Opakování základů teorie pravděpodobnosti

Petr Pošík

Části dokumentu jsou převzaty (i doslovně) z
Mirko Navara: Pravděpodobnost a matematická statistika,
`https://cw.felk.cvut.cz/lib/exe/fetch.php/courses/a6m33ssl/pms_print.pdf`
s laskavým svolením autora.

Jev

Elementární jevy jsou všechny možné vzájemně se vylučující výsledky nějakého experimentu. Jejich množinu označme Ω.

Jev je podmnožina množiny elementárních jevů, $A \subseteq \Omega$.

- Jakýkoli výrok o výsledku experimentu, u něhož lze vždy rozhodnout, zda platí nebo ne (jev nastal nebo nenastal).
- Ekvivalentně lze místo výroků a výrokových operací používat jim příslušné množiny elementárních jevů a množinové operace.

Některé zvláštní jevy a jejich kombinace:

- **Jev jistý:** Ω, **1**
- **Jev nemožný:** $\emptyset$, **0**
- **Konjunkce jevů („and“):** $A \cap B$
- **Disjunkce jevů („or“):** $A \cup B$
- **Jev opačný** k A: $\bar{A} = \Omega \setminus A$
- $A \Rightarrow B$: $A \subseteq B$
- **Jevy neslučitelné**: $A_1, \ldots, A_n : \bigcap_{i \le n} A_i = \emptyset$
- **Jevy po dvou neslučitelné** (=**vzájemně se vylučující**): $A_1, \ldots, A_n : \forall i, j \in \{1, \ldots, n\},\ i \neq j : A_i \cap A_j = \emptyset$

Úplný systém jevů

Úplný systém jevů tvoří jevy B_i, $i \in I$, jestliže jsou po dvou neslučitelné a $\bigcup_{i \in I} B_i = \Omega$.

- Speciální případ pro 2 jevy: $\{C, \bar{C}\}$.

Je-li $\{B_1, \ldots, B_n\}$ *úplný systém jevů*, pak

- $\sum_{i=1}^{n} P(B_i) = 1$ a
- pro libovolný jev A platí

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i)\ .$$

Speciálně pro $\{C, \bar{C}\}$:

$$P(A) = P(A \cap C) + P(A \cap \bar{C})\ .$$

Definice pravděpodobnosti

Klasická **Laplaceova definice pravděpodobnosti** založená na relativních četnostech výskytu jevu trpí mnoha neduhy:

- Platí jen pro n *stejně možných* elementárních jevů. (Stejně možných?)
- Nedovoluje nekonečné množiny jevů, geometrickou pravděpodobnost, ...
- Nedovoluje iracionální hodnoty pravděpodobnosti.

Kolmogorovova definice pravděpodobnosti:

- Množina elementárních jevů Ω může být nekonečná a el. jevy nemusí být stejně pravděpodobné.
- **Jevy** jsou podmnožiny množiny Ω, ale *ne nutně všechny*; tvoří podmnožinu $\mathcal{A} \subseteq \Omega^2$. ($\Omega^2$ je potenční množina, powerset, množina všech podmnožin množiny Ω.)

Chceme být schopni definovat pravděpodobnost pro jakoukoli konjunkci a disjunkci jevů; **jevové pole $\mathcal{A}$** proto nemůže být jakékoli, musí to být **σ-algebra**, tj. musí splňovat podmínky:

1. $\emptyset \in \mathcal{A}$.
2. $A \in \mathcal{A} \Rightarrow \bar{A} \in \mathcal{A}$.
3. $(\forall n \in \mathbb{N} : A_n \in \mathcal{A}) \Rightarrow \bigcup\limits_{n \in \mathbb{N}} A_n \in \mathcal{A}$. (Uzavřenost na *spočetná* sjednocení.)

Borelova σ-algebra je nejmenší σ-algebra podmnožin $\mathbb{R}$, která obsahuje všechny intervaly.

Pravděpodobnost

Pravděpodobnost (=pravděpodobnostní míra) je funkce $P : \mathcal{A} \to \langle 0, 1 \rangle$ splňující podmínky

1. $P(\Omega) = 1$
2. $P\left(\bigcup\limits_{n \in \mathbb{N}} A_n\right) = \sum\limits_{n \in \mathbb{N}} P(A_n)$,
   pokud jsou množiny (=jevy) $A_n, n \in \mathbb{N}$, po dvou neslučitelné (*spočetná aditivita*).

