Pravděpodobnost a matematická statistika

prof. RNDr. Gejza Dohnal, CSc.
dohnal@nipax.cz

V. Náhodný vektor

https://sms.nipax.cz/pas

Náhodný vektor $Z = (X,Y)$

Dvourozměrné rozdělení pravděpodobnosti:
sdružená distribuční funkce $F(x,y) = \mathrm{P}(X \leq x,\ Y \leq y)$

Jsou-li X a Y spojité náhodné veličiny, potom je i $F(x,y)$ spojitou diferencovatelnou funkcí a její druhou smíšenou parciální derivaci nazýváme sdruženou hustotou náhodného vektoru (X,Y).

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial xy}$$

Objem pod plochou $z = f(x,y)$ určuje pravděpodobnost nad odpovídající oblastí v rovině xy a zřejmě musí platit

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1$$

Tedy je

$$P\big((X,Y) \in W\big) = \iint_W f(x,y)dxdy$$

Speciálně

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)dudv$$

Náhodný vektor $Z = (X,Y)$

Dvourozměrné rozdělení pravděpodobnosti:
sdružená distribuční funkce $F(x,y) = \mathrm{P}(X \leq x,\ Y \leq y)$

$$F(x,y) = P(X \leq x \wedge Y \leq y) = P(X \leq x)P(Y \leq y|X \leq x)$$

$$F(x,y) = F(x)F(y|x)$$

Pokud jsou X a Y stochasticky nezávislé, potom je

$$F(x,y) = P(X \leq x \wedge Y \leq y) = P(X \leq x)P(Y \leq y)$$

$$F(x,y) = F(x)F(y)$$

Platí to i opačně: pokud je $F(x,y) = F(x)F(y)$, potom jsou X a Y stochasticky nezávislé.

Náhodný vektor $Z = (X,Y)$

Dvourozměrné rozdělení pravděpodobnosti:
sdružená distribuční funkce $F(x,y) = \mathrm{P}(X \le x,\ Y \le y)$

$$F(x,y) = P(X \le x \wedge Y \le y) = P(X \le x)P(Y \le y|X \le x)$$

$$F(x,y) = F(x)F(y|x)$$

Pokud jsou X a Y stochasticky nezávislé, potom je

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial xy} = \frac{\partial^2 F(x)F(y)}{\partial xy} = \frac{\partial F(x)}{\partial x}\frac{\partial F(y)}{\partial y}$$

$$f(x,y) = f(x)f(y)$$

Platí to i opačně: pokud je $f(x,y) = f(x)\,f(y)$, potom jsou X a Y stochasticky nezávislé.

Spojitý náhodný vektor $Z = (X_1,\ldots,X_n)$

Jsou-li $X_1,\ldots, X_n$ spojité náhodné veličiny, potom je i jejich sdružená distribuční funkce

$$P(X_1 \leq x_1, \ldots, X_n \leq x_n) = F(x_1, \ldots, x_n)$$

spojitou diferencovatelnou funkcí a její n-tou smíšenou parciální derivaci nazýváme sdruženou hustotou náhodného vektoru Z

$$\frac{\partial^n F(x_1, \ldots, x_n)}{\partial x_1 \ldots x_n} = f(x_1, \ldots, x_n)$$

Tedy je

$$P(Z \in W) = \underset{W}{\int \cdots \int} f(x_1, \ldots, x_n) \mathrm{d}x_1 \ldots \mathrm{d}x_n$$

a zřejmě musí platit

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_n) \mathrm{d}x_1 \ldots \mathrm{d}x_n = 1$$

Spojitý náhodný vektor $Z = (X,Y)$

Jaké je rozdělení pravděpodobnosti složky X?

$$P\big(X \leq x\big) = P\big(X \leq x \wedge Y \leq \infty\big) = \int_{-\infty}^{x}\int_{-\infty}^{\infty} f(x,y)dydx$$

$$= \int_{-\infty}^{x}\left(\int_{-\infty}^{\infty} f(x,y)dy\right)dx = \int_{-\infty}^{x} f_X(u)du$$

Hustotu

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$$

nazýváme marginální hustotou složky X vektoru $Z=(X,Y)$.

Podobně

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx$$

je marginální hustotou složky Y.

