# M3121 Pravděpodobnost a statistika I

# 8. Marginální náhodné vektory, nezávislost

Jan Koláček (kolacek@math.muni.cz)

Ústav matematiky a statistiky, Přírodovědecká fakulta, Masarykova univerzita, Brno

# Motivační příklad

## Příklad 1 (Chodec a semafor)

*Roztržitý profesor přechází silnici aniž by sledoval semafory na křižovatce. Pravděpodobnost, že bude zasažen autem je $0,01$, pokud svítí červená, $0,1$ pokud svítí oranžová a $0,8$ pokud svítí zelená. Zelená na semaforu svítí $20\%$ času, oranžová $10\%$ a červená $70\%$. Určete pravděpodobnostní rozdělení náhodné veličiny, která značí zasažení chodce autem.*

$S$ ... barva na semaforu, $S \in \{0,1,2\}$, $0 =$ 🔴, $1 =$ 🟠, $2 =$ 🟢
$Z$ ... zasažení chodce autem, $Z \in \{0,1\}$, $0 =$ „ne“, $1 =$ „ano“

$P(Z|S) \Rightarrow$

| $Z$ \ $S$ | 🔴 | 🟠 | 🟢 |
|---|---|---|---|
| „ano“ | $0,01$ | $0,1$ | $0,8$ |
| „ne“ | $0,99$ | $0,9$ | $0,2$ |

$P(S) \Rightarrow$

| $S$ | 🔴 | 🟠 | 🟢 |
|---|---|---|---|
| $P(S)$ | $0,7$ | $0,1$ | $0,2$ |

$P(Z,S) = P(Z|S)P(S) \Rightarrow$

| $Z$ \ $S$ | 🔴 | 🟠 | 🟢 | $P(Z)$ |
|---|---|---|---|---|
| „ano“ | $0,007$ | $0,01$ | $0,16$ | $0,177$ |
| „ne“ | $0,693$ | $0,09$ | $0,04$ | $0,823$ |
| $P(S)$ | $0,7$ | $0,1$ | $0,2$ | $1$ |

$\sum_S P(Z,S)$

# Marginální náhodné vektory

## Definice 1

Nechť $\mathbf{X} = (X_1, \ldots, X_n)'$ je náhodný vektor. Zvolme přirozené $k < n$ a libovolnou $k$-tici indexů $\{i_1, \ldots, i_k\} \subset \{1, \ldots, n\}$, pak $\mathbf{X}^* = \left(X_{i_1}, \ldots, X_{i_k}\right)'$ nazveme **marginálním náhodným vektorem**.

## Věta 2 (obecně)

*Všechna marginální rozdělení náhodného vektoru* $\mathbf{X} = (X_1, \ldots, X_n)'$ *jsou jednoznačně určena rozdělením náhodného vektoru* $\mathbf{X}$*, přitom pro marginální distribuční funkci* $F^*(\mathbf{x}^*)$ *marginálního náhodného vektoru* $\mathbf{X}^* = \left(X_{i_1}, \ldots, X_{i_k}\right)'$ *platí*

$$F^*(\mathbf{x}^*) = F^*(x_{i_1}, \ldots, x_{i_k}) = \lim_{\substack{x_{j_1} \to \infty \\ \vdots \\ x_{j_{n-k}} \to \infty}} F(x_1, \ldots, x_n),$$

*kde*

$$\{j_1, \ldots, j_{n-k}\} = \{1, \ldots, n\} \setminus \{i_1, \ldots, i_k\}.$$

# Marginální náhodné vektory

**Věta 3 (n=2)**

*Všechna marginální rozdělení náhodného vektoru $(X, Y)$ jsou jednoznačně určena a pro marginální distribuční funkce $F_X(x)$ a $F_Y(y)$ platí*

$$F_X(x) = \lim_{y \to \infty} F(x, y), \qquad F_Y(y) = \lim_{x \to \infty} F(x, y).$$

