Kapitola 1

Základy pravděpodobnosti

1.1 Základní pojmy

Definice 1.1 *Nechť $\mathcal{A}$ je neprázdný systém podmnožin množiny $\Omega \neq \emptyset$ takový, že*

a) $\emptyset \in \mathcal{A}$,

b) je-li $A \in \mathcal{A}$, pak $A^c \in \mathcal{A}$, kde A^c značí doplněk množiny A do Ω.

c) jsou-li $A_i \in \mathcal{A}$, $i = 1,2,\ldots$, pak $\cup_{i=1}^{\infty} A_i \in \mathcal{A}$.

Pak $\mathcal{A}$ nazýváme σ-algebrou.

Definice 1.2 *Nechť $\Omega \neq \emptyset$ a $\mathcal{A}$ je σ-algebra definovaná na Ω. Pak pravděpodobností nazveme reálnou funkci P definovanou na $\mathcal{A}$, která splňuje*

a) $P(\Omega) = 1$, $P(\emptyset) = 0$,

b) $P(A) \geqq 0$ *pro všechna* $A \in \mathcal{A}$,

c) pro každou posloupnost **po dvou** *disjunktních jevů $\{A_n\}_{n=1}^{\infty}$ platí*

$$P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i).$$

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

Terminologie:

1) $\emptyset$ ... jev nemožný
2) Ω ... jev jistý
3) $A \cup B$ ... sjednocení jevů A, B (jev, který nastane právě tehdy, nastane-li aspoň jeden z jevů A, B)
4) $A \cap B$ ... průnik jevů A, B (jev, který nastane právě tehdy, nastanou-li oba dva jevy současně)
5) $B - A$ ... rozdíl jevu B a A (jev, který nastane právě tehdy, když nastane jev B a zároveň nenastane jev A)
6) $A \subset B$ ... A je podjev jevu B
7) $A^c = \Omega - A$ ... doplněk jevu A (jev, který nastane právě tehdy, když nenastane jev A)
8) $A \cap B = \emptyset$ ... jevy A, B jsou disjunktní (nemohou nastat současně)

Vlastnosti pravděpodobnosti:

1) $0 \leq P(A) \leq 1, \quad \forall A \in \mathcal{A}$,
2) P je monotónní: $A, B \in \mathcal{A}, A \subset B \Rightarrow P(A) \leq P(B)$,
3) $P(A^c) = 1 - P(A), \quad \forall A \in \mathcal{A}$,
4) $P(A \cup B) = P(A) + P(B) - P(A \cap B), \quad \forall A, B \in \mathcal{A}$,
5) $A, B \in \mathcal{A}, A \subset B \Rightarrow P(B - A) = P(B) - P(A)$,
6) pro každou posloupnost disjunktních jevů $\{A_i\}_{i=1}^{\infty}$ takových, že $\cup_{i=1}^{\infty} A_i = \Omega$, platí $P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i) = 1$.

1.1.1 Speciální typy pravděpodobnostních prostorů

I. Klasický pravděpodobnostní prostor

Pravděpodobnostní prostor $(\Omega, \mathcal{A}, P)$ nazveme klasickým pravděpodobnostním prostorem, jestliže

a) množina Ω je konečná a všechny možné výsledky jsou stejně pravděpodobné, tzn. označíme-li postupně $\mathrm{p}_1, \ldots, \mathrm{p}_m$ pravděpodobobnosti jednotlivých výsledků elementárních jevů, pak $\mathrm{p}_1 = \mathrm{p}_2 = \ldots = \mathrm{p}_m = \frac{1}{m}$ (je-li možných výsledků m),

b) za σ-algebru $\mathcal{A}$ vezmeme systém všech podmnožin množiny Ω,

c) pravděpodobnost P náhodného jevu A je rovna

$$P(A) = \frac{m_A}{m},$$

kde m_A je počet výsledků příznivých jevů A a m je počet všech možných výsledků náhodného pokusu.

II. Geometrický pravděpodobnostní prostor

Geometrickým pravděpodobnostním prostorem nazveme pravděpodobnostní prostor $(\Omega, \mathcal{A}, P)$ takový, že

a) $\Omega \subset \mathbb{R}^d$ (obvykle $d = 1, 2, 3$), neboli všechny elementární jevy lze vyjádřit jako body nějaké množiny,

b) $\mathcal{A} = \mathcal{B}(\Omega)$ je Borelovská σ-algebra na Ω (tj. nejmenší σ-algebra obsahující všechny otevřené podmnožiny Ω, tudíž i všechny uzavřené podmnožiny a kombinace obou těchto typů),

c) $P(A) = \frac{\mu^d(A)}{\mu^d(\Omega)}$, kde μ^d je d-rozměrná Lebesqueova míra. Pro naše účely postačí, pokud si pod $\mu^1(A)$ představíme délku množiny A, pod $\mu^2(A)$ obsah A a pod $\mu^3(A)$ objem A.

III. Obecný diskrétní pravděpodobnostní prostor

Obecný diskrétní pravděpodobnostní prostor je pravděpodobnostní prostor $(\Omega, \mathcal{A}, P)$ takový, že

a) $\Omega = \{\omega_1, \omega_2, \ldots\}$,

b) $\mathcal{A}$ je množina všech podmnožin Ω,

c) jsou dány pravděpodobnosti elementárních jevů $P(\omega_i)$, které splňují: $\sum_{i=1}^{\infty} P(\omega_i) = 1$. Pak pravděpodobnost libovolného jevu je dána jednoznačně vztahem $P(A) = \sum_{\omega_i \in A} P(\omega_i)$.

IV. Obecný spojitý pravděpodobnostní prostor

O spojitém pravděpodobnostním prostoru mluvíme tehdy, když

a) $\Omega = \mathbb{R}$, neboli všechny elementární jevy lze vyjádřit jako body na reálné ose,

b) $\mathcal{A} = \mathcal{B}(\mathbb{R})$ je Borelovská σ-algebra nad $\mathbb{R}$,

c) Je dána funkce f: $\mathbb{R} \to [0, \infty]$ taková, že $\int_{\mathbb{R}} f(x)dx = 1$. Pak pravděpodobnost libovolného jevu $A \in \mathcal{A}$ je dána jednoznačně vztahem

$$P(A) = \int_A f(x)dx.$$

1.1.2 Podmíněná pravděpodobnost

Definice 1.3 *Nechť je dán pravděpodobnostní prostor $(\Omega, \mathcal{A}, P)$ a náhodné jevy A, B, kde $P(B) > 0$. Podmíněnou pravděpodobnost jevu A za podmínky, že nastal jev B, definujeme vztahem*

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{1.1}$$

Věta 1.1 *Nechť je dán pravděpododnostní prostor $(\Omega, \mathcal{A}, P)$ a náhodný jev B, kde $P(B) > 0$. Potom pro libovolný jev $A \in \mathcal{A}$ platí:*

a) $P(A|B) \geq 0$,

b) $P(\Omega|B) = 1$,

c) $P(\cup_{n=1}^{\infty} A_n|B) = \sum_{n=1}^{\infty} P(A_n|B)$ *pro každou posloupnost* $\{A_n\}$ *disjunktních jevů.*

Důkaz.

a) zřejmé,

b) z definice 1.3 plyne, že

$$P(\Omega|B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1,$$

c) protože $A_1, A_2, \ldots$ jsou disjunktní, tak i $A_1 \cap B, A_2 \cap B, \ldots$ jsou disjunktní. Z axiomu c) definice 1.2 a z definice 1.3 plyne

$$\begin{aligned} P(\cup_{n=1}^{\infty} A_n|B) &= \frac{P(\cup_{n=1}^{\infty} A_n \cap B)}{P(B)} = \frac{\sum_{n=1}^{\infty} P(A_n \cap B)}{P(B)} = \\ &= \sum_{n=1}^{\infty} P(A_n|B) \end{aligned}$$

$\square$

Poznámka 1.1 *Věta 1.1 říká, že podmíněná pravděpodobnost má všechny vlastnosti pravděpodobnosti nepodmíněné.*

Věta 1.2 (o násobení pravděpodobnosti)*:*
Pro libovolnou posloupnost náhodných jevů $A_1, A_2, \ldots, A_n$*,* $P(A_1 \cap A_2 \cap \ \ldots \ \cap A_{n-1}) > 0$ *platí*

$$\begin{aligned} P(\cap_{i=1}^{n} A_i) &= P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2)\ldots \\ &\quad \ldots P(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1}). \end{aligned} \tag{1.2}$$

Důkaz. Opakovaným použitím definice 1.3 podmíněné pravděpodobnosti dostáváme:

$$\begin{aligned} P(\cap_{i=1}^{n-1} A_i \cap A_n) &= P(\cap_{i=1}^{n-1} A_i)P(A_n| \cap_{i=1}^{n-1} A_i) = \\ &= P(\cap_{i=1}^{n-2} A_i)P(A_{n-1}| \cap_{i=1}^{n-2} A_i)P(A_n| \cap_{i=1}^{n-1} A_i)\ldots \\ &= P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2)\ldots P(A_n| \cap_{i=1}^{n-1} A_i). \end{aligned}$$

Vzhledem k monotonnii pravděpodobnosti a předpokladu věty máme

$$P(A_1) \geq P(A_1 \cap A_2) \geq \ldots \geq P(A_1 \cap \ldots \cap A_{n-1}) > 0$$

a tedy všechny podmíněné pravděpodobnosti v tvrzení věty jsou dobře definovány.

Věta 1.3 (O celkové pravděpodobnosti) *Nechť* A_1*,* A_2*, …jsou náhodné jevy tvořící rozklad jevu jistého, tzn.*

$$A_i \cap A_j = \emptyset, \ \forall i \neq j \ \ a \ \ \cup_{i=1}^{\infty} A_i = \Omega.$$

Nechť tyto náhodné jevy mají postupně pravděpodobnosti $P(A_1), P(A_2), \ldots,$ *přičemž* $P(A_i) > 0, \ \forall i = 1, 2, \ldots$ *Uvažujme libovolný náhodný jev* B*, u něhož známe podmíněné pravděpodobnosti*

$$P(B|A_i), \ \forall i = 1, 2, \ldots$$

Potom

$$P(B) = \sum_{i=1}^{\infty} P(A_i) \cdot P(B|A_i). \tag{1.3}$$

Důkaz. Jevy $A_1, \ldots, A_n$ tvoří disjunktní rozklad $\Rightarrow$ $(A_i \ \cap B) \cap (A_j \cap B) = \emptyset \quad \forall i \neq j$, $\cup_{i=1}^{\infty}(A_i \cap B) = B$. Potom

$$P(B) = P(\cup_{i=1}^{\infty}(A_i \cap B)) = \sum_{i=1}^{\infty} P(A_i \cap B) = \sum_{i=1}^{\infty} P(A_i) \cdot P(B|A_i).$$

□

Věta 1.4 (Bayesova věta) *Nechť jsou splněny předpoklady věty 1.3. Pak*

$$P(A_i|B) = \frac{P(B|A_i) \cdot P(A_i)}{\sum_{j=1}^{\infty} P(A_j) \cdot P(B|A_j)}, \quad i = 1, 2, \dots \tag{1.4}$$

Důkaz. Podle vzorce pro podmíněnou pravděpodobnost je

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)}. \tag{1.5}$$

Po dosazení (1.3) do (1.5) dostáváme

$$P(A_i|B) = \frac{P(A_i \cap B)}{\sum_{j=1}^{\infty} P(A_j) \cdot P(B|A_j)} = \frac{P(B|A_i) \cdot P(A_i)}{\sum_{j=1}^{\infty} P(A_j) \cdot P(B|A_j)}. \tag{1.6}$$

□

1.1.3 Nezávislost

Definice 1.4 *Náhodné jevy A a B jsou nezávislé, jestliže platí*

$$P(A \cap B) = P(A) \cdot P(B). \tag{1.7}$$

Pojem nezávislosti můžeme rozšířit i na skupinu náhodných jevů.

