Souhrn teorie pravděpodobnosti

Pro obor Finanční matematika

Michal Kulich

Katedra pravděpodobnosti a matematické statistiky
Matematicko-fysikální fakulta University Karlovy

Tento dokument poskytuje souhrn základních poznatků z teorie pravděpodobnosti, jež jsou potřebné pro výuku předmětu „Statistika pro finanční matematiky“ v rámci bakalářského studia oboru „Finanční matematika“ na MFF UK.

Autor bude povděčen za upozornění na případné překlepy a nejasnosti, které laskavý čtenář nalezne kdekoli v tomto dokumentu.

Michal Kulich
kulich@karlin.mff.cuni.cz

Dáno v Karlíně dne 4. října 2013

1 Úvod

1.1 Kolmogorovova definice pravděpodobnosti

Nechť je dána libovolná množina Ω.

Definice 1.1. Systém $\mathcal{A}$ podmnožin množiny Ω nazveme *σ-algebrou* pokud platí
(a) $\emptyset \in \mathcal{A}$;
(b) $A \in \mathcal{A} \Rightarrow A^{\mathsf{c}} \in \mathcal{A}$;
(c) $A_1, A_2, A_3, \ldots \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

Definice 1.2. Nechť Ω je nějaká množina a $\mathcal{A}$ σ-algebra jejích podmnožin. Funkci $P : \mathcal{A} \to \langle 0, 1 \rangle$ nazveme *pravděpodobností*, právě když splňuje následující podmínky:
(a) $P(A) \geq 0$, $P(\Omega) = 1$;
(b) $A_1, A_2, A_3, \ldots \in \mathcal{A}$ a $A_i \cap A_j = \emptyset \ \forall i \neq j \Rightarrow P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

Definice 1.3. Množinu Ω nazýváme *prostor elementárních jevů*, její prvky $\omega \in \Omega$ nazýváme elementární jevy. Prvky σ-algebry $\mathcal{A}$ nazýváme *měřitelné množiny* nebo také *náhodné jevy*. Trojici $(\Omega, \mathcal{A}, P)$ nazýváme *pravděpodobnostní prostor*.

1.2 Náhodná veličina

Nechť je dán pravděpodobnostní prostor $(\Omega, \mathcal{A}, P)$.

Definice 1.4. Měřitelné zobrazení $X : (\Omega, \mathcal{A}) \to (\mathcal{X}, \mathcal{B})$, kde $\mathcal{X}$ je nějaká množina a $\mathcal{B}$ nějaká σ-algebra na $\mathcal{X}$, nazveme *náhodnou veličinou*. Množinu $\mathcal{X}$ nazýváme *výběrový prostor*.

Poznámka. Nechť jsou dány σ-algebry $\mathcal{A}$ na množině Ω a $\mathcal{B}$ na množině $\mathcal{X}$. Zobrazení $X : \Omega \to \mathcal{X}$ je měřitelné vzhledem k σ-algebrám $\mathcal{A}$ a $\mathcal{B}$, právě když $\forall B \in \mathcal{B}$ platí $\{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A}$ (tj. vzory měřitelných množin jsou měřitelné).

Příklad. (Reálná) náhodná veličina, náhodný vektor, náhodná posloupnost, náhodný proces.

1.3 Rozdělení náhodné veličiny, hustota

Definice 1.5. *Rozdělením* náhodné veličiny $X : (\Omega, \mathcal{A}) \to (\mathcal{X}, \mathcal{B})$ rozumíme indukovanou pravděpodobnostní míru P_X na $(\mathcal{X}, \mathcal{B})$ definovanou vztahem

$$P_X(B) \stackrel{\mathsf{df}}{=} P(\{\omega \in \Omega : X(\omega) \in B\}), \quad B \in \mathcal{B}.$$

Pravděpodobnost $P_X(B)$ značíme také $\mathrm{P}\left[X \in B\right]$.

Poznámka. Pravděpodobnostní prostor $(\Omega, \mathcal{A}, P)$ se pro danou náhodnou veličinu transformuje na pravděpodobnostní prostor $(\mathcal{X}, \mathcal{B}, P_X)$.

Tvrzení 1.1 (Věta o přenosu integrace)**.** Nechť h jest měřitelná funkce z $(\mathcal{X}, \mathcal{B})$ do $(\mathbb{R}, \mathcal{B}_0)$. Pak platí

$$\int_\Omega h(X(\omega))\, dP(\omega) = \int_\mathcal{X} h(x)\, dP_X(x).$$

Poznámka.
- Míra μ na $(\mathcal{X}, \mathcal{B})$ je σ-konečná, právě když existují množiny $B_1, B_2, B_3, \ldots \in \mathcal{B}$ takové, že $\mu(B_i) < \infty$ a $\cup_{i=1}^\infty B_i = \mathcal{X}$.
- Míra P_X je absolutně spojitá vzhledem k míře μ na $(\mathcal{X}, \mathcal{B})$ právě když $\forall B \in \mathcal{B}$ $\mu(B) = 0 \Rightarrow P_X(B) = 0$.

Tvrzení 1.2 (Radon-Nikodymova věta)**.** Nechť $X : (\Omega, \mathcal{A}) \to (\mathcal{X}, \mathcal{B})$ je náhodná veličina, nechť μ je σ-konečná míra na $\mathcal{X}$ a nechť P_X je absolutně spojitá vzhledem k μ. Pak existuje reálná měřitelná nezáporná funkce $f_X(x)$ taková, že pro každou měřitelnou funkci $h : (\mathcal{X}, \mathcal{B}) \to (\mathbb{R}, \mathcal{B}_0)$ platí

$$\int_\mathcal{X} h(x)\, dP_X(x) = \int_\mathcal{X} h(x) f_X(x)\, d\mu(x).$$

Funkce $f_X(x)$ je určena jednoznačně μ-skoro všude.

Definice 1.6. Funkce f_X z předchozí věty se nazývá *hustotou* náhodné veličiny X vzhledem k míře μ.

Poznámka. Zvolme nějaké $B \in \mathcal{B}$ a dosaďme za funkci h indikátor množiny B (tj. $h(x) \equiv \mathbb{I}_B(x) = 1$ pokud $x \in B$, 0 jinak). Pak máme z věty o přenosu integrace

$$\int_\Omega \mathbb{I}_B(X(\omega))\, dP(\omega) = \int_B 1\, dP_X(x) = P[X \in B]$$

a z Radon-Nikodymovy věty

$$P[X \in B] = \int_\mathcal{X} \mathbb{I}_B(x)\, dP_X(x) = \int_\mathcal{X} \mathbb{I}_B(x) f_X(x)\, d\mu(x) = \int_B f_X(x)\, d\mu(x).$$

Hustota tedy jednoznačně určuje rozdělení náhodné veličiny X.

2 Reálná náhodná veličina a její rozdělení

Nechť je dán pravděpodobnostní prostor $(\Omega, \mathcal{A}, P)$. V této kapitole se zabýváme reálnými náhodnými veličinami, tj. $X : (\Omega, \mathcal{A}) \to (\mathbb{R}, \mathcal{B}_0)$.

2.1 Charakterizace rozdělení reálné náhodné veličiny

Uveďme si několik způsobů, jak specifikovat rozdělení reálné náhodné veličiny. Výčet nebude úplný, existují i jiné způsoby (charakteristická funkce).

Hustota

Zvolme σ-konečnou míru μ na $\mathbb{R}$ tak, aby P_X byla absolutně spojitá vzhledem k μ. Podle tvrzení 1.2 a poznámky pod definicí 1.6 existuje nezáporná měřitelná $f_X : \mathbb{R} \to \mathbb{R}$ (jednoznačně určená skoro všude) taková, že $P[X \in B] = \int_B f_X(x)\, d\mu(x)$ $\forall B \in \mathcal{B}_0$. Vezmeme-li $B = \mathbb{R}$, máme $\int_{-\infty}^{\infty} f_X(x)\, d\mu(x) = 1$.

Příklad.

- P_X absolutně spojitá vzhledem k Lebesgueově míře λ: X je *spojitá náhodná veličina* [náhodná veličina se *spojitým rozdělením*]
- P_X absolutně spojitá vzhledem k čítací míře μ_S (S nejvýše spočetná množina v $\mathbb{R}$): X je *diskrétní náhodná veličina* [náhodná veličina s *diskrétním rozdělením*]
- P_X absolutně spojitá vzhledem k $\lambda + \mu_{\{0\}}$: náhodná veličina s diskrétní i spojitou složkou

Distribuční funkce

Definice 2.1. Funkci $F_X : \mathbb{R} \to \mathbb{R}$ definovanou vztahem $F_X(x) = \mathrm{P}[X \leq x]$ nazýváme *distribuční funkcí* náhodné veličiny X.

Poznámka. Distribuční funkce F_X jednoznačně charakterizuje rozdělení X [jedním směrem zřejmé, druhým směrem plyne z toho, že množiny $(-\infty, x\rangle$ generují borelovskou σ-algebru $\mathcal{B}_0$].

Poznámka.

- U spojité náhodné veličiny máme $F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt$, z čehož plyne $f_X(x) = dF_X(x)/dx$.
- U diskrétní náhodné veličiny s hodnotami v S máme $F_X(x) = \sum_{t \in S,\, t \le x} \mathrm{P}\,[X = t]$, z čehož plyne $\mathrm{P}\,[X = x] = \Delta F_X(x)$.

Tvrzení 2.1 (Vlastnosti distribuční funkce)**.**

1. F_X je neklesající, zprava spojitá
2. $\lim_{x \to -\infty} F_X(x) = 0$, $\lim_{x \to \infty} F_X(x) = 1$
3. Pro libovolnou měřitelnou $h : \mathbb{R} \to \mathbb{R}$ platí

$$\int h(x) f_X(x)\, d\mu(x) = \int h(x)\, dF_X(x)$$

Poznámka. $\int h(x)\, dF_X(x)$ je Lebesgueův-Stieltjesův integrál. Tvrzení 1.1, 1.2 a 2.1 dohromady dávají

$$\int_\Omega h(X(\omega))\, dP(\omega) = \int h(x)\, dP_X(x) = \int h(x) f_X(x)\, d\mu(x) = \int h(x)\, dF_X(x).$$

Kvantilová funkce

Definice 2.2. Nechť F_X je distribuční funkce reálné náhodné veličiny X. Funkce

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

se nazývá *kvantilová funkce* náhodné veličiny X.

Poznámka. Kvantilová funkce je neklesající a zprava spojitá. Z kvantilové funkce lze jednoznačně určit funkci distribuční. Je-li F_X rostoucí a spojitá, pak F_X^{-1} je inversní funkcí k F_X.

Definice 2.3. Nechť $\alpha \in (0,1)$. *α-kvantil* $u_X(\alpha)$ rozdělení F_X je kterékoli reálné číslo splňující $\lim_{h \searrow 0} F_X(u_X(\alpha) - h) \le \alpha$ a $F_X(u_X(\alpha)) \ge \alpha$.

Poznámka. Definicí kvantilu je více, tato jej neurčuje vždy jednoznačně. $F_X^{-1}(\alpha)$ je vždy jeden z α-kvantilů.

Definice 2.4.

- 0.5-kvantil se zove *medián* náhodné veličiny X; budeme jej značit m_X
- 0.25- a 0.75-kvantily se zovou *kvartily* náhodné veličiny X

2.2 Momenty reálné náhodné veličiny

Definice 2.5. *Střední hodnotou* $\mathsf{E}\,X$ (reálné) náhodné veličiny X rozumíme reálné číslo $\mathsf{E}\,X$ dané výrazem

$$\mathsf{E}\,X \stackrel{\text{df}}{=} \int_{\Omega} X(\omega)\,dP(\omega),$$

pokud integrál na pravé straně existuje.

Poznámka. Tuto definici lze snadno použít i v obecnějších výběrových prostorech.

