostávame

$$f_{\mathbf{X},\mathbf{Y}}(x|y) = \begin{cases} 2\frac{x+y}{2y+1} & \text{ak } 0 \le x \le 1, 0 \le y \le 1, \\ 0 & \text{ináč.} \end{cases}$$

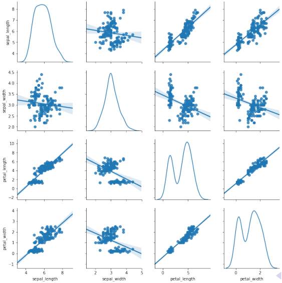
$$f_{\mathbf{X},\mathbf{Y}}(y|x) = \begin{cases} 2\frac{x+y}{2x+1} & \text{ak } 0 \le x \le 1, 0 \le y \le 1, \\ 0 & \text{ináč.} \end{cases}$$

Podmienenou strednou hodnotou diskrétneho resp. spojitého náhodného vektora rozumieme funkciu

$$E(\mathbf{Y}|\mathbf{X}=x_i) = \sum_j y_j p_{\mathbf{X},\mathbf{Y}}(y_j|x_i) \text{ resp.}$$

$$E(\mathbf{Y}|\mathbf{X}=x) = \int_{-\infty}^{\infty} y f_{\mathbf{X},\mathbf{Y}}(y|x)\mathrm{d}y.$$

Podmienená stredná hodnota náhodnej premennej **Y** závislá na hodnotách náhodnej premennej **X** sa nazýva regresná funkcia.



Príklad 6.3

Nech má náhodný vektor $(\mathbf{X},\mathbf{Y})$ združenú hustotu

$$f_{\mathbf{X},\mathbf{Y}}(x,y)=\begin{cases}\frac{2}{3} & \text{ak } 1\le x\le 2, 0\le y\le x,\\ 0 & \text{ináč.}\end{cases}$$

Nájdite regresnú funkciu.

Regresná funkcia je podmienená stredná hodnota náhodnej premennej Y za podmienky, že náhodná premenná X nadobudla niektorú z množiny svojich možných hodnôt.

$$\bar{y}(x)=E(\mathbf{Y}|\mathbf{X}=x)=\int_{-\infty}^{\infty} y f_{\mathbf{X},\mathbf{Y}}(y|x)\mathrm{d}y,$$

$$f_{\mathbf{X}}(x)=\int_{-\infty}^{\infty} f_{\mathbf{X},\mathbf{Y}}(y|x)\mathrm{d}y=\int_0^x \frac{2}{3}\mathrm{d}y=\frac{2}{3}x,$$

$$f_{\mathbf{X},\mathbf{Y}}(y|x)=\frac{f_{\mathbf{X},\mathbf{Y}}(x,y)}{f_{\mathbf{X}}(x)}=\frac{\frac{2}{3}}{\frac{2}{3}x}=\frac{1}{x}, 1\le x\le 2,$$

a tak $\bar{y}(x)=\int_0^x y\frac{1}{x}\mathrm{d}y=\frac{x}{2}, 1\le x\le 2.$

Nezávislosť náhodných premenných

O dvoch náhodných premenných $\mathbf{X}$ a a $\mathbf{Y}$ hovoríme, že sú nezávislé, ak pre každé dva intervaly A a B platí

$$\mathcal{P}(\mathbf{X} \in A, \mathbf{Y} \in B) = \mathcal{P}(\mathbf{X} \in A)\mathcal{P}(\mathbf{Y} \in B). \tag{21}$$

Premenné náhodného vektora $(\mathbf{X}, \mathbf{Y})$ nazývame nezávislé, ak

$$F_{\mathbf{X},\mathbf{Y}}(x, y) = F_{\mathbf{X}}(x)F_{\mathbf{Y}}(y), \tag{22}$$

pre každý bod (x, y) t.j. pre diskrétny vektor platí

$$p_{\mathbf{X},\mathbf{Y}}(x, y) = p_{\mathbf{X}}(x)p_{\mathbf{Y}}(y) \tag{23}$$

a pre spojitý vektor platí

$$f_{\mathbf{X},\mathbf{Y}}(x, y) = f_{\mathbf{X}}(x)f_{\mathbf{Y}}(y). \tag{24}$$

Príklad 6.2 – pokračovanie

Pre $0 \leq x \leq 1, 0 \leq y \leq 1$ je

$$f_{\mathbf{X},\mathbf{Y}}(x,y) = x + y \neq \left(x + \frac{1}{2}\right)\left(y + \frac{1}{2}\right) = f_{\mathbf{X}}(x) f_{\mathbf{Y}}(y).$$

A tak premenné **X** a **Y** náhodného vektora $(\mathbf{X}, \mathbf{Y})$ nie sú nezávislé.