Definice 1.5 *Nechť $A_1, A_2, \dots, A_n$ jsou náhodné jevy. Řekneme, že jsou skupinově (totálně) nezávislé, jestliže pro libovolnou posloupnost indexů $\{k_1, k_2, \dots, k_r\} \subset \{1, \dots, n\}$, $r = 2, \dots, n$ platí*

$$P(A_{k_1} \cap A_{k_2} \cap \dots \cap A_{k_r}) = P(A_{k_1}) \cdot P(A_{k_2}) \cdot \dots \cdot P(A_{k_n}). \tag{1.8}$$

Definice 1.6 *Nechť $A_1, \dots, A_n$ jsou náhodné jevy. Řekneme, že jsou po dvou nezávislé, jestliže jevy A_i, A_j jsou nezávislé pro všechna $i, j = 1, \dots, n,\ i \neq j$.*

Věta 1.5 *Nechť A, B jsou nezávislé náhodné jevy. Pak dvojice jevů (A, B^c), (A^c, B), (A^c, B^c) jsou nezávislé.*

Důkaz.

$$\begin{aligned} P(A^c \cap B) &= P(B - A) = P(B - [A \cap B]) = P(B) - P(A \cap B) = \\ &= P(B) - P(B) \cdot P(A) = P(B) \cdot (1 - P(A)) = P(B) \cdot P(A^c). \end{aligned}$$

Nezávislost jevů A, B^c se dokáže analogicky.

Jsou-li nezávislé jevy A, B, pak podle předchozího jsou nezávislé jevy A, B^c, ale odtud opět podle předchozího jsou nezávislé i jevy A^c, B^c.

□

1.2 Náhodná veličina

Definice 1.7 *Nechť $(\Omega, \mathcal{A}, P)$ je pravděpodobnostní prostor. Reálnou funkci X definovanou na Ω nazýváme náhodnou veličinou, jestliže X je měřitelné zobrazení $X : (\Omega, \mathcal{A}) \to (\mathbb{R}, \mathcal{B})$, tj.*

$$\{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A} \tag{1.9}$$

pro libovolnou borelovskou množinu $B \in \mathcal{B}$ ($\mathcal{B}$ je σ-algebra borelovských podmnožin, tj. nejmenší σ-algebra obsahující systém všech otevřených podmnožin $\mathbb{R}$).

Poznámka 1.2 *Náhodné veličiny budeme značit velkými písmeny $X, Y, Z \ldots$ Hodnoty, kterých mohou náhodné veličiny nabývat, budeme značit malými písmeny $x, y, z \ldots$*

Poznámka 1.3 *Místo $\{\omega \in \Omega : X(\omega) \in B\}$ budeme zjednodušeně psát $\{X \in B\}$ a místo $\{\omega \in \Omega : X(\omega) < x\}$ budeme zjednodušeně psát $\{X < x\}$.*

Poznámka 1.4 *Součty, součiny, podíly, minima a maxima náhodných veličin jsou opět náhodné veličiny; umocnění náhodné veličiny přirozeným číslem, násobení náhodné veličiny skalárem jsou také náhodné veličiny.*

Definice 1.8 *Nechť X je náhodná veličina. Její distribuční funkcí nazýváme reálnou funkci F_X reálné proměnné x definovanou*

$$F_X(x) = P(X \leq x) = P(\{\omega : X(\omega) \leq x\}), \quad x \in \mathbb{R}. \tag{1.10}$$

Vlastnosti distribuční funkce:

Distribuční funkce $F_X(x)$ náhodné veličiny X je

a) neklesající, tj. pro libovolné $a, b \in \mathbb{R}, a \leq b$, platí $F_X(a) \leq F_X(b)$,

b) zprava spojitá v libovolném bodě $x \in \mathbb{R}$,

c) $\lim_{x \to -\infty} F_X(x) = 0, \lim_{x \to \infty} F_X(x) = 1$.

Diskrétní náhodná veličina

Definice 1.9 *Náhodná veličina X se nazývá diskrétní (nebo také říkáme, že X má diskrétní rozdělení), jestliže existuje (konečná nebo spočetná) posloupnost reálných čísel $\{x_n\}$ a odpovídající posloupnost nezáporných čísel $\{p_n\} = P(X = x_n)$ (obvykle $p_n > 0$ pro všechna n) taková, že*

$$\sum_{n=1}^{\infty} p_n = 1. \tag{1.11}$$

Distribuční funkce diskrétní náhodné veličiny X má tvar

$$F_X(x) = P(X \le x) = \sum_{\{n: x_n \le x\}} P(X = x_n) = \sum_{\{n: x_n \le x\}} p_n \tag{1.12}$$

a

$$P(a < X \le b) = F_X(b) - F_X(a) = \sum_{\{n: a < x_n \le b\}} P(X = x_n) = \sum_{\{n: a < x_n \le b\}} p_n$$

pro libovolná reálná čísla a, b, kde $a \le b$.

Distribuční funkce je tedy schodovitá funkce se skoky v bodech $x_1, x_2, \ldots$ a je konstantní na intervalech $[x_n, x_{n+1})$. Velikost skoku v bodě x_n je $p_n = P(X = x_n)$.

Absolutně spojitá náhodná veličina

Definice 1.10 *Náhodná veličina X se nazývá absolutně spojitá (nebo také říkáme, že X má absolutně spojité rozdělení), jestliže existuje nezáporná integrovatelná funkce f_X taková, že platí*

$$F_X(x) = P(X \le x) = \int_{-\infty}^{x} f_X(t)dt, \quad x \in (-\infty, \infty). \tag{1.13}$$

Funkce f_X se nazývá hustotou rozdělení pravděpodobnosti.

Poznámka 1.5 *Místo "$P(X$ má vlastnost $V) = 1$" budeme říkat "X má vlastnost V skoro jistě." Často budeme užívat zkratku s.j.*

Vlastnosti hustoty:

a) $f_X(x) = \frac{d}{dx} F_X(x)$ s.j.,

b) $\int_{-\infty}^{\infty} f_X(x)dx = 1$,

c) $P(a < X \le b) = F_X(b) - F_X(a) = \int_a^b f_X(x)dx$ pro libovolná reálná čísla a, b, kde $a \le b$.

Zobecnění

Definice 1.11 *Míra je množinová fce na $(\Omega, \mathcal{A})$, tj.*

(i) $\mu : \mathcal{A} \to [0, \infty]$,

(ii) $\mu(\emptyset) = 0$,

(iii) jsou-li $A_n \in \mathcal{A}, n \geq 1$ disjunktní, pak $\mu(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$.

Je-li $\mu(\Omega) = 1$, říkáme μ pravděpodobnostní míra.

Každé náhodné veličině X a borelovské množině $B \in \mathcal{B}$ pak lze připsat pravděpodobnostní míru na $(\mathbb{R}, \mathcal{B})$
$$\mu_X(B) = P(\{\omega \in \Omega : X(\omega) \in B\}),$$
kterou nazýváme *rozdělení pravděpodobnosti* náhodné veličiny X. Položíme-li speciálně $B = (-\infty, x]$, dostáváme

$$\mu_X(B) = P(\{\omega \in \Omega : X(\omega) \leq x\}) = F_X(x),$$

což není nic jiného než distribuční funkce. Vidíme tedy, že mezi distribuční funkcí a rozdělením pravděpodobností existuje vzájemně jednoznačný vztah. Speciálně pak, položíme-li $B = (a, b];\ -\infty < a \leq b < \infty$, dostáváme

$$P(X \in (a, b]) = F_X(b) - F_X(a) = \mu_X((a, b]).$$

Uvedený vztah je definován pouze pro intervaly, ovšem to nám stačí pro jednoznačné definování míry pro všechny borelovské množiny, tudíž platí

$$P(X \in B) = \mu_X(B) = \int_B 1 d\mu_X(x) = \int_B 1 dF_X(x), \quad \forall B \in \mathcal{B}.$$

Tento integrál je tedy zobecněním klasického integrálu $\int 1dx$, v němž ale x mají různou váhu, a sumy, kde jednotlivé sčítance mají také různou váhu.

Tyto úvahy nás proto vedly k rozdělení náhodných veličin na dva základní typy - na diskrétní a absolutní spojité náhodné veličiny. V praxi se však mohou vyskytovat i jejich kombinace.

1.2.1 Charakteristiky náhodných veličin

Střední hodnota

Definice 1.12 *Nechť X je náhodná veličina definovaná na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$.*

a) Nechť X je diskrétní náhodná veličina nabývající reálných hodnot $x_1, x_2, x_3, \ldots,$ tzn. taková, že $P(X = x_i) = p_i$. Pak střední hodnota $\mathbb{E}X$ náhodné veličiny X je tvaru
$$\mathbb{E}X = \sum_{i=1}^{\infty} x_i \cdot p_i, \tag{1.14}$$
pokud řada v (1.14) konverguje.

b) Nechť X je absolutně spojitá náhodná veličina s hustotou f_X. Pak střední hodnota náhodné veličiny X je

$$\mathbb{E}X = \int_{-\infty}^{\infty} x f_X(x) dx, \tag{1.15}$$

pokud integrál existuje.

Obecně řečeno, střední hodnotou $\mathbb{E}X$ náhodné veličiny X nazveme integrál

$$\mathbb{E}X = \int_{-\infty}^{\infty} x dF_X(x), \tag{1.16}$$

pokud tento integrál existuje.

Základní vlastnosti střední hodnoty

1) $\mathbb{E}a = a$
2) $\mathbb{E}(aX + bY) = a\mathbb{E}X + b\mathbb{E}Y$
3) $X_1 \leq X \leq X_2$ *s.j.* $\Rightarrow \mathbb{E}X_1 \leq \mathbb{E}X \leq \mathbb{E}X_2$, speciálně $X \geq 0$ *s.j.* $\Rightarrow \mathbb{E}X \geq 0$

Věta 1.6 *Nechť X je náhodná veličina definovaná na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ a nechť $\varphi : \mathbb{R} \to \mathbb{R}$.*

a) Má-li náhodná veličina X diskrétní rozdělení $\{x_n, p_n\}_{n \in N_0}$, pak

$$\mathbb{E}\varphi(X) = \sum_{n \in N_0} \varphi(x_n) p_n, \tag{1.17}$$

pokud jedna ze stran rovnosti existuje.

b) Má-li náhodná veličina X absolutně spojité rozdělení s hustotou f, potom

$$\mathbb{E}\varphi(X) = \int_{-\infty}^{\infty} \varphi(x) f(x) dx, \tag{1.18}$$

pokud jeden z integrálů existuje.

Obecně napsáno,

$$\mathbb{E}\varphi(X) = \int_{-\infty}^{\infty} \varphi(x) dF_X(x), \tag{1.19}$$

pokud jeden z integrálů existuje.

Momenty, rozptyl a kovariance

Definice 1.13 *Nechť X je náhodná veličina. Hodnota*
$\mathbb{E}X^n$ *se nazývá n-tý moment náhodné veličiny X,*
$\mathbb{E}(X-\mathbb{E}X)^n$ *se nazývá n-tý centrální moment náhodné veličiny X,*
$\mathbb{E}|X-\mathbb{E}X|$ *se nazývá absolutní moment náhodné veličiny X.*

Definice 1.14 *Druhý centrální moment se nazývá rozptyl, značí se* $\mathrm{var}\ X = \mathbb{E}(X-\mathbb{E}X)^2$.

Definice 1.15 *Nechť X, Y jsou náhodné veličiny takové, že $\mathbb{E}X^2 < \infty$ a $\mathbb{E}Y^2 < \infty$. Pak jejich kovariance je definována jako*

$$\mathrm{cov}(X,Y) = \mathbb{E}(X-\mathbb{E}X)(Y-\mathbb{E}Y).$$

Poznámka 1.6 *Z předešlých dvou definic vyplývá souvislost mezi rozptylem a kovariancí, a to* $\mathrm{cov}(X,X) = var(X)$.

