Pravděpodobnost pro finanční matematiky

RNDr. Jitka Zichová, Dr.

Matematicko-fyzikální fakulta
Univerzita Karlova
2021

Obsah

Úvod

Tento text je studijní pomůckou k předmětu NMFM202 Pravděpodobnost pro finanční matematiky, který je součástí bakalářského studia finanční matematiky na MFF UK. Pokrývá kompletně látku přednášky. Literaturu, ze které přednáška čerpá, najdou posluchači u předmětu NMFM202 v SISu.

Za pečlivé přepsání rukopisných poznámek k přednášce vyjadřuji srdečné poděkování Janu Hanouskovi, studentovi 2. ročníku bakalářského programu Finanční matematika na MFF.

V Praze 30. 7. 2021 Jitka Zichová

1. Náhodné jevy a pravděpodobnost

1.1. Klasická definice pravděpodobnosti

Značení. Základní pojmy.

- *Náhodný pokus* je experiment, jehož výsledek není jednoznačně určen podmínkami, za kterých pokus provádíme. Může mít konečně, spočetně nebo nespočetně mnoho výsledků.
- *Náhodný jev* je výsledek náhodného pokusu. Značíme $A,\ B,\ C, ...$
- *Elementární jevy* ω jsou všechny dále nerozložitelné výsledky náhodného pokusu.
- *Prostor elementárních jevů* Ω je množina všech elementárních jevů $\omega \in \Omega$. Náhodné jevy jsou podmnožiny Ω, např. $A \subset \Omega$.
- Jev, který nastane tehdy, když nenastane jev A, nazýváme *doplňkem* jevu A a značíme $\bar{A}$.
- Náhodný jev nazveme *jevem jistým*, pokud je roven prostoru elementárních jevů, např. $B = A \cup \bar{A} = \Omega$.
- Náhodný jev nazveme *jevem nemožným*, pokud neobsahuje žádný elementární jev, např. $B = A \cap \bar{A} = \emptyset$.
- Náhodný jev A nazveme *podjevem* jevu B, pokud $A \subseteq B,\ A \Rightarrow B$.
- Dva jevy A a B nazveme *ekvivalentními jevy*, pokud $A \subseteq B$ a zároveň $B \subseteq A,\ A \Leftrightarrow B$.
- Dva jevy A a B nazveme *neslučitelnými jevy*, pokud nemohou nastat současně, tedy $A \cap B = \emptyset$.

Poznámka. Nechť $n \in \mathbb{N}$. Potom de Morganovými pravidly rozumíme vztahy

$$\overline{\bigcap_{i=1}^{n} A_i} = \bigcup_{i=1}^{n} \bar{A}_i, \qquad \overline{\bigcup_{i=1}^{n} A_i} = \bigcap_{i=1}^{n} \bar{A}_i.$$

Definice 1.1 (Klasická definice pravděpodobnosti)**.** Nechť prostor Ω obsahuje konečný počet stejně možných elementárních jevů. Pak definujeme *pravděpodobnost náhodného jevu A* předpisem

$$P[A] = \frac{|A|}{|\Omega|},$$

kde $|\Omega|$ je počet všech možných výsledků náhodného pokusu (počet elementárních jevů) a $|A|$ je počet výsledků, při kterých nastane jev A.

Dále definujeme *četnost jevu A* jako počet jeho nastání v n pokusech a značíme n_A. Procento pokusů, ve kterých nastane jev A, nazýváme *relativní četnost.* Je to podíl

$$\frac{n_A}{n}.$$

1.2. Kolmogorovova definice pravděpodobnosti

Úmluva. *V této sekci uvažujme prostor Ω jako libovolnou (konečnou, spočetnou nebo nespočetnou) množinu.*

Definice 1.2 (σ-algebra)**.** Systém $\mathcal{A}$ podmnožin Ω nazveme *σ-algebra*, pokud platí

(i) $\Omega \in \mathcal{A}$,

(ii) $A \in \mathcal{A} \implies \bar{A} \in \mathcal{A}$,

(iii) $A_1, A_2, ... \in \mathcal{A} \implies \bigcup_{i=1}^{\infty} \in \mathcal{A}$.

Věta 1.1 (Vlastnosti σ-algebry)**.** *Nechť $\mathcal{A}$ je σ-algebra podmnožin Ω. Pak platí*

(i) $\emptyset \in \mathcal{A}$,

(ii) $A_1, A_2, \dots \in \mathcal{A} \implies \bigcap_{i=1}^{\infty} A_i \in \mathcal{A}$,

(iii) $A, B \in \mathcal{A} \implies A \setminus B \in \mathcal{A}$.

Důkaz.

(i) Plyne z toho, že $\emptyset = \bar{\Omega}$, a z uzavřenosti $\mathcal{A}$ na doplňky.

(ii) Plyne z de Morganových pravidel a z uzavřenosti $\mathcal{A}$ na doplňky a spočetná sjednocení.

(iii) Plyne z toho, že $A \setminus B = A \cap \bar{B}$ a z uzavřenosti $\mathcal{A}$ na doplňky a průniky. □

Definice 1.3 (Kolmogorovova definice pravděpodobnosti)**.** Nechť Ω je libovolná množina a $\mathcal{A}$ je σ-algebra jejích podmnožin. Funkci $P : \mathcal{A} \to [0, 1]$ nazveme *pravděpodobnost*, pokud platí

(i) $P[A] \geq 0, \ A \in \mathcal{A}$,

(ii) $P[\Omega] = 1$,

(iii) $P\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} P[A_i]$ za přepodkladu, že $A_1, A_2, \ldots \in \mathcal{A}$ a že $A_i \cap A_j = \emptyset$ pro $i \neq j$. (σ-aditivita)

Věta 1.2 (Základní vlastnosti pravděpodobnosti)**.** *Nechť funkce P je dána Kolmogorovovou definicí, $A, B \in \mathcal{A}$. Pak platí*

(i) $P[\bar{A}] = 1 - P[A], \quad P[\emptyset] = 0$,

(ii) $P[A] \leq 1$,

(iii) pokud $A \subseteq B$, *pak* $P[A] \leq P[B]$ *a* $P[B \setminus A] = P[B] - P[A]$,

(iv) $P[A \cup B] = P[A] + P[B] - P[A \cap B]$,

(v) $P[A \cap B] = P[A] - P[A \cap \bar{B}]$.

Důkaz.

(i) Máme $A \cup \bar{A} = \Omega, \ A \cap \bar{A} = \emptyset$, pak $P[\Omega] = 1 = P[A] + P[\bar{A}]$. Navíc $\emptyset = \bar{\Omega}, \ P[\Omega] = 1$, a potom $P[\emptyset] = 1 - P[\Omega] = 0$.

(ii) Jelikož $0 \leq P[\bar{A}] = 1 - P[A]$, tak zřejmě $P[A] \leq 1$.

(iii) Jednoduše odvodíme, že $A \subseteq B \implies P[B] = P\left[A \cup (B \cap \bar{A})\right] = P[A] + P[B \setminus A] \implies P[B] \geq P[A]$.

(iv) Víme, že $A \cup B = A \cup \left[B \cap (\overline{A \cap B})\right]$. Potom $P[A \cup B] = P[A] + P\left[B \setminus (A \cap B)\right] = P[A] + P[B] - P[A \cap B]$.

(v) Ze vztahu $A = (A \cap B) \cup (A \cap \bar{B})$ plyne, že $P[A] = P[A \cap B] + P[A \cap \bar{B}]$. □

Věta 1.3 (Další vlastnosti pravděpodobnosti)**.** *Nechť funkce P je dána Kolmogorovovou definicí, $A, B \in \mathcal{A}$. Pak platí*

(i) pro $A_1 \subseteq A_2 \subseteq \ldots$ *je* $P\left[\bigcup_{i=1}^{\infty} A_i\right] = \lim_{i \longrightarrow \infty} P[A_i]$,

(ii) pro $A_1 \supseteq A_2 \supseteq \ldots$ *je* $P\left[\bigcap_{i=1}^{\infty} A_i\right] = \lim_{i \longrightarrow \infty} P[A_i]$,

(iii) $P\left[\bigcup_{i=1}^{\infty} A_i\right] \leq \sum_{i=1}^{\infty} P[A_i]$,

(iv) $P\left[\bigcap_{i=1}^{\infty} A_i\right] \geq 1 - \sum_{i=1}^{\infty} P[\bar{A}_i]$.

Důkaz.

(i) Definujme náhodné jevy

$$B_1 = A_1, \quad B_2 = A_2 \setminus A_1, \quad B_3 = A_3 \setminus A_2, \dots$$

Zřejmě $B_i \cap B_j = \emptyset$ pro $i \neq j$, takže $\bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} A_i$. Potom

$$\begin{aligned} P\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} P[B_i] &= \lim_{n \longrightarrow \infty} \sum_{i=1}^{n} P[B_i] \\ &= \lim_{n \longrightarrow \infty} [P[A_1] + P[A_2] - P[A_1] + \dots + P[A_n] - P[A_{n-1}]] \\ &= \lim_{n \longrightarrow \infty} P[A_n]. \end{aligned}$$

(ii) Dle přepokladu $A_1 \supseteq A_2 \supseteq \cdots \implies \bar{A}_1 \subseteq \bar{A}_2 \subseteq \cdots$ Potom

$$P\left[\bigcap_{i=1}^{\infty} A_i\right] = P\left[\overline{\bigcup_{i=1}^{\infty} \bar{A}_i}\right] = 1 - P\left[\bigcup_{i=1}^{\infty} \bar{A}_i\right] = 1 - \lim_{i \longrightarrow \infty} P[\bar{A}_i] = 1 - 1 + \lim_{i \longrightarrow \infty} P[A_i] = \lim_{i \longrightarrow \infty} P[A_i].$$

(iii) Definujme náhodné jevy

$$B_1 = A_1, \quad B_2 = A_2 \setminus A_1, \quad B_3 = A_3 \setminus (A_1 \cup A_2), \dots \quad B_i = A_i \setminus \bigcup_{j<i} A_j.$$

Zřejmě $B_i \cap B_j = \emptyset$ pro $i \neq j$ a $\bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} A_i$. Pak

$$P\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} P[B_i] = \sum_{i=2}^{\infty} P\left[A_i \setminus \bigcup_{j<i} A_j\right] + P[A_1] \leq \sum_{i=1}^{\infty} P[A_i].$$

(iv) Za pomoci komplementárních jevů potom

$$P\left[\bigcap_{i=1}^{\infty} A_i\right] = 1 - P\left[\bigcup_{i=1}^{\infty} \bar{A}_i\right] \geq 1 - \sum_{i=1}^{\infty} P[\bar{A}_i].$$

□

Definice 1.4 (Pravděpodobnostní prostor). Nechť Ω je libovolná množina, $\mathcal{A}$ je σ-algebra jejích podmnožin a P je pravděpodobnost daná Kolmogorovovou definicí. Trojice $(\Omega, \mathcal{A}, P)$ se nazývá *pravděpodobnostní prostor*.

1.3. Podmíněná pravděpodobnost

Definice 1.5 (Podmíněná pravděpodobnost). Pravděpodobnost jevu A za podmínky, že nastane jev B, je definována předpisem

$$P[A \mid B] = \frac{P[A \cap B]}{P[B]} \qquad \text{pro } P[B] > 0.$$

Poznámka. Z definice podmíněné pravděpodobnosti plyne

$$P[A \cap B] = P[A \mid B]P[B] = P[B \mid A]P[A].$$

Zde lze i $P[A] = 0,\ P[B] = 0$.

Definice 1.6 (Úplný systém jevů). Náhodné jevy $A_1, A_2, \dots$ na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ tvoří *úplný systém jevů*, pokud jsou splněny podmínky

(i) $A_i \cap A_j = \emptyset,\ i \neq j$,

(ii) $\bigcup_{i=1}^{\infty} A_i = \Omega$.

Poznámka. Nejjednodušší úplný systém obsahuje jevy $A, \bar{A}$. Nejjemnější úplný systém je prostor elementárních jevů Ω.

Věta 1.4 (O úplné pravděpodobnosti)**.** *V úplném systému jevů* $A_1, A_2, \dots$ *platí*

(i) $\sum_{i=1}^{\infty} P[A_i \mid B] = 1, \quad$ *je-li* $P[B] > 0$,

(ii) $P[B] = \sum_{i=1}^{\infty} P[B \mid A_i] P[A_i], \quad$ *je-li* $P[A_i] > 0$ *pro všechna* $i \in \mathbb{N}$.

Důkaz. Uvažujme úplný systém jevů $A_1, A_2, \dots$

(i) Vzhledem k tomu, že $B = \bigcup_{i=1}^{\infty} (A_i \cap B)$, tak

$$\sum_{i=1}^{\infty} P[A_i \mid B] = \frac{1}{P[B]} \sum_{i=1}^{\infty} P[A_i \cap B] = \frac{1}{P[B]} P[B] = 1.$$

(ii) Z definice 1.5 plyne, že

$$P[B] = \sum_{i=1}^{\infty} P[A_i \cap B] = \sum_{i=1}^{\infty} P[B \mid A_i] P[A_i].$$

□

Poznámka. Speciálně, pro $P[B] > 0$ a $0 < P[A] < 1$, platí

(i) $P[A \mid B] + P[\bar{A} \mid B] = 1$,

(ii) $P[B] = P[B \mid A] P[A] + P[B \mid \bar{A}] P[\bar{A}]$.

Věta 1.5 (Bayesův vzorec)**.** *V úplném systému jevů* $A_1, A_2, \dots$ *platí (pro fixní* j*)*

$$P[A_j \mid B] = \frac{P[B \mid A_j] P[A_j]}{\sum_{i=1}^{\infty} P[B \mid A_i] P[A_i]},$$

je-li $P[B] > 0$, $P[A_i] > 0$ *pro všechna* $i \in \mathbb{N}$.

Důkaz. Čitatel dle definice 1.5, jmenovatel dle věty 1.4. □

Věta 1.6 (O násobení pravděpodobností)**.** *Nechť pro náhodné jevy* $A_1, A_2, \dots, A_n$ *platí*

$$P\left[\bigcap_{i=1}^{n-1} A_i\right] > 0.$$

Potom

$$P\left[\bigcap_{i=1}^{n} A_i\right] = P[A_1] P[A_2 \mid A_1] P[A_3 \mid A_1 \cap A_2] \cdots P\left[A_n \middle| \bigcap_{i=1}^{n-1} A_i\right].$$

Důkaz. Zřejmě $P\left[\bigcap_{i=1}^{n-1} A_i\right] > 0 \implies P\left[\bigcap_{i=1}^{n-2} A_i\right] > 0$ atd. Dále indukcí podle n dostaneme

$$\underline{n = 2}: \quad P[A_1 \cap A_2] = P[A_1] P[A_2 \mid A_1],$$

$$\underline{n \longrightarrow n+1}: \quad P\left[\bigcap_{i=1}^{n+1} A_i\right] = P\left[\bigcap_{i=1}^{n} A_i \cap A_{n+1}\right] = P\left[\bigcap_{i=1}^{n} A_i\right] P\left[A_{n+1} \middle| \bigcap_{i=1}^{n} A_i\right]$$

a na první činitel uplatníme indukční předpoklad. □

1.4. Nezávislost náhodných jevů

Definice 1.7 (Nezávislé náhodné jevy)**.** Náhodné jevy A, B nazveme *nezávislé*, pokud platí

$$P[A \cap B] = P[A] P[B].$$

Definice 1.8 (Obecnější definice nezávislých náhodných jevů)**.** Náhodné jevy $A_1, A_2, \dots$ nazveme *sdruženě (vzájemně) nezávislé*, pokud pro libovolné $k \in \mathbb{N}$ a libovolnou k-tici $A_{i_1}, \dots, A_{i_k}$ platí

$$P\left[\bigcap_{j=1}^{k} A_{i_j}\right] = \prod_{j=1}^{k} P[A_{i_j}].$$

Poznámka. Vlastnost, že $P[A_i \cap A_j] = P[A_i]P[A_j]$ pro všechna $(i,j) \in \mathbb{N}^2$, $i \neq j$, se nazývá nezávislost po dvou. Analogicky nezávislost po třech, čtyřech apod. Zřejmě, pokud jsou $A_1, A_2, \ldots$ sdruženě nezávislé, pak jsou také po dvou nezávislé. Opačná implikace neplatí.

Definice 1.9 (Nezávislé náhodné pokusy)**.** Sérii n pokusů na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ nazveme *nezávislé náhodné pokusy*, jsou-li náhodné jevy

$$A_i \times \bigtimes_{j \neq i} \Omega_j, \quad A_i \in \mathcal{A}_i, \quad i = 1, \ldots, n$$

nezávislé.

Poznámka. Je-li pravděpodobnost P definována klasicky (dle definice 1.1), máme

$$P[A_i] = \frac{|A_i|}{|\Omega_i|}, \quad i = 1, \ldots, n,\ A_i \in \mathcal{A}_i.$$

Pro $A = A_1 \times \cdots \times A_n \in \mathcal{A}$ je při nezávislosti

$$P[A] = \frac{|A|}{|\Omega|} = \frac{|A_1 \times \cdots \times A_n|}{|\Omega_1 \times \cdots \times \Omega_n|} = \frac{|A_1| \cdots |A_n|}{|\Omega_1| \cdots |\Omega_n|} = \prod_{i=1}^{n} P[A_i].$$

Věta 1.7 (Nezávislost s doplňkovými jevy)**.** *Nechť $A_1, \ldots A_n$ jsou nezávislé náhodné jevy. Pro $1 \leq k \leq n$ a pro některou k-tici definujme $B_{i_1} = \bar{A}_{i_1}$, $B_{i_2} = \bar{A}_{i_2}, \ldots, B_{i_k} = \bar{A}_{i_k}$, $B_i = A_i$ pro $i \notin \{i_1, \ldots, i_k\}$. Pak jsou náhodné jevy $B_1, \ldots, B_n$ také nezávislé.*

Důkaz. Indukcí dle k při pevném n.

(i) $\underline{k = 1}$: Nechť bez újmy na obecnosti $i_k = i_1 = n$, tedy $B_n = \bar{A}_n$ a $B_i = A_i$ jinak. Pak lze psát

$$P\left[\bigcap_{i=1}^{n} B_i\right] = P[A_1 \cap \cdots \cap A_{n-1} \cap \bar{A}_n] = P\left[\bigcap_{i=1}^{n-1} A_i\right] - P\left[\bigcap_{i=1}^{n} A_i\right]$$

dle bodu (v) věty 1.2, kde $A = \bigcap_{i=1}^{n-1} A_i$, $B = \bar{A}_n$. Z předpokladu nezávislosti dále dostáváme

$$P\left[\bigcap_{i=1}^{n-1} A_i\right] - P\left[\bigcap_{i=1}^{n} A_i\right] = \prod_{i=1}^{n-1} P[A_i] - \prod_{i=1}^{n} P[A_i] = \prod_{i=1}^{n-1} P[A_i]\,(1 - P[A_n]) = \prod_{i=1}^{n} P[B_i].$$

Pro libovolnou podmnožinu množiny jevů $\{B_1, \ldots, B_n\}$ platí totéž, neboť buď neobsahuje $B_n = \bar{A}_n$, a pak tvrzení plyne přímo z nezávislosti, anebo obsahuje $B_n = \bar{A}_n$, a pak použijeme důkaz výše.

(ii) $\underline{k \longrightarrow k+1}$: Nechť bez újmy na obecnosti $i_1 = n-k, \ldots, i_{k+1} = n$. Potom lze psát

$$\begin{aligned}
P\left[\bigcap_{i=1}^{n} B_i\right] &= P[A_1 \cap \cdots \cap A_{n-k-1} \cap \bar{A}_{n-k} \cap \cdots \cap \bar{A}_n] \\
&= P\left[\bigcap_{i=1}^{n-k-1} A_i \cap \bigcap_{i=n-k}^{n-1} \bar{A}_i\right] - P\left[\bigcap_{i=1}^{n-k-1} A_i \cap \bigcap_{i=n-k}^{n-1} \bar{A}_i \cap A_n\right] \\
&= \prod_{i=1}^{n-k-1} P[A_i] \prod_{i=n-k}^{n-1} P[\bar{A}_i]\,(1 - P[A_n]) \\
&= \prod_{i=1}^{n} P[B_i].
\end{aligned}$$

Třetí rovnost plyne z nezávislosti a indukčního předpokladu. Pro libovolnou podmnožinu jevů $B_1, \ldots, B_n$ platí totéž. $\square$

Poznámka. Speciálně platí, že pokud jsou A, B nezávislé, potom jsou také

(i) $\bar{A}, B$ nezávislé, (ii) $A, \bar{B}$ nezávislé, (iii) $\bar{A}, \bar{B}$ nezávislé.

Věta 1.8 (Pravděpodobnost sjednocení nezávislých jevů)**.** *Nechť* $A_1, \dots A_n$ *jsou nezávislé náhodné jevy. Pak platí*

$$P\left[\bigcup_{i=1}^{n} A_i\right] = 1 - \prod_{i=1}^{n} P[\bar{A}_i].$$

Důkaz. Z de Morganových pravidel dostáváme

$$P\left[\bigcup_{i=1}^{n} A_i\right] = P\left[\overline{\bigcap_{i=1}^{n} \bar{A}_i}\right] = 1 - P\left[\bigcap_{i=1}^{n} \bar{A}_i\right] = 1 - \prod_{i=1}^{n} P[\bar{A}_i].$$

Poslední rovnost plyne z věty 1.7 a definice 1.8. □

2. Náhodné veličiny

2.1. Pravděpodobnostní rozdělení

Definice 2.1 (Měřitelný prostor)**.** *Měřitelný prostor* $(M, \mathcal{B})$ je neprázdná množina M se σ-algebrou $\mathcal{B}$.

Úmluva. *Nadále mějme dán pravděpodobnostní prostor* $(\Omega, \mathcal{A}, P)$ *a měřitelný prostor* $(M, \mathcal{B})$, *který lze nazvat výběrovým prostorem.*

Definice 2.2 (Měřitelné zobrazení)**.** Zobrazení $X : (\Omega, \mathcal{A}) \longrightarrow (M, \mathcal{B})$ je *měřitelné* (vzhledem k σ-algebrám $\mathcal{A}$ a $\mathcal{B}$), když $X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A}$ pro všechny $B \in \mathcal{B}$. Pokud toto platí, říkáme, že vzory měřitelných množin jsou měřitelné množiny.

Definice 2.3 (Náhodná veličina)**.** Měřitelné zobrazení $X : (\Omega, \mathcal{A}, P) \longrightarrow (M, \mathcal{B})$ nazýváme *náhodná veličina.*

Značení. Nadále pro pevné $\omega \in \Omega$ budeme konkrétní realizaci náhodné veličiny značit jako $X(\omega)$.

Definice 2.4 (Borelovská σ-algebra)**.** *Borelovská σ-algebra* $\mathcal{B}_0$ je nejmenší σ-algebra obsahující otevřené množiny na prostoru reálných čísel $\mathbb{R}$.

Poznámka. $\mathcal{B}_0$ zřejmě kromě otevřených intervalů obsahuje i všechny uzavřené a polouzavřené intervaly na $\mathbb{R}$, neboť pro $a < b$ lze psát

$$[a, b] = (-\infty, b] \setminus (-\infty, a) = [\mathbb{R} \setminus (b, \infty)] \setminus (-\infty, a).$$

Analogicky $[a, b) = (-\infty, b) \setminus (-\infty, a)$ a $(a, b] = (a, \infty) \setminus (b, \infty)$.

Příklad 2.1. Volby množiny M.

(i) $M = \mathbb{R} \dots$ reálná náhodná veličina $X = X(\omega)$ přiřazuje elementárním jevům $\omega \in \Omega$ reálná čísla;

(ii) $M = \mathbb{R}^n \dots$ zde máme náhodný vektor $\boldsymbol{X} = (X_1, \dots, X_n)$ s náhodnými veličinami $X_i = X_i(\omega)$, $\omega \in \Omega$, kde $i = 1, \dots, n$;

(iii) $M = \mathbb{R}^\infty \dots$ prostor posloupností reálných čísel s náhodnou posloupností $\boldsymbol{X} = \{X_1, X_2, \dots\}$ a náhodnými veličinami $X_i = X_i(\omega)$, $\omega \in \Omega$, $i = 1, 2, \dots$;

(iv) $M = C(\mathbb{R}^+) \dots$ prostor spojitých funkcí na $[0, \infty)$ s náhodným procesem $\boldsymbol{X} = \{X_t,\, t \geq 0\}$, kde $X_t = X_t(\omega)$, $\omega \in \Omega$. ▲

Definice 2.5 (Míra)**.** Nechť $(M, \mathcal{B})$ je měřitelný prostor. Funkce $\mu : \mathcal{B} \longrightarrow \mathbb{R}^+$ se nazývá *míra*, pokud platí, že

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

(ii) pokud $B_i \cap B_j = \emptyset$ pro $B_1, B_2, \dots \in \mathcal{B}$, $i \neq j$, potom

$$\mu\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} \mu(B_i).$$

Poznámka. Míra je zřejmě nezáporná, σ-aditivní funkce. Kolmogorovova definice (definice 1.3) definuje pravděpodobnost jako míru na $(\Omega, \mathcal{A})$.

Definice 2.6 (Vlastnosti míry)**.** Uvažujme měřitelný prostor $(M, \mathcal{B})$ s mírami μ, P_X.

(i) μ se nazývá *σ-konečná*, pokud existují $B_1, B_2, \dots \in \mathcal{B}$ takové, že $\mu(B_i) < \infty$ pro všechna i a platí $\bigcup_{i=1}^{\infty} B_i = M$.

(ii) P_X se nazývá *absolutně spojitá vzhledem k μ*, pokud platí, že když $\mu(B) = 0$, potom $P_X(B) = 0$ pro všechny $B \in \mathcal{B}$.

Definice 2.7 (Pravděpodobnostní rozdělení)**.** *Pravděpodobnostním rozdělením* náhodné veličiny $X : (\Omega, \mathcal{A}) \longrightarrow (M, \mathcal{B})$ rozumíme míru P_X indukovanou náhodnou veličinou X na $(M, \mathcal{B})$ ve smyslu

$$P_X(B) = P\left[\{\omega : X(\omega) \in B\}\right] = P[X \in B], \quad B \in \mathcal{B}.$$

Poznámka. Zřejmě $\{\omega : X(\omega) \in B\} = X^{-1}(B) \in \mathcal{A}$ je náhodný jev, kterému při zobrazení X odpovídá množina $B \in \mathcal{B}$. Tedy X transformuje prostor $(\Omega, \mathcal{A}, P)$ na prostor $(M, \mathcal{B}, P_X)$.

Věta 2.1 (O přenosu integrace)**.** *Nechť $X : (\Omega, \mathcal{A}) \longrightarrow (M, \mathcal{B})$ je náhodná veličina a $h : (M, \mathcal{B}) \longrightarrow (\mathbb{R}, \mathcal{B}_0)$ je měřitelná funkce. Pak platí*

$$\int_\Omega h\left(X(\omega)\right) \mathrm{d}P(\omega) = \int_M h(x) \, \mathrm{d}P_X(x).$$

Důkaz. Vynechán. □

Věta 2.2 (Radonova-Nikodymova věta)**.** *Nechť $X : (\Omega, \mathcal{A}) \longrightarrow (M, \mathcal{B})$ je náhodná veličina a $h : (M, \mathcal{B}) \longrightarrow (\mathbb{R}, \mathcal{B}_0)$ je měřitelná funkce. Dále nechť μ je σ-konečná míra na $(M, \mathcal{B})$ a nechť P_X je absolutně spojitá vzhledem k μ. Pak existuje nezáporná měřitelná funkce $f_X : (M, \mathcal{B}) \longrightarrow (\mathbb{R}, \mathcal{B}_0)$ taková, že platí*

$$\int_M h(x) \, \mathrm{d}P_X(x) = \int_M h(x) f_X(x) \, \mathrm{d}\mu(x).$$

Důkaz. Vynechán. □

Poznámka. Funkce $f_X(x)$ je určena jednoznačně μ-skoro všude, tedy skoro všude vzhledem k μ. Existuje-li tedy funkce $g(x)$ s vlastnostmi funkce $f_X(x)$ z věty 2.2, tak $g(x) = f_X(x)$ pro všechna $x \in (M \setminus N)$, kde $\mu(N) = 0$.

Definice 2.8 (Hustota náhodné veličiny)**.** Funkce $f_X(x)$ z věty 2.2 se nazývá *hustota* náhodné veličiny X vzhledem k μ.

Poznámka. Speciálně,

$$h(x) = \mathbb{I}_B(x) = \begin{cases} 1; & x \in B, \\ 0; & x \notin B \end{cases}$$

pro nějakou množinu $B \in \mathcal{B}$.

Důsledek 2.1. *Hustota jednoznačně určuje pravděpodobnostní rozdělení náhodné veličiny X.*

Důkaz. Dle věty 2.1 máme

$$\int_\Omega \mathbb{I}_B\left(X(\omega)\right) \mathrm{d}P(\omega) = \int_M \mathbb{I}_B(x) \, \mathrm{d}P_X(x) = \int_B 1 \, \mathrm{d}P_X(x) = P_X(B) = P\left[X \in B\right]$$

a dle věty 2.2 potom

$$P\left[X \in B\right] = \int_M \mathbb{I}_B(x) f_X(x) \, \mathrm{d}\mu(x) = \int_B f_X(x) \, \mathrm{d}\mu(x).$$ □

Úmluva. *Nadále uvažujme reálnou náhodnou veličinu $X : (\Omega, \mathcal{A}) \longrightarrow (\mathbb{R}, \mathcal{B}_0)$.*

Poznámka (Míry na $(\mathbb{R}, \mathcal{B}_0)$)**.**

(i) Lebesgueova míra je definována skrze $\mu_L((a, b]) = b - a$ pro všechna $a < b \in \mathbb{R}$. (Stačí definovat na generátoru borelovské σ-algebry.)

(ii) Nechť F je daná neklesající zprava spojitá reálná funkce taková, že

$$\lim_{x \longrightarrow -\infty} F(x) = 0, \quad \lim_{x \longrightarrow \infty} F(x) = 1.$$

Lebesgueovu-Stieltjesovu míru pak definujeme jako $\mu_F((a, b]) = F(b) - F(a)$ pro všechna $a < b \in \mathbb{R}$.

(iii) Nechť $S = \{x_1, x_2, \dots\} \subset \mathbb{R}$ je nejvýše spočetná množina reálných čísel. Čítací míru definujeme jako

$$\mu_S(B) = \#\{i : x_i \in B\} = \sum_{i=1}^{\infty} \mathbb{I}_B(x_i)$$

pro $B \in \mathcal{B}_0$, tedy počet bodů z nosiče S, které se nacházejí v množině B.

Definice 2.9 (Diskrétní a spojitá náhodná veličina)**.** Je-li P_X absolutně spojitá vzhledem k čítací míře μ_S, pak udává diskrétní pravděpodobnostní rozdělení a X nazýváme *diskrétní náhodná veličina*. Pro diskrétní veličiny a pro $S = \{x_1, x_2, \dots\} \subset \mathbb{R}$ je hustota vzhledem k μ_S charakterizována posloupností pravděpodobností

$$0 < p_i = P[X = x_i] < 1, \ i = 1, 2, \dots$$

Pro $B \in \mathcal{B}_0$ pak platí

$$P_X(B) = P[X \in B] = \sum_{x_i \in S} \mathbb{I}_B(x_i) p_i = \sum_{x_i \in B \cap S} p_i$$

a pro $B = S$ máme

$$\sum_{x_i \in S} p_i = 1.$$

Je-li P_X absolutně spojitá vzhledem k Lebesgueově míře μ_L, pak udává spojité pravděpodobnostní rozdělení a X nazýváme *spojitá náhodná veličina*. Pro spojité náhodné veličiny je hustota vzhledem k μ_L nezáporná funkce $f_X : \mathbb{R} \longrightarrow \mathbb{R}$. Pro $B \in \mathcal{B}_0$ pak platí

$$P_X(B) = P[X \in B] = \int_{-\infty}^{\infty} \mathbb{I}_B(x) f_X(x) \, \mathrm{d}x = \int_B f_X(x) \, \mathrm{d}x,$$

a pro $B = \mathbb{R}$ máme

$$\int_{-\infty}^{\infty} f_X(x) \, \mathrm{d}x = 1.$$

Definice 2.10 (Distribuční funkce)**.** Funkce $F_X(x) = P[X \leq x]$, $x \in \mathbb{R}$ se nazývá *distribuční funkce* náhodné veličiny X.

Poznámky.

(i) Pravděpodobnostní rozdělení reálné náhodné veličiny X lze chápat jako systém pravděpodobností

$$\{P_X(B) : B \in \mathcal{B}_0\}.$$

Je-li tedy dáno pravděpodobnostní rozdělení, je jím určena distribuční funkce, neboť

$$P[X \leq \ x] = P_X[(-\infty, x]],$$

kde interval $(-\infty, x]$ je speciální případ $B \in \mathcal{B}_0$.

(ii) Opačná implikace, tedy že distribuční funkce F_X jednoznačně určuje pravděpodobnostní rozdělení P_X, souvisí s tím, že intervaly $(-\infty, x]$ lze položit jako generátory borelovské σ-algebry $\mathcal{B}_0$. Míru P_X lze jednoznačně rozšířit z těchto generátorů na celou $\mathcal{B}_0$.

(iii) Zřejmě

(a) pro diskrétní náhodné veličiny je $F_X(x) = \sum_i p_i$ po částech konstantní (sčítáme přes i splňující $x_i \leq x$);

(b) pro spojité náhodné veličiny je $F_X(x) = \int_{-\infty}^{x} f_X(u) \, \mathrm{d}u$ spojitá.

(iv) Pro vztah distribuční funkce F_X a hustoty f_X platí

$$\int_\Omega h(X(\omega)) \, \mathrm{d}P(\omega) = \int_M h(x) \, \mathrm{d}P_X(x) = \int_M h(x) f_X(x) \, \mathrm{d}\mu(x) = \int_{-\infty}^{\infty} h(x) \, \mathrm{d}F_X(x).$$

První rovnost plyne z věty 2.1, druhá rovnost plyne z věty 2.2 a ve třetí rovnosti použijeme $M = \mathbb{R}$. Poslední integrál je Lebesgueův-Stieltjesův integrál. Navíc, $f_X(x) = F'_X(x)$ pro spojitou náhodnou veličinu X a $x \in \mathbb{R}$.

Věta 2.3 (Základní vlastnosti distribuční funkce)**.** *Pro distribuční funkci $F_X(x)$, $x \in \mathbb{R}$ platí, že*

(i) *je neklesající,*

(ii) *je zprava spojitá,*

(iii) $\lim_{x \longrightarrow -\infty} F_X(x) = 0$,

(iv) $\lim_{x \longrightarrow +\infty} F_X(x) = 1$.

Důkaz.

(i) Položme $x \leq y$. Pak z nerovnosti $X \leq x$ plyne, že $X \leq y$, a dle věty 1.2 potom

$$P[X \leq x] = F_X(x) \leq F_X(y) = P[X \leq y].$$

Tedy F_X je neklesající.

(ii) Pro $n \longrightarrow \infty$ máme

$$x_n = x + \frac{1}{n} \longrightarrow x^+,$$

a proto

$$\lim_{x_n \longrightarrow x^+} F_X(x_n) = \lim_{n \longrightarrow \infty} P\left[X \leq x + \frac{1}{n}\right] = P\left[\bigcap_{n=1}^{\infty}\left(X \leq x + \frac{1}{n}\right)\right] = P[X \leq x] = F_X(x).$$

Tedy F_X je zprava spojitá. Druhá rovnost plyne z věty 1.3.

(iii) Počítejme:

$$\lim_{x \longrightarrow -\infty} F_X(x) = \lim_{n \longrightarrow \infty} P[X \leq -n] = P\left[\bigcap_{n=1}^{\infty}(X \leq -n)\right] = P[\emptyset] = 0.$$

Druhá rovnost plyne z věty 1.3.

(iv) Počítejme:

$$\lim_{x \longrightarrow \infty} F_X(x) = \lim_{n \longrightarrow \infty} P[X \leq n] = P\left[\bigcup_{n=1}^{\infty}(X \leq n)\right] = P[\Omega] = 1.$$

Druhá rovnost plyne z věty 1.3. □

Poznámka. V důkazu výše je pro $n \longrightarrow \infty$

(i) $X \leq x + \frac{1}{n}$ nerostoucí posloupnost náhodných jevů,

(ii) $X \leq -n$ nerostoucí posloupnost náhodných jevů,

(iii) $X \leq n$ neklesající posloupnost náhodných jevů.

Věta 2.4 (Pravděpodobnost jednobodové množiny)**.** *Nechť $F_X(x)$ je distribuční funkce a označme*

$$F_X(x^-) = \lim_{y \longrightarrow x^-} F_X(y).$$

Potom pro $x \in \mathbb{R}$ platí

$$P[X = x] = F_X(x) - F_X(x^-).$$

Důkaz. Můžeme psát

$$F_X(x^-) = \lim_{n \longrightarrow \infty} F_X\left(x - \frac{1}{n}\right) = \lim_{n \longrightarrow \infty} P\left[X \leq x - \frac{1}{n}\right] = P\left[\bigcup_{n=1}^{\infty}\left(X \leq x - \frac{1}{n}\right)\right] = P[X < x].$$

Potom zřejmě

$$P[X \leq x] = F_X(x) = P[X = x] + P[X < x] \implies P[X = x] = F_X(x) - P[X < x].$$ □

Důsledek 2.2.

(i) Distribuční funkce F_X *je spojitá právě tehdy, když* $P[X = x] = 0$ *pro všechna* $x \in \mathbb{R}$.

(ii) Má-li F_X *body nespojitosti, je funkcí skoků, přičemž velikost skoku v bodě* x *je* $P[X = x]$.

Věta 2.5 (Body nespojitosti distribuční funkce)**.** *Distribuční funkce* $F_X(x)$, $x \in \mathbb{R}$ *má nejvýše spočetně mnoho bodů nespojitosti.*

Důkaz. Označme
$$N_n = \left\{x \in \mathbb{R} : P[X = x] > \frac{1}{n}\right\},$$
kde $n = 2, 3, \dots$ Protože $0 \le F_X(x) \le 1$, množina N_n obsahuje nejvýše $n-1$ bodů, v nichž je skok o velikosti větší než $\frac{1}{n}$. Položme
$$N = \bigcup_{n=2}^{\infty} N_n$$
jako množinu bodů, v nichž má F_X skok, $N = \{x \in \mathbb{R} : P[X = x] \neq 0\}$. Vzhledem k tomu, že N je spočetné sjednocení konečných množin, obsahuje nejvýše spočetně mnoho bodů $x \in \mathbb{R}$. □

Shrnutí.
Diskrétní náhodná veličina má hustotu vzhledem k čítací míře μ_S (kde $S = \{x_1, x_2, \dots\} \subset \mathbb{R}$ je nejvýše spočetná), takže nabývá hodnot $x_1, x_2, \dots$, jež jsou body nespojitosti funkce F_X, která má skoky o velikosti $p_i = P[X = x_i]$, $i = 1, 2, \dots$
Spojitá náhodná veličina má hustotu $f_X(x)$ vzhledem k Lebesgueově míře μ_L, takže nabývá hodnot $x \in (a, b)$, kde $-\infty \le a < b \le \infty$, přičemž $f_X(x) > 0$ pro $x \in (a, b)$ a $f_X(x) = 0$ jinak. Zároveň $F_X(x)$ je spojitá a rostoucí na (a, b) a platí
$$\int_a^b f_X(x)\,\mathrm{d}x = 1, \quad f_X(x) = F'_X(x),\ x \in (a, b).$$

Věta 2.6 (Pravděpodobnost intervalu)**.** *Nechť* X *je reálná náhodná veličina s distribuční funkcí* $F_X(x)$ *a nechť* $a, b \in \mathbb{R}$, $a < b$. *Pak platí*
$$P[a < X \le b] = F_X(b) - F_X(a).$$

Důkaz. Označme náhodný jev $A : X \le a$ a náhodný jev $B : X \le b$. Zřejmě $A \subseteq B$ a platí
$$P[a < X \le b] = P[\bar{A} \cap B] = P[B \setminus A] = P[B] - P[A] = F_X(b) - F_X(a).$$
Třetí rovnost platí podle bodu (iii) věty 1.2. □

Poznámka. Vzhledem k důsledku 2.2 platí u spojité náhodné veličiny X
$$\begin{aligned} P[a < X < b] &= P[a \le X < b] = P[a < X \le b] = P[a \le X \le b] \\ &= F_X(b) - F_X(a) = \int_{-\infty}^b f_X(x)\,\mathrm{d}x - \int_{-\infty}^a f_X(x)\,\mathrm{d}x \\ &= \int_a^b f_X(x)\,\mathrm{d}x. \end{aligned}$$

Definice 2.11 (Kvantilová funkce)**.** Nechť X je reálná náhodná veličina a $F_X(x)$, $x \in \mathbb{R}$ její distribuční funkce. Funkce $F_X^{-1}(u) = \inf\{x : F_X(x) \ge u\}$, $u \in (0, 1)$ se nazývá *kvantilová funkce.*

Poznámka. Kvantilová funkce jednoznačně určuje distribuční funkci, a tedy i pravděpodobnostní rozdělení náhodné veličiny X. Je neklesající a zleva spojitá. U spojitých náhodných veličin je F_X^{-1} inverzní funkcí k F_X. V takovém případě zřejmě pro $\alpha \in (0, 1)$ platí
$$P\left[X \le F_X^{-1}(\alpha)\right] = F_X\left[F_X^{-1}(\alpha)\right] = \alpha.$$

Definice 2.12 (Kvantil)**.** Nechť X je reálná náhodná veličina, F_X její distribuční funkce a $\alpha \in (0, 1)$. Reálné číslo $u_X(\alpha)$ splňující
$$F_X[u_X(\alpha)] \ge \alpha, \quad F_X[u_X(\alpha)^-] \le \alpha$$
nazýváme *α-kvantil.*

Poznámky.

(i) Při označení

$$F_X[u_X(\alpha)^-] = \lim_{y \longrightarrow u_X(\alpha)^-} F_X(y)$$

máme dle věty 2.4

$$P[X \leq u_X(\alpha)] \geq \alpha, \quad P[X < u_X(\alpha)] \leq \alpha.$$

(ii) Definicí 2.12 není α-kvantil určen jednoznačně. Hodnota kvantilové funkce $F_X^{-1}(\alpha)$ je jedním z α-kvantilů. U spojitých náhodných veličin je $u_X(\alpha) = F_X^{-1}(\alpha)$ jediný α-kvantil.

