

# Pravděpodobnost a matematická statistika (NMSA202)

Petr Velička \*

přednášející: doc. RNDr. Michal Pešta, Ph.D. †

LS 2024/25

---

\*petrvel@matfyz.cz

†pesta@karlin.mff.cuni.cz

# 1 Náhodné jevy

Začneme nejdříve základními definicemi, bez nichž vůbec nemůžeme mluvit o pravděpodobnosti.

**Definice 1.1.** *Výběrovým prostorem* rozumíme množinu  $\Omega$  všech možných výsledků nějakého experimentu. Prvky  $\omega \in \Omega$  této množiny nazýváme *elementárními jevy*. Podmnožině  $A \subset \Omega$  říkáme (*náhodný*) *jev*.

Pro ilustraci uvedeme následující motivační příklad, kde podrobně popíšeme souvislosti s právě zdefinovanými pojmy.

**Příklad 1.2.** Házíme dvakrát férovou mincí. Naším výběrovým prostorem bude množina  $\Omega = \{PP, PO, OP, OO\}$ . Událost, že první hod je panna, je tedy  $A = \{PP, PO\}$ . V tomto zápise písmeno  $P$  odpovídá tomu, že padla panna, kdežto písmeno  $O$  odpovídá orlu.

Dále uvažujme jevy  $H_1$  – při prvním hodu padne panna, a  $H_2$  – při druhém hodu padne panna. Necht jsou všechny výsledky stejně pravděpodobné (jinými slovy, mince je férová), potom pravděpodobnost, že padne alespoň jedna panna (tj. nastane jev  $H_1 \cup H_2$ ) je  $\frac{3}{4}$ .

*Důkaz.* Zřejmě z předchozího máme  $H_1 = \{PP, PO\}$  a  $H_2 = \{OP, PP\}$ . Pravděpodobnost spočteme jako podíl velikosti  $|H_1 \cup H_2| = 3$  a velikosti celého prostoru  $|\Omega| = 4$ .  $\square$

Tato jednoduchá intuice však selže v případě nekonečné (nespočetné) množiny  $\Omega$ , neboť jak již čtenář jistě ví z přednášky základů teorie míry, na nespočetné množině neexistuje “rozumný” způsob, jak měřit množiny. Musíme proto pracovat pouze s jistou třídou podmnožin  $\Omega$ , které budeme říkat  $\sigma$ -algebra.

**Definice 1.3.** Necht  $\Omega \neq \emptyset$  je množina a  $\mathcal{A} \subset 2^\Omega$  soubor jejích podmnožin. Této množině  $\mathcal{A}$  říkáme  $\sigma$ -algebra, jestliže jsou splněny následující podmínky:

- (i)  $\emptyset \in \mathcal{A}$ ,
- (ii) Pokud  $A \in \mathcal{A}$ , pak  $A^C := \Omega \setminus A \in \mathcal{A}$ ,
- (iii) Pokud  $A_1, A_2, \dots \in \mathcal{A}$ , pak  $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ .

Dvojici  $(\Omega, \mathcal{A})$  nazýváme *měřitelný prostor*.

Každé události  $A \in \mathcal{A}$  přiřadíme číslo  $\mathbb{P}(A)$ , které nazýváme *pravděpodobnost* jevu  $A$ . Jelikož chceme, aby se zachovala intuice z předchozího příkladu, musíme tuto představu náležitým způsobem formalizovat.

**Definice 1.4.** Necht  $(\Omega, \mathcal{A})$  je měřitelný prostor. Zobrazení  $P : \mathcal{A} \rightarrow [0, 1]$  nazýváme *pravděpodobnostní mírou* (*pravděpodobností*), jestliže:

- (i)  $P(\Omega) = 1$ ,
- (ii) Pro libovolné po dvou disjunktní měřitelné množiny  $A_i \in \mathcal{A}$ ,  $i \in \mathbb{N}$  platí  $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ .

Trojici  $(\Omega, \mathcal{A}, P)$  nazýváme *pravděpodobnostní prostor*.

Přímo z této definice již můžeme odvodit pár základních vlastností pravděpodobnosti, se kterými dále budeme pracovat. Ve všech následujících tvrzeních pracujeme s pravděpodobnostním prostorem  $(\Omega, \mathcal{A}, P)$ .

**Pozorování 1.5** (Základní vlastnosti pravděpodobnostní míry). *Pro výše jmenovaný pravděpodobnostní prostor platí následující tvrzení:*

1.  $P(\emptyset) = 0$ ,
2. Pro  $A, B \in \mathcal{A}$  disjunktní platí  $P(A \cup B) = P(A) + P(B)$ .
3. Pro  $A \in \mathcal{A}$  platí  $P(A^C) = 1 - P(A)$ ,
4. Pro  $A, B \in \mathcal{A}, A \subset B$  platí  $P(A) \leq P(B)$ .