Vlastnosti rozptylu a kovariance

1) Nechť X je náhodná veličina. Pak var $X = \mathbb{E}(X^2) - (\mathbb{E}X)^2$
2) Nechť c je konstanta. Pak var $c = 0$.
3) Nechť X je náhodná veličina a a je reálné číslo. Pak var $(aX) = a^2$var X.
4) Nechť X je náhodná veličina a c je konstanta. Pak var $(X+c) =$ var X.
5) Nechť X je náhodná veličina, která má konečnou střední hodnotu a konečný nenulový rozptyl. Nechť
$$Z = \frac{X-\mathbb{E}X}{\sqrt{\mathrm{var}\ X}}.$$
Pak $\mathbb{E}Z = 0$ a var $Z = 1$.
6) Nechť X, Y jsou náhodné veličiny. Pak var $(X+Y) =$ var $X +$ var $Y + 2\mathrm{cov}(X,Y)$.
7) Nechť X, Y jsou náhodné veličiny. Pak $\mathrm{cov}(X,Y) = \mathbb{E}(XY)$ - $\mathbb{E}X\mathbb{E}Y$.

Věta 1.7 (Čebyševova nerovnost) *Nechť X je náhodná veličina s konečným rozptylem. Pak pro libovolné $\varepsilon > 0$ platí*

$$P[|X-\mathbb{E}X| \geq \varepsilon] \leq \frac{\mathrm{var}\ X}{\varepsilon^2}.$$

Důkaz. Uvažujme náhodnou veličinu $Y = X - \mathbb{E}X$. Pak

$$\begin{aligned}\mathrm{var}\ X &= \mathbb{E}Y^2 = \int_{-\infty}^{\infty} x^2 dF_Y(x) \geq \int_{|x|\geq\varepsilon} x^2 dF_Y(x) \geq \\ &\geq \varepsilon^2 \int_{|x|\geq\varepsilon} dF_Y(x) = \varepsilon^2 P[|Y| \geq \varepsilon].\end{aligned}$$

□

1.2.2 Příklady diskrétních náhodných veličin

1. **Alternativní rozdělení** ($Alt(p)$)
   má náhodná veličina X, která nabývá jen hodnot 0 a 1 s pravděpodobnostmi $1-p$ a p. Číslo p se nazývá parametr alternativního rozdělení, $0 < p < 1$.
   Distribuční funkce je

$$F(x) = \begin{cases} 0 & \text{pro } x < 0 \\ 1-p & \text{pro } 0 \leq x < 1 \\ 1 & \text{pro } x \geq 1 \end{cases}$$

   Snadno se z definic spočte, že střední hodnota $\mathbb{E}X = p$ a rozptyl $\operatorname{var} X = p(1-p)$.

   Použití: Příkladem alternativně rozdělené náhodné veličiny je např. počet jedniček, které padnou při jednom hodu kostkou, počet zmetků při náhodném výběru jednoho výrobku atd.

2. **Binomické rozdělení** ($Bi(n,p)$)
   je rozdělení náhodné veličiny X, která nabývá hodnot $k = 0, 1, 2, \ldots, n$. Binomické rozdělení je jednoznačně určeno dvěma parametry: přirozeným číslem n a číslem $p \in (0,1)$.
   Pravděpodobnosti p_k jsou tvaru

$$p_k = \tbinom{n}{k} p^k (1-p)^{n-k}, \text{ pro } k = 0, 1, \ldots, n.$$

   Distribuční funkce $F(x)$ je

$$F(x) = \begin{cases} 0 & x < 0 \\ \sum_{0 \leq k \leq x} \tbinom{n}{k} p^k (1-p)^{n-k} & 0 \leq x < n \\ 1 & x \geq n. \end{cases}$$

   Rozdělení $Bi(n,p)$ má střední hodnotu np, neboť

$$\begin{aligned} \mathbb{E}X &= \sum_{k=0}^{n} k \tbinom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^{n} k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\ &= \sum_{k=1}^{n} k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} (1-p)^{n-k} \\ &= np \sum_{k=0}^{n-1} \frac{(n-1)!}{k!(n-k-1)!} p^k (1-p)^{n-k-1} = np(p + (1-p))^{n-1} = np. \end{aligned}$$

   K výpočtu rozptylu použijeme vztah

$$\operatorname{var} X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \mathbb{E}X(X-1) + \mathbb{E}X - (\mathbb{E}X)^2$$

   a provedeme výpočet $\mathbb{E}X(X-1) = \sum_{k=0}^{n} k(k-1) \tbinom{n}{k} p^k (1-p)^{n-k}$ analogický tomu pro střední hodnotu. Dosazením pak získáme hodnotu pro rozptyl $np(1-p)$.

Použití: Binomickým rozdělením se řídí např. náhodná veličina X, která je rovna počtu úspěchů v posloupnosti n nezávislých alternativních pokusů, kde pravděpodobnost úspěchu v každém pokuse je p, $0 < p < 1$. Tedy

$$X = \sum_{i=1}^{n} X_i,$$

kde

$$X_i = \begin{cases} 1 & \text{pokud v i-tém pokuse nastal úspěch,} \\ 0 & \text{pokud úspěch nenastal.} \end{cases}$$

X je součtem n alternativních náhodných veličin.

3. Poissonovo rozdělení ($Po(\lambda)$)
je rozdělení náhodné veličiny X, která nabývá hodnot $k = 0, 1, 2, \ldots$ s pravděpodobnostmi

$$p_k = e^{-\lambda} \frac{\lambda^k}{k!}.$$

Číslo $\lambda > 0$ je parametr Poissonova rozdělení.
Distribuční funkce je

$$F(x) = \begin{cases} 0 & \text{pro } x \leq 0 \\ \sum_{0 \leq j \leq x} e^{-\lambda} \frac{\lambda^j}{j!} & \text{pro } 0 \leq x < \infty. \end{cases}$$

Střední hodnota je λ. Rozptyl je rovněž roven λ. Jejich výpočet je obdobou výpočtu charakteristik binomického rozdělení

Poznámka 1.7 *Poissonovo rozdělení je limitním případem binomického rozdělení pro $n \to \infty, p \to 0, np \to \lambda$ (=konstanta).*

Použití: Tímto rozdělením se řídí např. náhodná veličina, kterou je počet výskytů sledovaného jevu v určitém časovém intervalu délky t (předpokládejme, že jev může nastat v kterémkoliv okamžiku a počet výskytů během časového intervalu závisí jen na jeho délce a ne na jeho počátku ani na tom, kolikrát jev nastal před jeho počátkem), např. počet telefonních zavolání během nějakého časového intervalu atd.

4. Geometrické rozdělení ($Geom(p)$)
je rozdělení náhodné veličiny X, která nabývá hodnot $k = 0, 1, 2, \ldots$ s pravděpodobnostmi $p_k = p(1-p)^k$. Parametr p je z intervalu (0,1). Distribuční funkce je

$$F(x) = \begin{cases} 0 & \text{pro } x < 0 \\ \sum_{0 \leq k \leq x} p(1-p)^k & \text{pro } x \geq 0. \end{cases}$$

Pomocí vzorců pro geometrické řady se spočte střední hodnota $\frac{1-p}{p}$ a rozptyl $\frac{1-p}{p^2}$.

Použití: Předpokládejme, že provádíme nezávislé pokusy a že pravděpodobnost úspěchu v jednom pokuse je pro všechny pokusy stejná a je rovna p. Pak vidíme, že náhodná veličina (počet pokusů do prvního úspěchu) se řídí geometrickým rozdělením. Název tohoto rozdělení vyplývá ze skutečnosti, že s rostoucími hodnotami k pravděpodobnosti p_k geometricky klesají.

1.2.3 Příklady spojitých náhodných veličin

1. **Rovnoměrné rozdělení na intervalu** $[a, b]$ **(**$Ro(a, b)$**)**
   je dáno hustotou
   $$f(x) = \begin{cases} \frac{1}{b-a} & a \le x \le b, \\ 0 & x < a, \quad x > b. \end{cases}$$
   Distribuční funkce je
   $$F(x) = \begin{cases} 0 & x < a, \\ \frac{x-a}{b-a} & a \le x \le b, \\ 1 & x \ge b. \end{cases}$$
   Lehce se spočte, že střední hodnota a rozptyl jsou
   $$\mathbb{E}X = \frac{a+b}{2}, \text{var } X = \frac{1}{12}(b-a)^2.$$
   Použití: S rovnoměrným rozdělením se setkáváme např. při vyšetřování chyb ze zaokrouhlování v numerických výpočtech. Jsou-li čísla vstupující do výpočtů nekonečné desetinné zlomky, jež se zaokrouhlují na k desetinných míst, pak lze chybu ze zaokrouhlení považovat za náhodnou veličinu s rovnoměrným rozdělením na intervalu $[-5 \cdot 10^{-k-1}, 5 \cdot 10^{-k-1}]$, ale také u spousty dalších jevů.

2. **Exponenciální rozdělení (**$Exp(\lambda)$**)**
   je rozdělení s hustotou
   $$f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & \text{jinak}, \end{cases}$$
   kde λ je parametrem rozdělení.
   Distribuční funkce je
   $$F(x) = \begin{cases} 0 & \text{pro } x \le 0 \\ 1 - e^{-\lambda x} & x > 0. \end{cases}$$
   Integrací per partes se spočte střední hodnota
   $$\mathbb{E}X = \int_0^\infty x\lambda e^{-\lambda x} dx = \frac{1}{\lambda}.$$
   Dvojitým použitím per partes získáme
   $$\mathbb{E}X^2 = \int_0^\infty x^2 \lambda e^{-\lambda x} dx = \frac{2}{\lambda^2}$$
   a následně rozptyl
   $$\text{var } X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \frac{1}{\lambda^2}.$$
   Použití: Exponenciální rozdělení $Exp(\lambda)$ je vhodným modelem "doby čekání" do nastoupení určitého jevu, např. doby životnosti určitého zařízení a to tehdy, jestliže rozdělení zbývající doby čekání nezávisí na tom, jak dlouho již čekáme. Říká se

tomu, že exponenciální rozdělení nemá paměť. Přesně je tato vlastnost popsána tvrzením: Má-li náhodná veličina X exponenciální rozdělení, pak

$$P(X > x + y|X > y) = P(X > x) \quad \forall x > 0, y > 0,$$

neboť $P(X > x + y|X > y)$ můžeme podle definice podmíněné pravděpodobnosti přepsat na

$$\frac{P(X > x + y)}{P(X > y)} = \frac{e^{-\lambda(x+y)}}{e^{-\lambda y}} = e^{-\lambda x}.$$

Poznámka 1.8 *Souvislost s Poissonovým rozdělením: Má-li náhodná veličina X popisující dobu čekání na událost rozdělení $Exp(\lambda)$, pak náhodná veličina Y popisující počet těchto událostí v časovém intervalu délky T má rozdělení $Po(\lambda T)$.*

3. **Obecné normální rozdělení ($N(\mu, \sigma^2)$)**
   je definováno hustotou

$$f(x) = \frac{1}{\sqrt{2\pi\ \sigma^2}}\ e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty,$$

kde μ reálné a σ^2 kladné jsou parametry.

Distribuční funkce je

$$F(x) = \frac{1}{\sqrt{2\pi}\ \sigma} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt, \quad -\infty < x < \infty. \tag{1.20}$$

Střední hodnota $\mathbb{E}X = \mu$ a rozptyl var $X = \sigma^2$.

4. **Normované normální rozdělení ($N(0, 1)$)**
   je speciálním případem obecného normální rozdělení pro $\mu = 0$ a $\sigma^2 = 1$, tj. je definováno hustotou

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad -\infty < x < \infty.$$

Jeho distribuční funkce se tradičně značí písmenem Φ a platí

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt, \ -\infty < x < \infty.$$

Je zřejmé, že střední hodnota je rovna 0 a rozptyl je roven 1.