# Marginální náhodné vektory

## Věta 4 (obecně)

*Pro přirozené $k < n$ mějme indexy $\{i_1, \ldots, i_k\} \subset \{1, \ldots, n\}$ a $\{j_1, \ldots, j_{n-k}\} = \{1, \ldots, n\} \setminus \{i_1, \ldots, i_k\}$.*

- *Nechť $\mathbf{X} \sim (M, p)$. Pak marginální náhodný vektor $\mathbf{X}^*$ má marginální pravděpodobnostní funkci rovnu*

$$p^*(\mathbf{x}^*) = p^*(x_{i_1}, \ldots, x_{i_k}) = P(\mathbf{X}^* = \mathbf{x}^*) = \sum_{x_{j_1} \in M_{j_1}} \cdots \sum_{x_{j_{n-k}} \in M_{j_{n-k}}} p(x_1, \ldots, x_n),$$

  *kde $M = M_1 \times \cdots \times M_n$, přičemž $M_i$ je obor hodnot náhodné veličiny $X_i$, $i = 1, \ldots, n$.*

- *Nechť $\mathbf{X}$ je náhodný vektor absolutně spojitého typu s hustotou $f(\mathbf{x})$. Pak marginální náhodný vektor $\mathbf{X}^*$ má marginální hustotu tvaru*

$$f^*(\mathbf{x}^*) = f^*(x_{i_1}, \ldots, x_{i_k}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_n)\, dx_{j_1} \ldots dx_{j_{n-k}}.$$

# Marginální náhodné vektory

## Věta 5 (n=2)

- *Nechť* $(X, Y) \sim (M, p)$. *Pak marginální náhodné veličiny* $X$ *a* $Y$ *mají marginální pravděpodobnostní funkce*

$$p_X(x) = \sum_{y \in M_Y} p(x, y), \qquad p_Y(y) = \sum_{x \in M_X} p(x, y),$$

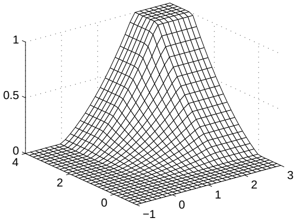
*kde* $M = M_X \times M_Y$.

- *Nechť* $(X, Y) \sim f(x, y)$. *Pak marginální náhodné veličiny* $X$ *a* $Y$ *mají marginální hustoty tvaru*

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

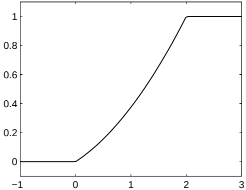
# Příklad

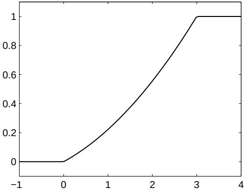
## Sdružená distribuční funkce




## Marginální distribuční funkce

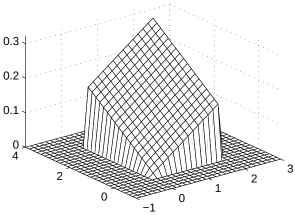
$$F_X(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{4}\left(\frac{x^2}{2} + x\right) & x \in \langle 0, 2 \rangle, \\ 1 & x > 2. \end{cases} \qquad F_Y(y) = \begin{cases} 0 & y < 0 \\ \frac{1}{6}\left(\frac{y^2}{3} + y\right) & y \in \langle 0, 3 \rangle, \\ 1 & y > 3. \end{cases}$$



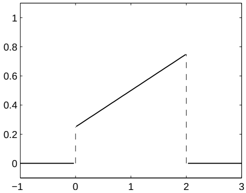

## Sdružená hustota

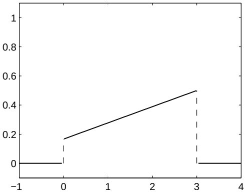
$$f(x,y) = \begin{cases} \frac{1}{6}\left(\frac{x}{2} + \frac{y}{3}\right) & x \in \langle 0, 2 \rangle,\ y \in \langle 0, 3 \rangle, \\ 0 & \text{jinak.} \end{cases}$$