*Důkaz.* 1. Uvažujme posloupnost  $A_1 = \Omega, A_2 = A_3 = \dots = \emptyset$ . Potom z vlastnosti (ii) z definice máme, že  $P(\Omega) = P(\Omega \cup \emptyset \cup \emptyset \dots) = P(\Omega) + \sum_{n=2}^{\infty} P(\emptyset)$ . Tedy  $\sum_{n=2}^{\infty} P(\emptyset) = 0$ , což může nastat pouze v případě  $P(\emptyset) = 0$  (jde o součet nekonečně mnoha nezáporných čísel).

2. Necht  $A_1 = A, A_2 = B, A_i = \emptyset$  pro  $i > 2$ . Tvrzení plyne přímo z vlastnosti (ii) z definice pravděpodobnostní míry a již dokázané vlastnosti 1.
3.  $1 = P(\Omega) = P(A \cup A^C) = P(A) + P(A^C)$ . Tato rovnost platí, neboť množina je vždy disjunktní se svým komplementem.
4.  $P(B) = P(A \cup B \setminus A) = P(A) + P(B \setminus A)$ . Jelikož funkce  $P$  je nezáporná, snadno vidíme, že  $P(B) \geq P(A)$ .

□

**Lemma 1.6** (Pravděpodobnost sjednocení). *Pro libovolné  $A, B \in \mathcal{A}$  platí  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ .*

*Důkaz.* Rozepíšeme  $A \cup B = (A \cap B^C) \cup (A \cap B) \cup (A^C \cap B)$ . Tyto tři množiny jsou zřejmě po dvou disjunktní. Dále díky aditivitě pravděpodobnosti máme  $P(A \cup B) = P(A \cap B^C) + P(A \cap B) + P(A^C \cap B) = P(A \cap B) + P(A^C \cap B) + P(A \cap B) = P(A) + P(B) - P(A \cap B)$ . □

**Věta 1.7** (Spojitost pravděpodobnosti). *Bud  $A_n \uparrow A$  nebo  $A_n \downarrow A$  pro  $A_n, A \in \mathcal{A}$ . Potom platí  $P(A_n) \rightarrow P(A)$ .*

*Důkaz.* Necht  $A_n \uparrow A$ . Potom z definice  $A_1 \subset A_2 \dots$  a platí  $A = \bigcup_{i=1}^{\infty} A_i$ . Definujme posloupnost  $B_n$ :  $B_1 = A_1, B_n = A_n \setminus A_{n-1}$ . Potom  $B_i$  jsou po dvou disjunktní a platí  $A_n = \bigcup_{i=1}^n B_i$ . Zřejmě také platí  $A \equiv \bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n$ . Pak  $P(A_n) = P(\bigcup_{i=1}^n B_i) = \sum_{i=1}^n P(B_i)$ . Z toho již můžeme odvodit  $\lim_{n \rightarrow \infty} P(A_n) = \lim_{n \rightarrow \infty} \sum_{i=1}^n P(B_i) = \sum_{i=1}^{\infty} P(B_i) = P(\bigcup_{i=1}^{\infty} B_i) = P(A)$ .

Případ klesající  $A_n$  se dokáže analogicky, stačí uvažovat  $C_n = A_n^C$ . □

*konec 1. přednášky (17.2.2025)*

Uvedeme si ještě jeden příklad ilustrující intuitivní chápání pravděpodobnosti a zavedeme první takzvané pravděpodobnostní rozdělení. Uvažujme případ, že prostor  $\Omega$  je konečný. Necht všechny výsledky jsou stejně pravděpodobné, pak platí

$$P(A) = \frac{|A|}{|\Omega|}.$$

V tomto případě mluvíme o *rovnoměrném rozdělení pravděpodobnosti*.

**Příklad 1.8** (Hod dvěma kostkami). Výběrový prostor  $\Omega = \{(i, j) : i, j \in \{1 \dots 6\}\}$  má 36 prvků. Jestliže všechny výsledky jsou stejně pravděpodobné, pak platí  $P(A) = \frac{|A|}{36}$ . Například, pravděpodobnost toho, že součet na kostkách je přesně 11, je  $2/36$ , protože pouze dva výsledky  $(5, 6)$  a  $(6, 5)$  odpovídají této události.

V praxi často chceme odlišit, zda pravděpodobnost výskytu jedné události nějakým způsobem závisí na výskytu jiné události. K tomu nám poslouží pojem nezávislosti jevů.

**Definice 1.9.** Dvě události  $A, B \in \mathcal{A}$  jsou *nezávislé*, jestliže platí  $P(A \cap B) = P(A)P(B)$ . Obdobně, množina událostí  $\{A_i : i \in I\}$  (kde indexová množina  $I$  je nejvýše spočetná) je *nezávislá*, jestliže platí

$$P\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} P(A_j)$$

pro každou konečnou podmnožinu  $J \subset I$ .