Použití: Normální rozdělení má mimořádný význam v teorii pravděpodobnosti a matematické statistice, přestože se tímto rozdělením řídí přesně jen málo náhodných veličin. V následujících odstavcích bude řečeno, že součet velkého počtu nezávislých náhodných veličin (o jejichž rozdělení se činí jen velmi obecné předpoklady) má přibližně normální rozdělení, tím lze vysvětlit klíčovou roli tohoto rozdělení v teorii pravděpodobnosti a matematické statistice. Náhodné veličiny, s nimiž se v reálném světě setkáváme, lze velmi často považovat za výslednice působení velkého počtu drobných náhodných vlivů. Pak lze očekávat, že normální rozdělení bude vhodným modelem pro takové náhodné veličiny. Nejběžnějším typem takových veličin jsou náhodné chyby (chyby měření, způsobené velkým počtem neznámých a vzájemně nezávislých příčin). Normální rozdělení je vhodným modelem pro řadu fyzikálních, technických a biologických veličin jako například tělesná výška jedinců homogenní populace, roční částka, kterou pojišťovna vyplatí za pojistné příhody atd.

Poznámka 1.9 *Jelikož se s normálním rozdělením velmi často pracuje a výpočet distribuční funkce je zdlouhavý, jsou hodnoty distribuční funkce $N(0,1)$ tabelovány. Vzhledem k symetrii funkce ($\Phi(x) = 1 - \Phi(-x)$) se tabelují hodnoty Φ pouze pro nezáporné x.*

Věta 1.8 (Transformace veličin s normálním rozdělením)

a) *Má-li X normované normální rozdělení a $Y = \mu + \sigma X$, pak Y má normální rozdělení s parametry μ a σ^2.*

b) *Má-li X normální rozdělení s parametry μ, σ^2 a je-li $Y = a + bX$, pak Y má opět normální rozdělení s parametry $a + b\mu$ a $b^2\sigma^2$.*

c) *Nechť X, Y jsou náhodné veličiny, X má rozdělení $N(\mu_1, \sigma_1^2)$, Y má rozdělení $N(\mu_2, \sigma_2^2)$ a $\operatorname{cov}(X, Y) = 0$. Potom $Z = X + Y$ má rozdělení $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.*

1.3 Funkce náhodných veličin

1.3.1 Jedna veličina funkcí druhé

Při řešení některých pravděpodobnostních úloh se setkáváme se situací, kdy známe rozdělení náhodné veličiny X a hledáme rozdělení náhodné veličiny Y, která je funkcí veličiny X

$$Y = \varphi(X).$$

Věta 1.9 *Nechť X je náhodná veličina s distribuční funkcí F a nechť $\varphi : \mathbb{R} \to \mathbb{R}$. Označme $Y = \varphi(X)$ a G její distribuční funkci. Potom*

$$G(y) = \int_{\{x; \varphi(x) \le y\}} dF(x), \quad \forall y \in \mathbb{R}. \tag{1.21}$$

Speciálně, je-li F diskrétní $\{x_n, p_n\}$, je

$$G(y) = \sum_{\{x_n; \varphi(x_n) \le y\}} p_n, \quad \forall y \in \mathbb{R} \tag{1.22}$$

a je-li absolutně spojitá s hustotou f, je

$$G(y) = \int_{\{x; \varphi(x) \le y\}} f(x)\ dx, \quad \forall y \in \mathbb{R}. \tag{1.23}$$

Důkaz. Označme $B_y = \{x; \varphi(x) \leq y\}$. Pak

$$G(y) = P(Y \leq y) = P(\varphi(X) \leq y) = P(X \in B_y) = \int_{\{x;\varphi(x)\leq y\}} dF(x).$$

□

1.3.2 Součet náhodných veličin

Mějme dvě nezávislé (viz kapitola 1.3.4) náhodné veličiny X a Y s distribučními funkcemi $F(x)$ a $G(y)$. Zajímá nás rozdělení součtu $Z = X+Y$. Distribuční funkci náhodné veličiny Z označme $H(z)$. Pak obecně platí

$$\begin{aligned} H(z) &= \int\int_{x+y\leq z} dF(x)dG(y) = \int_{-\infty}^{\infty} F(z-y)dG(y) = \\ &= \int_{-\infty}^{\infty} G(z-x)dF(x). \end{aligned} \tag{1.24}$$

Definice 1.16 *Rozdělení s distribuční funkcí $H(z)$ se nazývá konvoluce rozdělení s distribučními funkcemi $F(x)$ a $G(y)$. H se nazývá konvoluce distribučních funkcí F a G.*

Operaci konvoluce budeme značit $H = F * G$.

Mají-li náhodné veličiny X a Y diskrétní, resp. absolutně spojité, rozdělení, pak s rozdělením jejich součtů pracujeme následovně:

Věta 1.10 *Nechť F, G jsou diskrétní distribuční funkce se skoky v přirozených číslech o velikosti $\{p_n\}, \{q_n\}$, tj.*

$$F(x) = \sum_{0\leq n\leq x} p_n, \quad G(y) = \sum_{0\leq n\leq y} q_n.$$

Nechť $H = F * G$. Potom H je diskrétní distribuční funkce se skoky v přirozených číslech a platí

$$H(z) = \sum_{0\leq n\leq z} h_n, \text{ kde } h_n = \sum_{k=0}^{n} p_k\, q_{n-k}.$$

Věta 1.11 *Nechť náhodné veličiny X a Y jsou nezávislé a mají absolutně spojité distribuční funkce $F(x)$ a $G(y)$ s hustotami $f(x)$ a $g(y)$. Potom také $H = F * G$ je absolutně spojitá a pro její hustotu $h(z)$ (tj. pro hustotu náhodné veličiny $Z = X + Y$) platí*

$$h(z) = \int_{-\infty}^{\infty} f(x)g(z-x)dx = \int_{-\infty}^{\infty} f(z-y)g(y)dy. \tag{1.25}$$

Poznámka 1.10 *Funkce $h(z)$ definovaná vztahem (1.25) se nazývá* **konvoluce hustot** *$f(x)$ a $g(y)$ a budeme ji značit $h = f * g$. Je to skutečně hustota, neboť z (1.25) plyne, že $h(z) \geq 0$ a*

$$\begin{aligned}\int_{-\infty}^{\infty} h(z)dz &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x-y)g(y)dydx = \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(x-y)dx\right)\, g(y)dy = \\ &= \int_{-\infty}^{\infty} 1 \cdot\; g(y)dy = 1.\end{aligned}$$

Konvoluce některých základních pravděpodobnostních rozdělení:

1. **Konvoluce binomických rozdělení**
   Nechť X, Y jsou nezávislé náhodné veličiny $X \sim Binom(n_1, p)$ a $Y \sim Binom(n_2, p)$. Potom náhodná veličina $Z = X + Y$ má rozdělení $Binom(n_1 + n_2, p)$.

2. **Konvoluce Poissonových rozdělení**
   Nechť $X \sim Po(\lambda_1)$ a $Y \sim Po(\lambda_2)$ jsou nezávislé. Potom náhodná veličina $Z = X+Y$ má rozdělení $Po(\lambda_1 + \lambda_2)$.

3. **Konvoluce rovnoměrných rozdělení**
   Nechť

   $$f(x) = \begin{cases} \frac{1}{b-a} & \text{pro } a \leq x \leq b \\ 0 & \text{jinak} \end{cases}$$

   a

   $$g(y) = \begin{cases} \frac{1}{d-c} & \text{pro } c \leq y \leq d \\ 0 & \text{jinak.} \end{cases}$$

   Předpokládejme, že $d - c \geq b - a$. Pak pro konvoluci hustot $h(z)$ náhodných veličin s hustotami $f(x)$ a $g(y)$ platí

   $$h(z) = \begin{cases} 0 & \text{pro } z \leq a + c \text{ nebo } b + d \leq z \\ \frac{z-(a+c)}{(b-a)(d-c)} & \text{pro } a + c \leq z \leq b + c \\ \frac{1}{d-c} & \text{pro } b + c \leq z \leq a + d \\ \frac{(b+d)-z}{(b-a)(d-c)} & \text{pro } a + d \leq z \leq b + d \end{cases}$$

   Grafem je lichoběžník se základnou ve vodorovné ose. Vidíme, že hustota $h(z)$ je všude spojitá, ačkoliv $f(x)$ a $g(y)$ mají body nespojitosti (konvoluce "vyhlazuje" nespojitosti). Ve speciálním případě, kdy obě náhodné veličiny X a Y mají stejné rozdělení (tj. a = c, b = d), má hustota $h(z)$ tvar trojúhelníku. Toto rozdělení se nazývá *Simpsonovo rozdělení*.

4. **Konvoluce normálních rozdělení**
   Nechť X, Y jsou nezávislé náhodné veličiny, $X \sim N(\mu_1, \sigma_1^2)$ a $Y \sim N(\mu_2, \sigma_2^2)$. Potom $Z = X + Y$ má rozdělení $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

5. **Konvoluce exponenciálních rozdělení**
   Jsou-li X, Y nezávislé náhodné veličiny s týmž exponenciálním rozdělením s parametrem $\lambda > 0$, pak náhodná veličina $Z = X + Y$ má rozdělení s hustotou

$$h(z) = \begin{cases} \lambda^2 z \exp\{-z\lambda\} & z > 0, \\ 0 & z \le 0. \end{cases}$$

1.3.3 Náhodný vektor

Definice 1.17 *Nechť $(\Omega, \mathcal{A}, P)$ je pravděpodobnostní prostor a nechť na tomto prostoru jsou definovány náhodné veličiny $X_1,\ X_2,\ \ldots,\ X_n$. Pak vektor $\mathbb{X} = (X_1, \ldots, X_n)^T$ nazýváme náhodný vektor.*

Poznámka 1.11 *Náhodný vektor je tedy zobrazení z Ω do $\mathbb{R}^n$. Hodnoty náhodného vektoru je možno geometricky interpretovat jako bod v n-rozměrném prostoru.*

Definice 1.18 *Nechť $\mathbb{X} = (X_1, \ldots, X_n)^T$ je náhodný vektor definovaný na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$. (Sdruženou) distribuční funkcí $F_{\mathbb{X}}$ náhodného vektoru $\mathbb{X}$ nazveme reálnou funkci n proměnných definovanou na $\mathbb{R}^n$ vztahem*

$$\begin{aligned} F_{\mathbb{X}}(x_1, \ldots, x_n) &= P(X_1 \le x_1, X_2 \le x_2, \ldots, X_n \le x_n) = \\ &= P(\cap_{i=1}^n \{\omega : X_i(\omega) \le x_i\}),\ -\infty < x_i < \infty,\ i = 1, \ldots, n. \end{aligned} \tag{1.26}$$

Vlastnosti distribuční funkce náhodného vektoru

1. $F_{\mathbb{X}}(x_1, \ldots, x_n)$ je neklesající funkce v každé ze svých proměnných při pevných hodnotách ostatních proměnných.

2. $F_{\mathbb{X}}(x_1, \ldots, x_n)$ je zprava spojitá v každé proměnné.

3. $\lim_{x_i \to -\infty} F_{\mathbb{X}}(x_1, \ldots, x_n) = 0,\ i = 1, \ldots, n$, hodnoty x_j jsou pevné, $j \ne i,\ j = 1, \ldots, n$.

