# Pravděpodobnost a matematická statistika

LS 2022/23

Michaela Prokešová

# O čem to bude

- budeme matematicky popisovat náhodné jevy
- vyvíjet matematický model náhody (pravděpodobnost)
- zkoumat, jak pro pozorovaná data (výsledek nějakého náhodného procesu) odpovědět na otázky, co nás zajímají (statistika)
- přechod od matematického modelu k vytěžení informace z dat pomocí limitních vět

# Co k tomu budeme potřebovat umět

- matematickou analýzu
- lineární algebru
- teorii míry

# Technickosti

- poznámky k přednášce jsou na mém webu
- také v Moodle kurzu ke cvičení
- co ještě bude v Moodle kurzu
- jak je to s cvičeními
- zápočet:
  - zápočtová písemka středa 12.4.
  - domácí zápočtová písemka (týden 14)
- dotazy a nejasnosti?

# Pravděpodobnostní prostor $(\Omega, \mathcal{A}, \mathrm{P})$

## Definice

Buď $\Omega$ neprázdná množina a $\mathcal{A}$ $\sigma$-algebra na množině $\Omega$.
Pravděpodobnost $\mathrm{P}$ je množinová funkce $\mathrm{P} : \mathcal{A} \to [0, 1]$ splňující

(i) $\mathrm{P}(A) \geq 0, \forall A \in \mathcal{A}$

(ii) $\mathrm{P}(\Omega) = 1$

(iii) jsou-li $A_1, A_2, \dots \in \mathcal{A}$ po dvou disjunktní, pak
$$\mathrm{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathrm{P}(A_i).$$

Trojice $(\Omega, \mathcal{A}, \mathrm{P})$ se nazývá pravděpodobnostní prostor.

# Příklady

- klasický pravděpodobnostní prostor
- diskrétní pravděpodobnostní prostor
- spojitý pravděpodobnostní prostor
- geometrická pravděpodobnost

# Terminologie

Buď $\omega \in \Omega$ přesně ten elementární jev, který nastal.

- $\omega \in A$ ... nastal jev $A$
- $\Omega$ ... jev jistý
- $\emptyset$ ... jev nemožný
- $A \subset B$ ... $A$ je podjev jevu $B$
- $A \cap B$ ... nastaly oba jevy $A$ i $B$
- $A \cup B$ ... nastal alespoň jeden z jevů $A$ a $B$
- $A^c = \Omega \backslash A$ ... jev opačný (doplňkový) k jevu $A$
- $A \cap B = \emptyset$ ... jevy vzájemně neslučitelné

# Základní vlastnosti pravděpodobnosti

## Věta

Buď $(\Omega, \mathcal{A}, \mathrm{P})$ pravděpodobnostní prostor, $A, B \in \mathcal{A}$, $\{A_n\}_{n=1}^{\infty} \subset \mathcal{A}$. Pak platí:

(i) $\mathrm{P}(\emptyset) = 0$

(ii) P je konečně aditivní

(iii) $\mathrm{P}(A^c) = 1 - \mathrm{P}(A)$

(iv) $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$

(v) P je monotónní, tj. $A \subset B \;\Rightarrow\; \mathrm{P}(A) \leq \mathrm{P}(B)$

(vi) je-li $A_1 \subseteq A_2 \subseteq \ldots$, pak $\lim\limits_{n\to\infty} \mathrm{P}(A_n) = \mathrm{P}\left(\bigcup\limits_{i=1}^{\infty} A_i\right)$

(vii) je-li $A_1 \supseteq A_2 \supseteq \ldots$, pak $\lim\limits_{n\to\infty} \mathrm{P}(A_n) = \mathrm{P}\left(\bigcap\limits_{i=1}^{\infty} A_i\right)$

(viii) $B \subset A \;\Rightarrow\; \mathrm{P}(A \backslash B) = \mathrm{P}(A) - \mathrm{P}(B)$

## Věta - pokračování

Buď $(\Omega, \mathcal{A}, \mathrm{P})$ pravděpodobnostní prostor, $A, B \in \mathcal{A}$, $\{A_n\}_{n=1}^{\infty} \subset \mathcal{A}$. Pak platí:

.

.

.

(ix) princip inkluze a exkluze, tj.

$$\mathrm{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathrm{P}(A_i) - \sum_{1 \leq i < j \leq n}\sum \mathrm{P}(A_i \cap A_j) + \sum_{1 \leq i < j < k \leq n}\sum\sum \mathrm{P}(A_i \cap A_j \cap A_k) - \ldots + (-1)^{n-1} \mathrm{P}\left(\bigcap_{i=1}^{n} A_i\right).$$

# Podmíněná pravděpodobnost

## Definice

Buď $(\Omega, \mathcal{A}, \mathrm{P})$ pravděpodobnostní prostor a $A, B \in \mathcal{A}$ splňující $\mathrm{P}(B) > 0$. Podmíněnou pravděpodobnost jevu $A$ za podmínky (jevu) $B$ definujeme vztahem

$$\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}.$$

## Věta

Buď $(\Omega, \mathcal{A}, \mathrm{P})$ pravděpodobnostní prostor a $B \in \mathcal{A}$ takové, že $\mathrm{P}(B) > 0$. Pak zobrazení $\mathrm{P}(\cdot|B) : \mathcal{A} \to [0, 1]$ splňuje definici pravděpodobnosti.

# Vlastnosti podmíněné pravděpodobnosti

## Věta

Buď $(\Omega, \mathcal{A}, \mathrm{P})$ pravděpodobnostní prostor a $B \in \mathcal{A}$ takové, že $\mathrm{P}(B) > 0$. Pro podmíněnou pravděpodobnost platí:

(i) $\mathrm{P}(A \cup C|B) = \mathrm{P}(A|B) + \mathrm{P}(C|B) - \mathrm{P}(A \cap C|B)$

(ii) $B \subset A \;\Rightarrow\; \mathrm{P}(A|B) = 1$

(iii) $A \cap B = \emptyset \;\Rightarrow\; \mathrm{P}(A|B) = 0$

(iv) $\mathrm{P}(A|\Omega) = \mathrm{P}(A)$

(v) pokud $\mathrm{P}(\{\omega\}) > 0$ pro nějaké $\omega \in \Omega$, pak $\forall A \in \mathcal{A}$ platí $\mathrm{P}(A|\{\omega\}) \in \{0, 1\}$.

POZOR! obecně $\mathrm{P}(A|B \cup C) \neq \mathrm{P}(A|B) + \mathrm{P}(A|C)$

## Věta o násobení pravděpodobností

Buď $(\Omega, \mathcal{A}, \mathrm{P})$ pravděpodobnostní prostor, jevy $A_1, \dots A_n \in \mathcal{A}$ takové, že $\mathrm{P}(A_1 \cap \dots \cap A_{n-1}) > 0$. Pak platí

$$\mathrm{P}(A_1 \cap \dots \cap A_n) = \\ \mathrm{P}(A_n | A_1 \cap \dots \cap A_{n-1}) \mathrm{P}(A_{n-1} | A_1 \cap \dots \cap A_{n-2}) \dots \mathrm{P}(A_2 | A_1) \mathrm{P}(A_1).$$

## Věta o celkové pravděpodobnosti

Buď $(\Omega, \mathcal{A}, \mathrm{P})$ pravděpodobnostní prostor, $A \in \mathcal{A}$ a $\{B_n\} \subset \mathcal{A}$ konečná nebo spočetná posloupnost vzájemně neslučitelných jevů splňujících $\mathrm{P}(B_n) > 0 \; \forall n$ a $\mathrm{P}(\bigcup_n B_n) = 1$. Pak platí

$$\mathrm{P}(A) = \sum_n \mathrm{P}(A|B_n)\mathrm{P}(B_n).$$

## Bayesova věta

Buď $(\Omega, \mathcal{A}, \mathrm{P})$ pravděpodobnostní prostor, $A \in \mathcal{A}$ a $\{B_n\} \subset \mathcal{A}$ konečná nebo spočetná posloupnost vzájemně neslučitelných jevů splňujících $\mathrm{P}(B_n) > 0 \; \forall n$ a $\mathrm{P}(\bigcup_n B_n) = 1$. Buď navíc $\mathrm{P}(A) > 0$. Pak platí

$$\mathrm{P}(B_i|A) = \frac{\mathrm{P}(A|B_i)\mathrm{P}(B_i)}{\sum_n \mathrm{P}(A|B_n)\mathrm{P}(B_n)} \quad \text{pro všechna } i.$$

# Nezávislé jevy

## Definice

Náhodné jevy $A$ a $B$ jsou nezávislé, pokud platí $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$.

## Věta

Jsou-li jevy $A$ a $B$ nezávislé, pak také jevy $A$ a $B^c$ jsou nezávislé. Pokud navíc $\mathrm{P}(B) > 0$, pak $\mathrm{P}(A|B) = \mathrm{P}(A)$.

# Vzájemně nezávislé jevy

## Definice

Buď $\{A_l, l \in \Lambda\}$ systém náhodných jevů. Jevy nazveme vzájemně nezávislé, pokud pro každé $n \in \mathbb{N}$ a každou $n$-prvkovou množinu $I \subset \Lambda$ platí

$$\mathrm{P}\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \mathrm{P}(A_i).$$

## Věta

Buď $C = \{B_1, B_2, \dots B_k\}, k \in \mathbb{N}$, systém vzájemně nezávislých jevů. Nahradíme-li libovolnou podmnožinu těchto jevů jejich doplňky, dostaneme opět systém vzájemně nezávislých jevů.

## Věta

Jsou-li jevy $A_1, \dots A_n, B_1, \dots B_m$, $n, m \in \mathbb{N}$ vzájemně nezávislé a $\mathrm{P}(B_1 \cap \dots \cap B_m) > 0$, pak

$$\mathrm{P}(A_1 \cap \dots \cap A_n | B_1 \cap \dots \cap B_m) = \mathrm{P}(A_1 \cap \dots \cap A_n) = \mathrm{P}(A_1) \dots \mathrm{P}(A_n).$$

# Náhodná veličina

## Definice

Buď $(\Omega, \mathcal{A}, \mathrm{P})$ pravděpodobnostní prostor a $(\Omega', \mathcal{A}')$ měřitelný prostor. Každé měřitelné zobrazení $X : \Omega \to \Omega'$ nazveme náhodný element z $\Omega'$.

## Definice

Buď $(\Omega, \mathcal{A}, \mathrm{P})$ pravděpodobnostní prostor. Měřitelné zobrazení $X : (\Omega, \mathcal{A}) \to (\mathbb{R}, \mathcal{B})$ nazveme (reálnou) náhodnou veličinou.

# $\sigma$-algebra indukovaná náhodnou veličinou

## Značení

Buď $A \in \mathcal{B}$. Místo $\{\omega \in \Omega : X(\omega) \in A\}$ budeme psát $\{X \in A\}$.

## Definice

Buď $X$ náhodná veličina. Množinový systém

$$\{X^{-1}(B), B \in \mathcal{B}\}$$

nazveme $\sigma$-algebrou náhodných jevů generovaných náhodnou veličinou $X$ nebo $\sigma$-algebrou indukovanou náhodnou veličinou $X$, a značíme ji $\sigma(X)$.

# Rozdělení náhodné veličiny

## Definice

Rozdělením náhodné veličiny $X : (\Omega, \mathcal{A}) \to (\mathbb{R}, \mathcal{B})$ nazveme indukovanou pravděpodobnostní míru $\mathsf{P}_X$ na $(\mathbb{R}, \mathcal{B})$ definovanou jako

$$\mathsf{P}_X(B) = \mathsf{P}(\{\omega \in \Omega : X(\omega) \in B\}), \ B \in \mathcal{B}.$$

## Definice (TMI)

Buď $(M, \mathcal{A}, \mu)$ prostor s mírou, $g : (M, \mathcal{A}) \to (N, \mathcal{F})$ měřitelné zobrazení. Pak množinová funkce

$$\mu g^{-1} : B \to \mu(g^{-1}(B)), \ B \in \mathcal{F},$$

je míra na $(N, \mathcal{F})$ a nazýváme ji obrazem míry $\mu$ při zobrazení $g$.