Poznámka. Nechť h je reálná měřitelná funkce. Poznámka pod tvrzením 2.1 říká, že

$$\mathsf{E}\,h(X) = \int_{-\infty}^{\infty} h(x)\,dP_X(x) = \int_{-\infty}^{\infty} h(x)f_X(x)\,d\mu(x) = \int_{-\infty}^{\infty} h(x)\,dF_X(x)$$

Integrál uprostřed umíme v principu počítat pro μ Lebesgueovu nebo čítací míru. Integrál vpravo slouží k pohodlnému zápisu střední hodnoty (je kratší a nemusíme specifikovat míru μ).

Značení. Značkou $\mathcal{L}^p$ budeme značit množinu všech reálných náhodných veličin na $(\Omega, \mathcal{A}, P)$ takových, že $\mathsf{E}\,|X|^p < \infty$.

Tvrzení 2.2 (Vlastnosti střední hodnoty)**.** Nechť $X, Y \in \mathcal{L}^1$. Pak platí

1. $\mathsf{E}\,(a + bX) = a + b\mathsf{E}\,X \;\forall a, b \in \mathbb{R}$
2. $\mathsf{E}\,(X + Y) = \mathsf{E}\,X + \mathsf{E}\,Y$
3. $\mathrm{P}\,[X \leq Y] = 1 \Rightarrow \mathsf{E}\,X \leq \mathsf{E}\,Y$
4. Jestliže $\exists \mu \in \mathbb{R}\; \forall x \in \mathbb{R}\; f_X(\mu - x) = f_X(\mu + x)$ pak $\mathsf{E}\,X = \mu$

Definice 2.6.

- $\mu'_k \stackrel{\text{df}}{=} \mathsf{E}\,X^k$ se nazývá *k-tý moment* náhodné veličiny X (typicky je k přirozené, ale nemusí to tak nutně být)
- $\mu_k \stackrel{\text{df}}{=} \mathsf{E}\,(X - \mathsf{E}\,X)^k$ se nazývá *k-tý centrální moment* náhodné veličiny X
- $\mathsf{E}\,|X|^k$ se nazývá *k-tý absolutní moment* náhodné veličiny X

Definice 2.7.

- *Rozptyl* $\mathsf{var}\,X$ náhodné veličiny X je její druhý centrální moment, tj. $\mathsf{var}\,X = \mathsf{E}\,(X - \mathsf{E}\,X)^2$. Rozptyl se může také značit σ_X^2 nebo σ^2.
- *Směrodatná odchylka* σ_X náhodné veličiny X je odmocnina z jejího rozptylu, $\sigma_X = \sqrt{\mathsf{var}\,X}$.
- *Šikmost* γ_3 náhodné veličiny X je definována jako $\gamma_3 \stackrel{\text{df}}{=} \mu_3/\sigma^3$.
- *Špičatost* γ_4 náhodné veličiny X je definována jako $\gamma_4 \stackrel{\text{df}}{=} \mu_4/\sigma^4$.

Tvrzení 2.3 (Vlastnosti rozptylu)**.** Nechť X je náhodná veličina taková, že $\mathsf{var}\, X < \infty$. Pak platí

1. $\mathsf{var}\, X \geq 0$; navíc $\mathsf{var}\, X = 0 \Leftrightarrow \exists\, c \in \mathbb{R} : \mathrm{P}\,[X = c] = 1$
2. $\mathsf{var}\, X = \mathsf{E}\, X^2 - (\mathsf{E}\, X)^2$
3. $\mathsf{var}\,(a + bX) = b^2\, \mathsf{var}\, X$ pro $a, b \in \mathbb{R}$

Věta 2.4 (Jensenova nerovnost)**.** Nechť X je náhodná veličina s hodnotami v intervalu $\mathcal{I} \subseteq \mathbb{R}$ (může být nekonečný), tj. $\mathrm{P}\,[X \in \mathcal{I}] = 1$. Nechť g je [neostře] konvexní funkce na $\mathcal{I}$ taková, že existuje $\mathsf{E}\, g(X)$. Pak

$$\mathsf{E}\, g(X) \geq g(\mathsf{E}\, X)$$

a rovnost nastává právě když $g(x) = a + bx$ nebo X je konstanta.

Důsledky.

1. $\mathsf{E}\, X^2 \geq (\mathsf{E}\, X)^2$.
2. $\mathsf{E}\, \log X \leq \log \mathsf{E}\, X$ pro $X \in \mathcal{L}^1$ takovou, že $\mathrm{P}\,[X > 0] = 1$.
3. Nechť $p > q > 0$. Pak $\left(\mathsf{E}\, |X|^p\right)^{1/p} \geq \left(\mathsf{E}\, |X|^q\right)^{1/q}$.
4. Nechť $p > q > 0$ a $\mathsf{E}\, |X|^p < \infty$. Pak $\mathsf{E}\, |X|^q < \infty$.

Věta 2.5 (Markovova nerovnost)**.** Nechť $X \in \mathcal{L}^r$, kde $r > 0$. Pak pro libovolné $\varepsilon > 0$

$$\mathrm{P}\,[|X| \geq \varepsilon] \leq \frac{\mathsf{E}\, |X|^r}{\varepsilon^r}.$$

Důsledek (Čebyševova nerovnost)**.** Pro $X \in \mathcal{L}^2$ a pro libovolné $\varepsilon > 0$ platí

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

Důsledek. Pro $X \in \mathcal{L}^2$ s rozptylem $\mathsf{var}\, X = \sigma^2$ platí (například)

$$\mathrm{P}\,[|X - \mathsf{E}\, X| \geq 3\sigma] \leq \frac{1}{9}.$$

3 Náhodný vektor a mnohorozměrné rozdělení

Nechť je dán pravděpodobnostní prostor $(\Omega, \mathcal{A}, P)$. V této kapitole se zabýváme náhodnými vektory, tj. $\boldsymbol{X} : (\Omega, \mathcal{A}) \to (\mathbb{R}^n, \mathcal{B}_0^n)$.

3.1 Rozdělení náhodného vektoru

Poznámka. Náhodný vektor je (do sloupce) uspořádaná n-tice náhodných veličin, tj.

$$\boldsymbol{X}(\omega) = (X_1(\omega), \dots, X_n(\omega))^{\mathsf{T}}.$$

Definice 3.1. $\mathcal{B}_0^n$ je borelovská σ-algebra v $\mathbb{R}^n$ definovaná jako

$$\mathcal{B}_0^n = \sigma\{(a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n); a_1 < b_1, a_2 < b_2, \dots, a_n < b_n \in \mathbb{R}\}$$

Poznámka. Míru na $(\mathbb{R}^n, \mathcal{B}_0^n)$ stačí definovat na některém generátoru borelovské σ-algebry, např. na otevřených nebo uzavřených n-rozměrných kvádrech.

Hustota náhodného vektoru

Poznámka. Podle Radon-Nikodymovy věty (Tvrzení 1.2) platí: Jestliže $P_{\boldsymbol{X}}$ je absolutně spojitá vzhledem k σ-konečné míře μ na $(\mathbb{R}^n, \mathcal{B}_0^n)$, tj. $\mu(B) = 0 \Rightarrow \mathrm{P}[\boldsymbol{X} \in B] = 0$ pro $B \in \mathcal{B}_0^n$, pak existuje jednoznačně (až na množiny s nulovou mírou μ) daná nezáporná měřitelná funkce $f_{\boldsymbol{X}}(\boldsymbol{x}) : \mathbb{R}^n \to \mathbb{R}$, zvaná *hustota* náhodného vektoru $\boldsymbol{X}$ taková, že

$$\int_\Omega h(\boldsymbol{X}(\omega))\, dP(\omega) = \int_{\mathbb{R}^n} h(\boldsymbol{x})\, dP_{\boldsymbol{X}}(\boldsymbol{x}) = \int_{\mathbb{R}^n} h(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x})\, d\mu(\boldsymbol{x})$$

pro každou měřitelnou funkci $h : \mathbb{R}^n \to \mathbb{R}$.

Značení. V dalším výkladu používáme v argumentech funkcí definovaných na $\mathbb{R}^n$ záměnně značení $\boldsymbol{x}$ a $(x_1, \dots, x_n)$.

Poznámka.

- Jestliže je rozdělení $\boldsymbol{X}$ absolutně spojité vzhledem k Lebesgueově míře λ^n na $\mathbb{R}^n$, pak rozdělení náhodného vektoru $\boldsymbol{X}$ nazýváme *spojité* a $\mathrm{P}[\boldsymbol{X} \in B]$ počítáme jako
$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathbb{I}_B(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x})\, dx_1\, dx_2 \ldots dx_n.$$
- Nechť je rozdělení $\boldsymbol{X}$ absolutně spojité vzhledem k čítací míře μ_S na $\mathbb{R}^n$, kde S je nejvýše spočetná množina bodů v $\mathbb{R}^n$ tvaru $S_1 \times S_2 \times \cdots S_n$ a $S_k = \{t_{k,1}, t_{k,2}, \ldots\}$. Pak rozdělení náhodného vektoru $\boldsymbol{X}$ nazýváme *diskrétní* a $\mathrm{P}[\boldsymbol{X} \in B]$ počítáme jako
$$\sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} \cdots \sum_{i_n=1}^{\infty} \mathbb{I}_B(t_{1,i_1}, t_{2,i_2}, \ldots, t_{n,i_n}) \mathrm{P}[\boldsymbol{X} = (t_{1,i_1}, t_{2,i_2}, \ldots, t_{n,i_n})] .$$
- Jestliže náhodný vektor obsahuje diskrétní i spojité složky, pak jeho rozdělení není ani diskrétní, ani spojité. Přesto pro něj máme použitelnou hustotu, s jejíž pomocí můžeme vyjádřit $\mathrm{P}[\boldsymbol{X} \in B]$
- Jestliže všechny složky náhodného vektoru jsou spojité, neznamená to nutně, že vektor jako celek má spojité rozdělení. Příklad: rozdělení na jednotkové kružnici v $\mathbb{R}^2$.

Distribuční funkce náhodného vektoru

Definice 3.2. Funkci
$$F_{\boldsymbol{X}}(\boldsymbol{x}) = \mathrm{P}[X_1 \le x_1, \ldots, X_n \le x_n]$$
nazýváme *distribuční funkcí* náhodného vektoru.

Tvrzení 3.1. Jestliže je rozdělení $\boldsymbol{X}$ absolutně spojité vzhledem k Lebesgueově míře λ^n, pak
$$F_{\boldsymbol{X}}(\boldsymbol{x}) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{\boldsymbol{X}}(x_1, \ldots, x_n)\, dx_1 \ldots dx_n$$
a naopak,
$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{\partial^n F_{\boldsymbol{X}}(x_1, \ldots, x_n)}{\partial x_1 \cdots \partial x_n} \quad \text{skoro všude.}$$

Poznámka.
1. Distribuční funkce jednoznačně určuje rozdělení náhodného vektoru $\boldsymbol{X}$.

2. Kvůli jednoduššímu značení budeme psát

$$\int_{\mathbb{R}^n} h(\boldsymbol{x})\, dF_{\boldsymbol{X}}(\boldsymbol{x}) \overset{\mathsf{df}}{=} \int_{\mathbb{R}^n} h(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x})\, d\mu(\boldsymbol{x}).$$

Příklad. Nechť je dán náhodný vektor $\boldsymbol{X} = (X_1, X_2)^{\mathsf{T}}$. Pomocí jeho distribuční funkce vyjádřete pravděpodobnost

$$\mathrm{P}\left[a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2\right].$$

Sdružené a marginální rozdělení

Definice 3.3.
- Rozdělení celého náhodného vektoru $\boldsymbol{X} = (X_1, \ldots, X_n)^{\mathsf{T}}$ se říká *sdružené rozdělení*. Jeho distribuční funkce a hustota se nazývají *sdružená distribuční funkce* a *sdružená hustota*.
- Rozdělením jednotlivých náhodných veličin $X_1, \ldots, X_n$ se říká *marginální rozdělení*. Jejich distribuční funkce a hustoty se nazývají *marginální distribuční funkce* a *marginální hustoty*.