4. $\lim_{\substack{x_1 \to \infty \\ x_2 \to \infty \\ \vdots \\ x_n \to \infty}} F_{\mathbb{X}}(x_1, \ldots, x_n) = 1.$

Definice 1.19 *Náhodný vektor* $\mathbb{X}$ *má diskrétní rozdělení, jestliže existuje posloupnost* $\{\mathbf{x}_k\}_{k=1}^{\infty}$, $\mathbf{x}_k \in \mathbb{R}^n$, *a odpovídající posloupnost kladných čísel* $\{p_k\}_{k=1}^{\infty}$ *taková, že*

$$\sum_{k=1}^{\infty} p_k = 1, \qquad \text{kde } p_k = P(\mathbb{X} = \mathbf{x}_k) = P(\{\omega \in \Omega : \mathbb{X}(\omega) = \mathbf{x}_k\}). \tag{1.27}$$

Distribuční funkce náhodného vektoru $\mathbb{X}$ diskrétního typu má tvar

$$F_{\mathbb{X}}(\mathbf{x}) = \sum_{\{k: \mathbf{x}_k \leq \mathbf{x}\}} p_k, \quad \forall \mathbf{x} \in \mathbb{R}^n, \tag{1.28}$$

kde nerovnost $\mathbf{x}_k < \mathbf{x}$ je uvažována po složkách, tj. $x_k^i < x^i$ pro všechna $i = 1, \ldots, n$.

Definice 1.20 *Náhodný vektor* $\mathbb{X} = (X_1, \ldots, X_n)^T$ *má absolutně spojité rozdělení, jestliže existuje nezáporná funkce* $f_{\mathbb{X}}$ n *reálných proměnných taková, že*

$$F_{\mathbb{X}}(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_n} f_{\mathbb{X}}(t_1, \ldots, t_n) dt_1, \ldots, dt_n, \tag{1.29}$$

kde funkci $f_{\mathbb{X}}$ *nazýváme hustotou rozdělení pravděpodobnosti náhodného vektoru* $\mathbb{X}$, *nebo též sdruženou hustotou náhodných veličin* X_1, ..., X_n.

Poznámka 1.12 *Stejně jako v případě náhodných veličin, i pro náhodný vektor lze jeho rozdělení zobecnit pomocí pravděpodobnostních měr. Pro naše účely ale úplně postačí základní dělení na náhodné vektory diskrétní a spojité.*

Definice 1.21 *Rozdělení (distribuční funkce, hustota, pravděpodobnosti) náhodného vektoru* $(X_1, \ldots, X_k)^T$, *který je podvektorem náhodného vektoru* $\mathbb{X} = (X_1, \ldots, X_n)^T$, *se nazývá marginální rozdělení (distribuční funkce, hustota, pravděpodobnosti).*

Způsob výpočtu marginálního rozdělení:

1. Má-li náhodný vektor $\mathbb{X} = (X_1, \ldots, X_n)^T$ diskrétní rozdělení se sdruženými pravděpodobnostmi $P(X_1 = ., ..., X_{i-1} = ., X_i = ., X_{i+1} = ., ..., X_n = .)$, kde náhodné veličiny X_i nabývají hodnot $x_{i,1}, ..., x_{i,k_i}$, pak marginální pravděpodobnosti jsou

$$P(X_i = x) = \sum_{j_1=1}^{k_1} \ldots \sum_{j_{i-1}=1}^{k_{i-1}} \sum_{j_{i+1}=1}^{k_{i+1}} \ldots \sum_{j_n=1}^{k_n} P(X_1 = x_{1,j_1}, ..., X_{i-1} = x_{i-1,j_{i-1}}, \\ X_i = x, X_{i+1} = x_{i+1,j_{i+1}}, ..., X_n = x_{n,j_n}).$$

2. Má-li náhodný vektor $\mathbb{X} = (X_1, \ldots, X_n)^T$ spojité rozdělení se sdruženou hustotou $f_{\mathbb{X}}$, pak marginální hustota náhodné veličiny X_i je $(n-1)$-násobný integrál
$$f_{X_i}(x) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbb{X}}(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) dx_1, \ldots, dx_{i-1} dx_{i+1}, \ldots, dx_n.$$

Nejčastěji používané číselné charakteristiky náhodného vektoru $\mathbb{X} = (X_1, \ldots, X_n)^T$:

a) Vektor středních hodnot
$$\mathbb{E}\mathbb{X} = (\mathbb{E}X_1, \ldots, \mathbb{E}X_n)^T.$$

b) Varianční matice var $\mathbb{X}$ definována jako matice typu $n \times n$ s prvky
$$\mathrm{cov}(X_i, X_j) = \mathbb{E}(X_i - \mathbb{E}X_i)(X_j - \mathbb{E}X_j), \quad 1 \leq i, j \leq n.$$

c) Korelační matice corr$\mathbb{X}$ s prvky
$$\mathrm{corr}(X_i, X_j) = \frac{\mathrm{cov}(X_i, X_j)}{\sqrt{varX_i}\sqrt{varX_j}}, \quad 1 \leq i, j \leq n.$$

Poznámka 1.13 *Korelace je vynormovaná kovariance v tom smyslu, že zatímco kovariance může navývat libovolně velkých hodnot podle toho, jaké náhodné veličiny popisuje, pro korelaci platí*
$$-1 \leq corr(X, Y) \leq 1.$$

1.3.4 Nezávislost náhodných veličin

Definice 1.22 *Náhodné veličiny X_1, X_2 …, X_n jsou vzájemně nezávislé, jestliže*
$$P(\cap_{j=1}^{r}\{\omega : X_{i_j}(\omega) < x_{i_j}\}) = \Pi_{j=1}^{r} P(\{\omega : X_{i_j}(\omega) < x_{i_j}\}) \tag{1.30}$$
$$\forall\{i_1, i_2, \ldots, i_r\} \subset \{1, 2, \ldots, n\}, 1 \leq r \leq n, \forall x_{i_j} \in \mathbb{R}.$$

Poznámka 1.14 *Podobně jako u náhodných jevů můžeme zde definovat nezávislost náhodných veličin X_1, X_2 …, X_n po dvou. Definici nezávislosti po dvou bychom dostali z definice 1.22 pro $r = 2$.*

Věta 1.12 (Ověřování nezávislosti náhodných veličin v praxi)

a) Nechť $\mathbb{X} = (X_1, X_2 \ldots, X_n)^T$ je náhodný vektor diskrétního typu. Náhodné veličiny $X_1, X_2 \ldots, X_n$ jsou vzájemně nezávislé právě tehdy, když platí
$$P(X_1 = x_1^{(i)}, \ldots, X_n = x_n^{(i)}) = \Pi_{j=1}^{n} P(X_j = x_j^{(i)})$$
pro všechna $\mathbf{x}^{(i)} = (x_1^{(i)}, x_2^{(i)}, \ldots, x_n^{(i)})$, $i = 1, 2, \ldots,$ kterých může $\mathbb{X}$ nabývat.

b) Nechť $\mathbb{X} = (X_1, X_2 \ldots, X_n)^T$ *je náhodný vektor absolutně spojitého typu. Náhodné veličiny* $X_1, X_2 \ldots, X_n$ *jsou vzájemně nezávislé právě tehdy, platí-li*

$$f_{\mathbb{X}}(x_1, x_2 \ldots, x_n) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \ldots f_{X_n}(x_n), \; \forall (x_1, x_2 \ldots, x_n) \in \mathbb{R}^n.$$

Věta 1.13 *Jsou-li* X, Y *nezávislé náhodné veličiny s konečnými středními hodnotami, pak*

a) $\mathbb{E}XY = (\mathbb{E}X)(\mathbb{E}Y)$.

b) Jsou-li navíc $\mathbb{E}X^2 < \infty$ *a* $\mathbb{E}Y^2 < \infty$*, pak* $\operatorname{cov}(X, Y) = 0$.

Platí-li $\operatorname{cov}(X, Y) = 0$, pak říkáme, že náhodné veličiny jsou nekorelované. Z nekorelovanosti však obecně ještě neplyne nezávislost!

1.4 Zákony velkých čísel, centrální limitní věta

Definice 1.23 *Mějme posloupnost náhodných veličin* X_1, X_2, X_3, … *a náhodnou veličinu* X*. Nechť jsou všechny tyto veličiny definovány na témže pravděpodobnostním prostoru* $(\Omega, \mathcal{A}, P)$.
Říkáme, že X_n *konverguje k* X *skoro jistě, jestliže*

$$P\{\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\} = 1.$$

Jestliže pro každé $\varepsilon > 0$ *platí*

$$\lim_{n\to\infty} P\{\omega : |X_n(\omega) - X(\omega)| > \varepsilon\} = 0,$$

pak říkáme, že X_n *konverguje k* X *v pravděpodobnosti.*

Věta 1.14 *Z konvergence skoro jistě plyne konvergence v pravděpodobnosti.*

Implikaci nelze bez dodatečných předpokladů obrátit!

1.4.1 Zákony velkých čísel

Věta 1.15 (Slabý zákon velkých čísel)
Nechť $\{X_n\}_{n=1}^{\infty}$ *je posloupnost nezávislých náhodných veličin se stejnými středními hodnotami* μ *a stejnými rozptyly* $\sigma^2 < \infty$*. Pak pro* $n \to \infty$ *platí*

$$\frac{1}{n}(X_1 + X_2 + \ldots + X_n) \to \mu$$

v pravděpodobnosti.

Věta 1.16 (Silný zákon velkých čísel)
Nechť $\{X_n\}_{n=1}^{\infty}$ *je posloupnost nezávislých stejně rozdělených náhodných veličin s konečnou střední hodnotou* $\mathbb{E}X_1 = \mu$. *Pak pro* $n \to \infty$ *platí*

$$\frac{1}{n}(X_1 + X_2 + \ldots + X_n) \to \mu$$

v pravděpodobnosti i skoro jistě.

1.4.2 Centrální limitní věta

Podstatou centrální limitní věty (CLV) je tvrzení, že náhodná veličina X, která vznikla jako součet velkého počtu vzájemně nezávislých náhodných veličin $X_1, X_2, \ldots X_n$, má za velmi obecných podmínek přibližně normální rozdělení. Budeme říkat, že náhodná veličina X má tzv. asymptoticky normální rozdělení.

Centrální limitní věta má několik verzí s různými předpoklady. Asi nejčastěji používanou verzí je tzv. Lévy-Lindebergova CLV.

Věta 1.17 (Lévy-Lindebergova CLV)
Nechť X_1, X_2, $\ldots$ *jsou nezávislé stejně rozdělené náhodné veličiny se střední hodnotou* μ *a konečným rozptylem* σ^2. *Označme*

$$Z_n = \frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n\sigma^2}} \quad n = 1, 2, \ldots$$

a označme $F_n(x)$ *distribuční funkci* Z_n. *Potom* $\lim_{n\to\infty} F_n(x) = \Phi(x)$ *pro všechna* $-\infty < x < \infty$, *kde* $\Phi(x)$ *je distribuční funkce* $N(0, 1)$.

Poznámka 1.15 *Předpoklad, že veličiny jsou nezávislé a stejně rozdělené, se v literatuře často označuje zkratkou i.i.d. (z anglického "independent, identically disributed").*

Kapitola 2

Základy statistiky

Úlohu matematické statistiky lze zjednodušeně popsat takto: Pro danou reálnou situaci (nasbíraná data) vytvoříme model (např. nějaké pravděpodobnostní rozdělení), z dat odhadneme jeho parametry a na závěr testujeme hypotézy o těchto parametrech. Zatímco pro první krok stačí srovnat graficky znázorněná data (např. histogram) s grafickým znázorněním teoretického rozdělení (např. tvar grafu hustoty), k dalším dvěma krokům je již zapotřebí statistických výpočetních metod, které lze zjednodušeně rozdělit do tří skupin, a to:

1. Bodové odhady parametrů
2. Intervalové odhady parametrů
3. Testování hypotéz

A právě těmto metodám bude věnována tato kapitola.

Předtím, než popíšeme výše zmíněné statistické metody, uvedeme si další pravděpodobnostní rozdělení, kterých se využívá - na základě věty 2.1 - zejména při tvorbě intervalových odhadů a při testování hypotéz (viz níže). Hodnoty jejich distribučních funkcí jsou stejně jako hodnoty distribuční funkce normovaného normálního rozdělení uvedeny ve statistických tabulkách.