Příklad 2.2.

(i) $\alpha = 0.75 \dots\ u_X(0.75) \dots$ 3. (horní) kvartil,

(ii) $\alpha = 0.50 \dots\ u_X(0.50) \dots$ medián,

(iii) $\alpha = 0.25 \dots\ u_X(0.25) \dots$ 1. (dolní) kvartil,

(iv) $\alpha = 0.10 \dots\ u_X(0.10) \dots$ 1. decil,

(v) $\alpha = 0.01 \dots\ u_X(0.01) \dots$ 1. percentil,

(vi) percentil lze použít jako synonymum pro kvantil. ▲

2.2. Momenty reálné náhodné veličiny

Definice 2.13 (Střední hodnota)**.** *Střední hodnotu* náhodné veličiny X definujeme předpisem

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

pokud takový integrál existuje.

Poznámky.

(i) Pro $X : (\Omega, \mathcal{A}) \longrightarrow (\mathbb{R}, \mathcal{B}_0)$ a $h(x) = x$, $x \in \mathbb{R}$ plyne z věty 2.1 a věty 2.2, že

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}\mu(x).$$

(ii) Pro měřitelnou funkci $h : (\mathbb{R}, \mathcal{B}_0) \longrightarrow (\mathbb{R}, \mathcal{B}_0)$ máme

$$E[h(X)] = \int_\Omega h(X(\omega))\,\mathrm{d}P(\omega) = \int_{-\infty}^{\infty} h(x) f_X(x)\,\mathrm{d}\mu(x).$$

(iii) Pro diskrétní náhodnou veličinu, kdy $\mu = \mu_S$, máme

$$E[X] = \sum_i x_i p_i,$$

tedy střední hodnota je vážený průměr hodnot, kterých X nabývá, s vahami $p_i = P[X = x_i]$, $i = 1, 2, \dots$, přičemž $S = \{x_1, x_2, \dots\}$ je nosič. V tomto případě pak

$$E[h(X)] = \sum_i h(x_i) p_i.$$

(iv) Pro spojitou náhodnou veličinu, kdy $\mu = \mu_L$, máme

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}x.$$

Zde integrujeme přes obor, na kterém je $f(x) > 0$, což je obor hodnot, kterých X nabývá. Platí pak tedy

$$E[h(X)] = \int_{-\infty}^{\infty} h(x) f_X(x)\,\mathrm{d}x.$$

(v) Střední hodnota má význam úrovně, kolem které náhodná veličina kolísá. Je tzv. měrou polohy. Značení souvisí s anglickým expectation, expected value. Samotná střední hodnota neposkytuje dostatečnou informaci o hodnotách náhodné veličiny, protože náhodná veličina může kolem své střední hodnoty kolísat s různým rozsahem.

Úmluva. *Nadále budeme značit $L^p(\Omega)$ jako množinu reálných náhodných veličin, pro které $E|X|^p < \infty$.*

Věta 2.7 (Vlastnosti střední hodnoty)**.** *Nechť $X, Y \in L^1(\Omega)$. Potom platí*

(i) $E[a + bX + cY] = a + bE[X] + cE[Y],\ a, b, c \in \mathbb{R}$;

(ii) pokud $P[X \leq Y] = 1$, *potom* $E[X] \leq E[Y]$;

(iii) pokud existuje $\mu \in \mathbb{R}$, že pro všechna $x \in \mathbb{R}$ platí $f_X(\mu - x) = f_X(\mu + x)$, potom $E[X] = \mu$. (symetrie)

Důkaz.

(i) Plyne z toho, že integrál je lineární funkcionál.

(ii) Vynechán.

(iii) Využijeme symetrie hustoty kolem svislé osy v bodě μ. Předpokládejme bez újmy na obecnosti, že $\mu = 0$. Potom

$$
\begin{aligned}
E[X] &= \int_{-\infty}^{0} x f_X(x) \,\mathrm{d}\mu(x) + \int_{0}^{\infty} x f_X(x) \,\mathrm{d}\mu(x) \\
&= \int_{\infty}^{0} -x^* f_X(-x^*)(-1) \,\mathrm{d}\mu(x^*) + \int_{0}^{\infty} x f_X(x) \,\mathrm{d}\mu(x) \\
&= -\int_{0}^{\infty} x^* f_X(x^*) \,\mathrm{d}\mu(x^*) + \int_{0}^{\infty} x f_X(x) \,\mathrm{d}\mu(x) \\
&= 0.
\end{aligned}
$$

V první úpravě integrálu byla použita substituce $x^* = -x$, ve druhé potom symetrie f_X. □

Poznámka. Speciálně pro střední hodnotu platí za předpokladu z věty 2.7 následující vztahy.

(i) $E[a] = a,\ a \in \mathbb{R}$;

(ii) $E[a + X] = a + E[X],\ a \in \mathbb{R}$;

(iii) $E[aX] = aE[X],\ a \in \mathbb{R}$;

(iv) $E[X + Y] = E[X] + E[Y]$;

(v) $E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X]_i$, pokud $X_i \in L^1(\Omega)$, $i = 1, \ldots, n$;

(vi) $E\left[\sum_{i=1}^{\infty} X_i\right] = \sum_{i=1}^{\infty} E[X]_i$, pokud $X_i \in L^1(\Omega)$, $i = 1, \ldots, n$ a $\sum_{i=1}^{\infty} E[X]_i < \infty$.

Poznámka. Je-li X spojitá náhodná veličina a f_X symetrická kolem μ, pak podle věty 2.7 platí

$$E[X] = \mu$$

a dle poznámky za definicí 2.11 pro medián $\tilde{x} = u_X(\frac{1}{2})$ platí

$$F_X\left[u_X\left(\frac{1}{2}\right)\right] = \frac{1}{2} = \int_{-\infty}^{\tilde{x}} f_X(x) \,\mathrm{d}x = \int_{\tilde{x}}^{\infty} f_X(x) \,\mathrm{d}x,$$

neboť $\int_{-\infty}^{\infty} f_X(x) \,\mathrm{d}x = 1$. Z toho plyne, že f_X musí být symetrická kolem $\tilde{x}$, a tedy $\mu = E[X] = \tilde{x} = u_X(\frac{1}{2})$.

Definice 2.14 (Momenty)**.** Střední hodnota $E[X]^k = \mu'_k$ se nazývá *k-tý moment* náhodné veličiny X. Definujeme také $E|X|^k$ jako *k-tý absolutní moment* a $E[X - E[X]]^k = \mu_k$ jako *k-tý centrální moment.*

Poznámky. Momenty nemusí vždy existovat.

(i) Uvažujme diskrétní náhodnou veličinu s pravděpodobnostním rozdělením

$$x_i = 2^i, \quad p_i = P[X = x_i] = \frac{1}{2^i}, \quad i = 1, 2, \dots$$

Potom

$$\sum_{i=1}^{\infty} p_i = \frac{\frac{1}{2}}{1 - \frac{1}{2}} = 1, \quad E[X] = \sum_{i=1}^{\infty} 2^i \frac{1}{2^i}.$$

Zde tedy $E[X]$ není konečná.

(ii) Uvažujme spojitou náhodnou veličinu s pravděpodobnostním rozdělením s hustotou

$$f_X(x) = \frac{1}{\pi} \frac{1}{1 + x^2}, \quad x \in \mathbb{R}.$$

Pak platí

$$\int_{-\infty}^{\infty} f_X(x) \,\mathrm{d}x = \frac{1}{\pi} \left(\int_{-\infty}^{\infty} \frac{1}{1 + x^2} \,\mathrm{d}x \right) = \frac{1}{\pi} [\arctan(x)]_{-\infty}^{\infty} = \frac{1}{\pi} \left(\frac{\pi}{2} - \left(-\frac{\pi}{2} \right) \right) = 1,$$

ale střední hodnota

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) \,\mathrm{d}x = \frac{1}{\pi} \left(\int_{-\infty}^{\infty} \frac{x}{1 + x^2} \,\mathrm{d}x \right) = \left[\text{Substituce: } 1 + x^2 = t,\ x = \sqrt{t-1},\ \mathrm{d}x = \frac{1}{2\sqrt{t-1}} \,\mathrm{d}t \right]$$

$$= \frac{1}{2\pi} \left(\int_1^{\infty} \frac{\sqrt{t-1}}{t} \frac{1}{\sqrt{t-1}} \,\mathrm{d}t \right) = \frac{1}{2\pi} \left(\int_1^{\infty} \frac{1}{t} \,\mathrm{d}t \right) = \frac{1}{2\pi} \left[\ln(t) \right]_1^{\infty}$$

není konečná.

Definice 2.15 (Rozptyl, směrodatná odchylka)**.** Druhý centrální moment

$$\operatorname{Var}[X] = \sigma^2 = E[X - E[X]]^2$$

se nazývá *rozptyl* náhodné veličiny X. *Směrodatná odchylka* je potom definována jako

$$\sqrt{\operatorname{Var}[X]} = \sigma > 0.$$

Poznámka. Rozptyl měří rozsah kolísání náhodné veličiny kolem její střední hodnoty. Je tzv. měrou variability. V případě, že chceme mít míru variability ve stejných jednotkách jako náhodnou veličinu, použijeme směrodatnou odchylku σ nebo střední absolutní odchylku $E|X - E[X]|$.

Věta 2.8 (Vlastnosti rozptylu)**.** *Nechť $X \in L^2(\Omega)$. Potom platí*

(i) $\operatorname{Var}[a + bX] = b^2 \operatorname{Var}[X],\ a, b \in \mathbb{R}$;

(ii) $\operatorname{Var}[X] = E[X]^2 - (E[X])^2$;

(iii) $\operatorname{Var}[X] \geq 0$;

(iv) $\operatorname{Var}[X] = 0$ *právě tehdy, když existuje $\mu \in \mathbb{R}$ takové, že* $P[X = \mu] = 1$.

Důkaz.

(i) $\operatorname{Var}[a + bX] = E[a + bX - E[a + bX]]^2 = E[b^2(X - E[X])^2] = b^2 E[X - E[X]]^2 = b^2 \operatorname{Var}[X]$.
Druhá a třetí rovnost plyne z věty 2.7.

(ii) $\operatorname{Var}[X] = E[X^2 - 2XE[X] + (E[X])^2] = E[X]^2 - 2E[X]E[X] + (E[X])^2 = E[X]^2 - (E[X])^2$.
Druhá rovnost opět plyne z věty 2.7.

(iii) $\operatorname{Var}[X] = E[X - E[X]]^2 \geq 0$ zřejmě z vlastnosti integrálu (integrál z nezáporných hodnot je nezáporný).

(iv) $\operatorname{Var}[X] = 0$ je ekvivalentní s $E[X - E[X]]^2 = 0$, a to platí právě tehdy (podle věty 2.1 a definice 2.13), když $\int_{-\infty}^{\infty} (X - E[X])^2 \,\mathrm{d}P_X(x) = 0$, což je ekvivalentní s $P[X = E[X]] = 1$ a $E[X] = \mu$. □

Definice 2.16 (Šikmost a špičatost)**.** *Šikmost* náhodné veličiny X je definována předpisem

$$\gamma_3 = \frac{\mu_3}{\sigma^3}.$$

Špičatost je definována předpisem

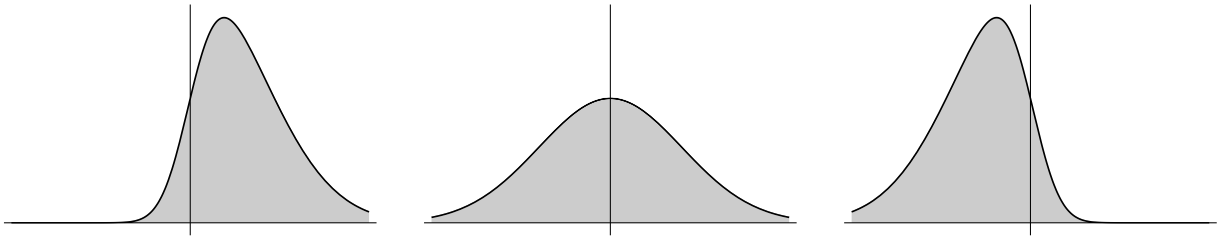
$$\gamma_4 = \frac{\mu_4}{\sigma^4}.$$

Poznámky. Šikmost má vliv na symetrii hustoty.

(i) Pro $\gamma_3 > 0$ je hustota zešikmená doprava a X s nezanedbatelnou pravděpodobností nabývá extrémně velkých hodnot.

(ii) Pro $\gamma_3 = 0$ je hustota symetrická.

(iii) Pro $\gamma_3 < 0$ je hustota zešikmená doleva a X s nezanedbatelnou pravděpodobností nabývá extrémně malých hodnot.



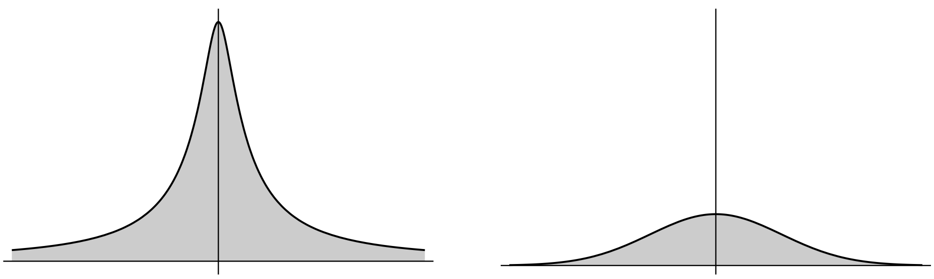
(a) Šikmost kladná, $\gamma_3 > 0$ (b) Šikmost nulová, $\gamma_3 = 0$ (c) Šikmost záporná, $\gamma_3 < 0$

Obrázek 2.1: Ilustrace šikmosti pro různá γ_3

Poznámky. Špičatost má vliv na koncentraci hustoty.

(i) Pro γ_4 velké má rozdělení těžké konce (chvosty), jedná se o tzv. špičaté neboli leptokurtické rozdělení.

(ii) Pro γ_4 malé má rozdělení plochou hustotu, jedná se o tzv. ploché neboli platykurtické rozdělení.



(a) Špičaté rozdělení, $\gamma_4 >> 0$ (b) Ploché rozdělení, $\gamma_4 \approx 0$

Obrázek 2.2: Ilustrace špičatosti pro různá γ_4

Věta 2.9 (Jensenova nerovnost)**.** *Nechť g je konvexní funkce na intervalu (a, b), $-\infty \leq a < b \leq \infty$, a $X \in L^1(\Omega)$ je náhodná veličina taková, že $P[a < X < b] = 1$ a $E[g(X)]$ existuje. Pak platí $E[g(X)] \geq g(E[X])$, přičemž rovnost nastává právě tehdy, když g je lineární (nebo konstantní) funkcí na (a, b).*

Důkaz. Platí, že pokud $g(x)$ je konvexní na (a, b), potom $g(x)$ je spojitá na (a, b). Označme $E[X] = \mu$, $b_\mu = g'(\mu)$, pokud tato první derivace existuje, a $b_\mu = \frac{1}{2}[g'(\mu-) - g'(\mu+)]$ jinak. Z konvexity $g(x)$ plyne

$$g(x) \geq g(\mu) + b_\mu(x - \mu), \quad x \in (a, b). \tag{2.1}$$

Nechť nyní X je náhodná veličina a aplikujeme nerovnost (2.1) na střední hodnotu. S využitím věty 2.7 dostaneme

$$E[g(X)] \geq g(E[X]) + b_\mu(E[X] - \mu),$$

kde $E[X] - \mu = 0$. Rovnost $E[g(X)] = g(E[X])$ nastává při rovnosti v (2.1). Nechť tedy v (2.1) platí rovnost. Potom je $g(x)$ lineární funkcí ve tvaru

$$g(x) = g(\mu) + b_\mu x - b_\mu \mu.$$

Nechť naopak $g(x) = c + dx$, $x \in (a, b)$. Potom

$$E[g(X)] = c + dE[X] = c + d\mu = g(\mu) = g(E[X]).$$

Při $d = 0$ je $g(x) = c$, $x \in (a, b)$ konstantní funkce. □

Důsledek 2.3. *$E[X]^2 \geq (E[X])^2$, tedy $\mathrm{Var}[X] \geq 0$ pro $X \in L^2(\Omega)$.*

Důkaz. Položíme $g(x) = x^2$, dále dle vět 2.8 a 2.9. □

Důsledek 2.4. *$E[\ln X] \leq \ln(E[X])$ pro $X \in L^1(\Omega)$, $\ln X \in L^1(\Omega)$, $P[X > 0] = 1$.*

Důkaz. Položíme $g(x) = -\ln x$, což je konvexní, a podle vět 2.7 a 2.9 máme

$$E[-\ln X] = -E[\ln X] \geq -\ln(E[X]) \implies \ln(E[X]) \geq E[\ln X].$$ □

Důsledek 2.5. *$(E[|X|^p])^{\frac{1}{p}} \geq (E[|X|^q])^{\frac{1}{q}}$ pro $p > q > 0$.*

Důkaz. Položíme $g(y) = y^{\frac{p}{q}}$, což je konvexní pro $p > q$, a $Y = |X|^q$. Potom dle věty 2.9

$$E[g(Y)] = E\left[|X|^{q\frac{p}{q}}\right] = E[|X|^p] \geq g(E[Y]) = (E[|X|^q])^{\frac{p}{q}}.$$

Odtud

$$(E[|X|^p])^{\frac{1}{p}} \geq (E[|X|^q])^{\frac{1}{q}}.$$ □

Důsledek 2.6. *Pokud $E[|X|^p] < \infty$, pak také $E[|X|^q] < \infty$ pro $p > q > 0$.*

Důkaz. Dle důsledku 2.5 máme

$$E[|X|^p] \geq (E[|X|^q])^{\frac{p}{q}}$$

pro $p > q > 0$. Tedy jestliže $X \in L^p(\Omega)$, potom též $X \in L^q(\Omega)$ pro $p > q > 0$. □

Věta 2.10 (Markovova nerovnost)**.** *Nechť $X \in L^p(\Omega)$. Potom pro libovolné $\varepsilon > 0$ platí*

$$P[|X| \geq \varepsilon] \leq \frac{E[|X|^p]}{\varepsilon^p}.$$

Důkaz. Počítejme:

$$E[|X|^p] = \int_{-\infty}^{\infty} |x|^p f_X(x)\,\mathrm{d}\mu(x) = \int_{-\infty}^{\infty} |x|^p\,\mathrm{d}P_X(x) \geq \varepsilon^p \int_{|x|\geq\varepsilon} 1\,\mathrm{d}P_X(x) = \varepsilon^p P_X[\{x : |x| \geq \varepsilon\}] = \varepsilon^p P[|X| \geq \varepsilon].$$

První rovnost platí dle poznámky (ii) za definicí 2.13, druhá rovnost plyne z věty 2.2, předposlední rovnost potom z důsledku 2.1. □

Důsledek 2.7 (Čebyševova nerovnost)**.** *Nechť $X \in L^2(\Omega)$. Potom pro libovolné $\varepsilon > 0$ platí*

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

Důkaz. Ve větě 2.10 zvolíme $p = 2$ a $Y = X - E[X]$. Zřejmě pak $\mathrm{Var}[X] = E\left[|X - E[X]|^2\right] = E[|Y|^2]$. □

Důsledek 2.8 (Pravidlo 2σ a 3σ)**.** *Nechť $X \in L^2(\Omega)$, $\mathrm{Var}[X] = \sigma^2$. Potom platí*

$$P[|X - E[X]| \geq 2\sigma] \leq \frac{1}{4}, \quad P[|X - E[X]| \geq 3\sigma] \leq \frac{1}{9}.$$

Důkaz. V Čebyševově nerovnosti zvolíme $\varepsilon = 2\sigma$, resp. $\varepsilon = 3\sigma$. □

Příklad 2.3. Vidíme, že σ jako míra variability souvisí s kolísáním X kolem $E[X]$ jakožto míry polohy. Například

$$P[E[X] - 3\sigma < X < E[X] + 3\sigma] = 1 - P[|X - E[X]| \geq 3\sigma] \geq \frac{8}{9}.$$ ▲

Definice 2.17 (Vytvořující funkce posloupnosti pravděpodobností)**.** Uvažujme diskrétní rozdělení s posloupností pravděpodobností $p_0, p_1, p_2, \ldots$, kde $0 < p_k < 1$ pro všechna $k = 0, 1, 2, \ldots$, $\sum_{k=0}^{\infty} p_k = 1$. Funkce

$$p(x) = \sum_{k=0}^{\infty} p_k x^k$$

se nazývá *vytvořující funkce* posloupnosti pravděpodobností.

Věta 2.11 (Výpočet střední hodnoty diskrétní náhodné veličiny). *Nechť náhodná veličina* $X \in L^1(\Omega)$ *má diskrétní rozdělení definované pravděpodobnostmi* $p_k = P[X = k]$, *kde* $k = 0, 1, 2, \ldots$ *Pak platí* $E[X] = p'(1)$.

Důkaz. Po derivaci vytvořující funkce dostaneme

$$p'(x) = \sum_{k=1}^{\infty} k p_k x^{k-1}$$

a ihned získáme

$$p'(1) = \sum_{k=0}^{\infty} k p_k = E[X]. \qquad \square$$

Poznámka. Postupným derivováním lze ukázat, že $p^{(k)}(0) = k!\, p_k$, kde $k = 1, 2, \ldots$, tedy že se vytvoří posloupnost pravděpodobností

$$\begin{aligned}
p'(0) &= \sum_{k=1}^{\infty} k p_k 0^{k-1} = p_1, \\
p''(0) &= \sum_{k=2}^{\infty} k(k-1) p_k 0^{k-2} = 2!\, p_2, \\
&\vdots \\
p^{(k)}(0) &= \sum_{j=k}^{\infty} j(j-1)\cdots(j-k+1) p_j 0^{k-j} = k(k-1)\cdots 1 p_k = k!\, p_k.
\end{aligned}$$

Pro $j = k$ v úpravách klademe $0^{k-j} = 0^0 = 1$.

Definice 2.18 (Momentová vytvořující funkce). Nechť X je reálná náhodná veličina. Funkce $M_X(t) = E\left[e^{tX}\right]$, $t \in \mathbb{R}$ se nazývá *momentová vytvořující funkce* náhodné veličiny X.

Věta 2.12 (Vlastnosti momentové vytvořující funkce). *Nechť existuje* $b > 0$ *takové, že* $M_X(t) < \infty$, $t \in (-b, b)$. *Potom platí*

(i) M_X *má spojité derivace všech řádů na* $(-b, b)$;

(ii) $M^{(k)}(0) = \mu'_k < \infty$ *pro* $k = 1, 2, \ldots$;

(iii) *pro všechna* $t \in (-b, b)$ *platí*

$$M_X(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \mu'_k,$$

kde řada konverguje absolutně;

(iv) *pokud pro všechna* $t \in (-b, b)$ *platí* $M_X(t) = M_Y(t)$, *potom platí také* $F_X(x) = F_Y(x)$ *pro všechna* $x \in \mathbb{R}$;

(v) $M_{r+sX}(t) = e^{rt} M_X(st)$ *pro* $r, s \in \mathbb{R}$.

Důkaz.

(i) - (iv) jsou vynechány.

(v) Z definice 2.18 a věty 2.7 plyne

$$M_{r+sX}(t) = E\left[e^{t(r+sX)}\right] = e^{rt} E\left[e^{stX}\right] = e^{rt} M_X(st). \qquad \square$$

Poznámka. Ve větě 2.12 body (ii) a (iii) ukazují, že pomocí M_X se vytváří posloupnost momentů, a to derivováním nebo rozvojem v řadu. Bod (iv) ukazuje, že shoda momentových vytvořujících funkcí náhodných veličin X a Y znamená shodu pravděpodobnostních rozdělení těchto náhodných veličin.

Definice 2.19 (Charakteristická funkce). Nechť X je reálná náhodná veličina. Funkce $\Psi_X(t) = E\left[e^{itX}\right]$, $t \in \mathbb{R}$ se nazývá *charakteristická funkce* náhodné veličiny X.

Poznámka. Charakteristická funkce Ψ_X (na rozdíl od M_X) jako integrál vždy existuje. Dá se pro ni dokázat analogie věty 2.12. Jak M_X (pokud existuje), tak Ψ_X jednoznačně určují pravděpodobnostní rozdělení náhodné veličiny X.

Definice 2.20 (Kovariance)**.** Pro náhodné veličiny $X, Y \in L^2(\Omega)$ definujeme *kovarianci* předpisem

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

Věta 2.13 (Vlastnosti kovariance)**.** *Nechť $X, Y, Z \in L^2(\Omega)$. Potom platí*

(i) $\operatorname{cov}(X,X) = \operatorname{Var}[X]$*;*

(ii) $\operatorname{cov}(X,Y) = \operatorname{cov}(Y,X)$*;*

(iii) $\operatorname{cov}(X,Y) = E[XY] - E[X]E[Y]$*;*

(iv) $\operatorname{cov}(a + bX, c + dY) = bd\operatorname{cov}(X,Y)$ *pro* $a, b, c, d \in \mathbb{R}$*;*

(v) $\operatorname{cov}(X \pm Y, Z) = \operatorname{cov}(X,Z) \pm \operatorname{cov}(Y,Z)$.

Důkaz.

(i) Zřejmé dle definice 2.15.

(ii) Zřejmé dle definice 2.20.

(iii) Analogicky jako v důkazu bodu (ii) věty 2.8.

(iv) Analogicky jako v důkazu bodu (i) věty 2.8.

(v) Zřejmé dle definice 2.20 a věty 2.7. □

Věta 2.14 (Rozptyl součtu a rozdílu)**.** *Nechť $X, Y \in L^2(\Omega)$. Potom platí*

$$\operatorname{Var}[X \pm Y] = \operatorname{Var}[X] + \operatorname{Var}[Y] \pm 2\operatorname{cov}(X,Y).$$

Důkaz. Pro součet lze psát

$$\begin{aligned}
\operatorname{Var}[X + Y] &= E[(X + Y) - E[X + Y]]^2 = E[(X - E[X]) + (Y - E[Y])]^2 \\
&= E(X - E[X])^2 + E(Y - E[Y])^2 + 2E[(X - E[X])(Y - E[Y])] \\
&= \operatorname{Var}[X] + \operatorname{Var}[Y] + 2\operatorname{cov}(X,Y).
\end{aligned}$$

První rovnost platí dle definice 2.15, druhá rovnost plyne z věty 2.7. Pro rozdíl je důkaz analogický. □

Poznámka. Indukcí lze pro $X_1, \ldots, X_n \in L^2(\Omega)$ dokázat, že

$$\operatorname{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \operatorname{Var}[X_i] + \sum_{\substack{i=1 \\ i \neq j}}^{n} \sum_{j=1}^{n} \operatorname{cov}(X_i, X_j) = \sum_{i=1}^{n} \sum_{j=1}^{n} \operatorname{cov}(X_i, X_j).$$

Definice 2.21 (Korelační koeficient)**.** Nechť $X, Y \in L^2(\Omega)$, $\operatorname{Var}[X] > 0$, $\operatorname{Var}[Y] > 0$. Potom *korelační koeficient* definujeme předpisem

$$\operatorname{cor}(X,Y) = \rho_{X,Y} = \frac{\operatorname{cov}(X,Y)}{\sqrt{\operatorname{Var}[X]\operatorname{Var}[Y]}}.$$

Cvičení. Uvažujme $X, Y \in L^2(\Omega)$ takové, že $\operatorname{Var}[X] > 0$, $\operatorname{Var}[Y] > 0$.

(i) Dokažte, že

$$\operatorname{Var}\left[\frac{X - E[X]}{\sqrt{\operatorname{Var}[X]}}\right] = 1.$$

(ii) Dokažte, že

$$E\left[\frac{X - E[X]}{\sqrt{\operatorname{Var}[X]}}\right] = 0.$$

(iii) Dokažte, že

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

Řešení.

(i) Uvědomíme si, že $E[X] < \infty$ a $\operatorname{Var}[X] > 0$ jsou reálná čísla. Použijeme tedy bod (i) z věty 2.8 a dostaneme

$$\operatorname{Var}\left[\frac{X - E[X]}{\sqrt{\operatorname{Var}[X]}}\right] = \frac{1}{\operatorname{Var}[X]}\operatorname{Var}[X - E[X]] = \frac{1}{\operatorname{Var}[X]}\operatorname{Var}[X] = 1.$$

(ii) Podle bodu (i) věty 2.7 máme

$$E\left[\frac{X - E[X]}{\sqrt{\operatorname{Var}[X]}}\right] = \frac{1}{\sqrt{\operatorname{Var}[X]}}(E[X] - E[X]) = \frac{0}{\sqrt{\operatorname{Var}[X]}} = 0.$$

(iii) K důkazu první rovnosti využijeme definici kovariance a vlastnosti střední hodnoty a získáme

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

Ve druhé rovnosti uplatníme bod (iii) věty 2.13 a také bod (ii) tohoto cvičení a dostaneme

$$\begin{aligned}\operatorname{cov}\left(\frac{X - E[X]}{\sqrt{\operatorname{Var}[X]}}, \frac{Y - E[Y]}{\sqrt{\operatorname{Var}[Y]}}\right) &= E\left[\frac{X - E[X]}{\sqrt{\operatorname{Var}[X]}}\cdot\frac{Y - E[Y]}{\sqrt{\operatorname{Var}[Y]}}\right] - E\left[\frac{X - E[X]}{\sqrt{\operatorname{Var}[X]}}\right]E\left[\frac{Y - E[Y]}{\sqrt{\operatorname{Var}[Y]}}\right] \\ &= E\left[\frac{(X - E[X])(Y - E[Y])}{\sqrt{\operatorname{Var}[X]\operatorname{Var}[Y]}}\right].\end{aligned}$$

▲

Věta 2.15 (Vlastnosti korelačního koeficientu)**.** *Nechť* $X, Y \in L^2(\Omega)$, $\operatorname{Var}[X] > 0$, $\operatorname{Var}[Y] > 0$. *Potom platí*

(i) $\rho_{X,X} = 1$;

(ii) $\rho_{a+bX,c+dY} = \operatorname{sgn}(bd)\rho_{X,Y}$ *pro* $a, b, c, d \in \mathbb{R}$ *a* $bd \neq 0$;

(iii) $-1 \leq \rho_{X,Y} \leq 1$;

(iv) $\rho_{X,Y} = \pm 1$ *právě tehdy, když* $Y = a \pm bX$, *kde* $b > 0$.

Důkaz.

(i) Zřejmé dle definice 2.21 a bodu (i) věty 2.13.

(ii) Za použití věty 2.13 a věty 2.8 počítejme:

$$\rho_{a+bX,c+dY} = \frac{bd\operatorname{cov}(X,Y)}{\sqrt{b^2\operatorname{Var}[X]d^2\operatorname{Var}[Y]}} = \frac{bd}{\sqrt{(bd)^2}}\rho_{X,Y} = \frac{bd}{|bd|}\rho_{X,Y} = \operatorname{sgn}(bd)\rho_{X,Y}.$$

(iii) Použijeme Cauchyho-Schwarzovu nerovnost, podle které

$$(E[(X - E[X])(Y - E[Y])])^2 \leq E[X - E[X]]^2 E[Y - E[Y]]^2.$$

Z toho plyne, že platí také

$$|\operatorname{cov}(X,Y)| \leq \sqrt{\operatorname{Var}[X]\operatorname{Var}[Y]},$$

a to je ekvivalentní s tím, že $|\rho_{X,Y}| \leq 1$. Je navíc podle důsledku 2.6 zřejmé, že pokud $E[X]^2 < \infty$, tak také $E[X] < \infty$.

(iv) Uvažujme opět Cauchyho-Schwarzovu nerovnost

$$(E[(X - E[X])(Y - E[Y])])^2 \leq E[X - E[X]]^2 E[Y - E[Y]]^2.$$

Rovnost zde nastává právě tehdy, když

$$Y - E[Y] = b(X - E[X]),$$

neboli právě tehdy, když

$$Y = E[Y] - \ bE[X] + bX.$$

Označme $a = E[Y] - bE[X]$. Potom platí, že

$$\rho_{X,a+bX} = \operatorname{sgn}(b)\rho_{X,X} = \begin{cases} 1; & b > 0, \\ -1; & b < 0 \end{cases}$$

podle bodu (i) a (ii). □

Poznámka. Vlastnost (ii) znamená, že korelační koeficient nezávisí na jednotkách, ve kterých X a Y měříme. Zároveň z vlastnosti (iii) plyne, že pokud $X, Y \in L^2(\Omega)$, tak $\mathrm{Var}[X] < \infty$, $\mathrm{Var}[Y] < \infty$, a proto také $\mathrm{cov}(X, Y) < \ \infty$.

Definice 2.22 (Nekorelované náhodné veličiny)**.** Nechť $X, Y \in L^2(\Omega)$, $\mathrm{Var}[X] > 0$, $\mathrm{Var}[Y] > 0$. Je-li $\rho_{X,Y} = 0$, označujeme náhodné veličiny X, Y jako *nekorelované.*

Poznámka. Korelační koeficient měří sílu lineární závislosti mezi veličinami X a Y. Nekorelovanost znamená, že buď mezi X a Y není žádná závislost, nebo je mezi nimi závislost nelineární.

2.3. Nezávislost náhodných veličin

Definice 2.23 (Nezávislé náhodné veličiny)**.** Reálné náhodné veličiny $X_1, X_2, \dots$ nazveme *nezávislé*, pokud jsou pro libovolné množiny $B_1, B_2, \dots \in \mathcal{B}_0$ náhodné jevy $X_i \in B_i$ nezávislé. Pro pevné $n \in \mathbb{N}$ řekneme, že jsou náhodné veličiny $X_1, \dots, X_n$ nezávislé, pokud pro $B_1, \dots, B_n \in \mathcal{B}_0$ platí, že

$$P\left[\bigcap_{i=1}^{n}(X_i \in B_i)\right] = \prod_{i=1}^{n} P[X_i \in B_i].$$

Definice 2.24 (Sdružená a marginální distribuční funkce)**.** Funkce $F(x_1, \dots, x_n) = P[X_1 \leq x_1, \dots, X_n \leq \ x_n]$ se nazývá *sdružená distribuční funkce* náhodných veličin $X_1, \dots, X_n$. Distribuční funkce $F_{X_1}(x), \dots, F_{X_n}(x)$ jednotlivých veličin nazýváme *marginální distribuční funkce.*

Věta 2.16 (Nezávislost náhodných veličin pomocí distribučních funkcí)**.** *Náhodné veličiny $X_1, \dots, X_n$ jsou nezávislé právě tehdy, když platí*

$$F(x_1, \dots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n)$$

pro všechna $x_1, \dots, x_n \in \mathbb{R}$.

Důkaz. Pro ověření implikace zleva doprava stačí zvolit $B_i = (-\infty, x_i]$ v definici 2.23. Implikace zprava doleva pak plyne z toho, že distribuční funkce jednoznačně určuje pravděpodobnostní rozdělení (ve smyslu poznámky (ii) za definicí 2.10). □

Definice 2.25 (Sdružené a marginální pravděpodobnosti)**.** Nechť $X_1, \dots, X_n$ jsou diskrétní náhodné veličiny s hodnotami $x_i \in S_i$, kde $i = 1, \dots, n$. Potom $P[X_1 = x_1, \dots, X_n = x_n]$ nazýváme *sdružené pravděpodobnosti* a $P[X_i = x_i]$ nazýváme *marginální pravděpodobnosti.*

Poznámky.

(i) Pro sdruženou pravděpodobnost platí

$$0 < P[X_1 = x_1, \dots, X_n = x_n] < 1,$$

kde $x_i \in S_i$ pro $i = 1, \dots, n$, a také

$$\sum_{x_1 \in S_1} \cdots \sum_{x_n \in S_n} P[X_1 = x_1, \dots, X_n = x_n] = 1.$$

(ii) Distribuční funkce je ve tvaru

$$F(x_1,\dots,x_n)=\sum_{\substack{k_1\in S_1\\ k_1\le x_1}}\cdots\sum_{\substack{k_n\in S_n\\ k_n\le x_n}}P[X_1=k_1,\dots,X_n=k_n].$$

(iii) Uvažujme $B_i\in\mathcal{B}_0$, $B_i\subset S_i$, kde $i=1,\dots,n$. Potom

$$P[X_1\in B_1,\dots,X_n\in B_n]=\sum_{x_1\in B_1}\cdots\sum_{x_n\in B_n}P[X_1=x_1,\dots,X_n=x_n].$$

(iv) Uvažujme funkci $g:\mathbb{R}^n\longrightarrow\mathbb{R}$ a $B\in\mathcal{B}_0$. Pak lze psát

$$P[g(X_1,\dots,X_n)\in B]=\sum_{x_1}\cdots\sum_{x_n}P[X_1=x_1,\dots,X_n=x_n],$$

kde pro $i=1,\dots,n$ sčítáme přes x_i z S_i taková, že $g(x_1,\dots,x_n)\in B$.

(v) Uvažujme funkci $g:\mathbb{R}^n\longrightarrow\mathbb{R}$ jako v předchozím případě. Potom platí, že

$$E[g(X_1,\dots,X_n)]=\sum_{x_1\in S_1}\cdots\sum_{x_n\in S_n}g(x_1,\dots,x_n)P[X_1=x_1,\dots,X_n=x_n].$$

Věta 2.17 (Nezávislost diskrétních veličin)**.** *Nechť X,Y jsou diskrétní náhodné veličiny, $P[X=x_i]=p_i$ pro $i=1,2,\dots$ a $P[Y=y_j]=q_j$ pro $j=1,2,\dots$ Veličiny X a Y jsou nezávislé právě tehdy, když platí*

$$P[X=x_i,Y=y_j]=p_iq_j$$

pro všechna uvažovaná i,j.

Důkaz. Pro ověření implikace zleva doprava stačí zvolit $B_1=x_i$, $B_2=y_j$ v definici 2.23. Pro důkaz opačné implikace uvažujme $B_1,B_2\in\mathcal{B}_0$. Pak lze psát

$$P[X\in B_1,Y\in B_2]=\sum_{x_i\in B_1}\sum_{y_j\in B_2}P[X=x_i,Y=y_j]=\sum_{x_i\in B_1}p_i\sum_{y_j\in B_2}q_j=P[X\in B_1]P[Y\in B_2].$$

Druhá rovnost plyne z předpokladu. □

Poznámka. Indukcí lze rozšířit na n diskrétních náhodných veličin, takže potom $X_1,\dots,X_n$ jsou nezávislé právě tehdy, když $P[\bigcap_{i=1}^n(X_i=x_i)]=\prod_{i=1}^n P[X_i=x_i]$ pro všechna $x_i\in S_i$, $i=1,\dots,n$.

Definice 2.26 (Sdružená a marginální hustota)**.** Nechť $X_1,\dots,X_n$ jsou spojité náhodné veličiny se sdruženou distribuční funkcí $F(x_1,\dots,x_n)$. Funkce

$$f(x_1,\dots,x_n)=\frac{\partial^n F(x_1,\dots,x_n)}{\partial x_1\cdots\partial x_n}$$

se nazývá *sdružená hustota* veličin $X_1,\dots,X_n$. Hustoty $f_{X_1}(x),\dots,f_{X_n}(x)$ jednotlivých veličin se nazývají *marginální hustoty.*

Poznámky.

(i) Uvažujme $x_i\in\mathbb{R}$ pro $i=1,\dots,n$ a funkci sdružené hustoty $f(x_1,\dots,x_n)\ge 0$. Potom platí

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}f(x_1,\dots,x_n)\,\mathrm{d}x_1\cdots\mathrm{d}x_n=1.$$

(ii) Uvažujme stejné předpoklady jako v předchozí poznámce. Pak platí

$$F(x_1,\dots,x_n)=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n}f(u_1,\dots,u_n)\,\mathrm{d}u_1\cdots\mathrm{d}u_n.$$

(iii) Uvažujme $B_i\in\mathcal{B}_0$, kde $i=1,\dots,n$. Potom

$$P[X_1\in B_1,\dots,X_n\in B_n]=\int_{B_1}\cdots\int_{B_n}f(x_1,\dots,x_n)\,\mathrm{d}x_1\cdots\mathrm{d}x_n.$$

(iv) Uvažujme funkci $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ a $B \in \mathcal{B}_0$. Pak lze psát

$$P[g(X_1, \ldots, X_n) \in B] = \int_{x_1} \cdots \int_{x_n} f(x_1, \ldots, x_n)\, \mathrm{d}x_1 \cdots \mathrm{d}x_n,$$

kde pro $i = 1, \ldots, n$ integrujeme přes $x_i \in \mathbb{R}$ taková, že $g(x_1, \ldots, x_n) \in B$.

(v) Uvažujme funkci $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ jako v předchozím případě. Potom platí, že

$$E[g(X_1, \ldots, X_n)] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, \ldots, x_n) f(x_1, \ldots, x_n)\, \mathrm{d}x_1 \cdots \mathrm{d}x_n.$$

Věta 2.18 (Nezávislost spojitých veličin)**.** *Nechť X, Y jsou spojité náhodné veličiny s hustotami $f_X(x)$, $f_Y(y)$ a se sdruženou hustotou $f_{X,Y}(x, y)$. Potom X a Y jsou nezávislé právě tehdy, když pro všechna $x, y \in \mathbb{R}$ platí*

$$f_{X,Y}(x, y) = f_X(x) f_Y(y).$$

Důkaz. Nejdříve dokážeme implikaci zleva doprava pomocí definice 2.26 a věty 2.16:

$$f_{X,Y}(x, y) = \frac{\partial^2 F_{X,Y}(x, y)}{\partial x\, \partial y} = \frac{\partial^2 [F_X(x) F_Y(y)]}{\partial x\, \partial y} = F'_X(x) F'_Y(y) = f_X(x) f_Y(y).$$

V důkazu opačné implikace si pomůžeme druhou poznámkou za definicí 2.26:

$$F_{X,Y}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u, v)\, \mathrm{d}u\, \mathrm{d}v = \int_{-\infty}^{x} f_X(u)\, \mathrm{d}u \int_{-\infty}^{y} f_Y(v)\, \mathrm{d}v = F_X(x) F_Y(y). \qquad \square$$

Poznámky.

(i) Indukcí lze rozšířit na n spojitých náhodných veličin, takže potom $X_1, \ldots, X_n$ jsou nezávislé právě tehdy, když $f(x_1, \ldots, x_n) = f_{X_1}(x_1) \ldots f_{X_n}(x_n)$ pro všechna $x_i \in \mathbb{R}$, $i = 1, \ldots, n$.