# Věta o přenosu integrace

## Věta o přenosu integrace (TMI)

Buď $(M, \mathcal{A}, \mu)$ prostor s mírou, $g : (M, \mathcal{A}) \to (N, \mathcal{F})$ měřitelné zobrazení a $\nu = \mu g^{-1}$ obraz míry $\mu$ v zobrazení $g$. Buď $h : (N, \mathcal{F}) \to (\mathbb{R}, \mathcal{B})$ měřitelná funkce. Pak platí

$$\int_N h(y)\,\mathrm{d}\nu(y) = \int_M h(g(x))\,\mathrm{d}\mu(x)$$

kdykoli má jedna strana smysl.

## Věta o přenosu integrace pro $\mathrm{P}_X$

Buď $X$ náhodná veličina a buď $h : (\mathbb{R}, \mathcal{B}) \to (\mathbb{R}, \mathcal{B})$ měřitelná funkce. Pak platí

$$\int_\Omega h(X(\omega))\,\mathrm{dP}(\omega) = \int_{\mathbb{R}} h(x)\,\mathrm{dP}_X(x),$$

pokud má alespoň jedna strana smysl.

## Radonova-Nikodýmova věta (TMI)

Buďte $\mu, \nu$ $\sigma$-konečné míry na $(M, \mathcal{A})$ a $\nu \ll \mu$. Pak existuje právě jedna (až na modifikace na množinách $\mu$-míry 0) nezáporná měřitelná funkce $f$ na $M$ taková, že

$$\nu(A) = \int_A f \, \mathrm{d}\mu, \ \forall A \in \mathcal{A}.$$

Říkáme, že $f = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}$ je Radonova-Nikodýmova hustota $\nu$ vzhledem k $\mu$. Pokud je $\nu$ konečná, je $f$ integrovatelná.

# Hustota náhodné veličiny

## Definice

Buď $X$ náhodná veličina, $\mathsf{P}_X$ její rozdělení a $\mu$ $\sigma$-konečná míra na $(\mathbb{R}, \mathcal{B})$ taková, že $\mathsf{P}_X \ll \mu$. Potom

$$f_X = \frac{\mathrm{d}\mathsf{P}_X}{\mathrm{d}\mu}$$

se nazývá hustota náhodné veličiny $X$ vzhledem k míře $\mu$.

# Jak počítat $\mathsf{P}(X \in B)$

## Věta

Buď $X$ náhodná veličina, $\mathsf{P}_X$ její rozdělení a $B \in \mathcal{B}$. Pak platí následující rovnosti

$$\begin{aligned} \mathsf{P}(X \in B) &= \mathsf{P}(\{\omega \in \Omega : X(\omega) \in B\}) \\ &= \int_\Omega \mathbf{1}_B(X(\omega))\,\mathrm{d}\mathsf{P}(\omega) = \int_\mathbb{R} \mathbf{1}_B(x)\,\mathrm{d}\mathsf{P}_X(x) = \int_B 1\,\mathrm{d}\mathsf{P}_X(x) = \mathsf{P}_X(B). \end{aligned}$$

Pokud je navíc $f_X$ hustota $X$ vzhledem k $\sigma$-konečné míře $\mu$, pak platí i

$$\mathsf{P}(X \in B) = \int_\mathbb{R} \mathbf{1}_B(x)\,\mathrm{d}\mathsf{P}_X(x) = \int_\mathbb{R} \mathbf{1}_B(x) f_X(x)\,\mathrm{d}\mu(x) = \int_B f_X(x)\,\mathrm{d}\mu(x).$$

# Distribuční funkce náhodné veličiny

## Definice

Buď $X$ náhodná veličina. Funkci $F_X : \mathbb{R} \to [0, 1]$ definovanou jako

$$F_X(x) = \mathrm{P}(X \leq x), \ x \in \mathbb{R},$$

nazveme distribuční funkcí náhodné veličiny $X$.

## Věta o vlastnostech distribuční funkce

Buď $F_X$ distribuční funkce náhodné veličiny $X$. Pak

(i) $F_X$ je neklesající;

(ii) $F_X$ je zprava spojitá;

(iii) $\lim\limits_{x \to -\infty} F_X(x) = 0, \ \lim\limits_{x \to \infty} F_X(x) = 1.$

## Věta (TMI)

Buď $F : \mathbb{R} \to \mathbb{R}$ funkce splňující body (i) – (iii) z věty výše. Pak existuje právě jedna konečná borelovská (Lebesgueova-Stieltjesova) míra $\mu$ na $\mathbb{R}$ splňující

$$\mu((-\infty, a]) = F(a), \ \forall a \in \mathbb{R}.$$

## Věta

Buď $F : \mathbb{R} \to \mathbb{R}$ funkce splňující body (i) – (iii) z věty nahoře. Pak existuje pravděpodobnostní prostor $(\Omega, \mathcal{A}, \mathrm{P})$ a náhodná veličina $X : (\Omega, \mathcal{A}) \to (\mathbb{R}, \mathcal{B})$ taková, že $F = F_X$.

## Definice - Lebesgueův-Stieltjesův integrál (TMI)

Je-li $F$ distribuční funkcí konečné borelovské (Lebesgueovy-Stieltjesovy) míry $\mu$ na $\mathbb{R}$ a $g \in L^1(\mu)$, pak píšeme

$$\int g(x)\,\mathrm{d}F(x) = \int g(x)\,\mathrm{d}\mu(x).$$

Tedy pro distribuční funkci $F_X$ náhodné veličiny $X$ s rozdělením $\mathrm{P}_X$

$$\int g(x)\,\mathrm{d}F_X(x) = \int g(x)\,\mathrm{d}\mathrm{P}_X(x)$$

a speciálně platí

$$\mathrm{P}(X \in B) = \int_{\mathbb{R}} \mathbf{1}_B(x)\,\mathrm{d}\mathrm{P}_X(x) = \int_{\mathbb{R}} \mathbf{1}_B(x)\,\mathrm{d}F_X(x),\ \ B \in \mathcal{B}.$$

# Diskrétní náhodná veličina

## Definice

Náhodnou veličinu $X$ nazveme diskrétní náhodnou veličinou, pokud existuje (konečná nebo spočetná) posloupnost bodů $\{x_i\}_{i \in I} \subset \mathbb{R}$ a posloupnost čísel $\{p_i\}_{i \in I} \subset (0, 1]$, splňujících $\sum_{i \in I} p_i = 1$, takových, že

$$\mathrm{P}_X = \sum_{i \in I} p_i \, \delta_{x_i},$$

kde $\delta_u$ značí Diracovu míru v bodě $u$.

Odpovídající distribuční funkce tvaru

$$F_X(x) = \sum_{i \in I} p_i \mathbf{1}_{[x_i, \infty)}(x) = \sum_{x_i \leq x, i \in I} p_i$$

se pak nazývá diskrétní distribuční funkce.

# Absolutně spojitá náhodná veličina

## Definice

Náhodnou veličinu $X$ nazveme absolutně spojitou náhodnou veličinou, pokud $P_X \ll \lambda$.

Odpovídající distribuční funkci $F_X$ pak nazveme absolutně spojitou distribuční funkcí.

# Kvantilová funkce

## Definice

Buď $F_X$ distribuční funkce náhodné veličiny $X$. Funkce

$$F_X^{-1}(u) = \inf\{x : F_X(x) \geq u\}, \quad u \in (0,1),$$

se nazývá kvantilová funkce náhodné veličiny $X$.

# Střední hodnota

## Definice

Střední hodnota náhodné veličiny $X$ je číslo

$$\mathrm{E}\,X = \int_\Omega X(\omega)\,\mathrm{d}\mathrm{P}(\omega),$$

pokud má tento integrál smysl.

## Značení

Prostor všech reálných náhodných veličin na $(\Omega, \mathcal{A}, \mathrm{P})$ s konečnou střední hodnotou značíme $L^1(\Omega, \mathcal{A}, \mathrm{P}) = L^1$.

## Definice

Medián rozdělení náhodné veličiny $X$ je číslo $q_{\frac{1}{2}}$ splňující

$$\mathrm{P}\left(X \le q_{\frac{1}{2}}\right) \ge \frac{1}{2} \qquad \text{a} \qquad \mathrm{P}\left(X \ge q_{\frac{1}{2}}\right) \ge \frac{1}{2}.$$

## Věta

Buď $X$ náhodná veličina a $g : \mathbb{R} \to \mathbb{R}$ měřitelná funkce. Pak $g(X)$ je také náhodná veličina a platí

$$\mathrm{E}\, g(X) = \int_{\mathbb{R}} g(x)\, \mathrm{d}\mathrm{P}_X(x) = \int_{\mathbb{R}} g(x)\, \mathrm{d}F_X(x),$$

pokud jeden z nich existuje.

Speciálně pro diskrétní náhodnou veličinu

$$\mathrm{E}\, g(X) = \sum_{i \in I} g(x_i) p_i,$$

a pro absolutně spojitou náhodnou veličinu

$$\mathrm{E}\, g(X) = \int_{\mathbb{R}} g(x) f_X(x)\, \mathrm{d}x.$$

## Věta (o vlastnostech střední hodnoty)

Buďte $X$, $Y$ náhodné veličiny a $a, b \in \mathbb{R}$. Pak platí:

(i) $\mathrm{E}\,(a + bX) = a + b\mathrm{E}\,X, \quad$ pro $X \in L^1$,

(ii) $\mathrm{E}\,(X + Y) = \mathrm{E}\,X + \mathrm{E}\,Y, \quad$ pro $X, Y \in L^1$,

(iii) $\mathrm{P}(X \geq 0) = 1 \;\Rightarrow\; \mathrm{E}\,X \geq 0$,

(iv) $X \in L^1 \;\Rightarrow\; |X| \in L^1$,

(v) $X \leq Y$ P-s.j. $\;\Rightarrow\; \mathrm{E}\,X \leq \mathrm{E}\,Y$ (pokud existují).

## Definice

Řekneme, že něco platí P$-$skoro jistě (P-s.j.), pokud to platí pro všechna $\omega \in \Omega \backslash N$, kde $N$ je jev s $\mathrm{P}(N) = 0$.

# Vyšší momenty

## Definice

Buď $n \in \mathbb{N}$. Pak

n-tý moment náhodné veličiny $X$ definujeme jako $\mathrm{E}\,X^n$

n-tý absolutní moment náhodné veličiny $X$ definujeme jako $\mathrm{E}\,|X|^n$

n-tý centrální moment náhodné veličiny $X$ definujeme jako $\mathrm{E}\,(X - \mathrm{E}\,X)^n$, existuje-li $\mathrm{E}\,X$

n-tý absolutní centrální moment náhodné veličiny $X$ definujeme jako $\mathrm{E}\,|X - \mathrm{E}\,X|^n$, existuje-li $\mathrm{E}\,X$.

## Značení

Je-li $n$-tý moment konečný, píšeme $X \in L^n(\Omega, \mathcal{A}, \mathrm{P})$, resp. $X \in L^n$.

# Rozptyl

## Definice

Rozptyl náhodné veličiny $X$ je definován jako

$$\operatorname{var} X = \operatorname{E}(X - \operatorname{E} X)^2.$$

## Věta

Buď $X$ náhodná veličina s konečným rozptylem a $a, b \in \mathbb{R}$. Pak platí

$$\operatorname{var}(a + bX) = b^2 \operatorname{var} X.$$

## Čebyševova nerovnost

Buď $X$ náhodná veličina z $L^1$ a $a > 0$. Pak platí

$$P(|X - E\,X| \geq a) \leq \frac{\operatorname{var} X}{a^2}.$$

## Markovova nerovnost

Buď $X$ náhodná veličina z $L^n, n \in \mathbb{N}$ a $a > 0$. Pak platí

$$P(|X| \geq a) \leq \frac{E\,|X|^n}{a^n}.$$

## Nerovnost mezi $L^p$ normami na pstních prostorech

Buď $X$ náhodná veličina, $0 < \alpha < \beta$ a $\mathsf{E}\,|X|^\beta < \infty$. Pak

$$\left(\mathsf{E}\,|X|^\alpha\right)^{\frac{1}{\alpha}} \leq \left(\mathsf{E}\,|X|^\beta\right)^{\frac{1}{\beta}},$$

a speciálně $\mathsf{E}\,|X| \leq \sqrt{\mathsf{E}\,X^2}$.