## Marginální hustoty

$$f_X(x) = \begin{cases} \frac{1}{4}(x+1) & x \in \langle 0, 2 \rangle, \\ 0 & \text{jinak.} \end{cases} \qquad f_Y(y) = \begin{cases} \frac{1}{6}\left(\frac{2y}{3} + 1\right) & y \in \langle 0, 3 \rangle, \\ 0 & \text{jinak.} \end{cases}$$



# Příklad

**Příklad 2**

*Nechť* $\mathbf{X} = (X, Y)' \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, *kde*

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \rho \in \langle 0, 1 \rangle.$$

*Určete marginální hustotu náhodné veličiny* $X$.

$$X \sim N(\mu_1, \sigma_1^2)$$

# Motivační příklad

## Příklad 3 (Piráti silnic)

*Na daném úseku měříme rychlost aut. Zaznamenáváme barvu auta a překročení povolené rychlosti. Pozorovali jsme 100 aut, z nichž 30 bylo modrých, 20 zelených a 50 červených. Tabulka uvádí relativní četnosti aut, která překročila povolenou rychlost. Zjistěte, zda překročení rychlosti* **závisí** *na barvě auta.*

| | *modrá* | *zelená* | *červená* |
|---|---|---|---|
| *nepřekročí* | $0,18$ | $0,12$ | $0,3$ |
| *překročí* | $0,12$ | $0,08$ | $0,2$ |

$X$ ... překročení rychlosti, $Y$ ... barva auta
$X$ **nezávisí** na $Y$, pak

$$P(X|Y) \overset{?}{=} P(X) \Leftrightarrow \frac{P(X \cap Y)}{P(Y)} = P(X)$$
$$\Leftrightarrow P(X \cap Y) = P(X)P(Y)$$
$$\Leftrightarrow \boxed{p(x,y) = p_X(x)p_Y(y)}$$

# Příklad

| $X \backslash Y$ | modrá | zelená | červená | $p_X(x)$ |
|---|---|---|---|---|
| nepřekročí | $0,18$ | $0,12$ | $0,3$ | $0,6$ |
| překročí | $0,12$ | $0,08$ | $0,2$ | $0,4$ |
| $p_Y(y)$ | $0,3$ | $0,2$ | $0,5$ | $1$ |

Je třeba ověřit $p(x,y) = p_X(x)p_Y(y)$, tj.

$$
\begin{aligned}
0,18 &= 0,3 \cdot 0,6 \\
0,12 &= 0,2 \cdot 0,6 \\
\vdots \quad & \quad \vdots
\end{aligned}
$$

# Nezávislé náhodné veličiny

## Definice 6

Řekneme, že náhodné veličiny $X_1, \ldots, X_n$ jsou (stochasticky) **nezávislé** (**independent**), jestliže jsou nezávislé náhodné jevy $\{X_1 \leq x_1\}, \ldots, \{X_n \leq x_n\}$ pro libovolné $\mathbf{x} = (x_1, \ldots, x_n)' \in \mathbb{R}^n$.

## Věta 7

*Nechť náhodný vektor* $\mathbf{X} = (X_1, \ldots, X_n)'$ *má sdruženou distribuční funkci* $F(\mathbf{x}) = F(x_1, \ldots, x_n)$ *a nechť pro* $i = 1, \ldots, n$ *je* $F_i(x)$ *marginální distribuční funkce náhodné veličiny* $X_i$*. Pak náhodné veličiny* $X_1, \ldots, X_n$ *jsou (stochasticky)* **nezávislé**, *právě když*

$$F(\mathbf{x}) = F(x_1, \ldots, x_n) = \prod_{i=1}^{n} F_i(x_i) \quad \text{pro} \quad \forall\, \mathbf{x} = (x_1, \ldots, x_n)' \in \mathbb{R}^n.$$