(ii) Termín marginální se používá i pro libovolnou podmnožinu $X_{i_1}, \ldots, X_{i_k}$ množiny $X_1, \ldots X_n$ pro $1 \leq k < n$.

Věta 2.19 (Nezávislost a střední hodnota součinu)**.** *Nechť $X_1, \ldots, X_n$ jsou nezávislé náhodné veličiny. Pak platí*

$$E\left[\prod_{i=1}^{n} X_i\right] = \prod_{i=1}^{n} E\left[X_i\right].$$

Důkaz. Zřejmě z poznámek za definicemi 2.26, 2.25 a 2.10 plyne, že

$$E\left[\prod_{i=1}^{n} X_i\right] = \int_{\mathbb{R}^n} x_1 \cdots x_n\, \mathrm{d}F(x_1, \ldots, x_n).$$

Nyní dle věty 2.16 a Fubiniovy věty můžeme psát

$$\begin{aligned}
\int_{\mathbb{R}^n} x_1 \cdots x_n\, \mathrm{d}F(x_1, \ldots, x_n) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1 \cdots x_n\, \mathrm{d}F_{X_1}(x_1) \cdots \mathrm{d}F_{X_n}(x_n) \\
&= \int_{-\infty}^{\infty} x_1\, \mathrm{d}F_{X_1}(x_1) \cdots \int_{-\infty}^{\infty} x_n\, \mathrm{d}F_{X_n}(x_n) \\
&= E[X_1] \cdots E[X_n] = \prod_{i=1}^{n} E\left[x_i\right].
\end{aligned} \qquad \square$$

Důsledek 2.9.

(i) Pokud jsou náhodné veličiny $X, Y \in L^2(\Omega)$ nezávislé, potom z bodu (iii) věty 2.13 plyne

$$\mathrm{cov}(X, Y) = E[XY] - E[X]E[Y] = 0.$$

(ii) Pokud jsou náhodné veličiny $X, Y \in L^2(\Omega)$ nezávislé, potom z věty 2.14 plyne

$$\mathrm{Var}[X \pm Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] \pm 2\,\mathrm{cov}(X, Y) = \mathrm{Var}[X] + \mathrm{Var}[Y].$$

(iii) Pokud jsou náhodné veličiny $X_1, \dots, X_n \in L^2(\Omega)$ po dvou nezávislé, pak

$$\operatorname{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \operatorname{Var}[X_i].$$

Tento důsledek plyne z poznámky za větou 2.14 a platí, i pokud jsou uvažované náhodné veličiny pouze po dvou nekorelované.

(iv) Pokud jsou náhodné veličiny $X, Y \in L^2(\Omega)$ nezávislé, potom jsou nekorelované. Tento důsledek plyne z definic 2.21 a 2.22. Opačná implikace neplatí.

Věta 2.20 (Nezávislost funkcí náhodných veličin)**.** *Nechť X, Y jsou nezávislé náhodné veličiny a h, g jsou reálné měřitelné funkce. Pak jsou $h(X)$, $g(Y)$ také nezávislé.*

Důkaz. Nechť $B_1, B_2 \in \mathcal{B}_0$. Z měřitelnosti h a g lze na základě definice 2.2 psát

$$P[h(x) \in B_1, g(Y) \in B_2] = P[X \in h^{-1}(B_1), Y \in g^{-1}(B_2)],$$

kde $h^{-1}(B_1)$, $g^{-1}(B_2) \in \mathcal{B}_0$. Dále z nezávislosti X a Y a z definice 2.23 máme

$$P[X \in h^{-1}(B_1), Y \in g^{-1}(B_2)] = P[X \in h^{-1}(B_1)]P[Y \in g^{-1}(B_2)] = P[h(X) \in B_1]P[g(Y) \in B_2],$$

což dokazuje nezávislost $h(X)$ a $g(Y)$. □

Poznámka. Věta 2.20 platí také pro komplexní funkce.

Věta 2.21 (Nezávislost a momentová vytvořující funkce)**.** *Nechť náhodné veličiny X, Y mají momentové vytvořující funkce $M_X(t), M_Y(t)$ a označme*

$$M_{X,Y}(t_1, t_2) = E[e^{t_1 X + t_2 Y}],$$

kde $t, t_1, t_2 \in \mathbb{R}$. Pak X a Y jsou nezávislé právě tehdy, když $M_{X,Y}(t_1, t_2) = M_X(t_1)M_Y(t_2)$.

Důkaz. K ověření implikace zleva doprava využijeme věty 2.20 a 2.19. Protože X, Y jsou nezávislé, tak e^{tX} a e^{tY} jsou podle věty 2.20 také nezávislé. Potom však dle věty 2.19 platí, že

$$E\left[e^{t_1 X} e^{t_2 X}\right] = E\left[e^{t_1 X}\right] E\left[e^{t_2 Y}\right] = M_X(t_1)M_Y(t_2) = M_{X,Y}(t_1, t_2).$$ □

Poznámka. Indukcí lze větu rozšířit na n náhodných veličin. Analogické tvrzení platí i pro charakteristické funkce (viz definice 2.19).

Věta 2.22 (Momentová vytvořující funkce součtu nezávislých veličin)**.** *Nechť $X_1, \dots, X_n$ jsou nezávislé náhodné veličiny a $S_n = \sum_{i=1}^{n} X_i$. Potom pro $t \in \mathbb{R}$ platí*

$$M_{S_n}(t) = \prod_{i=1}^{n} M_{X_i}(t).$$

Důkaz. Počítejme:

$$M_{S_n}(t) = E[e^{tS_n}] = E\left[\prod_{i=1}^{n} e^{tX_i}\right] = \prod_{i=1}^{n} E\left[e^{tX_i}\right] = \prod_{i=1}^{n} M_{X_i}(t).$$

Třetí rovnost plyne z vět 2.20 a 2.19. □

Poznámka. Analogické tvrzení platí i pro charakteristické funkce.

2.4. Přehled pravděpodobnostních rozdělení

V této sekci si uvedeme několik příkladů s charakterizacemi diskrétních a spojitých pravděpodobnostních rozdělení. Začneme rozděleními *diskrétními*.

1. Rovnoměrné diskrétní rozdělení

$$\begin{aligned} P[X=k] &= \frac{1}{n} \text{ pro } n \in \mathbb{N} \text{ a } k = 1, \dots, n; \\ E[X] &= \frac{1}{n}\sum_{k=1}^{n} k = \frac{1}{n}\frac{n}{2}(1+n) = \frac{n+1}{2}; \\ E[X]^2 &= \frac{1}{n}\sum_{k=1}^{n} k^2 = \frac{1}{n}\frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}; \\ \mathrm{Var}[X] &= E[X]^2 - (E[X])^2 = \frac{n^2-1}{12}. \end{aligned}$$

   Použití: modelování náhodných pokusů s konečně mnoha stejně pravděpodobnými výsledky (např. hod kostkou).

2. Alternativní rozdělení $X \sim \mathrm{Alt}(p)$ s parametrem p (Bernoulliho rozdělení, $0 < p < 1$)

$$\begin{aligned} P[X=1] &= p; \\ P[X=0] &= 1-p; \\ E[X] &= 1p + 0(1-p) = p; \\ E[X]^2 &= 1^2 p + 0^2(1-p) = p; \\ \mathrm{Var}[X] &= E[X]^2 - (E[X])^2 = p(1-p). \end{aligned}$$

   Použití: modelování náhodných pokusů se dvěma výsledky (např. hod mincí).

3. Binomické rozdělení $X \sim \mathrm{Bi}(n,p)$ s parametry n, p

$$\begin{aligned} X &= \sum_{i=1}^{n} X_i \text{ pro } X_i \text{ nezávislé s alternativním rozdělením}; \\ P[X=k] &= \binom{n}{k} p^k (1-p)^{n-k} \text{ pro } k = 0, 1, \dots, n \text{ a } 0 < p < 1; \\ E[X] &= np; \\ \mathrm{Var}[X] &= np(1-p). \end{aligned}$$

   Vztah pro střední hodnotu plyne z poznámky (v) za větou 2.7, vztah pro rozptyl potom z definice 2.23 a bodu (iii) důsledku 2.9.
   Použití: modelování počtu úspěchů v sérii n nezávislých pokusů, kde každý pokus má 2 možné výsledky, a to úspěch (nastane jev A, 1), nebo neúspěch (nenastane jev A, 0). Jinými slovy se jedná modelování četnosti jevu A v n nezávislých pokusech, kterou může být například počet infikovaných lidí při vyšetření n pacientů.

4. Poissonovo rozdělení $X \sim \mathrm{Po}(\lambda)$ s parametrem λ
   Zde si příslušné vztahy odvodíme. Nechť $n \longrightarrow \infty$, $p \longrightarrow 0$ takové, že $np \longrightarrow \lambda$. Potom za pomoci limitního

přechodu

$$\lim_{n\longrightarrow\infty}\binom{n}{k}p^k(1-p)^{n-k} = \lim_{n\longrightarrow\infty}\frac{n(n-1)\cdots(n-k+1)}{k!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}$$
$$= \frac{\lambda^k}{k!}\lim_{n\longrightarrow\infty}\frac{n(n-1)\cdots(n-k+1)}{n\cdots n}\left(1-\frac{\lambda}{n}\right)^{-k}\left(1-\frac{\lambda}{n}\right)^{n}$$
$$= \frac{\lambda^k}{k!}\lim_{n\longrightarrow\infty}\left(1-\frac{\lambda}{n}\right)^{n} = \frac{\lambda^k}{k!}e^{-\lambda}.$$

Definujeme tedy

$$P[X=k] = \frac{\lambda^k}{k!}e^{-\lambda} \text{ pro všechna } k=0,1,\dots \text{ a } \lambda>0.$$

Zaměřme se nyní na střední hodnotu, kterou získáme pomocí vytvořující funkce posloupnosti pravděpodobností (definice 2.17 a věta 2.11). Mějme

$$p(x) = \sum_{k=0}^{\infty}\frac{(\lambda x)^k}{k!}e^{-\lambda} = e^{\lambda x}e^{-\lambda} = e^{\lambda(x-1)}.$$

Zderivujme

$$p'(x) = e^{\lambda(x-1)}\lambda$$

a dosaďme jedničku

$$p'(1) = \lambda = E[X].$$

Nakonec, protože platí

$$E[X(X-1)] = \sum_{k=2}^{\infty}k(k-1)\frac{\lambda^k}{k!}e^{-\lambda} = \lambda^2\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!}e^{-\lambda} = \lambda^2,$$

lze snadno spočítat rozptyl jako

$$\mathrm{Var}[X] = E[X]^2 - (E[X])^2 = E[X(X-1)] + E[X] - (E[X])^2 = \lambda.$$

Použití: aproximace binomického rozdělení, nebo modelování počtu událostí v daném časovém, délkovém nebo jiném úseku (např. příchody hovorů do call centra, poruchy stroje, kazy v pásu látky aj.).

5. Negativně binomické rozdělení s parametry r, p

$$P[X=k] = \binom{k+r-1}{r-1}p^r(1-p)^k \text{ pro } 0<p<1 \text{ a } k=0,1,\dots$$

Použití: modelování počtu neúspěchů před r-tým úspěchem v sérii nezávislých pokusů. Počet pokusů předem nespecifikujeme, čekáme na nastání r-tého úspěchu (libovolně dlouho). Speciálně, pro $r=1$ dostáváme geometrické rozdělení.

6. Geometrické rozdělení s parametrem p

$$P[X=k] = p(1-p)^k \text{ pro } 0<p<1 \text{ a } k=0,1,\dots$$

7. Hypergeometrické rozdělení

$$P[X=k] = \frac{\binom{A}{k}\binom{N-A}{n-k}}{\binom{N}{n}}, \text{ kde } \max(0,\, A+n-N) \le k \le \min(A,n).$$

Použití: modelování počtu předmětů se sledovanou vlastností ve výběru o rozsahu n ze základního souboru N předmětů, z nichž A má sledovanou vlastnost (např. vadné výrobky při kontrole kvality).

Ve druhé části této sekce se zaměříme na *spojitá* rozdělení, mezi která mimo jiné patří také normální a normované normální rozdělení.

1. <u>Rovnoměrné rozdělení</u> $X \sim R(a,b)$ na intervalu (a,b)
   Zde platí, že
$$f(x) = \frac{1}{b-a}$$
   pro $a < x < b$ a $f(x) = 0$ jinak. Pak tedy
$$\begin{aligned}
F(x) &= \int_a^x \frac{1}{b-a}\,\mathrm{d}u = \frac{x-a}{b-a} \text{ pro } a < x < b;\\
F(x) &= 0 \text{ pro } x \leq a;\\
F(x) &= 1 \text{ pro } x \geq b;\\
E[X] &= \int_a^b x\frac{1}{b-a}\,\mathrm{d}x = \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2};\\
E[X]^2 &= \int_a^b x^2\frac{1}{b-a}\,\mathrm{d}x = \frac{b^3-a^3}{3(b-a)} = \frac{a^2+ab+b^2}{3};\\
\mathrm{Var}[X] &= E[X]^2 - (E[X])^2 = \frac{(b-a)^2}{12}.
\end{aligned}$$
   Speciálně, pro $R(0,1)$ máme
$$\begin{aligned}
f(x) &= 1 \text{ pro } x \in (0,1);\\
F(x) &= x \text{ pro } x \in (0,1);\\
E[X] &= \frac{1}{2};\\
\mathrm{Var}[X] &= \frac{1}{12}.
\end{aligned}$$
   Použití: modelování doby nastání události, která může nastat kdykoli během daného časového intervalu, případně doby čekání na tuto událost od začátku časového intervalu.

   Na tomto místě si uvedeme větu vztahující se ke speciálnímu případu rovnoměrného rozdělení.

   Věta 2.23 (Metoda inverzní transformace)**.** *Nechť $X \sim R(0,1)$ a nechť F je spojitá distribuční funkce. Potom má náhodná veličina $Y = F^{-1}(X)$ rozdělení s distribuční funkcí F.*

   Důkaz. Lze jednoduše spočítat, že $P[Y \leq y] = P[F^{-1}(X) \leq y] = P[X \leq F(y)] = F(y)$. □

   Na základě této věty lze generovat posloupnosti nezávislých realizací ze spojitého rozdělení s danou distribuční funkcí F. Softwarové produkty mají zabudované generátory náhodných čísel z $R(0,1)$. Jejich opakovaným voláním lze získat posloupnost nezávislých realizací $x_1, \ldots x_n$ z rozdělení $R(0,1)$ a inverzní transformací je možno je převést na nezávislé realizace $y_i = F^{-1}(x_i)$, $i = 1, \ldots, n$.

2. <u>Exponenciální rozdělení</u> $X \sim \mathrm{Exp}(\lambda)$ s parametrem λ
$$\begin{aligned}
f(x) &= \frac{1}{\lambda}e^{-\frac{x}{\lambda}} \text{ pro } x > 0,\ \lambda > 0;\\
F(x) &= \int_0^x \frac{1}{\lambda}e^{-\frac{u}{\lambda}}\,\mathrm{d}u = \frac{1}{\lambda}\left[-\lambda e^{-\frac{u}{\lambda}}\right]_0^x = 1 - e^{-\frac{x}{\lambda}} \text{ pro } x > 0;\\
E[X] &= \int_0^\infty \frac{x}{\lambda}e^{-\frac{x}{\lambda}}\,\mathrm{d}x = \left[-xe^{-\frac{x}{\lambda}}\right]_0^\infty + \int_0^\infty e^{-\frac{x}{\lambda}}\,\mathrm{d}x = \lambda;\\
E[X]^2 &= \int_0^\infty \frac{x^2}{\lambda}e^{-\frac{x}{\lambda}}\,\mathrm{d}x = 2\int_0^\infty xe^{-\frac{x}{\lambda}}\,\mathrm{d}x = 2\lambda E[X] = 2\lambda^2;\\
\mathrm{Var}[X] &= E[X]^2 - (E[X])^2 = \lambda^2.
\end{aligned}$$

Někdy se hustota uvádí ve tvaru $f(x) = ae^{ax}$, kde $x > 0$, $a > 0$. Pak je zřejmě

$$E[X] = \frac{1}{a}, \quad \mathrm{Var}[X] = \frac{1}{a^2}.$$

O exponenciálním rozdělení se říká, že nemá paměť. Pro $x > 0$, $y > 0$ máme

$$P[X > x + y \mid X > y] = \frac{P[X > x + y,\, X > y]}{P[X > y]} = \frac{P[X > x + y]}{P[X > y]} = \frac{\exp[-\frac{x+y}{\lambda}]}{\exp[-\frac{y}{\lambda}]} = e^{-\frac{x}{\lambda}} = P[X > x].$$

Použití: modelování doby mezi dvěma událostmi v posloupnosti událostí téhož typu (např. poruchy stroje).

3. <u>Normované normální rozdělení</u> $Z \sim \mathrm{N}(0,1)$

$$\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2}\right] \text{ pro } -\infty < z < \infty;$$

$$\Phi(z) = \int_{-\infty}^{z} \phi(u)\,\mathrm{d}u.$$

Funkci $\Phi(z)$ nelze vyjádřit pomocí elementárních funkcí, je tabelována a počítá ji statistický software. Platí pro ni, že $\Phi(-z) = 1 - \Phi(z)$, neboť $\phi(z)$ je symetrická kolem nuly.

Dále z lichosti funkce $z\phi(z)$ a symetrie $\phi(z)$ a věty 2.7 plyne

$$E[Z] = \int_{-\infty}^{\infty} z\phi(z)\,\mathrm{d}z = 0.$$

Zřejmě pokud je k liché, je $E[Z]^k = 0$ a šikmost $\gamma_3 = 0$. Pro rozptyl navíc platí

$$\begin{aligned}\mathrm{Var}[Z] = E[Z]^2 &= \int_{-\infty}^{\infty} z^2\phi(z)\,\mathrm{d}z = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 \exp\left[-\frac{z^2}{2}\right] \mathrm{d}z \\ &= \frac{1}{\sqrt{2\pi}} \left(\left[-z\exp\left[-\frac{z^2}{2}\right]\right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \exp\left[-\frac{z^2}{2}\right] \mathrm{d}z \right) \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2}\right] \mathrm{d}z = 1.\end{aligned}$$

4. <u>Normální rozdělení</u> $X \sim \mathrm{N}(\mu, \sigma^2)$ s parametry μ, σ^2
Položme $X = \mu + \sigma Z$, $Z \sim \mathrm{N}(0,1)$, $\mu \in \mathbb{R}$, $\sigma > 0$. Potom máme

$$F_X(x) = P[X \le x] = P[\mu + \sigma Z \le x] = P\left[Z \le \frac{x-\mu}{\sigma}\right] = \Phi\left(\frac{x-\mu}{\sigma}\right) \text{ pro } x \in \mathbb{R};$$

$$f_X(x) = F_X'(x) = \phi\left(\frac{x-\mu}{\sigma}\right)\frac{1}{\sigma} = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \text{ pro } x \in \mathbb{R};$$

$$E[X] = \mu + \sigma E[Z] = \mu \text{ podle věty 2.7;}$$

$$\mathrm{Var}[X] = \sigma^2 \,\mathrm{Var}[Z] = \sigma^2 \text{ podle věty 2.8;}$$

$$\mu_{2k-1} = E[X - E[X]]^{2k-1} = 0 \text{ pro } k = 1, 2, \ldots, \text{ odtud dle definice 2.16 pak } \gamma_3 = 0;$$

$$\mu_{2k} = \frac{(2k)!}{k!}\frac{\sigma^{2k}}{2^k} \text{ pro } k = 1, 2, \ldots, \text{ odtud dle definice 2.16 pak } \gamma_4 = \frac{4!}{2!}\frac{\sigma^4}{2^2}\frac{1}{\sigma^4} = 3;$$

$$u_X(0.5) = \mu, \text{ tedy vzhledem k symetrii hustoty se medián rovná střední hodnotě.}$$

Pro momentovou vytvořující funkci a charakteristickou funkci navíc platí, že

$$M_X(t) = \exp\left[t\mu + \frac{1}{2}t^2\sigma^2\right] = E\left[e^{tX}\right];$$

$$\Psi_X(t) = \exp\left[it\mu - \frac{1}{2}t^2\sigma^2\right] = E\left[e^{itX}\right].$$

Použití: měření fyzikálních veličin v podobě $\hat{Y} = Y + e,\ e \sim \mathrm{N}(0, \sigma^2)$, kde $\hat{Y}$ je naměřená hodnota, Y je skutečná hodnota a e je náhodná chyba měření. Dále také v lineárním regresním modelu

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k + e,\ e \sim \mathrm{N}(0, \sigma^2)$$

nebo autoregresním modelu pro časovou řadu $AR(p)$

$$X_t = \varphi_0 + \varphi_1 X_{t-1} + \cdots + \varphi_p X_{t-p} + e_t,\ e_t \sim \mathrm{N}(0, \sigma^2).$$

Normální rozdělení se také nazývá Gaussovo rozdělení, hustotě se říká Gaussova křivka. Normální rozdělení vykazují četné znaky v přírodě a v lidské populaci.

5. Logaritmicko-normální rozdělení
   Nechť $X \sim \mathrm{N}(\mu, \sigma^2)$ a položme $Y = e^x$. Potom

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma y} \exp\left[-\frac{(\log y - \mu)^2}{2\sigma^2}\right],\ y > 0,\ \mu \in R,\ \sigma^2 > 0.$$

   Použití: experimentálně bylo zjištěno, že intenzita úroku $\delta \sim \mathrm{N}(\mu, \sigma^2)$, a proto úrokovací faktor $1 + i = e^\delta$ má logaritmicko-normální rozdělení.

Cvičení.

(i) Odvoďte momentovou vytvořující funkci normovaného normálního rozdělení.

(ii) Odvoďte pomocí momentové vytvořující funkce střední hodnotu a rozptyl $\mathrm{N}(0, 1)$.

(iii) Odvoďte momentovou vytvořující funkci obecného normálního rozdělení.

Řešení.
O momentové vytvořující funkci náhodné veličiny X víme, že

- je definována jako $M_X(t) = E[e^{tX}]$ pro $t \in \mathbb{R}$ dle definice 2.18;
- k-tá derivace $M_X(t)$ v bodě nula je rovna $E[X]^k$ dle bodu (ii) věty 2.12;
- existuje-li rozvoj $M_X(t)$ v řadu, má tvar

$$\sum_{k=0}^{\infty} \frac{t^k}{k!} E[X]^k$$

  dle bodu (iii) věty 2.12;
- pro náhodnou veličinu $Y = r + sX$, kde $r,\ s \in \mathbb{R}$, platí $M_Y(t) = e^{rt} M_X(st)$ dle bodu (v) věty 2.12.

Nyní tedy k samotnému postupu.

(i) Náhodná veličina s normovaným normálním rozdělením $\mathrm{N}(0, 1)$ má hustotu

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x^2}{2}\right]$$

pro $x \in \mathbb{R}$. Momentová vytvořující funkce je pak

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x^2}{2}\right] \mathrm{d}x = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(x-t)^2}{2}\right] \exp\left[\frac{t^2}{2}\right] \mathrm{d}x = \exp\left[\frac{t^2}{2}\right],$$

neboť v integrálu zůstane po substituci $x - t = y$ hustota rozdělení $\mathrm{N}(0, 1)$, a tedy je integrál roven 1.

(ii) Střední hodnota je první derivace momentové vytvořující funkce v nule. Derivováním dostaneme

$$M_X^{(1)}(t) = \left(\exp\left[\frac{t^2}{2}\right]\right)' = \exp\left[\frac{t^2}{2}\right] t$$

a v bodě 0 máme $E[X] = 0$.
Rozptyl, čili druhý moment, je druhá derivace momentové vytvořující funkce v nule. Derivováním dostaneme

$$M_X^{(2)}(t) = \exp\left[\frac{t^2}{2}\right] (t^2 + 1)$$

a v bodě 0 máme $\mathrm{Var}[X] = E[X]^2 = 1$.

(iii) Chceme momentovou vytvořující funkci pro $\mathrm{N}(\mu, \sigma^2)$. Označme $Y = \mu + \sigma X$. Potom Y má momentovou vytvořující funkci

$$M_Y(t) = e^{\mu t} M_X(\sigma t) = \exp\left[\mu t + \frac{t^2\sigma^2}{2}\right]. \qquad \blacktriangle$$

2.5. Limitní věty pro náhodné veličiny

V této sekci se budeme zabývat konvergencí náhodných veličin a uvedeme si i s důkazem nejjednodušší verzi centrální limitní věty (CLV).

Definice 2.27 (Konvergence v pravděpodobnosti)**.** Řekneme, že posloupnost náhodných veličin $\{X_n\}_{n=1}^{\infty}$ *konverguje v pravděpodobnosti* ke konstantě μ, pokud pro libovolné $\varepsilon > 0$ a $n \longrightarrow \infty$ platí

$$P[|X_n - \mu| \geq \varepsilon] \longrightarrow 0.$$

Značíme $X_n \xrightarrow{P} \mu$. Limitou může být i náhodná veličina.

Věta 2.24 (Postačující podmínky konvergence v pravděpodobnosti)**.** *Nechť $\{X_n\}_{n=1}^{\infty}$ je posloupnost náhodných veličin taková, že $X_n \in L^2(\Omega)$ pro všechna $n \in \mathbb{N}$. Dále mějme $E[X_n] \longrightarrow \mu \in \mathbb{R}$ a $\mathrm{Var}[X_n] \longrightarrow 0$ pro $n \longrightarrow \infty$. Potom $X_n \xrightarrow{P} \mu$ pro $n \longrightarrow \infty$.*

Důkaz. Podle Markovovy nerovnosti (věta 2.10) pro $X = X_n - \mu$ a $p = 2$ máme

$$0 \leq P[|X_n - \mu| \geq \varepsilon] \leq \frac{E[|X_n - \mu|^2]}{\varepsilon^2} \longrightarrow 0$$

pro $n \longrightarrow \infty$, neboť

$$E[|X_n - \mu|^2] = \mathrm{Var}[X_n - \mu] + \left(E[X_n - \mu]\right)^2 = \mathrm{Var}[X_n] + (E[X_n] - \mu)^2 \longrightarrow 0$$

pro $n \longrightarrow \infty$. V poslední rovnosti jsme využili věty 2.8 a 2.7. $\square$

Uvažujme posloupnost náhodných veličin $\{X_n\}_{n=1}^{\infty}$, kde $X_1, X_2, \ldots$ jsou po dvou nezávislé, $E[X_n] = \mu < \infty$ pro všechna $n \in \mathbb{N}$, a posloupnost průměrů $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, kde $n = 1, 2, \ldots$ Za těchto podmínek můžeme formulovat Zákon velkých čísel (ZVČ).

Zákon velkých čísel

Průměry konvergují ke střední hodnotě, zvětšujeme-li počet průměrovaných veličin.

Konkrétní formulaci ZVČ si uvedeme v několika podobách níže.

Věta 2.25 (Čebyševův ZVČ)**.** *Nechť $\{X_n\}_{n=1}^{\infty}$ je posloupnost po dvou nezávislých náhodných veličin. Dále nechť*

$$E[X_n] = \mu, \quad \mathrm{Var}[X_n] < c < \infty$$

pro $n = 1, 2, \ldots$ Potom $\bar{X}_n \xrightarrow{P} \mu$ pro $n \longrightarrow \infty$.

Důkaz. Protože dle předpokladu $\mathrm{Var}[X_n] < c < \infty$, pak $E[X_n] = \mu < \infty$ dle důsledku 2.6. Dále

$$E[\bar{X}_n] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n} n\mu = \mu$$

podle věty 2.7. Podobně

$$\mathrm{Var}[\bar{X}_n] = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}[X_i] < \frac{1}{n^2} nc = \frac{c}{n}$$

podle věty 2.8 a bodu (iii) důsledku 2.9. Vzhledem k tomu, že $E[\bar{X}_n] = \mu$ a $0 \leq \mathrm{Var}[\bar{X}_n] < \frac{c}{n} \longrightarrow 0$, máme $\mathrm{Var}[\bar{X}_n] \longrightarrow 0$ pro $n \longrightarrow \infty$, a dle věty 2.24 dostáváme $\bar{X}_n \xrightarrow{P} \mu$ pro $n \longrightarrow \infty$. $\square$

Poznámka. Tvrzení věty plyne i z Čebyševovy nerovnosti (důsledek 2.7). Protože

$$0 \leq P[|\bar{X}_n - E[\bar{X}_n]| \geq \varepsilon] = P[|\bar{X}_n - \mu| \geq \varepsilon] \leq \frac{\mathrm{Var}[\bar{X}_n]}{\varepsilon^2} < \frac{c}{n\varepsilon^2} \longrightarrow 0$$

pro $n \longrightarrow \infty$, tak $P[|\bar{X}_n - \mu| \geq \varepsilon] \longrightarrow 0$ pro $n \longrightarrow \infty$. Odtud pak dle definice 2.27 máme $\bar{X}_n \xrightarrow{P} \mu$ pro $n \longrightarrow \infty$.

Věta 2.26 (Chinčinův ZVČ)**.** *Nechť* $\{X_n\}_{n=1}^{\infty}$ *je posloupnost nezávislých stejně rozdělených náhodných veličin a*

$$E[X_n] = \mu < \infty$$

pro $n = 1, 2, \ldots$ *Potom* $\bar{X}_n \xrightarrow{P} \mu$ *pro* $n \longrightarrow \infty$.

Důkaz. Vynechán. □

Poznámka. Věta 2.26 nevyžaduje konečný rozptyl veličin X_n. V případě $\mathrm{Var}[X_n] = \sigma^2 < \infty$, $n = 1, 2, \ldots$ ji lze dokázat analogicky jako větu 2.25. Z nezávislosti veličin X_n navíc plyne dle definice 2.23 jejich nezávislost po dvou. Stačilo by předpokládat nekorelovanost po dvou.

Věta 2.27 (Bernoulliho ZVČ)**.** *Uvažujme posloupnost nezávislých náhodných pokusů, z nichž každý má možné výsledky* $A, \bar{A}$, *a označme* n_A *četnost jevu* A *v* n *pokusech. Potom platí*

$$\frac{n_A}{n} \xrightarrow{P} P[A]$$

pro $n \longrightarrow \infty$.

Důkaz. Zaveďme náhodné veličiny

$$X_i = \begin{cases} 1; & \text{nastane-li v i-tém pokusu } A,\ P[X_i = 1] = P[A] = p_A, \\ 0; & \text{nastane-li v i-tém pokusu } \bar{A},\ P[X_i = 0] = P[\bar{A}] = 1 - p_A. \end{cases}$$

Potom

$$\sum_{i=1}^{n} X_i = n_A \sim B_i(n, p_A)$$

a střední hodnota $E[X_i] = p_A$ pro všechna i. Z toho plyne, že

$$\frac{n_A}{n} = \frac{1}{n} \sum_{i=1}^{n} X_i = \bar{X}_n \xrightarrow{P} p_A$$

pro $n \longrightarrow \infty$. □

Poznámky.

(i) Čebyševův, Chinčinův a Bernoulliho ZVČ jsou tzv. slabé zákony velkých čísel, tedy ZVČ s konvergencí v pravděpodobnosti. Existují také tzv. silné zákony velkých čísel, tedy ZVČ s konvergencí skoro jistě.

(ii) Důsledkem ZVČ je, že průměr je kvalitním odhadem střední hodnoty a že relativní četnost je kvalitním odhadem pravděpodobnosti. Teorie odhadu charakteristik náhodných veličin na základě jejich pozorovaných realizací je jednou ze součástí matematické statistiky.

Definice 2.28 (Konvergence v distribuci)**.** Řekneme, že posloupnost náhodných veličin $\{X_n\}_{n=1}^{\infty}$ *konverguje v distribuci* k náhodné veličině X, pokud platí, že $F_{X_n}(x) \longrightarrow F_X(x)$ v každém bodě spojitosti x limitní distribuční funkce F_X. Značíme $X_n \xrightarrow{D} X$, $F_{X_n} \longrightarrow F_X$.

Poznámky.

(i) Konvergence v distribuci je ekvivalentní konvergencím $M_{X_n}(t) \longrightarrow M_X(t)$ a $\Psi_{X_n}(t) \longrightarrow \Psi_X(t)$ pro $n \longrightarrow \infty$, $t \in \mathbb{R}$.

(ii) Pro rozdělení náhodné veličiny X se též používá značení $\mathcal{L}(X)$ (z anglického Law), tedy $X \sim \mathcal{L}(X)$. Při konvergenci v distribuci pak píšeme $X_n \xrightarrow{D} \mathcal{L}(X)$, $\mathcal{L}(X_n) \longrightarrow \mathcal{L}(X)$ nebo $X_n \overset{\text{as.}}{\sim} \mathcal{L}(X)$ (přičemž as. znamená asymptoticky, to jest ve smyslu konvergence v distribuci).

Dále se zaměříme na Centrální limitní větu (CLV).

Centrální limitní věta

Součty nezávislých náhodných veličin mají přibližně normální rozdělení při velkém počtu sčítanců.

Zformulujeme a dokážeme nyní nejjednodušší verzi CLV, tzv. Lindebergovu větu.

Věta 2.28 (CLV pro nezávislé stejně rozdělené náhodné veličiny)**.** *Nechť* $\{X_n\}_{n=1}^{\infty}$ *je posloupnost nezávislých stejně rozdělených náhodných veličin,* $E[X_n] = \mu$, $\mathrm{Var}[X_n] = \sigma^2 < \infty$ *pro* $n = 1, 2, \ldots$ *Potom platí*

$$S_n = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \xrightarrow{D} \mathrm{N}(0,1)$$

pro $n \longrightarrow \infty$.

Důkaz. Budeme uvažovat náhodné veličiny s konečnou momentovou vytvořující funkcí $M_X(t)$. Označme

$$Z_i = \frac{X_i - \mu}{\sigma}, \quad i = 1, \ldots, n.$$

Potom $E[Z_i] = 0$ a $\mathrm{Var}[Z_i] = E[Z_i]^2 = 1$ dle vět 2.7 a 2.8. Z bodu (v) věty 2.12 o vlastnostech momentové vytvořující funkce plyne, že Z_i mají momentovou vytvořující funkci

$$M_Z(t) = \exp\left[-\frac{t\mu}{\sigma}\right] M_X\left(\frac{t}{\sigma}\right) < \infty$$

pro $M_X(t) < \infty,\ t \in \mathbb{R}$. Lze také ukázat, že $\frac{Z_i}{\sqrt{n}}$ mají momentovou vytvořující funkci

$$M_Z\left(\frac{t}{\sqrt{n}}\right) < \infty.$$

Tedy dle věty 2.22 má $S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i$ momentovou vytvořující funkci

$$M_{S_n}(t) = \left[M_Z\left(\frac{t}{\sqrt{n}}\right)\right]^n.$$

Podle bodu (iii) věty 2.12 získáme rozvoj

$$M_Z\left(\frac{t}{\sqrt{n}}\right) = \sum_{k=0}^{\infty} \frac{1}{k!}\left(\frac{t}{\sqrt{n}}\right)^k E[Z]^k = \sum_{k=0}^{2} \frac{1}{k!}\left(\frac{t}{\sqrt{n}}\right)^k E[Z]^k + o\left(\frac{t^2}{n}\right) \approx 1 + \frac{t}{\sqrt{n}} \cdot 0 + \frac{t^2}{2n}\mathrm{E}[Z]^2 = 1 + \frac{t^2}{2n}.$$

Druhá rovnost plyne z konečnosti druhého momentu a z vlastností Taylorova rozvoje. Dále lze psát

$$\lim_{n \longrightarrow \infty} M_{S_n}(t) = \lim_{n \longrightarrow \infty} \left(1 + \frac{t^2}{2n}\right)^n = \exp\left[\frac{t^2}{2}\right],$$

což je momentová vytvořující funkce rozdělení $\mathrm{N}(0,1)$. Podle poznámky (i) za definicí 2.28 jsme dokázali konvergenci momentových vytvořujících funkcí náhodných veličin S_n k momentové vytvořující funkci rozdělení $\mathrm{N}(0,1)$. Tedy jsme dokázali konvergenci distribučních funkcí, a proto $S_n \xrightarrow{D} \mathrm{N}(0,1)$ pro $n \longrightarrow \infty$. □

Důsledek 2.10. *Nechť* $\{X_n\}_{n=1}^{\infty}$ *je posloupnost nezávislých stejně rozdělených náhodných veličin a*

$$E[X_n] = \mu, \quad \mathrm{Var}[X_n] = \sigma^2 < \infty$$

pro $n = 1, 2, \ldots$ *Potom platí, že*

$$\frac{\bar{X}_n - \mu}{\sigma}\sqrt{n} \sim \mathrm{N}(0,1).$$

Důkaz. Podle věty 2.20 jsou $\frac{X_i}{n}$ pro $i = 1, \ldots, n$ nezávislé a stejně rozdělené, přičemž

$$\bar{X}_n = \sum_{i=1}^{n} \frac{X_i}{n}, \quad E[\bar{X}_n] = \mu, \quad \mathrm{Var}[\bar{X}_n] = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}.$$

Potom tedy dle věty 2.28 máme, že

$$\frac{\bar{X}_n - \mu}{\sigma}\sqrt{n} \xrightarrow{D} \mathrm{N}(0, 1)$$

pro $n \longrightarrow \infty$. □

Poznámka. Standardizace náhodných veličin $\sum_{i=1}^{n} X_i$, resp. $\sum_{i=1}^{n} \frac{X_i}{n}$, tj. jejich centrování (odečtení střední hodnoty) a normování (vydělení směrodatnou odchylkou), zajistí podle věty 2.7 nulovou střední hodnotu a podle věty 2.8 jednotkový rozptyl. CLV přidá normalitu limitního rozdělení.

Díky CLV je normální rozdělení klíčovým rozdělením v matematické statistice. Například výsledná chyba různých měření je součtem mnoha nezávislých malých nepřesností a dle CLV má tato chyba normální rozdělení.

3. Náhodné vektory

3.1. Mnohorozměrné rozdělení

Značení. Náhodný vektor.

- *Náhodný vektor* je sloupec reálných náhodných veličin $\boldsymbol{X} = (X_1, \ldots, X_n)^T$.
- Vektor se všemi složkami rovnými 0 značíme $\mathbf{0} = (0, 0, \ldots, 0)^T$.
- Vektor se všemi složkami rovnými 1 značíme $\mathbf{1} = (1, 1, \ldots, 1)^T$.
- Matici $m \times n$ se všemi prvky rovnými 0 značíme $\mathbf{0}_{m\times n}$. Jsou-li rozměry z kontextu jasné, píšeme pouze $\mathbf{0}$.
- Jednotkovou matici $n \times n$ značíme I_n.

Definice 3.1 (n-rozměrná borelovská σ-algebra)**.** Borelovská σ-algebra $\mathcal{B}_0^n$ je nejmenší σ-algebra obsahující množiny vzniklé kartézským součinem

$$\bigtimes_{i=1}^{n} (a_i, b_i),$$

kde $a_i, b_i \in \mathbb{R},\ a_i \le b_i$ a $i = 1, \ldots, n$.

Definice 3.2 (Náhodný vektor)**.** Měřitelné zobrazení $\boldsymbol{X} : (\Omega, \mathcal{A}) \longrightarrow (\mathbb{R}^n, \mathcal{B}_0^n)$ nazýváme *reálný náhodný vektor*.

Poznámky.

(i) Zřejmě platí, že

$$\boldsymbol{X} = \boldsymbol{X}(\omega) = (X_1(\omega), \ldots, X_n(\omega))^T,\ \omega \in \Omega,$$

$$\boldsymbol{X}^{-1}(B) = \{\omega \in \Omega : X_i(\omega) \in B_i,\ i = 1, \ldots, n\} \in \mathcal{A} \text{ pro } B = \bigtimes_{i=1}^{n} B_i \in \mathcal{B}_0^n.$$

Na $(\mathbb{R}^n, \mathcal{B}_0^n)$ definujeme míry.

(ii) V analogii s definicí 2.7 zavádíme *pravděpodobnostní rozdělení* náhodného vektoru $\boldsymbol{X}$ jako indukovanou míru $P_{\boldsymbol{X}}$ s vlastností $P_{\boldsymbol{X}}(B) = P[\boldsymbol{X} \in B],\ B \in \mathcal{B}_0^n$. Nechť μ je σ-konečná míra na $(\mathbb{R}^n, \mathcal{B}_0^n)$ a $P_{\boldsymbol{X}}$ je absolutně spojitá vzhledem k μ. Dále nechť $h(\boldsymbol{x}) : \mathbb{R}^n \longrightarrow \mathbb{R}$ je měřitelná funkce. Potom podle vět 2.1 a 2.2 platí, že

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

kde nezáporná a měřitelná funkce $f_{\boldsymbol{X}}(\boldsymbol{x}) : \mathbb{R}^n \longrightarrow \mathbb{R}$ je hustota náhodného vektoru $\boldsymbol{X}$ vzhledem k μ jednoznačně určující pravděpodobnostní rozdělení $\boldsymbol{X}$.