Tvrzení 3.2. Ze sdruženého rozdělení $\boldsymbol{X}$ lze jednoznačně určit marginální rozdělení $X_1, \ldots, X_n$. Platí

$$F_{X_i}(u) = \lim_{x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n \to \infty} F_{\boldsymbol{X}}(x_1, \ldots, x_{i-1}, u, x_{i+1}, \ldots, x_n)$$

a pro spojitý náhodný vektor navíc

$$f_{X_i}(u) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\boldsymbol{X}}(x_1, \ldots, x_{i-1}, u, x_{i+1}, \ldots, x_n)\, dx_1 \ldots dx_{i-1}\, dx_{i+1} \ldots dx_n. \tag{3.1}$$

3.2 Momenty

Střední hodnota

Poznámka. Podle definice 2.5 a poznámek na str. 9 a 10 máme pro libovolnou měřitelnou funkci $h : \mathbb{R}^n \to \mathbb{R}$

$$\mathsf{E}\, h(\boldsymbol{X}) = \int_{\Omega} h(\boldsymbol{X}(\omega))\, dP(\omega) = \int_{\mathbb{R}^n} h(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x})\, d\mu(\boldsymbol{x}) = \int_{\mathbb{R}^n} h(\boldsymbol{x})\, dF_{\boldsymbol{X}}(\boldsymbol{x}).$$

Definice 3.4. Pro měřitelnou $g : \mathbb{R}^n \to \mathbb{R}^m$ definujeme

$$\mathsf{E}\, g(\boldsymbol{X}) = (\mathsf{E}\, g_1(\boldsymbol{X}), \ldots, \mathsf{E}\, g_m(\boldsymbol{X}))^{\mathsf{T}}.$$

Poznámka. Střední hodnota náhodného vektoru je tedy vektorem středních hodnot jejích složek. Střední hodnota matice náhodných veličin je maticí středních hodnot jednotlivých prvků.

Rozptyl

V této části nechť $X_i \in \mathcal{L}^2$, $i = 1, \dots, n$.

Značení. Nechť $\boldsymbol{a}$ je sloupcový vektor v $\mathbb{R}^n$. Pak definujeme $\boldsymbol{a}^{\otimes 2} \stackrel{\mathsf{df}}{=} \boldsymbol{a}\boldsymbol{a}^\mathsf{T}$ (matice součinů prvků a_i a a_j).

Definice 3.5.

1. Matice
$$\mathsf{var}\,\boldsymbol{X} \stackrel{\mathsf{df}}{=} \mathsf{E}\,(\boldsymbol{X} - \mathsf{E}\,\boldsymbol{X})^{\otimes 2} = \mathsf{E}\,(\boldsymbol{X} - \mathsf{E}\,\boldsymbol{X})(\boldsymbol{X} - \mathsf{E}\,\boldsymbol{X})^\mathsf{T}$$
se nazývá *rozptylová* (varianční) matice náhodného vektoru $\boldsymbol{X}$.

2. (i, j)-tý prvek matice $\mathsf{var}\,\boldsymbol{X}$ jest $\mathsf{E}\,(X_i - \mathsf{E}\,X_i)(X_j - \mathsf{E}\,X_j)$ a nazývá se *kovariance* náhodných veličin X_i a X_j.

3. Rozdělíme-li $\boldsymbol{X}$ na $\boldsymbol{X} = \binom{\boldsymbol{X}_1}{\boldsymbol{X}_2}$, pak matice
$$\mathsf{cov}\,(\boldsymbol{X}_1, \boldsymbol{X}_2) \stackrel{\mathsf{df}}{=} \mathsf{E}\,(\boldsymbol{X}_1 - \mathsf{E}\,\boldsymbol{X}_1)(\boldsymbol{X}_2 - \mathsf{E}\,\boldsymbol{X}_2)^\mathsf{T}$$
se nazývá *kovarianční matice* vektorů $\boldsymbol{X}_1$ a $\boldsymbol{X}_2$.

Tvrzení 3.3.

1. i-tý diagonální prvek matice $\mathsf{var}\,\boldsymbol{X}$ je $\mathsf{var}\,X_i$.
2. $\mathsf{var}\,\boldsymbol{X}$ je positivně semidefinitní matice [značíme $\mathsf{var}\,\boldsymbol{X} \geq 0$], tj. $\forall \boldsymbol{c} \in \mathbb{R}^n$ $\boldsymbol{c}^\mathsf{T}(\mathsf{var}\,\boldsymbol{X})\boldsymbol{c} \geq 0$.
3. $\mathsf{var}\,\boldsymbol{X} = \mathsf{E}\,\boldsymbol{X}^{\otimes 2} - (\mathsf{E}\,\boldsymbol{X})^{\otimes 2}$, $\mathsf{cov}\,(\boldsymbol{X}_1, \boldsymbol{X}_2) = \mathsf{E}\,\boldsymbol{X}_1\boldsymbol{X}_2^\mathsf{T} - \mathsf{E}\,\boldsymbol{X}_1\,\mathsf{E}\,\boldsymbol{X}_2^\mathsf{T}$.
4. $\mathsf{cov}\,(\boldsymbol{X}_1, \boldsymbol{X}_2) = \mathsf{cov}\,(\boldsymbol{X}_2, \boldsymbol{X}_1)^\mathsf{T}$, $\mathsf{cov}\,(\boldsymbol{X}, \boldsymbol{X}) = \mathsf{var}\,\boldsymbol{X}$.
5. Pro vektory $\boldsymbol{a}$, $\boldsymbol{c}$ a matice B, D vhodných dimenzí platí
$$\mathsf{cov}\,(\boldsymbol{a} + B\boldsymbol{X}_1, \boldsymbol{c} + D\boldsymbol{X}_2) = B\,\mathsf{cov}\,(\boldsymbol{X}_1, \boldsymbol{X}_2)D^\mathsf{T}.$$
Speciálně: $\mathsf{var}\,(\boldsymbol{a} + B\boldsymbol{X}) = B\,(\mathsf{var}\,\boldsymbol{X})B^\mathsf{T}$.

Důsledek. Dosadíme-li v 5. části předchozího tvrzení $\boldsymbol{X}_1 = \boldsymbol{X}_2 = (X_1, \dots, X_n)^\mathsf{T}$, $\boldsymbol{a} = \boldsymbol{c} = \boldsymbol{0}$ a $B = D = (1, \dots, 1)$, dostaneme vztah pro rozptyl součtu n náhodných veličin:
$$\mathsf{var}\sum_{i=1}^{n} X_i = \sum_{i=1}^{n} \mathsf{var}\,X_i + 2\sum_{i=2}^{n}\sum_{j=1}^{i-1} \mathsf{cov}\,(X_i, X_j) \tag{3.2}$$

Tvrzení 3.4. Nechť $\boldsymbol{X}$ a $\boldsymbol{Y}$ jsou náhodné vektory v $\mathbb{R}^n$, jejichž složky mají konečné druhé momenty. Pak platí
$$\mathsf{var}\,(\boldsymbol{X} + \boldsymbol{Y}) = \mathsf{var}\,\boldsymbol{X} + \mathsf{cov}\,(\boldsymbol{X}, \boldsymbol{Y}) + \mathsf{cov}\,(\boldsymbol{X}, \boldsymbol{Y})^\mathsf{T} + \mathsf{var}\,\boldsymbol{Y}$$

3.3 Nezávislost

Definice 3.6.

- Náhodné veličiny $X_1, \dots, X_n$ nazveme *(vzájemně) nezávislé* právě když pro každý bod $\boldsymbol{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ platí
$$F_{\boldsymbol{X}}(\boldsymbol{x}) = F_{X_1}(x_1) \cdot \dots \cdot F_{X_n}(x_n).$$
- Náhodné veličiny $X_1, X_2, \dots$ nazveme *(vzájemně) nezávislé* právě když
$$\forall k > 1 \quad \forall n_1, \dots, n_k \in \mathbb{N} \quad X_{n_1}, \dots, X_{n_k} \text{ jsou nezávislé.}$$
- Náhodné vektory $\boldsymbol{X}_1$ s n_1 složkami a $\boldsymbol{X}_2$ s n_2 složkami nazveme *nezávislé* právě když pro každý bod $\boldsymbol{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ platí
$$F_{\boldsymbol{X}}(\boldsymbol{x}) = F_{\boldsymbol{X}_1}(\boldsymbol{x}_1) \cdot F_{\boldsymbol{X}_2}(\boldsymbol{x}_2),$$
kde $n = n_1 + n_2$, $\boldsymbol{X} = (\boldsymbol{X}_1^\mathsf{T}, \boldsymbol{X}_2^\mathsf{T})^\mathsf{T}$ a $\boldsymbol{x} = (\boldsymbol{x}_1^\mathsf{T}, \boldsymbol{x}_2^\mathsf{T})^\mathsf{T}$.

Poznámka. Pro nezávislé náhodné veličiny platí, že vezmeme-li libovolné borelovské množiny $B_1, \dots, B_n \in \mathcal{B}_0$, pak
$$\mathrm{P}\left[\boldsymbol{X} \in B_1 \times \dots \times B_n\right] = \mathrm{P}\left[X_1 \in B_1\right] \cdot \dots \cdot \mathrm{P}\left[X_n \in B_n\right],$$
neboli náhodné jevy $[X_i \in B_i]$ jsou vzájemně nezávislé. Dále máme např.
$$\mathrm{P}\left[\, X_1 \in B_1 \mid X_2 \in B_2, \dots, X_n \in B_n \,\right] = \mathrm{P}\left[X_1 \in B_1\right].$$
Pro nezávislé náhodné vektory platí, že vezmeme-li libovolné borelovské množiny $B_1 \in \mathcal{B}_0^{n_1}$ a $B_2 \in \mathcal{B}_0^{n_2}$, pak
$$\mathrm{P}\left[\boldsymbol{X} \in B_1 \times B_2\right] = \mathrm{P}\left[\boldsymbol{X}_1 \in B_1\right] \cdot \mathrm{P}\left[\boldsymbol{X}_2 \in B_2\right],$$
neboli náhodné jevy $[\boldsymbol{X}_i \in B_i]$ jsou nezávislé.

Tvrzení 3.5. Nechť náhodná veličina X_i má hustotu f_{X_i} vzhledem k σ-konečné míře μ_i, $i = 1, \dots, n$. Pak jsou náhodné veličiny $X_1, \dots, X_n$ vzájemně nezávislé právě když vektor $\boldsymbol{X} = (X_1, \dots, X_n)^\mathsf{T}$ má hustotu $f_{\boldsymbol{X}}$ vzhledem k součinové míře $\mu = \mu_1 \otimes \dots \otimes \mu_n$ a platí
$$f_{\boldsymbol{X}}(x_1, \dots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i).$$

Tvrzení 3.6. Nechť $\boldsymbol{X}_1$ a $\boldsymbol{X}_2$ jsou nezávislé náhodné vektory a $g_1 : \mathbb{R}^{n_1} \to \mathbb{R}^q$ a $g_2 : \mathbb{R}^{n_2} \to \mathbb{R}^s$ jsou libovolné měřitelné funkce. Pak $g_1(\boldsymbol{X}_1)$ a $g_2(\boldsymbol{X}_2)$ jsou nezávislé náhodné vektory.