Spojitý náhodný vektor $Z = (X_1,\ldots,X_n)$

Jaké je rozdělení pravděpodobnosti složky X_i v náhodném vektoru Z?

$$P(X_i \leq x) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{x} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_i, \ldots, x_n) \mathrm{d}x_1 \ldots \mathrm{d}x_i \ldots \mathrm{d}x_n$$

$$= \int_{-\infty}^{x} \left(\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_i, \ldots, x_n) \mathrm{d}x_1 \ldots \mathrm{d}x_n \right) \mathrm{d}x_i = \int_{-\infty}^{x} f_i(x_i) \mathrm{d}x_i$$

Hustotu

$$f_i(x_i) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_i, \ldots, x_n) \mathrm{d}x_1 \ldots \mathrm{d}x_{i-1} \mathrm{d}x_{i+1} \ldots \mathrm{d}x_n$$

nazýváme marginální hustotou složky X_i vektoru $Z=(X_1, \ldots, X_n)$.

Složky náhodného vektoru $Z=(X_1, \ldots, X_n)$ jsou stochasticky nezávislé právě když platí jedna z následujících rovností:

$$f(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_i(x_i) \quad \text{nebo} \quad F(x_1, \ldots, x_n) = \prod_{i=1}^{n} F_i(x_i)$$

Spojitý náhodný vektor $Z = (X,Y)$

Příklad: Nechť náhodný vektor (X,Y) má rovnoměrné rozdělení na jednotkovém kruhu. Jsou jeho složky X a Y stochasticky nezávislé náhodné veličiny?

$$f(x,y) = \begin{cases} const, & (x,y) \in K = \{(x,y) : x^2 + y^2 \leq 1\} \\ 0, & \text{jinde} \end{cases}$$

$$const = \frac{1}{\pi}$$

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi}dy = \frac{2\sqrt{1-x^2}}{\pi}, \quad x \in \langle -1, 1 \rangle$$

Spojitý náhodný vektor $Z = (X, Y)$

Charakteristiky náhodného vektoru:

Střední hodnota: $\mathrm{E}(Z) = \mathrm{E}(X, Y) = (\mathrm{E}X,\ \mathrm{E}Y)$

Druhý centrální moment: $\mathrm{E}((Z - EZ)'(Z - EZ)) =$

$$= \mathrm{E}\begin{pmatrix} X - \mathrm{E}X \\ Y - \mathrm{E}Y \end{pmatrix} \cdot (X - \mathrm{E}X, Y - \mathrm{E}Y)$$

$$= \begin{pmatrix} \mathrm{E}(X - \mathrm{E}X)^2 & \mathrm{E}(X - \mathrm{E}X)(Y - \mathrm{E}Y) \\ \mathrm{E}(Y - \mathrm{E}Y)(X - \mathrm{E}X) & \mathrm{E}(Y - \mathrm{E}Y)^2 \end{pmatrix}$$

Označíme $\mathrm{E}(X - \mathrm{E}X)(Y - \mathrm{E}Y) = \mathrm{Cov}(X, Y)$ a potom je

$$D(X, Y) = \begin{pmatrix} \mathrm{Var}(X) & \mathrm{Cov}(X, Y) \\ \mathrm{Cov}(X, Y) & \mathrm{Var}(Y) \end{pmatrix}$$

kovarianční matice vektoru Z.

Spojitý náhodný vektor $Z = (X,Y)$

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

Korelační koeficient: $$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$$

$$\rho(X,Y) \in \langle -1,\ 1 \rangle$$

Kovariance a korelační koeficient jsou míry lineární závislosti - korelovanosti

Označíme $\mathrm{E}(X - \mathrm{E}X)(Y - \mathrm{E}Y) = \mathrm{Cov}(X,Y)$ a potom je

$$D(X,Y) = \begin{pmatrix} \mathrm{Var}(X) & \mathrm{Cov}(X,Y) \\ \mathrm{Cov}(X,Y) & \mathrm{Var}(Y) \end{pmatrix}$$

kovarianční matice vektoru Z.

Spojitý náhodný vektor $Z = (X,Y)$

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

Korelační koeficient: $$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$$

$$\rho(X,Y) \in \langle -1,\ 1\rangle$$

Kovariance a korelační koeficient jsou míry lineární závislosti - korelovanosti

Pozor! Nekorelovanost není totéž, co stochastická nezávislost !

$\mathrm{Cov}(X,Y) = 0 \ \Rightarrow\ X$ a Y jsou nekorelované (ale mohou být stochasticky závislé !