χ^2_n-rozdělení

Nechť $X_1, X_2 \ldots, X_n$ jsou nezávislé stejně rozdělené náhodné veličiny s rozděleními $N(0,1)$. Pak náhodná veličina

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

má tzv. χ^2_n-rozdělení (čteno "chí-kvadrát rozdělení o n stupních volnosti).

Studentovo t-rozdělení

Nechť X je náhodná veličina s rozdělením $N(0,1)$ a Y je náhodná veličina s rozdělením χ^2_n. Pak náhodná veličina

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

má tzv. t_n-rozdělení nazváno také Studentovo t-rozdělení o n stupních volnosti.

Definice 2.1 *Náhodnému vektoru $\mathbb{X} = (X_1, X_2 \ldots, X_n)^T$ nezávislých stejně rozdělených náhodných veličin s distribuční funkcí F_θ závisející na parametru θ ve statistice říkáme náhodný výběr. Číslu n říkáme rozsah výběru.*

Definice 2.2 *Funkce náhodného výběru*

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

se nazývá výběrový průměr a funkce

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$$

se nazývá výběrový rozptyl. $S_n = \sqrt{S_n^2}$ je pak výběrová směrodatná odchylka.

Věta 2.1 *Nechť $\mathbb{X} = (X_1, X_2 \ldots, X_n)^T$ je náhodný výběr z rozdělení $N(\mu, \sigma^2), \mu \in \mathbb{R}, \sigma^2 > 0$. Pak*

1. *výběrový průměr $\bar{X}_n$ a výběrový rozptyl S_n^2 jsou nezávislé náhodné veličiny,*
2. *rozdělení výběrového průměru $\bar{X}_n$ je $N(\mu, \sigma^2/n)$,*
3. *náhodná veličina $(n-1)S_n^2/\sigma^2$ má χ^2-rozdělení o $(n-1)$ stupních volnosti,*
4. *náhodná veličina $T = \frac{\bar{X}_n - \mu}{S_n}\sqrt{n}$ má t-rozdělení o $(n-1)$ stupních volnosti.*

Definice 2.3 *Nechť distribuční funkce F je spojitá, ryze monotonní a nechť $0 < \beta < 1$. Pak číslo z_β takové, že $F(z_\beta) = \beta$ se nazývá β-kvantil tohoto rozdělení.*

Označení kvantilů používaných v této přednášce

Kvantily v této přednášce budeme často označovat v souladu s označením rozdělení příslušné náhodné veličiny. Používat budeme

- u_β ... β-kvantil normovaného normálního rozdělení,

- $t_{\beta,n}$ ... β-kvantil Studentova $t-$rozdělení o n stupních volnosti,
- $\chi^2_{\beta,n}$ ... β-kvantil χ^2 rozdělení o n stupních volnosti.

Poznámka 2.1 *Je-li X náhodná veličina s distribuční funkcí F a kvantily z_β, pak*

$$P(z_{\alpha/2} < X < z_{1-\alpha/2}) = F(z_{1-\alpha/2}) - F(z_{\alpha/2}) = 1 - \alpha.$$

Definice 2.4 *Nechť $(x_1, x_2 \ldots, x_n)^T$ je realizace náhodného výběru $\mathbb{X} = (X_1, X_2 \ldots, X_n)^T$ Pak*

$$F_{emp}(x) = \frac{\#\{x_i : x_i \leq x\}}{n},$$

kde # značí počet prvků, se nazývá empirická distribuční funkce.

Definice 2.5 *Nechť $(x_1, x_2 \ldots, x_n)^T$ je realizace náhodného výběru $\mathbb{X} = (X_1, X_2 \ldots, X_n)^T$ Pak*

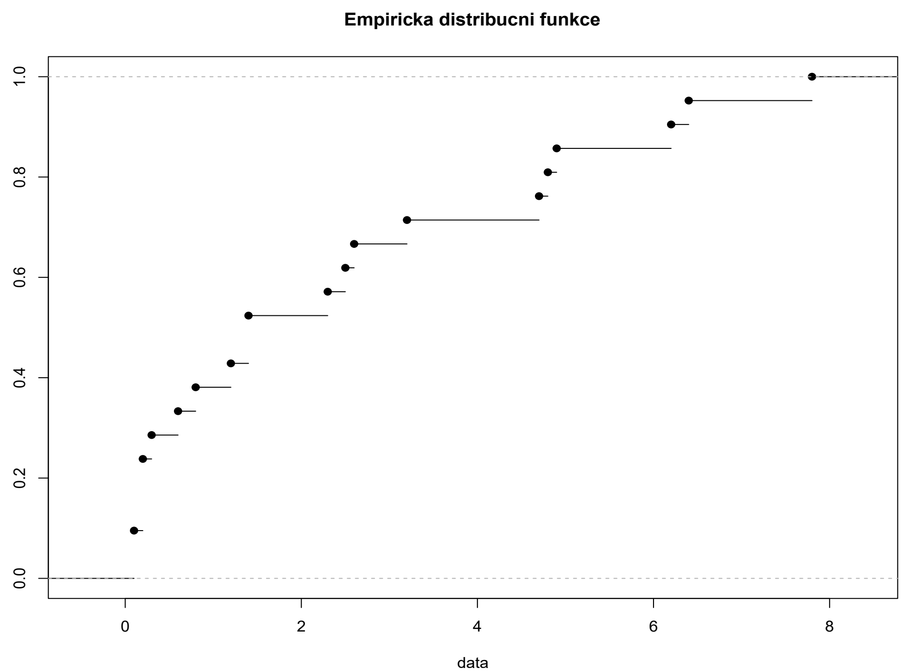
- $z = min(x_i : F_{emp}(x_i) \geq 1/4)$ *se nazývá 1.kvartil,*
- $z = min(x_i : F_{emp}(x_i) \geq 3/4)$ *se nazývá 3.kvartil,*
- $z = min(x_i : F_{emp}(x_i) \geq 1/2)$ *se nazývá medián (2.kvartil),*
- *nejčastěji zastoupený prvek se nazývá modus.*

Poznámka 2.2 *Grafickým znázorněním kvartilů je tzv. krabicový graf (boxplot) - viz obrázek 2.2 vpravo.*
Dalším často používaným grafickým znázorněním dat pak je tzv. histogram znázorňující počet hodnot spadající do (vhodně zvolených) ekvidistantních intervalů - viz obrázek 2.2 vlevo.

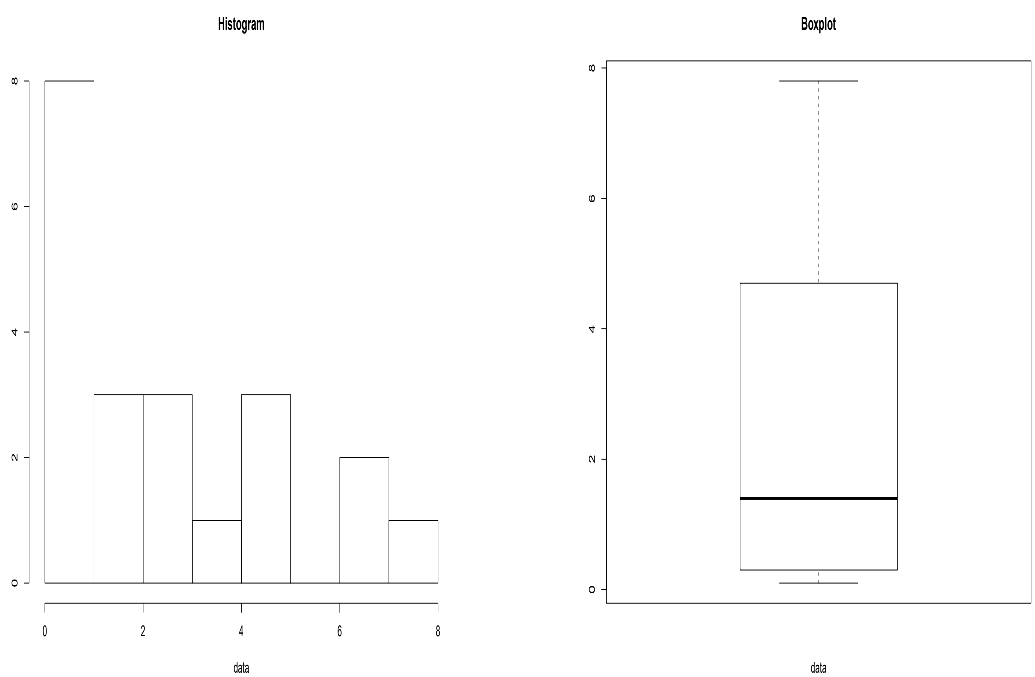
Příklad 2.1 *21× po sobě byly sledovány doby (v měsících) do poruchy daného přístroje. Naměřeny byly hodnoty:*
4.9, 6.2, 2.6, 0.6, 0.3, 2.3, 3.2, 1.4, 6.4, 4.8, 1.2
2.5, 0.2, 0.2, 0.8, 0.1, 0.1, 1.4, 7.8, 0.2, 4.7.
Uspořádáme-li pro lepší přehled tato data od nejmenší hodnoty k největší, dostaneme řadu hodnot:
0.1, 0.1, 0.2, 0.2, 0.2, 0.3, 0.6, 0.8, 1.2, 1.4, 1.4,
2.3, 2.5, 2.6, 3.2, 4.7, 4.8, 4.9, 6.2, 6.4, 7.8.

Graficky znázorněná empirická distribuční funkce je na obrázku 2.1.

Z dat pak máme:
výběrový průměr $\bar{X}_{21} = 2.471$,
výběrový rozptyl $S^2_{21} = 5.81$ *a výběrovou směrodatnou odchylku* $S_{21} = \sqrt{5.81} = 2.21$,
1.kvartil = 0.3, medián (tj. 2.kvartil) = 1.4, 3.kvartil = 4.7 a modus = 0.2.



Obrázek 2.1: Distribuční funkce pro data z příkladu 2.1.



Obrázek 2.2: Histogram a boxplot pro data z příkladu 2.1.

2.1 Bodový odhad

Definice 2.6 *Mějme náhodný výběr* $\mathbb{X} = (X_1, X_2 \ldots, X_n)^T$*, kde rozdělení náhodné veličiny* X_1 *závisí na parametru* θ*. Bodový odhad parametru* θ *je jakákoliv funkce náhodného výběru* $\theta^*(\mathbb{X})$*, jejíž funkční předpis nezávisí na* θ*. Tento odhad se nazývá nestranný, jestliže* $\mathbb{E}\theta^*(\mathbb{X}) = \theta$*.*

Výše zmíněná definice je sice matematicky přesná, avšak pro naše potřeby příliš obecná a nepřehledná. Budeme si proto pod bodovým odhadem představovat nikoliv funkci náhodného výběru $\theta^*(\mathbb{X})$, nýbrž samotné číslo $\hat{\theta}$, které získáme z realizace $(x_1, x_2 \ldots, x_n)^T$ náhodného výběru $\mathbb{X} = (X_1, X_2 \ldots, X_n)^T$ a které co nejpřesněji popisuje hledaný parametr θ.

2.1.1 Metoda momentů

Mějme $(x_1, x_2 \ldots, x_n)^T$ realizaci náhodného výběru $\mathbb{X} = (X_1, X_2 \ldots, X_n)^T$, kde s distribuční funkce F náhodné veličiny X_1 závisí na parametrech $\theta_1, ..., \theta_k \in \Theta$, kde Θ je množina, z níž může parametr pocházet (např. nezáporná reálná čísla). Předpokládejme, že tzv. i-té momenty $\mathbb{E}X_1^i$ jsou konečné pro všechna $i = 1, ...k$. Tyto momenty rovněž závisejí na $\theta_1, ..., \theta_k$. Pak položením

$$\mathbb{E}X_1^i = m_i,$$

kde m_i je $i-$tý výběrový moment získaný jako

$$m_i = \frac{1}{n}\sum_{j=1}^{n} x_j^i$$

pro všechna $i = 1, ...k$, získáme soustavu k rovnic o k neznámých $\theta_1, ..., \theta_k$, jejímž řešením jsou odhady $\hat{\theta}_1, ..., \hat{\theta}_k$.

