MATIKA IV: CVIČENÍ 10:

nové pojmy:
- směrodatná odchylka $\sqrt{D(X)}$
- kovariance $C(X_1,X_2) = E([X_1-E(X_1)][X_2-E(X_2)])$
- koeficient korelace

pro diskrétní i spojité náhodné veličiny

$$R(X_1,X_2) = E\left(\frac{X_1-E(X_1)}{\sqrt{D(X_1)}}\cdot\frac{X_2-E(X_2)}{\sqrt{D(X_2)}}\right)$$

(je-li $\sqrt{D(X_1)}\cdot\sqrt{D(X_2)} = 0$, pak definujeme $R(X_1,X_2)=0$)

(s příslušným výpočtem

SPOJITÁ X: $E(X) = \int_{-\infty}^{\infty} x\cdot\varphi(x)\,dx$

DISKRÉTNÍ X: $E(X) = \sum_{\forall x} \pi(x)\cdot x$

- transformace náhodného vektoru $(X_1,\dots,X_n)$ na skalární náhodnou veličinu Y:

$$Y = g(X_1,\dots,X_n)$$

- je-li $(X_1,\dots,X_n)$ diskrétní náhodný vektor s pravděpodobnostní funkcí $\pi(x_1,\dots,x_n)$, pak definujeme množinu $S(y) = \{(x_1,\dots,x_n) \mid g(x_1,\dots,x_n) = y\}$
a spočítáme pravděpodobnostní funkci π_* náhodné veličiny Y takto:

$$\underbrace{\pi_*(y)}_{\text{to je } P(Y=y)} = \sum_{(x_1,\dots,x_n)\in S(y)} \pi(x_1,\dots,x_n)$$

- je-li $(X_1,\dots,X_n)$ spojitý náhodný vektor s hustotou pravděpodobnosti $\varphi(x_1,\dots,x_n)$ a distribuční funkcí $\underbrace{\Phi(x_1,\dots,x_n)}_{P(X_1\le x_1,\dots,X_n\le x_n)}$

definujeme $S(y) = \{(x_1,\dots,x_n) \mid g(x_1,\dots,x_n)\le y\}$

a spočítáme distribuční funkci $\Phi_*(y)$ náhodné veličiny Y

$$\Phi_*(y) = \int\cdots\int_{S(y)} \varphi(x_1,\dots,x_n)\,dx_1\dots dx_n$$

a hustotu pravděpodobnosti $\varphi_*(y) = \dfrac{d\Phi_*(y)}{dy}$ (pokud derivace existuje)

- marginální a simultánní vlastnosti náhodného vektoru:

- diskrétní náhodný vektor:

(simultánní) pravděpodobnostní funkce $\pi(x_1,\dots,x_n) = P(X_1=x_1\wedge\dots\wedge X_n=x_n)$

$i\in\{1,\dots,n\}$: marginální pravděpodobnostní funkce náhodné veličiny X_i

$$\pi_i(x_i) = \sum_{\substack{\text{všechny možné}\\ \text{kombinace}\\ x_1,\dots,x_{i-1},x_{i+1},\dots,x_n}} \pi(x_1,\dots,x_{i-1},x_i,x_{i+1},\dots,x_n) = P(X_i=x_i)$$

- spojitý náhodný vektor:

(simultánní) hustota pravděpodobnosti $\varphi(x_1,\dots,x_n)$ – po částech spojitá, nezáporná, normovaná $\left(\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\varphi(x_1,\dots,x_n)\,dx_1\dots dx_n = 1\right)$

distribuční funkce $\Phi(x_1,\dots,x_n)$:

$$\underbrace{\Phi(x_1,\dots,x_n)}_{P(X_1\le x_1,\dots,X_n\le x_n)} = \int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n}\varphi(t_1,\dots,t_n)\,dt_1\dots dt_n$$

naopak: $\varphi(x_1,\dots,x_n) = \dfrac{\partial^n\Phi(x_1,\dots,x_n)}{\partial x_1\dots\partial x_n}$ (n-tá parciální derivace) ve $\forall$ bodech spojitosti fce φ