(iii) Pro $B \in \mathcal{B}_0^n$ lze psát

$$P[\boldsymbol{X} \in B] = \int_{\mathbb{R}^n} \mathbb{I}_B(\boldsymbol{x})\,\mathrm{d}P_{\boldsymbol{X}}(\boldsymbol{x}) = \int_B 1\,\mathrm{d}P_{\boldsymbol{X}}(\boldsymbol{x}) = \int_B f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}\mu(\boldsymbol{x}).$$

Dále pro měřitelnou funkci $g(\boldsymbol{x}) : \mathbb{R}^n \longrightarrow \mathbb{R}$ a $B \in \mathcal{B}_0$ platí, že

$$g^{-1}(B) = \{\boldsymbol{x} \in \mathbb{R}^n : g(\boldsymbol{x}) \in B\} \in \mathcal{B}_0^n$$

a také

$$P[g(\boldsymbol{X}) \in B] = P[\boldsymbol{X} \in g^{-1}(B)] = \int_{g^{-1}(B)} 1\,\mathrm{d}P_{\boldsymbol{X}}(\boldsymbol{x}) = \int_{g^{-1}(B)} f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}\mu(\boldsymbol{x}).$$

(iv) Je-li $P_{\boldsymbol{X}}$ absolutně spojitá vzhledem k součinové čítací míře μ_S na $(\mathbb{R}^n, \mathcal{B}_0^n)$, má náhodný vektor $\boldsymbol{X}$ *mnohorozměrné diskrétní rozdělení.* Nechť $S \subset \mathbb{R}^n$, $S = \times_{i=1}^n S_i$ a $B \in \mathcal{B}_0^n$. Máme pak

$$\mu_S(B) = \sum_{x_1 \in S_1} \cdots \sum_{x_n \in S_n} \mathbb{I}_B(x_1, \ldots, x_n),$$

kde

$$\mathbb{I}_B(\boldsymbol{x}) = \mathbb{I}_B(x_1, \ldots, x_n) = \begin{cases} 1 & \text{pro } \boldsymbol{x} = (x_1, \ldots, x_n)^T \in B, \\ 0 & \text{jinak.} \end{cases}$$

Dále dle definice 2.25 a poznámek za ní lze psát

$$\begin{aligned} P[\boldsymbol{X} \in B] &= \sum_{x_1 \in S_1} \cdots \sum_{x_n \in S_n} \mathbb{I}_B(x_1, \ldots, x_n) P[X_1 = x_1, \ldots, X_n = x_n] \\ &= \underbrace{\sum_{x_1} \cdots \sum_{x_n}}_{\boldsymbol{x} \in (S \cap B)} P[X_1 = x_1, \ldots, X_n = x_n] \end{aligned}$$

a také

$$\sum_{x_1 \in S_1} \cdots \sum_{x_n \in S_n} P[X_1 = x_1, \ldots, X_n = x_n] = 1.$$

Zobecněním (iii) na měřitelnou funkci

$$g : \mathbb{R}^n \longrightarrow \mathbb{R}^m,\ g(X_1, \ldots, X_n) = [g_1(X_1, \ldots, X_n), \ldots, g_m(X_1, \ldots, X_n)]^T$$

máme pro $B_1, \ldots, B_m \subseteq \mathcal{B}_0$

$$P[g_1(\boldsymbol{X}) \in B_1, \ldots, g_m(\boldsymbol{X}) \in B_m] = \sum_{x_1} \cdots \sum_{x_n} P[X_1 = x_1, \ldots, X_n = x_n],$$

kde sčítáme přes množinu $\{\boldsymbol{x} = (x_1, \ldots, x_n)^T : g_1(\boldsymbol{x}) \in B_1, \ldots, g_m(\boldsymbol{x}) \in B_m\}$. Například pro $m = 1$ a vektor $(X, Y)^T$ je

$$P[X + Y < 1] = \sum_x \sum_{y < 1 - x} P[X = x, Y = y].$$

(v) Je-li $P_{\boldsymbol{X}}$ absolutně spojitá vzhledem k součinové Lebesgueově míře μ_L na $(\mathbb{R}^n, \mathcal{B}_0^n)$, má náhodný vektor $\boldsymbol{X}$ *mnohorozměrné spojité rozdělení.* Zřejmě pro $B \in \mathcal{B}_0^n$ vzhledem k poznámkám za definicí 2.26

$$\begin{aligned} P[\boldsymbol{X} \in B] &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathbb{I}_B(x_1, \ldots, x_n) f_{\boldsymbol{X}}(x_1, \ldots, x_n) \, \mathrm{d}x_1 \cdots \mathrm{d}x_n \\ &= \int_{x_1 \in B} \cdots \int_{x_n \in B} f_{\boldsymbol{X}}(x_1, \ldots, x_n) \, \mathrm{d}x_1 \cdots \mathrm{d}x_n \end{aligned}$$

a také

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\boldsymbol{X}}(x_1, \ldots, x_n) \, \mathrm{d}x_1 \cdots \mathrm{d}x_n = 1.$$

Zobecněním (iii) na měřitelnou funkci

$$g : \mathbb{R}^n \longrightarrow \mathbb{R}^m,\ g(X_1, \ldots, X_n) = [g_1(X_1, \ldots, X_n), \ldots, g_m(X_1, \ldots, X_n)]^T$$

máme pro $B_1, \ldots, B_m \subseteq \mathcal{B}_0$

$$P[g_1(\boldsymbol{X}) \in B_1, \ldots, g_m(\boldsymbol{X}) \in B_m] = \int_{x_1} \cdots \int_{x_n} f_{\boldsymbol{X}}(x_1, \ldots, x_n) \, \mathrm{d}x_1 \cdots \mathrm{d}x_n,$$

kde integrujeme přes množinu $\{\boldsymbol{x} = (x_1, \ldots, x_n)^T : g_1(\boldsymbol{x}) \in B_1, \ldots, g_m(\boldsymbol{x}) \in B_m\}$ a podle potřeby zohledníme pořadí integrace. Například pro $m = 1$ a vektor $(X, Y)^T$ je

$$P[X + Y < 1] = \int_x \left(\int_{y < 1 - x} f_{X,Y}(x, y) \, \mathrm{d}y \right) \mathrm{d}x.$$

Spojité rozdělení všech složek nemusí zaručit, že náhodný vektor má spojité rozdělení. Náhodný vektor se spojitými i diskrétními složkami nemá ani spojité, ani diskrétní rozdělení. Má hustotu vzhledem k součinu Lebesgueových a čítacích měr.

(vi) Distribuční funkce náhodného vektoru $\boldsymbol{X}$ je

$$F_{\boldsymbol{X}}(x_1, \ldots, x_n) = P\left[\bigcap_{i=1}^{n}(X_i \le x_i)\right],$$

viz sdružená distribuční funkce z definice 2.24. Tato funkce jednoznačně určuje pravděpodobnostní rozdělení $\boldsymbol{X}$. Pro diskrétně rozdělené náhodné vektory máme podle poznámky (ii) za definicí 2.25

$$F_{\boldsymbol{X}}(\boldsymbol{x}) = \sum_{\substack{k_1 \in S_1 \\ k_1 \le x_1}} \cdots \sum_{\substack{k_n \in S_n \\ k_n \le x_n}} P[X_1 = k_1, \ldots, X_n = k_n].$$

V případě spojitě rozdělených náhodných vektorů platí podle poznámky (ii) za definicí 2.26

$$F_{\boldsymbol{X}}(\boldsymbol{x}) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{\boldsymbol{X}}(u_1, \ldots, u_n)\,\mathrm{d}u_1 \cdots \mathrm{d}u_n$$

a navíc podle definice 2.26

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{\partial^n F_{\boldsymbol{X}}(\boldsymbol{x})}{\partial x_1 \cdots \partial x_n}.$$

(vii) Nakonec lze ukázat, že $F_{\boldsymbol{X}}(\boldsymbol{x})$ má následující vlastnosti (analogie věty 2.3):

(a) $F_{\boldsymbol{X}}(\boldsymbol{x})$ je neklesající a zprava spojitá v každé proměnné;

(b) pokud existuje $1 \le i \le n$, takové, že $x_i \longrightarrow -\infty$, potom $F_{\boldsymbol{X}}(\boldsymbol{x}) \longrightarrow 0$;

(c) pokud pro všechna $1 \le i \le n$ platí, že $x_i \longrightarrow +\infty$, potom $F_{\boldsymbol{X}}(\boldsymbol{x}) \longrightarrow 1$.

Definice 3.3 (Marginální rozdělení)**.** Nechť $\boldsymbol{X} = (X_1, \ldots, X_n)^T$ je náhodný vektor. Pro libovolné $1 \le r < n$ a libovolné indexy $i_1, \ldots, i_r$ označujeme pravděpodobnostní rozdělení náhodného vektoru $(X_{i_1}, \ldots, X_{i_r})^T$ jako *marginální rozdělení.*

Poznámka. Sdružené rozdělení náhodného vektoru $\boldsymbol{X}$ jednoznačně určuje všechna marginální rozdělení. Tento vztah opačně neplatí.

Věta 3.1 (Marginální distribuční funkce)**.** *Nechť $\boldsymbol{X} = (X_1, \ldots, X_n)^T$ je náhodný vektor s distribuční funkcí $F_{\boldsymbol{X}}(\boldsymbol{x})$. Pro $1 \le r < n$ má náhodný vektor $\boldsymbol{Y} = (X_1, \ldots, X_r)^T$ distribuční funkci*

$$F_{\boldsymbol{Y}}(x_1, \ldots, x_r) = \lim_{\substack{x_{r+1} \longrightarrow \infty \\ \cdots \\ x_n \longrightarrow \infty}} F_{\boldsymbol{X}}(x_1, \ldots, x_r, x_{r+1}, \ldots, x_n).$$

Důkaz. Z definice 2.24 plyne

$$\begin{aligned}
F_{\boldsymbol{Y}}(x_1, \ldots, x_r) &= P[X_1 \le x_1, \ldots, X_r \le x_r] \\
&= P[X_1 \le x_1, \ldots, X_r \le x_r, X_{r+1} < \infty, \ldots, X_n < \infty] \\
&= \lim_{\substack{x_{r+1} \longrightarrow \infty \\ \cdots \\ x_n \longrightarrow \infty}} P[X_1 \le x_1, \ldots, X_r \le x_r, X_{r+1} \le x_{r+1}, \ldots, X_n \le x_n] \\
&= \lim_{\substack{x_{r+1} \longrightarrow \infty \\ \cdots \\ x_n \longrightarrow \infty}} F_{\boldsymbol{X}}(x_1, \ldots, x_n).
\end{aligned}$$

□

Poznámka. Z věty 3.1 plyne, že pokud jsou náhodné veličiny $X_1, \ldots, X_n$ nezávislé, pak jsou také po dvou nezávislé, neboť pro libovolnou dvojici indexů (i, j) máme

$$\begin{aligned}
F_{X_i, X_j}(x_i, x_j) &= P[X_i \le x_i, X_j \le x_j] = \lim_{\substack{x_k \longrightarrow \infty \\ k \ne i,j}} P[X_1 \le x_1, \ldots, X_n \le x_n] \\
&= \lim_{\substack{x_k \longrightarrow \infty \\ k \ne i,j}} [P[X_1 \le x_1] \cdots P[X_n \le x_n]] = P[X_i \le x_i] P[X_j \le x_j] \lim_{x_k \longrightarrow \infty} \prod_{k \ne i,j} P[X_k \le x_k] \\
&= F_{X_i}(x_i) F_{X_j}(x_j) \prod_{k \ne i,j} \lim_{x_k \longrightarrow \infty} F_{X_k}(x_k) = F_{X_i}(x_i) F_{X_j}(x_j).
\end{aligned}$$

Třetí rovnost plyne z definice 2.23, závěr potom z věty 2.3 a věty 2.16.

Důsledek 3.1. *Vzhledem k bodu (iii) důsledku 2.9 dostáváme, že pokud jsou $X_1, \dots, X_n \in L^2(\Omega)$ nezávislé, potom*

$$\operatorname{Var}\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n \operatorname{Var}[X_i].$$

Věta 3.2 (Marginální pravděpodobnosti)**.** *Nechť $\boldsymbol{X}$ je náhodný vektor s diskrétním rozdělením daným pravděpodobnostmi $P[X_1 = x_1, \dots, X_n = x_n]$, kde $x_i \in S_i$, $i = 1, \dots, n$. Pro $1 \le r < n$ má náhodný vektor $\boldsymbol{Y} = (X_1, \dots, X_r)^T$ rozdělení dané pravděpodobnostmi*

$$P[X_1 = x_1, \dots, X_r = x_r] = \sum_{x_{r+1} \in S_{r+1}} \cdots \sum_{x_n \in S_n} P[X_1 = x_1, \dots, X_r = x_r,\, X_{r+1} = x_{r+1}, \dots, X_n = x_n].$$

Důkaz. Zjevně platí

$$\begin{aligned} P[X_1 = x_1, \dots, X_r = x_r] &= P[X_1 = x_1, \dots, X_r = x_r, X_{r+1} < \infty, \dots, X_n < \infty] \\ &= \sum_{x_{r+1} \in S_{r+1}} \cdots \sum_{x_n \in S_n} P[X_1 = x_1, \dots, X_r = x_r,\, X_{r+1} = x_{r+1}, \dots, X_n = x_n]. \end{aligned}$$

□

Věta 3.3 (Marginální hustota)**.** *Nechť $\boldsymbol{X} = (X_1, \dots, X_n)^T$ je náhodný vektor se spojitým rozdělením s hustotou $f_{\boldsymbol{X}}(\boldsymbol{x})$. Pro $1 \le r < n$ má náhodný vektor $\boldsymbol{Y} = (X_1, \dots, X_r)^T$ spojité rozdělení s hustotou*

$$f_{\boldsymbol{Y}}(x_1, \dots, x_r) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\boldsymbol{X}}(x_1, \dots, x_r, x_{r+1}, \dots, x_n)\, \mathrm{d}x_{r+1} \cdots \mathrm{d}x_n.$$

Důkaz. Opět jednoduše rozepíšeme (se zohledněním pořadí integrace):

$$\begin{aligned} F_{\boldsymbol{Y}}(x_1, \dots, x_r) &= P[X_1 \le x_1, \dots, X_r \le x_r, X_{r+1} < \infty, \dots, X_n < \infty] \\ &= \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_r} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\boldsymbol{X}}(u_1, \dots, u_r, u_{r+1}, \dots, u_n)\, \mathrm{d}u_n \cdots \mathrm{d}u_{r+1}\, \mathrm{d}u_r \cdots \mathrm{d}u_1 \\ &= \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_r} f_{\boldsymbol{Y}}(u_1, \dots, u_r)\, \mathrm{d}u_r \cdots \mathrm{d}u_1, \end{aligned}$$

z čehož vyplývá naše tvrzení. □

Poznámky.

(i) Věty 3.1, 3.2 a 3.3 lze dokázat pro libovolný r-složkový podvektor vektoru $\boldsymbol{X}$, kde $1 \le r < n$.

(ii) Věty 3.2 a 3.3 lze shrnout do obecnějšího tvrzení. Nechť

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

má hustotu $f_{\boldsymbol{X}}(\boldsymbol{x})$ vzhledem k součinové míře $\mu_1 \times \mu_2$ na $(\mathbb{R}^n, \mathcal{B}_0^n)$. Dále nechť

$$\boldsymbol{Y} = (X_1, \dots, X_r)^T = (Y_1, \dots, Y_r)^T$$

má hustotu $f_{\boldsymbol{Y}}(\boldsymbol{y})$ vzhledem k σ-konečné míře μ_1 na $(\mathbb{R}^r, \mathcal{B}_0^r)$. Nakonec nechť

$$\boldsymbol{Z} = (X_{r+1}, \dots, X_n)^T = (Z_1, \dots, Z_{n-r})^T$$

má hustotu $f_{\boldsymbol{Z}}(\boldsymbol{z})$ vzhledem k σ-konečné míře μ_2 na $(\mathbb{R}^{n-r}, \mathcal{B}_0^{n-r})$. Potom platí

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \int_{\mathbb{R}^{n-r}} f_{\boldsymbol{X}}(\boldsymbol{y}, \boldsymbol{z})\, \mathrm{d}\mu_2(\boldsymbol{z}).$$

3.2. Momenty náhodného vektoru

Definice 3.4 (Střední hodnota náhodného vektoru)**.** Nechť $\boldsymbol{X} = (X_1, \ldots, X_n)^T$ je náhodný vektor takový, že $X_i \in L^1(\Omega)$ pro $i = 1, \ldots, n$. *Střední hodnotu náhodného vektoru* $\boldsymbol{X}$ definujeme předpisem

$$E[\boldsymbol{X}] = \boldsymbol{\mu} = (E[X_1], \ldots, E[X_n])^T.$$

Poznámky.

(i) Pro měřitelnou funkci $g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ definujeme

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

je-li $g_i(\boldsymbol{X}) \in L^1(\Omega)$ pro $i = 1, \ldots, m$. Vzhledem k poznámce (ii) za definicí 3.2 máme pro diskrétně rozdělený náhodný vektor $\boldsymbol{X}$

$$E[g(\boldsymbol{X})] = \sum_{x_1} \cdots \sum_{x_n} g(x_1, \ldots, x_n) P[X_1 = x_1, \ldots, X_n = x_n],$$

přičemž k-tá složka m-rozměrného vektoru $E[g(\boldsymbol{X})]$ je

$$\sum_{x_1} \cdots \sum_{x_n} g_k(x_1, \ldots, x_n) P[X_1 = x_1, \ldots, X_n = x_n].$$

Sčítáme přes $x_i \in S_i$, $i = 1, \ldots, n$.

Pro spojitě rozdělený náhodný vektor $\boldsymbol{X}$ máme

$$E[g(\boldsymbol{X})] = \int_{x_1} \cdots \int_{x_n} g(x_1, \ldots, x_n) f_{\boldsymbol{X}}(x_1, \ldots, x_n) \, \mathrm{d}x_1 \cdots \mathrm{d}x_n,$$

přičemž k-tá složka m-rozměrného vektoru $E[g(\boldsymbol{X})]$ je

$$\int_{x_1} \cdots \int_{x_n} g_k(x_1, \ldots, x_n) f_{\boldsymbol{X}}(x_1, \ldots, x_n) \, \mathrm{d}x_1 \cdots \mathrm{d}x_n.$$

Integrujeme přes $x_i \in (-\infty, \infty)$, $i = 1, \ldots, n$ a podle potřeby zohledníme pořadí integrace.

(ii) Střední hodnota matice náhodných veličin je matice středních hodnot jednotlivých veličin.

Věta 3.4 (Vlastnosti střední hodnoty náhodných vektorů)**.** *Pro náhodné vektory* $\boldsymbol{X}, \boldsymbol{Y} \in \mathbb{R}^n$, *vektor* $\boldsymbol{a} \in \mathbb{R}^m$ *a matici* $\boldsymbol{B}_{m \times n}$ *platí*

(i) $E[\boldsymbol{a} + \boldsymbol{B}\boldsymbol{X}] = \boldsymbol{a} + \boldsymbol{B}E[\boldsymbol{X}]$, *je-li* $X_i \in L^1(\Omega)$ *pro* $i = 1, \ldots, n$;

(ii) $E[\boldsymbol{X} + \boldsymbol{Y}] = E[\boldsymbol{X}] + E[\boldsymbol{Y}]$, *je-li* $X_i, Y_i \in L^1(\Omega)$ *pro* $i = 1, \ldots, n$.

Důkaz. V obou případech budeme postupovat po složkách.

(i) Podle věty 2.7 máme

$$E\left[a_i + \sum_{j=1}^{n} b_{ij} X_j\right] = a_i + \sum_{j=1}^{n} b_{ij} E[X_j],$$

přičemž $\boldsymbol{B} = (b_{ij})$.

(ii) Opět z vlastnosti střední hodnoty plyne

$$E[X_i + Y_i] = E[X_i] + E[Y_i]. \qquad \square$$

Poznámka. Bod (ii) z věty 3.4 lze zobecnit pro k n-rozměrných vektorů.

Definice 3.5 (Varianční matice)**.** Nechť $\boldsymbol{X} = (X_1, \ldots, X_n)^T$ je náhodný vektor takový, že $X_i \in L^2(\Omega)$, $i = 1, \ldots, n$. Matici

$$\mathrm{Var}[\boldsymbol{X}] = \boldsymbol{V} = E[(\boldsymbol{X} - E[\boldsymbol{X}])(\boldsymbol{X} - E[\boldsymbol{X}])^T]$$

nazýváme *varianční matice* náhodného vektoru $\boldsymbol{X}$. Případně lze použít označení *rozptylová matice*.

Věta 3.5 (Vlastnosti varianční matice)**.** *Nechť* $\boldsymbol{X} \in \mathbb{R}^n$ *je náhodný vektor takový, že* $X_i \in L^2(\Omega)$, $i = 1, \dots, n$ *a jeho varianční matice je* $\boldsymbol{V} = (\sigma_{ij})_{i,j=1}^n$. *Potom platí*

(i) $\sigma_{ij} = \operatorname{cov}(X_i, X_j)$, $\sigma_{ii} = \operatorname{Var}[X_i]$ *a* $\boldsymbol{V}$ *je symetrická;*

(ii) $\boldsymbol{V} = E[\boldsymbol{X}\boldsymbol{X}^T] - E[\boldsymbol{X}](E[\boldsymbol{X}])^T$;

(iii) $\operatorname{Var}[\boldsymbol{a} + \boldsymbol{B}\boldsymbol{X}] = \boldsymbol{B}\boldsymbol{V}\boldsymbol{B}^T$, *kde* $\boldsymbol{a} \in \mathbb{R}^m$ *a* $\boldsymbol{B}$ *je typu* $m \times n$;

(iv) $\boldsymbol{V}$ *je pozitivně semidefinitní.*

Důkaz.

(i) Matice $E[(\boldsymbol{X} - E[\boldsymbol{X}])(\boldsymbol{X} - E[\boldsymbol{X}])^T]$ typu $n \times n$ má na místě (i, j) prvek $E[(X_i - E[X_i])(X_j - E[X_j])]$, takže podle definice kovariance (2.20) máme $\sigma_{ij} = \operatorname{cov}(X_i, X_j)$. Dále pro $i = j$ dostáváme dle věty 2.13 rovnost $\sigma_{ii} = \operatorname{cov}(X_i, X_i) = \operatorname{Var}[X_i]$. Dle stejné věty platí, že $\operatorname{cov}(X_i, X_j) = \operatorname{cov}(X_j, X_i)$, takže $\boldsymbol{V}$ musí být symetrická.

(ii) Matice $E[\boldsymbol{X}\boldsymbol{X}^T]$ má na místě (i, j) prvek $E[X_i X_j]$. Matice $E[\boldsymbol{X}](E[\boldsymbol{X}])^T$ má na místě (i, j) prvek $E[X_i]E[X_j]$. Pak tedy matice $E[\boldsymbol{X}\boldsymbol{X}^T] - E[\boldsymbol{X}](E[\boldsymbol{X}])^T$ má vzhledem k větě 2.13 na místě (i, j) prvky

$$E[X_i X_j] - E[X_i]E[X_j] = \operatorname{cov}(X_i, X_j).$$

(iii) Počítejme:

$$\begin{aligned}
\operatorname{Var}[\boldsymbol{a} + \boldsymbol{B}\boldsymbol{X}] &= E\left[(\boldsymbol{a} + \boldsymbol{B}\boldsymbol{X} - E[\boldsymbol{a} + \boldsymbol{B}\boldsymbol{X}])\,(\boldsymbol{a} + \boldsymbol{B}\boldsymbol{X} - E[\boldsymbol{a} + \boldsymbol{B}\boldsymbol{X}])^T\right] \\
&= E\left[(\boldsymbol{B}(\boldsymbol{X} - E[\boldsymbol{X}]))(\boldsymbol{B}(\boldsymbol{X} - E[\boldsymbol{X}]))^T\right] = E\left[\boldsymbol{B}(\boldsymbol{X} - E[\boldsymbol{X}])(\boldsymbol{X} - E[\boldsymbol{X}])^T\boldsymbol{B}^T\right] \\
&= \boldsymbol{B}(\operatorname{Var}[\boldsymbol{X}])\boldsymbol{B}^T.
\end{aligned}$$

Druhá rovnost plyne z věty 3.4, poslední rovnost potom z věty 3.4 a definice 3.5.

(iv) Nechť $\boldsymbol{c} \in \mathbb{R}^n$. Potom

$$\operatorname{Var}[\boldsymbol{c}^T\boldsymbol{X}] = \boldsymbol{c}^T\boldsymbol{V}\boldsymbol{c}$$

dle bodu (iii). Dále zřejmě dle věty 2.8

$$\operatorname{Var}[\boldsymbol{c}^T\boldsymbol{X}] = \operatorname{Var}\left[\sum_{i=1}^n c_i X_i\right] \geq 0.$$

Pak tedy $\boldsymbol{c}^T\boldsymbol{V}\boldsymbol{c} \geq 0$ pro všechna $\boldsymbol{c} \in \mathbb{R}^n$, takže $\boldsymbol{V}$ je pozitivně semidefinitní. □

Značení. Nechť nyní $\boldsymbol{X} = (X_1, \dots, X_n)^T$ je rozdělen na podvektory

(i) $\boldsymbol{Y} = (X_1, \dots, X_r)^T = (Y_1, \dots, Y_r)^T$,

(ii) $\boldsymbol{Z} = (X_{r+1}, \dots, X_n)^T = (Z_1, \dots, Z_{n-r})^T$.

Definice 3.6 (Kovarianční matice)**.** *Kovarianční matice* náhodných vektorů $\boldsymbol{Y} \in \mathbb{R}^r$ a $\boldsymbol{Z} \in \mathbb{R}^{n-r}$, pro něž platí, že

$$Y_i, Z_j \in L^2(\Omega) \quad \text{pro } i = 1, \dots, r \text{ a pro } j = 1, \dots, n - r,$$

je definována předpisem

$$\operatorname{cov}(\boldsymbol{Y}, \boldsymbol{Z}) = E[(\boldsymbol{Y} - E[\boldsymbol{Y}])(\boldsymbol{Z} - E[\boldsymbol{Z}])^T].$$

Věta 3.6 (Vlastnosti kovarianční matice)**.** *Nechť náhodné vektory* $\boldsymbol{Y}, \boldsymbol{Z}$ *splňují předpoklady z definice 3.6. Pro jejich kovarianční matici pak platí*

(i) $\sigma_{ij} = \operatorname{cov}(Y_i, Z_j)$ *pro všechny dvojice* (i, j);

(ii) $\operatorname{cov}(\boldsymbol{Y}, \boldsymbol{Z}) = E[\boldsymbol{Y}\boldsymbol{Z}^T] - E[\boldsymbol{Y}](E[\boldsymbol{Z}])^T$;

(iii) $\operatorname{cov}(\boldsymbol{Y}, \boldsymbol{Y}) = \operatorname{Var}[\boldsymbol{Y}]$, $\operatorname{cov}(\boldsymbol{Z}, \boldsymbol{Y}) = (\operatorname{cov}(\boldsymbol{Y}, \boldsymbol{Z}))^T$;

(iv) $\operatorname{cov}(\boldsymbol{a}+\boldsymbol{B}\boldsymbol{Y}, \boldsymbol{c}+\boldsymbol{D}\boldsymbol{Z}) = \boldsymbol{B}\operatorname{cov}(\boldsymbol{Y},\boldsymbol{Z})\boldsymbol{D}^T$, *přičemž* $\boldsymbol{a} \in \mathbb{R}^m$, $\boldsymbol{B}$ *je matice typu* $m \times r$, $\boldsymbol{c} \in \mathbb{R}^q$ *a* $\boldsymbol{D}$ *je matice typu* $q \times s$;

(v) $\operatorname{cov}(\boldsymbol{Y}_1 \pm \boldsymbol{Y}_2, \boldsymbol{Z}) = \operatorname{cov}(\boldsymbol{Y}_1, \boldsymbol{Z}) \pm \operatorname{cov}(\boldsymbol{Y}_2, \boldsymbol{Z})$, *přičemž* $\boldsymbol{Y}_1, \boldsymbol{Y}_2 \in \mathbb{R}^r$ *a mají konečné druhé momenty.*

Důkaz.

(i) Matice $E[(\boldsymbol{Y} - E[\boldsymbol{Y}])(\boldsymbol{Z} - E[\boldsymbol{Z}])^T]$ typu $r \times s$ má na místě (i,j) prvek $E[(Y_i - E[Y_i])(Z_j - E[Z_j])]$, takže dle definice 2.20 máme $\sigma_{ij} = \operatorname{cov}(Y_i, Z_j)$.

(ii) Matice $E[\boldsymbol{Y}\boldsymbol{Z}^T]$ má na místě (i,j) prvek $E[Y_i Z_j]$. Matice $E[\boldsymbol{Y}](E[\boldsymbol{Z}])^T$ má na místě (i,j) prvek $E[Y_i]E[Z_j]$. Pak tedy matice $E[\boldsymbol{Y}\boldsymbol{Z}^T] - E[\boldsymbol{Y}](E[\boldsymbol{Z}])^T$ má na místě (i,j) prvky $E[Y_iZ_j] - E[Y_i]E[Z_j] = \operatorname{cov}(Y_i, Z_j)$ podle věty 2.13.

(iii) Podle definice varianční matice (3.5) máme $\operatorname{cov}(\boldsymbol{Y},\boldsymbol{Y}) = E[(\boldsymbol{Y} - E[\boldsymbol{Y}])(\boldsymbol{Y} - E[\boldsymbol{Y}])^T] = \operatorname{Var}[\boldsymbol{Y}]$. Dále počítejme:

$$\begin{aligned}\operatorname{cov}(\boldsymbol{Z},\boldsymbol{Y}) &= E[(\boldsymbol{Z} - E[\boldsymbol{Z}])(\boldsymbol{Y} - E[\boldsymbol{Y}])^T] = E[((\boldsymbol{Y} - E[\boldsymbol{Y}])(\boldsymbol{Z} - E[\boldsymbol{Z}])^T)^T] \\ &= (E[(\boldsymbol{Y} - E[\boldsymbol{Y}])(\boldsymbol{Z} - E[\boldsymbol{Z}])^T])^T = (\operatorname{cov}(\boldsymbol{Y},\boldsymbol{Z}))^T.\end{aligned}$$

(iv) Analogicky jako v důkazu bodu (iii) věty 3.5:

$$\begin{aligned}\operatorname{cov}(\boldsymbol{a}+\boldsymbol{B}\boldsymbol{Y}, \boldsymbol{c}+\boldsymbol{D}\boldsymbol{Z}) &= E[(\boldsymbol{B}(\boldsymbol{Y} - E[\boldsymbol{Y}]))(\boldsymbol{D}(\boldsymbol{Z} - E[\boldsymbol{Z}]))^T] \\ &= \boldsymbol{B}E[(\boldsymbol{Y} - E[\boldsymbol{Y}])(\boldsymbol{Z} - E[\boldsymbol{Z}])^T]\boldsymbol{D}^T = \boldsymbol{B}\operatorname{cov}(\boldsymbol{Y},\boldsymbol{Z})\boldsymbol{D}^T.\end{aligned}$$

(v) Uvažujme případ součtu. Rozepsáním definice 3.6 dostaneme

$$\begin{aligned}\operatorname{cov}(\boldsymbol{Y}_1 + \boldsymbol{Y}_2, \boldsymbol{Z}) &= E[(\boldsymbol{Y}_1 + \boldsymbol{Y}_2 - E[\boldsymbol{Y}_1] - E[\boldsymbol{Y}_2])(\boldsymbol{Z} - E[\boldsymbol{Z}])^T] \\ &= E[(\boldsymbol{Y}_1 - E[\boldsymbol{Y}_1])(\boldsymbol{Z} - E[\boldsymbol{Z}])^T] + E[(\boldsymbol{Y}_2 - E[\boldsymbol{Y}_2])(\boldsymbol{Z} - E[\boldsymbol{Z}])^T] \\ &= \operatorname{cov}(\boldsymbol{Y}_1, \boldsymbol{Z}) + \operatorname{cov}(\boldsymbol{Y}_2, \boldsymbol{Z}).\end{aligned}$$

Pro rozdíl analogicky. □

Důsledek 3.2. *Zvolíme-li v bodě (iv) věty 3.6*

$$\boldsymbol{Y} = \boldsymbol{Z} = \boldsymbol{X} = (X_1, \ldots, X_n)^T, \quad \boldsymbol{a} = \boldsymbol{c} = 0 \in \mathbb{R}, \quad \boldsymbol{B} = \boldsymbol{D} = \boldsymbol{1}^T = (1, \ldots, 1) \in \mathbb{R}^n,$$

dostaneme

$$\begin{aligned}\operatorname{cov}\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i\right) &= \operatorname{Var}\left[\sum_{i=1}^n X_i\right] = \boldsymbol{1}^T \operatorname{Var}[\boldsymbol{X}] \cdot \boldsymbol{1} = \sum_{i=1}^n \sum_{j=1}^n \operatorname{cov}(X_i, X_j) \\ &= \sum_{i=1}^n \operatorname{Var}[X_i] + \sum_{\substack{i=1 \\ i \neq j}}^n \sum_{j=1}^n \operatorname{cov}(X_i, X_j).\end{aligned}$$

Dokázali jsme vztah z poznámky za větou 2.14. Podobně pro

$$\boldsymbol{a} = \boldsymbol{c} = 0 \in \mathbb{R}, \quad \boldsymbol{B} = \boldsymbol{1}^T \in \mathbb{R}^r, \quad \boldsymbol{D} = \boldsymbol{1}^T \in \mathbb{R}^s$$

dostáváme

$$\operatorname{cov}\left(\sum_{i=1}^r Y_i, \sum_{j=1}^s Z_j\right) = \sum_{i=1}^r \sum_{j=1}^s \operatorname{cov}(Y_i, Z_j).$$

Věta 3.7 (Rozptyl součtu náhodných vektorů). *Nechť* $\boldsymbol{Y}, \boldsymbol{Z} \in \mathbb{R}^n$ *jsou náhodné vektory a* $Y_i, Z_i \in L^2(\Omega)$ *pro* $i = 1, \ldots, n$. *Pak platí*

$$\operatorname{Var}[\boldsymbol{Y} + \boldsymbol{Z}] = \operatorname{Var}[\boldsymbol{Y}] + \operatorname{Var}[\boldsymbol{Z}] + \operatorname{cov}(\boldsymbol{Y},\boldsymbol{Z}) + \operatorname{cov}(\boldsymbol{Z},\boldsymbol{Y}).$$

Důkaz. Počítejme za využití vlastností střední hodnoty a rozptylu:

$$\begin{aligned}\operatorname{Var}[\boldsymbol{Y}+\boldsymbol{Z}] &= E[(\boldsymbol{Y}+\boldsymbol{Z}-E[\boldsymbol{Y}]-E[\boldsymbol{Z}])(\boldsymbol{Y}+\boldsymbol{Z}-E[\boldsymbol{Y}]-E[\boldsymbol{Z}])^T] \\ &= E[((\boldsymbol{Y}-E[\boldsymbol{Y}])+(\boldsymbol{Z}-E[\boldsymbol{Z}]))((\boldsymbol{Y}-E[\boldsymbol{Y}])+(\boldsymbol{Z}-E[\boldsymbol{Z}]))^T] \\ &= E[(\boldsymbol{Y}-E[\boldsymbol{Y}])(\boldsymbol{Y}-E[\boldsymbol{Y}])^T]+E[(\boldsymbol{Y}-E[\boldsymbol{Y}])(\boldsymbol{Z}-E[\boldsymbol{Z}])^T] \\ &+ E[(\boldsymbol{Z}-E[\boldsymbol{Z}])(\boldsymbol{Y}-E[\boldsymbol{Y}])^T]+E[(\boldsymbol{Z}-E[\boldsymbol{Z}])(\boldsymbol{Z}-E[\boldsymbol{Z}])^T] \\ &= \operatorname{Var}[\boldsymbol{Y}]+\operatorname{cov}(\boldsymbol{Y},\boldsymbol{Z})+\operatorname{cov}(\boldsymbol{Z},\boldsymbol{Y})+\operatorname{Var}[\boldsymbol{Z}]. \end{aligned}$$

□

Poznámka. Pro náhodné vektory $\boldsymbol{X}_1,\dots,\boldsymbol{X}_k \in \mathbb{R}^n$ s konečnými druhými momenty platí, že

$$\operatorname{Var}\left[\sum_{i=1}^k \boldsymbol{X}_i\right] = \sum_{i=1}^k \operatorname{Var}[\boldsymbol{X}_i] + \sum_{\substack{i=1\\ i\neq j}}^k \sum_{j=1}^k \operatorname{cov}(\boldsymbol{X}_i,\boldsymbol{X}_j).$$

Definice 3.7 (Korelační matice)**.** Nechť $\boldsymbol{X} \in \mathbb{R}^n$ je náhodný vektor takový, že $X_i \in L^2(\Omega)$, $i=1,\dots,n$ a jeho varianční matice je $\boldsymbol{V} = (\sigma_{ij})_{i,j=1}^n$. *Korelační matici* vektoru $\boldsymbol{X}$ definujeme jako

$$\operatorname{cor}(\boldsymbol{X}) = \boldsymbol{D}^{-1}\boldsymbol{V}\boldsymbol{D}^{-1},$$

kde $\boldsymbol{D} = \operatorname{diag}\{\sqrt{\sigma_{11}},\dots,\sqrt{\sigma_{nn}}\}$.

Věta 3.8 (Vlastnosti korelační matice)**.** *Nechť $\boldsymbol{X} \in \mathbb{R}^n$ je náhodný vektor takový, že $X_i \in L^2(\Omega)$, $i=1,\dots,n$ a jeho korelační matice je $\operatorname{cor}(\boldsymbol{X}) = (\rho_{ij})_{i,j=1}^n$. Pak platí*

(i) $\rho_{ij} = \rho_{X_i,X_j}$ *pro všechny dvojice* (i,j);

(ii) $\rho_{ii} = 1$ *pro všechna* $i=1,\dots,n$;

(iii) $\operatorname{cor}(\boldsymbol{X})$ *je symetrická.*

Důkaz. Tvrzení plyne přímo z definic 3.7 a 2.21 a z věty 2.15. □

Poznámka. Pro náhodné vektory $\boldsymbol{Y} \in \mathbb{R}^r, \boldsymbol{Z} \in \mathbb{R}^s$ definujeme korelační matici předpisem

$$\operatorname{cor}(\boldsymbol{Y},\boldsymbol{Z}) = \boldsymbol{D}_Y^{-1}\operatorname{cov}(\boldsymbol{Y},\boldsymbol{Z})\boldsymbol{D}_Z^{-1},$$

kde

$$\boldsymbol{D}_Y = \operatorname{diag}\{\sqrt{\operatorname{Var}[Y_1]},\dots,\sqrt{\operatorname{Var}[Y_r]}\}, \quad \boldsymbol{D}_Z = \operatorname{diag}\{\sqrt{\operatorname{Var}[Z_1]},\dots,\sqrt{\operatorname{Var}[Z_s]}\},$$

přičemž $Y_i, Z_j \in L^2(\Omega)$, $i=1,\dots,r$ a $j=1,\dots,s$.
Matice $\operatorname{cor}(\boldsymbol{Y},\boldsymbol{Z})$ má zřejmě prvky $\rho_{ij} = \rho_{Y_i,Z_j}$ pro všechny dvojice (i,j) a platí, že $\operatorname{cor}(\boldsymbol{Z},\boldsymbol{Y}) = (\operatorname{cor}(\boldsymbol{Y},\boldsymbol{Z}))^T$.

3.3. Nezávislost náhodných vektorů

Značení. Uvažujme náhodný vektor $\boldsymbol{X} = (X_1,\dots,X_n)^T$ rozdělený na podvektory

$$\begin{aligned}\boldsymbol{Y} &= (X_1,\dots,X_r)^T = (Y_1,\dots,Y_r)^T, \\ \boldsymbol{Z} &= (X_{r+1},\dots,X_n)^T = (Z_1,\dots,Z_s)^T,\end{aligned}$$

kde $1 \le r < n$ a $s = n-r$.

Definice 3.8 (Nezávislé náhodné vektory)**.** Náhodné vektory $\boldsymbol{Y} \in \mathbb{R}^r$ a $\boldsymbol{Z} \in \mathbb{R}^s$ nazveme *nezávislé náhodné vektory*, když pro libovolné množiny $B \in \mathcal{B}_0^r$ a $C \in \mathcal{B}_0^s$ platí, že

$$P[\boldsymbol{Y} \in B, \boldsymbol{Z} \in C] = P[\boldsymbol{Y} \in B]P[\boldsymbol{Z} \in C].$$

Stejně jako u náhodných veličin lze formulovat několik různých kritérií nezávislosti.

Věta 3.9 (Kritéria nezávislosti)**.**

(i) *Náhodné vektory* $\boldsymbol{Y}$ *a* $\boldsymbol{Z}$ *jsou nezávislé právě tehdy, když pro distribuční funkci* $F_{\boldsymbol{X}}$ *a marginální distribuční funkce* $F_{\boldsymbol{Y}}, F_{\boldsymbol{Z}}$ *platí*

$$F_{\boldsymbol{X}}(\boldsymbol{x}) = F_{\boldsymbol{Y}}(\boldsymbol{y})F_{\boldsymbol{Z}}(\boldsymbol{z})$$

pro $\boldsymbol{y} \in R^r$, $\boldsymbol{z} \in \mathbb{R}^s$ *a* $\boldsymbol{x} = (\boldsymbol{y}^T, \boldsymbol{z}^T)^T$.

(ii) *Je-li* $P_{\boldsymbol{X}}$ *absolutně spojitá vzhledem k součinové čítací míře na* $(\mathbb{R}^n, \mathcal{B}_0^n)$, *jsou* $\boldsymbol{Y}$ *a* $\boldsymbol{Z}$ *nezávislé právě tehdy, když* $P[\boldsymbol{X} = \boldsymbol{x}] = P[\boldsymbol{Y} = \boldsymbol{y}]P[\boldsymbol{Z} = \boldsymbol{z}]$ *pro všechny* $(\boldsymbol{y}, \boldsymbol{z})$.

(iii) *Je-li* $P_{\boldsymbol{X}}$ *absolutně spojitá vzhledem k součinové Lebesgueově míře na* $(\mathbb{R}^n, \mathcal{B}_0^n)$, *jsou* $\boldsymbol{Y}$ *a* $\boldsymbol{Z}$ *nezávislé právě tehdy, když pro hustotu* $f_{\boldsymbol{X}}$ *a marginální hustoty* $f_{\boldsymbol{Y}}$ *a* $f_{\boldsymbol{Z}}$ *platí, že* $f_{\boldsymbol{X}}(\boldsymbol{x}) = f_{\boldsymbol{Y}}(\boldsymbol{y})f_{\boldsymbol{Z}}(\boldsymbol{z})$ *pro všechny* $(\boldsymbol{y}, \boldsymbol{z})$.

(iv) *Náhodné vektory* $\boldsymbol{Y} \in \mathbb{R}^r$ *a* $\boldsymbol{Z} \in \mathbb{R}^s$ *jsou nezávislé právě tehdy, když pro momentové vytvořující funkce platí*

$$E\left[e^{\boldsymbol{t}^T\boldsymbol{X}}\right] = M_{\boldsymbol{X}}(\boldsymbol{t}) = M_{\boldsymbol{Y}}(\boldsymbol{t}_1)M_{\boldsymbol{Z}}(\boldsymbol{t}_2),$$

kde $\boldsymbol{t} = (\boldsymbol{t}_1^T, \boldsymbol{t}_2^T)^T$, $\boldsymbol{t}_1 \in \mathbb{R}^r$, $\boldsymbol{t}_2 \in \mathbb{R}^s$, *pokud momentové vytvořující funkce existují.*

(v) *Náhodné vektory* $\boldsymbol{Y} \in R^r$ *a* $\boldsymbol{Z} \in \mathbb{R}^s$ *jsou nezávislé právě tehdy, když pro charakteristické funkce platí*

$$E\left[e^{i\boldsymbol{t}^T\boldsymbol{X}}\right] = \Psi_{\boldsymbol{X}}(\boldsymbol{t}) = \Psi_{\boldsymbol{Y}}(\boldsymbol{t}_1)\Psi_{\boldsymbol{Z}}(\boldsymbol{t}_2),$$

kde $\boldsymbol{t} = (\boldsymbol{t}_1^T, \boldsymbol{t}_2^T)^T$, $\boldsymbol{t}_1 \in \mathbb{R}^r$, $\boldsymbol{t}_2 \in \mathbb{R}^s$.