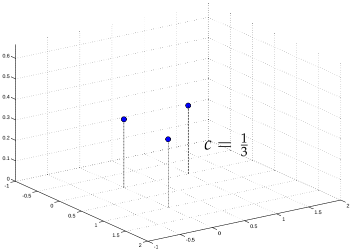
# Nezávislé náhodné veličiny

**Věta 8**

- *Mějme* **diskrétní** *náhodný vektor* $\mathbf{X} = (X_1, \ldots, X_n)' \sim (M, p)$. *Pak* $X_1, \ldots, X_n$ *jsou nezávislé, právě když*
$$p(x_1, \ldots, x_n) = \prod_{i=1}^{n} p_i(x_i) \quad \text{pro} \quad \forall\, \mathbf{x} = (x_1, \ldots, x_n)' \in \mathbb{R}^n,$$
*kde pro* $i = 1, \ldots, n$ *je* $p_i(x_i)$ *marginální pravděpodobnostní funkce náhodné veličiny* $X_i$.
- *Nechť* $\mathbf{X} = (X_1, \ldots, X_n)'$ *je* **absolutně spojitý** *náhodný vektor se sdruženou hustotou* $f(x_1, \ldots, x_n)$. *Pak* $X_1, \ldots, X_n$ *jsou nezávislé, právě když*
$$f(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_i(x_i) \quad \text{pro} \quad s.v.\ \mathbf{x} = (x_1, \ldots, x_n)' \in \mathbb{R}^n,$$
*kde pro* $i = 1, \ldots, n$ *je* $f_i(x_i)$ *marginální hustota náhodné veličiny* $X_i$.

## Příklad 4

*Náhodný vektor* $(X,Y)$ *má rovnoměrné diskrétní rozdělení na množině* $G = \{[0,0];[1,0];[0,1]\}$. *Jsou náhodné veličiny* $X$ *a* $Y$ *nezávislé?*



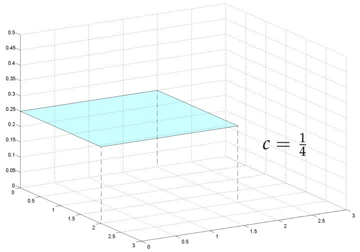

$$p(x,y) = \begin{cases} \frac{1}{3} & \text{pro } (x,y) \in G \\ 0 & \text{jinak} \end{cases}$$

| $X \backslash Y$ | $0$ | $1$ | $p_X(x)$ |
|---|---|---|---|
| $0$ | $1/3$ | $1/3$ | $2/3$ |
| $1$ | $1/3$ | $0$ | $1/3$ |
| $p_Y(y)$ | $2/3$ | $1/3$ | $1$ |

# Příklad

## Příklad 5

*Náhodný vektor* $(X,Y)$ *má rovnoměrné spojité rozdělení na množině* $G = \langle 0;2\rangle \times \langle 0;2\rangle$. *Jsou náhodné veličiny* $X$ *a* $Y$ *nezávislé?*




$$f(x,y) = \begin{cases} \boxed{\frac{1}{4}} & \text{pro } (x,y) \in G \\ 0 & \text{jinak} \end{cases}$$

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_0^2 \frac{1}{4}dy = \boxed{\frac{1}{2}}$$

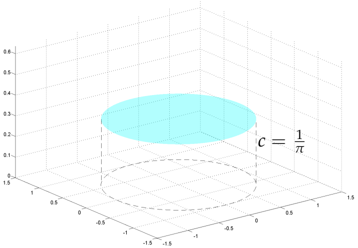
$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx = \int_0^2 \frac{1}{4}dx = \boxed{\frac{1}{2}}$$

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

# Příklad

## Příklad 6

*Náhodný vektor* $(X, Y)$ *má rovnoměrné spojité rozdělení na množině* $G = \{(x,y) \in \mathbb{R}^2; x^2 + y^2 \leq 1\}$. *Jsou náhodné veličiny* $X$ *a* $Y$ *nezávislé?*




$$f(x,y) = \begin{cases} \boxed{\frac{1}{\pi}} & \text{pro } (x,y) \in G \\ 0 & \text{jinak} \end{cases}$$

$$f_X(x) = \int\limits_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi} dy = \boxed{\frac{2}{\pi}\sqrt{1-x^2}}$$

$$f_Y(y) = \int\limits_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \frac{1}{\pi} dx = \boxed{\frac{2}{\pi}\sqrt{1-y^2}}$$

$$f(x,y) \neq f_X(x) f_Y(y)$$