Je důležité si uvědomit, že disjunktní události s kladnou pravděpodobností nejsou nezávislé (neboť součin jejich pravděpodobností není roven 0 – pravděpodobnost výskytu jejich prázdného průniku). Obecně se pracuje se dvěma typy nezávislosti – předpokládanou (plyne z podstaty zkoumané úlohy) a odvozenou (dokázaná pomocí jiných vlastností úlohy). Následující příklad ilustruje praktické použití právě zavedeného pojmu.

**Příklad 1.10.** Házíme férovou mincí 10krát. Necht  $A$  je událost “padla aspoň jedna panna”. Pak platí  $P(A) = 1 - (1/2)^{10}$ .

*Důkaz.* Necht  $T_j$  je událost, že při  $j$ -tém hodu padne orel. Můžeme psát  $P(A) = 1 - P(A^C) = 1 - P(\text{samé orly}) = 1 - P(T_1 \cap \dots \cap T_{10})$ . Dále díky nezávislosti (v tomto případě jde o nezávislost předpokládanou) jevů  $T_j$  máme  $1 - P(T_1 \cap \dots \cap T_{10}) = 1 - P(T_1) \cdot \dots \cdot P(T_{10}) = 1 - (1/2)^{10} \approx 0.999$ .  $\square$

Dalším silným nástrojem v teorii pravděpodobnosti je podmíněná pravděpodobnost, která nám poskytuje odpověď na otázku “Pokud vím, že nastala událost  $B$ , jaká je pravděpodobnost události  $A$ ?”.

**Definice 1.11.** Mějme jevy  $A, B \in \mathcal{A}$ . Pokud  $P(B) > 0$ , pak *podmíněná pravděpodobnost*  $A$  za podmínky  $B$  je definována vztahem

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

Poznamenejme si několik základních vlastností podmíněné pravděpodobnosti, jejichž důkaz snadno plyne z příslušných definic.

**Pozorování 1.12** (Vlastnosti podmíněné pravděpodobnosti).

- (i) Pro pevné  $B \in \mathcal{A}, P(B) > 0$  je  $P(\cdot|B)$  pravděpodobnostní míra.
- (ii) Obecně platí  $P(A|B) \neq P(B|A)$ , platí totiž  $P(A|B) = P(B|A) \frac{P(A)}{P(B)}$  (pokud obě strany rovnosti dávají smysl).
- (iii) Události  $A$  a  $B$  jsou nezávislé právě tehdy, když  $P(A|B) = P(A)$  (předpokládáme nenulovost  $P(B)$ ).
- (iv)  $P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$  v případě, že  $P(A)P(B) > 0$ .

*Důkaz.* Vlastnosti (iii) a (iv) plynou přímo z definice vynásobením vhodnou konstantou.

Vlastnost (ii) se dokáže následujícím protipříkladem, uvažujme hod dvěma férovými mincemi. Nechť  $H_1$  je událost “padla aspoň jedna panna” a  $H_2$  událost “padly dvě panny”. Potom  $P(H_1|H_2) = 1$  ale  $P(H_2|H_1) = \frac{1}{3}$ . Důkaz obecného vztahu je ponechán čtenáři jako snadné (ale užitečné) cvičení.

Nakonec, vlastnost (i) je důsledkem toho, že pro libovolnou množinu  $A \in \mathcal{A}$  je  $A \cap B$  měřitelná, a navíc pro libovolný systém po dvou disjunktních množin  $A_i, i \in \mathbb{N}$  platí  $P(\bigcup_{i=1}^{\infty} A_i|B) = \frac{1}{P(B)} P((\bigcup_{i=1}^{\infty} A_i) \cap B) = \frac{1}{P(B)} P(\bigcup_{i=1}^{\infty} (A_i \cap B)) = \frac{1}{P(B)} \sum_{i=1}^{\infty} P(A_i \cap B) = \sum_{i=1}^{\infty} P(A_i|B)$ .  $\square$

Použití podmíněné pravděpodobnosti v praxi však někdy může vést k neintuitivním výsledkům, které ilustruje následující příklad.

**Příklad 1.13.** Uvažujme nemoc  $D$  a test, který má dva možné výsledky. Pravděpodobnosti výsledků tohoto testu jsou uvedeny v následující tabulce. Zde sloupce odpovídají přítomnosti/absenci nemoci a řádky výsledkům testu.

	$D$	$D^C$
+	0.009	0.099
−	0.001	0.891

Z definice spočteme následující podmíněné pravděpodobnosti:

$$P(+|D) = \frac{P(+ \cap D)}{P(D)} = \frac{0.009}{0.009 + 0.001} = 0.9.$$

$$P(-|D^C) = \frac{P(- \cap D^C)}{P(D^C)} = \frac{0.891}{0.891 + 0.099} \approx 0.9.$$

Vychází nám, že test je docela přesný, neboť nemocní lidé mají test v 90% případů pozitivní, stejně tak zdraví lidé jsou v 90% případů negativní.