Tvrzení 3.7. Nechť $X_1, \dots, X_n$ jsou nezávislé.
(i) Jsou-li $X_i \in \mathcal{L}^1$, pak $\mathsf{E}\,(X_1 \cdot \dots \cdot X_n) = \mathsf{E}\,X_1 \cdot \dots \cdot \mathsf{E}\,X_n$.
(ii) Jsou-li $X_i \in \mathcal{L}^2$, pak $\mathsf{cov}\,(X_i, X_j) = 0\ \forall i \neq j$.
(iii) Jsou-li $X_i \in \mathcal{L}^2$ a $\sigma_i^2 = \mathsf{var}\,X_i$, pak $\mathsf{var}\,\boldsymbol{X} = \mathrm{diag}\,(\sigma_1^2, \dots, \sigma_n^2)$.

Poznámka. Z vlastností (i) – (iii) předchozího tvrzení neplyne bez dalších podmínek nezávislost.

3.4 Korelace

Definice 3.7. Nechť X, Y jsou náhodné veličiny s kladnými a konečnými rozptyly. *Korelační koeficient* veličin X a Y se značí $\varrho(X, Y)$ nebo $\mathsf{cor}\,(X, Y)$ a je definován vztahem

$$\varrho(X, Y) = \frac{\mathsf{cov}\,(X, Y)}{\sqrt{\mathsf{var}\,X\,\mathsf{var}\,Y}}.$$

Tvrzení 3.8 (Cauchyova-Schwartzova nerovnost)**.**
Nechť $X, Y \in \mathcal{L}^2$. Pak $(\mathsf{E}\,XY)^2 \leq \mathsf{E}\,X^2\,\mathsf{E}\,Y^2$ a rovnost platí, právě když $X = bY$ s.j. pro nějaké $b \neq 0$.

Důsledek. Pro jakékoli veličiny $X, Y \in \mathcal{L}^2$ máme $|\mathsf{cov}\,(X, Y)| \leq \sqrt{\mathsf{var}\,X\,\mathsf{var}\,Y}$ a tudíž, pokud mají nenulový rozptyl, také $|\varrho(X, Y)| \leq 1$.

Tvrzení 3.9 (Vlastnosti korelačního koeficientu)**.** Nechť $X, Y \in \mathcal{L}^2$, $\mathsf{var}\,X > 0$, $\mathsf{var}\,Y > 0$.

1. $\varrho(X, Y) = \varrho(Y, X)$;
2. $-1 \leq \varrho(X, Y) \leq 1$,
   - $\varrho(X, Y) = 1$ právě když $X = a + bY$ s.j., kde $b > 0$;
   - $\varrho(X, Y) = -1$ právě když $X = a + bY$ s.j., kde $b < 0$;
3. $\varrho(a + bX, c + dY) = \mathrm{sgn}\,(bd)\varrho(X, Y)$.

Poznámka.

- Je-li $\varrho(X, Y) = 0$ (nebo $\mathsf{cov}\,(X, Y) = 0$), náhodným veličinám X, Y se říká *nekorelované veličiny.* Nezávislé veličiny jsou i nekorelované, opak nutně neplatí.
- Korelační koeficient měří sílu *lineárního* vztahu mezi X a Y.

Definice 3.8. Nechť $\boldsymbol{X} = (X_1, \ldots, X_n)^{\mathsf{T}}$ a $\boldsymbol{Y} = (Y_1, \ldots, Y_m)^{\mathsf{T}}$ jsou dva náhodné vektory se složkami, jež mají konečné a kladné rozptyly. *Korelační maticí* $\mathsf{cor}\,(\boldsymbol{X}, \boldsymbol{Y})$ vektorů $\boldsymbol{X}$ a $\boldsymbol{Y}$ rozumíme matici typu $n \times m$ se složkami $\varrho(X_i, Y_j)$ na místě (i, j).

Poznámka. Korelační matice $\mathsf{cor}\,(\boldsymbol{X}, \boldsymbol{X})$ má tvar

$$\mathsf{cor}\,(\boldsymbol{X}, \boldsymbol{X}) = \begin{pmatrix} 1 & \varrho_{12} & \ldots & \varrho_{1n} \\ \varrho_{12} & 1 & \ldots & \varrho_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \varrho_{1n} & \varrho_{2n} & \ldots & 1 \end{pmatrix},$$

kde $\varrho_{jk} = \varrho(X_i, X_j)$. Je-li $V = \mathsf{var}\,\boldsymbol{X}$, $\sigma_i = \sqrt{\mathsf{var}\,X_i}$ a $D = \operatorname{diag}(\sigma_1, \ldots, \sigma_n)$, pak máme $\mathsf{cor}\,(\boldsymbol{X}, \boldsymbol{X}) = D^{-1} V D^{-1}$.

4 Podmíněné rozdělení

Nechť je dán pravděpodobnostní prostor $(\Omega, \mathcal{A}, P)$. V této kapitole uvažujeme náhodné veličiny a náhodné vektory definované na tomto prostoru. Připomeňme si nejdříve definici podmíněné pravděpodobnosti.

Poznámka. Nechť A a B jsou náhodné jevy a $\mathrm{P}(B) > 0$. Podmíněná pravděpodobnost $\mathrm{P}\left(A \mid B \right)$ jevu A za podmínky, že nastal jev B, je definována podílem

$$\mathrm{P}\left(A \mid B \right) \overset{\mathsf{df}}{=} \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}.$$

Jsou-li oba jevy nezávislé, pak $\mathrm{P}\left(A \mid B \right) = \mathrm{P}(A)$.

4.1 Podmíněná hustota

Uvažujme náhodný vektor $\boldsymbol{X} = (X_1, \ldots, X_n)^\mathsf{T}$, který je rozdělen na dva podvektory $\boldsymbol{Y} = (X_1, \ldots, X_r)^\mathsf{T}$ a $\boldsymbol{Z} = (X_{r+1}, \ldots, X_n)^\mathsf{T}$, $1 \leq r < n$. Chceme zkoumat rozdělení náhodného vektoru $\boldsymbol{Y}$ v situaci, kdy víme, že náhodný vektor $\boldsymbol{Z}$ nabyl konkrétní hodnoty $\boldsymbol{z} \in \mathbb{R}^{n-r}$.

Definice 4.1. Nechť náhodný vektor $\boldsymbol{Y}$ má hustotu $f_{\boldsymbol{Y}}(\boldsymbol{y})$ vzhledem k σ-konečné míře μ_1 na $(\mathbb{R}^r, \mathcal{B}_0^r)$. Nechť náhodný vektor $\boldsymbol{Z}$ má hustotu $f_{\boldsymbol{Z}}(\boldsymbol{z})$ vzhledem k σ-konečné míře μ_2 na $(\mathbb{R}^{n-r}, \mathcal{B}_0^{n-r})$. Nechť náhodný vektor $\boldsymbol{X} = (\boldsymbol{Y}^\mathsf{T}, \boldsymbol{Z}^\mathsf{T})^\mathsf{T}$ má hustotu $f_{\boldsymbol{X}}(\boldsymbol{y}, \boldsymbol{z})$ vzhledem k součinové míře $\mu = \mu_1 \times \mu_2$ na $(\mathbb{R}^n, \mathcal{B}_0^n)$.

Podmíněnou hustotou náhodného vektoru $\boldsymbol{Y}$, je-li dáno $\boldsymbol{Z} = \boldsymbol{z}$ nazveme libovolnou nezápornou měřitelnou funkci $f(\boldsymbol{y} \mid \boldsymbol{z})$, která pro všechna $B \in \mathcal{B}_0^r$ a $C \in \mathcal{B}_0^{n-r}$ splňuje rovnost

$$\mathrm{P}\left[\boldsymbol{Y} \in B, \boldsymbol{Z} \in C\right] = \int_C \left[\int_B f(\boldsymbol{y} \mid \boldsymbol{z}) \, d\mu_1(\boldsymbol{y}) \right] f_{\boldsymbol{Z}}(\boldsymbol{z}) \, d\mu_2(\boldsymbol{z}). \tag{4.1}$$

Poznámka. Podmíněná hustota za daných předpokladů existuje a je jednoznačně určena μ_1-skoro všude. Předpoklad existence hustoty $f_{\boldsymbol{X}}$ vzhledem k součinové míře $\mu = \mu_1 \times \mu_2$ je závažný (někdy neplatí) a nutný (jinak nelze podmíněnou hustotu rovností (4.1) definovat).

Poznámka (Výpočet podmíněné hustoty)**.** Levá strana rovnosti (4.1) je vlastně

$$\int_{B\times C} f_{\boldsymbol{X}}(\boldsymbol{y},\boldsymbol{z})\,d\mu(\boldsymbol{y},\boldsymbol{z}).$$

Pravá strana dává

$$\int_{B\times C} f(\boldsymbol{y}\mid\boldsymbol{z})f_{\boldsymbol{Z}}(\boldsymbol{z})\,d\mu(\boldsymbol{y},\boldsymbol{z}).$$

Rovnost pro každé B a C nastane právě když

$$f_{\boldsymbol{X}}(\boldsymbol{y},\boldsymbol{z}) = f(\boldsymbol{y}\mid\boldsymbol{z})f_{\boldsymbol{Z}}(\boldsymbol{z})$$

μ-skoro všude. Podmíněnou hustotu tudíž můžeme počítat vztahem

$$f(\boldsymbol{y}\mid\boldsymbol{z}) = \frac{f_{\boldsymbol{X}}(\boldsymbol{y},\boldsymbol{z})}{f_{\boldsymbol{Z}}(\boldsymbol{z})}$$

pro $\boldsymbol{z}$ taková, že $f_{\boldsymbol{Z}}(\boldsymbol{z}) \neq 0$.

Věta 4.1 (Bayesova)**.** Platí-li podmínky definice 4.1, pak podmíněná hustota $p(\boldsymbol{z}\mid\boldsymbol{y})$ náhodného vektoru $\boldsymbol{Z}$, je-li dáno $\boldsymbol{Y}=\boldsymbol{y}$ je rovna

$$p(\boldsymbol{z}\mid\boldsymbol{y}) = \begin{cases} \dfrac{f(\boldsymbol{y}\mid\boldsymbol{z})f_{\boldsymbol{Z}}(\boldsymbol{z})}{\displaystyle\int_{\mathbb{R}^{n-r}} f(\boldsymbol{y}\mid\boldsymbol{z})f_{\boldsymbol{Z}}(\boldsymbol{z})\,d\mu_2(\boldsymbol{z})} & \text{pokud jmenovatel není roven 0,} \\ 0 & \text{jinak.} \end{cases}$$

4.2 Podmíněná střední hodnota

Stále se zabýváme náhodným vektorem $\boldsymbol{X} = (X_1,\dots,X_n)^\mathsf{T}$ rozděleným na dva podvektory $\boldsymbol{Y} = (X_1,\dots,X_r)^\mathsf{T}$ a $\boldsymbol{Z} = (X_{r+1},\dots,X_n)^\mathsf{T}$, $1 \le r < n$. Máme tedy $\boldsymbol{X} = (\boldsymbol{Y}^\mathsf{T},\boldsymbol{Z}^\mathsf{T})^\mathsf{T}$.

Definice 4.2. Nechť $h(\boldsymbol{y},\boldsymbol{z})$ je měřitelná funkce $\mathbb{R}^n \to \mathbb{R}^m$. Označme $\boldsymbol{U} = h(\boldsymbol{X}) = h(\boldsymbol{Y},\boldsymbol{Z})$.

1. *Podmíněná střední hodnota* $\mathsf{E}(\boldsymbol{U}\mid\boldsymbol{Z}=\boldsymbol{z})$ náhodného vektoru $\boldsymbol{U} \equiv h(\boldsymbol{Y},\boldsymbol{Z})$, je-li dáno $\boldsymbol{Z}=\boldsymbol{z}$ je definována výrazem

   $$\mathsf{E}(\boldsymbol{U}\mid\boldsymbol{Z}=\boldsymbol{z}) = \int_{\mathbb{R}^r} h(\boldsymbol{y},\boldsymbol{z})f(\boldsymbol{y}\mid\boldsymbol{z})\,d\mu_1(\boldsymbol{y})$$

   (pokud existuje).