$i \in \{1,..,n\}$: marginální hustota náhodné veličiny X_i (ze spojitého náhodného vektoru)

$$\varphi_i(x_i) = \underbrace{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty}}_{\substack{\text{integrál} \\ \text{podle} \\ x_1,\dots,x_{i-1},\, x_{i+1},\dots,x_n}} \varphi(x_1,\dots,x_n)\,dx_1\dots dx_{i-1}\,dx_{i+1}\dots dx_n$$

Pravděpodobnost, že se náhodný vektor bude realizovat v oblasti B:

- diskrétní náhodný vektor: $P((x_1,\dots,x_n) \in B) = \sum\limits_{(x_1,\dots,x_n) \in B} \pi(x_1,\dots,x_n)$

- spojitý náhodný vektor: $P((x_1,\dots,x_n) \in B) = \int \dots \int_B \varphi(x_1,\dots,x_n)\,dx_1\dots dx_n$

Př.: Náhodné veličiny A, B mají diskrétní rozdělení určené tabulkou, hledáme marginální rozdělení A a B a korelační koeficient A a B.

B \ A	0	1
3	0,5	0,1
4	0,2	0,2

Řešení:

$\pi_A(0) = P(A=0) = 0,5 + 0,2 = 0,7$
$\pi_A(1) = P(A=1) = 0,1 + 0,2 = 0,3$ } marginální rozdělení A

$\pi_B(3) = P(B=3) = 0,5 + 0,1 = 0,6$
$\pi_B(4) = P(B=4) = 0,2 + 0,2 = 0,4$ } marginální rozdělení B

korelační koeficient A a B:

$$R(A,B) = E\left(\frac{A - E(A)}{\sqrt{D(A)}} \cdot \frac{B - E(B)}{\sqrt{D(B)}}\right)$$

$E(A) = 0 \cdot 0,7 + 1 \cdot 0,3 = 0,3$

$E(B) = 3 \cdot 0,6 + 4 \cdot 0,4 = 3,4$

$D(A) = E((A-0,3)^2) = (0-0,3)^2 \cdot 0,7 + (1-0,3)^2 \cdot 0,3 =$
$= 0,063 + 0,147 = 0,21$

$\sqrt{D(A)} \doteq 0,458$

$D(B) = E((B-3,4)^2) = (3-3,4)^2 \cdot 0,6 + (4-3,4)^2 \cdot 0,4 =$
$= 0,096 + 0,144 = 0,24$

$\sqrt{D(B)} \doteq 0,490$

$$R(A,B) = E\left(\frac{A-0,3}{0,458} \cdot \frac{B-3,4}{0,490}\right) = \frac{1}{0,458 \cdot 0,490} \cdot \Big((0-0,3)(3-3,4) \cdot 0,5$$
$$+ (0-0,3)(4-3,4) \cdot 0,2 + (1-0,3)(3-3,4) \cdot 0,1 + (1-0,3)(4-3,4) \cdot 0,2\Big) =$$
$$\doteq 0,356$$

Př.:

Náhodné veličiny X, Y mají rovnoměrné rozdělení pravděpodobnosti na intervalu $\langle 1,2 \rangle$.
takže jsou spojité

Veličina
$$V = X + 2Y.$$

Hledáme hustotu pravděpodobnosti veličiny V.

Řešení: (Připomenutí k rovnoměrnému rozdělení:
Pravděpodobnost, že X je z intervalu $I \subseteq \langle 1,2 \rangle$ je přímo úměrná délce I. (Děleno délkou $\langle 1,2 \rangle$.)
Pravděpodobnost, že (X,Y) je z množiny $B \subseteq \langle 1,2 \rangle \times \langle 1,2 \rangle$ je přímo úměrná obsahu oblasti B (děleno obsahem $\langle 1,2 \rangle \times \langle 1,2 \rangle$.))