Důkaz. Analogicky jako u vět 2.16, 2.17, 2.18 a 2.21. □

Věta 3.10 (Nezávislost funkcí náhodných vektorů)**.** *Nechť* $\boldsymbol{Y} \in \mathbb{R}^r$ *a* $\boldsymbol{Z} \in \mathbb{R}^s$ *jsou nezávislé náhodné vektory a funkce* $h : \mathbb{R}^r \longrightarrow \mathbb{R}^p$, $g : \mathbb{R}^s \longrightarrow \mathbb{R}^q$ *jsou měřitelné. Potom jsou náhodné vektory* $h(\boldsymbol{Y})$ *a* $g(\boldsymbol{Z})$ *také nezávislé.*

Důkaz. Stejný jako pro náhodné veličiny (věta 2.20). □

Důsledek 3.3. *Nechť* $\boldsymbol{Y} \in \mathbb{R}^r$ *a* $\boldsymbol{Z} \in \mathbb{R}^s$ *jsou nezávislé. Potom pro*

$$h(\boldsymbol{Y}) = \boldsymbol{a}^T\boldsymbol{Y}, \quad g(\boldsymbol{Z}) = \boldsymbol{b}^T\boldsymbol{Z},$$

kde $\boldsymbol{a} \in \mathbb{R}^r$ *a* $\boldsymbol{b} \in \mathbb{R}^s$, *jsou* $\boldsymbol{a}^T\boldsymbol{Y}$ *a* $\boldsymbol{b}^T\boldsymbol{Z}$ *nezávislé náhodné veličiny. Speciálně,*

(i) *pro případ* $\boldsymbol{a} = (0, \dots, 0, 1, 0, \dots)^T$ *s jedničkou na* i*-té pozici a* $\boldsymbol{b} = (0, \dots, 0, 1, 0, \dots, 0)^T$ *s jedničkou na* j*-té pozici platí, že pokud jsou vektory* $\boldsymbol{Y}$ *a* $\boldsymbol{Z}$ *nezávislé, potom náhodné veličiny* Y_i *a* Z_j *jsou nezávislé pro všechny dvojice* (i, j).

(ii) *pro případ*

$$\boldsymbol{a} = \left(\frac{1}{r}, \dots, \frac{1}{r}\right)^T, \quad \boldsymbol{b} = \left(\frac{1}{s}, \dots, \frac{1}{s}\right)^T$$

platí, že pokud jsou vektory $\boldsymbol{Y}$ *a* $\boldsymbol{Z}$ *nezávislé, potom náhodné veličiny*

$$\overline{Y} = \frac{1}{r}\sum_{i=1}^{r} Y_i, \quad \overline{Z} = \frac{1}{s}\sum_{j=1}^{s} Z_j$$

jsou nezávislé.

Poznámka. Definici 3.8 a věty 3.9, 3.10 lze rozšířit na libovolný konečný počet náhodných vektorů.

Věta 3.11 (Nezávislost a diagonální matice)**.** *Nechť má náhodný vektor* $\boldsymbol{X}$ *složky* $X_i, X_j \in L^2(\Omega)$ *po dvou nezávislé pro všechny dvojice* (i, j). *Potom jsou matice* $\mathrm{Var}[\boldsymbol{X}]$ *a* $\mathrm{cor}(\boldsymbol{X})$ *diagonální.*

Důkaz. Dle věty 3.5 platí, že $\mathrm{Var}[\boldsymbol{X}]$ má prvky $\mathrm{cov}(X_i, X_j)$. Dle bodu (i) důsledku 2.9 získáme, že pokud jsou X_i, X_j nezávislé, pak $\mathrm{cov}(X_i, X_j) = 0$ pro $i \neq j$. Dle věty 3.8 má $\mathrm{cor}(\boldsymbol{X})$ prvky $\mathrm{cor}(X_i, X_j)$. Dle bodu (iv) důsledku 2.9 platí, že pokud jsou X_i, X_j nezávislé, potom $\mathrm{cor}(X_i, X_j) = 0$ pro $i \neq j$. □

Poznámka. Ve větě 3.11 je též možné předpokládat $X_1, \dots, X_n$ po dvou nekorelované nebo $X_1, \dots, X_n$ nezávislé.

Věta 3.12 (Nezávislost a blokově diagonální matice)**.** *Nechť* $\boldsymbol{X} = (\boldsymbol{Y}^T, \boldsymbol{Z}^T)^T$*, kde* $\boldsymbol{Y}, \boldsymbol{Z}$ *jsou nezávislé a* $X_i \in L^2(\Omega)$ *pro všechna* $i = 1, \dots, n$*. Pak jsou matice* $\operatorname{Var}[\boldsymbol{X}]$ *a* $\operatorname{cor}(\boldsymbol{X})$ *blokově diagonální.*

Důkaz. Zřejmě

$$\operatorname{Var}[\boldsymbol{X}] = \begin{pmatrix} \operatorname{Var}[\boldsymbol{Y}] & \operatorname{cov}(\boldsymbol{Y}, \boldsymbol{Z}) \\ \operatorname{cov}(\boldsymbol{Z}, \boldsymbol{Y}) & \operatorname{Var}[\boldsymbol{Z}] \end{pmatrix}.$$

Dle bodu (i) věty 3.6 platí, že $\operatorname{cov}(\boldsymbol{Y}, \boldsymbol{Z})$ má prvky $\operatorname{cov}(Y_i, Z_j)$. Podle důsledku 3.3 máme, že pokud jsou $\boldsymbol{Y}, \boldsymbol{Z}$ nezávislé, potom také Y_i a Z_j jsou nezávislé pro všechny dvojice (i, j). Pomocí bodu (i) důsledku 2.9 lze zjistit, že $\operatorname{cov}(Y_i, Z_j) = 0$ pro všechny dvojice (i, j). Nakonec dle bodů (i) a (iii) věty 3.6 dostaneme, že $\operatorname{cov}(\boldsymbol{Y}, \boldsymbol{Z}) = \mathbf{0}_{r\times(n-r)}$ a $\operatorname{cov}(\boldsymbol{Z}, \boldsymbol{Y}) = \mathbf{0}_{(n-r)\times r}$.
Dále

$$\operatorname{cor}(\boldsymbol{X}) = \begin{pmatrix} \operatorname{cor}(\boldsymbol{Y}) & \operatorname{cor}(\boldsymbol{Y}, \boldsymbol{Z}) \\ \operatorname{cor}(\boldsymbol{Z}, \boldsymbol{Y}) & \operatorname{cor}(\boldsymbol{Z}) \end{pmatrix}.$$

Dle věty 3.8 platí, že $\operatorname{cor}(\boldsymbol{Y}, \boldsymbol{Z})$ má prvky $\operatorname{cor}(Y_i, Z_j)$. Podle důsledku 3.3 máme, že Y_i a Z_j jsou nezávislé pro všechny dvojice (i, j). Pomocí bodu (iv) důsledku 2.9 dostaneme, že pokud pro všechny dvojice (i, j) je $\operatorname{cor}(Y_i, Z_j) = 0$, potom $\operatorname{cor}(\boldsymbol{Y}, \boldsymbol{Z}) = \mathbf{0}_{r\times(n-r)}$ a $\operatorname{cor}(\boldsymbol{Z}, \boldsymbol{Y}) = \mathbf{0}_{(n-r)\times r}$. □

3.4. Podmíněné rozdělení

Úmluva. *V této sekci uvažujme náhodný vektor* $\boldsymbol{X} = (X_1, \dots, X_n)^T$*. Dále pro* $1 \le r < n$ *mějme náhodné vektory*

(i) $\boldsymbol{Y} = (X_1, \dots, X_n)^T = (Y_1, \dots, Y_r)^T$ *s hustotou* $f_{\boldsymbol{Y}}(\boldsymbol{y})$ *vzhledem k* σ*-konečné míře* μ_1 *na* $(\mathbb{R}^r, \mathcal{B}_0^r)$*;*

(ii) $\boldsymbol{Z} = (X_{r+1}, \dots, X_n)^T = (Z_1, \dots, Z_{n-r})^T$ *s hustotou* $f_{\boldsymbol{Z}}(\boldsymbol{z})$ *vzhledem k* σ*-konečné míře* μ_2 *na* $(\mathbb{R}^{n-r}, \mathcal{B}_0^{n-r})$*.*

Nechť $\boldsymbol{X}$ *má hustotu* $f_{\boldsymbol{X}}(\boldsymbol{x})$ *vzhledem k součinové míře* $\mu_1 \times \mu_2$*. (Tento předpoklad neplyne z existence marginálních hustot, je proto třeba ho uvést samostatně.) V této sekci se budeme zabývat rozdělením vektoru* $\boldsymbol{Y}$ *za podmínky, že vektor* $\boldsymbol{Z}$ *nabyl konkrétní hodnoty* $\boldsymbol{z} \in \mathbb{R}^{n-r}$*.*

Definice 3.9 (Podmíněná hustota)**.** *Podmíněná hustota* náhodného vektoru $\boldsymbol{Y}$ při pevném $\boldsymbol{Z} = \boldsymbol{z}$ je libovolná nezáporná měřitelná funkce, pro kterou platí

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

kde $B \in \mathcal{B}_0^r$ a $C \in \mathcal{B}_0^{n-r}$.

Poznámka. Podmíněná hustota $f(\boldsymbol{y} \mid \boldsymbol{z})$ za výše uvedených předpokladů existuje a je určena jednoznačně μ_1-skoro všude.

Věta 3.13 (Výpočet podmíněné hustoty)**.** *Pro podmíněnou hustotu platí*

$$f(\boldsymbol{y} \mid \boldsymbol{z}) = \begin{cases} \frac{f_{\boldsymbol{X}}(\boldsymbol{y}, \boldsymbol{z})}{f_{\boldsymbol{Z}}(\boldsymbol{z})} & \textit{pro } \boldsymbol{z} \in \mathbb{R}^{n-r} \textit{ taková, že } f_{\boldsymbol{Z}}(\boldsymbol{z}) \neq 0, \\ 0 & \textit{jinak.} \end{cases}$$

Důkaz. Podle poznámek za definicí 3.2 máme

$$P[\boldsymbol{Y} \in B, \boldsymbol{Z} \in C] = P[\boldsymbol{X} \in (B \times C)] = \int_{B\times C} f_{\boldsymbol{X}}(\boldsymbol{y}, \boldsymbol{z}) \, \mathrm{d}\mu(\boldsymbol{y}, \boldsymbol{z}),$$

kde $B \in \mathcal{B}_0^r$, $C \in \mathcal{B}_0^{n-r}$. Dále podle definice podmíněné hustoty (3.9) dostaneme

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

Aby se obě vyjádření rovnala, musí platit, že

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

μ-skoro všude. Tedy lze položit

$$f(\boldsymbol{y} \mid \boldsymbol{z}) = f_{\boldsymbol{X}}(\boldsymbol{y},\boldsymbol{z}) \left(f_{\boldsymbol{Z}}(\boldsymbol{z})\right)^{-1}$$

pro $f_{\boldsymbol{Z}}(\boldsymbol{z}) \neq 0$ a dodefinovat $f(\boldsymbol{y} \mid \boldsymbol{z}) = 0$ pro $f_{\boldsymbol{Z}}(\boldsymbol{z})$. □

Poznámka. Pro $\boldsymbol{Y}, \boldsymbol{Z}$ s diskrétním rozdělením dostaneme

$$P[\boldsymbol{Y} = \boldsymbol{y} \mid \boldsymbol{Z} = \boldsymbol{z}] = \frac{P[\boldsymbol{Y} = \boldsymbol{y}, \boldsymbol{Z} = \boldsymbol{z}]}{P[\boldsymbol{Z} = \boldsymbol{z}]}$$

pro $P[\boldsymbol{Z} = \boldsymbol{z}] \neq 0$. Zřejmě také platí, že

$$P[\boldsymbol{Y} \in B, \boldsymbol{Z} \in C] = \sum_{\boldsymbol{y} \in B} \sum_{\boldsymbol{z} \in C} P[\boldsymbol{Y} = \boldsymbol{y}, \boldsymbol{Z} = \boldsymbol{z}] = \sum_{\boldsymbol{z} \in C} \left(\sum_{\boldsymbol{y} \in B} P[\boldsymbol{Y} = \boldsymbol{y} \mid \boldsymbol{Z} = \boldsymbol{z}] \right) P[\boldsymbol{Z} = \boldsymbol{z}],$$

kde sčítáme přes všechny složky y_i a z_j, přičemž $i = 1, \ldots, r$ a $j = 1, \ldots, n-r$.

Věta 3.14 (Bayesova věta)**.** *Pro podmíněné hustoty $f(\boldsymbol{z} \mid \boldsymbol{y})$ a $f(\boldsymbol{y} \mid \boldsymbol{z})$ platí*

$$f(\boldsymbol{z} \mid \boldsymbol{y}) = \frac{f(\boldsymbol{y} \mid \boldsymbol{z}) f_{\boldsymbol{Z}}(\boldsymbol{z})}{\int_{\mathbb{R}^{n-r}} f(\boldsymbol{y} \mid \boldsymbol{z}) f_{\boldsymbol{Z}}(\boldsymbol{z}) \, \mathrm{d}\mu_2(\boldsymbol{z})}, \textit{ pokud jmenovatel } \neq 0,$$

$$f(\boldsymbol{z} \mid \boldsymbol{y}) = 0 \textit{ jinak.}$$

Důkaz. Dle věty 3.13 máme, že

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

pokud $f_{\boldsymbol{Z}}(\boldsymbol{z}) \neq 0$. Dále dle poznámek za větou 3.3 platí, že

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \int_{\mathbb{R}^{n-r}} f_{\boldsymbol{X}}(\boldsymbol{y},\boldsymbol{z}) \, \mathrm{d}\mu_2(\boldsymbol{z}).$$

Nakonec vyjádříme $f(\boldsymbol{z} \mid \boldsymbol{y})$ dle věty 3.13. □

Poznámka. Pro $\boldsymbol{Y}, \boldsymbol{Z}$ s diskrétním rozdělením dostáváme

$$P[\boldsymbol{Z} = \boldsymbol{z} \mid \boldsymbol{Y} = \boldsymbol{y}] = \frac{P[\boldsymbol{Y} = \boldsymbol{y} \mid \boldsymbol{Z} = \boldsymbol{z}] P[\boldsymbol{Z} = \boldsymbol{z}]}{\sum_{\boldsymbol{z}} P[\boldsymbol{Y} = \boldsymbol{y} \mid \boldsymbol{Z} = \boldsymbol{z}] P[\boldsymbol{Z} = \boldsymbol{z}]}.$$

Definice 3.10 (Podmíněná střední hodnota)**.** Nechť $h(\boldsymbol{y},\boldsymbol{z}) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ je měřitelná funkce a položme

$$\boldsymbol{H} = h(\boldsymbol{Y},\boldsymbol{Z}) = h(\boldsymbol{X}) \in \mathbb{R}^m.$$

Potom *podmíněná střední hodnota* náhodného vektoru $\boldsymbol{H}$ při pevném $\boldsymbol{Z} = \boldsymbol{z}$ je definována předpisem

$$E[\boldsymbol{H} \mid \boldsymbol{Z} = \boldsymbol{z}] = \int_{\mathbb{R}^r} h(\boldsymbol{y},\boldsymbol{z}) f(\boldsymbol{y} \mid \boldsymbol{z}) \, \mathrm{d}\mu_1(\boldsymbol{y}),$$

pokud integrál existuje. Označíme-li

$$\varphi(\boldsymbol{z}) = E[\boldsymbol{H} \mid \boldsymbol{Z} = \boldsymbol{z}], \; \boldsymbol{z} \in \mathbb{R}^{n-r},$$

nazýváme náhodný vektor $\varphi(\boldsymbol{Z}) \in \mathbb{R}^m$ *podmíněnou střední hodnotou* vektoru $\boldsymbol{H}$ při *pevném, ale nespecifikovaném* $\boldsymbol{Z}$, a píšeme $\varphi(\boldsymbol{Z}) = E[\boldsymbol{H} \mid \boldsymbol{Z}]$.

Poznámky.

(i) Podmíněná střední hodnota je definována po složkách. Je-li

$$\boldsymbol{H} = (H_1, \dots, H_m)^T = (h_1(\boldsymbol{Y}, \boldsymbol{Z}), \dots, h_m(\boldsymbol{Y}, \boldsymbol{Z}))^T, \quad h_i : \mathbb{R}^n \longrightarrow \mathbb{R}, \quad i = 1, \dots, m,$$

pak i-tá složka vektoru $E[\boldsymbol{H} \mid \boldsymbol{Z} = \boldsymbol{z}]$ je rovna

$$\int_{\mathbb{R}^r} h_i(\boldsymbol{y}, \boldsymbol{z}) f(\boldsymbol{y} \mid \boldsymbol{z}) \,\mathrm{d}\mu_1(\boldsymbol{y}), \quad i = 1, \dots, m.$$

(ii) Podmíněná střední hodnota je funkcí podmínky. Při konkrétně dané podmínce $\boldsymbol{Z} = \boldsymbol{z} \in \mathbb{R}^{n-r}$ je

$$E[\boldsymbol{H} \mid \boldsymbol{Z} = \boldsymbol{z}]$$

konstanta z $\mathbb{R}^m$. Při nespecifikované podmínce $\boldsymbol{Z}$ je

$$E[\boldsymbol{H} \mid \boldsymbol{Z}] = (E[H_1 \mid \boldsymbol{Z}], \dots, E[H_m \mid \boldsymbol{Z}])^T$$

náhodný vektor (pro $m = 1$ náhodná veličina).

(iii) Podmíněná střední hodnota $E[\boldsymbol{H} \mid \boldsymbol{Z}]$ odstranila z $\boldsymbol{H} = h(\boldsymbol{Y}, \boldsymbol{Z})$ náhodnost v $\boldsymbol{Y}$, ale ponechala náhodnost v $\boldsymbol{Z}$.

(iv) Podmíněná střední hodnota je určena jednoznačně μ_2-skoro všude.

(v) V teorii pravděpodobnosti se zavádí obecná abstraktní definice podmíněné střední hodnoty, která nevyžaduje existenci podmíněné hustoty. Například $E[\boldsymbol{Z} \mid \boldsymbol{Z}]$ nelze spočítat dle definice 3.10.

(vi) Pro diskrétně rozdělené vektory $\boldsymbol{Y}, \boldsymbol{Z}$ dostáváme podmíněnou střední hodnotu ve tvaru

$$E[\boldsymbol{H} \mid \boldsymbol{Z} = \boldsymbol{z}] = \sum_{y_1} \cdots \sum_{y_r} h(\boldsymbol{y}, \boldsymbol{z}) P[\boldsymbol{Y} = \boldsymbol{y} \mid \boldsymbol{Z} = \boldsymbol{z}],$$

kde i-tá složka je

$$E[H_i \mid \boldsymbol{Z} = \boldsymbol{z}] = \sum_{y_1} \cdots \sum_{y_r} h_i(\boldsymbol{y}, \boldsymbol{z}) P[\boldsymbol{Y} = \boldsymbol{y} \mid \boldsymbol{Z} = \boldsymbol{z}]$$

pro $i = 1, \dots, m$.

Věta 3.15 (Vlastnosti podmíněné střední hodnoty)**.** *Nechť $h : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, $g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, $p : \mathbb{R}^{n-r} \longrightarrow \mathbb{R}$ jsou měřitelné funkce. Označme*

$$\boldsymbol{H} = h(\boldsymbol{Y}, \boldsymbol{Z}) = (H_1, \dots, H_m)^T, \quad \boldsymbol{G} = g(\boldsymbol{Y}, \boldsymbol{Z}) = (G_1, \dots, G_m)^T$$

a předpokládejme, že $H_i \in L^1(\Omega)$ pro $i = 1, \dots, m$. Potom platí

(i) $E[\boldsymbol{a} \mid \boldsymbol{Z}] = \boldsymbol{a}$ *skoro jistě,* $\boldsymbol{a} \in \mathbb{R}^m$,

(ii) $E[E[\boldsymbol{H} \mid \boldsymbol{Z}]] = E[\boldsymbol{H}]$,

(iii) $E[a\boldsymbol{H} + b\boldsymbol{G} \mid \boldsymbol{Z}] = aE[\boldsymbol{H} \mid \boldsymbol{Z}] + bE[\boldsymbol{G} \mid \boldsymbol{Z}]$ *pro* $a, b \in \mathbb{R}$,

(iv) $E[p(\boldsymbol{Z})\boldsymbol{H} \mid \boldsymbol{Z}] = p(\boldsymbol{Z})E[\boldsymbol{H} \mid \boldsymbol{Z}]$,

pokud všechny podmíněné střední hodnoty existují.

Důkaz.

(i) Dle definice 3.10 máme

$$E[\boldsymbol{a} \mid \boldsymbol{Z} = \boldsymbol{z}] = \int_{\mathbb{R}^r} \boldsymbol{a} f(\boldsymbol{y} \mid \boldsymbol{z}) \,\mathrm{d}\mu_1(\boldsymbol{y}).$$

Pro i-tou složku platí podle vět 3.3 a 3.13

$$E[a_i \mid \boldsymbol{Z} = \boldsymbol{z}] = a_i \int_{\mathbb{R}^r} \frac{f_{\boldsymbol{X}}(\boldsymbol{y}, \boldsymbol{z})}{f_{\boldsymbol{Z}}(\boldsymbol{z})} \,\mathrm{d}\mu_1(\boldsymbol{y}) = \frac{a_i f_{\boldsymbol{Z}}(\boldsymbol{z})}{f_{\boldsymbol{Z}}(\boldsymbol{z})} = a_i,$$

za podmínky, že $f_{\boldsymbol{Z}}(\boldsymbol{z}) \neq 0$. Označme $N = \{\boldsymbol{z} \in \mathbb{R}^{n-r} : f_{\boldsymbol{Z}}(\boldsymbol{z}) = 0.\}$. Dle poznámek za definicemi 2.7 a 3.2 dostaneme, že

$$P_{\boldsymbol{Z}}(N) = P[\boldsymbol{Z} \in N] = \int_N f_{\boldsymbol{Z}}(\boldsymbol{z})\,\mathrm{d}\mu_2(\boldsymbol{z}) = 0,$$

z čehož plyne

$$P_{\boldsymbol{Z}}\left(\left\{\boldsymbol{z} \in \mathbb{R}^{n-r} : f_{\boldsymbol{Z}}(\boldsymbol{z}) \neq 0\right\}\right) = P_{\boldsymbol{Z}}\left(\left\{\boldsymbol{z} \in \mathbb{R}^{n-r} : E[\boldsymbol{a} \mid \boldsymbol{Z} = \boldsymbol{z}] = \boldsymbol{a}\right\}\right) = 1,$$

tedy

$$P[E[\boldsymbol{a} \mid \boldsymbol{Z}] = \boldsymbol{a}] = 1,$$

což znamená, že $E[\boldsymbol{a} \mid \boldsymbol{Z}] = \boldsymbol{a}$ skoro jistě.

(ii) Podle poznámky (ii) za definicí 3.2 a poznámky (i) za definicí 3.4 lze psát

$$E[\boldsymbol{H}] = \int_{\mathbb{R}^n} h(\boldsymbol{y},\boldsymbol{z}) f_{\boldsymbol{X}}(\boldsymbol{y},\boldsymbol{z})\,\mathrm{d}\mu(\boldsymbol{y},\boldsymbol{z}),$$

kde $\mu = \mu_1 \times \mu_2$. Podle věty 3.13 a Fubiniovy věty dostaneme

$$\begin{aligned} E[\boldsymbol{H}] &= \int_{\mathbb{R}^{n-r}} \left(\int_{\mathbb{R}^r} h(\boldsymbol{y},\boldsymbol{z}) f(\boldsymbol{y} \mid \boldsymbol{z})\,\mathrm{d}\mu_1(\boldsymbol{y}) \right) f_{\boldsymbol{Z}}(\boldsymbol{z})\,\mathrm{d}\mu_2(\boldsymbol{z}) \\ &= \int_{\mathbb{R}^{n-r}} \varphi(\boldsymbol{z}) f_{\boldsymbol{Z}}(\boldsymbol{z})\,\mathrm{d}\mu_2(\boldsymbol{z}) \\ &= E[\varphi(\boldsymbol{Z})] = E[E[\boldsymbol{H} \mid \boldsymbol{Z}]]. \end{aligned}$$

Ve druhé rovnosti jsme použili vztah z definice 3.10.

(iii) Plyne z toho, že integrál je lineární funkcionál.

(iv) Počítejme:

$$E[p(\boldsymbol{Z})\boldsymbol{H} \mid \boldsymbol{Z} = \boldsymbol{z}] = \int_{\mathbb{R}^r} p(\boldsymbol{z}) h(\boldsymbol{y},\boldsymbol{z}) f(\boldsymbol{y} \mid \boldsymbol{z})\,\mathrm{d}\mu_1(\boldsymbol{y}) = p(\boldsymbol{z}) E[\boldsymbol{H} \mid \boldsymbol{Z} = \boldsymbol{z}],$$

z čehož plyne, že

$$E[p(\boldsymbol{Z})\boldsymbol{H} \mid \boldsymbol{Z}] = p(\boldsymbol{Z}) E[\boldsymbol{H} \mid \boldsymbol{Z}]. \qquad \square$$

Důsledek 3.4. *Pro měřitelné funkce* $p : \mathbb{R}^{n-r} \longrightarrow \mathbb{R}^m$, $h : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *platí při existenci podmíněné střední hodnoty* $E[\boldsymbol{H} \mid \boldsymbol{Z}]$

(i) $E[\boldsymbol{P}\boldsymbol{H}^T \mid \boldsymbol{Z}] = \boldsymbol{P} E[\boldsymbol{H}^T \mid \boldsymbol{Z}]$,

(ii) $E[\boldsymbol{H}\boldsymbol{P}^T \mid \boldsymbol{Z}] = E[\boldsymbol{H} \mid \boldsymbol{Z}]\boldsymbol{P}^T$.

Důkaz. Označme

$$\begin{aligned} \boldsymbol{P} &= (P_1, \dots, P_m)^T = (p_1(\boldsymbol{Z}), \dots, p_m(\boldsymbol{Z}))^T, \\ \boldsymbol{H} &= (H_1, \dots, H_m)^T = (h_1(\boldsymbol{Y},\boldsymbol{Z}), \dots, h_m(\boldsymbol{Y},\boldsymbol{Z}))^T. \end{aligned}$$

Potom $\boldsymbol{P}\boldsymbol{H}^T$ a $\boldsymbol{H}\boldsymbol{P}^T$ jsou matice typu $m \times m$. Dále $\boldsymbol{P}\boldsymbol{H}^T$ má na místě (i,j) prvek $P_i H_j = p_i(\boldsymbol{Z}) h_j(\boldsymbol{Y},\boldsymbol{Z})$ a máme

$$E[p_i(\boldsymbol{Z}) h_j(\boldsymbol{Y},\boldsymbol{Z}) \mid \boldsymbol{Z}] = p_i(\boldsymbol{Z}) E[H_j \mid \boldsymbol{Z}].$$

Z toho plyne, že

$$E[\boldsymbol{P}\boldsymbol{H}^T \mid \boldsymbol{Z}] = \boldsymbol{P} E[\boldsymbol{H}^T \mid \boldsymbol{Z}].$$

Analogicky pro druhý případ, kde matice $\boldsymbol{H}\boldsymbol{P}^T$ má na místě (i,j) prvek $H_i P_j$. $\square$

Poznámka. Mějme měřitelné funkce $p : \mathbb{R}^{n-r} \longrightarrow \mathbb{R}^m$, $h : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, pro něž platí, že

$$\begin{aligned} \boldsymbol{P} &= (P_1, \ldots, P_m)^T = (p_1(\boldsymbol{Z}), \ldots, p_m(\boldsymbol{Z}))^T, \\ \boldsymbol{H} &= (H_1, \ldots, H_m)^T = (h_1(\boldsymbol{Y}, \boldsymbol{Z}), \ldots, h_m(\boldsymbol{Y}, \boldsymbol{Z}))^T, \end{aligned}$$

kde $H_i, P_i \in L^2(\Omega)$ pro všechna $i = 1, \ldots, m$. Potom pro tyto funkce platí

$$\operatorname{Var}[\boldsymbol{H} - \boldsymbol{P}] - \operatorname{Var}[\boldsymbol{H} - E[\boldsymbol{H} \mid \boldsymbol{Z}]] \geq 0,$$

což znamená, že rozdíl variančních matic je pozitivně semidefinitní. Speciálně pro $m = 1$ máme

$$\begin{aligned} \operatorname{Var}[H - P] \geq \operatorname{Var}[H - E[H \mid \boldsymbol{Z}]] &= E[H - E[H \mid \boldsymbol{Z}]]^2 - (E[H] - E[E[H \mid \boldsymbol{Z}]])^2 \\ &= E[H - E[H \mid \boldsymbol{Z}]]^2, \end{aligned}$$

přičemž poslední rovnost plyne z věty 3.15.
Získali jsme tedy interpretaci, že podmíněná střední hodnota $E[H \mid \boldsymbol{Z}]$ minimalizuje vzdálenost $P = p(\boldsymbol{Z})$ od $H = h(\boldsymbol{Y}, \boldsymbol{Z})$ měřenou rozptylem rozdílu. Jinak řečeno ze všech funkcí $P = p(\boldsymbol{Z})$ je náhodné veličině $H = h(\boldsymbol{Y}, \boldsymbol{Z})$ nejblíže její podmíněná střední hodnota při pevném $\boldsymbol{Z}$, to jest $E[H \mid \boldsymbol{Z}]$.
Speciálně pro $r = 1$ a $H = h(Y, \boldsymbol{Z}) = Y$ je $E[Y \mid \boldsymbol{Z}]$ nejlepší aproximací náhodné veličiny Y pomocí náhodného vektoru $\boldsymbol{Z} \in \mathbb{R}^{n-1}$ ve smyslu minimalizace rozptylu jejich rozdílu. Tento poznatek se využívá ve statistických regresních modelech.

Definice 3.11 (Podmíněná pravděpodobnost)**.** *Podmíněnou pravděpodobnost* definujeme předpisem

$$P[\boldsymbol{X} \in B \mid \boldsymbol{Z} = \boldsymbol{z}] = \int_{\mathbb{R}^r} \mathbb{I}_B(\boldsymbol{y}, \boldsymbol{z}) f(\boldsymbol{y} \mid \boldsymbol{z}) \, \mathrm{d}\mu_1(\boldsymbol{y})$$

pro $B \in \mathcal{B}_0^n$, pokud integrál existuje.

Poznámka. Máme $\mathbb{I}_B(\boldsymbol{y}, \boldsymbol{z}) = 1$ pro $\boldsymbol{x} = (\boldsymbol{y}^T, \boldsymbol{z}^T)^T \in B$ a $\mathbb{I}_B(\boldsymbol{y}, \boldsymbol{z}) = 0$ jinak. Zřejmě platí

$$P[\boldsymbol{X} \in B \mid \boldsymbol{Z} = \boldsymbol{z}] = E[\mathbb{I}_B(\boldsymbol{Y}, \boldsymbol{Z}) \mid \boldsymbol{Z} = \boldsymbol{z}].$$

Pro porovnání připomeňme

$$P[\boldsymbol{X} \in B] = \int_{\mathbb{R}^n} \mathbb{I}_B(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) \, \mathrm{d}\mu(\boldsymbol{x}) = E[\mathbb{I}_B(\boldsymbol{Y}, \boldsymbol{Z})].$$

Definice 3.12 (Podmíněný rozptyl)**.** Nechť $\boldsymbol{H} = h(\boldsymbol{Y}, \boldsymbol{Z}) = (H_1, \ldots, H_m)^T$. Potom *podmíněný rozptyl náhodného vektoru $\boldsymbol{H}$ při pevném $\boldsymbol{Z}$* je definován předpisem

$$\operatorname{Var}[\boldsymbol{H} \mid \boldsymbol{Z}] = E\left[(\boldsymbol{H} - E[\boldsymbol{H} \mid \boldsymbol{Z}])(\boldsymbol{H} - E[\boldsymbol{H} \mid \boldsymbol{Z}])^T \mid \boldsymbol{Z}\right],$$

pokud podmíněné střední hodnoty existují.

Poznámka. Podmíněný rozptyl je tedy náhodná matice typu $m \times m$. Pro $m = 1$ je to náhodná veličina $E[(H - E[H \mid \boldsymbol{Z}])^2 \mid \boldsymbol{Z}]$. Rozptyl (varianční matice) náhodného vektoru $\boldsymbol{H}$ lze rozložit pomocí podmíněného rozptylu a podmíněné střední hodnoty.

Věta 3.16 (Rozklad varianční matice)**.** *Za předpokladů z definice 3.12 lze psát varianční matici náhodného vektoru $\boldsymbol{H} = h(\boldsymbol{Y}, \boldsymbol{Z})$ ve tvaru*

$$\operatorname{Var}[\boldsymbol{H}] = E[\operatorname{Var}[\boldsymbol{H} \mid \boldsymbol{Z}]] + \operatorname{Var}[E[\boldsymbol{H} \mid \boldsymbol{Z}]].$$

Důkaz. S využitím bodu (iii) věty 3.15 a důsledku 3.4 počítejme:

$$\begin{aligned} \operatorname{Var}[\boldsymbol{H} \mid \boldsymbol{Z}] &= E\left[((\boldsymbol{H} - E[\boldsymbol{H}]) + (E[\boldsymbol{H}] - E[\boldsymbol{H} \mid \boldsymbol{Z}]))((\boldsymbol{H} - E[\boldsymbol{H}]) + (E[\boldsymbol{H}] - E[\boldsymbol{H} \mid \boldsymbol{Z}]))^T \mid \boldsymbol{Z}\right] \\ &= E[(\boldsymbol{H} - E[\boldsymbol{H}])(\boldsymbol{H} - E[\boldsymbol{H}])^T \mid \boldsymbol{Z}] \\ &+ E[(\boldsymbol{H} - E[\boldsymbol{H}]) \mid \boldsymbol{Z}] \cdot (E[\boldsymbol{H}] - E[\boldsymbol{H} \mid \boldsymbol{Z}])^T \\ &+ (E[\boldsymbol{H}] - E[\boldsymbol{H} \mid \boldsymbol{Z}]) \cdot E[(\boldsymbol{H} - E[\boldsymbol{H}])^T \mid \boldsymbol{Z}] \\ &+ (E[\boldsymbol{H}] - E[\boldsymbol{H} \mid \boldsymbol{Z}]) \cdot (E[\boldsymbol{H}] - E[\boldsymbol{H} \mid \boldsymbol{Z}])^T \cdot E[1 \mid \boldsymbol{Z}] \\ &= E[(\boldsymbol{H} - E[\boldsymbol{H}])(\boldsymbol{H} - E[\boldsymbol{H}])^T \mid \boldsymbol{Z}] - (E[\boldsymbol{H}] - E[\boldsymbol{H} \mid \boldsymbol{Z}]) \cdot (E[\boldsymbol{H}] - E[\boldsymbol{H} \mid \boldsymbol{Z}])^T. \end{aligned}$$

Po aplikaci střední hodnoty dostaneme s použitím bodu (ii) věty 3.15

$$\begin{aligned} E[\operatorname{Var}[\boldsymbol{H} \mid \boldsymbol{Z}]] &= E[(\boldsymbol{H}-E[\boldsymbol{H}])(\boldsymbol{H}-E[\boldsymbol{H}])^T] - E\left[(E[\boldsymbol{H} \mid \boldsymbol{Z}] - E[E[\boldsymbol{H} \mid \boldsymbol{Z}]]) \cdot (E[\boldsymbol{H} \mid \boldsymbol{Z}] - E[E[\boldsymbol{H} \mid \boldsymbol{Z}]])^T\right] \\ &= \operatorname{Var}[\boldsymbol{H}] - \operatorname{Var}[E[\boldsymbol{H} \mid \boldsymbol{Z}]]. \end{aligned}$$ □

Poznámka. Obdobně jako podmíněný rozptyl lze zavést *podmíněnou kovarianci*. Nechť

$$\boldsymbol{H} = h(\boldsymbol{Y}, \boldsymbol{Z}) = (H_1, \dots, H_m)^T, \quad \boldsymbol{G} = g(\boldsymbol{Y}, \boldsymbol{Z}) = (G_1, \dots, G_p)^T.$$

Potom

$$\operatorname{cov}(\boldsymbol{H}, \boldsymbol{G} \mid \boldsymbol{Z}) = E[(\boldsymbol{H} - E[\boldsymbol{H} \mid \boldsymbol{Z}])(\boldsymbol{G} - E[\boldsymbol{G} \mid \boldsymbol{Z}])^T \mid \boldsymbol{Z}]$$

a zároveň

$$\operatorname{cov}(\boldsymbol{H}, \boldsymbol{G}) = E[\operatorname{cov}(\boldsymbol{H}, \boldsymbol{G} \mid \boldsymbol{Z})] + \operatorname{cov}(E[\boldsymbol{H} \mid \boldsymbol{Z}], E[\boldsymbol{G} \mid \boldsymbol{Z}]),$$

pokud podmíněné střední hodnoty existují.

Příklad 3.1 (Podmíněné a nepodmíněné momenty)**.** Nechť $X_1 \sim \mathrm{N}(\mu, \sigma^2)$ a $X_2 \sim \mathrm{N}(\mu, \sigma^2)$, kde $\sigma^2 > 0$, jsou při pevném μ nezávislé, přičemž obecně je $\mu \sim \mathrm{N}(\mu_0, \sigma_0^2)$, kde $\sigma_0^2 > 0$. Zřejmě $f(x \mid \mu)$ je hustota $\mathrm{N}(\mu, \sigma^2)$.