Pravděpodobnostní prostor je trojice $(\Omega, \mathcal{A}, P)$, kde Ω je neprázdná množina, $\mathcal{A}$ je σ-algebra podmnožin množiny Ω a $P : \mathcal{A} \to \langle 0, 1 \rangle$ je pravděpodobnost.

Vlastnosti pravděpodobnosti:

- $P(A) \in \langle 0, 1 \rangle$
- $P(\mathbf{0}) = 0, \qquad P(\mathbf{1}) = 1$
- $P(\bar{A}) = 1 - P(A)$
- $A \subseteq B \Rightarrow P(A) \leq P(B)$
- $A \subseteq B \Rightarrow P(B \setminus A) = P(B) - P(A)$
- $A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$ (*aditivita*)
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Nezávislé jevy

Definice: Jevy A a B jsou **nezávislé**, pokud platí

$$P(A \cap B) = P(A) \cdot P(B).$$

Jsou-li A, B nezávislé, pak

- $P(A \cup B) = P(A) + P(B) - P(A) \cdot P(B)$,
- jsou nezávislé taky dvojice $A, \overline{B}$ a $\overline{A}, B$ a $\overline{A}, \overline{B}$.

Jevy $A_1, \ldots, A_n$ se nazývají **po dvou nezávislé**, jestliže každé dva z nich jsou nezávislé.

Množina jevů $\mathcal{M}$ se nazývá **nezávislá**, jestliže

$$P\left(\bigcap_{A \in \mathcal{K}} A\right) = \prod_{A \in \mathcal{K}} P(A)$$

pro všechny konečné podmnožiny $\mathcal{K} \subseteq \mathcal{M}$.

Podmíněná pravděpodobnost

Podmíněná pravděpodobnost jevu A za podmínky B je definována jako

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \qquad P(B) \neq 0.$$

- $P(A)$ známe z pravděpodobnostního popisu systému. Dostaneme-li navíc informaci o tom, že nastal jev B, podmíněná pravděpodobnost $P(A|B)$ je naše aktualizovaná znalost o pravděpodobnosti jevu A.
- $P(\overline{B}|B) = 0$, což odpovídá naší znalosti, že jev $\overline{B}$ nemůže nastat, když nastal jev B.
- Podm. pravděpodobnost je chápána též jako funkce $P(.|B) : \mathcal{A} \to \langle 0, 1 \rangle$ a je to pravděpodobnost v původním smyslu.

Vlastnosti:

- $P(\Omega|B) = 1$, $P(\emptyset|B) = 0$.
- $B \subseteq A \Rightarrow P(A|B) = 1$.
- $P(A \cap B) = 0 \Rightarrow P(A|B) = 0$.
- Pokud se jevy $A_1, \ldots, A_n$ vzájemně vylučují, pak

  $$P\left(\bigcup_{n \in \mathbb{N}} A_n | B\right) = \sum_{n \in \mathbb{N}} P(A_n|B).$$

- Je-li $P(A|B)$ definována, jsou **jevy A, B nezávislé** právě tehdy, když $P(A|B) = P(A)$.

Bayesova věta

Věta o úplné pravděpodobnosti: Je-li $B_i, i \in I$, (spočetný) úplný systém jevů a $\forall i \in I : P(B_i) \neq 0$, pak pro každý jev A platí

$$P(A) = \sum_{i \in I} P(A|B_i) \cdot P(B_i)$$

Platí:

$$P(A|B) \cdot P(B) = P(A \cap B) = P(B|A) \cdot P(A)$$

$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)}$$

Bayesova věta: Je-li $B_i, i \in I$, (spočetný) úplný systém jevů a $\forall i \in I : P(B_i) \neq 0$, pak pro každý jev $A, P(A) \neq 0$, platí

$$P(B_i|A) = \frac{P(A|B_i) \cdot P(B_i)}{\sum_{j \in I} P(A|B_j) \cdot P(B_j)}.$$

Význam: Pravděpodobnosti $P(A|B_i)$ odhadneme z pokusů nebo modelu, pomocí nich určíme pravděpdobnosti $P(B_i|A)$, které slouží k „optimálnímu“ odhadu, který z jevů B_i nastal.

Problém: Ke stanovení **aposteriorních pravděpodobností** $P(B_i|A)$ potřebujeme znát **apriorní pravděpodobnosti** $P(B_i)$.