# Příklady rozdělení náhodné veličiny

## Alternativní (Bernoullího) rozdělení Alt(p)

Náhodná veličina $X$ s alternativním rozdělením nabývá pouze hodnot 1 a 0 s pravděpodobnostmi $p$ a $1-p$, kde $p \in (0,1)$ se nazývá parametr alternativního rozdělení. Takže

$$\mathrm{P}_X = p\delta_1 + (1-p)\delta_0$$

Pro momenty platí

$$\mathrm{E}\, X^k = p, \;\; k \in \mathbb{N} \qquad\qquad \operatorname{var} X = p\,(1-p)$$

## Binomické rozdělení Binom(n,p)

Náhodná veličina $X$ má binomické rozdělení s parametry $p$, $n$, kde $p \in (0,1)$, $n \in \mathbb{N}$, pokud platí

$$\mathrm{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad k \in \{0, 1, \dots n\}$$

Pro momenty platí

$$\mathrm{E}\, X = np \qquad\qquad \operatorname{var} X = np(1-p)$$

## Geometrické rozdělení Geom(p)

Náhodná veličina $X$ má geometrické rozdělení s parametrem $p$, kde $p \in (0, 1)$, pokud platí

$$\mathrm{P}(X = k) = p(1-p)^k, \quad k \in \mathbb{N}_0$$

Pro momenty platí

$$\mathrm{E}\, X = \frac{(1-p)}{p} \qquad\qquad \operatorname{var} X = \frac{(1-p)}{p^2}$$

## Poissonovo rozdělení Pois($\lambda$)

Náhodná veličina $X$ má *Poissonovo rozdělení s parametrem* $\lambda$, kde $\lambda \in (0, \infty)$, pokud platí

$$\mathsf{P}(X = k) = \mathrm{e}^{-\lambda} \frac{\lambda^k}{k!}, \qquad k \in \mathbb{N}_0$$

Pro momenty platí

$$\mathsf{E}\, X = \operatorname{var} X = \lambda$$

## Rovnoměrné rozdělení R(a,b)

Náhodná veličina $X$ má rovnoměrné rozdělení na intervalu $(a, b)$, kde $a < b \in \mathbb{R}$, pokud má hustotu

$$f_X(x) = \frac{1}{b-a} \mathbf{1}_{(a,b)}(x), \qquad x \in \mathbb{R}$$

tj. distribuční funkci

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,\mathrm{d}t = \begin{cases} 0 & x \le a \\ \frac{x-a}{b-a} & a \le x \le b \\ 1 & x \ge b \end{cases}$$

Pro momenty platí

$$\mathrm{E}\,X = \frac{b+a}{2}, \qquad \operatorname{var} X = \frac{(b-a)^2}{12}$$

## Exponenciální rozdělení Exp($\lambda$)

Náhodná veličina $X$ má exponenciální rozdělení s parametrem $\lambda$, kde $\lambda > 0$, pokud má hustotu

$$f_X(x) = \lambda \mathrm{e}^{-\lambda x}\,\mathbf{1}_{(0,\infty)}(x), \qquad x \in \mathbb{R}$$

tj. distribuční funkci

$$F_X(x) = \begin{cases} 1 - \mathrm{e}^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Pro momenty platí

$$\mathrm{E}\,X = \frac{1}{\lambda} \qquad \operatorname{var} X = \frac{1}{\lambda^2}$$

## Normované normální rozdělení N(0,1)

Náhodná veličina $X$ má normované normální rozdělení, pokud má hustotu

$$f_X(x) = \varphi(x) = \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\frac{x^2}{2}}, \qquad x \in \mathbb{R}$$

a distribuční funkci

$$F_X(x) = \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t, \qquad x \in \mathbb{R}.$$

Pro momenty platí

$$\mathsf{E}\,X^{2k+1} = 0, \quad k \in \mathbb{N}_0 \qquad\qquad \operatorname{var} X = 1$$

## Obecné normální rozdělení $\mathrm{N}(\mu, \sigma^2)$

Náhodná veličina $Y$ má normální rozdělení s parametry $\mu$ a $\sigma^2$, kde $\mu \in \mathbb{R}$ a $\sigma^2 > 0$, pokud má hustotu

$$f_Y(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \, \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad x \in \mathbb{R}$$

### Věta

Buď náhodná veličina $Z \sim N(0,1)$. Buď $Y = aZ + b$, kde $a \neq 0, b \in \mathbb{R}$. Pak $Y$ má normální rozdělení $N(b, a^2)$.

### Důsledky

$$\mathrm{E}\, Y = \mu \qquad \qquad \mathrm{var}\, Y = \sigma^2$$

$$F_Y(t) = \Phi\left(\frac{t-\mu}{\sigma}\right), \qquad t \in \mathbb{R}$$

# Rozdělení funkce náhodné veličiny

## Věta o rozdělení funkce náhodné veličiny

Buď $X$ náhodná veličina s distribuční funkcí $F_X$ a $g : \mathbb{R} \to \mathbb{R}$ měřitelná funkce. Pak $Y = g(X)$ je náhodná veličina s distribuční funkcí

$$F_Y(y) = \int_{\{x : g(x) \leq y\}} \mathrm{d}F_X(x).$$

## Věta o monotónní transformaci

Buď $X$ náhodná veličina s distribuční funkcí $F_X$ a buď $S_X$ nosič rozdělení $\mathsf{P}_X$ (tj. $\mathsf{P}(X \in S_X) = 1$). Buď $g : S_X \to \mathbb{R}$ měřitelná funkce a položme $Y = g(X)$. Pak platí:

(i) pro $g$ ryze rostoucí $F_Y(y) = F_X(g^{-1}(y))$, pro $y \in g(S_X)$,

(ii) pro $g$ ryze klesající $F_Y(y) = 1 - F_X(g^{-1}(y)_-)$, pro $y \in g(S_X)$.

## Důsledek

Buď $X$ absolutně spojitá náhodná veličina s hustotou $f_X$ a $g : S_X \to \mathbb{R}$ ryze monotónní funkce diferencovatelná $\lambda$-s.v. na $S_X$. Pak náhodná veličina $Y = g(X)$ má hustotu

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{\mathrm{d}g^{-1}(y)}{\mathrm{d}y} \right| \mathbf{1}_{g(S_X)}(y),$$

vzhledem k $\lambda$ na $\mathbb{R}$.

# Náhodný vektor

## Definice

Buď $(\Omega, \mathcal{A}, \mathrm{P})$ pravděpodobnostní prostor. Měřitelné zobrazení $\mathbf{X} : (\Omega, \mathcal{A}) \to (\mathbb{R}^n, \mathcal{B}^n)$ nazveme ($n$-rozměrným) náhodným vektorem.

## Definice

Rozdělením náhodného vektoru $\mathbf{X} : (\Omega, \mathcal{A}) \to (\mathbb{R}^n, \mathcal{B}^n)$ nazveme indukovanou pravděpodobnostní míru $\mathrm{P}_{\mathbf{X}}$ na $(\mathbb{R}^n, \mathcal{B}^n)$ definovanou jako

$$\mathrm{P}_{\mathbf{X}}(B) = \mathrm{P}(\{\omega \in \Omega : \mathbf{X}(\omega) \in B\}), B \in \mathcal{B}^n$$

## Definice

(Sdružená) distribuční funkce náhodného vektoru $\mathbf{X}$ je definována jako

$$F_{\mathbf{X}}(\mathbf{x}) = \mathrm{P}\left(\bigcap_{i=1}^{n} \{X_i \leq x_i\}\right), \quad \mathbf{x} = (x_1, \ldots x_n)^{\mathsf{T}} \in \mathbb{R}^n$$

# Marginální rozdělení

## Definice

Rozdělení $\mathrm{P}_{\mathbf{Y}}$ podvektoru $\mathbf{Y} = (X_j)_{j \in J}, J \subset \{1, \ldots, n\} = I$, se nazývá marginální rozdělení.

## Věta o marginální distribuční funkci

Buď $\mathbf{X}$ $n$-rozměrný náhodný vektor s distribuční funkcí $F_{\mathbf{X}}$. Pak pro každé $\mathbf{x} \in \mathbb{R}^n$ platí

$$\lim_{x_n \to \infty} F_{\mathbf{X}}(x_1, \ldots, x_n) = F_{(X_1, \ldots X_{n-1})^{\mathsf{T}}}(x_1, \ldots, x_{n-1}),$$

kde $F_{(X_1, \ldots X_{n-1})^{\mathsf{T}}}$ je distribuční funkce $(n-1)$-rozměrného náhodného vektoru $(X_1, \ldots, X_{n-1})^{\mathsf{T}}$.

## Věta o vlastnostech sdružené distribuční funkce

Distribuční funkce $F_{\mathbf{X}}$ náhodného vektoru $\mathbf{X}$ splňuje:

(i) $\displaystyle\lim_{x_i \to \infty, 1 \le i \le n} F_{\mathbf{X}}(x_1, \dots, x_n) = 1$

(ii) $\displaystyle\lim_{x_j \to -\infty} F_{\mathbf{X}}(x_1, \dots, x_n) = 0, \ \forall j \in \{1, \dots, n\}$

$$\forall x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n \in \mathbb{R}$$

(iii) $F_{\mathbf{X}}$ je zprava spojitá v každé proměnné

(iv) pro $(\mathbf{a}, \mathbf{b}]$ neprázdný interval v $\mathbb{R}^n$ platí

$$\sum_{k=0}^{n} (-1)^k \sum_{\mathbf{c} \in \Delta_{n,k}} F_{\mathbf{X}}(\mathbf{c}) \ge 0.$$

## Značení

Buď $(\mathbf{a}, \mathbf{b}]$ neprázdný interval v $\mathbb{R}^n$, $\Delta_{n,k}$ značíme množinu bodů $\mathbf{c}$ z $\mathbb{R}^n$ takových, že $c_i \in \{a_i, b_i\}$ $\forall i \in \{1, \dots, n\}$ a $|\{i : c_i = a_i\}| = k$ (tedy $\mathbf{c}$ se rovná $\mathbf{a}$ přesně v $k$ souřadnicích).

## Věta

Nechť $F : \mathbb{R}^n \to [0, 1]$ splňuje body (i)–(iv) z věty výše. Pak existuje

## Definice

Náhodný vektor $\mathbf{X}$ má diskrétní rozdělení, pokud existuje (konečná nebo spočetná) posloupnost bodů $\{\mathbf{x}_j\}_{j\in I} \subset \mathbb{R}^n$ a posloupnost čísel $\{p_j\}_{j\in I} \subset (0,1]$, splňujících $\sum_{j\in I} p_j = 1$, takových, že

$$\mathrm{P}_{\mathbf{X}} = \sum_{j\in I} p_j\ \delta_{\mathbf{x}_j}.$$

Odpovídající distribuční funkce $F_{\mathbf{X}}$ se pak nazývá diskrétní distribuční funkce.

## Definice

Náhodný vektor $\mathbf{X}$ má absolutně spojité rozdělení, pokud existuje nezáporná měřitelná funkce $f_{\mathbf{X}}$ splňující

$$F_{\mathbf{X}}(\mathbf{x}) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f_{\mathbf{X}}(t_1,\dots,t_n)\,\mathrm{d}t_n \dots \mathrm{d}t_1, \quad \forall\, \mathbf{x} \in \mathbb{R}^n.$$

Funkce $f_{\mathbf{X}}$ se nazývá hustota rozdělení $\mathbf{X}$. Odpovídající distribuční funkce $F_{\mathbf{X}}$ se pak nazývá absolutně spojitá distribuční funkce.