Alternativa: Pokud $k = 2$, pak místo vztahů pro i-té momenty, $i = 1, 2$, můžeme uvažovat rovnice $\mathbb{E}X_1 = \bar{X}_n$ a var $X_1 = S_n^2$.

Nevýhoda: tento odhad má velký rozptyl.

2.1.2 Metoda maximální věrohodnosti

Předpokládejme, že máme $(x_1, x_2 \ldots, x_n)^T$ realizaci náhodného výběru $\mathbb{X} = (X_1, X_2 \ldots, X_n)^T$ z rozdělení s pravděpodobnostmi $P_\theta(X_1 = .)$ nebo s hustotou f_θ a nechť tyto pravděpodobnosti, resp. hustota, závisí na nějakém parametru $\theta \in \Theta$. Odhad $\hat{\theta}$ je maximálně věrohodným odhadem, jestliže

$$\prod_{i=1}^{n} P_{\hat{\theta}}(X_1 = x_i) = \max_{\theta\in\Theta} \prod_{i=1}^{n} P_\theta(X_1 = x_i)),$$

resp.

$$\prod_{i=1}^{n} f_{\hat{\theta}}(x_i) = \max_{\theta \in \Theta} \prod_{i=1}^{n} f_\theta(x_i).$$

Většinou je však výhodnější pracovat s logaritmem součinu, neboť tak získáme součet logaritmů. Postup je tudíž následující:

a) Diskrétní případ:

1. Vyjádřit věrohodnostní funkci $L(\theta) = \prod_{i=1}^{n} P_\theta(X_1 = x_i)$.
2. Zlogaritmovat a tím získat logaritmicko-věrohodnostní funkci $l(\theta) = \sum_{i=1}^{n} \log P_\theta(X_1 = x_i))$.
3. Položit $\frac{\partial l(\theta)}{\partial \theta} = 0$.
4. Řešení rovnice $\frac{\partial l(\theta)}{\partial \theta} = 0$ je hledaný maximálně věrohodný odhad $\hat{\theta}$.

b) Spojitý případ:

1. Vyjádřit věrohodnostní funkci $L(\theta) = \prod_{i=1}^{n} f_\theta(x_i)$.
2. Zlogaritmovat a tím získat logaritmicko-věrohodnostní funkci $l(\theta) = \sum_{i=1}^{n} \log f_\theta(x_i)$.
3. Položit $\frac{\partial l(\theta)}{\partial \theta} = 0$.
4. Řešení rovnice $\frac{\partial l(\theta)}{\partial \theta} = 0$ je hledaný maximálně věrohodný odhad $\hat{\theta}$.

2.2 Intervalový odhad

Definice 2.7 *Mějme náhodný výběr* $\mathbb{X} = (X_1, X_2 \ldots, X_n)^T$ *a číslo* $\alpha \in (0, 1)$.

1. *Dvojice* $(\theta_L^*(X_1, \ldots, X_n), \theta_U^*(X_1, \ldots, X_n))$ *se nazývá intervalový odhad parametru* θ *o spolehlivosti* $1 - \alpha$, *jestliže*

$$P(\theta_L^*(X_1, \ldots, X_n) < \theta < \theta_U^*(X_1, \ldots, X_n)) = 1 - \alpha.$$

2. $(\theta_D^*(X_1, \ldots, X_n))$ *se nazývá dolní odhad parametru* θ *o spolehlivosti* $1 - \alpha$, *jestliže*

$$P(\theta_D^*(X_1, \ldots, X_n) < \theta) = 1 - \alpha.$$

3. $(\theta_H^*(X_1, \ldots, X_n))$ *se nazývá horní odhad parametru* θ *o spolehlivosti* $1 - \alpha$, *jestliže*

$$P(\theta_H^*(X_1, \ldots, X_n) > \theta) = 1 - \alpha.$$

Konstrukce intervalového odhadu závisí na typu odhadovaného parametru a samozřejmě také na rozdělení, z něhož náhodný výběr pochází. Existuje sice obecná metoda pro konstrukci intervalového odhadu. Ta je však pro naše účely zbytečně komplikovaná, a proto se raději zaměříme jen na odhady nejčastěji používané v praxi.

2.2.1 Intervalové odhady parametrů normálního rozdělení

Věta 2.2 *Nechť* $\mathbb{X} = (X_1, X_2 \dots, X_n)^T$ *je náhodný výběr z rozdělení* $N(\mu, \sigma^2), \mu \in \mathbb{R}$ *je neznámý parametr,* $\sigma^2 > 0$ *známá konstanta. Pak*

1. $(\bar{X}_n - u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{X}_n + u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}})$ *je intervalový odhad* μ *o spolehlivosti* $1-\alpha$,
2. $\bar{X}_n - u_{1-\alpha}\frac{\sigma}{\sqrt{n}}$ *je dolní intervalový odhad* μ *o spolehlivosti* $1-\alpha$,
3. $\bar{X}_n + u_{1-\alpha}\frac{\sigma}{\sqrt{n}}$ *je horní intervalový odhad* μ *o spolehlivosti* $1-\alpha$.

Věta 2.3 *Nechť* $\mathbb{X} = (X_1, X_2 \dots, X_n)^T$ *je náhodný výběr z rozdělení* $N(\mu, \sigma^2), \mu \in \mathbb{R}$, $\sigma^2 > 0$, *oba parametry neznámé. Pak*

1. $(\bar{X}_n - t_{1-\alpha/2,n-1}\frac{S_n}{\sqrt{n}}, \bar{X}_n + t_{1-\alpha/2,n-1}\frac{S_n}{\sqrt{n}})$ *je intervalový odhad* μ *o spolehlivosti* $1-\alpha$,
2. $\bar{X}_n - t_{1-\alpha,n-1}\frac{S_n}{\sqrt{n}}$ *je dolní intervalový odhad* μ *o spolehlivosti* $1-\alpha$,
3. $\bar{X}_n + t_{1-\alpha,n-1}\frac{S_n}{\sqrt{n}}$ *je horní intervalový odhad* μ *o spolehlivosti* $1-\alpha$,
4. $(\frac{(n-1)S_n^2}{\chi^2_{1-\alpha/2,n-1}}, \frac{(n-1)S_n^2}{\chi^2_{\alpha/2,n-1}})$ *je intervalový odhad* σ^2 *o spolehlivosti* $1-\alpha$,
5. $\frac{(n-1)S_n^2}{\chi^2_{1-\alpha,n-1}}$ *je dolní intervalový odhad* σ^2 *o spolehlivosti* $1-\alpha$,
6. $\frac{(n-1)S_n^2}{\chi^2_{\alpha,n-1}}$ *je horní intervalový odhad* σ^2 *o spolehlivosti* $1-\alpha$.

2.2.2 Intervalové odhady založené na CLV

Věta 2.4 *Nechť* $\mathbb{X} = (X_1, X_2 \dots, X_n)^T$ *je náhodný výběr z libovolného rozdělení, pro které* $0 < \sigma^2 < \infty$. *Pak intervalovým odhadem* $\mu = \mathbb{E}X$ *o asymptotické spolehlivosti* $1-\alpha$ *je*

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

Důkaz. Víme, že pro velká n platí $\frac{S_n}{\sigma} \to 1$, tj. S_n se asymptoticky blíží hodnotě σ. Z CLV víme, že $\frac{\sum X_i - n\mu}{\sqrt{n\sigma^2}}$ má přibližně normální rozdělení. To znamená, že

$$
\begin{aligned}
P(u_{\frac{\alpha}{2}} \leq \frac{\sum X_i - n\mu}{\sqrt{n\sigma^2}} \leq u_{1-\frac{\alpha}{2}}) &= 1-\alpha \\
P(u_{\frac{\alpha}{2}} \leq \frac{\sum X_i - n\mu}{\sqrt{n}S_n} \leq u_{1-\frac{\alpha}{2}}) &= 1-\alpha \\
P(\frac{\sum X_i}{n} + u_{1-\frac{\alpha}{2}}\frac{S_n}{\sqrt{n}} \geq \mu \geq \frac{\sum X_i}{n} + u_{\frac{\alpha}{2}}\frac{S_n}{\sqrt{n}}) &= 1-\alpha \\
P(\bar{X}_n + u_{1-\frac{\alpha}{2}}\frac{S_n}{\sqrt{n}} \geq \mu \geq \bar{X}_n - u_{1-\frac{\alpha}{2}}\frac{S_n}{\sqrt{n}}) &= 1-\alpha,
\end{aligned}
$$

což je definice intervalového odhadu o spolehlivosti $1-\alpha$.

Speciálními případy Věty 2.4 jsou následující dvě tvrzení:

Věta 2.5 *1. Nechť $\mathbb{X} = (X_1, X_2 \ldots, X_n)^T$ je náhodný výběr z alternativního rozdělení s parametrem $0 < p < 1$. Pak intervalovým odhadem p o asymptotické spolehlivosti $1-\alpha$ je*

$$(\bar{X}_n - u_{1-\alpha/2}\sqrt{\frac{\bar{X}_n(1-\bar{X}_n)}{n}}, \bar{X}_n + u_{1-\alpha/2}\sqrt{\frac{\bar{X}_n(1-\bar{X}_n)}{n}}).$$

2. Nechť $\mathbb{X} = (X_1, X_2 \ldots, X_n)^T$ je náhodný výběr z Poissonova rozdělení s parametrem $0 < \lambda < \infty$. Pak intervalovým odhadem λ o asymptotické spolehlivosti $1-\alpha$ je

$$(\bar{X}_n - u_{1-\alpha/2}\sqrt{\frac{\bar{X}_n}{n}}, \bar{X}_n + u_{1-\alpha/2}\sqrt{\frac{\bar{X}_n}{n}}).$$

Důkaz. Myšlenka důkazu plyne z charakteristik daných rozdělení: pro alternativní rozdělení máme $\mathbb{E}X = p$ a $\operatorname{var} X = p(1-p)$ a v Poissonově rozdělení platí $\mathbb{E}X = \operatorname{var} X = \lambda$. □

2.3 Testování hypotéz

Matematická statistika popisuje obrovské množství testů o hodnotách získaných odhadů, shodnosti parametrů různých rozdělení, nezávislosti apod. V této přednášce si pouze popíšeme obecný princip testování hypotéz a jen pro ukázku jejich aplikace zde uvedeme ty nejčastěji používané testy.

2.3.1 Princip testování hypotéz

Nechť $\mathbb{X} = (X_1, X_2 \dots, X_n)^T$ je náhodný výběr z rozdělení, které závisí na parametru $\theta \in \Theta$. Chceme-li zjistit, jestli θ patří do nějaké užší podmnožiny parametrů Θ_0, nazveme tuto domněnku nulovou hypotézou, označme $H_0 : \theta \in \Theta_0$, a na základě výběru $\mathbb{X} = (X_1, X_2 \dots, X_n)^T$ testujeme tuto hypotézu proti alternativě $H_A : \theta \in \Theta \setminus \Theta_0$. To se zpravidla provede tak, že se určí množina W taková, že pokud $\mathbb{X} \in W$, pak H_0 zamítáme.

Poznámka 2.3 *Většinou testujeme $H_0 : \theta = \theta_0$, kde θ_0 je nějaká konkrétní hodnota. Přirozenou alternativou je pak $H_A : \theta \neq \theta_0$. Někdy je však smysluplnější testovat např. proti alternativě $H_A : \theta > \theta_0$. I když totiž teoreticky může být $\theta < \theta_0$, nemá smysl se tou situací zabývat, např. pokud při testování střední hodnoty $H_0 : \mu = 3$ z dat víme, že $\bar{X}_n = 5$, pak nemá smysl uvažovat situaci $\mu < 3$.*

Přitom mohou nastat následující situace:

- H_0 ve skutečnosti platí a test ji nezamítá. $\surd$
- H_0 ve skutečnosti neplatí a test ji zamítá. $\surd$
- H_0 ve skutečnosti platí, ale test ji zamítá $\rightarrow$ nastala chyba 1.druhu.
- H_0 ve skutečnosti neplatí, ale test ji nezamítá $\rightarrow$ nastala chyba 2.druhu.