2. Označme $\phi(\boldsymbol{z}) = \mathsf{E}\,(\boldsymbol{U} \mid \boldsymbol{Z} = \boldsymbol{z}\,)$ (je to nějaká měřitelná funkce z $\mathbb{R}^{n-r}$ do $\mathbb{R}^m$). Náhodný vektor $\phi(\boldsymbol{Z})$ značíme $\mathsf{E}\,(\boldsymbol{U} \mid \boldsymbol{Z}\,)$ a nazýváme jej *podmíněnou střední hodnotou* náhodného vektoru $\boldsymbol{U} = h(\boldsymbol{Y}, \boldsymbol{Z})$ při daném (leč neurčeném) $\boldsymbol{Z}$.

Poznámka. Jak je řečeno výše, podmíněná střední hodnota $\mathsf{E}\,(\cdot \mid \boldsymbol{Z} = \boldsymbol{z}\,)$ je funkce argumentu $\boldsymbol{z}$ zobrazující z $\mathbb{R}^{n-r}$ do $\mathbb{R}^m$. Pro pevné $\boldsymbol{z}$ je to konstanta (v $\mathbb{R}^m$). Podmíněná střední hodnota $\mathsf{E}\,(\cdot \mid \boldsymbol{Z}\,)$ je náhodný vektor o m složkách; jeho realizovaná hodnota závisí na realizované hodnotě náhodného vektoru $\boldsymbol{Z}$.

Nyní přibereme do úvahy ještě další měřitelné funkce $h_1, h_2 : \mathbb{R}^n \to \mathbb{R}^m$ a $\psi : \mathbb{R}^{n-r} \to \mathbb{R}$. Označme $\boldsymbol{U}_1 = h_1(\boldsymbol{Y}, \boldsymbol{Z})$ a $\boldsymbol{U}_2 = h_2(\boldsymbol{Y}, \boldsymbol{Z})$. Nechť všechny složky $\boldsymbol{U}$, $\boldsymbol{U}_1$ a $\boldsymbol{U}_2$ mají konečné první momenty.

Věta 4.2 (Vlastnosti podmíněné střední hodnoty)**.**

1. $\mathsf{E}\,(\boldsymbol{a} \mid \boldsymbol{Z}\,) = \boldsymbol{a}$ pro jakékoli $\boldsymbol{a} \in \mathbb{R}^m$.
2. $\mathsf{E}\,\big[\mathsf{E}\,(\boldsymbol{U} \mid \boldsymbol{Z}\,)\big] = \mathsf{E}\,\boldsymbol{U}$.
3. $\mathsf{E}\,(a_1\boldsymbol{U}_1 + a_2\boldsymbol{U}_2 \mid \boldsymbol{Z}\,) = a_1\mathsf{E}\,(\boldsymbol{U}_1 \mid \boldsymbol{Z}\,) + a_2\mathsf{E}\,(\boldsymbol{U}_2 \mid \boldsymbol{Z}\,)$ pro jakékoli $a_1, a_2 \in \mathbb{R}$.
4. $\mathsf{E}\,(\psi(\boldsymbol{Z})\boldsymbol{U} \mid \boldsymbol{Z}\,) = \psi(\boldsymbol{Z})\mathsf{E}\,(\boldsymbol{U} \mid \boldsymbol{Z}\,)$.

Věta 4.3. Nechť všechny složky $\boldsymbol{U} = h(\boldsymbol{Y}, \boldsymbol{Z})$ mají konečný rozptyl a nechť τ je jakákoli měřitelná funkce z $\mathbb{R}^{n-r}$ do $\mathbb{R}^m$ taková, že všechny složky $\tau(\boldsymbol{Z})$ mají konečný rozptyl. Pak platí

$$\mathsf{var}\,[\boldsymbol{U} - \tau(\boldsymbol{Z})] \geq \mathsf{var}\,[\boldsymbol{U} - \mathsf{E}\,(\boldsymbol{U} \mid \boldsymbol{Z}\,)].$$

Poznámka. Pracujeme-li s rozptylovými maticemi ($m > 1$), rozumíme výševuvedené nerovnosti tak, že rozdíl levé a pravé strany je positivně semidefinitní matice.

Poznámka. Věta 4.3 říká, že chceme-li aproximovat náhodný vektor $\boldsymbol{U}$ pomocí funkce náhodného vektoru $\boldsymbol{Z}$, poskytuje podmíněná střední hodnota $\mathsf{E}\,(\boldsymbol{U} \mid \boldsymbol{Z}\,)$ nejlepší aproximaci (co do rozptylu) mezi všemi možnými funkcemi $\boldsymbol{Z}$.

Podmíněná střední hodnota se dá (obrazně leč poněkud nepřesně) vysvětlit tímto způsobem: Podmíněná střední hodnota odstraňuje z $\boldsymbol{U}$ náhodnost související s náhodným vektorem $\boldsymbol{Y}$, ale ponechává náhodnost způsobenou náhodným vektorem $\boldsymbol{Z}$.

Poznámka. V teorii pravděpodobnosti se zavádí obecná abstraktní definice podmíněné střední hodnoty, která nespoléhá na existenci podmíněné hustoty. O podmíněné střední hodnotě pak lze mluvit i tam, kde neexistuje podmíněná hustota. Příklad: $\mathsf{E}\,(Z \mid Z\,)$ nelze podle definice 4.1 a 4.2 spočítat.

4.3 Podmíněný rozptyl

Nechť $\mathsf{E}\,\boldsymbol{U}^{\mathsf{T}}\boldsymbol{U} < \infty$, čili všech m složek náhodného vektoru $\boldsymbol{U} \equiv h(\boldsymbol{Y}, \boldsymbol{Z})$ má konečné rozptyly.

Definice 4.3. *Podmíněný rozptyl* $\mathsf{var}\,(\boldsymbol{U} \mid \boldsymbol{Z}\,)$ náhodného vektoru $\boldsymbol{U}$, je-li dáno $\boldsymbol{Z}$, jest definován výrazem

$$\mathsf{var}\,(\boldsymbol{U} \mid \boldsymbol{Z}\,) = \mathsf{E}\left(\left[\boldsymbol{U} - \mathsf{E}\,(\boldsymbol{U} \mid \boldsymbol{Z}\,)\right]^{\otimes 2} \,\middle|\, \boldsymbol{Z} \right).$$

Poznámka. Podmíněný rozptyl z definice 4.3 je náhodná matice (náhodná veličina, pokud $m = 1$). Podobně lze definovat podmíněný rozptyl $\mathsf{var}\,(\boldsymbol{U} \mid \boldsymbol{Z} = \boldsymbol{z}\,)$ pro konkrétní realizovanou hodnotu $\boldsymbol{z}$ vektoru $\boldsymbol{Z}$.

Věta 4.4 (Rozklad nepodmíněného rozptylu)**.**

$$\mathsf{var}\,\boldsymbol{U} = \mathsf{E}\,\mathsf{var}\,(\boldsymbol{U} \mid \boldsymbol{Z}\,) + \mathsf{var}\,\mathsf{E}\,(\boldsymbol{U} \mid \boldsymbol{Z}\,)\,.$$

5 Transformace náhodných veličin a vektorů

Na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ uvažujme daný náhodný vektor $\boldsymbol{X} = (X_1, \ldots, X_n)^\mathsf{T}$, jehož rozdělení známe, a měřitelnou funkci $g : \mathbb{R}^n \to \mathbb{R}^n$. Naším úkolem je zjistit rozdělení náhodného vektoru $\boldsymbol{Y} = g(\boldsymbol{X})$.

Definice 5.1 (Nosič rozdělení)**.** Nechť X je (obecná) náhodná veličina, která nabývá hodnot z výběrového prostoru $\mathcal{X}$. Nechť rozdělení X je absolutně spojité vzhledem k nějaké σ-konečné míře μ. Množinu $S_X \subseteq \mathcal{X}$ nazveme *nosičem rozdělení* náhodné veličiny X právě když platí:

1. $\mathrm{P}[X \in S_X] = 1$;
2. $\forall A \subset S_X : \mu(S_X \setminus A) > 0 \Rightarrow \mathrm{P}[X \in A] < 1$.

5.1 Transformace náhodných veličin

Nejprve uvažujme případ $\mathcal{X} = \mathbb{R}$, tj. transformujeme reálnou náhodnou veličinu.

Tvrzení 5.1 (Věta o monotonní transformaci)**.** Nechť X má distribuční funkci F_X a nosič S_X. Nechť funkce g zobrazuje S_X na $S_0 \subseteq \mathbb{R}$. Označme $Y = g(X)$.

1. Je-li g ryze rostoucí, pak distribuční funkce náhodné veličiny Y je $F_Y(y) = F_X(g^{-1}(y))$ pro $y \in S_0$.
2. Je-li g ryze klesající, pak distribuční funkce náhodné veličiny Y je $F_Y(y) = 1 - F_X(g^{-1}(y)-)$ pro $y \in S_0$.

Značení. Je-li g reálná funkce s limitami zleva ve všech bodech, pak výraz $g(x-)$ značí zleva spojitou verzi funkce g, tj. $g(x-) \overset{\mathsf{df}}{=} \lim\limits_{h \searrow 0} g(x - h)$.

Důsledky.

1. Nechť X je spojitá reálná veličina s hustotou $f_X(x)$ a nechť g je ryze monotonní a diferencovatelná skoro všude. Hustota náhodné veličiny $Y = g(X)$ je pak rovna

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right| & \text{pro } y \in g(S_X); \\ 0 & \text{pro } y \notin g(S_X). \end{cases}$$

2. Nechť X je diskrétní reálná veličina s rozdělením $\mathrm{P}[X = x] = q_x$, $x \in S_X$. Pak $\mathrm{P}[Y = y] = q_{g^{-1}(y)}$, $y \in g(S_X)$.

Nyní prozkoumáme nemonotonní transformace. Budeme předpokládat, že existují intervaly $G_k \subseteq \mathbb{R}$, $k = 1, 2, \ldots$, takové, že $\bigcup_{k=1}^{\infty} G_k \supseteq S_X$, $G_i \cap G_j = \emptyset$ pro $i \neq j$, a g je ostře monotonní na každém G_k.

Značení.

- Označme $\mathcal{K}^+$ množinu všech indexů k takových, že g roste na G_k a $\mathcal{K}^-$ množinu všech indexů k takových, že g klesá na G_k.
- Označme g_k funkci g restriktovanou na G_k, třeba $g_k(x) = g(x)\mathbb{I}_{G_k}(x)$. Pak existuje $g_k^{-1}(y)$ pro $y \in g_k(G_k)$.
- Označme $X_k = X\,\mathbb{I}_{G_k}(X)$, $Y_k = g_k(X_k)$. Máme $X = \sum_{k=1}^{\infty} X_k$ a $Y = \sum_{k=1}^{\infty} Y_k$.

Tvrzení 5.2. Za daných předpokladů platí

$$F_Y(y) = \sum_{k \in \mathcal{K}^+} \mathrm{P}\left[X_k \leq g_k^{-1}(y), X \in G_k\right] + \sum_{k \in \mathcal{K}^-} \mathrm{P}\left[X_k \geq g_k^{-1}(y), X \in G_k\right].$$

Tvrzení 5.3. Nechť má navíc X hustotu vzhledem k Lebesgueově míře a nechť je každá g_k diferencovatelná (skoro všude) v G_k. Pak Y má hustotu

$$f_Y(y) = \sum_{k=1}^{\infty} f_X(g_k^{-1}(y)) \left| \frac{dg_k^{-1}(y)}{dy} \right| \mathbb{I}_{g_k(G_k)}(y).$$

Poznámka. Chceme-li pouze spočítat střední hodnotu $\mathsf{E}\,Y \equiv \mathsf{E}\,g(X)$, je obvykle snazší použít přímý vzorec $\mathsf{E}\,g(X) = \int g(x) f_X(x)\, d\mu(x)$ než počítat nejprve hustotu Y a pak integrovat $\mathsf{E}\,g(X) = \int y f_Y(y)\, d\mu(y)$.