# Věta o hustotě $P_{\mathbf{X}}$ vzhledem k součinové referenční míře

Buď $P_{\mathbf{X}}$ rozdělení $n$-rozměrného náhodného vektoru $\mathbf{X}$.
Nechť $P_{\mathbf{X}} \ll \nu_1 \otimes \cdots \otimes \nu_n$ (součin $\sigma$-konečných měr na $\mathbb{R}$).
Pak $P_{X_i} \ll \nu_i$, $\forall i \in \{1, \ldots, n\}$ a existují nezáporné měřitelné funkce $f_{\mathbf{X}} : \mathbb{R}^n \to [0, \infty)$, $f_i : \mathbb{R} \to [0, \infty), i \in \{1, \ldots, n\}$ takové, že:

$$P_{\mathbf{X}}((-\infty, \mathbf{x}]) = F_{\mathbf{X}}(\mathbf{x}) = \int_{(-\infty,\mathbf{x}]} f_{\mathbf{X}}(t_1, \ldots, t_n)\, \mathrm{d}(\nu_1 \otimes \cdots \otimes \nu_n)(\mathbf{t}), \quad \forall \mathbf{x} \in \mathbb{R}^n$$

a

$$P_{\mathbf{X}}(B) = \int_B f_{\mathbf{X}}(\mathbf{t})\, \mathrm{d}(\nu_1 \otimes \cdots \otimes \nu_n)(\mathbf{t}), \quad \forall B \in \mathcal{B}^n.$$

Navíc pro každé $i \in \{1, \ldots, n\}$ platí

$$F_{X_i}(x_i) = \int_{-\infty}^{x_i} f_i(t)\, \mathrm{d}\nu_i(t), \quad \forall x_i \in \mathbb{R},$$

kde $f_i(y_i) =$

$$\int_{\mathbb{R}^{n-1}} f_{\mathbf{X}}(y_1, \ldots, y_n)\, \mathrm{d}(\nu_1 \otimes \cdots \otimes \nu_{i-1} \otimes \nu_{i+1} \otimes \cdots \otimes \nu_n)(y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_n)$$

pro $\nu_i$-s.v. $y_i \in \mathbb{R}$.

# Nezávislé náhodné veličiny

## Definice

Buď $\{X_i, i \in I\}$ systém náhodných veličin na $(\Omega, \mathcal{A}, \mathrm{P})$, kde $I \neq \emptyset$ je libovolná indexová množina. Náhodné veličiny $\{X_i, i \in I\}$ nazveme (vzájemně) nezávislé, pokud pro každou konečnou množinu $J \subset I$ platí

$$\mathrm{P}\left(\bigcap_{i \in J} \{X_i \in B_i\}\right) = \prod_{i \in J} \mathrm{P}(X_i \in B_i), \quad \forall B_i \in \mathcal{B}, i \in J$$

## Věta o rozdělení vektoru s nezávislými složkami

Buď $\mathbf{X} = (X_1, \ldots, X_n)^\mathsf{T}$ náhodný vektor. Pak $\{X_i\}_{i=1}^n$ jsou nezávislé náhodné veličiny právě tehdy, když

$$\mathrm{P}_\mathbf{X} = \mathrm{P}_{X_1} \otimes \cdots \otimes \mathrm{P}_{X_n}.$$

## Věta o distribuční funkci vektoru s nezávislými složkami

Buď $\mathbf{X} = (X_1, \dots, X_n)^\mathsf{T}$ náhodný vektor. Pak $\{X_i\}_{i=1}^n$ jsou nezávislé náhodné veličiny právě tehdy, když

$$F_\mathbf{X}(\mathbf{x}) = \prod_{i=1}^n F_{X_i}(x_i), \ \forall \mathbf{x} \in \mathbb{R}^n.$$

## Věta o hustotě náhodného vektoru s nezávislými složkami

Buď $\mathsf{P}_\mathbf{X}$ rozdělení $n$-rozměrného náhodného vektoru $\mathbf{X}$ splňující $\mathsf{P}_\mathbf{X} \ll \nu_1 \otimes \cdots \otimes \nu_n$ (součin $\sigma$-konečných měr na $\mathbb{R}$) a $f_\mathbf{X}$ jeho hustota. Pak náhodné veličiny $\{X_i\}_{i=1}^n$ jsou nezávislé právě tehdy, když

$$f_\mathbf{X}(x_1, \dots, x_n) = \prod_{i=1}^n f_{X_i}(x_i), \quad \text{pro } \nu_1 \otimes \cdots \otimes \nu_n - \text{s.v. } \ \mathbf{x} \in \mathbb{R}^n$$

kde $f_{X_i} = \frac{\mathrm{d}\mathsf{P}_{X_i}}{\mathrm{d}\nu_i}, i \in \{1, \dots, n\}$.

### Věta

Buď $\{X_i, i \in I\}$ systém nezávislých náhodných veličin a $g_i : \mathbb{R} \to \mathbb{R}, i \in I$, měřitelné funkce. Pak $\{g_i(X_i), i \in I\}$ je opět systém nezávislých náhodných veličin.

# Momenty náhodného vektoru

## Značení

Vektor středních hodnot náhodného vektoru $\mathbf{X}$ značíme

$$\mathrm{E}\,\mathbf{X} = (\mathrm{E}\,X_1, \ldots, \mathrm{E}\,X_n)^{\mathsf{T}}$$

## Tvrzení

Buď $\mathbf{X}$ náhodný vektor a $g : \mathbb{R}^n \to \mathbb{R}$ měřitelná funkce. Pak je $g(\mathbf{X})$ náhodná veličina a pro $\mathrm{P}_{\mathbf{X}} \ll \nu_1 \otimes \cdots \otimes \nu_n$ platí

$$\begin{aligned}\mathrm{E}\,g(\mathbf{X}) = \int_{\mathbb{R}} y \,\mathrm{d}\mathrm{P}_{g(\mathbf{X})} = \int_{\mathbb{R}^n} g(\mathbf{x})\,\mathrm{d}\mathrm{P}_{\mathbf{X}}(\mathbf{x}) \\ = \int_{\mathbb{R}^n} g(\mathbf{x})\; f_{\mathbf{X}}(\mathbf{x})\,\mathrm{d}(\nu_1 \otimes \cdots \otimes \nu_n)(\mathbf{x})\end{aligned}$$

## Definice

Buďte $X, Y$ náhodné veličiny definované na tomtéž pravděpodobnostním prostoru. Pak kovariance $X$ a $Y$ je definována jako

$$\operatorname{cov}(X, Y) = \operatorname{E}\left[(X - \operatorname{E}X)(Y - \operatorname{E}Y)\right].$$

Koeficient korelace je definován jako

$$\operatorname{cor}(X, Y) = \frac{\operatorname{cov}(X, Y)}{\sqrt{\operatorname{var}X \operatorname{var}Y}},$$

pokud $\operatorname{var}X > 0$ a $\operatorname{var}Y > 0$.

## Tvrzení

Kovariance splňuje

$$\operatorname{cov}(aX + b, cY + d) = ac\operatorname{cov}(X, Y) \qquad \forall a, b, c, d \in \mathbb{R}$$

## Věta - Hölderova nerovnost

Buďte $X, Y$ náhodné veličiny na tomtéž pravděpodobnostním prostoru a nechť $\mathrm{E}\,|X|^p \leq \infty$, $\mathrm{E}\,|Y|^q \leq \infty$, kde $p, q \geq 1$, $\frac{1}{p} + \frac{1}{q} = 1$. Potom

$$\mathrm{E}\,|XY| \leq \left(\mathrm{E}\,|X|^p\right)^{\frac{1}{p}} \left(\mathrm{E}\,|Y|^q\right)^{\frac{1}{q}},$$

a rovnost nastává právě tehdy, když existují $a, b \in [0, \infty)$ (alespoň jedno z jich nenulové) takové, že $a|X|^p = b|Y|^q$ P-skoro jistě.

## Důsledek

Speciálně pro $p = q = 2$ dostaneme

$$\mathrm{E}\,|XY| \leq \sqrt{\mathrm{E}\,X^2\;\mathrm{E}\,Y^2}$$

takže

$$\begin{aligned} |\mathrm{cov}\,(X, Y)| &\leq \mathrm{E}\,|(X - \mathrm{E}\,X)(Y - \mathrm{E}\,Y)| \\ &\leq \sqrt{\mathrm{E}\,(X - \mathrm{E}\,X)^2\;\mathrm{E}\,(Y - \mathrm{E}\,Y)^2} = \sqrt{\mathrm{var}\,X\;\mathrm{var}\,Y} \end{aligned}$$

a $\mathrm{cor}\,(X, Y) \in [-1, 1]$.

## Věta

Buďte $X_1, X_2$ nezávislé náhodné veličiny a buď $\mathrm{E}\,|X_i| < \infty, i = 1, 2$. Pak $\mathrm{E}\,|X_1 X_2| < \infty$ a platí $\mathrm{E}\,X_1 X_2 = \mathrm{E}\,X_1 \mathrm{E}\,X_2$.

## Věta o cov a var pro nezávislé náhodné veličiny

Buďte $X_1, \dots, X_n$ nezávislé náhodné veličiny a buď $\mathrm{E}\,|X_i| < \infty, i = 1, \dots, n$. Pak

$$\operatorname{cov}(X_i, X_j) = 0, \ \forall i \neq j$$

a pro každé $\mathbf{a} \in \mathbb{R}^n$ platí

$$\operatorname{var}\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 \operatorname{var} X_i.$$

## Definice

Náhodné veličiny $X$ a $Y$ definované na tomtéž pravděpodobnostním prostoru nazveme nekorelované, pokud platí $\operatorname{cov}(X, Y) = 0$.

## Důsledek (důkazu předchozí věty)

Buďte $X_1, \dots, X_n$ náhodné veličiny a nechť $\mathrm{E}\, X_i^2 < \infty$, $\forall i \in \{1, \dots, n\}$. Pak platí rovnost

$$\operatorname{var}\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 \operatorname{var} X_i + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} a_i a_j \operatorname{cov}(X_i, X_j).$$

pro každé $\mathbf{a} \in \mathbb{R}^n$.

# Varianční matice

### Definice

Varianční matice náhodného vektoru $\mathbf{X}$ je matice $n \times n$ s prvky

$$a_{i,j} = \operatorname{cov}(X_i, X_j), \qquad i, j = 1, \ldots, n$$

tj.

$$\operatorname{Var} \mathbf{X} = \operatorname{E}(\mathbf{X} - \operatorname{E}\mathbf{X})(\mathbf{X} - \operatorname{E}\mathbf{X})^{\mathsf{T}}$$

Korelační matice náhodného vektoru $\mathbf{X}$ je matice $n \times n$ s prvky

$$a_{i,j} = \operatorname{cor}(X_i, X_j), \qquad i, j = 1, \ldots, n.$$

## Věta o vlastnostech varianční matice

Buď $\mathbf{X} = (X_1, \dots, X_n)^\mathsf{T}$ náhodný vektor takový, že $\mathrm{E}\, X_i^2 < \infty\ \forall i = 1, \dots, n$. Pak

(i) $\mathrm{Var}\, \mathbf{X}$ je symetrická a positivně semidefinitní

(ii) pro libovolná $\mathbf{a} \in \mathbb{R}^m$ a matici $B$ typu $m \times n$ je

$$\mathrm{Var}\,(\mathbf{a} + B\mathbf{X}) = B\ \mathrm{Var}\, \mathbf{X}\ B^\mathsf{T}$$

(iii) $|\mathrm{cov}\,(X_i, X_j)| \leq \sqrt{\mathrm{var}\, X_i\ \mathrm{var}\, X_j}$
a rovnost nastává právě tehdy, když existují $a, b \in \mathbb{R}$ takové, že $X_i = a + bX_j$ skoro jistě, nebo $X_j = a + bX_i$ skoro jistě

(iv) jsou-li $X_1, \dots, X_n$ vzájemně nezávislé, pak je $\mathrm{Var}\, \mathbf{X}$ diagonální

(v) $\mathrm{Var}\, \mathbf{X}$ je singulární právě tehdy, když existují $a_1, \dots a_n \in \mathbb{R}$, alespoň jedno z nich nenulové, a $k \in \mathbb{R}$ takové, že

$$\sum_{i=1}^{n} a_i X_i = k \quad \text{skoro jistě.}$$

## Věta o momentech výběrového průměru

Buďte $X_1, \dots X_n$ nezávislé (a nebo jen nekorelované) náhodné veličiny a buďte

$$\mathrm{E}\, X_i = \mu \qquad \operatorname{var} X_i = \sigma^2 \qquad \forall i = 1, \dots n,\ \mu \in \mathbb{R},\ \sigma^2 > 0.$$

Pak pro

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

platí

$$\mathrm{E}\, \overline{X}_n = \mu \qquad \text{a} \qquad \operatorname{var} \overline{X}_n = \frac{\sigma^2}{n}.$$

# Rozdělení transformovaného náhodného vektoru

## Věta

Buďte $X, Y$ nezávislé náhodné veličiny a $\Psi : \mathbb{R}^2 \to \mathbb{R}$ měřitelné zobrazení. Pak náhodná veličina $U = \Psi(X, Y)$ má distribuční funkci

$$\begin{aligned} G_U(u) &= \int_{\mathbb{R}} \int_{\{y:\ \Psi(x,y) \le u\}} \mathrm{d}F_Y(y)\,\mathrm{d}F_X(x) \\ &= \int_{\mathbb{R}} \int_{\{y:\ \Psi(x,y) \le u\}} \mathrm{dP}_Y(y)\,\mathrm{dP}_X(x) \\ &= \int_{\mathbb{R}} \int_{\{x:\ \Psi(x,y) \le u\}} \mathrm{d}F_X(x)\,\mathrm{d}F_Y(y) \\ &= \int_{\mathbb{R}} \int_{\{x:\ \Psi(x,y) \le u\}} \mathrm{dP}_X(x)\,\mathrm{dP}_Y(y), \quad u \in \mathbb{R}, \end{aligned}$$

a

$$\mathrm{E}\,U = \int_{\mathbb{R}} \int_{\mathbb{R}} \Psi(x, y)\,\mathrm{dP}_Y(y)\,\mathrm{dP}_X(x),$$

pokud má $\mathrm{E}\,U$ smysl.