Príklad 6.4

Možno :-) odvodiť, že v náhodnom vektore $(\mathbf{X}, \mathbf{Y})$, ktorý má nenulové hodnoty pre $x > 0, y > 0$ hustoty pp.

$$f_{\mathbf{X},\mathbf{Y}}(x, y) = 4xye^{-x^2-y^2},$$
$$f_{\mathbf{X}}(x) = 2xe^{-x^2},$$
$$f_{\mathbf{Y}}(y) = 2ye^{-y^2}.$$

sú premenné $\mathbf{X}$ a $\mathbf{Y}$ nezávislé, lebo platí (24)

$$f_{\mathbf{X},\mathbf{Y}}(x, y) = f_{\mathbf{X}}(x) f_{\mathbf{Y}}(y).$$

Stredná hodnota funkcie náhodného vektora

Nech je $(\mathbf{X}, \mathbf{Y})$ diskrétny náhodný vektor so združenou pravdepodobnostnou funkciou $p_{\mathbf{X},\mathbf{Y}}, \in H_{\mathbf{Y}}, \mathbf{Y} \in H_{\mathbf{Y}}$. Nech $g : H_{\mathbf{X}} \times H_{\mathbf{Y}} \to \Re$ je reálna funkcia. Strednou hodnotou funkcie $g(\mathbf{X}, \mathbf{Y})$ rozumieme súčet

$$E(g\mathbf{X}, \mathbf{Y})) = \sum_{x \in H_{\mathbf{X}}} \sum_{y \in H_{\mathbf{Y}}} g(x, y) p_{X,Y}(x, y). \qquad (25)$$

Nech je $(\mathbf{X}, \mathbf{Y})$ je náhodný vektor so združenou hustotou pravdepodobnosti $f_{\mathbf{X},\mathbf{Y}}$. Nech $g : \Re^2 \to \Re$ je reálna funkcia. Strednou hodnotou funkcie $g(\mathbf{X}, \mathbf{Y})$ rozumieme integrál

$$E(g(\mathbf{X}, \mathbf{Y})) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{\mathbf{X},\mathbf{Y}}(x, y) \mathrm{d}x \mathrm{d}y, \qquad (26)$$

pokiaľ daný integrál existuje.

Príklad 6.5

Majme náhodný vektor $(\mathbf{X},\mathbf{Y})$ určený združenou hustotou pp.

$$f_{\mathbf{X},\mathbf{Y}}(x,y) = \begin{cases} 1-(x+y)/3 & \text{ak } 0 \le x \le 1, 0 \le y \le 2, \\ 0 & \text{ináč.} \end{cases}$$

Chceme nájsť $E(\mathbf{XY})$.

$$E(\mathbf{XY}) = \int_0^1 \int_0^2 xy\Big(1-\frac{x+y}{3}\Big)\mathrm{d}x\mathrm{d}y = \frac{1}{3}\int_0^1 \Big[\int_0^2 (3xy - x^2y - xy^2)\mathrm{d}y\Big]\mathrm{d}x$$

$$= \frac{1}{3}\int_0^1 \Big[\frac{3xy^2}{2} - \frac{x^2y^2}{2} - \frac{xy^3}{3})\Big]_{y=0}^{y=2}\mathrm{d}x = \frac{1}{3}\int_0^1 \Big(6x - 2x^2 - \frac{8x}{3}\Big)\mathrm{d}x$$

$$= \frac{1}{3}\Big[3x^2 - \frac{2x^3}{3} - \frac{4x^2}{3}\Big]_0^1 = \frac{1}{3}\Big(3-\frac{2}{3}-\frac{4}{3}\Big) = \frac{1}{3}.$$