Náhodná veličina

Náhodná veličina na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ je *měřitelná* funkce $X : \Omega \to \mathbb{R}$, tj. taková, že pro každý interval I platí

$$X^{-1}(I) = \{\omega \in \Omega : X(\omega) \in I\} \in \mathcal{A}.$$

Rozdělení náhodné veličiny je popsáno praděpodobnostmi

$$P_X(I) = P[X \in I] = P(\{\omega \in \Omega : X(\omega) \in I\})$$

definovanými pro lib. interval I. Funkce P_X je **pravděpodobnostní míra** na Borelově σ-algebře a splňuje

- $P_X(\mathbb{R}) = 1$, $P_X(\emptyset) = 0$,
- $P_X(\bigcup_{n \in \mathbb{N}} I_n) = \sum_{n \in \mathbb{N}} P_X(I_n)$, pokud jsou množiny $I_n, n \in \mathbb{N}$, navzájem disjunktní,
- $P_X(\mathbb{R} \setminus I) = 1 - P_X(I)$,
- jestliže $I \subseteq J$, pak $P_X(I) \leq P_X(J)$ a $P_X(J \setminus I) = P_X(J) - P_X(I)$.

Distribuční funkce náhodné veličiny X je funkce $F_X : \mathbb{R} \to \langle 0, 1 \rangle$ definovaná jako

$$F_X(t) = P[X \in (-\infty, t)] = P[X \leq t] = P_X((-\infty, t\rangle).$$

Distribuční funkce je

- neklesající,
- zprava spojitá,
- $\lim_{t \to -\infty} F_X(t) = 0$, $\lim_{t \to \infty} F_X(t) = 1$.

Náhodný vektor

Náhodný vektor (n-rozměrný) na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ je *měřitelná* funkce $\boldsymbol{X} : \Omega \to \mathbb{R}^n$, tj. taková, že pro každý n-rozměrný interval $\boldsymbol{I}$ platí

$$\boldsymbol{X}^{-1}(\boldsymbol{I}) = \{\omega \in \Omega : \boldsymbol{X}(\omega) \in \boldsymbol{I}\} \in \mathcal{A}.$$

Náhodný vektor lze považovat za vektor náhodných veličin $\boldsymbol{X} = (X_1, \ldots, X_n)$, tj. lze psát $\boldsymbol{X}(\omega) = (X_1(\omega), \ldots, X_n(\omega))$, kde zobrazení $X_k : \Omega \to \mathbb{R}, k = 1, \ldots, n$, jsou náhodné veličiny.

Rozdělení náhodného vektoru je popsáno pravděpodobnostmi

$$P_{\boldsymbol{X}}(I_1 \times \ldots \times I_n) = P[X_1 \in I_1, \ldots, X_n \in I_n] = P(\{\omega \in \Omega : X_1(\omega) \in I_1, \ldots, X_n(\omega) \in I_n\}),$$

kde $I_1, \ldots, I_n$ jsou intervaly v $\mathbb{R}$, tj.

$$P_{\boldsymbol{X}}(\boldsymbol{I}) = P[\boldsymbol{X} \in \boldsymbol{I}] = P\left(\{\omega \in \Omega : \boldsymbol{X}(\omega) \in \boldsymbol{I}\}\right),$$

definovanými pro libovolnou borelovskou množinu $\boldsymbol{I} \in \mathbb{R}^n$.

Distribuční funkce náhodného vektoru $\boldsymbol{X}$ je funkce $F_{\boldsymbol{X}} : \mathbb{R}^n \to \langle 0, 1\rangle$ definovaná jako

$$F_{\boldsymbol{X}}(t_1, \ldots, t_n) = P[X_1 \leq t_1, \ldots, X_n \leq t_n] = P_{\boldsymbol{X}}\left((-\infty, t_1\rangle \times \ldots \times (-\infty, t_n\rangle\right),$$

která má tyto vlastnosti:

- neklesající a zprava spojitá (ve všech proměnných),
- $\lim\limits_{t_1\to\infty,\ldots,t_n\to\infty} F_{\boldsymbol{X}}(t_1, \ldots, t_n) = 1,$
- $\forall k \in \{1, \ldots, n\} \forall t_1, \ldots, t_{k-1}, t_{k+1}, \ldots, t_n : \lim\limits_{t_k\to-\infty} F_{\boldsymbol{X}}(t_1, \ldots, t_n) = 0.$