Dále předpokládejme, že pacient šel na test a získal pozitivní výsledek. Spočteme, s jakou pravděpodobností je opravdu nakažený.

$$P(D|+) = \frac{P(D \cap +)}{P(+)} = \frac{0.009}{0.009 + 0.099} \approx 0.08.$$

Vyšlo nám, že na první pohled zdánlivě precizní test ve skutečnosti má méně než 10% úspěšnost. Jedním z důvodů této diskrepance může být například velký nepoměr zdravých lidí vůči nakaženým (pouze jedno procento) ve zdrojových datech, což je jev který se obecně vyskytuje u většiny nemocí. V praxi se proto často pracuje s domněnkami – například testujeme jen pacienty, kteří vykazují nějaké symptomy apod.

Na závěr uvedeme dvě velmi užitečné věty, které se často používají v nej-různějších úlohách a týkají se podmíněné pravděpodobnosti. Zformulujeme je pro spočetné rozklady, ale obdobná tvrzení platí i pro konečné rozklady s velmi podobným důkazem.

**Věta 1.14** (Zákon úplné pravděpodobnosti). *Nechť  $A_1, A_2, \dots$  je spočetný disjunkttní rozklad  $\Omega$  takový, že  $P(A_i) > 0$  pro každé  $i \in \mathbb{N}$ . Potom pro libovolnou událost  $B \in \mathcal{A}$  platí:*

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

*Důkaz.* Definujme posloupnost množin  $C_i = B \cap A_i$  pro  $i \in \mathbb{N}$ . Zjevně  $\{C_i, i \in \mathbb{N}\}$  je disjunkttní pokrytí  $B$ . Potom  $P(B) = \sum_{i=1}^{\infty} P(C_i) = \sum_{i=1}^{\infty} P(B \cap A_i) = \sum_{i=1}^{\infty} P(B|A_i)P(A_i)$ .  $\square$

**Věta 1.15** (Bayes). *Nechť  $A_1, A_2, \dots$  je spočetný disjunkttní rozklad  $\Omega$  takový, že  $P(A_i) > 0$  pro každé  $i \in \mathbb{N}$ . Mějme událost  $B \in \mathcal{A}$  s nenulovou pravděpodobností. Potom platí:*

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

*Důkaz.* Příímým výpočtem dostáváme

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(B|A_i)P(A_i)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{\infty} P(B|A_j)P(A_j)},$$

kde poslední rovnost získáme aplikací *Věty 1.14*.  $\square$

Použití Bayesovy věty si ukážeme na následujícím příkladu.

**Příklad 1.16.** Uvažujme e-mailovou schránku. Máme tři kategorie e-mailů:  $A_1$  – spam,  $A_2$  – nízká priorita,  $A_3$  – vysoká priorita. Na základě předchozích zkušeností víme, že  $P(A_1) = 0.7$ ,  $P(A_2) = 0.2$ ,  $P(A_3) = 0.1$ . Nechť  $B$  je událost, že daný e-mail obsahuje slovo “zdarma”. Platí  $P(B|A_1) = 0.9$ ,  $P(B|A_2) =$

$0.01$ ,  $P(B|A_3) = 0.01$ <sup>1</sup>. Jaká je pravděpodobnost, že příchozí e-mail obsahující slovo “zdarma” je spam?

Přímým výpočtem z Bayesovy věty získáme

$$P(A_1|B) = \frac{0.9 \cdot 0.7}{0.9 \cdot 0.7 + 0.01 \cdot 0.2 + 0.01 \cdot 0.1} = 0.995.$$

Tedy pravděpodobnost, že tento e-mail je spam je přes 99%!

**Věta 1.17** (O postupném podmiňování). *Nechť  $\{A_i\}_{i=1}^n$  jsou náhodné jevy takové, že  $P(\bigcap_{i=1}^n A_i) > 0$ . Pak platí*

$$P\left(\bigcap_{i=1}^n A_i\right) = P\left(A_n \mid \bigcap_{i=1}^{n-1} A_i\right) \cdot P(A_2|A_1) \cdot P(A_1).$$

*Důkaz.* Dokazujeme indukcí podle počtu náhodných jevů. Z definice podmíněné pravděpodobnosti víme, že  $P(A_2 \cap A_1) = P(A_2|A_1)P(A_1)$ . Dále

$$P\left(\bigcap_{i=1}^n A_i\right) = P\left(A_n \cap \left(\bigcap_{i=1}^{n-1} A_i\right)\right) = P\left(A_n \mid \bigcap_{i=1}^{n-1} A_i\right) P\left(\bigcap_{i=1}^{n-1} A_i\right),$$

čímž je důkaz ukončen. □

---

<sup>1</sup>Tyto hodnoty se nutně nemusí sečíst na 1

## 2 Náhodné veličiny

V této kapitole se budeme věnovat náhodným veličinám, což bude formalizovat (a zobecňovat) jakýsi intuitivní chápání toho, že nějaká proměnná nabývá různých hodnot s určitými pravděpodobnostmi. Začneme ústřední definicí celé statistiky – náhodnou veličinou.