## Věta o rozdělení součtu nezávislých náhodných veličin

Buďte $X, Y$ nezávislé náhodné veličiny. Pak jejich součet $U = X + Y$ má distribuční funkci

$$F_U(u) = \int_{\mathbb{R}} F_X(u-y)\,\mathrm{d}F_Y(y) = \int_{\mathbb{R}} F_Y(u-x)\,\mathrm{d}F_X(x),\ u \in \mathbb{R}.$$

$F_U$ se nazývá konvoluce distribučních funkcí $F_X$ a $F_Y$ a značíme ji $F_U = F_X * F_Y$.

## Definice

Buď $\Psi : (x, y) \mapsto (x + y)$. Pravděpodobnostní rozdělení $(\mathrm{P}_X \otimes \mathrm{P}_Y)\Psi^{-1}$ se nazývá konvoluce pravděpodobnostních rozdělení $\mathrm{P}_X$ a $\mathrm{P}_Y$ a značíme ji $\mathrm{P}_U = \mathrm{P}_X * \mathrm{P}_Y$.

## Důsledek věty o rozdělení součtu nezávislých náhodných veličin

Buďte $X$ a $Y$ nezávislé absolutně spojité náhodné veličiny. Pak náhodná veličina $U = X + Y$ je také absolutně spojitá a její hustota je

$$f_U(u) = \int_{\mathbb{R}} f_X(u-y) f_Y(y)\,\mathrm{d}y = \int_{\mathbb{R}} f_Y(u-x) f_X(x)\,\mathrm{d}x, \ \ u \in \mathbb{R}.$$

## Věta

Buďte náhodné veličiny $X$ a $Y$ nezávislé a čítací (to jest $\mathrm{P}(X \in \mathbb{N}_0) = 1 = \mathrm{P}(Y \in \mathbb{N}_0)$). Pak $U = X + Y$ je také čítací a

$$\mathrm{P}(U = u) = \sum_{n=0}^{u} \mathrm{P}(X = n)\mathrm{P}(Y = u - n), \ \ u \in \mathbb{N}_0.$$

## Věta o transformaci hustot

Buď $\mathbf{X}$ $n$-rozměrný absolutně spojitý náhodný vektor s hustotou $f_{\mathbf{X}}$ vzhledem k $\lambda^n$. Buď $S_{\mathbf{X}}$ otevřená taková, že $\mathsf{P}(\mathbf{X} \in S_{\mathbf{X}}) = 1$ a $g : S_{\mathbf{X}} \to \mathbb{R}^n$ difeomorfismus. Pak rozdělení náhodného vektoru $\mathbf{Y} = g(\mathbf{X})$ má vzhledem k $\lambda^n$ hustotu

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(g^{-1}(\mathbf{y})) \left|\mathrm{J}g^{-1}(\mathbf{y})\right| \mathbf{1}_{g(S_{\mathbf{X}})}(\mathbf{y}), \ \mathbf{y} \in \mathbb{R}^n.$$

## Věta o substituci (TMI)

Buď $M \subseteq \mathbb{R}^n$ otevřená a $\phi : M \to \mathbb{R}^n$ difeomorfismus. Buď $h : \phi(M) \to \mathbb{R}$ lebesgueovsky měřitelná funkce a $N \subset \phi(M)$ lebesgueovsky měřitelná množina. Pak

$$\int_N h(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\phi^{-1}(N)} h(\phi(\mathbf{t}))\ |\mathrm{J}\,\phi(\mathbf{t})|\ \mathrm{d}\mathbf{t},$$

pokud má alespoň jedna strana smysl.

## Věta o rozdělení součinu a podílu nezávislých náhodných veličin

Buďte $X$, $Y$, nezávislé náhodné veličiny s hustotami $f_X, f_Y$ vzhledem k $\lambda$ na $\mathbb{R}$. Pak $U = XY$ je také absolutně spojitá náhodná veličina s hustotou

$$f_U(u) = \int_{\mathbb{R}} f_X\left(\frac{u}{x}\right) f_Y(x) \frac{1}{|x|} \mathbf{1}(x \neq 0)\, \mathrm{d}x.$$

Pokud navíc $\mathrm{P}(Y > 0) = 1$, pak $V = \frac{X}{Y}$ je absolutně spojitá náhodná veličina s hustotou

$$f_V(v) = \int_0^\infty f_X(vx) f_Y(x) x\, \mathrm{d}x.$$

# Mnohorozměrné normální rozdělení

## Definice

Buď $\mathbf{Z} = (Z_1, \dots, Z_r)^\mathsf{T}$, $r \in \mathbb{N}$, kde $Z_i$ jsou vzájemně nezávislé a mají $N(0,1)$ rozdělení. Buď $A_{n \times r}$, $n \in \mathbb{N}$, matice a $\mu \in \mathbb{R}^n$ pevný vektor. Náhodný vektor definovaný jako

$$\mathbf{X} = A\mathbf{Z} + \mu$$

má $n$-rozměrné normální rozdělení s parametry $\mu$ a $\Sigma = AA^\mathsf{T}$. Značíme $N_n(\mu, \Sigma)$.

## Důsledky definice $n$-rozměrného normálního rozdělení:

Buďte $\mathbf{X}$ a $\mathbf{Z}$ jako v definici.

(i) Volíme-li v definici $A = \mathbb{I}_r$ jednotková matice a $\mu = \mathbf{0}$ nulový vektor, pak dostaneme, že $\mathbf{Z}$ má rozdělení $N_r(\mathbf{0}, \mathbb{I}_r)$.

(ii) Platí $\mathrm{E}\,\mathbf{X} = \mu$ a $\mathrm{Var}\,\mathbf{X} = \Sigma$.

(iii) Pro $k \in \mathbb{N}$ a matici $B_{k\times n}$ platí, že náhodný vektor $\mathbf{Y} = B\mathbf{X}$ má normální rozdělení $N_k(B\mu, B\Sigma B^{\mathsf{T}})$.

(iv) Speciálně pro vektor $\mathbf{c} = (c_1, \ldots, c_n) \in \mathbb{R}^n$ má náhodný vektor $\mathbf{cX}$ jednorozměrné normální rozdělení $N(\mathbf{c}\mu, \mathbf{c}\Sigma\mathbf{c}^{\mathsf{T}})$.

(v) Marginální rozdělení $X_i$ je $N(\mu_i, \sigma_i^2)$, kde $\sigma_i^2 = \Sigma_{i,i} = \mathrm{var}\,X_i$, $i = 1, \ldots, n$.

(vi) Pokud je $\rho_{ij} = \mathrm{cor}\,(X_i, X_j)$ definována, pak má podvektor $(X_i, X_j)^{\mathsf{T}}$ rozdělení

$$N_2\left(\begin{pmatrix}\mu_i\\ \mu_j\end{pmatrix}, \begin{pmatrix}\sigma_i^2 & \rho_{ij}\sigma_i\sigma_j\\ \rho_{ij}\sigma_i\sigma_j & \sigma_j^2\end{pmatrix}\right).$$

### Věta o hustotě $n$-rozměrného normálního rozdělení

Buď $\mathbf{X}$ náhodný vektor s rozdělením $N_n(\mu, \Sigma)$, kde $\Sigma$ je regulární matice. Pak $\mathsf{P}_{\mathbf{X}} \ll \lambda^n$ a

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\det \Sigma}}\, \mathrm{e}^{-\frac{1}{2}(\mathbf{x}-\mu)^{\mathsf{T}}\Sigma^{-1}(\mathbf{x}-\mu)}, \quad \mathbf{x} \in \mathbb{R}^n.$$

### Tvrzení

Nechť má náhodný vektor $(X, Y)^{\mathsf{T}}$ dvourozměrné normální rozdělení. Pak mají i $X$ a $Y$ (jednorozměrné) normální rozdělení. Pokud je navíc $\operatorname{cor}(X, Y) = 0$, pak jsou $X$ a $Y$ nezávislé.

# Rozdělení odvozená od normálního rozdělení

## Definice

Buďte $X_1, \dots, X_n, n \in \mathbb{N}$ nezávislé náhodné veličiny s rozdělením $N(0,1)$. Pak náhodná veličina

$$Y_n = \sum_{i=1}^{n} X_i^2$$

má $\chi^2$-rozdělení o $n$ stupních volnosti (značíme $\chi_n^2$) s hustotou

$$g_n(y) = \frac{y^{\frac{n}{2}-1}}{2^{\frac{n}{2}}\,\Gamma\left(\frac{n}{2}\right)}\,\mathrm{e}^{-\frac{y}{2}}\,\mathbf{1}(y>0), \quad y \in \mathbb{R}.$$

## Definice

Buďte $X \sim N(0,1)$ a $Y \sim \chi^2_n, n \in \mathbb{N}$, nezávislé náhodné veličiny. Pak náhodná veličina

$$T = \frac{X}{\sqrt{\frac{Y}{n}}}$$

má Studentovo $t_n$-rozdělení (neboli $t$-rozdělení o $n$ stupních volnosti) s hustotou

$$h_n(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\,\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}, \quad t \in \mathbb{R}.$$

Speciálně pro $n = 1$ rozdělení s

$$h_1(t) = \frac{1}{\pi(1+t^2)}, \quad t \in \mathbb{R}$$

a

$$F_T(x) = \frac{1}{\pi}\left(\arctan(x) + \frac{\pi}{2}\right), \quad x \in \mathbb{R}$$

se nazývá Cauchyho rozdělení.

# Limitní věty

## Značení

Buďte $X_1, X_2, \dots, X_n$ náhodné veličiny definované na tomtéž pravděpodobnostním prostoru. Značíme

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

## Značení

Buďte $A_1, A_2, \dots$ náhodné jevy (ze $\sigma$-algebry $\mathcal{A}$). Značíme

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k, \qquad \liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k,$$

což jsou také jevy ze $\sigma$-algebry $\mathcal{A}$.