5.2 Transformace náhodných vektorů

Uvažujme náhodný vektor $\boldsymbol{X} = (X_1, \ldots, X_n)^\mathsf{T}$ s nosičem rozdělení $S_{\boldsymbol{X}} \subseteq \mathbb{R}^n$ a spojitým rozdělením (má hustotu vzhledem k Lebesgueově míře). Nechť je dána transformace $g : \mathbb{R}^n \to \mathbb{R}^n$, vlastně vektor n funkcí $g_1, \ldots, g_n$, každá z nichž zobrazuje $\mathbb{R}^n$ do $\mathbb{R}$.

Zajímá nás rozdělení náhodného vektoru $\boldsymbol{Y} = g(\boldsymbol{X})$. Budeme předpokládat, že transformace g je diferencovatelná skoro všude v $S_{\boldsymbol{X}}$, tj. existuje matice

$$\frac{\partial g(\boldsymbol{x})}{\partial \boldsymbol{x}} = \begin{pmatrix} \frac{\partial g_1(\boldsymbol{x})}{\partial x_1} & \cdots & \frac{\partial g_1(\boldsymbol{x})}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial g_n(\boldsymbol{x})}{\partial x_1} & \cdots & \frac{\partial g_n(\boldsymbol{x})}{\partial x_n} \end{pmatrix}.$$

Determinant této matice (*jakobián* transformace g) budeme značit $\det \dfrac{\partial g(\boldsymbol{x})}{\partial \boldsymbol{x}}$.

Tvrzení 5.4. Nechť $\boldsymbol{X}$ má hustotu $f_{\boldsymbol{X}}(\boldsymbol{x})$ vzhledem k Lebesgueově míře. Nechť g je prosté zobrazení a $\det \dfrac{\partial g(\boldsymbol{x})}{\partial \boldsymbol{x}} \neq 0$ pro skoro všechna $\boldsymbol{x} \in S_{\boldsymbol{X}}$. Pak $\boldsymbol{Y} = g(\boldsymbol{X})$ má hustotu

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = f_{\boldsymbol{X}}(g^{-1}(\boldsymbol{y})) \cdot \left| \det \frac{\partial g^{-1}(\boldsymbol{y})}{\partial \boldsymbol{y}} \right| \mathbb{I}_{g(S_{\boldsymbol{X}})}(\boldsymbol{y})$$

vzhledem k Lebesgueově míře.

Poznámka. Platí

$$\frac{\partial g^{-1}(\boldsymbol{y})}{\partial \boldsymbol{y}} = \left(\left. \frac{\partial g(\boldsymbol{x})}{\partial \boldsymbol{x}} \right|_{\boldsymbol{x}=g^{-1}(\boldsymbol{y})} \right)^{-1}$$

a

$$\det \frac{\partial g^{-1}(\boldsymbol{y})}{\partial \boldsymbol{y}} = \frac{1}{\det \left. \frac{\partial g(\boldsymbol{x})}{\partial \boldsymbol{x}} \right|_{\boldsymbol{x}=g^{-1}(\boldsymbol{y})}}.$$

Tvrzení 5.5. Nechť $\boldsymbol{X}$ má hustotu $f_{\boldsymbol{X}}(\boldsymbol{x})$ vzhledem k Lebesgueově míře. Nechť existují množiny $G_k \subseteq \mathbb{R}^n$, $k = 1, 2, \ldots$, takové, že $\bigcup_{k=1}^{\infty} G_k \supseteq S_{\boldsymbol{X}}$, $G_i \cap G_j = \emptyset$ pro $i \neq j$, $g_k(\boldsymbol{x}) \overset{\mathsf{df}}{=} g(\boldsymbol{x}) \mathbb{I}_{G_k}(\boldsymbol{x})$ je prostá na každém G_k, a $\det \dfrac{\partial g_k(\boldsymbol{x})}{\partial \boldsymbol{x}} \neq 0$ pro skoro všechna $\boldsymbol{x} \in G_k$. Pak $\boldsymbol{Y} = g(\boldsymbol{X})$ má hustotu

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \sum_{k=1}^{\infty} f_{\boldsymbol{X}}(g_k^{-1}(\boldsymbol{y})) \cdot \left| \det \frac{\partial g_k^{-1}(\boldsymbol{y})}{\partial \boldsymbol{y}} \right| \mathbb{I}_{g_k(G_k)}(\boldsymbol{y})$$

vzhledem k Lebesgueově míře.

Poznámka. Nechť $\boldsymbol{X} = (X_1, , \ldots, X_n)^{\mathsf{T}}$ je náhodný vektor a t nějaká hladká měřitelná funkce $\mathbb{R}^n \to \mathbb{R}$. Jaké je rozdělení náhodné veličiny $T = t(\boldsymbol{X})$?

Zvolme vhodně transformaci $g : \mathbb{R}^n \to \mathbb{R}^n$ tak, aby $g_1(\boldsymbol{x}) = t(\boldsymbol{x})$. Platí-li předpoklady tvrzení 5.5, můžeme podle něj spočítat sdruženou hustotu náhodného vektoru $\boldsymbol{Y} = g(\boldsymbol{X})$. Marginální hustotu náhodné veličiny $T \equiv Y_1$ zjistíme vyintegrováním ostatních složek podle (3.1).

Věta 5.6 (o konvoluci)**.** Nechť X a Y jsou *nezávislé* náhodné veličiny, nechť X má hustotu f_X vzhledem k míře μ_1 a Y má hustotu f_Y vzhledem k míře μ_2. Pak $Z = X + Y$ má distribuční funkci

$$F_Z(z) = \int_{-\infty}^{\infty} f_Y(y) F_X(z - y)\, d\mu_2(y) = \int_{-\infty}^{\infty} f_X(x) F_Y(z - x)\, d\mu_1(x).$$

Jsou-li X a Y spojité, pak Z má hustotu

$$f_Z(z) = \int_{-\infty}^{\infty} f_Y(y) f_X(z-y)\, dy = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\, dx$$

vzhledem k Lebesgueově míře. Jsou-li X a Y diskrétní, pak

$$\mathrm{P}\left[Z = z\right] = \sum_{y \in S_Y} \mathrm{P}\left[Y = y\right] \mathrm{P}\left[X = z - y\right] = \sum_{x \in S_X} \mathrm{P}\left[X = x\right] \mathrm{P}\left[Y = z - x\right].$$

Tvrzení 5.7. Nechť X a Y jsou *nezávislé* náhodné veličiny, nechť X má hustotu f_X vzhledem k míře μ_1 a Y má hustotu f_Y vzhledem k míře μ_2. Pak $Z = X/Y$ má distribuční funkci

$$F_Z(z) = \int_0^{\infty} f_Y(y) F_X(zy)\, d\mu_2(y) + \int_{-\infty}^{0} f_Y(y)[1 - F_X(zy)]\, d\mu_2(y).$$

Jsou-li X a Y spojité, pak Z má hustotu

$$f_Z(z) = \int_{-\infty}^{\infty} |y|\, f_Y(y) f_X(zy)\, dy.$$

6 Normální rozdělení

Poznámka (Normální rozdělení)**.**

- Náhodná veličina Z s hustotou $\varphi(z) = \dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-z^2/2}$ má *normované normální rozdělení*; značíme $Z \sim \mathsf{N}(0,1)$. Její distribuční funkci značíme
$$\Phi(z) \overset{\mathsf{df}}{=} \int_{-\infty}^{z} \varphi(t)\,dt.$$
- Jestliže $Z \sim \mathsf{N}(0,1)$ a $X = \sigma Z + \mu$, kde $\sigma > 0$ a $\mu \in \mathbb{R}$, pak X má *normální rozdělení* s parametry μ a σ^2, značíme $X \sim \mathsf{N}(\mu, \sigma^2)$. Její hustota je
$$f_X(x) = \frac{1}{\sigma}\varphi\Big(\frac{x-\mu}{\sigma}\Big) = \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$
Její distribuční funkce je $F_X(x) = \Phi\big(\frac{x-\mu}{\sigma}\big)$.
- Jestliže $X \sim \mathsf{N}(\mu, \sigma^2)$ pak $\mathsf{E}\, X = \mu$, $\mathsf{var}\, X = \sigma^2$, $\gamma_3 = 0$, $\gamma_4 = 3$.

6.1 Mnohorozměrné normální rozdělení

Definice 6.1. Nechť $\boldsymbol{Z} = (Z_1, \dots, Z_r)^\mathsf{T}$, kde $Z_i \sim \mathsf{N}(0,1)$ jsou nezávislé. Nechť $A_{n\times r}$ je matice a $\boldsymbol{\mu} \in \mathbb{R}^n$ je pevný vektor. Náhodný vektor $\boldsymbol{X}$ definovaný jako $\boldsymbol{X} = A\boldsymbol{Z} + \boldsymbol{\mu}$ pak má *n-rozměrné normální rozdělení* s parametry $\boldsymbol{\mu}$ a $\Sigma \overset{\mathsf{df}}{=} AA^\mathsf{T}$. Značíme $\boldsymbol{X} \sim \mathsf{N}_n(\boldsymbol{\mu}, \Sigma)$.

Poznámka.

- $\mathsf{E}\, \boldsymbol{X} = \boldsymbol{\mu}$, $\mathsf{var}\, \boldsymbol{X} = \Sigma$.
- $\boldsymbol{X}$ má n-rozměrné normální rozdělení $\Leftrightarrow$ pro libovolné $\boldsymbol{c} \in \mathbb{R}^n$ platí $\boldsymbol{c}^\mathsf{T}\boldsymbol{X} \sim \mathsf{N}(\,\cdot\,,\,\cdot\,)$.
- Libovolná symetrická positivně semidefinitní matice Σ se dá napsat jako AA^T pro nějaké $A_{n\times r}$, $r \leq n$. Platí: $r < n$ právě když Σ je singulární.

Věta 6.1. Nechť $\boldsymbol{X} \sim \mathsf{N}_n(\boldsymbol{\mu}, \Sigma)$ a Σ je regulární. Pak existuje hustota $\boldsymbol{X}$ vzhledem k Lebesgueově míře na $\mathbb{R}^n$ a její tvar je
$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{n/2}\sqrt{\det \Sigma}}\mathrm{e}^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\mathsf{T}\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}$$
pro $\boldsymbol{x} \in \mathbb{R}^n$.

Poznámka.
- Je-li Σ singulární, pak existuje nenulové $\boldsymbol{c} \in \mathbb{R}^n$ takové, že $\boldsymbol{c}^\mathsf{T} \boldsymbol{X} = 0$ (tj. složky $\boldsymbol{X}$ jsou lineárně závislé).
- Je-li Σ singulární, pak hustota $\boldsymbol{X}$ vzhledem k Lebesgueově míře na $\mathbb{R}^n$ neexistuje.