**Definice 2.1.** Necht  $(\Omega, \mathcal{A})$  je měřitelný prostor. *Náhodná veličina* je měřitelné zobrazení, které přiřazuje každému výsledku  $\omega$  reálné číslo  $X(\omega)$ . Jinými slovy,  $\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{A} \forall x \in \mathbb{R}$ .

*konec 2. přednášky (18.2.2025)*

**Úmluva 2.2.** Zavedeme značení  $[X \in B] = \{\omega : X(\omega) \in B\}$ ,  $[X \leq a] = \{\omega, X(\omega) \leq a\}$ . Platí tedy  $[X \in B], [X \leq a] \in \mathcal{A}$  pro všechna  $B \in \mathcal{B}$ ,  $a \in \mathbb{R}$ . Jde o náhodné jevy a jsou tedy dobře definované jejich pravděpodobnosti  $P[X \in B], P[X \leq a]$ .

**Příklad 2.3.** Házíme mincí desetkrát. Necht  $X(\omega)$  je počet orlů v posloupnosti  $\omega$ . Jestliže  $\omega = O O P O O P O O P P$  (kde  $O$  je orel a  $P$  je panna), platí  $X(\omega) = 6$ .

V předchozí kapitole jsme mluvili o pravděpodobnostním rozdělení, je na čase tento pojem formálně zdefinovat.

**Definice 2.4.** *Rozdělením náhodné veličiny*  $X : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$  nazýváme indukovanou pravděpodobnostní míru  $P_X$  na  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  definovanou jako

$$P_X(B) := P[\{\omega \in \Omega : X(\omega) \in B\}], B \in \mathcal{B}(\mathbb{R}).$$

Máme tedy jakýsi obraz míry  $P$  v zobrazení  $P_X$  čímž se  $(\Omega, \mathcal{A}, P)$  zobrazí na pravděpodobnostní prostor  $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$ . V opačném směru můžeme použít takzvané kanonické vnoření do prostoru  $(\mathbb{R}, \mathcal{B}, P_X)$ , kde naší zvolenou měřitelnou funkcí bude identita, tedy není potřeba se bát, že by příslušný prostor nemusel existovat. Následující věta říká, že nezáleží ve kterém z těchto dvou prostorů integrujeme libovolnou funkci.

**Věta 2.5** (O přenosu integrace). *Bud'  $g$  měřitelná funkce na měřitelném prostoru  $(\mathbb{M}, \mathcal{M})$  a  $X : (\Omega, \mathcal{A}, P) \rightarrow (\mathbb{M}, \mathcal{M})$ . Necht  $P_X$  je míra na  $\mathcal{M}$  indukovaná zobrazením  $X$ , tedy  $P_X(M) = P[X^{-1}(M)]$  pro  $M \in \mathcal{M}$ . Potom, je-li aspoň jedna strana definována, platí*

$$\int_{\Omega} g[X(\omega)] dP(\omega) = \int_{\mathbb{M}} g(x) dP_X(x).$$

*Důkaz.* Důkaz této věty je poměrně technický, hlavní ideou je “klasický” postup z teorie míry postupným důkazem nejdříve pro charakteristickou funkci, poté pro jednoduchou měřitelnou (nabývající jen konečně mnoha hodnot), pak pro nezápornou měřitelnou a na závěr pro obecnou měřitelnou funkci.



Nechť  $g = \chi_B, B \in \mathcal{M}$ . Tedy  $g(X(\omega)) = 1$  pro  $X(\omega) \in B$  (a všude jinde nulová), tedy pro  $\omega \in X^{-1}(B)$ . Potom máme

$$\int_{\Omega} g(X(\omega)) dP(\omega) = \int_{X^{-1}(B)} dP(\omega) = P[X^{-1}(B)].$$

Pro pravou stranu máme

$$\int_{\mathbb{M}} g(x) dP_X(x) = \int_B dP_X(x) = P_X(B) = P[X^{-1}(B)].$$

Dále necht  $g$  je jednoduchá měřitelná, tedy  $g(\cdot) = \sum_{k=1}^n c_k \chi_{B_k}(\cdot)$  pro  $n \in \mathbb{N}$ ,  $c_k \in \mathbb{R}$  a  $B_k \in \mathcal{M}$  pro všechna  $k$ . Z linearity integrálu plyne (vytkneme sumu)  $\int_{\Omega} g(X(\omega)) dP(\omega) = \int_{X^{-1}(B)} dP(\omega) = P[X^{-1}(B)]$ .