Nezávislost náhodných veličin

Náh. veličiny $X_1, \ldots, X_n$ jsou **nezávislé**, pokud pro libovolné intervaly $I_1, \ldots, I_n$ platí

$$P[X_1 \in I_1, \ldots, X_n \in I_n] = P[X_1 \in I_1] \cdot \ldots \cdot P[X_n \in I_n] = \prod_{i=1}^{n} P[X_i \in I_i].$$

Ekvivalentně stačí pro všechna $t_1, \ldots, t_n \in \mathbb{R}$ požadovat

$$P[X_1 \leq t_1, \ldots, X_n \leq t_n] = \prod_{i=1}^{n} P[X_i \leq t_i],$$

takže pro sdruženou distribuční funkci *nezávislých* náhodných veličin musí platit

$$F_{\boldsymbol{X}}(t_1, \ldots, t_n) = \prod_{i=1}^{n} F_{X_i}(t_i).$$

Náhodné veličiny $X_1, \ldots, X_n$ jsou **po dvou nezávislé**, pokud jsou každé dvě z nich nezávislé. To je slabší podmínka než **nezávislost všech veličin** $X_1, \ldots, X_n$.

Druhy náhodných veličin

Diskrétní náhodná veličina má *po částech konstantní distribuční funkci*.

Existuje pro ně nejvýše spočetná množina O_X taková, že $P[X \notin O_X] = P_X(\mathbb{R} \setminus O_X) = 0$. Nejmenší taková množina (pokud existuje) je $\Omega_X = \{t \in \mathbb{R} : P_X(\{t\}) \neq 0\} = \{t \in \mathbb{R} : P[X = t] \neq 0\}$.

Popisuje ji pravděpodobnostní funkce $p_X(t) = P_X(\{t\}) = P[X = t]$, která nabývá nenulových hodnot na nejvýše spočetné množině Ω_X a která splňuje

$$\sum_{t \in \mathbb{R}} p_X(t) = \sum_{t \in \Omega_X} p_X(t) = 1.$$

Spojitá náhodná veličina má *spojitou distribuční funkci*.

Absolutně spojité náhodné veličiny jsou ty, které mají **hustotu pravděpodobnosti**, což je nezáporná funkce $f_X : \mathbb{R} \to \langle 0, \infty)$ taková, že

$$F_X(t) = \int_{-\infty}^{t} f_X(u) du.$$

- Hustota splňuje $\int_{-\infty}^{\infty} f_X(u) du = 1$.
- Není určena jednoznačně, lze volit $f_X(t) = \frac{\mathrm{d} F_X(t)}{\mathrm{d} t}$, pokud existuje.
- $P_X(\{t\}) = 0$ pro všechna t.

Kvantilová funkce

Distribuční funkce $F_X(t)$ říká, jak velká část populace má hodnotu proměnné X menší nebo rovnu určitému limitu t (např. kolik procent studentů FEL má vážený průměr 1.5 nebo lepší).

Obráceně se lze ptát, jaká hodnota náhodné veličiny odpovídá určité části populace (např. jaký vážený průměr je třeba, aby se student dostal mezi 5 % nejlepších). Pro určité $\alpha \in \langle 0, 1 \rangle$ tak hledáme takové t, pro které $F_X(t) = \alpha$. Protože ale distribuční funkce může být na nějakém intervalu konstantní, nemusí být hledané t jediné, mohou tvořit omezený interval. Proto:

Kvantilová funkce $q_X : (0, 1) \to \mathbb{R}$ je definována jako

$$q_X(\alpha) = \frac{1}{2}(\sup\{t \in \mathbb{R} : P[X < t] \leq \alpha\} + \inf\{t \in \mathbb{R} : P[X \leq t] \geq \alpha\})$$

- Číslo $q_X(\alpha)$ se nazývá **α-kvantil** náhodné veličiny X.
- **Medián** náhodné veličiny je $q_X(0.5)$.
- **Dolní**, $q_X(0.25)$, a **horní kvartil**, $q_X(0.75)$, dále **decily**, **centily** neboli **percentily**, ...
- q_X je neklesající.
- F_X a q_X jsou navzájem inverzní tam, kde jsou spojité a rostoucí.