Příklad (Dvourozměrné normální rozdělení)**.** Nechť $n = 2$, Σ je regulární, $\sigma_1 = \mathsf{var}\, X_1$, $\sigma_2 = \mathsf{var}\, X_2$ a $\varrho = \mathsf{cor}\,(X_1, X_2)$. Hustotu náhodného vektoru $\boldsymbol{X} = (X_1, X_2)^\mathsf{T}$ pak lze vyjádřit ve tvaru

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\varrho^2}} \, \mathrm{e}^{-\frac{1}{2(1-\varrho^2)}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\varrho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]}.$$

Věta 6.2 (Vlastnosti mnohorozměrného normálního rozdělení)**.**
Nechť $\boldsymbol{X} \sim \mathsf{N}_n(\boldsymbol{\mu}, \Sigma)$, kde $\boldsymbol{X} = (\boldsymbol{X}_1^\mathsf{T}, \boldsymbol{X}_2^\mathsf{T})^\mathsf{T}$, $\boldsymbol{\mu} = (\boldsymbol{\mu}_1^\mathsf{T}, \boldsymbol{\mu}_2^\mathsf{T})^\mathsf{T}$, $\Sigma = \left(\begin{smallmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{smallmatrix}\right)$ a dimenze jednotlivých složek jsou $k \times 1$ pro $\boldsymbol{X}_1$ a $\boldsymbol{\mu}_1$ a $k \times k$ pro Σ_{11}. Pak platí:

1. $\boldsymbol{X}_1 \sim \mathsf{N}_k(\boldsymbol{\mu}_1, \Sigma_{11})$.
2. Jestliže $\Sigma_{12} = 0$, pak $\boldsymbol{X}_1$ a $\boldsymbol{X}_2$ jsou nezávislé.
3. Je-li Σ_{22} regulární, pak podmíněné rozdělení $\boldsymbol{X}_1$, je-li dáno $\boldsymbol{X}_2 = \boldsymbol{x}_2$, je k-rozměrné normální se střední hodnotou
$$\boldsymbol{\mu}_{1.2} = \boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2)$$
a rozptylem
$$\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}.$$

Poznámka. Z předchozí věty plyne:
- Mají-li X_1 a X_2 sdružené normální rozdělení, pak mají marginální normální rozdělení.
- Mají-li X_1 a X_2 sdružené normální rozdělení a jsou-li nekorelované, pak jsou nezávislé.

Poznámka. Mají-li X_1 a X_2 marginální normální rozdělení, pak $\boldsymbol{X} = (X_1, X_2)^\mathsf{T}$ nemusí mít sdružené normální rozdělení (protipříklad).

6.2 Rozdělení χ^2, t a F

Poznámka. Náhodná veličina X má χ^2 rozdělení o r stupních volnosti, značíme $X \sim \chi_r^2$, právě když její hustota vzhledem k Lebesgueově míře je

$$f_X(x) = \frac{1}{2^{r/2}\Gamma\left(\frac{r}{2}\right)} x^{r/2-1} \mathrm{e}^{-x/2} \, \mathbb{I}_{(0,\infty)}(x).$$

Rozdělení χ_r^2 je speciální případ gama rozdělení: $\Gamma(\frac{1}{2}, \frac{r}{2})$.

Věta 6.3 (o χ^2-rozdělení)**.**

1. Nechť $X_1, \dots, X_n$ jsou nezávislé náhodné veličiny s rozdělením $\mathsf{N}(0,1)$. Pak $Y = \sum_{i=1}^n X_i^2 \sim \chi^2_n$.
2. Nechť $\boldsymbol{X} \sim \mathsf{N}_n(\boldsymbol{\mu}, \Sigma)$, kde Σ je regulární. Pak
$$Y = (\boldsymbol{X} - \boldsymbol{\mu})^{\mathsf{T}} \Sigma^{-1} (\boldsymbol{X} - \boldsymbol{\mu}) \sim \chi^2_n.$$
3. Nechť $\boldsymbol{X} \sim \mathsf{N}_n(\boldsymbol{0}, \Sigma)$ a nechť A je taková matice typu $n \times n$, že $A\Sigma$ je idempotentní. Pak
$$Y = \boldsymbol{X}^{\mathsf{T}} A \boldsymbol{X} \sim \chi^2_{\operatorname{tr} AV}.$$

Poznámka (něco o maticích)**.**
- Čtvercovou matici D nazveme *idempotentní* právě když $DD = D$.
- $\operatorname{tr} D$ značí *stopu matice* D, tj. součet jejích diagonálních prvků.
- Je-li matice D idempotentní, pak $\operatorname{tr} D = r(D)$ (hodnost je rovna stopě).

Věta 6.4 (o t-rozdělení)**.** Nechť $X \sim \mathsf{N}(0,1)$ a $Z \sim \chi^2_k$ jsou nezávislé. Pak náhodná veličina $T \overset{\mathsf{df}}{=} \frac{X}{\sqrt{Z/k}}$ má rozdělení s hustotou

$$f_{T,k}(t) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{\pi k}} \left(1 + \frac{t^2}{k}\right)^{-\frac{k+1}{2}}$$

vzhledem k Lebesgueově míře. Rozdělení náhodné veličiny T se nazývá [Studentovo] *t rozdělení* s k stupni volnosti, značíme $T \sim t_k$.

Poznámka (Vlastnosti t rozdělení)**.**
- Hustota t rozdělení je symetrická kolem 0.
- Pro $k = 1$ jest $f_{T,1}$ hustotou Cauchyova rozdělení $\mathsf{C}(0,1)$. Rozdělení t_1 nemá střední hodnotu.
- Obecně má T konečné momenty do řádu $k-1$, $\mathsf{E}\, T = 0$ pro $k > 1$, $\mathsf{var}\, T = \frac{k}{k-2}$ pro $k > 2$.
- Pro velké k se hustota t rozdělení blíží hustotě normovaného normálního rozdělení: $\lim_{k\to\infty} |f_{T,k}(t) - \varphi(t)| = 0$ pro každé $t \in \mathbb{R}$. Tudíž α-kvantil rozdělení t_k konverguje k α-kvantilu rozdělení $\mathsf{N}(0,1)$ pro $k \to \infty$.

Věta 6.5 (o F-rozdělení)**.** Nechť $X \sim \chi^2_m$ a $Y \sim \chi^2_n$ jsou nezávislé. Pak náhodná veličina
$$Z = \frac{X/m}{Y/n}$$

má hustotu

$$f_{F;m,n}(z) = \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\left(\frac{m}{n}\right)^{\frac{m}{2}} z^{\frac{m}{2}-1}\left(1+\frac{m}{n}z\right)^{-\frac{m+n}{2}} \mathbb{I}_{(0,\infty)}(z)$$

vzhledem k Lebesgueově míře. Rozdělení náhodné veličiny Z se nazývá [Fisherovo-Snedecorovo] F *rozdělení* s m a n stupni volnosti.

7 Limitní věty

Na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ máme danou posloupnost náhodných vektorů $\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3, \ldots$, kde $\boldsymbol{X}_i : (\Omega, \mathcal{A}) \to (\mathbb{R}^k, \mathcal{B}_0^k)$ a $\boldsymbol{X}_i = (X_{i1}, \ldots, X_{ik})^\mathsf{T}$.

7.1 Konvergence náhodných veličin a vektorů

Definice 7.1 (konvergence v pravděpodobnosti)**.** Říkáme, že posloupnost $\{\boldsymbol{X}_n\}_{n=1}^\infty$ *konverguje v pravděpodobnosti* k náhodnému vektoru $\boldsymbol{X}$ pro $n \to \infty$ právě když

$$\forall \varepsilon > 0 : \lim_{n\to\infty} \mathrm{P}\left[\|\boldsymbol{X}_n - \boldsymbol{X}\| > \varepsilon\right] = 0.$$

Konvergenci v pravděpodobnosti značíme $\boldsymbol{X}_n \xrightarrow{\mathsf{P}} \boldsymbol{X}$.

Poznámka. $\|\boldsymbol{a}\|$ značí eukleidovskou normu vektoru $\boldsymbol{a}$, tj. $\|\boldsymbol{a}\| = \sqrt{\boldsymbol{a}^\mathsf{T}\boldsymbol{a}}$.

Definice 7.2 (konvergence v distribuci)**.** Říkáme, že posloupnost $\{\boldsymbol{X}_n\}_{n=1}^\infty$ *konverguje v distribuci* k náhodnému vektoru $\boldsymbol{X}$ pro $n \to \infty$ právě když

$$\lim_{n\to\infty} F_{\boldsymbol{X}_n}(\boldsymbol{x}) = F_{\boldsymbol{X}}(\boldsymbol{x})$$

v každém bodě $\boldsymbol{x}$, v němž je $F_{\boldsymbol{X}}(\boldsymbol{x})$ spojitá. Konvergenci v distribuci značíme $\boldsymbol{X}_n \xrightarrow{\mathsf{D}} \boldsymbol{X}$ nebo $F_{\boldsymbol{X}_n} \to F_{\boldsymbol{X}}$ nebo $\mathcal{L}(\boldsymbol{X}_n) \to \mathcal{L}(\boldsymbol{X})$.

Poznámka. Symbolem $\mathcal{L}(\boldsymbol{X}_n)$ se rozumí rozdělení náhodného vektoru $\boldsymbol{X}_n$ (z angl. *Law*). Výraz $\mathcal{L}(\boldsymbol{X}_n) \to \mathcal{L}(\boldsymbol{X})$ čteme *„rozdělení $\boldsymbol{X}_n$ konverguje k rozdělení $\boldsymbol{X}$“*. Můžeme také říkat, že $\boldsymbol{X}_n$ má asymptotické (limitní) rozdělení $F_{\boldsymbol{X}}$ a psát $\boldsymbol{X}_n \overset{\mathsf{as}}{\sim} \mathcal{L}(\boldsymbol{X})$.

Tvrzení 7.1.

$$\boldsymbol{X}_n \xrightarrow{\mathsf{P}} \boldsymbol{X} \;\Rightarrow\; \boldsymbol{X}_n \xrightarrow{\mathsf{D}} \boldsymbol{X}$$

Poznámka. Opačná implikace neplatí. Nicméně pokud limitní vektor je konstanta, tj. $\boldsymbol{X}_n \xrightarrow{\mathsf{D}} \boldsymbol{c}$, pak $\boldsymbol{X}_n \xrightarrow{\mathsf{P}} \boldsymbol{c}$ a obě konvergence jsou ekvivalentní.

Tvrzení 7.2 (vlastnosti konvergence v distribuci)**.**

1. *(Cramér-Woldova věta)* $\boldsymbol{X}_n \xrightarrow{\mathsf{D}} \boldsymbol{X} \;\Leftrightarrow\; \forall \boldsymbol{c} \in \mathbb{R}^k : \boldsymbol{c}^\mathsf{T}\boldsymbol{X}_n \xrightarrow{\mathsf{D}} \boldsymbol{c}^\mathsf{T}\boldsymbol{X}$.

2. *(Helly-Brayova věta)* $X_n \xrightarrow{\mathsf{D}} X \;\Leftrightarrow\; \mathsf{E}\, g(X_n) \to \mathsf{E}\, g(X)$ pro každou spojitou omezenou funkci $g : \mathbb{R} \to \mathbb{R}$.

3. *(Fatouovo lemma)* $X_n \xrightarrow{\mathsf{D}} X \;\Rightarrow\; \mathsf{E}\, X \leq \liminf\limits_{n\to\infty} \mathsf{E}\, X_n$.

Tvrzení 7.3 (Věta o spojité transformaci)**.** Nechť $g : \mathbb{R}^k \to \mathbb{R}^m$ je spojitá funkce.

1. $\boldsymbol{X}_n \xrightarrow{\mathsf{P}} \boldsymbol{X} \;\Rightarrow\; g(\boldsymbol{X}_n) \xrightarrow{\mathsf{P}} g(\boldsymbol{X})$.

2. $\boldsymbol{X}_n \xrightarrow{\mathsf{D}} \boldsymbol{X} \;\Rightarrow\; g(\boldsymbol{X}_n) \xrightarrow{\mathsf{D}} g(\boldsymbol{X})$.