Je-li  $g$  nezáporná měřitelná, potom existuje posloupnost  $g_n$  jednoduchých měřitelných funkcí takových, že  $g_n \nearrow g$ . Potom dle Léviho věty o monotonní konvergenci máme

$$\begin{aligned} \int_{\Omega} g[X(\omega)] dP(\omega) &= \lim_{n \rightarrow \infty} \int_{\Omega} g_n[X(\omega)] dP(\omega) \\ &= \lim_{n \rightarrow \infty} \int_{\mathbb{M}} g_n(x) dP_X(x) = \int_{\mathbb{M}} g(x) dP_X(x), \end{aligned}$$

kde třetí rovnost plyne z již dokázané části pro jednoduché měřitelné funkce.

Nakonec, pro  $g$  měřitelnou existuje rozklad  $g = g^+ - g^-$  takový, že  $g^+, g^-$  jsou nezáporné měřitelné, tedy požadované tvrzení plyne z části pro nezáporné měřitelné funkce.  $\square$

Na závěr poznamenejme, že se nám budou obzvlášť hodit volby  $(\mathbb{M}, \mathcal{M}) = (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$  pro  $n \geq 1$ .

Připomeňme si, že jsou-li  $\mu, \nu$  dvě  $\sigma$ -konečné míry na  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  a je-li  $\nu \ll \mu$  (tedy  $\mu(B) = 0$  implikuje  $\nu(B) = 0$ ), potom z Radonovy-Nikodymovy věty plyne existence nezáporné měřitelné funkce  $f$  takové, že  $\nu(B) = \int_{\mathbb{R}} f d\mu$  pro všechna  $B \in \mathcal{B}$ . Této funkci  $f$  říkáme Radonova-Nikodymova derivace a píšeme  $f = \frac{d\nu}{d\mu}$ . Taková funkce  $f$  je navíc určena jednoznačně až na množinu  $\mu$ -míry 0.

Využijeme těchto poznatků tak, že zvolíme vhodnou referenční míru na  $\mathbb{R}$  a rozdělení  $P_X$  pak bude popsáno právě zavedenou Radonovou-Nikodymovou derivací. Vhodné referenční míry jsou např.

- Lebesgueova míra  $\lambda$ ,
- Čítací míra na spočetné podmnožině  $\mathbb{R}$ , platí  $\mu_S(B) = |B \cap S|$  kde  $S$  je nejvýše spočetná podmnožina  $\mathbb{R}$ .

**Definice 2.6.** Buď  $X$  náhodná veličina a  $P_X$  její rozdělení. Necht  $P_X$  je absolutně spojitě vůči  $\mu$ , kde  $\mu$  je  $\sigma$ -konečná míra na  $\mathbb{R}$ . Pak funkci  $f_X$  splňující  $P_X(B) = \int_B f_X d\mu$  pro všechny  $B \in \mathcal{B}$  nazveme *hustotou* rozdělení náhodné veličiny  $X$  vůči míře  $\mu$ .

Je třeba si dát pozor na to, aby zvolená referenční míra opravdu byla absolutně spojitá, například při hodu kostkou má výsledek 1 nenulovou pravděpodobnost, ale  $\lambda(\{1\}) = 0$ .

**Věta 2.7.** *Buď  $X$  náhodná veličina a  $P_X$  její rozdělení. Je-li  $f_X$  hustota (rozdělení) vůči  $\sigma$ -konečné míře  $\mu$ , pak*

$$P[X \in B] = \int_B f_X d\mu.$$

*Důkaz.* Jde o přímý důsledek Radonovy-Nikodymovy věty a vztahu mezi  $P_X$  a  $P$ .  $\square$

Další funkcí, která plně charakterizuje rozdělení náhodné veličiny je tzv. distribuční funkce.

**Definice 2.8.** Buď  $X$  náhodná veličina na  $(\Omega, \mathcal{A}, P)$  a  $P_X$  její rozdělení. *Distribuční funkce  $F_X$  náhodné veličiny  $X$  je definována vztahem*

$$F_X(a) := P((-\infty, a]) = P[X \leq a].$$

Uvedeme si několik užitečných vlastností distribučních funkcí:

**Důsledek 2.9** (Základní vlastnosti distribučních funkcí).

- (i) *Distribuční funkce jednoznačně určuje rozdělení (jinými slovy,  $F_X = F_Y$  implikuje  $P_X = P_Y$ ).*
- (ii) *Různé náhodné veličiny mohou mít stejné distribuční funkce, tedy stejné rozdělení.*

*konec 3. přednášky (24.2.2025)*

**Příklad 2.10.** Hodíme dvěma kostkami, označme  $Y$  počet sudých čísel na těchto dvou kostkách. Potom  $Y \in \{0, 1, 2\}$ . Z definice  $F_Y(a) = P[Y \leq a]$ , tedy

$$F_Y(a) = \begin{cases} 0, & a < 0, \\ \frac{1}{4}, & 0 \leq a < 1, \\ \frac{3}{4}, & 1 \leq a < 2, \\ 1, & a \geq 2. \end{cases}$$