Zmiešané momenty náhodného vektora

Počiatočným momentom $\nu_{r,s}$ (r, s)-tého rádu náhodného vektora $(\mathbf{X}, \mathbf{Y})$ rozumieme hodnotu

$$\nu_{rs} = E(\mathbf{X}^r \mathbf{Y}^s). \tag{27}$$

Centrálnym momentom $\mu_{r,s}$ (r, s)-tého rádu náhodného vektora $(\mathbf{X}, \mathbf{Y})$ rozumieme hodnotu

$$\mu_{rs} = E((\mathbf{X} - E(\mathbf{X}))^r (\mathbf{Y} - E(\mathbf{Y}))^s). \tag{28}$$

Významné momenty

Počiatočné momenty náhodného vektora $(\mathbf{X}, \mathbf{Y})$, v ktorých je jeden index rovný 1 a druhý 0, zodpovedá stredným hodnotám jednotlivých premenných

$$\nu_{1,0} = E(\mathbf{X}), \quad \nu_{0,1} = E(\mathbf{Y}). \tag{29}$$

Centrálne momenty náhodného vektora $(\mathbf{X}, \mathbf{Y})$, v ktorých je jeden index rovný 2 a druhý 0, zodpovedá rozptylom jednotlivých premenných

$$\mu_{2,0} = D(\mathbf{X}), \quad \mu_{0,2} = D(\mathbf{Y}). \tag{30}$$

Marginálnymi momentami náhodného vektora $(\mathbf{X}, \mathbf{Y})$ rozumieme strednú hodnotu náhodného vektora $(E(\mathbf{X}), E(\mathbf{Y}))$ a disperziu náhodného vektora $(D(\mathbf{X}), D(\mathbf{Y}))$.

Kovariancia

Pod kovarianciou premených náhodného vektora $(\mathbf{X}, \mathbf{Y})$ rozumieme centrálny moment v ktorom sú oba indexy 1. Označujeme ju

$$cov(\mathbf{X}, \mathbf{Y}) = \mu_{1,1} = E((\mathbf{X} - E(\mathbf{X}))(\mathbf{Y} - E(\mathbf{Y}))). \qquad (31)$$

Vlastnosti:

1) $cov(\mathbf{X}, \mathbf{Y}) = cov(\mathbf{Y}, \mathbf{X})$,
2) $cov(\mathbf{X}, \mathbf{X}) = D(\mathbf{X}), cov(\mathbf{Y}, \mathbf{Y}) = D(\mathbf{Y})$,
3) $cov(a\mathbf{X}, b\mathbf{Y}) = abcov(\mathbf{X}, \mathbf{Y})$ pre ľubovoľné $a, b \in \Re$,
4) $cov(\mathbf{X} + a, \mathbf{Y} + b) = cov(\mathbf{X}, \mathbf{Y})$ pre ľubovoľné $a, b \in \Re$.

Kovariančnou maticou $\mathbb{C}$ rozumieme maticu

$$\mathbb{C} = \begin{pmatrix} D(\mathbf{X}) & cov(\mathbf{X}, \mathbf{Y}) \\ cov(\mathbf{Y}, \mathbf{X}) & D(\mathbf{Y}) \end{pmatrix}.$$

Koeficient korelácie

Koeficientom korelácie rozumieme podiel

$$\rho(\mathbf{X}, \mathbf{Y}) = \frac{cov(\mathbf{X}, \mathbf{Y})}{\sqrt{D(\mathbf{X})D(\mathbf{Y})}}, \tag{32}$$

je bezrozmernou charakteristikou, ktorá vyjadruje mieru lineárnej závislosti premenných $\mathbf{X}$ a $\mathbf{Y}$. Ak $\rho(\mathbf{X}, \mathbf{Y}) = 0$, potom hovoríme, že premenné $\mathbf{X}$ a $\mathbf{Y}$ sú nekorelované.

1) $-1 \leq \rho(\mathbf{X}, \mathbf{Y}) \leq 1$,
2) $\rho(\mathbf{X}, \mathbf{X}) = \rho(\mathbf{Y}, \mathbf{Y}) = 1$,
3) $\rho(\mathbf{X}, a) = \rho(\mathbf{Y}, b) = 0$ pre ľubovoľné $a, b \in \Re$,
4) ak $|\rho(\mathbf{X}, \mathbf{Y})| = 1$ potom pre ľubovoľné $a, b \in \Re$, $a \neq 1$ je $\mathbf{Y} = a\mathbf{X} + b$.