Tvrzení 7.4. Nechť pro posloupnost $\{\boldsymbol{X}_n\}_{n=1}^{\infty}$ platí $X_{nj} \xrightarrow{\mathsf{P}} X_j$ pro $n \to \infty$ a $j = 1, \ldots, k$. Pak $\boldsymbol{X}_n \xrightarrow{\mathsf{P}} \boldsymbol{X} = (X_1, \ldots, X_k)^{\mathsf{T}}$.

Poznámka. Pro konvergenci v distribuci tato vlastnost neplatí.

Tvrzení 7.5. Nechť $X_1, X_2, \ldots$ je posloupnost náhodných veličin takových, že $\mathsf{E}\, X_n \to \mu$ a $\mathsf{var}\, X_n \to 0$. Pak $X_n \xrightarrow{\mathsf{P}} \mu$.

Tvrzení 7.6 (Cramérova-Sluckého věta)**.** Nechť $X_n \xrightarrow{\mathsf{D}} X$, $A_n \xrightarrow{\mathsf{P}} a$ a $B_n \xrightarrow{\mathsf{P}} b$, kde X_n, X, A_n, B_n jsou náhodné veličiny a a, b jsou konstanty. Pak platí

$$A_n X_n + B_n \xrightarrow{\mathsf{D}} aX + b.$$

Poznámka. Cramérově-Sluckého větě se často říká Sluckého věta. Tato věta platí i pro vektory, tj. pokud $\boldsymbol{X}_n \xrightarrow{\mathsf{D}} \boldsymbol{X}$, $A_n \xrightarrow{\mathsf{P}} A$ a $\boldsymbol{B}_n \xrightarrow{\mathsf{P}} \boldsymbol{b}$, kde $\boldsymbol{X}_n$ a $\boldsymbol{X}$ jsou k-rozměrné náhodné vektory, A_n je náhodná matice o dimenzích $m \times k$, A je matice konstant o dimenzích $m \times k$, $\boldsymbol{B}_n$ jsou m-rozměrné náhodné vektory a $\boldsymbol{b}$ je m-rozměrný vektor konstant, pak

$$A_n \boldsymbol{X}_n + \boldsymbol{B}_n \xrightarrow{\mathsf{D}} A\boldsymbol{X} + \boldsymbol{b}.$$

Tvrzení 7.7. Nechť $a_n(\boldsymbol{X}_n - \boldsymbol{\mu}) \xrightarrow{\mathsf{D}} \boldsymbol{X}$, kde $a_n > 0$ je posloupnost reálných čísel splňující $a_n \to \infty$ a $\boldsymbol{\mu}$ je vektor konstant. Pak $\boldsymbol{X}_n \xrightarrow{\mathsf{P}} \boldsymbol{\mu}$.

7.2 Zákon velkých čísel

Uvažujme náhodnou posloupnost $\{\boldsymbol{X}_n\}_{n=1}^{\infty}$. Označme $\overline{\boldsymbol{X}}_n \overset{\mathsf{df}}{=} \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{X}_i$ (průměr z prvních n vektorů).

Věta 7.8 (Čebyševův slabý zákon velkých čísel)**.** Nechť $X_1, X_2, \ldots$ je posloupnost nezávislých náhodných veličin se střední hodnotou $\mathsf{E}\, X_i = \mu$ a rozptylem $\mathsf{var}\, X_i \leq C$ pro nějaké $C \in \mathbb{R}$. Pak platí

$$\overline{X}_n \xrightarrow{\mathsf{P}} \mu.$$

Tvrzení 7.9 (Chinčinův slabý zákon velkých čísel)**.** Nechť $X_1, X_2, \ldots$ je posloupnost nezávislých stejně rozdělených náhodných veličin se střední hodnotou $\mathsf{E}\, X_i = \mu < \infty$. Pak platí

$$\overline{X}_n \xrightarrow{\mathsf{P}} \mu.$$

Poznámka.

- Oba zákony velkých čísel platí i pro náhodné vektory, pokud všechny jejich složky splňují stanovené předpoklady (viz tvrzení 7.4).
- Čebyševův zákon velkých čísel nevyžaduje, aby byly všechny veličiny stejně rozdělené, ale vyžaduje, aby měly omezený (tj. nutně konečný) rozptyl. Chinčinův zákon velkých čísel vyžaduje, aby byly všechny veličiny stejně rozdělené, ale nevyžaduje, aby měly konečný rozptyl.
- Zákony velkých čísel lze zobecnit i na závislé veličiny, pokud nejsou závislé „příliš“. Např. u Čebyševova zákona velkých čísel stačí nahradit nezávislost podmínkou $n^{-2} \sum \sum \mathsf{cov}\,(X_i, X_j) \to 0$.
- Existují i „silné“ zákony velkých čísel, které udávají podmínky pro konvergenci $\overline{X}_n$ k μ *skoro jistě* (silnější typ konvergence než konvergence v pravděpodobnosti).

Příklady.

1. Jestliže $X_i \sim \mathsf{C}(0,1)$, pak $\overline{X}_n \sim \mathsf{C}(0,1)$ pro libovolné n. Průměr nekonverguje ke konstantě.
2. Empirická četnost vs. pravděpodobnost jevu.

7.3 Centrální limitní věta

Nadále uvažujme náhodnou posloupnost k-rozměrných vektorů $\{\boldsymbol{X}_n\}_{n=1}^{\infty}$.

Tvrzení 7.10 (centrální limitní věta pro nezávislé stejně rozdělené náhodné vektory)**.** Nechť $\{\boldsymbol{X}_n\}_{n=1}^{\infty}$ jsou nezávislé a stejně rozdělené náhodné vektory se střední hodnotou $\boldsymbol{\mu} \equiv \mathsf{E}\, \boldsymbol{X}_i$ a konečnou rozptylovou maticí $\Sigma \equiv \mathsf{var}\, \boldsymbol{X}_i$. Pak platí

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\boldsymbol{X}_i - \boldsymbol{\mu}) = \sqrt{n}(\overline{\boldsymbol{X}}_n - \boldsymbol{\mu}) \xrightarrow{\mathsf{D}} \mathsf{N}_k(\boldsymbol{0}, \Sigma).$$

Poznámka. Neformální zápis tvrzení centrální limitní věty: $\overline{\boldsymbol{X}}_n \overset{\mathsf{as}}{\sim} \mathsf{N}_k(\boldsymbol{\mu}, n^{-1}\Sigma)$.

Příklady.
1. Aproximace binomického rozdělení normálním
2. Aproximace χ^2 rozdělení normálním

Věta 7.11 (Δ-metoda)**.** Nechť $\{\boldsymbol{T}_n\}_{n=1}^{\infty}$ splňuje

$$\sqrt{n}(\boldsymbol{T}_n - \boldsymbol{\mu}) \xrightarrow{\mathsf{D}} \mathsf{N}_k(\boldsymbol{0}, \Sigma)$$

pro nějaký vektor konstant $\boldsymbol{\mu} \in \mathbb{R}^k$ a matici Σ. Nechť g je spojitě diferencovatelná funkce $\mathbb{R}^k \to \mathbb{R}^p$. Označme $D(\boldsymbol{x}) = \frac{\partial g(\boldsymbol{x})}{\partial \boldsymbol{x}}$. Pak platí

$$\sqrt{n}(g(\boldsymbol{T}_n) - g(\boldsymbol{\mu})) \xrightarrow{\mathsf{D}} \mathsf{N}_p(\boldsymbol{0}, D(\boldsymbol{\mu})\Sigma D(\boldsymbol{\mu})^{\mathsf{T}})$$

Příklad. Asymptotické rozdělení $\log \overline{X}_n$.

Obsah

1 Úvod **3**
1.1 Kolmogorovova definice pravděpodobnosti . . . 3
1.2 Náhodná veličina . . . 3
1.3 Rozdělení náhodné veličiny, hustota . . . 4

2 Reálná náhodná veličina a její rozdělení **5**
2.1 Charakterizace rozdělení reálné náhodné veličiny . . . 5
2.2 Momenty reálné náhodné veličiny . . . 7

3 Náhodný vektor a mnohorozměrné rozdělení **9**
3.1 Rozdělení náhodného vektoru . . . 9
3.2 Momenty . . . 11
3.3 Nezávislost . . . 13
3.4 Korelace . . . 14

4 Podmíněné rozdělení **16**
4.1 Podmíněná hustota . . . 16
4.2 Podmíněná střední hodnota . . . 17
4.3 Podmíněný rozptyl . . . 19

5 Transformace náhodných veličin a vektorů **20**
5.1 Transformace náhodných veličin . . . 20
5.2 Transformace náhodných vektorů . . . 21

6 Normální rozdělení **24**
6.1 Mnohorozměrné normální rozdělení . . . 24
6.2 Rozdělení χ^2, t a F . . . 25

7 Limitní věty **28**
7.1 Konvergence náhodných veličin a vektorů . . . 28
7.2 Zákon velkých čísel . . . 29
7.3 Centrální limitní věta . . . 30

Rejstřík

Δ-metoda, 31

χ^2 rozdělení, 25
asymptotické rozdělení, 28

centrální limitní věta, 30

distribuční funkce, 5, 10
    marginální, 11
    sdružená, 11

F rozdělení, 27

hustota, 4, 5
    marginální, 11
    podmíněná, 16
    sdružená, 11

konvergence
    v distribuci, 28
    v pravděpodobnosti, 28
korelační koeficient, 14
kvantil, 6
kvantilová funkce, 6
kvartil, 6

limitní rozdělení, 28

medián, 6
moment
    k-tý, 7
    k-tý absolutní, 7
    k-tý centrální, 7

náhodná veličina, 3
    diskrétní, 5
    distribuční funkce, 5
    hustota, 4, 5
    kvantilová funkce, 6
    moment, 7
    nekorelovanost, 14
    nezávislost, 13
    nosič, 20
    rozdělení, 4, 5
    rozptyl, 7
    směrodatná odchylka, 7
    spojitá, 5
    střední hodnota, 7
    šikmost, 7
    špičatost, 7
    transformace, 20
náhodný vektor, 9
    diskrétní, 10
    distribuční funkce, 10
    hustota, 9
    korelační matice, 15
    kovarianční matice, 12
    marginální rozdělení, 11
    nekorelovanost, 14
    nezávislost, 13
    rozdělení, 9
    rozptylová matice, 12
    sdružené rozdělení, 11
    spojitý, 10
    střední hodnota, 11
    transformace, 22
nekorelovanost, 14
nerovnost

Cauchyova-Schwartzova, 14
Čebyševova, 8
Jensenova, 8
Markovova, 8
nezávislé náhodné veličiny, 13
normální rozdělení, 24
mnohorozměrné, 24
normované, 24
nosič rozdělení, 20

podmíněná hustota, 16
podmíněná pravděpodobnost, 16
podmíněná střední hodnota, 18
podmíněný rozptyl, 19
pravděpodobnost, 3
podmíněná, 16
pravděpodobnostní prostor, 3

rozdělení
asymptotické, 28
χ^2, 25
F, 27
limitní, 28
marginální, 11
náhodné veličiny, 4, 5
náhodného vektoru, 9
normální, 24
mnohorozměrné, 24
normované, 24
nosič, 20
podmíněné, 16
sdružené, 11
t, 26
rozptyl, 7, 12
podmíněný, 19

směrodatná odchylka, 7
střední hodnota, 7, 11
podmíněná, 18
šikmost, 7
špičatost, 7

věta
Bayesova, 17
centrální limitní, 30
Cramérova-Sluckého, 29
Čebyševův zákon velkých čísel, 30
Δ-metoda, 31
Chinčinův zákon velkých čísel, 30
o χ^2 rozdělení, 26
o F rozdělení, 27
o konvoluci, 23
o rozdělení podílu, 23
o spojité transformaci, 29
o t rozdělení, 26
o transformaci, 20, 22
Radon-Nikodymova, 4
Sluckého, 29

zákon velkých čísel
Čebyševův, 30
Chinčinův, 30