(i) Podmíněné momenty:

$$\begin{aligned} E[X_i \mid \mu] &= \int_{-\infty}^{\infty} x f(x \mid \mu)\,\mathrm{d}x = \mu; \\ \operatorname{Var}[X_i \mid \mu] &= E[(X_i - E[X_i \mid \mu])^2 \mid \mu] = E[(X_i - \mu)^2 \mid \mu] = \sigma^2; \\ \operatorname{cov}(X_1, X_2 \mid \mu) &= E[(X_1 - E[X_1 \mid \mu])(X_2 - E[X_2 \mid \mu]) \mid \mu] = E[(X_1 - \mu)(X_2 - \mu) \mid \mu] = 0. \end{aligned}$$

(ii) Nepodmíněné momenty:

$$\begin{aligned} E[X_i] &= E[E[X_i \mid \mu]] = E[\mu] = \mu_0; \\ \operatorname{Var}[X_i] &= E[\operatorname{Var}[X_i \mid \mu]] + \operatorname{Var}[E[X_i \mid \mu]] = E[\sigma^2] + \operatorname{Var}[\mu] = \sigma^2 + \sigma_0^2; \\ \operatorname{cov}(X_1, X_2) &= \operatorname{cov}(E[X_1 \mid \mu], E[X_2 \mid \mu]) + E[\operatorname{cov}(X_1, X_2 \mid \mu)] = \operatorname{cov}(\mu, \mu) + E[0] = \operatorname{Var}[\mu] = \sigma_0^2. \end{aligned}$$

Zřejmě také

$$\rho_{X_1, X_2} = \frac{\operatorname{cov}(X_1, X_2)}{\sqrt{\operatorname{Var}[X_1] \operatorname{Var}[X_2]}} = \frac{\sigma_0^2}{\operatorname{Var}[X_1]} = \frac{\sigma_0^2}{\sigma^2 + \sigma_0^2}.$$

Je-li σ^2 výrazně menší než σ_0^2, pak $\rho_{X_1, X_2} \approx 1$, a tedy X_1, X_2 jsou silně korelované, ačkoli podmíněně při pevném μ jsou nezávislé, a tedy nekorelované. ▲

3.5. Přehled mnohorozměrných rozdělení

V této sekci si uvedeme přehled diskrétních a spojitých vícerozměrných rozdělení. V první části se zaměříme na *diskrétní rozdělení.*

1. Rozdělení dané kontingenční tabulkou
   Mějme náhodný vektor $\boldsymbol{X} = (Y, Z)^T$. Potom jeho kontingenční tabulku lze zapsat následujícím způsobem:

		Z			
		z_1	$\dots$	z_s	
Y	y_1	p_{11}	$\dots$	p_{1s}	$p_{1\cdot}$
	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
	y_r	p_{r1}	$\dots$	p_{rs}	$p_{r\cdot}$
		$p_{\cdot 1}$	$\dots$	$p_{\cdot s}$	1

Sdružené pravděpodobnosti:

$$p_{ij} = P[Y = y_i, Z = z_j], \quad i = 1, \dots, r;\ j = 1, \dots, s,$$
$$\sum_{i=1}^{r} \sum_{j=1}^{s} p_{ij} = 1.$$

Marginální pravděpodobnosti:

$$p_{i\cdot} = P[Y = y_i] = \sum_{j=1}^{s} p_{ij}, \quad i = 1, \dots, r,$$
$$p_{\cdot j} = P[Z = z_j] = \sum_{i=1}^{r} p_{ij}, \quad j = 1, \dots, s,$$
$$\sum_{i=1}^{r} p_{i\cdot} = 1,$$
$$\sum_{j=1}^{s} p_{\cdot j} = 1.$$

Podmíněné pravděpodobnosti:

$$p_{i|j} = P[Y = y_i \mid Z = z_j] = \frac{p_{ij}}{p_{\cdot j}}, \quad i = 1, \dots, r,$$
$$p_{j|i} = P[Z = z_j \mid Y = y_i] = \frac{p_{ij}}{p_{i\cdot}}, \quad j = 1, \dots, s,$$
$$\sum_{i=1}^{r} p_{i|j} = \frac{1}{p_{\cdot j}} \sum_{i=1}^{r} p_{ij} = \frac{1}{p_{\cdot j}} p_{\cdot j} = 1,$$
$$\sum_{j=1}^{s} p_{j|i} = \frac{1}{p_{i\cdot}} \sum_{j=1}^{s} p_{ij} = \frac{1}{p_{i\cdot}} p_{i\cdot} = 1.$$

2. Multinomické rozdělení s parametry $n, p_1, \dots, p_k$
   Jde o zobecnění binomického rozdělení. Modeluje počet výsledků typu i, $1 \leq i \leq k$, v sérii n nezávislých náhodných pokusů, z nichž každý má k možných výsledků, kde i-tý výsledek má pravděpodobnost p_i.
   Často se motivuje následující situací: Máme nádobu, v ní kuličky k barev a provedeme n tahů s vracením. Potom nechť $\boldsymbol{X} = (X_1, \dots, X_k)^T$, kde X_i je počet kuliček i-té barvy v n tazích a p_i je pravděpodobnost vytažení i-té barvy v jednom tahu.
   Zřejmě
   $$0 \leq X_i \leq n, \quad \sum_{i=1}^{k} X_i = n, \quad 0 < p_i < 1, \quad \sum_{i=1}^{k} p_i = 1.$$
   Sdružené pravděpodobnosti:
   $$P[X_1 = x_1, \dots, X_k = x_k] = \binom{n}{x_1}\binom{n - x_1}{x_2} \cdots \binom{n - x_1 - \cdots - x_{k-1}}{x_k} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}$$
   $$= \frac{n!}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k}$$
   pro $0 \leq x_i \leq n$, $\sum_{i=1}^{k} x_i = n$.
   Marginální pravděpodobnosti jsou rovněž multinomické. Nechť dále
   $$\boldsymbol{X} = (X_1, \dots, X_k)^T, \quad \boldsymbol{Y} = (X_1, \dots, X_r)^T, \quad \boldsymbol{Z} = (X_{r+1}, \dots, X_k)^T,$$

kde $1 \leq r < k$. Uvažujme barvy $r+1, \dots, k, \{1, \dots, r\}$. Pak počet kuliček s barvou z $\{1, \dots, r\}$ je

$$n - \sum_{i=r+1}^{k} X_i = \sum_{i=1}^{r} X_i$$

a pravděpodobnost vytažení kuličky s barvou z $\{1, \dots, r\}$ je rovna

$$1 - \sum_{i=r+1}^{k} p_i = \sum_{i=1}^{r} p_i.$$

Dále máme

$$\begin{aligned} P[X_{r+1} = x_{r+1}, \dots, X_k = x_k] &= P\left[X_{r+1} = x_{r+1}, \dots, X_k = x_k, \sum_{i=1}^{r} X_i = \sum_{i=1}^{r} x_i\right] \\ &= \frac{n!}{x_{r+1}! \cdots x_k! \left(\sum_{i=1}^{r} x_i\right)!} p_{r+1}^{x_{r+1}} \cdots p_k^{x_k} \left(\sum_{i=1}^{r} p_i\right)^{\sum_{i=1}^{r} x_i}. \end{aligned}$$

Pro $k = 2, r = 1$ dostáváme

$$P[X_2 = x_2] = \frac{n!}{x_2!(n - x_2)!} p_2^{x_2} (1 - p_2)^{n - x_2},$$

takže $X_2 \sim \mathrm{Bi}(n, p_2)$. Permutováním složek vektoru $\boldsymbol{X}$ dostaneme

$$X_i \sim \mathrm{Bi}(n, p_i), \quad E[X_i] = np_i, \quad \mathrm{Var}[X_i] = np_i(1 - p_i),$$

kde $i = 1, \dots, k$. Zřejmě

$$X_i = \sum_{t=1}^{n} \xi_{ti}, \quad \xi_{ti} \sim \mathrm{Alt}(p_i),$$

přičemž $\xi_{ti} = 1$, když v tahu t byla vytažena barva i, jinak je rovno nule. Dále $\mathrm{cov}(\xi_{ti}, \xi_{sj}) = 0$ pro $t \neq s$, neboť tahy jsou nezávislé, a

$$\mathrm{cov}(\xi_{ti}, \xi_{tj}) = E[\xi_{ti}\xi_{tj}] - E[\xi_{ti}]E[\xi_{tj}] = -E[\xi_{ti}]E[\xi_{tj}] = -p_i p_j.$$

Druhá rovnost plyne (pro $i \neq j$) z toho, že v tahu t nelze vytáhnout barvu i a barvu j. Potom podle důsledku 3.2

$$\mathrm{cov}(X_i, X_j) = \mathrm{cov}\left(\sum_{t=1}^{n} \xi_{ti}, \sum_{s=1}^{n} \xi_{sj}\right) = \sum_{t=1}^{n} \sum_{s=1}^{n} \mathrm{cov}(\xi_{ti}, \xi_{sj}) = \sum_{t=1}^{n} \mathrm{cov}(\xi_{ti}, \xi_{tj}) = -np_i p_j.$$

Kovariance je záporná, neboť při pevných X_l, kde $l \neq i, j$, platí, že zvýší-li se X_i, musí se snížit X_j vzhledem k tomu, že $\sum_{l=1}^{n} X_l = n$.
Pro náhodný vektor $\boldsymbol{X}$ s multinomickým rozdělením tedy platí:

(i) $E[\boldsymbol{X}]$ má složky $E[X_i] = np_i$, $i = 1, \dots, k$.

(ii) $\mathrm{Var}[\boldsymbol{X}]$ má diagonální prvky

$$\mathrm{Var}[X_i] = np_i(1 - p_i), \quad i = 1, \dots, k$$

a mimodiagonální prvky

$$\mathrm{cov}(X_i, X_j) = -np_i p_j, \quad i \neq j.$$

Podmíněné pravděpodobnosti jsou také multinomické:

$$\left(\frac{n!}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k}\right) \left(\left(\frac{n!}{x_{r+1}! \cdots x_k! \left(n - \sum_{i=r+1}^{k} x_i\right)!}\right) p_{r+1}^{x_{r+1}} \cdots p_k^{x_k} \left(1 - \sum_{i=r+1}^{k} p_i\right)^{n - \sum_{i=r+1}^{k} x_i}\right)^{-1}$$

$$= \frac{\left(\sum_{i=1}^{r} x_i\right)! p_1^{x_1} \cdots p_r^{x_r}}{x_1! \cdots x_r! \left(\sum_{i=1}^{r} p_i\right)^{\sum_{i=1}^{r} x_i}} = \left(\frac{\left(\sum_{i=1}^{r} x_i\right)!}{x_1! \cdots x_r!}\right) \prod_{j=1}^{r} \left(\frac{p_j}{\sum_{i=1}^{r} p_i}\right)^{x_j} = P[\boldsymbol{Y} = \boldsymbol{y} \mid \boldsymbol{Z} = \boldsymbol{z}],$$

což je multinomické rozdělení s parametry

$$\sum_{i=1}^{r} x_i, \quad \frac{p_j}{\sum_{i=1}^{r} p_i}, \quad j = 1, \dots, r.$$

Ve druhé části této sekce si uvedeme příklad *spojitých rozdělení.*

1. Mnohorozměrné normální rozdělení s parametry $\boldsymbol{\mu}, \boldsymbol{\Sigma}$
Toto rozdělení lze zkonstruovat z nezávislých náhodných veličin s rozdělením $\mathrm{N}(0,1)$. Uvažujme, že náhodné veličiny $\xi_1, \dots, \xi_k \sim \mathrm{N}(0,1)$ jsou nezávislé, a položme $\boldsymbol{\xi} = (\xi_1, \dots, \xi_k)^T$. Zvolme $\boldsymbol{\mu} \in \mathbb{R}^n$ a matici $\boldsymbol{A}_{n\times k}$. Definujme náhodný vektor $\boldsymbol{X} = \boldsymbol{A}\boldsymbol{\xi} + \boldsymbol{\mu}$. Potom dle vět 3.4 a 3.5 máme následující vztahy:

$$\begin{aligned} E[\boldsymbol{\xi}] &= \mathbf{0}, \\ \mathrm{Var}[\boldsymbol{\xi}] &= I_k, \\ E[\boldsymbol{X}] &= \boldsymbol{A}E[\boldsymbol{\xi}] + \boldsymbol{\mu} = \boldsymbol{\mu}, \\ \mathrm{Var}[\boldsymbol{X}] &= \boldsymbol{A}\,\mathrm{Var}[\boldsymbol{\xi}]\boldsymbol{A}^T = \boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{\Sigma}. \end{aligned}$$

Říkáme, že $\boldsymbol{X}$ má n-rozměrné normální rozdělení a značíme $\boldsymbol{X} \sim \mathrm{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Libovolná symetrická pozitivně semidefinitní matice $\boldsymbol{\Sigma}_{n\times n}$ taková, že $h(\boldsymbol{\Sigma}) = k \le n$, se dá rozložit na součin $\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{A}^T$, kde $\boldsymbol{A}$ je typu $n \times k$ a $h(\boldsymbol{A}) = k$. Speciálně, pokud $k = n$, pak $\boldsymbol{\Sigma}$ je regulární, a pokud $k < n$, pak $\boldsymbol{\Sigma}$ je singulární.

Věta 3.17 (Kritérium mnohorozměrné normality)**.** *Náhodný vektor* $\boldsymbol{X}$ *má rozdělení* $\mathrm{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ *právě tehdy, když* $\boldsymbol{c}^T\boldsymbol{X} \sim \mathrm{N}(\boldsymbol{c}^T\boldsymbol{\mu}, \boldsymbol{c}^T\boldsymbol{\Sigma}\boldsymbol{c})$ *pro libovolný vektor* $\boldsymbol{c} \in \mathbb{R}^n$.

Důkaz. Označme

$$X^* = \boldsymbol{c}^T\boldsymbol{X}, \quad \boldsymbol{A}^* = \boldsymbol{c}^T\boldsymbol{A}, \quad \mu^* = \boldsymbol{c}^T\boldsymbol{\mu}.$$

Potom $X^* = \boldsymbol{A}^*\boldsymbol{\xi} + \mu^*$ právě tehdy, když $\boldsymbol{X} = \boldsymbol{A}\boldsymbol{\xi} + \boldsymbol{\mu}$. Z toho plyne, dle definice mnohorozměrného normálního rozdělení, že $X^* \sim \mathrm{N}(\mu^*, \Sigma^*)$, kde $\Sigma^* = \boldsymbol{A}^*(\boldsymbol{A}^*)^T = \boldsymbol{c}^T\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{c} = \boldsymbol{c}^T\boldsymbol{\Sigma}\boldsymbol{c}$, a to je ekvivalentní s tím, že $\boldsymbol{X} \sim \mathrm{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. □

Důsledek 3.5. *Pro* $\boldsymbol{X} \sim \mathrm{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, *vektor* $\boldsymbol{a} \in \mathbb{R}^m$ *a matici* $\boldsymbol{B}_{m\times n}$ *platí*

$$\boldsymbol{a} + \boldsymbol{B}\boldsymbol{X} \sim \mathrm{N}_m(\boldsymbol{a} + \boldsymbol{B}\boldsymbol{\mu}, \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}^T),$$

tedy lineární transformace zachovává normalitu.

Důkaz. Dle věty 3.17 pro $\boldsymbol{d} \in \mathbb{R}^m$ položíme $\boldsymbol{c}^T = \boldsymbol{d}^T\boldsymbol{B}$ a dostaneme

$$\boldsymbol{c}^T\boldsymbol{X} = \boldsymbol{d}^T\boldsymbol{B}\boldsymbol{X} \sim \mathrm{N}(\boldsymbol{d}^T\boldsymbol{B}\boldsymbol{\mu}, \boldsymbol{d}^T\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}^T\boldsymbol{d}).$$

Posun o konstantu $\boldsymbol{d}^T\boldsymbol{a}$ nezmění typ rozdělení, tedy vzhledem k větám 3.4 a 3.5 máme, že pokud

$$\boldsymbol{d}^T(\boldsymbol{a} + \boldsymbol{B}\boldsymbol{X}) \sim \mathrm{N}(\boldsymbol{d}^T(\boldsymbol{a} + \boldsymbol{B}\boldsymbol{\mu}),\, \boldsymbol{d}^T\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}^T\boldsymbol{d}),$$

potom

$$\boldsymbol{a} + \boldsymbol{B}\boldsymbol{X} \sim \mathrm{N}_m(\boldsymbol{a} + \boldsymbol{B}\boldsymbol{\mu}, \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}^T).$$

□

Poznámka. Pokud $\boldsymbol{\Sigma}$ je singulární, pak existuje $\mathbf{0} \neq \boldsymbol{c} \in \mathbb{R}^n$ takový, že $\boldsymbol{c}^T\boldsymbol{X} = 0$.

Sdružená hustota vzhledem k Lebesgueově míře na $(\mathbb{R}^n, \mathcal{B}_0^n)$ existuje právě tehdy, když je $\boldsymbol{\Sigma}$ regulární, a platí

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \left(\frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\det(\boldsymbol{\Sigma})}}\right)\exp\left[-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right], \quad \boldsymbol{x} \in \mathbb{R}^n.$$

Nechť dále

$$\boldsymbol{X} = (\boldsymbol{Y}^T, \boldsymbol{Z}^T)^T, \quad \boldsymbol{Y} \in \mathbb{R}^r, \quad \boldsymbol{Z} \in \mathbb{R}^{n-r}, \quad \boldsymbol{X} \sim \mathrm{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}).$$

Označme

$$\boldsymbol{\mu} = (\boldsymbol{\mu}_{\boldsymbol{Y}}^T, \boldsymbol{\mu}_{\boldsymbol{Z}}^T)^T = (E[\boldsymbol{Y}^T], E[\boldsymbol{Z}^T])^T;$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{\boldsymbol{YY}} & \boldsymbol{\Sigma}_{\boldsymbol{YZ}} \\ \boldsymbol{\Sigma}_{\boldsymbol{ZY}} & \boldsymbol{\Sigma}_{\boldsymbol{ZZ}} \end{pmatrix} = \begin{pmatrix} \mathrm{Var}[\boldsymbol{Y}] & \mathrm{cov}(\boldsymbol{Y}, \boldsymbol{Z}) \\ \mathrm{cov}(\boldsymbol{Z}, \boldsymbol{Y}) & \mathrm{Var}[\boldsymbol{Z}] \end{pmatrix}.$$

Položme

$$\boldsymbol{a} = \boldsymbol{0} \in \mathbb{R}^{n-r},\ \boldsymbol{B} = (\boldsymbol{0}_{(n-r)\times r},\ I_{n-r}).$$

Potom je $\boldsymbol{a} + \boldsymbol{BX} = \boldsymbol{Z} = (X_{r+1}, \dots, X_n)^T$ a dle důsledku 3.5 máme, že $\boldsymbol{Z} \sim \mathrm{N}_{n-r}(\boldsymbol{\mu}_{\boldsymbol{Z}}, \boldsymbol{\Sigma}_{\boldsymbol{ZZ}})$, neboť zřejmě $\boldsymbol{a} + \boldsymbol{B\mu} = \boldsymbol{\mu}_{\boldsymbol{Z}}$, a nakonec

$$\boldsymbol{B\Sigma B}^T = \begin{pmatrix} \boldsymbol{0}_{(n-r)\times r} & I_{n-r} \end{pmatrix} \begin{pmatrix} \boldsymbol{\Sigma}_{\boldsymbol{YY}} & \boldsymbol{\Sigma}_{\boldsymbol{YZ}} \\ \boldsymbol{\Sigma}_{\boldsymbol{ZY}} & \boldsymbol{\Sigma}_{\boldsymbol{ZZ}} \end{pmatrix} \begin{pmatrix} \boldsymbol{0}_{r\times(n-r)} \\ I_{n-r} \end{pmatrix} = \boldsymbol{\Sigma}_{\boldsymbol{ZZ}}.$$

Marginální hustota je proto také normální.

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = \left(\frac{1}{(2\pi)^{\frac{n-r}{2}} \sqrt{\det(\boldsymbol{\Sigma}_{\boldsymbol{ZZ}})}} \right) \exp\left[-\frac{1}{2} (\boldsymbol{z} - \boldsymbol{\mu}_{\boldsymbol{Z}})^T \boldsymbol{\Sigma}_{\boldsymbol{ZZ}}^{-1} (\boldsymbol{z} - \boldsymbol{\mu}_{\boldsymbol{Z}}) \right], \quad \boldsymbol{z} \in \mathbb{R}^{n-r};$$

$$\boldsymbol{Z} \sim \mathrm{N}_{n-r}(\boldsymbol{\mu}_{\boldsymbol{Z}},\ \boldsymbol{\Sigma}_{\boldsymbol{ZZ}}).$$

Podmíněná hustota je rovněž normální. Podle věty 3.13 máme

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

Výpočet podmíněné hustoty tímto způsobem by byl obtížný. Odvodíme ji jiným postupem se zavedením pomocných náhodných vektorů a využitím jejich nezávislosti.

Momentová vytvořující funkce:

$$M_{\boldsymbol{X}}(\boldsymbol{t}) = E[e^{\boldsymbol{t}^T \boldsymbol{X}}] = \exp\left[\boldsymbol{t}^T \boldsymbol{\mu} + \frac{1}{2} \boldsymbol{t}^T \boldsymbol{\Sigma} \boldsymbol{t} \right].$$

Charakteristická funkce:

$$\Psi_{\boldsymbol{X}}(\boldsymbol{t}) = E[e^{i\boldsymbol{t}^T \boldsymbol{X}}] = \exp\left[i\boldsymbol{t}^T \boldsymbol{\mu} - \frac{1}{2} \boldsymbol{t}^T \boldsymbol{\Sigma} \boldsymbol{t} \right].$$

Věta 3.18 (Nezávislost a nekorelovanost v normálním rozdělení)**.** *Nechť* $\boldsymbol{X} = (\boldsymbol{Y}^T, \boldsymbol{Z}^T)^T \sim \mathrm{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. *Potom pokud* $\boldsymbol{\Sigma}_{\boldsymbol{YZ}} = \boldsymbol{0}$, *pak* $\boldsymbol{Y}$ *a* $\boldsymbol{Z}$ *jsou nezávislé.*

Důkaz. Nechť $\boldsymbol{\Sigma}_{\boldsymbol{YZ}} = \boldsymbol{0}$. Potom

$$\begin{aligned} M_{\boldsymbol{X}}(\boldsymbol{t}) &= \exp\left[\boldsymbol{t}_1^T \boldsymbol{\mu}_{\boldsymbol{Y}} + \boldsymbol{t}_2^T \boldsymbol{\mu}_{\boldsymbol{Z}} + \frac{1}{2} \boldsymbol{t}_1^T \boldsymbol{\Sigma}_{\boldsymbol{YY}} \boldsymbol{t}_1 + \frac{1}{2} \boldsymbol{t}_2^T \boldsymbol{\Sigma}_{\boldsymbol{ZZ}} \boldsymbol{t}_2 \right] \\ &= \exp\left[\boldsymbol{t}_1^T \boldsymbol{\mu}_{\boldsymbol{Y}} + \frac{1}{2} \boldsymbol{t}_1^T \boldsymbol{\Sigma}_{\boldsymbol{YY}} \boldsymbol{t}_1 \right] \exp\left[\boldsymbol{t}_2^T \boldsymbol{\mu}_{\boldsymbol{Z}} + \frac{1}{2} \boldsymbol{t}_2^T \boldsymbol{\Sigma}_{\boldsymbol{ZZ}} \boldsymbol{t}_2 \right] \\ &= M_{\boldsymbol{Y}}(\boldsymbol{t}_1) M_{\boldsymbol{Z}}(\boldsymbol{t}_2) \end{aligned}$$

pro $\boldsymbol{t} \in (\boldsymbol{t}_1^T, \boldsymbol{t}_2^T)^T \in \mathbb{R}^n$. Z toho dle bodu (iv) věty 3.9 plyne nezávislost vektorů $\boldsymbol{Y}$ a $\boldsymbol{Z}$. □

Důsledek 3.6. *Mějme* $\boldsymbol{X} \sim \mathrm{N}_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ *takový, že* $\boldsymbol{X} = (Y, Z)^T$. *Potom* Y, Z *jsou nekorelované právě tehdy, když jsou* Y, Z *nezávislé.*

Důkaz. Důsledek plyne jednoduše z vět 3.11 a 3.18. □

Vrátíme se k odvození podmíněné hustoty. Položme

$$\boldsymbol{U}_1 = \boldsymbol{Y} - \boldsymbol{\Sigma}_{\boldsymbol{YZ}}\boldsymbol{\Sigma}_{\boldsymbol{ZZ}}^{-1}\boldsymbol{Z}, \quad \boldsymbol{U}_2 = \boldsymbol{Z}, \quad \boldsymbol{U} = (\boldsymbol{U}_1^T, \boldsymbol{U}_2^T)^T, \quad \boldsymbol{B} = \begin{pmatrix} I_r & -\boldsymbol{\Sigma}_{\boldsymbol{YZ}}\boldsymbol{\Sigma}_{\boldsymbol{ZZ}}^{-1} \\ \mathbf{0} & I_{n-r} \end{pmatrix}.$$

Zřejmě $\boldsymbol{U} = \boldsymbol{B}\boldsymbol{X}$ a předpokládáme $\boldsymbol{X} \sim \mathrm{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Podle důsledku 3.5 máme

$$\boldsymbol{U} \sim \mathrm{N}_n(\boldsymbol{B}\boldsymbol{\mu}, \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}^T), \quad \boldsymbol{B}\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_{\boldsymbol{Y}} - \boldsymbol{\Sigma}_{\boldsymbol{YZ}}\boldsymbol{\Sigma}_{\boldsymbol{ZZ}}^{-1}\boldsymbol{\mu}_{\boldsymbol{Z}} \\ \boldsymbol{\mu}_{\boldsymbol{Z}} \end{pmatrix}.$$

Dále dostáváme

$$\boldsymbol{B}\boldsymbol{\Sigma} = \begin{pmatrix} I_r & -\boldsymbol{\Sigma}_{\boldsymbol{YZ}}\boldsymbol{\Sigma}_{\boldsymbol{ZZ}}^{-1} \\ \mathbf{0} & I_{n-r} \end{pmatrix} \begin{pmatrix} \boldsymbol{\Sigma}_{\boldsymbol{YY}} & \boldsymbol{\Sigma}_{\boldsymbol{YZ}} \\ \boldsymbol{\Sigma}_{\boldsymbol{ZY}} & \boldsymbol{\Sigma}_{\boldsymbol{ZZ}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma}_{\boldsymbol{YY}} - \boldsymbol{\Sigma}_{\boldsymbol{YZ}}\boldsymbol{\Sigma}_{\boldsymbol{ZZ}}^{-1}\boldsymbol{\Sigma}_{\boldsymbol{ZY}} & \mathbf{0} \\ \boldsymbol{\Sigma}_{\boldsymbol{ZY}} & \boldsymbol{\Sigma}_{\boldsymbol{ZZ}} \end{pmatrix},$$

$$\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}^T = \begin{pmatrix} \boldsymbol{\Sigma}_{\boldsymbol{YY}} - \boldsymbol{\Sigma}_{\boldsymbol{YZ}}\boldsymbol{\Sigma}_{\boldsymbol{ZZ}}^{-1}\boldsymbol{\Sigma}_{\boldsymbol{ZY}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{\boldsymbol{ZZ}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{\boldsymbol{ZZ}} \end{pmatrix}.$$

Podle věty 3.18 jsou $\boldsymbol{U}_1$ a $\boldsymbol{U}_2$ nezávislé. Z vět 3.9 a 3.13 plyne

$$f(\boldsymbol{u}_1 \mid \boldsymbol{u}_2) = \frac{f_{\boldsymbol{U}}(\boldsymbol{u}_1, \boldsymbol{u}_2)}{f_{\boldsymbol{U}_2}(\boldsymbol{u}_2)} = \frac{f_{\boldsymbol{U}_1}(\boldsymbol{u}_1) f_{\boldsymbol{U}_2}(\boldsymbol{u}_2)}{f_{\boldsymbol{U}_2}(\boldsymbol{u}_2)} = f_{\boldsymbol{U}_1}(\boldsymbol{u}_1).$$

Rozdělení náhodného vektoru $\boldsymbol{U}_1$ při pevném $\boldsymbol{U}_2 = \boldsymbol{u}_2$ je tedy $\mathrm{N}_r(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$, kde $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_{\boldsymbol{Y}} - \boldsymbol{\Sigma}_{\boldsymbol{YZ}}\boldsymbol{\Sigma}_{\boldsymbol{ZZ}}^{-1}\boldsymbol{\mu}_{\boldsymbol{Z}}$, čili rozdělení $\boldsymbol{Y} - \boldsymbol{\Sigma}_{\boldsymbol{YZ}}\boldsymbol{\Sigma}_{\boldsymbol{ZZ}}^{-1}\boldsymbol{Z}$ při pevném $\boldsymbol{Z} = \boldsymbol{z}$ je $\mathrm{N}_r(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$. Rozdělení $\boldsymbol{Y}$ při pevném $\boldsymbol{Z} = \boldsymbol{z}$ je pak

$$\mathrm{N}_r(\boldsymbol{\mu}_{\boldsymbol{Y}} + \boldsymbol{\Sigma}_{\boldsymbol{YZ}}\boldsymbol{\Sigma}_{\boldsymbol{ZZ}}^{-1}(\boldsymbol{z} - \boldsymbol{\mu}_{\boldsymbol{Z}}), \boldsymbol{\Sigma}_{11}).$$

Příklad 3.2 (Dvourozměrné normální rozdělení)**.** Položme $\boldsymbol{X} = (Y, Z)^T \sim \mathrm{N}_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ a předpokládejme, že $\boldsymbol{\mu} = \mathbf{0}$. Zřejmě

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_Y^2 & \rho\sigma_Y\sigma_Z \\ \rho\sigma_Y\sigma_Z & \sigma_Z^2 \end{pmatrix}, \text{ kde } \rho = \frac{\operatorname{cov}(Y, Z)}{\sqrt{\sigma_Y^2\sigma_Z^2}}, \quad \sigma_Y^2 > 0,\ \sigma_Z^2 > 0.$$

Pak pro $|\rho| < 1$ je $\det(\boldsymbol{\Sigma}) = \sigma_Y^2\sigma_Z^2(1 - \rho^2) \neq 0$, a tedy $\boldsymbol{\Sigma}$ je regulární. Můžeme tedy sestrojit inverzní matici ve tvaru

$$\boldsymbol{\Sigma}^{-1} = \frac{1}{\det(\boldsymbol{\Sigma})} \begin{pmatrix} \sigma_Z^2 & -\rho\sigma_Y\sigma_Z \\ -\rho\sigma_Y\sigma_Z & \sigma_Y^2 \end{pmatrix} = \frac{1}{1 - \rho^2} \begin{pmatrix} \frac{1}{\sigma_Y^2} & \frac{-\rho}{\sigma_Y\sigma_Z} \\ \frac{-\rho}{\sigma_Y\sigma_Z} & \frac{1}{\sigma_Z^2} \end{pmatrix}.$$

Označme

$$\begin{pmatrix} \frac{1}{\sigma_Y^2} & \frac{-\rho}{\sigma_Y\sigma_Z} \\ \frac{-\rho}{\sigma_Y\sigma_Z} & \frac{1}{\sigma_Z^2} \end{pmatrix} = C.$$

Pro hustotu potom platí

$$f_{\boldsymbol{X}}(y, z) = \left(\frac{1}{2\pi\sigma_Y\sigma_Z\sqrt{1 - \rho^2}}\right) \exp\left[\frac{-1}{2(1 - \rho^2)}(y, z)C(y, z)^T\right] =$$

$$= \left(\frac{1}{2\pi\sigma_Y\sigma_Z\sqrt{1 - \rho^2}}\right) \exp\left[\frac{-1}{2(1 - \rho^2)}\left(\frac{y^2}{\sigma_Y^2} - 2\rho\frac{yz}{\sigma_Y\sigma_Z} + \frac{z^2}{\sigma_Z^2}\right)\right].$$

Pokud bychom nepředpokládali, že $\boldsymbol{\mu} = \mathbf{0}$, ale místo toho $\boldsymbol{\mu} = (\mu_Y, \mu_Z)^T \neq \mathbf{0}$, dostaneme hustotu ve tvaru

$$f_{\boldsymbol{X}}(y, z) = \left(\frac{1}{2\pi\sigma_Y\sigma_Z\sqrt{1 - \rho^2}}\right) \exp\left[\frac{-1}{2(1 - \rho^2)}\left(\frac{(y - \mu_Y)^2}{\sigma_Y^2} - 2\rho\frac{(y - \mu_Y)(z - \mu_Z)}{\sigma_Y\sigma_Z} + \frac{(z - \mu_Z)^2}{\sigma_Z^2}\right)\right].$$

▲

4. Transformace náhodných veličin a vektorů

4.1. Transformace náhodných veličin

Nechť $X : (\Omega, \mathcal{A}) \longrightarrow (M, \mathcal{B})$ je obecná náhodná veličina (případně později náhodný vektor) a t je měřitelná funkce z M do $\mathbb{R}$ (případně do $\mathbb{R}^n$). Známe-li pravděpodobnostní rozdělení veličiny X, můžeme určit rozdělení veličiny $Y = t(X)$.

Definice 4.1 (Nosič rozdělení)**.** Nechť $X : (\Omega, \mathcal{A}) \longrightarrow (M, \mathcal{B})$ má absolutně spojité pravděpodobnostní rozdělení P_X vzhledem k σ-konečné míře μ. Množinu $S_X \subseteq M$ nazveme *nosičem rozdělení* náhodné veličiny X, pokud jsou splněny následující podmínky:

(i) $P[X \in S_X] = 1$;

(ii) pokud $\mu(S_X \setminus A) > 0$, pak $P[X \in A] < 1$ pro všechny $A \subset S_X$.

Úmluva. *Nadále budeme uvažovat výběrový prostor $M = \mathbb{R}$, případně $M = \mathbb{R}^n$.*

Příklad 4.1 (Nosiče rozdělení)**.**

Rozdělení	Nosič	Rozdělení	Nosič
$X \sim \mathrm{Alt}(p)$	$S_X = \{0, 1\}$	$X \sim \mathrm{Bi}(n, p)$	$S_X = \{0, 1, \dots, n\}$
$X \sim \mathrm{Po}(\lambda)$	$S_X = \{0, 1, \dots\}$	$X \sim \mathrm{Exp}(\lambda)$	$S_X = (0, \infty)$
$X \sim \mathrm{N}(\mu, \sigma^2)$	$S_X = \mathbb{R}$	$\boldsymbol{X} \sim \mathrm{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	$S_{\boldsymbol{X}} = \mathbb{R}^n$

▲

Věta 4.1 (O monotónní transformaci náhodné veličiny)**.** *Nechť náhodná veličina X má nosič S_X a distribuční funkci F_X. Nechť $t : S_X \longrightarrow S_0 \subseteq \mathbb{R}$ je ryze monotónní funkce. Označme $Y = t(X)$ a $\tau = t^{-1}$.*

(i) Je-li t rostoucí, má náhodná veličina Y distribuční funkci $F_Y(y) = F_X(\tau(y))$ pro $y \in S_0$.

(ii) Je-li t klesající, má náhodná veličina Y distribuční funkci $F_Y(y) = 1 - F_X(\tau(y)^-)$ pro $y \in S_0$, kde $F_X(\tau(y)^-) = \lim_{z \longrightarrow \tau(y)^-} F_X(z)$.

Důkaz.

(i) $F_Y(y) = P[Y \leq y] = P[t(X) \leq y] = P[X \leq \tau(y)] = F_X(\tau(y))$, přičemž třetí rovnost plyne z toho, že t je rostoucí.

(ii) $F_Y(y) = P[Y \leq y] = P[t(X) \leq y] = P[X \geq \tau(y)] = 1 - P[X < \tau(y)] = 1 - F_X(\tau(y)^-)$. Použili jsme klesající monotonii funkce t a větu 2.4 a její důkaz. □

Důsledek 4.1. *Nechť X je spojitá reálná náhodná veličina s hustotou f_X a t je ryze monotónní funkce mající derivaci skoro všude. Označme $Y = t(X)$ a $\tau = t^{-1}$. Potom náhodná veličina Y má hustotu*

$$f_Y(y) = \begin{cases} f_X(\tau(y))|\tau'(y)| & \text{pro } y \in S_0, \\ 0 & \text{jinak.} \end{cases}$$

Důkaz. Uvědomíme si, že pokud je X spojitá náhodná veličina, pak její distribuční funkce F_X je také spojitá. Pro t rostoucí máme

$$f_Y(y) = F_Y'(y) = F_X'(\tau(y))\tau'(y) = F_X'(\tau(y))|\tau'(y)|.$$

Vlastnost rostoucí monotonie jsme využili ve třetí rovnosti. Podobně pro t klesající

$$f_Y(y) = F_Y'(y) = -F_X'(\tau(y))\tau'(y) = F_X'(\tau(y))(-\tau'(y)) = F_X'(\tau(y))|\tau'(y)|.$$

Vlastnost klesající monotonie jsme využili v poslední rovnosti. □

Poznámka. Nechť X je diskrétní reálná náhodná veličina s rozdělením $P[X = x] = p_x$, $x \in S_X$. Mějme ryze monotónní funkci t, $Y = t(X)$ a $\tau = t^{-1}$. Náhodná veličina Y má potom rozdělení dané pravděpodobnostmi

$$P[Y = y] = P[t(X) = y] = P[X = \tau(y)], \quad y \in S_0.$$

Příklad 4.2 (Normalita lineární transformace náhodné veličiny)**.** Nechť

$$X \sim \mathrm{N}(\mu, \sigma^2), \quad a \in \mathbb{R}, \quad 0 \neq b \in \mathbb{R}, \quad Y = t(X) = a + bX.$$

Potom

$$X = \frac{1}{b}(Y - a) = \tau(Y), \quad \tau'(Y) = \frac{1}{b}, \quad S_X = S_0 = \mathbb{R}.$$

Zřejmě

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

pro $x \in \mathbb{R}$. Dále počítejme:

$$\begin{aligned} f_Y(y) &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2\sigma^2}\left(\frac{1}{b}(y-a) - \mu\right)^2\right] |b^{-1}| = \frac{1}{\sqrt{2\pi}\sigma|b|} \exp\left[-\frac{1}{2\sigma^2 b^2}\left(y - a - b\mu\right)^2\right] \\ &= \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2 b^2}} \exp\left[-\frac{1}{2\sigma^2 b^2}\left(y - (a + b\mu)\right)^2\right], \end{aligned}$$

z čehož plyne, že $Y \sim \mathrm{N}(a + b\mu, b^2\sigma^2)$. ▲

Příklad 4.3 (Transformace diskrétní náhodné veličiny)**.** Nechť

$$P[X = x] = p_x; \quad x = 0, 1, 2; \quad \sum_{x=0}^{2} p_x = 1; \quad S_X = \{0, 1, 2\}; \quad S_0 = \{0, 1, 4\}; \quad Y = t(X) = X^2.$$

Potom

(i) $X = \sqrt{Y} = \tau(Y)$,

(ii) $P[Y = 0] = P[X^2 = 0] = P[X = 0] = p_0$,

(iii) $P[Y = 1] = P[X = 1] = p_1$,

(iv) $P[Y = 4] = P[X = 2] = p_2$. ▲

Poznámka. Na následujících řádcích se budeme věnovat po částech monotónní transformaci. Uvažujme, že $S_X \subseteq \bigcup_{k=1}^{\infty} G_k$, kde $G_k \subseteq \mathbb{R}$ jsou intervaly splňující, že $G_i \cap G_j = \emptyset$ pro $i \neq j$. Nechť $t : S_X \longrightarrow S_0$ je ryze monotónní na každém G_k. Označme

$$K^+ = \{k : t \text{ roste na } G_k\}, \; K^- = \{k : t \text{ klesá na } G_k\};$$

$$t_k(x) = t(x) \cdot \mathbb{I}_{G_k}(x) = \begin{cases} t(x) & \text{pro } x \in G_k, \\ 0 & \text{jinak.} \end{cases}$$

$$\tau_k = t_k^{-1};$$

$$X_k = X \cdot \mathbb{I}_{G_k}(X);$$

$$Y_k = t_k(X_k).$$

Pokud $X \in G_k$, potom $X_k \neq 0$, $X_i = 0$ pro $i \neq k$, a tedy

$$X = \sum_{i=1}^{\infty} X_i, \quad Y = \sum_{i=1}^{\infty} Y_i.$$

Věta 4.2 (O nemonotónní transformaci náhodné veličiny). *Za předpokladů uvedených v předcházející poznámce platí pro distribuční funkci náhodné veličiny $Y = t(X)$*

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

pro všechna $y \in S_0$.

Důkaz. Protože $S_X \subseteq \bigcup_{k=1}^{\infty} G_k$ a G_k jsou po dvou disjunktní, náhodné jevy $X \in G_k$, $k = 1, 2, \ldots$ tvoří úplný systém jevů. Lze tedy psát

$$\begin{aligned}
F_Y(y) = P[Y \leq y] &= P\left[(Y \leq y) \cap \bigcup_{k=1}^{\infty}(X \in G_k)\right] = P\left[\bigcup_{k=1}^{\infty}((Y \leq y) \cap (X \in G_k))\right] \\
&= \sum_{k=1}^{\infty} P[(Y \leq y) \cap (X \in G_k)] = \sum_{k=1}^{\infty} P[Y_k \leq y,\, X \in G_k] = \sum_{k=1}^{\infty} P[t_k(X_k) \leq y,\, X \in G_k] \\
&= \sum_{k \in K^+} P[t_k(X_k) \leq y,\, X \in G_k] + \sum_{k \in K^-} P[t_k(X_k) \leq y,\, X \in G_k] \\
&= \sum_{k \in K^+} P[X_k \leq \tau_k(y),\, X \in G_k] + \sum_{k \in K^-} P[X_k \geq \tau_k(y),\, X \in G_k].
\end{aligned}$$

V poslední rovnosti využíváme toho, že v první sumě je t_k rostoucí a ve druhé klesající. □

Důsledek 4.2. *Nechť má navíc náhodná veličina X hustotu f_X vzhledem k Lebesgueově míře (je tedy spojitá) a nechť má každá funkce t_k derivaci skoro všude na G_k. Potom má $Y = t(X)$ hustotu*

$$f_Y(y) = \sum_{k=1}^{\infty} f_X(\tau_k(y))|\tau_k'(y)| \cdot \mathbb{I}_{t_k(G_k)}(y)$$

pro $y \in S_0$.