Korelačnou maticou $\mathbb{R}$ rozumieme maticu

$$\mathbb{R} = \begin{pmatrix} 1 & \rho(\mathbf{X}, \mathbf{Y}) \\ \rho(\mathbf{Y}, \mathbf{X}) & 1 \end{pmatrix}.$$

Alternatívne vyjadrenie kovariancie

Tvrdenie 15

Nech sú $\mathbf{X}$ a $\mathbf{Y}$ náhodné premenné také, že existujú $E(\mathbf{X}), E(\mathbf{Y})$ a $E(\mathbf{XY})$. Potom platí

$$cov(\mathbf{X},\mathbf{Y}) = E(\mathbf{XY}) - E(\mathbf{X})E(\mathbf{Y}). \tag{33}$$

Dôkaz:

$$\begin{aligned} cov(\mathbf{X},\mathbf{Y}) &= E((\mathbf{X}-E(\mathbf{X}))(\mathbf{Y}-E(\mathbf{Y})) \\ &= E(\mathbf{XY}-E(\mathbf{X})\mathbf{Y}-\mathbf{X}E(\mathbf{Y})+E(\mathbf{X})E(\mathbf{Y})) \\ &= E(\mathbf{XY})-E(\mathbf{X})E(\mathbf{Y})-E(\mathbf{X})E(\mathbf{Y})+E(\mathbf{X})E(\mathbf{Y}) \\ &= E(\mathbf{XY})-E(\mathbf{X})E(\mathbf{Y}). \end{aligned}$$

Príklad 6.6

Majme náhodný vektor $(\mathbf{X}, \mathbf{Y})$ určený združenou hustotou pp.

$$f_{\mathbf{X},\mathbf{Y}}(x, y) = \begin{cases} x + y & \text{ak } 0 \le x \le 1, 0 \le y \le 1, \\ 0 & \text{ináč.} \end{cases}$$

Chceme nájsť korelačnú a kovariačnú maticu.

Pre výpočet $cov(\mathbf{X}, \mathbf{Y})$ podľa (33) najskôr vypočítame marginálnu hustotu $f_{\mathbf{X}}(x) = \int_0^1 (x+y)\mathrm{d}y = x + \frac{1}{2}$ pre $x \in \{0, 1\}$. Odtiaľ $E(\mathbf{X}) = \int_0^1 x(x+\frac{1}{2})\mathrm{d}x = \frac{7}{12}$. $E(\mathbf{X}^2) = \int_0^1 x^2(x+\frac{1}{2})\mathrm{d}x = \frac{5}{12}$ a $D(\mathbf{X}) = E(\mathbf{X}^2) - E(\mathbf{X})^2 = \frac{11}{144}$, $E(\mathbf{XY}) = \int_0^1 \int_0^1 xy(x+y)\mathrm{d}x\mathrm{d}y = \frac{1}{3}$. Zo symetrie združenej hustotu dostaneme tie isté charakteristiky aj pre n.p. $\mathbf{Y}$ a tak máme

$$\mathbb{C} = \begin{pmatrix} \frac{11}{144} & -\frac{11}{144} \\ -\frac{11}{144} & \frac{11}{144} \end{pmatrix}, \ \mathbb{R} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Príklad 6.7

Náhodná premenná X má hustotu pp. $f_{\mathbf{X}}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, x \in \Re$.
Nech náhodná premenná $\mathbf{Y} = \mathbf{X}^2$. Vypočítajte $cov(\mathbf{X}, \mathbf{Y})$.

Náhodná premenná $\mathbf{X} \sim N(0,1)$ a tak $E(\mathbf{X}) = 0$.

$$
\begin{aligned}
cov(\mathbf{X}, \mathbf{Y}) &= E(\mathbf{XY}) - E(\mathbf{X})E(\mathbf{Y}) = E(\mathbf{X}^3) - E(\mathbf{X})E(\mathbf{X}^2) = \\
&= E(\mathbf{X}^3) - 0 \cdot E(\mathbf{X}^2) = E(\mathbf{X}^3) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^3 e^{-\frac{x^2}{2}} \mathrm{d}x.
\end{aligned}
$$

Ale funkcia $x^3 e^{-\frac{x^2}{2}}$ je nepárna a tak $cov(\mathbf{X}, \mathbf{Y}) = 0$.

Poznámka:
Máme príklad kedy $cov(\mathbf{X}, \mathbf{Y}) = 0$ hoci $\mathbf{X}$ a $\mathbf{Y}$ sú závislé premenné.