Střední hodnota

Střední hodnota náhodné proměnné X se značí $\mathrm{E}\,X$ nebo μ_X a je definována zvlášť pro

- *diskrétní* náhodnou veličinu X:

$$\mathrm{E}\,X = \sum_{t\in\mathbb{R}} t \cdot p_X(t) = \sum_{t\in\Omega_X} t \cdot p_X(t),$$

- *spojitou* náhodnou veličinu Y:

$$\mathrm{E}\,Y = \int_{-\infty}^{\infty} t \cdot f_Y(t)\,\mathrm{d}\,t.$$

Pro oba případy platí vzorec využívající kvantilovou funkci

$$\mathrm{E}\,Z = \int_0^1 q_Z(\alpha)\,\mathrm{d}\,\alpha.$$

Vlastnosti:

- $\mathrm{E}\,r = r$, $\mathrm{E}(\mathrm{E}\,X) = \mathrm{E}\,X$
- $\mathrm{E}(X+Y) = \mathrm{E}\,X + \mathrm{E}\,Y$, $\mathrm{E}(X+r) = \mathrm{E}\,X + r$, $\mathrm{E}(X-Y) = \mathrm{E}\,X - \mathrm{E}\,Y$
- $\mathrm{E}(rX + sY) = r\,\mathrm{E}\,X + s\,\mathrm{E}\,Y$
- Pouze pro *nezávislé* veličiny: $\mathrm{E}(X \cdot Y) = \mathrm{E}\,X \cdot \mathrm{E}\,Y$.

Rozptyl (disperze)

Rozptyl náhodné proměnné X se značí $\mathrm{D}\,X$, σ_X^2, var X, nebo $Var(X)$ a je definován jako

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

$$\mathrm{E}(X^2) = (\mathrm{E}\,X)^2 + \mathrm{D}\,X,$$

nebo také

$$\mathrm{D}\,X = \int_0^1 (q_X(\alpha) - \mathrm{E}\,X)^2\,\mathrm{d}\,\alpha.$$

Vlastnosti:

- $\mathrm{D}\,X \geq 0$
- $\mathrm{D}\,r = 0$
- $\mathrm{D}(X + r) = \mathrm{D}\,X$
- $\mathrm{D}(rX) = r^2\,\mathrm{D}\,X$
- Pouze pro *nezávislé* veličiny: $\mathrm{D}(X+Y) = \mathrm{D}\,X + \mathrm{D}(Y)$, $\mathrm{D}(X-Y) = \mathrm{D}\,X + \mathrm{D}(Y)$.

Směrodatná odchylka

Směrodatná odchylka náhodné proměnné X se značí σ_X, je definována jako

$$\sigma_X = \sqrt{\mathrm{D}\,X} = \sqrt{\mathrm{E}\left((X - \mathrm{E}\,X)^2\right)}$$

a *má stejný fyzikální rozměr* jako původní náhodná veličina (na rozdíl od rozptylu).

Vlastnosti:

- $\sigma_X \geq 0$
- $\sigma_r = 0$
- $\sigma_{X+r} = \sigma_X$
- $\sigma_{rX} = |r|\sigma_X$
- Pouze pro *nezávislé* náhodné veličiny: $\sigma_{X+Y} = \sqrt{\mathrm{D}\,X + \mathrm{D}\,Y} = \sqrt{\sigma_X^2 + \sigma_Y^2}$.

Normovaná náhodná veličina

Normovaná náhodná veličina je taková, která má nulovou střední hodnotu a jednotkový rozptyl:

$$\mathrm{norm}\,X = \frac{X - \mathrm{E}\,X}{\sigma_X},$$

má-li vzorec smysl. Zpětná transformace je

$$X = \mathrm{E}\,X + \sigma_X\,\mathrm{norm}\,X.$$

Diskrétní rozdělení

- **Diracovo:** jediný možný výsledek $r \in \mathbb{R}$.
- **Alternativní (Bernoulliho):** dva možné výsledky (obvykle označené 0 a 1), jeden z nich (1) má pravděpodobnost q.
- **Rovnoměrné:** m možných, stejně pravděpodobných výsledků.
- **Binomické:** počet úspěchů z m nezávislých pokusů, je-li v každém stejná pravděpodobnost úspěchu $q \in \langle 0, 1 \rangle$. Součet m nezávislých alternativních rozdělení.
- **Poissonovo:** limitní případ binomického rozdělení pro $m \to \infty$ při konstantním $mq = \lambda > 0$ (tedy $q \to 0$).
- **Geometrické:** počet úspěchů do prvního neúspěchu, je-li v každém pokusu stejná pravděpodobnost úspěchu $q \in (0, 1)$.
- **Hypergeometrické:** Počet výskytů v m vzorcích vybraných z M objektů, v nichž je celkem K výskytů ($1 \le m \le K \le M$).