Dále, z toho, že  $P_Y(0) = \frac{1}{4} > 0$ , plyne, že míra  $P_Y$  není absolutně spojitá vůči Lebesgueově míře  $\lambda$ , tedy musíme uvažovat čítací míru  $\mu_{\mathbb{Z}}$  na množině celých čísel. Potom hustota  $f_Y$  má následující tvar:

$$f_Y(a) = \begin{cases} \frac{1}{4}, & a = 0, \\ \frac{1}{2}, & a = 1, \\ \frac{1}{4}, & a = 2, \\ 0, & \text{jinak.} \end{cases}$$

Vidíme, že hustota odpovídá skokům distribuční funkce v daném bodě. V následující větě uvedeme charakterizaci distribučních funkcí.

**Věta 2.11** (Charakterizace distribučních funkcí). *Buď  $X$  náhodná veličina a  $F_X$  její distribuční funkce. Pak*

- (i)  $F_X$  je neklesající;
- (ii)  $\lim_{a \rightarrow -\infty} F_X(a) = 0$ ,  $\lim_{a \rightarrow +\infty} F_X(a) = 1$ ;
- (iii)  $F_X$  je zprava spojitá.

*Navíc, každá funkce  $F$  splňující body (i)-(iii) z této věty je distribuční funkcí nějaké náhodné veličiny.*

*Důkaz.* Dokážeme pouze implikaci o vlastnostech distribuční funkce, opačná implikace (existuje rozdělení) vyžaduje pokročilý matematický aparát z analýzy a teorie míry, který prozatím postrádáme.

- (i)  $F_X(a) = P[X \leq a]$ . Bez újmy na obecnosti nechť  $b > a$ . Potom  $F_X(b) = P[X \leq b] = P([X \leq a] \cup [a < X \leq b]) = P[X \leq a] + P[a < X \leq b]$  z aditivní míry, druhý sčítanec je nezáporný, tedy dostáváme požadované tvrzení.
- (ii) Platí  $\lim_{a \rightarrow -\infty} F_X(a) = \lim_{n \rightarrow \infty} F_X(-n) = \lim_{n \rightarrow \infty} P[X \in (-\infty, -n]] = \lim_{n \rightarrow \infty} P[X \in A_n] = 0$ . Poslední rovnost platí ze spojitosti míry (Věta 1.7), neboť platí  $A_n \searrow \emptyset$ . Obdobně se ukáže tvrzení pro  $a \rightarrow +\infty$  (cvičení).
- (iii) Stačí uvažovat posloupnost  $a_n = a + \frac{1}{n}$  pro  $n \in \mathbb{N}$ . Požadované tvrzení opět plyne z věty o spojitosti míry.

□

Pro každou funkci  $F$  splňující vlastnosti z předchozí věty existuje míra  $\mu_F$  na  $(\mathbb{R}, \mathcal{B})$  určená vztahem  $\mu_F((-\infty, a]) = F(a)$  pro všechna  $a$ . Tato míra je konečná a platí  $\mu_F((a, b]) = F(b) - F(a)$ .

**Definice 2.12** (Rozklad pravděpodobnostního rozdělení). Každou pravděpodobnostní míru  $P_X$  můžeme rozdělit na tři složky  $P_X = P_{X_{as}} + P_{X_{ds}} + P_{X_{sg}}$ , kde  $P_{X_{as}}$  je absolutně spojitá vůči Lebesgueově míře  $\lambda$ ,  $P_{X_{ds}}$  (diskrétní spojitá) je absolutně spojitá vůči číselní míře  $\mu$  na nějaké spočetné podmnožině  $\mathbb{R}$  a nakonec  $P_{X_{sg}}$  (singulární) není absolutně spojitá vůči  $\lambda$  ani ji nelze napsat jako spočetnou kombinaci Diracových měr  $\delta_x$ .

Příkladem singulární distribuční funkce je například integrál takzvaného Cantorova diskontinua. Obecně taková rozdělení nemají “hezké” vlastnosti, proto s nimi již nebudeme pracovat.

**Definice 2.13.** Náhodnou veličinu  $X$  nazveme *diskrétní*, jestliže existují  $\emptyset \neq I \subset \mathbb{N}$ ,  $\{x_i\}_{i \in I}$  a  $\{p_i \in (0, 1]\}_{i \in I}$  takové že  $P[X \in B] = \sum_{i, x_i \in B} p_i$  pro všechny borelovské  $B$ .

Platí  $P[X = x_i] = p_i$  a  $\sum_{i \in I} p_i = 1$ . Rozdělením takové veličiny je funkce  $P_X = \sum_{i \in I} p_i \delta_{x_i}$ , kde  $\delta_u$  je Diracova míra v bodě  $u$ . Toto rozdělení je absolutně spojitě vůči číselní míře na  $S = \{x_i\}_{i \in I} \subset \mathbb{R}$ . Potom funkce  $f_X(u) := \begin{cases} p_i, u = x_i, \\ 0, \text{jinak} \end{cases}$  je hustotou (občas také pravděpodobnostní funkcí) zkoumaného rozdělení.