X a Y jsou stochasticky nezávislé $\ \Rightarrow\ \mathrm{Cov}(X,Y) = 0$

$\mathrm{Cov}(X,Y) \neq 0 \Rightarrow\ X$ a Y jsou stochasticky závislé

Spojitý náhodný vektor $Z = (X,Y)$

Příklad: Nechť náhodný vektor (X,Y) má rovnoměrné rozdělení na jednotkovém kruhu. Jsou jeho složky X a Y stochasticky nezávislé náhodné veličiny?

$$f(x,y) = \begin{cases} const, & (x,y) \in K = \{(x,y) : x^2 + y^2 \leq 1\} \\ 0, & \text{jinde} \end{cases}$$

$$const = \frac{1}{\pi}$$

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi}dy = \frac{2\sqrt{1-x^2}}{\pi}, \quad x \in \langle -1, 1 \rangle$$

$$\mathrm{E}(X,Y) = (\mathrm{E}X, \mathrm{E}Y) = ? \qquad \mathrm{D}(X,Y) = ? \qquad \rho(X,Y) = ?$$

Spojitý náhodný vektor $Z = (X,Y)$

Příklad: Nechť náhodný vektor (X,Y) má rovnoměrné rozdělení na jednotkovém kruhu. Jsou jeho složky X a Y stochasticky nezávislé náhodné veličiny?

$$\mathrm{E}X = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f(x,y)\mathrm{d}x\mathrm{d}y = \iint_K \frac{x}{\pi}\mathrm{d}x\mathrm{d}y = \int_{-1}^{1} x \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi}\mathrm{d}y\mathrm{d}x =$$

$$= \int_{-1}^{1} x f_X(x)\mathrm{d}x = \int_{-1}^{1} \frac{2x\sqrt{1-x^2}}{\pi}\mathrm{d}x = 0 \qquad \Rightarrow \ \mathrm{E}(X,Y) = (0,0)$$

$$\mathrm{Cov}(X,Y) = \mathrm{E}XY = \iint_K \frac{xy}{\pi}\mathrm{d}x\mathrm{d}y = 0 \qquad \Rightarrow \ \rho(X,Y) = 0$$

Spojitý náhodný vektor $Z = (X_1,\dots,X_n)$

Charakteristiky náhodného vektoru:

Střední hodnota: $E(Z) = E(X_1,\dots, X_n) = (EX_1, \dots , EX_n)$

Kovarianční matice:

$$D(X_1,\dots,X_n) = \begin{pmatrix} \mathrm{Var}(X_1) & \mathrm{Cov}(X_1,X_2) & \cdots & \mathrm{Cov}(X_1,X_n) \\ \mathrm{Cov}(X_1,X_2) & \mathrm{Var}(X_2) & \cdots & \mathrm{Cov}(X_2,X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(X_1,X_n) & \mathrm{Cov}(X_2,X_n) & \cdots & \mathrm{Var}(X_n) \end{pmatrix}$$

Korelační matice:

$$R(X_1,\dots,X_n) = \begin{pmatrix} 1 & \rho(X_1,X_2) & \cdots & \rho(X_1,X_n) \\ \rho(X_1,X_2) & 1 & \cdots & \rho(X_2,X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(X_1,X_n) & \rho(X_2,X_n) & \cdots & 1 \end{pmatrix}$$

Spojitý náhodný vektor $Z = (X,Y)$

$$P(X \in A,\ Y \in B) = \iint_{A\times B} f(x,y)\mathrm{d}x\mathrm{d}y$$

$$= P(X \in A \,|\, Y \in B).P(Y \in B) = \int_B \left[\int_A f(x|y)\mathrm{d}x\right] f_Y(y)\mathrm{d}y$$

Podmíněné rozdělení $f(y|x) = \dfrac{f(x,y)}{f_Y(y)}$

$$P(X \le x \,|\, Y = y) = F(x|y) = \int_{-\infty}^{x} f(t|y)\mathrm{d}t$$

$$P(X \le x \,|\, Y \in B) = F(x|Y \in B) = \int_B \int_{-\infty}^{x} f(t|y) f_Y(y)\mathrm{d}t\mathrm{d}y$$

Podmíněná střední hodnota

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

Spojitý náhodný vektor $Z = (X, Y)$

Podmíněná střední hodnota

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

Vlastnosti podmíněné střední hodnoty:

- $E(X\,|\,Y)$
- $E(k\,|\,Y) = k$
- $E(E(X\,|\,Y)) = E(X)$
- $E(a.h(X) + b.g(X)\,|\,Y)) = a.E(h(X)\,|\,Y) + b.E(g(X)\,|\,Y)$
- $E(\phi(Y).h(X)\,|\,Y)) = \phi(Y).E(h(X)\,|\,Y)$

Podmíněný rozptyl:

$$Var(X\,|\,Y) = E(X^2|\,Y) - (E(X\,|\,Y))^2$$