# 0-1 zákony

## Cantelliho věta

Buďte $A_1, A_2, \ldots$ náhodné jevy (ze $\sigma$-algebry $\mathcal{A}$). Pokud $\sum_{n=1}^{\infty} \mathrm{P}(A_n) < \infty$, pak

$$\mathrm{P}\left(\limsup_{n\to\infty} A_n\right) = 0.$$

## Borelova věta (Borelův 0-1 zákon)

Buďte $A_1, A_2, \ldots$ <u>nezávislé</u> náhodné jevy (ze $\sigma$-algebry $\mathcal{A}$). Pak

$$\sum_{n=1}^{\infty} \mathrm{P}(A_n) < \infty \quad \Leftrightarrow \quad \mathrm{P}\left(\limsup_{n\to\infty} A_n\right) = 0$$

$$\sum_{n=1}^{\infty} \mathrm{P}(A_n) = \infty \quad \Leftrightarrow \quad \mathrm{P}\left(\limsup_{n\to\infty} A_n\right) = 1.$$

# Konvergence posloupnosti náhodných veličin

## Definice

Buďte $Y_1, Y_2, \ldots$ náhodné veličiny na stejném pravděpodobnostním prostoru $(\Omega, \mathcal{A}, \mathrm{P})$. Řekneme, že posloupnost náhodných veličin $\{Y_n\}_{n=1}^{\infty}$ konverguje v pravděpodobnosti k náhodné veličině $Y$ (značíme $Y_n \xrightarrow[n\to\infty]{\mathrm{P}} Y$), pokud

$$\lim_{n\to\infty} \mathrm{P}(|Y_n - Y| > \epsilon) = 0, \quad \forall \epsilon > 0.$$

A řekneme, že posloupnost náhodných veličin $\{Y_n\}_{n=1}^{\infty}$ konverguje skoro jistě k náhodné veličině $Y$ (značíme $Y_n \xrightarrow[n\to\infty]{\mathrm{sj}} Y$), pokud

$$\mathrm{P}\left(\{\omega \in \Omega : \lim_{n\to\infty} Y_n(\omega) = Y(\omega)\}\right) = 1.$$

### Tvrzení (TMI)

Buďte $Y_1, Y_2, \dots$ náhodné veličiny na stejném pravděpodobnostním prostoru. Pak $Y_n \xrightarrow[n\to\infty]{\text{sj}} Y \Rightarrow Y_n \xrightarrow[n\to\infty]{\text{P}} Y$.

### Tvrzení (TMI)

Buď $1 \leq p \leq \infty$ a buďte $Y_1, Y_2, \dots$ a $Y$ náhodné veličiny na stejném pravděpodobnostním prostoru, patřící do $L^p$. Pak

$$Y_n \xrightarrow[n\to\infty]{L^p} Y \Rightarrow Y_n \xrightarrow[n\to\infty]{\text{P}} Y.$$

## Věta o konvergenci v pravděpodobnosti a součtu

Buďte $Y_1, Y_2, \dots$ a $Z_1, Z_2, \dots$ náhodné veličiny na stejném pravděpodobnostním prostoru $(\Omega, \mathcal{A}, \mathrm{P})$. Pak platí

$$Y_n \xrightarrow[n\to\infty]{\mathrm{P}} 0 \quad \text{a} \quad Z_n \xrightarrow[n\to\infty]{\mathrm{P}} 0 \ \Rightarrow \ (Y_n + Z_n) \xrightarrow[n\to\infty]{\mathrm{P}} 0.$$

Buď navíc $\{a_n\}_{n=1}^{\infty}$ reálná omezená posloupnost. Pak platí

$$Y_n \xrightarrow[n\to\infty]{\mathrm{P}} 0 \ \Rightarrow \ a_n Y_n \xrightarrow[n\to\infty]{\mathrm{P}} 0.$$

## Věta o spojité transformaci

Buďte $Y_1, Y_2, \dots$ náhodné veličiny na stejném pravděpodobnostním prostoru $(\Omega, \mathcal{A}, \mathrm{P})$ a $g : \mathbb{R} \to \mathbb{R}$ funkce spojitá na otevřené množině $S_Y$, nosiči rozdělení náhodné veličiny $Y$ (tj. platí $\mathrm{P}(Y \in S_Y) = 1$). Pak

(i) $Y_n \xrightarrow[n\to\infty]{\mathrm{P}} Y \ \Rightarrow \ g(Y_n) \xrightarrow[n\to\infty]{\mathrm{P}} g(Y)$

(ii) $Y_n \xrightarrow[n\to\infty]{\mathrm{sj}} Y \ \Rightarrow \ g(Y_n) \xrightarrow[n\to\infty]{\mathrm{sj}} g(Y)$.

# Slabý zákon velkých čísel

## Čebyševův slabý zákon velkých čísel

Buďte $X_1, X_2, \ldots$ vzájemně nezávislé náhodné veličiny splňující $\operatorname{E} X_n^2 < \infty, \forall n \in \mathbb{N}$. Nechť

$$\lim_{n \to \infty} \left( \frac{1}{n^2} \sum_{i=1}^{n} \operatorname{var} X_i \right) = 0.$$

Potom platí

$$|\overline{X}_n - \operatorname{E} \overline{X}_n| \xrightarrow[n \to \infty]{\mathrm{P}} 0.$$

# Silný zákon velkých čísel

## Terminologie

Řekneme, že posloupnost náhodných veličin $\{X_n\}_{n=1}^{\infty}$ splňuje silný zákon velkých čísel (zkráceně SZVČ), pokud platí $|\overline{X}_n - \operatorname{E} \overline{X}_n| \xrightarrow[n\to\infty]{\text{sj}} 0$.

## Silný zákon velkých čísel pro nestejně rozdělené n.v.

Buď $\{X_n\}_{n=1}^{\infty}$ posloupnost vzájemně nezávislých náhodných veličin, splňujících $\operatorname{E} X_n^2 < \infty \ \forall n \in \mathbb{N}$. Nechť

$$\sum_{n=1}^{\infty} \frac{\operatorname{var} X_n}{n^2} < \infty.$$

Potom platí $|\overline{X}_n - \operatorname{E} \overline{X}_n| \xrightarrow[n\to\infty]{\text{sj}} 0$.

## Kolmogorovova nerovnost

Buďte $X_1, X_2, \dots$ vzájemně nezávislé náhodné veličiny, splňující $\mathrm{E}\, X_n^2 < \infty\ \forall n \in \mathbb{N}$. Označme

$$S_k = \sum_{j=1}^{k} X_j, k \in \mathbb{N}.$$

Pak

$$\mathrm{P}\left(\max_{1\le k\le n} |S_k - \mathrm{E}\, S_k| \ge \epsilon\right) \le \frac{\operatorname{var} S_n}{\epsilon^2} = \frac{\sum_{k=1}^{n} \operatorname{var} X_k}{\epsilon^2}, \quad \forall \epsilon > 0,\ \forall n \in \mathbb{N}.$$

## Silný zákon velkých čísel pro stejně rozdělené n.v., $L_1$ verze

Buď $\{X_n\}_{n=1}^{\infty}$ posloupnost vzájemně nezávislých, stejně rozdělených náhodných veličin. Pak

$$\overline{X}_n \xrightarrow[n\to\infty]{\text{sj}} \mu \ \text{ pro nějaké } \ \mu \in \mathbb{R} \quad \Leftrightarrow \quad \mathrm{E}\,|X_1| < \infty.$$

V takovém případě $\mathrm{E}\,X_1 = \mu$.

## Lemma

Pro libovolnou náhodnou veličinu $X$ platí

$$\sum_{n=0}^{\infty} \mathrm{P}(|X| \geq n) = \sum_{n=0}^{\infty} (n+1)\,\mathrm{P}(n \leq |X| < n+1)$$

a

$$\sum_{n=0}^{\infty} \mathrm{P}(|X| \geq n) - 1 \leq \mathrm{E}\,|X| \leq \sum_{n=0}^{\infty} \mathrm{P}(|X| \geq n).$$

# Centrální limitní věta

## Definice

Buď $\{Y_n\}_{n=1}^{\infty}$ posloupnost náhodných veličin na libovolných pravděpodobnostních prostorech a buď $Y$ náhodná veličina. Řekneme, že náhodné veličiny $Y_n$ konvergují v distribuci k náhodné veličině $Y$, pokud

$$\lim_{n\to\infty} F_{Y_n}(x) = F_Y(x)$$

v každém bodě $x \in \mathbb{R}$, ve kterém je $F_Y$ spojitá.
Značíme $Y_n \xrightarrow[n\to\infty]{\mathrm{d}} Y$.

## Značení

$Y_n \xrightarrow[n\to\infty]{\mathrm{d}} N(0,1)$ znamená, že $Y_n \xrightarrow[n\to\infty]{\mathrm{d}} Y$ a $Y \sim N(0,1)$

## Lemma (charakterizace konvergence v distribuci)

Buď $\{Y_n\}_{n=1}^{\infty}$ posloupnost náhodných veličin. Pak

$$Y_n \xrightarrow[n\to\infty]{\mathrm{d}} Y$$

právě tehdy, když

$$\mathrm{E}\, h(Y_n) \xrightarrow[n\to\infty]{} \mathrm{E}\, h(Y)$$

pro každou spojitou a omezenou funkci $h : \mathbb{R} \to \mathbb{R}$.

## Centrální limitní věta

Buďte $X_1, X_2, \dots$ vzájemně nezávislé, stejně rozdělené náhodné veličiny,

$$\mathsf{E}\, X_i = \mu \in \mathbb{R}, \quad \operatorname{var} X_i = \sigma^2 \in (0, \infty), \qquad \forall i \in \mathbb{N}.$$

Potom pro náhodné veličiny

$$Z_n = \frac{\sum_{i=1}^{n}(X_i - \mu)}{\sqrt{n\sigma^2}}$$

platí $Z_n \xrightarrow[n\to\infty]{\mathrm{d}} Z$, kde $Z \sim N(0, 1)$.

## de Moivreova–Laplaceova centrální limitní věta

Buďte $\{Y_n\}_{n=1}^{\infty}$ náhodné veličiny splňující

$$Y_n \sim Binom(n, p), \qquad 0 < p < 1.$$

Pak

$$\frac{Y_n - np}{\sqrt{np(1-p)}} \xrightarrow[n\to\infty]{\mathrm{d}} N(0, 1).$$

# Odpověď na zvědavý dotaz – netřeba si pamatovat

## Berryho-Esseénova nerovnost

Za předpokladů a značení CLV a pokud $X_1 \in L^3$, tak pro každé $n \in \mathbb{N}$ platí

$$\|F_{Z_n} - \Phi\|_\infty \leq 0.8 \, \frac{\mathrm{E}\left(|X_1 - \mathrm{E}\left(X_1\right)|^3\right)}{\sigma^3} \, \frac{1}{\sqrt{n}}.$$

## Lemma o spojité transformaci a konvergenci v distribuci

Buď $\{X_n\}_{n=1}^{\infty}$ posloupnost náhodných veličin taková, že

$$X_n \xrightarrow[n\to\infty]{\mathrm{d}} X$$

kde $X$ je nějaká náhodná veličina. A buď $g : \mathbb{R} \to \mathbb{R}$ spojitá funkce. Pak pro náhodné veličiny

$$Y_n = g(X_n), \quad n \in \mathbb{N}$$

platí

$$Y_n \xrightarrow[n\to\infty]{\mathrm{d}} Y$$

kde $Y = g(X)$.

## Věta

Buď $\{X_n\}_{n=1}^{\infty}$ posloupnost náhodných veličin definovaných na stejném $(\Omega, \mathcal{A}, \mathrm{P})$. Pak

$$(X_n - X) \xrightarrow[n\to\infty]{\mathrm{P}} 0 \quad \Rightarrow \quad X_n \xrightarrow[n\to\infty]{\mathrm{d}} X.$$

## Cramérova–Sluckého věta

Buďte $\{X_n\}_{n=1}^{\infty}, \{Y_n\}_{n=1}^{\infty}, X$ náhodné veličiny a nechť

$$X_n \xrightarrow[n\to\infty]{\mathrm{d}} X \qquad \text{a} \qquad Y_n \xrightarrow[n\to\infty]{\mathrm{P}} c \in \mathbb{R}$$

Potom

$$(X_n + Y_n) \xrightarrow[n\to\infty]{\mathrm{d}} (X + c)$$

a

$$X_n \cdot Y_n \xrightarrow[n\to\infty]{\mathrm{d}} c\, X.$$

# Mezi pravděpodobností a statistikou

- aneb co lze z pozorování realizace náhodné veličiny zjistit o jejím rozdělení
- aneb jakým způsobem pozorovat náhodné veličiny
- a co s tím mají společného limitní věty

## Definice

Buďte $X_1, \dots, X_n$ vzájemně nezávislé, stejně rozdělené náhodné veličiny. Empirická distribuční funkce je definována předpisem

$$\widehat{F}_n(x,\omega) = \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{(X_k \le x)}(\omega) = \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{(X_k(\omega) \le x)}, \qquad x \in \mathbb{R}, \omega \in \Omega.$$

# Statistika

Jak odpovídat na otázky, které nás zajímají o reálné situaci která zahrnuje náhodu?