Tvrdenie 16

Nech sú náhodné premenné $\mathbf{X}$ a $\mathbf{Y}$ nezávislé. Potom $\rho(\mathbf{X},\mathbf{Y}) = 0$.

Dôkaz:

Ak sú náhodné premenné X a Y nezávislé, potom $E(\mathbf{XY}) = E(\mathbf{X})E(\mathbf{Y})$. Podľa tvrdenia 15 je ale

$$cov(\mathbf{X},\mathbf{Y}) = E(\mathbf{XY}) - E(\mathbf{X})E(\mathbf{Y}) = 0$$

a tak podľa definície je $\rho(\mathbf{X},\mathbf{Y}) = 0$.

Poznámka:

Ak $\rho(\mathbf{X},\mathbf{Y}) \neq 0$ potom medzi premennými $\mathbf{X}$ a $\mathbf{Y}$ musí existovať nejaká závislosť.

Príklad 6.8

Majme náhodný vektor $(\mathbf{X}, \mathbf{Y})$, pre ktorý existuje $(E(\mathbf{X}), E(\mathbf{Y}))$. Hľadáme $D(\mathbf{X}+\mathbf{Y})$.

$$\begin{aligned}
D(\mathbf{X}+\mathbf{Y}) &= E((\mathbf{X}+\mathbf{Y})^2) - E^2(\mathbf{X}+\mathbf{Y}) \\
&= E(\mathbf{X}^2+\mathbf{Y}^2+2\mathbf{X}\mathbf{Y}) - (E(\mathbf{X})+E(\mathbf{Y}))^2 \\
&= E(\mathbf{X}^2)+E(\mathbf{Y}^2)+2E(\mathbf{X}\mathbf{Y}) \\
&\quad - E(\mathbf{X})^2 - E(\mathbf{Y})^2 - 2E(\mathbf{X})E(\mathbf{Y}) \\
&= D(\mathbf{X})+D(\mathbf{Y})+2E(\mathbf{X}\mathbf{Y}) - 2E(\mathbf{X})E(\mathbf{Y})) \\
&= D(\mathbf{X})+D(\mathbf{Y})+2cov(\mathbf{X},\mathbf{Y}).
\end{aligned}$$

Cvičenie

6.1 Majme náhodný vektor $(\mathbf{X}, \mathbf{Y})$ so združenou distribučnou funkciou $F_{\mathbf{X},\mathbf{Y}}(x, y) = (1 - e^{-\lambda x})(1 - e^{-\mu y})$ pre $x \geq 0, y \geq 0$. Vypočítajte a) $F_{\mathbf{X},\mathbf{Y}}(x, y)$, b) $\mathcal{P}(1 \leq \mathbf{X} \leq 2, 3 \leq \mathbf{Y} \leq 4)$, c) $F_{\mathbf{X}}(x)$, $F_{\mathbf{Y}}(y)$, d) Platí $F_{\mathbf{X},\mathbf{Y}}(x, y) = F_{\mathbf{X}}(x) \cdot F_{\mathbf{Y}}(y)$?

6.2* Nech D je zjednotenie pravej hornej a ľavej dolnej štvrtiny jednotkového štvorca. a) definujte hustotu $f_{\mathbf{X},\mathbf{Y}}(x, y)$ a distribučnú funkciu $F_{\mathbf{X},\mathbf{Y}}(x, y)$ rovnomerného rozdelenia na množine D, b) vypočítajte $\mathcal{P}((\mathbf{X}, \mathbf{Y}) \in \langle a, b\rangle \times \langle c, d\rangle)$.

6.3* Nech má združená hustota pp. $(\mathbf{X}, \mathbf{Y})$ tvar

$$f_{\mathbf{X},\mathbf{Y}}(x, y) = \begin{cases} c(x^2 + y) & \text{ak } 0 \leq x \leq a, 0 \leq y \leq b, \\ 0 & \text{ináč.} \end{cases}$$

Určte a) c ako funkciu a, b; b) marginálne hustoty pre $\mathbf{X}$ a $\mathbf{Y}$, c) podmienené hustoty $f_{\mathbf{X}}(x|y)$ a $f_{\mathbf{Y}}(y|x)$ d) $F_{\mathbf{X},\mathbf{Y}}(x, y)$, e) $D(\mathbf{X} + \mathbf{Y})$, f) $\rho(\mathbf{X}, \mathbf{Y})$.