Spočítáme hustotu pravděpodobnosti pro $V = X + 2Y$.
K tomu potřebujeme distribuční funkci Φ_V náhodné veličiny V.
Z návodu na počítání transformovaných náhodných veličin:

$$S(v) = \{(x,y) \mid x + 2y \le v\} = \{(x,y) \mid y \le \tfrac{1}{2}(v - x)\}$$

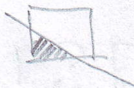
to je polorovina pod přímkou $y = \frac{1}{2}(v-x)$ včetně hraniční přímky.

$$\Phi_V(v) = \iint_{S(v)} \varphi(x,y)\,dx\,dy$$

pro rovnoměrné rozdělení na $\langle 1,2 \rangle \times \langle 1,2 \rangle$ je $\varphi(x,y) = \begin{cases} 1, & \text{je-li } x \in \langle 1,2 \rangle \text{ a } y \in \langle 1,2 \rangle \\ 0 & \text{jinak} \end{cases}$

integrál je tedy roven obsahu útvaru
$$S(v) \cap (\langle 1,2 \rangle \times \langle 1,2 \rangle)$$
(protože integrujeme přes $S(v)$ něco, co je nenulové a $=1$ právě na čtverci $\langle 1,2 \rangle \times \langle 1,2 \rangle$).

$$\Phi_V(v) = \begin{cases} 0 & \text{pro } v \le 3 \\ 1 & \text{pro } v \ge 6 \\ \text{obsah části čtverce pod přímkou } y = \frac{1}{2}(v-x) & \text{pro } v \in \langle 3, 6 \rangle. \end{cases}$$

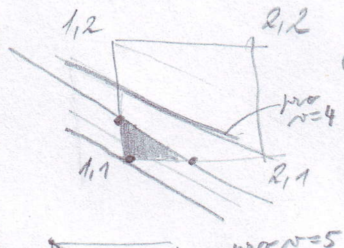


Vzoreček pro obsah:

$$\varphi_V(v) = \frac{d\Phi_V(v)}{dv}$$, kde má deriv. smysl:

$$\varphi_V(v) = \begin{cases} 0 & \text{pro } v < 3 \\ \frac{1}{2}v - \frac{6}{4} & \text{pro } v \in (3,4) \\ \frac{1}{2} & \text{pro } v \in (4,5) \\ -\frac{v}{2} + \frac{11}{4} & \text{pro } v \in (5,6) \\ 0 & \text{pro } v > 6 \end{cases}$$

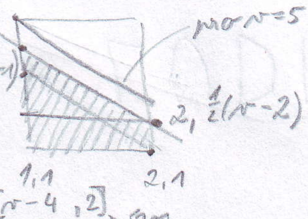
a)



obsah $[1, \frac{1}{2}(v-1)]$, $[1,1]$, $[v-2, 1]$ $= \underline{(v-3)\cdot(v-3)\cdot\frac{1}{4}}$

pro $\underline{v \in \langle 3,4 \rangle}$

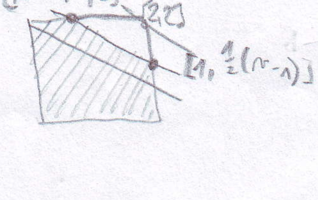
b)



$$1\cdot\left(\tfrac{1}{2}(v-2) - 1\right) + \tfrac{1}{2}\cdot 1 \cdot \tfrac{1}{2} =$$
$$= \underline{\tfrac{1}{2}v - 2 + \tfrac{1}{4}}$$

pro $\underline{v \in \langle 4,5 \rangle}$

c)



$$1 - \tfrac{1}{2}(2-(v-4))\cdot(2 - \tfrac{1}{2}(v-1)) =$$
$$= 1 - \tfrac{1}{2}\cdot\tfrac{1}{2}(6-v)(5-v) =$$
$$= \tfrac{1}{4}(4 - 30 + 11v - v^2) =$$
$$= \underline{\tfrac{1}{4}(-v^2 + 11v - 26)}$$

pro $\underline{v \in \langle 5,6 \rangle}$