Postup:

- vytvoříme pravděpodobnostní model
- na základě pozorování (empirických dat):
  - odhadneme parametry
  - testujeme hypotézy
  - posuzujeme shodu modelu

## Definice

Nechť $n \in \mathbb{N}$ a $X_1, \ldots, X_n$ jsou nezávislé a stejně rozdělené náhodné veličiny (nebo vektory) s rozdělením $\mathrm{P}_X$, resp. s distribuční funkcí $F$.
Pak $X_1, \ldots, X_n$ nazveme náhodný výběr z rozdělení $\mathrm{P}_X$, resp. náhodný výběr z rozdělení s distribuční funkcí $F$.
Číslo $n$ se nazývá rozsah výběru.

## Definice

Modelem pro pozorování $X_1, \ldots, X_n$ rozumíme předem stanovenou množinu pravděpodobnostních rozdělení $\mathcal{F}$, do níž neznámé rozdělení $\mathrm{P}_X$, resp. jeho distribuční funkce $F$, patří.

# Úloha bodového odhadu

## Definice

Borelovsky měřitelné zobrazení $g : \Theta \to \mathbb{R}$ nazveme parametrickou funkcí.

## Definice

Bodový odhad $\varphi_n$ parametrické funkce $g(\theta)$, je borelovsky měřitelné zobrazení $\varphi_n : \mathbb{R}^n \to \mathbb{R}$, jehož předpis nezávisí na $\theta$ (tedy ani na $\mathsf{P}_\theta$ či $F_\theta$) a jehož definiční obor obsahuje obor hodnot $(X_1, \ldots, X_n)$.

## Značení

Zkráceně značíme odhad $\varphi_n(X_1, \ldots, X_n)$ symbolem $\widehat{\theta}_n$.

# Vlastnosti dobrého odhadu

## Definice

Bodový odhad $\varphi_n$ parametrické funkce $g(\theta)$ se nazývá nestranný, pokud $\forall\theta \in \Theta$ platí

$$\mathrm{E}_{\theta}\, \varphi_n(X_1, \dots, X_n) = g(\theta).$$

$\mathrm{E}_{\theta}$ značí střední hodnotu počítanou vzhledem k rozdělení $\mathrm{P}_{\theta} \otimes \cdots \otimes \mathrm{P}_{\theta}$.

## Definice

Posloupnost bodových odhadů $\varphi_n$ parametrické funkce $g(\theta)$ se nazývá silně konzistentní, pokud $\forall\theta \in \Theta$ platí

$$\mathrm{P}(\lim_{n\to\infty} \varphi_n(X_1, \dots, X_n) = g(\theta)) = 1,$$

neboli když $\varphi_n(X_1, \dots, X_n)$ konverguje ke $g(\theta)$ skoro jistě.

# Vlastnosti dobrého odhadu

## Definice

Posloupnost bodových odhadů $\varphi_n$ parametrické funkce $g(\theta)$ se nazývá asymptoticky nestranná, pokud $\forall\theta \in \Theta$ platí

$$\mathsf{E}_{\theta}\, \varphi_n(X_1, \ldots, X_n) \xrightarrow[n\to\infty]{} g(\theta), \quad \forall\theta \in \Theta.$$

# Statistika

## Definice

Statistikou nazveme libovolnou borelovsky měřitelnou funkci $T_n : \mathbb{R}^n \to \mathbb{R}$, jejíž definiční obor obsahuje obor hodnot náhodného výběru $(X_1, \ldots, X_n)$.

## Definice

Statistika

$$\overline{X}_n = \frac{1}{n} \sum_{k=1}^{n} X_k$$

se nazývá výběrový průměr.
A statistika

$$S_n^2 = \frac{1}{n-1} \sum_{k=1}^{n} (X_k - \overline{X}_n)^2$$

se nazývá výběrový rozptyl.

## Věta o výběrovém průměru a rozptylu

Buď $X_1, \ldots, X_n$ náhodný výběr z rozdělení se střední hodnotou $\mu = \mathsf{E}\, X \in \mathbb{R}$ a rozptylem $\sigma^2 = \operatorname{var} X < \infty$.
Pak výběrový průměr je nestranný a konzistentní odhad $\mu = \mathsf{E}\, X$ a výběrový rozptyl je nestranný a konzistentní odhad rozptylu $\sigma^2 = \operatorname{var} X$.

# Konstrukce bodových odhadů

## Značení

Buď $X_1, \ldots, X_n$ náhodný výběr z rozdělení $\mathsf{P}_\theta$. Označme

$$m_r(\theta) = \mathsf{E}\, X^r, \quad r \in \mathbb{N},$$

$r$-tý moment náhodné veličiny $X$ s rozdělením $\mathsf{P}_\theta$ (pokud existuje).
A označme

$$\widehat{m_r} = \frac{1}{n} \sum_{k=1}^{n} X_k^r, \quad r \in \mathbb{N},$$

$r$-tý výběrový moment spočítaný z náhodného výběru $X_1, \ldots, X_n$.

## Tvrzení

Když pro náhodnou veličinu $X$ s rozdělením $\mathsf{P}_\theta$ platí $\mathsf{E}\,|X|^r < \infty$, pak $r$-tý výběrový moment $\widehat{m_r}$ je nestranný a silně konzistentní odhad $m_r(\theta)$.

# Metoda momentů

Momentový odhad $\hat{\theta}$ parametru $\theta$ najdeme jako řešení soustavy momentových odhadovacích rovnic

$$m_r(\theta) = \widehat{m_r}, \qquad r = 1, 2, \ldots, d.$$

Zde $d$ je typicky dimenze $\theta$, a nebo kolik je potřeba, abychom mohli soutavu vyřešit.

# Metoda maximální věrohodnosti

## Definice

Buď $\mathcal{F} = \{\mathrm{P}_\theta, \theta \in \Theta\}$ parametrický model, kde $\mathrm{P}_\theta$ mají hustoty $f_\theta$ vzhledem ke stejné referenční míře $\nu$.
Hustotu $f_\theta(x_1, \dots x_n)$ náhodného výběru $X_1, \dots, X_n$ z $\mathrm{P}_\theta$, nahlíženou jako funkci $\theta$, nazveme věrohodností $L(\theta; \mathbf{x})$ (případně věrohodností pro pozorování $\mathbf{x}$), tj.

$$L(\theta; \mathbf{x}) = f_\theta(x_1, \dots x_n) = \prod_{i=1}^{n} f_\theta(x_i), \quad \mathbf{x} = (x_1, \dots x_n) \in \mathbb{R}^n, \ \theta \in \Theta.$$

$l(\theta; \mathbf{x}) = \log(L(\theta; \mathbf{x}))$ se nazývá logaritmickou věrohodností.

Maximálně věrohodný odhad je definován jako

$$\varphi_n(\mathbf{x}) = \arg\max_{\theta \in \Theta} L(\theta; \mathbf{x}).$$

# Intervalový odhad

## Definice

Buď $\mathsf{P}_\theta$ z modelu $\mathcal{F}$, $\mathbf{X} = (X_1, \ldots, X_n)$ náhodný výběr o rozsahu $n$ z $\mathsf{P}_\theta$ a $\alpha \in (0, 1)$. Intervalovým odhadem parametrické funkce $g(\theta)$ o spolehlivosti $(1 - \alpha)$ nazveme dvojici borelovských funkcí $\eta_L : \mathbb{R}^n \to \mathbb{R}$ a $\eta_U : \mathbb{R}^n \to \mathbb{R}$, jejichž předpis nezávisí na $\theta$, a $\forall \theta \in \Theta$ platí

$$\mathsf{P}_\theta(\eta_L(\mathbf{X}) < g(\theta) < \eta_U(\mathbf{X})) \geq 1 - \alpha.$$

## Definice

Buď $\mathrm{P}_\theta$ z modelu $\mathcal{F}$, $\mathbf{X} = (X_1, \ldots, X_n)$ náhodný výběr o rozsahu $n$ z $\mathrm{P}_\theta$ a $\alpha \in (0,1)$. Dolním intervalovým odhadem parametrické funkce $g(\theta)$ o spolehlivosti $(1-\alpha)$ nazveme borelovskou funkci $\eta_D : \mathbb{R}^n \to \mathbb{R}$, jejíž předpis nezávisí na $\theta$ a $\forall \theta \in \Theta$ platí

$$\mathrm{P}_\theta(\eta_D(\mathbf{X}) < g(\theta)) \geq 1 - \alpha.$$

Horním intervalovým odhadem parametrické funkce $g(\theta)$ o spolehlivosti $(1-\alpha)$ nazveme borelovskou funkci $\eta_H : \mathbb{R}^n \to \mathbb{R}$, jejíž předpis nezávisí na $\theta$ a $\forall \theta \in \Theta$ platí

$$\mathrm{P}_\theta(g(\theta) < \eta_H(\mathbf{X})) \geq 1 - \alpha.$$

## Definice

Buď $F$ distribuční funkce, spojitá a ryze rostoucí na $F^{-1}(0,1)$, a buď $\beta \in (0,1)$. $\beta$-kvantilem rozdělení s distribuční funkcí $F$ nazveme hodnotu $q_\beta = F^{-1}(\beta)$.

## Značení

$\beta$-kvantil normovaného normálního rozdělení značíme $u_\beta$
$\beta$-kvantil $\chi^2$-rozdělení o $n$ stupních volnosti značíme $\chi^2_{\beta,n} = \chi^2_n(\beta)$
$\beta$-kvantil Studentova $t$-rozdělení o $n$ stupních volnosti značíme $t_{\beta,n} = t_n(\beta)$

# Intervalový odhad v normálním modelu

### Věta

Buď $\mathbf{X} = (X_1, \ldots, X_n)$ náhodný výběr z modelu $\mathcal{F} = \{N(\mu, \sigma^2), \mu \in \mathbb{R}\}$, kde $\sigma^2 > 0$ je známá konstanta. Pak

$$\left(\overline{X}_n - u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \overline{X}_n + u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right)$$

je oboustranný intervalový odhad parametru $\mu$ o spolehlivosti $(1 - \alpha)$.

### Věta

Buď $\mathbf{X}$ náhodný vektor s rozdělením $N_n(0, \sigma^2 \mathbb{I}_n)$, kde $\sigma^2 > 0$. Buď $C$ ortonormální matice rozměru $n \times n$ (tj. platí $CC^\mathsf{T} = C^\mathsf{T} C = \mathbb{I}_n$). Položme $\mathbf{Y} = C\mathbf{X}$. Pak $\mathbf{Y} \sim N_n(0, \sigma^2 \mathbb{I}_n)$.
Tedy $\{Y_i\}_{i=1}^n$ jsou vzájemně nezávislé se stejným rozdělením $N(0, \sigma^2)$.

### Věta

Buď $\mathbf{X} = (X_1, \ldots, X_n)$ náhodný výběr z rozdělení $N(\mu, \sigma^2), \mu \in \mathbb{R}$, $\sigma^2 \in \mathbb{R}^+$. Pak platí:

(i) $\overline{X}_n$ a $S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \overline{X}_n)^2$ jsou nezávislé náhodné veličiny,

(ii) náhodná veličina $\frac{n-1}{\sigma^2} S_n^2 = \sum_{i=1}^n \frac{(X_i - \overline{X}_n)^2}{\sigma^2}$ má $\chi^2_{n-1}$ rozdělení,

(iii) náhodná veličina $T_n = \frac{\overline{X}_n - \mu}{S_n} \sqrt{n}$ má Studentovo $t_{n-1}$ rozdělení.

## Věta

Buď $\mathbf{X} = (X_1, \ldots, X_n)$ náhodný výběr z modelu $\mathcal{F} = \{N(\mu, \sigma^2), \mu \in \mathbb{R}, \sigma^2 > 0\}$. Pak

(i)
$$\left(\overline{X}_n - t_{n-1}(1-\alpha/2)\frac{S_n}{\sqrt{n}}, \overline{X}_n + t_{n-1}(1-\alpha/2)\frac{S_n}{\sqrt{n}}\right)$$
je oboustranný intervalový odhad parametru $\mu$ o spolehlivosti $(1-\alpha)$.