Důkaz. Podle důsledku 4.1 máme

$$f_Y(y) = f_X(\tau_k(y))|\tau_k'(y)|$$

pro $y \in t_k(G_k)$, $y = t_k(x)$, kde $x \in G_k$. Víme, že když $x \in G_k$ a $y \in t_k(G_k)$ pro právě jeden index k, pak $f_Y(y) \neq 0$ pro $y \in t_k(G_k)$ a $f_Y(y) = 0$ jinak, takže

$$f_Y(y) = \sum_{k=1}^{\infty} f_X(\tau_k(y))|\tau_k'(y)| \cdot \mathbb{I}_{t_k(G_k)}(y)$$

pro $y \in S_0$. □

Příklad 4.4 (Po částech monotónní transformace). Mějme $X \sim R(-1, 1)$ a $f_X(x) = \frac{1}{2}$ pro $x \in (-1, 1)$. Dále nechť

$$G_1 = (-1, 0), \quad G_2 = [0, 1), \quad S_0 = G_2, \quad S_X = G_1 \cup G_2.$$

Položme $Y = t(X) = X^2$. Zřejmě je $K^+ = \{2\}$, a $K^- = \{1\}$. Nechť

$$t_1(x) = x^2 \text{ pro } x \in G_1 = (-1, 0),$$

$$t_2(x) = x^2 \text{ pro } x \in G_2 = [0, 1),$$

takže $t_1(G_1) = (0, 1)$ a $t_2(G_2) = [0, 1)$.
Uvažujme $y = x^2$ a funkce $\tau_1(y) = -\sqrt{y}$ pro $y \in t_1(G_1)$ a $\tau_2(y) = \sqrt{y}$ pro $y \in t_2(G_2)$. Potom

$$\tau_1'(y) = -\frac{1}{2\sqrt{y}}, \quad \tau_2'(y) = \frac{1}{2\sqrt{y}}$$

pokud $y \neq 0$. Uvědomme si, že

$$X_1 = X \text{ na } G_1, \quad X_2 = X \text{ na } G_2, \quad X = X_1 + X_2,$$

$$Y_1 = X^2 \text{ na } G_1, \quad Y_2 = X^2 \text{ na } G_2, \quad Y = Y_1 + Y_2 = X^2 = t(X) \text{ na } S_X = G_1 \cup G_2.$$

Nyní počítejme. Dle věty 4.2

$$\begin{aligned}F_Y(y) &= P[X_2 \le \tau_2(y),\, X \in G_2] + P[X_1 \ge \tau(y),\, X \in G_1]\\ &= P[X_2 \le \sqrt{y},\, X \in [0,1)] + P[X_1 \ge -\sqrt{y},\, X \in (-1,0)]\\ &= \int_0^{\sqrt{y}} \frac{1}{2}\,\mathrm{d}y + \int_{-\sqrt{y}}^{0} \frac{1}{2}\,\mathrm{d}y = \frac{1}{2}(\sqrt{y}+\sqrt{y}) = \sqrt{y}\end{aligned}$$

pro $y \in S_0 = [0,1)$. Dle důsledku 4.2 potom

$$\begin{aligned}f_Y(y) &= f_X(\tau_1(y))|\tau_1'(y)| \cdot \mathbb{I}_{t_1(G_1)}(y) + f_X(\tau_2(y))|\tau_2'(y)| \cdot \mathbb{I}_{t_2(G_2)}(y)\\ &= \frac{1}{2}\left|-\frac{1}{2\sqrt{y}}\right| \cdot \mathbb{I}_{(0,1)}(y) + \frac{1}{2}\left|\frac{1}{2\sqrt{y}}\right| \cdot \mathbb{I}_{(0,1)}(y) = \frac{1}{4\sqrt{y}} + \frac{1}{4\sqrt{y}} = \frac{1}{2\sqrt{y}}\end{aligned}$$

pro $y \in (0,1)$. ▲

Poznámka (Výpočet střední hodnoty)**.** Nechť náhodná veličina X má hustotu $f_X(x)$ vzhledem k σ-konečné míře μ. Dále nechť $Y = t(X)$, kde t je (ryze) monotónní, nebo po částech monotónní na S_X. Potom

$$E[Y] = E[t(X)] = \int_{S_0} y f_Y(y)\,\mathrm{d}\mu(y).$$

Pokud nás nezajímá hustota, ale pouze střední hodnota náhodné veličiny Y, pak lze počítat

$$E[Y] = E[t(X)] = \int_{S_X} t(x) f_X(x)\,\mathrm{d}\mu(x).$$

Příklad 4.5 (Výpočet střední hodnoty)**.** Mějme

$$X \sim \mathrm{Exp}(\lambda), \quad f_X(x) = \frac{1}{\lambda}\exp\left[-\frac{x}{\lambda}\right].$$

Dále nechť $S_X = (0,\infty) = S_0$. Položme

$$Y = t(X) = X^2, \quad X = \tau(Y) = \sqrt{Y}, \quad \tau'(Y) = \frac{1}{2\sqrt{Y}}.$$

Potom pro hustotu platí

$$f_Y(y) = \frac{f_X(\sqrt{y})}{2\sqrt{y}} = \frac{1}{2\lambda\sqrt{y}}\exp\left[-\frac{\sqrt{y}}{\lambda}\right]$$

pro $y > 0$. Střední hodnotu pak můžeme spočítat jako

$$E[Y] = \int_0^\infty y f_Y(y)\,\mathrm{d}y = \frac{1}{2\lambda}\int_0^\infty \frac{y}{\sqrt{y}}\exp\left[-\frac{\sqrt{y}}{\lambda}\right]\,\mathrm{d}y.$$

Použijeme substituci $\sqrt{y} = x,\ y = x^2,\ \mathrm{d}y = 2x\,\mathrm{d}x$ a dostaneme

$$\begin{aligned}E[Y] &= \frac{1}{2\lambda}\int_0^\infty x\exp\left[-\frac{x}{\lambda}\right] 2x\,\mathrm{d}x = \int_0^\infty \frac{x^2}{\lambda}\exp\left[-\frac{x}{\lambda}\right]\,\mathrm{d}x = \int_0^\infty t(x) f_X(x)\,\mathrm{d}x\\ &= E[X]^2 = \mathrm{Var}[X] + (E[X])^2 = \lambda^2 + \lambda^2 = 2\lambda^2.\end{aligned}$$

▲

4.2. Transformace náhodných vektorů

Nechť $\boldsymbol{X} : (\Omega, \mathcal{A}) \longrightarrow (\mathbb{R}^n, \mathcal{B}_0^n)$ je náhodný vektor s nosičem $S_{\boldsymbol{X}} \subseteq \mathbb{R}^n$ a hustotou $f_{\boldsymbol{X}}$ vzhledem k součinové Lebesgueově míře na $(\mathbb{R}^n, \mathcal{B}_0^n)$. Uvažujme měřitelné zobrazení $t : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ takové, že $\boldsymbol{Y} = t(\boldsymbol{X}) =$

$(t_1(\boldsymbol{X}),\ldots,t_n(\boldsymbol{X}))^T$, kde $t_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ a $i = 1,\ldots,n$. Nechť skoro všude v $S_{\boldsymbol{X}}$ existuje matice parciálních derivací

$$\frac{\partial t(\boldsymbol{x})}{\partial \boldsymbol{x}} = \begin{pmatrix} \frac{\partial t_1(\boldsymbol{x})}{\partial x_1} & \cdots & \frac{\partial t_1(\boldsymbol{x})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial t_n(\boldsymbol{x})}{\partial x_1} & \cdots & \frac{\partial t_n(\boldsymbol{x})}{\partial x_n} \end{pmatrix}$$

a označme jakobián transformace t

$$D_t(\boldsymbol{x}) = \det\left(\frac{\partial t(\boldsymbol{x})}{\partial \boldsymbol{x}}\right).$$

Pro $\boldsymbol{X} = t^{-1}(\boldsymbol{Y}) = \tau(\boldsymbol{Y}),\ \tau : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, platí

$$\frac{\partial \tau(\boldsymbol{y})}{\partial \boldsymbol{y}} = \left(\frac{\partial t(\boldsymbol{x})}{\partial \boldsymbol{x}}\right)^{-1},$$

a tedy jakobián inverzní transformace τ je

$$D_\tau(\boldsymbol{y}) = (D_t(\boldsymbol{x}))^{-1}.$$

Definice 4.2 (Regulární zobrazení)**.** Nechť $q : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, $q(\boldsymbol{y}) = (q_1(\boldsymbol{y}),\ldots,q_n(\boldsymbol{y}))^T$, $\boldsymbol{y} \in \mathbb{R}^n$ je zobrazení s jakobiánem D_q. Zobrazení q je *regulární* v množině $K \subseteq \mathbb{R}^n$, když platí

(i) K je otevřená množina,

(ii) funkce $q_1,\ldots,q_n$ mají spojité parciální derivace 1. řádu v K,

(iii) $D_q(\boldsymbol{y}) \neq 0$ pro všechny $\boldsymbol{y} \in K$.

Věta 4.3 (O substituci)**.** *Nechť q je regulární a prosté zobrazení otevřené množiny $K \subseteq \mathbb{R}^n$ na $L \subseteq \mathbb{R}^n$ s jakobiánem D_q. Nechť $M \subseteq L$ je borelovská množina a Q je měřitelná reálná funkce. Potom platí*

$$\int_M Q(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{q^{-1}(M)} Q(q(\boldsymbol{y}))|D_q(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y},$$

pokud jeden z integrálů existuje.

Důkaz. Vynechán. □

Věta 4.4 (O transformaci náhodného vektoru)**.** *Nechť náhodný vektor $\boldsymbol{X}$ má hustotu $f_{\boldsymbol{X}}$ vzhledem k Lebesgueově míře na $(\mathbb{R}^n,\mathcal{B}_0^n)$. Mějme zobrazení $t : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ regulární a prosté na otevřené množině G, pro kterou platí $\int_G f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 1$. Označme τ inverzní zobrazení k $t : G \longrightarrow t(G)$. Potom má náhodný vektor $\boldsymbol{Y} = t(\boldsymbol{X})$ hustotu vzhledem k Lebesgueově míře, pro niž platí*

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \begin{cases} f_{\boldsymbol{X}}(\tau(\boldsymbol{y}))|D_\tau(\boldsymbol{y})| & \text{pro } \boldsymbol{y} \in t(G), \\ 0 & \text{jinak.} \end{cases}$$

Důkaz. Pro libovolnou množinu $B \in \mathcal{B}_0^n$ lze psát

$$P[\boldsymbol{Y} \in B] = P[\boldsymbol{Y} \in B,\, \boldsymbol{Y} \in t(G)] + P[\boldsymbol{Y} \in B,\, \boldsymbol{Y} \notin t(G)].$$

Zřejmě platí

$$P[\boldsymbol{Y} \notin t(G)] = P[\boldsymbol{Y} \in (\mathbb{R}^n \setminus t(G))] = P[t(\boldsymbol{X}) \notin t(G)] = P[\boldsymbol{X} \notin G] = 0,$$

neboť

$$P[\boldsymbol{X} \in G] = \int_G f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 1.$$

Tedy

$$P[\boldsymbol{Y} \in B] = P[\boldsymbol{Y} \in B, \boldsymbol{Y} \in t(G)] = P[\boldsymbol{Y} \in (B \cap t(G))] = P[\boldsymbol{X} \in \tau(B \cap t(G))] = \int_{\tau(B\cap t(G))} f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

Zvolme nyní ve větě 4.3 (o substituci)

$$Q = f_{\boldsymbol{X}}, \quad q = \tau, \quad q^{-1} = t, \quad M = \tau(B \cap t(G)).$$

Vzhledem k tomu, že $P[\boldsymbol{Y} \in (\mathbb{R}^n \setminus t(G))] = 0$, tak také $P[\boldsymbol{Y} \in (B \cap (\mathbb{R}^n \setminus t(G)))] = 0$. Lze psát

$$P[\boldsymbol{Y} \in B] = \int_{B \cap t(G)} f_{\boldsymbol{X}}(\tau(\boldsymbol{y}))|D_\tau(\boldsymbol{y})| \, \mathrm{d}\boldsymbol{y} + \int_{B \cap (\mathbb{R}^n \setminus t(G))} 0 \, \mathrm{d}\boldsymbol{y} = \int_B f_{\boldsymbol{Y}}(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y},$$

protože lze integrovanou funkci $f_{\boldsymbol{Y}}(\boldsymbol{y})$ na $\mathbb{R}^n \setminus t(G)$ dodefinovat nulou. Dostali jsme

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \begin{cases} f_{\boldsymbol{X}}(\tau(\boldsymbol{y}))|D_\tau(\boldsymbol{y})| & \text{pro } \boldsymbol{y} \in t(G), \\ 0 & \text{pro } \boldsymbol{y} \notin t(G) \end{cases}$$

jako hustotu náhodného vektoru $\boldsymbol{Y}$. □

Poznámka. Obecně $S_{\boldsymbol{X}} \subseteq G$, ale pokud je nosič rozdělení náhodného vektoru $\boldsymbol{X}$ otevřená množina, pak $S_{\boldsymbol{X}} = G$.

Příklad 4.6 (Odvození hustoty mnohorozměrného normálního rozdělení)**.** Uvažujme $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)^T$, kde $\xi_i \sim \mathrm{N}(0,1)$ jsou nezávislé a $i = 1, \ldots, n$. Dále

$$f_i(\xi_i) = \phi(\xi_i) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{\xi_i^2}{2}\right]$$

pro $\xi_i \in \mathbb{R}$, $i = 1, \ldots, n$. Dle věty 2.18 dostaneme

$$f_{\boldsymbol{\xi}}(\boldsymbol{\xi}) = \prod_{i=1}^n f_i(\xi_i) = \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left[-\frac{1}{2}\sum_{i=1}^n \xi_i^2\right] = \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left[-\frac{1}{2}\boldsymbol{\xi}^T\boldsymbol{\xi}\right],$$

kde $\boldsymbol{\xi} \in \mathbb{R}^n$. Dále mějme $\boldsymbol{X} = \boldsymbol{A}\boldsymbol{\xi} + \boldsymbol{\mu} = t(\boldsymbol{\xi})$, kde $\boldsymbol{\mu} \in \mathbb{R}^n$ a $\boldsymbol{A} = (a_{ij})_{i,j=1}^n$ je regulární. Zobrazení t je prosté a regulární na $G = \mathbb{R}^n$, $t(G) = \mathbb{R}^n$. Pak

$$\int_{\mathbb{R}^n} f(\boldsymbol{\xi}) \, \mathrm{d}\boldsymbol{\xi} = 1,$$

neboť $S_{\boldsymbol{\xi}} = \mathbb{R}^n$. Zřejmě také

$$\boldsymbol{\xi} = \boldsymbol{A}^{-1}(\boldsymbol{X} - \boldsymbol{\mu}) = \tau(\boldsymbol{X}), \quad D_\tau(\boldsymbol{x}) = \frac{1}{D_t(\boldsymbol{\xi})}.$$

Pokud si rozepíšeme $\boldsymbol{X}$ po složkách, tak

$$X_i = \sum_{j=1}^n a_{ij}\xi_j + \mu_i, \quad i = 1, \ldots, n.$$

Zřejmě potom

$$t(\boldsymbol{\xi}) = (t_1(\boldsymbol{\xi}), \ldots, t_n(\boldsymbol{\xi}))^T = (X_1, \ldots, X_n)^T,$$

$$D_t(\boldsymbol{\xi}) = \det \begin{pmatrix} \frac{\partial x_1}{\partial \xi_1} & \cdots & \frac{\partial x_1}{\partial \xi_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial \xi_1} & \cdots & \frac{\partial x_n}{\partial \xi_n} \end{pmatrix} = \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \det(\boldsymbol{A}),$$

$$D_\tau(\boldsymbol{x}) = \frac{1}{\det(\boldsymbol{A})}.$$

Dále máme $\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{A}^T$. Pak $\det(\boldsymbol{\Sigma}) = \det(\boldsymbol{A}\boldsymbol{A}^T) = (\det(\boldsymbol{A}))^2$, takže $\sqrt{\det(\Sigma)} = |\det(\boldsymbol{A})|$. Upravme ještě skalární součin

$$\begin{aligned} \boldsymbol{\xi}^T\boldsymbol{\xi} &= (\tau(\boldsymbol{x}))^T\tau(\boldsymbol{x}) = (\boldsymbol{A}^{-1}(\boldsymbol{X} - \boldsymbol{\mu}))^T\boldsymbol{A}^{-1}(\boldsymbol{X} - \boldsymbol{\mu}) \\ &= (\boldsymbol{X} - \boldsymbol{\mu})^T(\boldsymbol{A}^{-1})^T\boldsymbol{A}^{-1}(\boldsymbol{X} - \boldsymbol{\mu}) \\ &= (\boldsymbol{X} - \boldsymbol{\mu})^T(\boldsymbol{A}\boldsymbol{A}^T)^{-1}(\boldsymbol{X} - \boldsymbol{\mu}) \\ &= (\boldsymbol{X} - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{X} - \boldsymbol{\mu}). \end{aligned}$$

Použili jsme pravidla maticové algebry:

(i) $(\boldsymbol{AB})^T = \boldsymbol{B}^T\boldsymbol{A}^T$,

(ii) $(\boldsymbol{A}^{-1})^T = (\boldsymbol{A}^T)^{-1}$,

(iii) $(\boldsymbol{AC})^{-1} = \boldsymbol{C}^{-1}\boldsymbol{A}^{-1}$

pro matice $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ vhodných rozměrů, $\boldsymbol{A}$, $\boldsymbol{C}$ regulární. Nakonec dle věty 4.4 dostáváme

$$\begin{aligned} f_{\boldsymbol{X}}(\boldsymbol{x}) = f_{\boldsymbol{\xi}}(\tau(\boldsymbol{x}))|D_\tau(\boldsymbol{x})| &= \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left[-\frac{1}{2}(\boldsymbol{X}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{X}-\boldsymbol{\mu})\right]\left|\frac{1}{\det(\boldsymbol{A})}\right| \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\det(\boldsymbol{\Sigma})}} \exp\left[-\frac{1}{2}(\boldsymbol{X}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{X}-\boldsymbol{\mu})\right] \end{aligned}$$

pro $\boldsymbol{x} \in \mathbb{R}^n$. ▲

Příklad 4.7 (Normalita lineární transformace náhodného vektoru)**.** Nechť

$$\boldsymbol{X} \sim \mathrm{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \quad \boldsymbol{\Sigma} \text{ je regulární}, \quad \boldsymbol{Y} = t(\boldsymbol{X}) = \boldsymbol{a} + \boldsymbol{BX}, \quad \boldsymbol{a} \in \mathbb{R}^n, \quad \boldsymbol{B}_{n\times n} \text{ je regulární}.$$

Zřejmě platí, že

$$\begin{aligned} &\boldsymbol{X} = \tau(\boldsymbol{Y}) = \boldsymbol{B}^{-1}(\boldsymbol{Y}-\boldsymbol{a}), \\ &D_\tau(\boldsymbol{y}) = \frac{1}{D_t(\boldsymbol{x})} = \frac{1}{\det(\boldsymbol{B})}, \\ &G = \mathbb{R}^n. \end{aligned}$$

Zobrazení t je prosté a regulární v G, $t(G) = \mathbb{R}^n$ a platí

$$\begin{aligned} &\int_{\mathbb{R}^n} f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 1 \\ &f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\det(\boldsymbol{\Sigma})}} \exp\left[-\frac{1}{2}(\boldsymbol{X}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{X}-\boldsymbol{\mu})\right] \end{aligned}$$

pro $\boldsymbol{x} \in \mathbb{R}^n$. Dle věty 4.4 pak dostaneme

$$\begin{aligned} f_{\boldsymbol{Y}}(\boldsymbol{y}) &= f_{\boldsymbol{X}}(\tau(\boldsymbol{y}))|D_\tau(\boldsymbol{y})| \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\det(\boldsymbol{\Sigma})}\sqrt{(\det(\boldsymbol{B}))^2}} \exp\left[-\frac{1}{2}(\boldsymbol{B}^{-1}(\boldsymbol{y}-\boldsymbol{a}-\boldsymbol{B\mu}))^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{B}^{-1}(\boldsymbol{y}-\boldsymbol{a}-\boldsymbol{B\mu}))\right] \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\det(\boldsymbol{\Sigma})\det(\boldsymbol{B})\det(\boldsymbol{B}^T)}} \exp\left[-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{a}-\boldsymbol{B\mu})^T(\boldsymbol{B}^{-1})^T\boldsymbol{\Sigma}^{-1}\boldsymbol{B}^{-1}(\boldsymbol{y}-\boldsymbol{a}-\boldsymbol{B\mu})\right] \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\det(\boldsymbol{B\Sigma B}^T)}} \exp\left[-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{a}-\boldsymbol{B\mu})^T(\boldsymbol{B\Sigma B}^T)^{-1}(\boldsymbol{y}-\boldsymbol{a}-\boldsymbol{B\mu})\right] \end{aligned}$$

pro $\boldsymbol{y} \in \mathbb{R}^n$. Takže $\boldsymbol{Y} \sim \mathrm{N}_n(\boldsymbol{a}+\boldsymbol{B\mu},\, \boldsymbol{B\Sigma B}^T)$. ▲

Nyní se podíváme na obecnou situaci. Nechť $S_{\boldsymbol{X}} \subseteq \bigcup_{k=1}^\infty G_k$, kde $G_k \subseteq \mathbb{R}^n$ jsou otevřené množiny a $G_i \cap G_j = \emptyset$ pro $i \neq j$. Dále mějme zobrazení $t : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ regulární a prosté na každém G_k a označme

$$\begin{aligned} &t_k(\boldsymbol{x}) = t(\boldsymbol{x})\mathbb{I}_{G_k}(\boldsymbol{x}) : G_k \longrightarrow t_k(G_k) \\ &\tau_k(\boldsymbol{y}) = t_k^{-1}(\boldsymbol{y}) : t_k(G_k) \longrightarrow G_k. \end{aligned}$$

Věta 4.5 (Zobecněná věta o transformaci náhodného vektoru)**.** *Nechť náhodný vektor $\boldsymbol{X}$ má hustotu $f_{\boldsymbol{X}}$ vzhledem k Lebesgueově míře. Potom má za výše uvedených předpokladů náhodný vektor $\boldsymbol{Y} = t(\boldsymbol{X})$ hustotu vzhledem k Lebesgueově míře*

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \sum_{k=1}^{\infty} f_{\boldsymbol{X}}(\tau_k(\boldsymbol{y}))|D_{\tau_k}(\boldsymbol{y})|\mathbb{I}_{t_k(G_k)}(\boldsymbol{y}), \quad \boldsymbol{y} \in \mathbb{R}^n.$$

Důkaz. Analogicky jako při odvození důsledku 4.2. □

Zaměřme se nyní na transformaci náhodného vektoru na náhodnou veličinu. Mějme náhodný vektor $\boldsymbol{X} = (X_1, \dots, X_n)^T$ s hustotou $f_{\boldsymbol{X}}$ vzhledem k Lebesgueově míře, měřitelnou funkci $t_1 : \mathbb{R}^n \longrightarrow \mathbb{R}$, která má derivace skoro všude na $S_{\boldsymbol{X}}$, a nechť $T = t_1(\boldsymbol{X})$, například

$$T = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Chceme zjistit, jaké rozdělení má náhodná veličina T. Zvolíme transformaci $\boldsymbol{Y} = t(\boldsymbol{X}) = (t_1(\boldsymbol{X}), \dots, t_n(\boldsymbol{X}))^T$. Jsou-li splněny předpoklady věty 4.4 nebo 4.5, určíme podle nich hustotu $f_{\boldsymbol{Y}}$. Podle věty 3.3 pak určíme marginální hustotu složky $Y_1 = t_1(\boldsymbol{X}) = T$, jež je rovna

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\boldsymbol{Y}}(y_1, \dots, y_n)\,\mathrm{d}y_2 \cdots \mathrm{d}y_n.$$

V dalším textu se zaměříme na rozdělení součtu nezávislých náhodných veličin.

Definice 4.3 (Konvoluce)**.** Nechť X, Y jsou nezávislé náhodné veličiny s distribučními funkcemi F_X, F_Y. Distribuční funkci náhodné veličiny $Z = X + Y$ nazýváme *konvolucí* distribučních funkcí F_X a F_Y. Značíme $F_Z = F_X * F_Y$.

Věta 4.6 (O konvoluci)**.** *Nechť X a Y jsou nezávislé náhodné veličiny, kde X má hustotu f_X vzhledem k σ-konečné míře μ_1 a distribuční funkci F_X a Y má hustotu f_Y vzhledem k σ-konečné míře μ_2 a distribuční funkci F_Y. Pak má náhodná veličina $Z = X + Y$ distribuční funkci*

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

Důkaz. Protože jsou X, Y nezávislé, platí, že sdružená hustota vzhledem k $\mu = \mu_1 \times \mu_2$ je $f_{X,Y}(x,y) = f_X(x) f_Y(y)$. Označme $B_z = \{(x,y) : x + y \leq z\} \subset \mathcal{B}_0^2$. Dále dle poznámek za definicí 3.2 máme

$$F_Z(z) = P[Z \leq z] = P[X + Y \leq z] = \int_{\mathbb{R}^2} \mathbb{I}_{B_z}(x,y) f_{X,Y}(x,y)\,\mathrm{d}\mu(x,y) = \iint\limits_{x+y\leq z} f_X(x) f_Y(y)\,\mathrm{d}\mu_1(x)\,\mathrm{d}\mu_2(y)$$

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

dle poznámek za definicí 2.10. □

Poznámka. Analogickým postupem lze dokázat, že platí

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

To znamená, že operace konvoluce je komutativní.

Věta 4.7 (Rozdělení součtu dvou diskrétních náhodných veličin)**.** *Nechť X, Y jsou nezávislé náhodné veličiny, $P[X = x] = p_x$ pro $x \in S_X$, $P[Y = y] = q_y$ pro $y \in S_Y$ a $S_X, S_Y \subseteq \{0, 1, 2, \dots\}$. Potom má náhodná veličina $Z = X + Y$ rozdělení dané pravděpodobnostmi*

$$P[Z = z] = \sum_{x \in S_X} p_x q_{z-x} = \sum_{y \in S_Y} p_{z-y} q_y.$$

Důkaz. S využitím věty 2.17 získáme

$$P[Z = z] = P[X + Y = z] = P\left[\bigcup_{x\in S_X}(X = x,\, Y = z - x)\right] = \sum_{x\in S_X} P[X = x,\, Y = z - x] = \sum_{x\in S_X} p_x q_{z-x}.$$

Druhé vyjádření má důkaz analogický. □

Poznámka. Větu lze formulovat i bez předpokladu nezávislosti; dostali bychom pak pouze

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

V sumách sčítáme přes všechna $x \in S_X$ taková, že $z - x \in S_Y$, resp. přes všechna $y \in S_Y$ taková, že $z - y \in S_X$.

Příklad 4.8 (Rozdělení součtu dvou binomických a poissonovských veličin)**.** Pomocí věty 4.7 lze ukázat

(i) pokud jsou $X \sim \mathrm{Bi}(n_1, p)$, $Y \sim \mathrm{Bi}(n_2, p)$ nezávislé, potom $Z = X + Y \sim \mathrm{Bi}(n_1 + n_2, p)$;

(ii) pokud jsou $X \sim \mathrm{Po}(\lambda_1)$, $Y \sim \mathrm{Po}(\lambda_2)$ nezávislé, potom $Z = X + Y \sim \mathrm{Po}(\lambda_1 + \lambda_2)$. ▲

Věta 4.8 (Rozdělení součtu dvou spojitých náhodných veličin)**.** *Nechť X, Y jsou nezávislé náhodné veličiny s hustotami f_X, f_Y vzhledem k Lebesgueově míře. Potom má náhodná veličina $Z = X + Y$ hustotu*

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x)\,\mathrm{d}x = \int_{-\infty}^{\infty} f_Y(y) f_X(z - y)\,\mathrm{d}y.$$

Důkaz. Použijeme distribuční funkci. Dle odvození v důkazu věty 4.6 máme

$$\begin{aligned} F_Z(z) &= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{z-x} f_Y(y)\,\mathrm{d}y\right) f_X(x)\,\mathrm{d}x = \int_{-\infty}^{\infty}\left(\int_{-\infty}^{z} f_Y(u - x)\,\mathrm{d}u\right) f_X(x)\,\mathrm{d}x \\ &= \int_{-\infty}^{z}\left(\int_{-\infty}^{\infty} f_Y(u - x) f_X(x)\,\mathrm{d}x\right)\mathrm{d}u = \int_{-\infty}^{z} f_Z(u)\,\mathrm{d}u, \end{aligned}$$

takže

$$f_Z(u) = \int_{-\infty}^{\infty} f_Y(u - x) f_X(x)\,\mathrm{d}x.$$

V úpravě jsou použili jsme substituci $y = u - x$. Druhé vyjádření se dokáže analogicky. □

Poznámky.

(i) Větu 4.8 lze opět formulovat i bez předpokladu nezávislosti; dostali bychom pak

$$f_Z(z) = \int_{-\infty}^{\infty} f_{X,Y}(x, z - x)\,\mathrm{d}x = \int_{-\infty}^{\infty} f_{X,Y}(z - y, y)\,\mathrm{d}y,$$

pokud existuje sdružená hustota. Integrujeme přes všechna $x \in S_X \subseteq \mathbb{R}$ taková, že $z - x \in S_Y$, resp. přes všechna $y \in S_Y \subset \mathbb{R}$ taková, že $z - y \in S_X$. Pokud $S_X = S_Y = \mathbb{R}$, integrujeme přes všechna $x \in \mathbb{R}$, resp. přes všechna $y \in \mathbb{R}$.

(ii) Větu 4.8 lze dokázat i pomocí věty 4.4 (o transformaci náhodného vektoru) a věty 3.3 (o marginální hustotě).

Mějme X, Y nezávislé náhodné veličiny s hustotami f_X, f_Y vzhledem k Lebesgueově míře. Nechť $Z = X+Y$ a $U = X$. Položme $t((X,Y)^T) = (Z,U)^T = (X+Y, X)^T$ prosté a regulární v $\mathbb{R}^2$ a $\tau((Z,U)^T) = (X,Y)^T = (U, Z - U)^T$. Potom

$$D_\tau = \det\begin{pmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial u} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial u} \end{pmatrix} = \det\begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} = -1.$$

Dle věty 4.4 máme

$$\begin{aligned} f_{Z,U}(z,u) &= f_{X,Y}(\tau(z,u))|D_\tau| \quad \text{pro } (z,u)^T \in \mathbb{R}^2; \\ f_{Z,U}(z,u) &= f_X(u) f_Y(z - u). \end{aligned}$$

Nakonec dle věty 3.3 dostaneme

$$f_Z(z) = \int_{-\infty}^{\infty} f_{ZU}(z,u)\,\mathrm{d}u = \int_{-\infty}^{\infty} f_X(u) f_Y(z - u)\,\mathrm{d}u.$$

Příklad 4.9 (Rozdělení součtu dvou normálních veličin)**.** Pomocí věty 4.8 lze dokázat, že pokud jsou náhodné veličiny $X \sim \mathrm{N}(\mu_1, \sigma_1^2)$, $Y \sim \mathrm{N}(\mu_2, \sigma_2^2)$ nezávislé, pak má náhodná veličina $Z = X + Y$ rozdělení

$$Z \sim \mathrm{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

▲

Příklad 4.10 (Součet rovnoměrných rozdělení)**.** Mějme $X \sim R(0,1)$, $Y \sim R(0,1)$ nezávislé, dále $f_X(x) = 1$, $x \in (0,1)$ a $f_Y(y) = 1$, $y \in (0,1)$. Pak $Z = X + Y$ má podle věty 4.8 hustotu

$$f_Z(z) = \int_{B_z} f_X(x) f_Y(z-x)\,\mathrm{d}x = \int_{B_z} 1\,\mathrm{d}x,$$

kde

$$\begin{aligned} B_z &= \{x : 0 < x < 1,\, 0 < z - x < 1\} = \{x : 0 < x < 1,\, z - 1 < x < z\} \\ &= \{x : \max(0, z-1) < x < \min(1, z)\}. \end{aligned}$$

Pro $0 < z \le 1$ máme $0 < x < z$ a

$$f_Z(z) = \int_0^z 1\,\mathrm{d}x = z;$$

pro $1 < z < 2$ máme $z - 1 < x < 1$ a

$$f_Z(z) = \int_{z-1}^1 1\,\mathrm{d}x = 2 - z.$$

Zřejmě

$$\int_0^2 f_Z(z)\,\mathrm{d}z = \int_0^1 z\,\mathrm{d}z + \int_1^2 2 - z\,\mathrm{d}z = \left[\frac{z^2}{2}\right]_0^1 + \left[2z - \frac{z^2}{2}\right]_1^2 = 1.$$

▲

Příklad 4.11 (Součet exponenciálních rozdělení, Gamma rozdělení)**.**

(i) Mějme $X \sim \mathrm{Exp}(\lambda)$, $Y \sim \mathrm{Exp}(\lambda)$ nezávislé a

$$f_X(x) = \frac{1}{\lambda} \exp\left[-\frac{x}{\lambda}\right]$$

pro $x > 0$. Potom $Z = X + Y$ má dle věty 4.8 hustotu

$$f_Z(z) = \int_0^z \frac{1}{\lambda} \exp\left[-\frac{x}{\lambda}\right] \frac{1}{\lambda} \exp\left[-\frac{z-x}{\lambda}\right] \mathrm{d}x = \frac{1}{\lambda^2} \exp\left[-\frac{z}{\lambda}\right] \int_0^z 1\,\mathrm{d}x = \frac{z}{\lambda^2} \exp\left[-\frac{z}{\lambda}\right],$$

kde $z > 0$, $x > 0$, $z - x > 0$.

(ii) Nechť X a Y jsou nezávislé, X má hustotu

$$f_X(x) = \frac{x}{\lambda^2} \exp\left[-\frac{x}{\lambda}\right]$$

pro $x > 0$ a $Y \sim \mathrm{Exp}(\lambda)$. Potom $Z = X + Y$ má podle věty 4.8 hustotu

$$f_Z(z) = \frac{1}{\lambda^3} \int_0^z x \exp\left[-\frac{x}{\lambda}\right] \exp\left[-\frac{z-x}{\lambda}\right] \mathrm{d}x = \frac{1}{\lambda^3} \exp\left[-\frac{z}{\lambda}\right] \frac{z^2}{2}$$

pro $z > 0$.

(iii) Indukcí lze s využitím věty 4.8 dokázat, že pokud jsou $X_1, \dots, X_n \sim \mathrm{Exp}(\lambda)$ nezávislé, potom náhodná veličina $Z_n = X_1 + \cdots + X_n$ má rozdělení s hustotou

$$f_{Z_n}(z) = \left(\frac{1}{\lambda}\right)^n \frac{z^{n-1}}{\Gamma(n)} \exp\left[-\frac{z}{\lambda}\right]$$

pro $z > 0$, kde $\Gamma(n)$ je *Gamma funkce* definovaná předpisem

$$\Gamma(n) = \int_0^\infty z^{n-1} e^{-z}\,\mathrm{d}z$$

pro $n > 0$. Shrneme si základní vlastnosti Gamma funkce:

$$\begin{aligned}
\Gamma\left(\frac{1}{2}\right) &= \sqrt{\pi}; \\
\Gamma(1) &= 1; \\
\Gamma(n+1) &= n! && \text{pro } n \in \mathbb{Z}; \\
\Gamma(n+1) &= n\Gamma(n) && \text{pro } n \in \mathbb{Z},\ n > 1.
\end{aligned}$$

Jedná se tedy o Gamma rozdělení s parametry $\frac{1}{\lambda}$, n; píšeme $Z_n \sim \Gamma(\frac{1}{\lambda}, n)$. ▲

Věta 4.9 (O rozdělení podílu náhodných veličin)**.** *Nechť X, Y jsou nezávislé náhodné veličiny, X má distribuční funkci F_X a hustotu f_X vzhledem k σ-konečné míře μ_1, Y má distribuční funkci F_Y a hustotu f_Y vzhledem k σ-konečné míře μ_2. Potom má náhodná veličina $Z = \frac{X}{Y}$ distribuční funkci*

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

Důkaz. Počítejme:

$$\begin{aligned}
F_Z(z) = P\left[\frac{X}{Y} \le z\right] &= \iint\limits_{\frac{x}{y} \le z} f_X(x) f_Y(y)\,\mathrm{d}\mu_1(x)\,\mathrm{d}\mu_2(y) \\
&= \iint\limits_{\substack{y>0 \\ x \le zy}} f_X(x) f_Y(y)\,\mathrm{d}\mu_1(x)\,\mathrm{d}\mu_2(y) + \iint\limits_{\substack{y<0 \\ x \ge zy}} f_X(x) f_Y(y)\,\mathrm{d}\mu_1(x)\,\mathrm{d}\mu_2(y) \\
&= \int_0^\infty \left(\int_{-\infty}^{zy} f_X(x)\,\mathrm{d}\mu_1(x)\right) f_Y(y)\,\mathrm{d}\mu_2(y) + \int_{-\infty}^0 \left(\int_{zy}^\infty f_X(x)\,\mathrm{d}\mu_1(x)\right) f_Y(y)\,\mathrm{d}\mu_2(y) \\
&= \int_0^\infty F_X(zy) f_Y(y)\,\mathrm{d}\mu_2(y) + \int_{-\infty}^0 [1 - F_X(zy)] f_Y(y)\,\mathrm{d}\mu_2(y).
\end{aligned}$$

□

Důsledek 4.3. *Jsou-li X, Y spojité, má $Z = \frac{X}{Y}$ hustotu vzhledem k Lebesgueově míře rovnou*

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

Důkaz. Opět lze jednoduše spočítat, že

$$f_Z(z) = F_Z'(z) = \int_0^\infty f_Y(y) F_X'(zy) y\,\mathrm{d}y + \int_{-\infty}^0 f_Y(y) F_X'(zy)(-y)\,\mathrm{d}y,$$

což dokazuje naše tvrzení. □

Příklad 4.12 (Rozdělení podílu spojitých veličin pomocí věty o transformaci)**.** Uvažujme $Z = \frac{X}{Y}$ a dodefinujme $U = Y$. Mějme

$$t((X,Y)^T) = (Z,U)^T = \left(\frac{X}{Y}, Y\right)^T$$

prosté a regulární na $G_1 \cup G_2$, přičemž $G_1 = (-\infty, \infty) \times (-\infty, 0)$ a $G_2 = (-\infty, \infty) \times (0, \infty)$. Položme

$$\tau((Z,U)^T) = (X,Y)^T = (ZU, U)^T,$$

přičemž $t(G_1) = \mathbb{R}^2$ a $t(G_2) = \mathbb{R}^2$. Dle věty 4.5

$$\begin{aligned}
f_{ZU}(z,u) &= f_{X,Y}(\tau(z,u))|D_\tau| && \text{pro } (z,u)^T \in \mathbb{R}^2; \\
f_{ZU}(z,u) &= f_{X,Y}(zu,u)|u| = f_X(zu) f_Y(u)|u|.
\end{aligned}$$

Nakonec dle věty 3.3 získáme

$$f_Z(z) = \int_{-\infty}^\infty f_{ZU}(z,u)\,\mathrm{d}u = \int_{-\infty}^\infty f_X(zu) f_Y(u)|u|\,\mathrm{d}u.$$

▲

Poznámky.

(i) Distribuční funkce součtu dvou nezávislých diskrétních náhodných veličin má podle věty 4.6 (o konvoluci) tvar $F_Z(z) = \sum_x F_Y(z-x)P[X=x]$, neboť

$$F_Z(z) = P[X+Y \le z] = \sum_{x \in S_X} P[X=x,\, Y \le z-x] = \sum_{x \in S_X} P[X=x]P[Y \le z-x]$$
$$= \sum_{x \in S_X} P[X=x]F_Y(z-x) = \sum_{y \in S_Y} P[Y=y]F_X(z-y).$$

(ii) Distribuční funkce podílu dvou nezávislých diskrétních veličin má podle věty 4.9 tvar

$$F_Z(z) = \sum_{\substack{y \in S_Y \\ y>0}} P[Y=y]F_X(zy) + \sum_{\substack{y \in S_Y \\ y<0}} P[Y=y][1-F_X(zy)],$$

což lze odvodit v analogii s důkazem věty 4.9 a podobně jako v předešlé poznámce.

4.3. Rozdělení odvozená od normálního

1. Chí-kvadrát rozdělení $X_n \sim \mathcal{X}_n^2$ o n stupních volnosti
Nechť $\xi_1, \ldots, \xi_n \sim \mathrm{N}(0,1)$ jsou nezávislé a položme $X_n = \sum_{i=1}^n \xi_i^2$. Pro $n=1$ má náhodná veličina $X_1 = \xi_1^2$ distribuční funkci

$$F_1(x) = P[\xi_1^2 \le x] = P[|\xi_1| \le \sqrt{x}] = P[-\sqrt{x} \le \xi_1 \le \sqrt{x}] = \Phi(\sqrt{x}) - \Phi(-\sqrt{x})$$
$$= \Phi(\sqrt{x}) - (1-\Phi(\sqrt{x})) = 2\Phi(\sqrt{x}) - 1$$

pro $x > 0$. Tato veličina má hustotu ve tvaru

$$f_1(x) = F_1'(x) = \frac{2\phi(\sqrt{x})}{2\sqrt{x}} = \frac{1}{\sqrt{2\pi x}} \exp\left[-\frac{x}{2}\right] = \left(\frac{1}{2}\right)^{\frac{1}{2}} \frac{x^{\frac{1}{2}-1}}{\Gamma\left(\frac{1}{2}\right)} \exp\left[-\frac{x}{2}\right]$$

pro $x > 0$, takže $X_1 \sim \Gamma\left(\frac{1}{2}, \frac{1}{2}\right)$. Indukcí lze s využitím věty 4.8 dokázat, že X_n má hustotu

$$f_n(x) = \left(\frac{1}{2}\right)^{\frac{n}{2}} \frac{x^{\frac{n}{2}-1}}{\Gamma\left(\frac{n}{2}\right)} \exp\left[-\frac{x}{2}\right]$$

pro $x > 0$, tedy že $X_n \sim \Gamma\left(\frac{1}{2}, \frac{n}{2}\right) = \mathcal{X}_n^2$.
Označme

$$X = X_n \sim \mathcal{X}_n^2, \quad Y = \xi_{n+1}^2 \sim \mathcal{X}_1^2,$$

které jsou nezávislé, a

$$Z = X + Y = \sum_{i=1}^{n+1} \xi_i^2 = X_{n+1}.$$

Dle věty 4.8 má Z hustotu

$$f_{n+1}(z) = \int_0^z f_n(x) f_1(z-x)\,\mathrm{d}x = \frac{\exp\left[-\frac{z}{2}\right]}{2^{\frac{n+1}{2}}\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{1}{2}\right)} \int_0^z x^{\frac{n}{2}-1}(z-x)^{-\frac{1}{2}}\,\mathrm{d}x$$
$$= \frac{\exp\left[-\frac{z}{2}\right] z^{\frac{n}{2}-1-\frac{1}{2}+1}}{2^{\frac{n+1}{2}}\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{1}{2}\right)} \int_0^1 u^{\frac{n}{2}-1}(1-u)^{-\frac{1}{2}}\,\mathrm{d}u.$$

Ve třetí rovnosti jsme použili substituci $x = uz$. Všimněme si, že poslední integrál představuje beta funkci

$$\int_0^1 u^{\frac{n}{2}-1}(1-u)^{-\frac{1}{2}}\,\mathrm{d}u = B\left(\frac{n}{2}, \frac{1}{2}\right) = \frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{n+1}{2}\right)}.$$

Tento vztah dosadíme a získáme

$$f_{n+1}(z) = \frac{\exp\left[-\frac{z}{2}\right] z^{\frac{n+1}{2}-1}}{2^{\frac{n+1}{2}}\Gamma\left(\frac{n+1}{2}\right)}$$

pro $z > 0$, takže $X_{n+1} \sim \mathcal{X}_{n+1}^2$.

Poznámka. Zřejmě, dle věty 4.8, platí, že pokud jsou veličiny $X_n \sim \mathcal{X}_n^2$ a $Y_m \sim \mathcal{X}_m^2$ nezávislé, potom $Z = X_n + Y_m \sim \mathcal{X}_{n+m}^2$.