Zvolíme si tzv. hladinu testu α (obvykle 0.05, někdy ale také 0.01 i méně) a W volíme tak, aby pravděpodobnost chyby 1.druhu nebyla větší než α (obvykle tak, aby byla rovna α).

Definice 2.8 *Nejmenší hladina α, pro kterou test hypotézu H_0 ve prospěch H_A zamítá, se nazývá p-hodnota.*

2.3.2 Testování pomocí intervalových odhadů

Intervalové odhady uvedené v kapitole 2.2 lze využít pro testování hypotéz o střední hodnotě rozdělení, z něhož náhodný výběr $\mathbb{X} = (X_1, X_2 \dots, X_n)^T$ pochází. Uvažujme hladinu testu α a předpokládejme, že intervalový odhad pro střední hodnotu μ o spolehlivosti $1 - \alpha$ vypočítaný z výběru $\mathbb{X}$ je (μ_L, μ_U). Pak hypotézu $H_0 : \mu = \mu_0$, kde μ_0 je libovolná konstanta, zamítáme ve prospěch hypotézy $H_A : \mu \neq \mu_0$, jestliže $\mu_0 \notin (\mu_L, \mu_U)$.

Podobně pak můžeme testovat také $H_0 : \mu = \mu_0$ oproti $H_A : \mu > \mu_0$. V tomto případě H_0 zamítáme ve prospěch hypotézy H_A, jestliže $\mu_0 \notin (-\infty, \mu_H)$, kde μ_H je horní odhad střední hodnoty μ. Analogicky se pak postupuje i při testování $H_0 : \mu = \mu_0$ oproti $H_A : \mu < \mu_0$.

2.3.3 Testování střední hodnoty normálního rozdělení - $t-$testy

Obvykle se však ve statistice používá postup "opačný", tj. místo hledání intervalů pro testovaný parametr se z dat a testované hodnoty vypočítá hodnota, která je realizací nějaké náhodné veličiny, a ta se poté porovná s příslušným kvantilem.

Tímto postupem lze samozřejmě testovat také hypotézy o střední hodnotě normálního rozdělení. Jelikož k výpočtu využijeme větu 2.1 o $t-$rozdělení trasformace výběrového průměru, nazýváme tyto testy tzv. $t-$testy.

Jednovýběrový $t-$test

Nechť $(X_1, X_2 \dots, X_n)^T$ je náhodný výběr z $N(\mu, \sigma^2)$, kde $\sigma^2 > 0$, a předpokládejme, že ani jeden parametr není znám. Testujeme-li $H_0 : \mu = \mu_0$ oproti $H_A : \mu \neq \mu_0$, je postup následující:

1. Vypočteme hodnotu $T_0 = \frac{\bar{X}_n - \mu_0}{S_n}\sqrt{n}$.
2. Pokud $|T_0| \geq t_{1-\alpha/2,n-1}$, zamítáme H_0. V opačném případě pak H_0 nezamítáme.

Jednostranný test $H_0 : \mu = \mu_0$ oproti $H_A : \mu > \mu_0$ pak probíhá podobně:

1. Vypočteme hodnotu $T_0 = \frac{\bar{X}_n - \mu_0}{S_n}\sqrt{n}$.
2. Pokud $T_0 \geq t_{1-\alpha,n-1}$, zamítáme H_0. V opačném případě pak H_0 nezamítáme.

$H_0 : \mu = \mu_0$ oproti $H_A : \mu < \mu_0$ pak zamítáme v případě, že $T_0 \leq t_{\alpha,n-1}$.

Párový $t-$test

Ve statistice se často vyskytuje situace, kdy u jednoho objektu sledujeme dva spolu související znaky najednou (např. výnosy dvou různých poboček jedné firmy apod.). Máme tedy náhodný výběr $(Y_1, Z_1), (Y_2, Z_2) \dots, (Y_n, Z_n)^T$ a chceme testovat $H_0 : \mathbb{E}Y_i - \mathbb{E}Z_i = \mu_0$ (nejčastěji $\mu_0 = 0$, tj. střední hodnoty se rovnají) oproti některé z výše uvedených alternativ. Pak postupujeme tak, že utvoříme rozdíly

$$X_1 = Y_1 - Z_1, \dots, X_n = Y_n - Z_n$$

a pokud veličiny $X_1, \dots, X_n$ pocházejí z normálního rozdělení, můžeme na ně použít výše popsaný jednovýběrový $t-$test.

Dvouvýběrový $t-$test

Mějme dva náhodné výběry $(X_1, X_2 \ldots, X_m)^T$ z $N(\mu_1, \sigma^2)$ a $(Y_1, Y_2 \ldots, Y_n)^T$ z $N(\mu_2, \sigma^2)$, kde $\sigma^2 > 0$, a předpokládejme, že tyto výběry na sobě nezávisí.

Označme $\bar{X}$ výběrový průměr výběru $(X_1, X_2 \ldots, X_m)^T$, $\bar{Y}$ výběrový průměr výběru $(Y_1, Y_2 \ldots, Y_n)^T$, S_X^2 výběrový rozptyl výběru $(X_1, X_2 \ldots, X_m)^T$ a S_Y^2 výběrový rozptyl výběru $(Y_1, Y_2 \ldots, Y_n)^T$.

Platí, že náhodná veličina

$$T = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{(m-1)S_X^2 + (n-1)S_Y^2}} \sqrt{\frac{mn(m+n-2)}{m+n}}$$

má t_{m+n-2} rozdělení.

Chceme-li tudíž testovat $H_0 : \mu_1 - \mu_2 = \mu_0$ oproti $H_A : \mu_1 - \mu_2 \neq \mu_0$, je postup následující:

1. Vypočteme hodnotu $T_0 = \frac{\bar{X} - \bar{Y} - \mu_0}{\sqrt{(m-1)S_X^2 + (n-1)S_Y^2}} \sqrt{\frac{mn(m+n-2)}{m+n}}$.
2. Pokud $|T_0| \geq t_{1-\alpha/2, m+n-2}$, zamítáme H_0. V opačném případě pak H_0 nezamítáme.

Testování normality

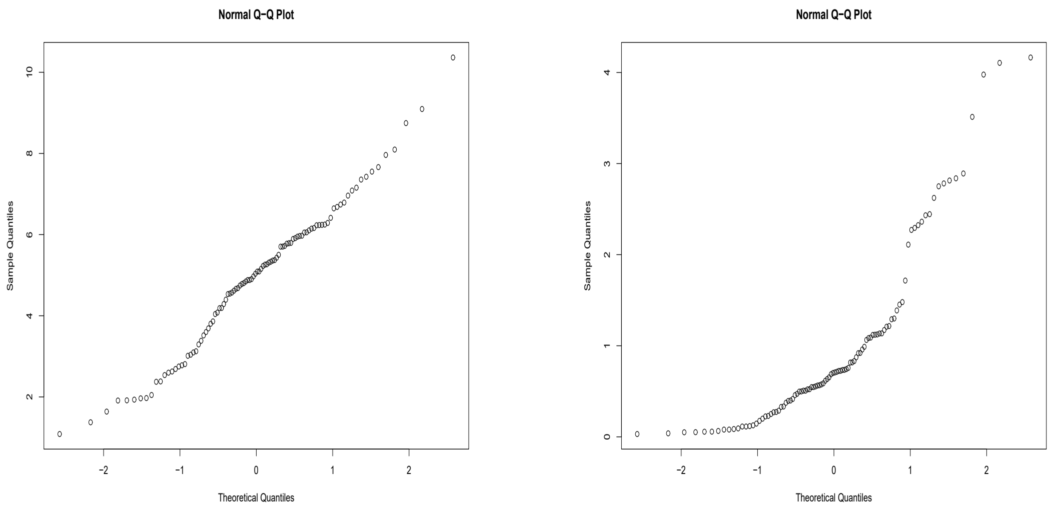
Základním předpokladem pro t-testy je normalita dat. Tu lze testovat

- analytickými testy (např. Shapiro-Wilkův test, D'Agostinův test apod. - v této přednášce vynecháme)
- graficky (např. histogram, Q-Q plot apod.)

Q-Q plot je graf, v němž jsou porovnávány teoretické kvantily testovaného - tj. v našem případě normálního - rozdělení (vodorovná osa) s empirickými kvantily získanými z dat (svislá osa). Data pak považujeme považujeme za výběr z normálního rozdělení, je-li graf přibližně lineární.

Poznámka 2.4 *Hodnoty parametrů testovaného rozdělení se přitom mohou lišit od parametrů dat, důležitý je typ rozdělení.*

Ukázka Q-Q plotu je na obrázku 2.3.



Obrázek 2.3: Q-Q plot pro data pocházející z rozdělení $N(5,4)$ (vlevo) a $Exp(1)$ (vpravo).

2.3.4 χ^2 test dobré shody

Nejprve si uveďme zobecnění binomického rozdělení pro více možných výsledků jednoho pokusu než jen úspěch nebo neúspěch.

Multinomické rozdělení

Předpokládejme, že v určitém pokusu může nastat právě jeden z jevů $A_1, A_2 \dots, A_k$ a označme pravděpodobnosti $p_i = P(A_i)$. Opakujme tento pokus $n-$krát a označme X_i počet výskytu jevu A_i v těchto n pokusech. Pak

$$P(X_1 = x_1, \dots, X_k = x_k) = \frac{n!}{x_1! \dots x_k!} p_1^{x_1} \dots p_k^{x_k}, \quad \sum_{i=1}^{k} p_i = 1, \quad \sum_{i=1}^{k} x_i = n$$

a rozdělení vektoru $(X_1, X_2 \dots, X_k)^T$ se nazývá multinomické.

Chceme-li testovat H_0, že skutečné hodnoty dílčích pravděpodobností jsou rovny předem zadaným číslům $p_1, \dots, p_k$, oproti alternativě, že aspoň jedna hodnota p_i je jiná, postupujeme následovně:

1. Vypočteme hodnotu $\chi^2 = \sum_{i=1}^{k} \frac{(X_i - np_i)^2}{np_i}$.

2. Pokud $\chi^2 > \chi^2_{1-\alpha,k-1}$, zamítáme H_0. V opačném případě pak H_0 nezamítáme.

2.3.5 Test nezávislosti v kontingenční tabulce

Tento test na rozdíl od předešlých netestuje hodnotu parametru, nýbrž vyhodnocuje, zda je možné dva výběry z veličin s diskrétním rozdělením považovat za nezávislé.

Předpokládejme, že máme výběr $(Y_1, Z_1), (Y_2, Z_2) \ldots, (Y_n, Z_n)$, kde Y_k může nabývat hodnot $1, \ldots, r$ a Z_k může nabývat hodnot $1, \ldots, c$. Označíme-li n_{ij} počet výskytů dvojice $(Y_k = i, Z_k = j)$, pak matici o rozměru $r \times c$ s prvky n_{ij} říkáme kontingenční tabulka a prvkům n_{ij} říkáme sdružené četnosti. Marginální četnosti jsou pak dány vztahy

$$n_{i.} = \sum_j n_{ij}, \quad n_{.j} = \sum_i n_{ij}.$$

K testování H_0: "Y a Z jsou nezávislé" oproti H_A: "Y a Z nejsou nezávislé" vypočteme hodnotu

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(n_{ij} - \frac{n_{i.}n_{.j}}{n})^2}{\frac{n_{i.}n_{.j}}{n}}.$$

Pokud $\chi^2 \geq \chi^2_{1-\alpha,(r-1)(c-1)}$, pak zamítáme H_0 ve prospěch H_A.