(ii)
$$\left(\frac{(n-1)S_n^2}{\chi^2_{n-1}(1-\frac{\alpha}{2})}, \frac{(n-1)S_n^2}{\chi^2_{n-1}(\frac{\alpha}{2})}\right)$$
je oboustranný intervalový odhad parametru $\sigma^2$ o spolehlivosti $(1-\alpha)$.

# Intervalový odhad založený na CLV

## Definice

Buď $\mathsf{P}_\theta$ z modelu $\mathcal{F}$, $\mathbf{X} = (X_1, \ldots, X_n)$ náhodný výběr o rozsahu $n$ z $\mathsf{P}_\theta$ a $\alpha \in (0, 1)$. Intervalovým odhadem parametrické funkce $g(\theta)$ o asymptotické spolehlivosti $(1 - \alpha)$ nazveme dvojici borelovských funkcí $\eta_L : \mathbb{R}^n \to \mathbb{R}$ a $\eta_U : \mathbb{R}^n \to \mathbb{R}$, jejichž předpis nezávisí na $\theta$, a $\forall \theta \in \Theta$ platí

$$\lim_{n\to\infty} \mathsf{P}_\theta(\eta_L(\mathbf{X}) < g(\theta) < \eta_U(\mathbf{X})) \geq 1 - \alpha.$$

## Věta

Buď $\mathbf{X} = (X_1, \ldots, X_n)$ náhodný výběr z rozdělení, jež je prvkem modelu $\mathcal{F} = \{\mathsf{P}_\theta, \theta \in \Theta\}$, ve kterém platí $\mathrm{var}_\theta X_1 \in \mathbb{R}^+$. Potom intervalovým odhadem střední hodnoty, tj. odhadem parametrické funkce $\mu = \mathsf{E}_\theta X_1$, o asymptotické spolehlivosti $(1 - \alpha)$ je interval

$$\left(\overline{X}_n - u_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}}, \overline{X}_n + u_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}}\right).$$

# Intervalový odhad v normálním modelu II

## Věta

Buďte $\mathbf{X} = (X_1, \ldots, X_n)$ náhodný výběr z $N(\mu_X, \sigma^2)$
a $\mathbf{Y} = (Y_1, \ldots, Y_m)$ na něm nezávislý náhodný výběr z $N(\mu_Y, \sigma^2)$,
kde $\sigma^2 > 0$ je známé.

Pak oboustranný intervalový odhad parametrické funkce $(\mu_X - \mu_Y)$ o spolehlivosti $(1 - \alpha)$ má tvar

$$\left( \overline{X}_n - \overline{Y}_m - u_{1-\frac{\alpha}{2}} \sigma \sqrt{\frac{m+n}{mn}},\ \overline{X}_n - \overline{Y}_m + u_{1-\frac{\alpha}{2}} \sigma \sqrt{\frac{m+n}{mn}} \right).$$

## Věta

Buďte $\mathbf{X} = (X_1, \ldots, X_n)$ náhodný výběr z $N(\mu_X, \sigma^2)$
a $\mathbf{Y} = (Y_1, \ldots, Y_m)$ na něm nezávislý náhodný výběr z $N(\mu_Y, \sigma^2)$.
Označme

$$S_X^2 = \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \overline{X}_n)^2, \qquad S_Y^2 = \frac{1}{m-1}\sum_{k=1}^{m}(Y_k - \overline{Y}_m)^2.$$

Pak oboustranný intervalový odhad parametrické funkce $(\mu_X - \mu_Y)$ o spolehlivosti $(1-\alpha)$ má tvar

$$\left(\overline{X}_n - \overline{Y}_m - t_{m+n-2}\left(1-\frac{\alpha}{2}\right) S^* \sqrt{\frac{m+n}{mn}},\ \overline{X}_n - \overline{Y}_m + t_{m+n-2}\left(1-\frac{\alpha}{2}\right) S^* \sqrt{\frac{m+n}{mn}}\right)$$

kde

$$S^* = \left(\frac{1}{m+n-2}\left((n-1)S_X^2 + (m-1)S_Y^2\right)\right)^{\frac{1}{2}}.$$

# Testování hypotéz — úvod

Postup:

1. Formulace statistického modelu $\mathbf{X} = (X_1, \dots, X_n) \sim \mathsf{P}_{\theta_0} \in \mathcal{F} = \{\mathsf{P}_\theta : \theta \in \Theta\}$
2. Formulace nulové hypotézy a alternativní hypotézy $\Theta = \Theta_0 \cup \Theta_1$

   princip: $\theta \in \Theta_0 \;\Leftrightarrow\; \theta$ uspokojivé
   
   $\theta \in \Theta_1 \;\Leftrightarrow\; \theta$ problematické

   testuji nulovou hypotézu $H_0 : \theta \in \Theta_0$ proti alternativní hypotéze $H_1 : \theta \in \Theta_1$
3. Volba hladiny testu $\alpha$, $0 < \alpha < 1$

   | | Rozhodnutí | |
   |---|---|---|
   | Skutečnost | nezamítáme $H_0$ | zamítáme $H_0$ |
   | $H_0$ platí | OK | chyba 1.druhu |
   | $H_1$ platí | chyba 2.druhu | OK |

   hladina testu $\alpha$ = maximální pst chyby I. druhu

   Asymetrie mezi $H_0$ a $H_1$ !
4. Volba rozhodovacího pravidla / kritického oboru testu $W \subset \mathbb{R}^n$

   náhodný výběr $\mathbf{X} \in W \;\Rightarrow\;$ zamítáme $H_0$
   
   $\mathbf{X} \notin W \;\Rightarrow\;$ nezamítáme $H_0$
5. Provedu experiment

- kritický obor $W \subset \mathbb{R}^n$ je nenáhodná množina
- → náhodný je jev $\{\mathbf{X} \in W\} = \{T_n(\mathbf{X}) \in C\}$
- → hladina testu je

$$\sup_{\theta \in \Theta_0} \mathsf{P}_\theta(\mathbf{X} \in W) = \sup_{\theta \in \Theta_0} \mathsf{P}_\theta(T_n(\mathbf{X}) \in C) \leq \alpha$$

- → Funkce

$$\beta(\theta) = \mathsf{P}_\theta(\mathbf{X} \in W), \quad \theta \in \Theta_1,$$

se nazývá síla testu proti alternativě $\theta$

Test s kritickým oborem $W^*$, který splňuje hladinu testu $\alpha$ a zároveň

$$\mathsf{P}_\theta(\mathbf{X} \in W^*) \geq \mathsf{P}_\theta(\mathbf{X} \in V) \quad \forall \theta \in \Theta_1,$$

pro každý jiný test na hladině $\alpha$ s kritickým oborem $V$,
je stejnoměrně nejsilnější test $H_0$ proti $H_1$ na hladině $\alpha$.

Jak zkonstruovat test, resp. jeho kritický obor?

- testová statistika $T_n(\mathbf{X})$:
  - za platnosti nulové hypotézy $H_0$ je rozdělení $\mathrm{P}_{T_n(\mathbf{X})}$ známé a nezávisí $\theta$
  - ale zároveň je rozdělení $\mathrm{P}_{T_n(\mathbf{X})}$ citlivé na skutečnou hodnotu $\theta$
- kritický obor $C$ volíme tak, aby:
  - byla dodržena hladina testu $\alpha$
  - a byly v něm zahrnuty ty hodnoty $T_n(\mathbf{X})$, které jsou za platnosti $H_0$ méně pravděpodobné, než za platnosti $H_1$

Příklad - chceme porovnat průměrnou výšku chlapců a děvčat:

1. Model: $X_i \sim N(\mu_X, \sigma^2)$, $Y_j \sim N(\mu_Y, \sigma^2)$ $\mu_X, \mu_Y \in \mathbb{R}, \sigma^2 \in \mathbb{R}^+$
   $\mathbf{X} = (X_1, \dots, X_n)$ a $\mathbf{Y} = (Y_1, \dots, Y_m)$ vzájemně nezávislé náhodné výběry
2. $H_0 : \mu_X = \mu_Y$ proti $H_1 : \mu_X \neq \mu_Y$
3. hladina např. $\alpha = 0.05$
4. Test: najít statistiku $T(\mathbf{X}, \mathbf{Y})$ a kritický obor $C$

Víme

$$\frac{\overline{X}_n - \overline{Y}_m - (\mu_X - \mu_Y)}{\sqrt{\frac{(n-1)S_X^2+(m-1)S_Y^2}{n+m-2}}}\sqrt{\frac{mn}{m+n}} \;\sim\; t_{m+n-2}$$

ale za platnosti $H_0$ je $\mu_X - \mu_Y = 0$, takže platí

$$T(\mathbf{X},\mathbf{Y}) = \frac{\overline{X}_n - \overline{Y}_m}{\sqrt{\frac{(n-1)S_X^2+(m-1)S_Y^2}{n+m-2}}}\sqrt{\frac{mn}{m+n}} \;\sim\; t_{m+n-2}$$

za kritický obor volíme

$$C = \left(-\infty, -t_{m+n-2}\left(1-\frac{\alpha}{2}\right)\right] \cup \left[t_{m+n-2}\left(1-\frac{\alpha}{2}\right), +\infty\right)$$

Test (dvouvýběrový t-test): $T(\mathbf{X},\mathbf{Y}) \in C \quad \Rightarrow$ zamítám $H_0$ (ve prospěch $H_1$)
$T(\mathbf{X},\mathbf{Y}) \notin C \quad \Rightarrow$ nezamítám $\mathrm{H}_0$

# Odvození testu z intervalového odhadu

Intervalový odhad pro rozdíl středních hodnot v normálním modelu splňuje

$$\mathsf{P}_{\mu_X,\mu_Y}\left((\mu_X-\mu_Y)\in\left(\overline{X}_n-\overline{Y}_m-t_{\mathbf{1}-\frac{\alpha}{\mathbf{2}},m+n-\mathbf{2}}S^*\sqrt{\frac{m+n}{mn}},\ \overline{X}_n-\overline{Y}_m+t_{\mathbf{1}-\frac{\alpha}{\mathbf{2}},m+n-\mathbf{2}}S^*\sqrt{\frac{m+n}{mn}}\right)\right)=\mathbf{1}-\alpha$$

za platnosti $H_0$ je $\mu_X-\mu_Y=0$, takže platí

$$\mathsf{P}_{H_\mathbf{0}}\left(\overline{X}_n-\overline{Y}_m-t_{1-\frac{\alpha}{2},m+n-2}S^*\sqrt{\frac{m+n}{mn}}<0<\overline{X}_n-\overline{Y}_m+t_{1-\frac{\alpha}{2},m+n-2}S^*\sqrt{\frac{m+n}{mn}}\right)=1-\alpha$$

$$\mathsf{P}_{H_\mathbf{0}}\left(|\overline{X}_n-\overline{Y}_m|<t_{m+n-2}\left(1-\frac{\alpha}{2}\right)S^*\sqrt{\frac{m+n}{mn}}\right)=1-\alpha$$

$$\mathsf{P}_{H_\mathbf{0}}\left(T(\mathbf{X},\mathbf{Y})\in\left(-t_{m+n-2}\left(1-\frac{\alpha}{2}\right),t_{m+n-2}\left(1-\frac{\alpha}{2}\right)\right)\right)=1-\alpha$$

Rozhodovací pravidlo testu tedy vypadá:
$T(\mathbf{X},\mathbf{Y})\in C \ \Rightarrow$ zamítáme $H_0$,
$T(\mathbf{X},\mathbf{Y})\notin C \ \Rightarrow$ nezamítáme $H_0$

# Odvození testu z intervalového odhadu - obecně

Buď $(\eta_L(\mathbf{X}), \eta_U(\mathbf{X}))$ intervalový odhad pro parametrickou funkci $g(\theta)$ se spolehlivostí $(1-\alpha)$ (přesnou nebo asymptotickou)

Pak test $H_0 : g(\theta) = g(\theta_0)$ proti $H_1 : g(\theta) \neq g(\theta_0)$

určený rozhodovacím pravidlem:

$g(\theta_0) \notin (\eta_L(\mathbf{X}), \eta_U(\mathbf{X})) \;\Rightarrow$ zamítáme $H_0$,
$g(\theta_0) \in (\eta_L(\mathbf{X}), \eta_U(\mathbf{X})) \;\Rightarrow$ nezamítáme $H_0$

má hladinu $\alpha$ (přesně nebo asymptoticky)