Nyní se podíváme na momenty rozdělení $\mathcal{X}_n^2$.

$$E[X_n] = E\left[\sum_{i=1}^n \xi_i^2\right] = \sum_{i=1}^n E\left[\xi_i^2\right] = \sum_{i=1}^n \mathrm{Var}[\xi_i] = n;$$

$$\mathrm{Var}[X_n] = \sum_{i=1}^n \mathrm{Var}[\xi_i^2] = n\left(E\left[\xi_1^4\right] - (E\left[\xi_1^2\right])^2\right) = n(3-1) = 2n.$$

Ve třetí rovnosti vztahu pro rozptyl jsme použili definici 2.16 a vlastnosti normálního rozdělení popsané v sekci 2.4. Podle nich se špičatost $\mathrm{N}(0,1)$ dá vyjádřit jako

$$\frac{E[\xi_1 - E[\xi_1]]^4}{(\mathrm{Var}[\xi_1])^2} = 3.$$

Věta 4.10 (Rozdělení kvadratické formy)**.** *Nechť* $\boldsymbol{X} \sim \mathrm{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, *kde* $\boldsymbol{\Sigma}$ *je regulární. Potom*

$$Y = (\boldsymbol{X} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{X} - \boldsymbol{\mu}) \sim \mathcal{X}_n^2.$$

Důkaz. Protože $\boldsymbol{X} \sim \mathrm{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ a $\boldsymbol{\Sigma}$ je regulární, tak, uvažujeme-li matici $\boldsymbol{A}$ typu $n \times n$, máme $\boldsymbol{X} = \boldsymbol{A}\boldsymbol{\xi} + \boldsymbol{\mu}$ a $\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{A}^T$, přičemž $\boldsymbol{\xi} = (\xi_1, \dots, \xi_n)^T$, kde $\xi_1, \dots, \xi_n \sim \mathrm{N}(0,1)$ jsou nezávislé. Potom

$$Y = (\boldsymbol{A}\boldsymbol{\xi})^T (\boldsymbol{A}\boldsymbol{A}^T)^{-1} \boldsymbol{A}\boldsymbol{\xi} = \boldsymbol{\xi}^T \boldsymbol{A}^T (\boldsymbol{A}^T)^{-1} \boldsymbol{A}^{-1} \boldsymbol{A}\boldsymbol{\xi} = \boldsymbol{\xi}^T \boldsymbol{\xi} = \sum_{i=1}^n \xi_i^2 \sim \mathcal{X}_n^2. \qquad \square$$

Úmluva. *Stopu matice budeme označovat jako* Tr *(anglicky trace). Hodnost matice budeme značit h.*

Věta 4.11 (Střední hodnota kvadratické formy)**.** *Nechť* $E[\boldsymbol{X}] = \boldsymbol{\mu}$ *a* $\mathrm{Var}[\boldsymbol{X}] = \boldsymbol{\Sigma}$, *kde* $\boldsymbol{X} = (X_1, \dots, X_n)^T$ *a* $X_i \in L^2(\Omega)$ *pro všechna* $i \in \mathbb{N}$. *Uvažujme matici* $\boldsymbol{A}$ *typu* $n \times n$. *Pak platí, že*

$$E\left[\boldsymbol{X}^T \boldsymbol{A} \boldsymbol{X}\right] = \mathrm{Tr}(\boldsymbol{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}^T \boldsymbol{A}^T \boldsymbol{\mu}.$$

Důkaz. Dle věty 3.5 máme

$$\boldsymbol{\Sigma} = \mathrm{Var}[\boldsymbol{X}] = E\left[\boldsymbol{X}\boldsymbol{X}^T\right] - E[\boldsymbol{X}](E[\boldsymbol{X}])^T = E[\boldsymbol{X}\boldsymbol{X}^T] - \boldsymbol{\mu}\boldsymbol{\mu}^T.$$

Dále počítejme

$$\begin{aligned} E[\boldsymbol{X}^T \boldsymbol{A} \boldsymbol{X}] &= E[\mathrm{Tr}(\boldsymbol{X}^T \boldsymbol{A} \boldsymbol{X})] = E\left[\mathrm{Tr}(\boldsymbol{A}\boldsymbol{X}\boldsymbol{X}^T)\right] = \mathrm{Tr}(\boldsymbol{A} E[\boldsymbol{X}\boldsymbol{X}^T]) = \mathrm{Tr}(\boldsymbol{A}\boldsymbol{\Sigma} + \boldsymbol{A}\boldsymbol{\mu}\boldsymbol{\mu}^T) \\ &= \mathrm{Tr}(\boldsymbol{A}\boldsymbol{\Sigma}) + \mathrm{Tr}(\boldsymbol{A}\boldsymbol{\mu}\boldsymbol{\mu}^T) = \mathrm{Tr}(\boldsymbol{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}^T \boldsymbol{A} \boldsymbol{\mu}. \qquad \square \end{aligned}$$

Poznámka. Pro $\boldsymbol{X} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$ a symetrickou pozitivně semidefinitní matici $\boldsymbol{A}_{n \times n}$ takovou, že $\boldsymbol{A}\boldsymbol{\Sigma} \neq \boldsymbol{0}$ je idempotentní (tj. $\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{\Sigma}$), platí

$$Y = \boldsymbol{X}^T \boldsymbol{A} \boldsymbol{X} \sim \mathcal{X}_r^2,$$

kde $r = h(\boldsymbol{A}\boldsymbol{\Sigma}) = \mathrm{Tr}(\boldsymbol{A}\boldsymbol{\Sigma})$. Zřejmě tedy $E\left[\boldsymbol{X}^T \boldsymbol{A} \boldsymbol{X}\right] = r = \mathrm{Tr}(\boldsymbol{A}\boldsymbol{\Sigma})$.

2. <u>T-rozdělení</u> $T_n \sim t_n$ o n stupních volnosti
   Mějme $Z \sim \mathrm{N}(0,1)$ a $X_n \sim \mathcal{X}_n^2$ nezávislé a položme

$$T_n = \frac{Z}{\sqrt{\frac{X_n}{n}}}.$$

Náhodná veličina T_n má hustotu

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

kde $t \in \mathbb{R}$. Lze ji odvodit pomocí vět 4.4 a 3.3. Položme $U = X_n$, $T = T_n$. Potom

$$t((Z, X_n)^T) = (T, U)^T = \left(\frac{Z\sqrt{n}}{\sqrt{X_n}}, X_n\right)^T,$$

kde $Z \in \mathbb{R}$ a $X_n > 0$, je prosté a regulární na $(-\infty, \infty) \times (0, \infty)$. Dále

$$\tau((T, U)^T) = (Z, X_n)^T = \left(\frac{T\sqrt{U}}{\sqrt{n}}, U\right)^T,$$

kde $T \in R$, $U > 0$. Determinant zobrazení τ je

$$D_\tau(t, u) = \det\begin{pmatrix} \frac{\partial z}{\partial t} & \frac{\partial z}{\partial u} \\ \frac{\partial x_n}{\partial t} & \frac{\partial x_n}{\partial u} \end{pmatrix} = \det\begin{pmatrix} \sqrt{\frac{u}{n}} & \frac{t}{2\sqrt{nu}} \\ 0 & 1 \end{pmatrix} = \sqrt{\frac{u}{n}} \neq 0.$$

Sdružené hustoty jsou ve tvaru

$$f_{Z,X_n}(z, x) = \frac{\exp\left[-\frac{z^2}{2}\right]}{\sqrt{2\pi}} \frac{\exp\left[-\frac{x}{2}\right] x^{\frac{n}{2}-1}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)};$$

$$f_{T,U}(t, u) = f_{Z,X_n}(\tau(t, u))|D_\tau(t, u)| = \frac{\exp\left[-\frac{t^2}{2}\frac{u}{n} - \frac{u}{2}\right] u^{\frac{n}{2}-1+\frac{1}{2}}}{2^{\frac{n+1}{2}}\Gamma(\frac{n}{2})\sqrt{\pi n}},$$

kde $z, t \in \mathbb{R}$, $x, u > 0$. Marginální hustotu získáme integrací

$$f_{T,n}(t) = \int_0^\infty \frac{\exp\left[-\frac{t^2}{2}\frac{u}{n} - \frac{u}{2}\right] u^{\frac{n}{2}-1+\frac{1}{2}}}{2^{\frac{n+1}{2}}\Gamma(\frac{n}{2})\sqrt{\pi n}}\,\mathrm{d}u,$$

přičemž lze použít substituci

$$y = \frac{u}{2}\left(1 + \frac{t^2}{n}\right).$$

Po úpravě dostaneme

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

pro $t \in \mathbb{R}$.

Speciálně pro $n = 1$ dostáváme hustotu

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

kde $t \in \mathbb{R}$, která definuje Cauchyovo rozdělení s parametry 0 a 1, píšeme $T \sim C(0, 1)$. Toto rozdělení nemá konečnou střední hodnotu. Obecné Cauchyovo rozdělení s parametry a, b má hustotu

$$f_X(x) = \frac{1}{\pi}\frac{b}{b^2 + (x-a)^2}$$

pro $a, x \in \mathbb{R}$, $b > 0$. Značíme $X \sim C(a, b)$.

Poznámka. Podle Williama S. Gosseta (1876-1937), který používal pseudonym Student, se t-rozdělení také označuje jako Studentovo rozdělení. Gosset pracoval v pivovaru Guiness v Dublinu a pseudonym používal při publikování svých matematických textů, které obsahovaly počátky aplikací matematické statistiky v průmyslu a zemědělství.

Nakonec si uvedeme čtyři vlastnosti t-rozdělení. Plyne z nich, že s rostoucím počtem stupňů volnosti n se rozdělení t_n blíží normovanému normálnímu rozdělení.

(a) Hustota $f_{T,n}$ je symetrická kolem nuly.

(b) $\lim_{n\longrightarrow\infty} |f_{T,n}(t) - \phi(t)| = 0$, kde $\phi(t) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{t^2}{2}\right]$ pro $t \in \mathbb{R}$.

(c) $\lim_{n\longrightarrow\infty} t_n(\alpha) = u(\alpha)$, kde $\alpha \in (0,1)$. Zde $t_n(\alpha)$ je α-kvantil t_n a $u(\alpha)$ je α-kvantil $\mathrm{N}(0,1)$.

(d) Náhodná veličina $T_n \sim t_n$ má konečné momenty do řádu $n-1$, $E[T_n] = 0$ pro $n > 1$ podle bodu (a) a $\mathrm{Var}[T_n] = \frac{n}{n-2}$ pro $n > 2$.

3. <u>F-rozdělení</u> $F \sim F_{n,m}$ o n a m stupních volnosti
Nechť $X_n \sim \mathcal{X}^2_n$ a $Y_m \sim \mathcal{X}^2_m$ jsou nezávislé a nechť pro náhodnou veličinu F platí

$$F = \frac{\frac{X_n}{n}}{\frac{Y_m}{m}}.$$

Potom F má hustotu

$$f_{n,m}(z) = \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)} \left(\frac{n}{m}\right)^{\frac{n}{2}} z^{\frac{n}{2}-1} \left(1 + \frac{nz}{m}\right)^{-\frac{n+m}{2}},$$

kde $z > 0$. Tuto hustotu lze odvodit pomocí věty 4.1 a důsledku 4.3. Veličina X_n má hustotu

$$f_n(x) = \frac{\exp\left[-\frac{x}{2}\right] x^{\frac{n}{2}-1}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}$$

pro $x > 0$. Dále

$$X = t(X_n) = \frac{X_n}{n}, \quad X_n = \tau(X) = nX, \quad \tau'(X) = n.$$

Dle věty 4.1 má X hustotu

$$f_n(\tau(x))|\tau'(x)| = \frac{n \exp\left[-\frac{nx}{2}\right] (nx)^{\frac{n}{2}-1}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} = n f_n(nx)$$

pro $x > 0$. Analogicky, $Y = \frac{Y_m}{m}$ má hustotu $m f_m(my)$ pro $y > 0$. Navíc, podle věty 2.20, jsou X a Y nezávislé. Podle důsledku 4.3 má

$$F = \frac{\frac{X_n}{n}}{\frac{Y_m}{m}} = \frac{X}{Y}$$

hustotu ve tvaru

$$\begin{aligned}\int_0^\infty y m f_m(my) n f_n(nzy)\,\mathrm{d}y &= \frac{nm(n)^{\frac{n}{2}-1}(m)^{\frac{m}{2}-1}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)} \int_0^\infty y y^{\frac{m}{2}-1} \exp\left[-\frac{my}{2}\right] (zy)^{\frac{n}{2}-1} \exp\left[-\frac{nzy}{2}\right]\,\mathrm{d}y \\ &= \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)} \left(\frac{n}{m}\right)^{\frac{n}{2}} z^{\frac{n}{2}-1} \left(1 + \frac{nz}{m}\right)^{-\frac{m+n}{2}}\end{aligned}$$

pro $z > 0$. V úpravách jsme použili substituci $u = \frac{y}{2}(m + nz)$ a vztah

$$\int_0^\infty u^{\frac{n+m}{2}-1} e^{-u}\,\mathrm{d}u = \Gamma\left(\frac{n+m}{2}\right).$$

Poznámka. F-rozdělení se také nazývá Fisherovo nebo Fisherovo-Snedecorovo rozdělení. Sir R. A. Fisher (1890-1962) byl významný britský statistik, zakladatel různých statistických metod např. v oblasti teorie odhadu či teorie navrhování experimentů. V letech 1943-1957 byl profesorem genetiky na univerzitě v Cambridgi. G. W. Snedecor (1881-1974) byl americký statistik, který se zabýval aplikacemi statistických metod v biologii a zemědělství.

5. Limitní věty pro náhodné vektory

5.1. Konvergence náhodných vektorů

Uvažujme posloupnost náhodných vektorů $\{\boldsymbol{X}_n\}_{n=1}^{\infty}$, kde $\boldsymbol{X}_n = (X_{n1}, \dots, X_{nk})^T : (\Omega, \mathcal{A}) \longrightarrow (\mathbb{R}^k, \mathcal{B}_0^k)$ pro $k \geq 1$ a $n = 1, 2, \dots$ Uvedeme různé způsoby konvergence a vztahy mezi nimi.

Definice 5.1 (Norma vektoru)**.** *Eukleidovská norma* vektoru $\boldsymbol{X}$ je definována předpisem

$$||\boldsymbol{X}|| = \sqrt{\boldsymbol{X}^T \boldsymbol{X}}.$$

Poznámka. Pro $\boldsymbol{X}_n \in \mathbb{R}^k$ je zřejmě

$$||\boldsymbol{X}_n|| = \sqrt{\sum_{i=1}^{k} X_{ni}^2}$$

a pro $k = 1$ máme $||\boldsymbol{X}|| = \sqrt{X_n^2} = |X_n|$.

Definice 5.2 (Konvergence v pravděpodobnosti)**.** Řekneme, že posloupnost náhodných vektorů $\{\boldsymbol{X}_n\}_{n=1}^{\infty}$ *konverguje v pravděpodobnosti* k náhodnému vektoru $\boldsymbol{X}$, pokud pro libovolné $\varepsilon > 0$ platí

$$\lim_{n \longrightarrow \infty} P[||\boldsymbol{X}_n - \boldsymbol{X}|| \geq \varepsilon] = 0.$$

Značíme $\boldsymbol{X}_n \xrightarrow{P} \boldsymbol{X}$.

Poznámky.

(i) Limitou může být i nenáhodný vektor $\boldsymbol{X}$.

(ii) Pro $k = 1$ dostáváme definici 2.27.

(iii) Platí, že $\boldsymbol{X}_n \xrightarrow{P} \boldsymbol{X}$ právě tehdy, když $||\boldsymbol{X}_n - \boldsymbol{X}|| \xrightarrow{P} 0$.

(iv) Postačující podmínky pro konvergenci v pravděpodobnosti náhodných veličin udává věta 2.24.

Definice 5.3 (Konvergence v distribuci)**.** Řekneme, že posloupnost náhodných vektorů $\{\boldsymbol{X}_n\}_{n=1}^{\infty}$ *konverguje v distribuci* k náhodnému vektoru $\boldsymbol{X}$, pokud platí

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

v každém bodě spojitosti $\boldsymbol{x}$ limitní funkce $F_{\boldsymbol{X}}$. Značíme

$$\boldsymbol{X}_n \xrightarrow{D} \boldsymbol{X},\ F_{\boldsymbol{X}_n} \longrightarrow F_{\boldsymbol{X}}; \quad \mathcal{L}(\boldsymbol{X}_n) \longrightarrow \mathcal{L}(\boldsymbol{X}),\ \boldsymbol{X}_n \overset{\text{as.}}{\sim} \mathcal{L}(\boldsymbol{X}); \quad \boldsymbol{X}_n \xrightarrow{D} \mathcal{L}(\boldsymbol{X}).$$

Poznámka. Pro $k = 1$ dostáváme definici 2.28.

Definice 5.4 (Konvergence skoro jistě)**.** Řekneme, že posloupnost náhodných vektorů $\{\boldsymbol{X}_n\}_{n=1}^{\infty}$ *konverguje skoro jistě* k náhodnému vektoru $\boldsymbol{X}$, pokud platí

$$P[X_{n1} \longrightarrow X_1, \dots, X_{nk} \longrightarrow X_k] = 1.$$

Značíme $\boldsymbol{X}_n \xrightarrow{s.j.} \boldsymbol{X}$ nebo $\boldsymbol{X}_n \longrightarrow \boldsymbol{X}$ *s.j.*

Poznámka. Limitou u konvergence v distribuci a u konvergence skoro jistě může být i nenáhodný vektor $\boldsymbol{X}$.

Věta 5.1 (O spojité transformaci)**.** *Uvažujme posloupnost náhodných vektorů* $\{\boldsymbol{X}_n\}_{n=1}^{\infty}$ *a spojitou funkci* $g: \mathbb{R}^k \longrightarrow \mathbb{R}^m$. *Pak platí* *(i)* $\boldsymbol{X}_n \xrightarrow{P} \boldsymbol{X} \implies g(\boldsymbol{X}_n) \xrightarrow{P} g(\boldsymbol{X})$,

(ii) $\boldsymbol{X}_n \xrightarrow{D} \boldsymbol{X} \implies g(\boldsymbol{X}_n) \xrightarrow{D} g(\boldsymbol{X})$,

(iii) $\boldsymbol{X}_n \xrightarrow{s.\,j.} \boldsymbol{X} \implies g(\boldsymbol{X}_n) \xrightarrow{s.\,j.} g(\boldsymbol{X})$.

Důkaz. Vynechán. □

Důsledek 5.1. *Pro $m = 1$ a $g(\boldsymbol{x}) = \sum_{i=1}^k x_i$ dostaneme, že pokud $\boldsymbol{X}_n \longrightarrow \boldsymbol{X}$, potom $\sum_{i=1}^k X_{ni} \longrightarrow \sum_{i=1}^k X_i$ pro všechny tři případy (P, D, s. j.). Podobně pro aritmetické průměry; pokud $\boldsymbol{X}_n \longrightarrow \boldsymbol{X}$, potom $\frac{1}{k}\sum_{i=1}^k X_{ni} \longrightarrow \frac{1}{k}\sum_{i=1}^k X_i$ pro všechny tři případy (P, D, s. j.). Analogicky pro $g(\boldsymbol{x}) = \prod_{i=1}^k x_i$ apod.*

Důsledek 5.2. *Uvažujme $m = 1$ a $g(\boldsymbol{x}) = \boldsymbol{c}_i^T \boldsymbol{x}$, kde $\boldsymbol{c}_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T$ s jedničkou na i-té pozici. Potom když $\boldsymbol{X}_n \longrightarrow \boldsymbol{X}$, pak $\boldsymbol{c}_i^T \boldsymbol{X}_n = X_{ni} \longrightarrow X_i = \boldsymbol{c}_i^T \boldsymbol{X}$ pro $i = 1, \ldots, k$. To znamená, že z konvergence náhodných vektorů $\boldsymbol{X}_n$ (P, D, s. j.) plyne konvergence jejich složek X_{ni}.*

Věta 5.2 (Ekvivalentní podmínky konvergence v distribuci)**.** *Mějme posloupnost náhodných vektorů $\{\boldsymbol{X}_n\}_{n=1}^\infty$ a posloupnost náhodných veličin $\{X_n\}_{n=1}^\infty$.*

(i) Pro posloupnost náhodných vektorů $\{\boldsymbol{X}_n\}_{n=1}^\infty$ platí

$$\boldsymbol{X}_n \xrightarrow{D} \boldsymbol{X} \text{ právě tehdy, když } \boldsymbol{c}^T \boldsymbol{X}_n \xrightarrow{D} \boldsymbol{c}^T \boldsymbol{X}$$

pro $n \longrightarrow \infty$ a pro všechny $\boldsymbol{c} \in \mathbb{R}^k$.

(ii) Pro posloupnost náhodných veličin $\{X_n\}_{n=1}^\infty$ platí

$$X_n \xrightarrow{D} X \text{ právě tehdy, když } E[g(X_n)] \longrightarrow E[g(X)]$$

pro $n \longrightarrow \infty$ a pro každou funkci $g : \mathbb{R} \longrightarrow \mathbb{R}$ spojitou a omezenou.

Důkaz. Vynechán. □

Nyní se podíváme na to, zda z konvergence složek plyne konvergence celých vektorů. Podle bodu (i) věty 5.2 je postačující podmínkou pro $\boldsymbol{X}_n \xrightarrow{D} \boldsymbol{X}$ to, že $\boldsymbol{c}^T \boldsymbol{X}_n \xrightarrow{D} \boldsymbol{c}^T \boldsymbol{X}$ pro všechny $\boldsymbol{c} \in \mathbb{R}^k$. To znamená, že máme-li vektor $\boldsymbol{c}_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T$ s jedničkou na i-té pozici, nestačí pouze pro tento jeden vektor $\boldsymbol{c}_i$

$$\boldsymbol{c}_i^T \boldsymbol{X}_n = X_{ni} \xrightarrow{D} \boldsymbol{c}_i^T \boldsymbol{X} = X_i,$$

kde $i = 1, \ldots, k$. Tedy <u>konvergenci v distribuci nelze definovat po složkách</u>.

Implikace když $X_{ni} \xrightarrow{D} X_i$ pro $i = 1, \ldots, k$, pak $\boldsymbol{X}_n \xrightarrow{D} \boldsymbol{X}$, platí pouze tehdy, když jsou složky X_{ni} navzájem nezávislé. Souvisí to s tím, že z marginálních rozdělení složek nelze jednoznačně určit sdružené rozdělení celého vektoru; to je možné pouze při nezávislých složkách.

Nechť $i = 1, \ldots, k$ a $n = 1, 2, \ldots$ Pokud jsou X_{ni} nezávislé, potom

$$F_{\boldsymbol{X}_n}(x_1, \ldots, x_k) = \prod_{i=1}^k F_{X_{ni}}(x_i).$$

Dále pokud $X_{ni} \xrightarrow{D} X_i$, potom $F_{X_{ni}}(x) \longrightarrow F_{X_i}(x)$ v bodech spojitosti F_{X_i}. Z toho plyne, že

$$\prod_{i=1}^k F_{X_{ni}}(x_i) \longrightarrow \prod_{i=1}^k F_{X_i}(x_i)$$

v bodech spojitosti $x_1, \ldots, x_k$ funkcí $F_{X_1}, \ldots, F_{X_k}$. Dostáváme

$$F_{\boldsymbol{X}_n}(x_1, \ldots, x_k) \longrightarrow F_{\boldsymbol{X}}(x_1, \ldots, x_k),$$

z čehož plyne $\boldsymbol{X}_n \xrightarrow{D} \boldsymbol{X}$, kde složky X_i limitního vektoru $\boldsymbol{X}$ jsou nezávislé.

Dále si uvědomíme, že $P[X_{ni} \longrightarrow X_i] = 1$ pro $i = 1, \ldots, k$ implikuje $P\left[\bigcap_{i=1}^k (X_{ni} \longrightarrow X_i)\right] = 1$. Tím jsme zjistili, že

$$\text{pokud } X_{ni} \xrightarrow{s.\,j.} X_i,\ i = 1, \ldots, k, \quad \text{pak } \boldsymbol{X}_n \xrightarrow{s.\,j.} \boldsymbol{X},$$

neboli že <u>konvergenci skoro jistě lze definovat po složkách</u>.

Věta 5.3 (Konvergence v pravděpodobnosti po složkách)**.** *Nechť* $X_{ni} \xrightarrow{P} X_i$ *pro* $i = 1, \ldots, k$. *Pak* $\boldsymbol{X}_n \xrightarrow{P} \boldsymbol{X}$.

Důkaz. Uvažujme $\varepsilon > 0$ a $i = 1, \ldots, k$. Potom z definice 2.27 máme

$$\lim_{n \longrightarrow \infty} P[|X_{ni} - X_i| \geq \varepsilon] = 0 \implies \lim_{n \longrightarrow \infty} P[(X_{ni} - X_i)^2 < \varepsilon^2] = 1.$$

Dle bodu (iv) věty 1.3, který platí i pro konečně mnoho jevů, máme

$$1 \geq P\left[\bigcap_{i=1}^{k}((X_{ni} - X_i)^2 < \varepsilon^2)\right] \geq 1 - \sum_{i=1}^{k} P[(X_{ni} - X_i)^2 \geq \varepsilon^2] \longrightarrow 1$$

pro $n \longrightarrow \infty$. Když $(X_{ni} - X_i)^2 < \varepsilon^2$ pro všechna $i = 1, \ldots, k$, tak $\sum_{i=1}^{k}(X_{ni} - X_i)^2 < k\varepsilon^2$ a dostaneme

$$P\left[\bigcap_{i=1}^{k}((X_{ni} - X_i)^2 < \varepsilon^2)\right] \leq P\left[\sum_{i=1}^{k}(X_{ni} - X_i)^2 < k\varepsilon^2\right] \leq 1.$$

Tedy $\lim_{n \longrightarrow \infty} P\left[||\boldsymbol{X}_n - \boldsymbol{X}||^2 < k\varepsilon^2\right] = 1$ a pro libovolné $\varepsilon^* = \sqrt{k}\varepsilon$ máme $\lim_{n \longrightarrow \infty} P\left[||\boldsymbol{X}_n - \boldsymbol{X}|| \geq \varepsilon^*\right] = 0$. Dle definice 5.2 jsme získali $\boldsymbol{X}_n \xrightarrow{P} \boldsymbol{X}$. □

Závěr tedy je, že konvergenci v pravděpodobnosti lze definovat po složkách.
Dále nás bude zajímat vztah mezi jednotlivými typy konvergencí.

Věta 5.4 (Vztah konvergence skoro jistě a v pravděpodobnosti)**.** *Pro posloupnost náhodných vektorů* $\{\boldsymbol{X}_n\}_{n=1}^{\infty}$ *platí, že když* $\boldsymbol{X}_n \xrightarrow{s.\,j.} \boldsymbol{X}$, *potom také* $\boldsymbol{X}_n \xrightarrow{P} \boldsymbol{X}$ *pro* $n \longrightarrow \infty$.

Důkaz. Vynechán. □

Věta 5.5 (Vztah konvergence v pravděpodobnosti a v distribuci)**.** *Pro posloupnost náhodných vektorů* $\{\boldsymbol{X}_n\}_{n=1}^{\infty}$ *platí, že když* $\boldsymbol{X}_n \xrightarrow{P} \boldsymbol{X}$, *potom také* $\boldsymbol{X}_n \xrightarrow{D} \boldsymbol{X}$ *pro* $n \longrightarrow \infty$.

Důkaz. Vynechán. □

Poznámka. Pokud je limitou nenáhodný vektor $\boldsymbol{X}$, platí ve větě 5.5 i opačná implikace.

Úmluva. *Náhodnou maticí rozumíme matici, která má jako prvky náhodné veličiny.*

Věta 5.6 (Cramérova-Sluckého věta)**.** *Nechť* $\boldsymbol{X}_n$, $\boldsymbol{X} \in \mathbb{R}^k$ *jsou náhodné vektory,* $\boldsymbol{B}_n$ *jsou náhodné matice typu* $m \times k$, $\boldsymbol{B}$ *je matice konstant typu* $m \times k$, $\boldsymbol{a}_n \in \mathbb{R}^m$ *jsou náhodné vektory,* $\boldsymbol{a} \in \mathbb{R}^m$ *je vektor konstant a* $n \longrightarrow \infty$. *Dále nechť platí*

$$\boldsymbol{X}_n \xrightarrow{D} \boldsymbol{X}, \quad \boldsymbol{a}_n \xrightarrow{P} \boldsymbol{a}, \quad \boldsymbol{B}_n \xrightarrow{P} \boldsymbol{B}.$$

Potom

$$\boldsymbol{a}_n + \boldsymbol{B}_n\boldsymbol{X}_n \xrightarrow{D} \boldsymbol{a} + \boldsymbol{B}\boldsymbol{X}.$$

Důkaz. Vynechán. □

5.2. Limitní věty

Pro náhodné vektory lze stejně jako pro náhodné veličiny formulovat zákon velkých čísel a centrální limitní větu. Uvažujme posloupnost náhodných vektorů $\{\boldsymbol{X}_n\}_{n=1}^{\infty}$ takovou, že $E[\boldsymbol{X}_n] = \boldsymbol{\mu} \in \mathbb{R}^k$ a $\mathrm{Var}[\boldsymbol{X}_n] = \boldsymbol{\Sigma}$ pro $n = 1, 2, \ldots$ Pro posloupnost průměrů $\{\overline{\boldsymbol{X}}_n\}_{n=1}^{\infty}$ platí

$$\overline{\boldsymbol{X}}_n = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{X}_i,$$

to jest

$$\overline{\boldsymbol{X}}_n = (\overline{X}_{n1}, \ldots, \overline{X}_{nk})^T, \text{ kde } \overline{X}_{nj} = \frac{1}{n}\sum_{i=1}^{n}X_{ij}.$$

Věta 5.7 (Chinčinův slabý zákon velkých čísel)**.** *Uvažujme posloupnost nezávislých stejně rozdělených náhodných vektorů* $\{\boldsymbol{X}_n\}_{n=1}^{\infty}$ *a nechť* $E[\boldsymbol{X}_n] = \boldsymbol{\mu}$ *má všechny složky konečné. Potom* $\overline{\boldsymbol{X}}_n \xrightarrow{P} \boldsymbol{\mu}$ *pro* $n \longrightarrow \infty$.

Důkaz. Dle věty 3.10 a důsledku 3.3 víme, že pokud jsou $\boldsymbol{X}_n$ nezávislé, potom jsou také X_{ni} nezávislé pro pevné $1 \leq i \leq k$. Pak dle věty 2.26 $\overline{X}_{ni} \xrightarrow{P} \mu_i$. Nakonec dle věty 5.3 dostáváme $\overline{\boldsymbol{X}}_n \xrightarrow{P} \boldsymbol{\mu}$. □

Věta 5.8 (Kolmogorovův silný zákon velkých čísel)**.** *Uvažujme posloupnost nezávislých stejně rozdělených náhodných vektorů* $\{\boldsymbol{X}_n\}_{n=1}^{\infty}$ *a nechť* $E[\boldsymbol{X}_n] = \boldsymbol{\mu}$ *má všechny složky konečné. Potom* $\overline{\boldsymbol{X}}_n \xrightarrow{s.j.} \boldsymbol{\mu}$ *pro* $n \longrightarrow \infty$.

Důkaz. Vynechán. □

Věta 5.9 (Centrální limitní věta)**.** *Uvažujme posloupnost nezávislých stejně rozdělených náhodných vektorů* $\{\boldsymbol{X}_n\}_{n=1}^{\infty}$, *pro které* $E[\boldsymbol{X}_n] = \boldsymbol{\mu}$ *a* $\mathrm{Var}[\boldsymbol{X}_n] = \boldsymbol{\Sigma} = (\sigma_{ij})_{i,j=1}^{k}$, *kde* $\sigma_{ij} < \infty$ *pro všechny dvojice* (i, j). *Potom*

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

pro $n \longrightarrow \infty$.

Důkaz. Vynechán. □

Poznámka. Speciálně pro $k = 1$ máme

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - \mu) \xrightarrow{D} \mathrm{N}(0, \sigma^2), \text{ neboli}$$

$$\frac{1}{\sqrt{n}\sigma} \left(\sum_{i=1}^{n} X_i - n\mu \right) \xrightarrow{D} \mathrm{N}(0, 1)$$

dle vět 2.8 a 5.1. Tím jsme dostali tvrzení věty 2.28.

Důsledek 5.3. *Nechť* $\{\boldsymbol{X}_n\}_{n=1}^{\infty}$ *splňuje předpoklady věty 5.9. Potom platí*

$$\sqrt{n}(\overline{\boldsymbol{X}}_n - \boldsymbol{\mu}) \xrightarrow{D} \mathrm{N}_k(\boldsymbol{0}, \boldsymbol{\Sigma})$$

pro $n \longrightarrow \infty$.

Důkaz. Důsledek plyne z věty 5.9 a z toho, že

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\boldsymbol{X}_i - \boldsymbol{\mu}) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \boldsymbol{X}_i - \frac{n}{\sqrt{n}} \boldsymbol{\mu} = \sqrt{n} \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{X}_i - \sqrt{n}\boldsymbol{\mu} = \sqrt{n}(\overline{\boldsymbol{X}}_n - \boldsymbol{\mu}).$$ □

Poznámka. Speciálně pro $k = 1$ máme

$$\sqrt{n}(\overline{X}_n - \mu) \xrightarrow{D} \mathrm{N}(0, \sigma^2), \text{ neboli}$$

$$\frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \xrightarrow{D} \mathrm{N}(0, 1).$$

Zřejmě platí za předpokladů věty 5.9

$$E[\overline{\boldsymbol{X}}_n] = \frac{1}{n} \sum_{i=1}^{n} E[\boldsymbol{X}_i] = \boldsymbol{\mu},$$

$$\mathrm{Var}[\overline{\boldsymbol{X}}_n] = \frac{1}{n^2} \mathrm{Var}\left[\sum_{i=1}^{n} \boldsymbol{X}_i \right] = \frac{1}{n^2} \left(\sum_{i=1}^{n} \mathrm{Var}\left[\boldsymbol{X}_i\right] + \sum_{i \neq j}^{n} \sum^{n} \mathrm{cov}(\boldsymbol{X}_i, \boldsymbol{X}_j) \right)$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} \mathrm{Var}[\boldsymbol{X}_i] = \frac{1}{n^2} n\boldsymbol{\Sigma} = \frac{1}{n}\boldsymbol{\Sigma}.$$

Pro pevné dostatečně velké n lze tedy symbolicky psát

$$\overline{\boldsymbol{X}}_n \stackrel{\cdot}{\sim} \mathrm{N}_k\left(\boldsymbol{\mu}, \frac{1}{n}\boldsymbol{\Sigma}\right),$$

přičemž pro $n \longrightarrow \infty$ platí

$$\overline{\boldsymbol{X}}_n \xrightarrow{P} \boldsymbol{\mu}, \quad \overline{\boldsymbol{X}}_n \xrightarrow{s.j.} \boldsymbol{\mu}, \quad \frac{1}{n}\boldsymbol{\Sigma} \longrightarrow \mathbf{0}.$$

Poznámka. Srovnejme ZVČ a CLV - pro nezávislé stejně rozdělené náhodné vektory platí

(i) ZVČ:

$$\overline{\boldsymbol{X}}_n - \boldsymbol{\mu} \longrightarrow \mathbf{0} \quad \ldots \text{ vektor konstant,}$$

$$\frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{X}_i - \boldsymbol{\mu}) \longrightarrow \mathbf{0} \quad \ldots \text{ vektor konstant,}$$

a to skoro jistě (věta 5.8), v pravděpodobnosti (věta 5.7 nebo 5.4), v distribuci (věta 5.5).

(ii) CLV:

$$\sqrt{n}(\overline{\boldsymbol{X}}_n - \boldsymbol{\mu}) \xrightarrow{D} N_k(\mathbf{0}, \boldsymbol{\Sigma}) \quad \ldots \text{ náhodný vektor,}$$

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\boldsymbol{X}_i - \boldsymbol{\mu}) \xrightarrow{D} N_k(\mathbf{0}, \boldsymbol{\Sigma}) \quad \ldots \text{ náhodný vektor.}$$

Příklad 5.1 (Aproximace $\mathcal{X}^2$ rozdělení normálním rozdělením)**.** Nechť $X_n \sim \mathcal{X}_n^2$. Zřejmě platí

$$X_n = \sum_{i=1}^{n}\xi_i^2, \quad \xi_i \sim \mathrm{N}(0,1) \text{ nezávislé,}$$

$$E[X_n] = n,$$

$$\mathrm{Var}[X_n] = 2n.$$

Dle CLV (věta 5.9 nebo 2.28) dostaneme

$$\frac{X_n - n}{\sqrt{2n}} \xrightarrow{D} \mathrm{N}(0,1).$$

Symbolicky pro pevné dostatečně velké n píšeme $X_n \stackrel{\cdot}{\sim} \mathrm{N}(n, 2n)$. Pravděpodobnosti typu $P[a < X_n < b]$ lze tedy namísto integrace hustoty rozdělení $\mathcal{X}_n^2$ počítat rozdílem $\Phi(b^*) - \Phi(a^*)$, kde Φ je distribuční funkce $\mathrm{N}(0,1)$ a

$$a^* = \frac{a-n}{\sqrt{2n}}, \quad b^* = \frac{b-n}{\sqrt{2n}}.$$ ▲

Příklad 5.2 (Aproximace Gamma rozdělení normálním rozdělením)**.** Uvažujme $Z_n \sim \Gamma\left(\frac{1}{\lambda}, n\right)$. Zřejmě platí

$$Z_n = \sum_{i=1}^{n} X_i, \quad X_i \sim \mathrm{Exp}(\lambda) \text{ nezávislé,}$$

$$E[Z_n] = n\lambda,$$

$$\mathrm{Var}[Z_n] = n\lambda^2.$$

Dle CLV (věta 5.9 nebo 2.28) máme

$$\frac{Z_n - n\lambda}{\sqrt{n}\lambda} \xrightarrow{D} \mathrm{N}(0,1).$$

Symbolicky pro pevné dostatečně velké n značíme $Z_n \stackrel{\cdot}{\sim} \mathrm{N}(n\lambda, n\lambda^2)$. Vzhledem k tomu, že $X_i \sim \mathrm{Exp}(\lambda)$ modelují například doby mezi poruchami stroje, má doba do n-té poruchy přibližně normální rozdělení. Díky CLV hraje normální rozdělení klíčovou roli v pravděpodobnostním modelování a matematické statistice. ▲

Příklad 5.3 (Aproximace t-rozdělení normálním rozdělením)**.** Nechť $T_n \sim t_n$. Zřejmě platí

$$T_n = \frac{Z}{\sqrt{\frac{X_n}{n}}}, \quad Z \sim \mathrm{N}(0,1), \quad X_n \sim \mathcal{X}_n^2 \text{ nezávislé }.$$

Dle ZVČ (věta 5.7 nebo 2.26) platí

$$\frac{1}{n}\sum_{i=1}^{n} \xi_i^2 \xrightarrow{P} 1$$

pro $n \longrightarrow \infty$, přičemž

$$\xi_i \sim \mathrm{N}(0,1) \text{ nezávislé},$$

$$X_n = \sum_{i=1}^{n} \xi_i^2,$$

$$E[\xi_i^2] = \mathrm{Var}[\xi_i] = 1.$$

Dle věty 5.1 platí

$$\frac{1}{\sqrt{\frac{X_n}{n}}} \xrightarrow{P} 1$$

pro $n \longrightarrow \infty$. Dále položme ve větě 5.6

$$\boldsymbol{X}_n = \boldsymbol{X} = Z \sim \mathrm{N}(0,1)$$

$$B_n = \frac{1}{\sqrt{\frac{X_n}{n}}}, \quad B = 1,$$

$$\boldsymbol{a}_n = \boldsymbol{a} = \boldsymbol{0}.$$

Odtud dostáváme $B_n \boldsymbol{X}_n = T_n \xrightarrow{D} Z \sim \mathrm{N}(0,1)$. Tedy pro velké n lze např. kvantily t-rozdělení nahradit kvantily normálního rozdělení s parametry 0, 1. ▲

Věta 5.10 (Delta metoda)**.** *Nechť posloupnost náhodných vektorů* $\{\boldsymbol{T}_n\}_{n=1}^{\infty}$ *splňuje* $\sqrt{n}(\boldsymbol{T}_n - \boldsymbol{\mu}) \xrightarrow{D} \mathrm{N}_k(\boldsymbol{0}, \boldsymbol{\Sigma})$ *pro* $n \longrightarrow \infty$*, kde* $\boldsymbol{\mu} \in \mathbb{R}^k$ *je vektor konstant a* $\boldsymbol{\Sigma}$ *je matice konstant. Dále nechť* $g : \mathbb{R}^k \longrightarrow \mathbb{R}^m$ *je spojitě diferencovatelná funkce a označme*

$$\boldsymbol{D}(\boldsymbol{x}) = \frac{\partial g(\boldsymbol{x})}{\partial \boldsymbol{x}} = \begin{pmatrix} \frac{\partial g_1(\boldsymbol{x})}{\partial x_1} & \cdots & \frac{\partial g_1(\boldsymbol{x})}{\partial x_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_m(\boldsymbol{x})}{\partial x_1} & \cdots & \frac{\partial g_m(\boldsymbol{x})}{\partial x_k} \end{pmatrix}.$$

Potom platí

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

pro $n \longrightarrow \infty$.

Důkaz. Podle Taylorova rozvoje v $\mathbb{R}^k$ máme

$$g(\boldsymbol{T}_n) = g(\boldsymbol{\mu}) + \boldsymbol{D}(\boldsymbol{T}_n^*)(\boldsymbol{T}_n - \boldsymbol{\mu}),$$

$$\boldsymbol{T}_n^* \in M = \{(1-t)\boldsymbol{T}_n + t\boldsymbol{\mu};\ t \in [0,1]\}.$$

Odtud dostáváme vztah

$$\sqrt{n}(g(\boldsymbol{T}_n) - g(\boldsymbol{\mu})) = \boldsymbol{D}(\boldsymbol{T}_n^*)\sqrt{n}(\boldsymbol{T}_n - \boldsymbol{\mu}).$$

Lze ukázat, že $\boldsymbol{D}(\boldsymbol{T}_n^*) \xrightarrow{P} D(\boldsymbol{\mu})$ pro $n \longrightarrow \infty$. Dále položme ve větě 5.6

$$\begin{aligned} \boldsymbol{X}_n &= \sqrt{n}(\boldsymbol{T}_n - \boldsymbol{\mu}), & \boldsymbol{X} &\sim \mathrm{N}_k(\boldsymbol{0}, \boldsymbol{\Sigma}), \\ \boldsymbol{B}_n &= \boldsymbol{D}(\boldsymbol{T}_n^*), & \boldsymbol{B} &= \boldsymbol{D}(\boldsymbol{\mu}), \\ \boldsymbol{a}_n &= \boldsymbol{0}, & \boldsymbol{a} &= \boldsymbol{0}. \end{aligned}$$

Dostaneme

$$\boldsymbol{B}_n \boldsymbol{X}_n = \boldsymbol{D}(\boldsymbol{T}_n^*)\sqrt{n}(\boldsymbol{T}_n - \boldsymbol{\mu}) \xrightarrow{D} \boldsymbol{D}(\boldsymbol{\mu})\boldsymbol{X},$$

přičemž dle věty 3.5 a důsledku 3.5 platí, že

$$\boldsymbol{X} \sim \mathrm{N}_k(\boldsymbol{0}, \boldsymbol{\Sigma}) \implies \boldsymbol{D}(\boldsymbol{\mu})\boldsymbol{X} \sim \mathrm{N}_m(\boldsymbol{0}, \boldsymbol{D}(\boldsymbol{\mu})\boldsymbol{\Sigma}\boldsymbol{D}(\boldsymbol{\mu})^T),$$

čímž je věta dokázána. □

Poznámka. Pro $\boldsymbol{T}_n = \overline{\boldsymbol{X}}_n$ a $g(\boldsymbol{x}) = \boldsymbol{x}$ dostáváme z věty 5.10 důsledek 5.3.

Příklad 5.4 (Rozdělení logaritmu průměru)**.** Nechť $X_1, X_2, \dots$ jsou nezávislé náhodné veličiny, pro které platí

$$X_i > 0, \quad E[X_i] = \mu, \quad \mathrm{Var}[X_i] = \sigma^2 < \infty$$

pro všechna $i = 1, 2, \dots$ Díky důsledku 5.3 víme, že $\sqrt{n}(\overline{X}_n - \mu) \xrightarrow{D} \mathrm{N}(0, \sigma^2)$ pro $n \longrightarrow \infty$. Nechť $g(x) = \log(x)$. Potom dle věty 5.10 máme

$$\sqrt{n}(\log(\overline{X}_n) - \log(\mu)) \xrightarrow{D} \mathrm{N}\left(0, \frac{\sigma^2}{\mu^2}\right)$$

pro $n \longrightarrow \infty$. ▲