Spojitá rozdělení

- **Rovnoměrné $R(a, b)$:** p_X je konstantní na intervalu $\langle a, b \rangle$. Hustota $f_{R(a,b)}$, distribuční funkce $F_{R(a,b)}$.
- **Normální (Gaussovo)**
  - **normované $\mathrm{N}(0, 1)$:** hustota ϕ, distribuční funkce Φ.
  - **obecné $\mathrm{N}(\mu, \sigma^2)$:** hustota $f_{\mathrm{N}(\mu,\sigma^2)}$, distribuční funkce $F_{\mathrm{N}(\mu,\sigma^2)}$.
- **Logaritmickonormální (lognormální) $\mathrm{LN}(\mu, \sigma^2)$:** rozdělení náhodné veličiny $X = e^Y$, kde Y má $\mathrm{N}(\mu, \sigma^2)$.
- **Exponenciální $\mathrm{Ex}(\tau)$:** např. rozdělení času do první poruchy, jestliže (podmíněná) pravděpodobnost poruchy za časový interval $\langle t, t + \delta \rangle$ závisí jen na δ, nikoli na t.

Náhodný vektor 2

Diskrétní náhodný vektor má všechny složky diskrétní.

- Lze jej popsat **sdruženou pravděpodobnostní funkcí** $p_{\boldsymbol{X}} : \mathbb{R}^n \to \langle 0, 1 \rangle$

$$p_{\boldsymbol{X}}(t_1, \ldots, t_n) = P[X_1 = t_1, \ldots, X_N = t_N],$$

  která je nenulová jen ve spočetně mnoha bodech.
- *Diskrétní* náh. veličiny $X_1, \ldots, X_n$ jsou **nezávislé** právě tehdy, když pro všechna $t_1, \ldots, t_n \in \mathbb{R}$ platí

$$p_{\boldsymbol{X}}(t_1, \ldots, t_n) = \prod_{i=1}^{n} p_{X_i}(t_i).$$

Spojitý náhodný vektor má všechny složky spojité.

- Lze jej popsat **sdruženou hustotou pravděpodobnosti**, což je každá nezáporná funkce $f_{\boldsymbol{X}} : \mathbb{R}^n \to \langle 0, \infty)$ taková, že pro všechny $t_1, \ldots, t_n \in \mathbb{R}$

$$F_{\boldsymbol{X}}(t_1, \ldots, t_n) = \int_{-\infty}^{t_1} \ldots \int_{-\infty}^{t_n} f_{\boldsymbol{X}}(u_1, \ldots, u_n) \, \mathrm{d}\, u_1 \ldots \mathrm{d}\, u_n.$$

  Pokud to jde, volíme $f_{\boldsymbol{X}}(u_1, \ldots, u_n) = \frac{\partial}{\partial t_1} \cdots \frac{\partial}{\partial t_n} F_{\boldsymbol{X}}(t_1, \ldots, t_n)$.
- *Spojité* náh. veličiny $X_1, \ldots, X_n$ jsou **nezávislé** právě tehdy, když pro *skoro* všechna $t_1, \ldots, t_n \in \mathbb{R}$ platí

$$f_{\boldsymbol{X}}(t_1, \ldots, t_n) = \prod_{i=1}^{n} f_{X_i}(t_i).$$

Číselné charakteristiky náhodného vektoru

Pro náhodný vektor $\boldsymbol{X} = (X_1, \ldots, X_n)$ definujeme

- **střední hodnotu** $\mathrm{E}\, \boldsymbol{X} = (\mathrm{E}\, X_1, \ldots, \mathrm{E}\, X_n)$,
- **rozptyl** $\mathrm{D}\, \boldsymbol{X} = (\mathrm{D}\, X_1, \ldots, \mathrm{D}\, X_n)$,
- **kovarianční matici**

$$\boldsymbol{\Sigma}_{\boldsymbol{X}} = \begin{pmatrix} \mathrm{D}\, X_1 & \mathrm{cov}(X_1, X_2) & \cdots & \mathrm{cov}(X_1, X_n) \\ \mathrm{cov}(X_2, X_1) & \mathrm{D}\, X_2 & \cdots & \mathrm{cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{cov}(X_n, X_1) & \mathrm{cov}(X_n, X_2) & \cdots & \mathrm{D}\, X_n \end{pmatrix},$$

  která je symetrická, pozitivně semidefinitní a na diagonále má rozptyly $\mathrm{D}\, X_i = \mathrm{cov}(X_i, X_i)$,
- **korelační matici**

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

  která je symetrická, pozitivně semidefinitní a na diagonále má jedničky ($\rho(X_i, X_i) = 1$).

Kovariance a korelace

Kovariance dvou náhodných veličin X, Y je míra toho, jak moc se proměnné X, Y společně mění. Je definována (existují-li rozptyly $\mathrm{D}\,X, \mathrm{D}\,Y$) jako

$$\begin{aligned}\mathrm{cov}(X,Y) &= \mathrm{E}((X - \mathrm{E}\,X)(Y - \mathrm{E}\,Y)) = \\ &= \mathrm{E}(XY) - \mathrm{E}\,X \cdot \mathrm{E}\,Y\end{aligned}$$

Poznámka: Pro rozptyl součtu 2 náhodných veličin (i závislých) platí

$$\mathrm{D}(X+Y) = \mathrm{D}\,X + \mathrm{D}\,Y + 2\,\mathrm{cov}(X,Y).$$

Vlastnosti kovariance:

- $\mathrm{cov}(X,X) = \mathrm{D}\,X$, $\mathrm{cov}(X,-X) = -\mathrm{D}\,X$
- $\mathrm{cov}(X,Y) = \mathrm{cov}(Y,X)$
- $\mathrm{cov}(aX+b, cY+d) = ac\,\mathrm{cov}(X,Y)$
- Pro *nezávislé* náhodné veličiny X, Y je $\mathrm{cov}(X,Y) = 0$.

Korelace (Pearsonův korelační koeficient) dvou náhodných veličin X, Y popisuje sílu lineární závislosti. Definujeme jej jako kovarianci normovaných veličin X, Y, tj.

$$\begin{aligned}\rho(X,Y) &= \mathrm{cov}(\mathrm{norm}\,X, \mathrm{norm}\,Y) = \\ &= \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} = \\ &= \mathrm{E}(\mathrm{norm}\,X \cdot \mathrm{norm}\,Y)\end{aligned}$$

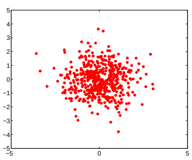
Vlastnosti korelace:

- $\rho(X,Y) \in \langle -1, 1\rangle$
- $\rho(X,X) = 1$, $\rho(X,-X) = -1$
- $\rho(X,Y) = \rho(Y,X)$
- $\rho(aX+b, cY+d) = \mathrm{sign}\,ac\,\rho(X,Y)$
- Pro *nezávislé* náhodné veličiny X, Y je $\rho(X,Y) = 0$.

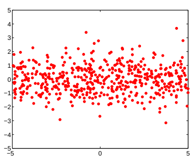
Pokud $\rho(X,Y) = 0$ a $\mathrm{cov}(X,Y) = 0$, *neznamená to*, že veličiny X, Y jsou nezávislé! Takové veličiny nazýváme **nekorelované**.

Korelace: příklady

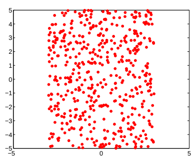
Korelační koeficienty pro náhodné výběry (viz později) ze sdruženého rozdělení náhodných veličin X a Y:



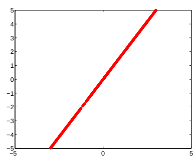
$r = 0$



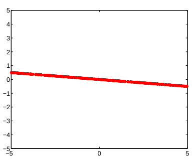
$r = 0$



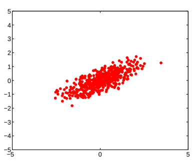
$r = 0$



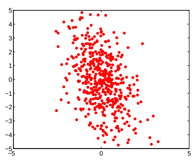
$r = 1$



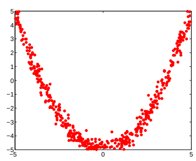
$r = -1$



$r = 0.76$



$r = -0.44$



$r = 0$

A dále nás čeká statistika...

Existují tři druhy lží: lži, naprosté lži a statistiky.

Benjamin Disraeli