**Definice 2.14.** Náhodná veličina  $X$  se nazývá (*absolutně*) *spojitá*, pokud její rozdělení  $P_X$  je absolutně spojitě vůči Lebesgueově míře  $\lambda$ .

Pro spojitou náhodnou veličinu  $X$  vždy existuje hustota  $f_X$  (nezáporná a jednoznačná až na množinu  $\lambda$ -míry 0) splňující  $P[X \in B] = \int_B f_X(t) dt$  a speciálně  $F_X(a) = \int_{-\infty}^a f_X(t) dt$  pro všechna  $a \in \mathbb{R}$ . Taková  $F_X$  má derivaci ve skoro všech bodech a platí  $F'_X(a) = f_X(a)$  pro s.v.  $a$ . Analogicky pro diskrétní náhodnou veličinu  $Y$  je hustota funkcí, která nabývá v bodě  $a$  hodnoty distribuční funkce v daném bodě.

Ne každá veličina, se kterou se běžně setkáme je ryze spojitá nebo ryze diskrétní. Příkladem veličiny, která má obě složky nenulové, je například úhrn denních srážek, s nenulovou pravděpodobností nenaprší vůbec, ale když už začne pršet, úhrn srážek je spojitá náhodná veličina.

**Lemma 2.15.** *Nechť  $F_X$  je distribuční funkce náhodné veličiny  $X$ . Pak pro  $a < b$  platí*

- (i)  $P[a < X \leq b] = P[X \in (a, b)] = F_X(b) - F_X(a)$ ,
- (ii)  $P[X > a] = 1 - F_X(a)$ ,
- (iii)  $P[X = a] = F_X(a) - F_X(a^-)$ , kde  $F_X(a^-)$  je limita zleva  $\lim_{h \rightarrow 0^+} F_X(a - h)$  a odtud  $P[a \leq X \leq b] = F_X(b) - F_X(a^-)$ .
- (iv) pro spojitou náhodnou veličinu platí  $P[a \leq X \leq b] = P[a \leq X < b] = P[a < X \leq b] = P[a < X < b] = F_X(b) - F_X(a)$ .

*Důkaz.* Důkaz je jednoduchý, plyne z příslušných definic. Uvedeme např. důkaz pro bod (iii).

$$P[X = a] = \lim_{h \rightarrow 0^+} P[a - h < X \leq a] = F_X(a) - \lim_{h \rightarrow 0^+} F_X(a - h). \quad \square$$

*konec 4. přednášky (25.2.2025)*

Dalším užitečným pojmem je funkce inverzní k distribuční funkci, které běžně říkáme kvantil.

**Definice 2.16.** Nechť  $X$  je náhodná veličina s distribuční funkcí  $F$ . *Inverzní distribuční funkce* neboli *kvantilová funkce* je definována jako

$$F^{-1}(q) = \inf \{x : F(x) > q\}$$

pro  $q \in (0, 1)$ . Hodnoty  $F^{-1}$  ve speciálních bodech mají své vlastní názvy:

- $F^{-1}(\frac{1}{4})$  je *první kvartil*,

- $F^{-1}(\frac{1}{2})$  je *medián*,
- $F^{-1}(\frac{3}{4})$  je *třetí kvartil*.

Je-li  $F$  ryze rostoucí a spojitá, je  $F^{-1}(q)$  to jediné  $x \in \mathbb{R}$  takové, že  $F(x) = q$ , jinými slovy,  $F$  je bijekce z  $\mathbb{R}$  do  $(0, 1)$ . Takto definovaná kvantilová funkce je neklesající a zprava spojitá. Dále z  $F^{-1}$  můžeme jednoznačně určit  $F$ , tedy také charakterizuje rozdělení  $P_X$ . Nakonec, o dvou náhodných veličinách  $X$  a  $Y$  říkáme, že jsou stejně rozdělené, zapisujeme  $X \stackrel{d}{=} Y$ , právě tehdy, když  $F_X(x) = F_Y(x)$  pro všechna  $x$ . To však neznamena, že  $X = Y$ .

Ukážeme si několik užitečných příkladů rozdělení (diskrétních a později i spojitých). Tato rozdělení se používají v praxi při modelování jednoduchých systémů, ale u komplikovanějších modelů se s těmito rozděleními bohužel nevystačíme.

## 2.1 Diskrétní náhodné veličiny

**Definice 2.17** (Bodové rozdělení). Náhodná veličina  $X$  má *bodové rozdělení* v bodě  $a$  právě tehdy, když  $P[X = x] = \chi_{\{x=a\}}, x \in \mathbb{R}$ . Zapisujeme  $X \sim \delta_a$ . Potom platí  $F_X(x) = \chi_{\{x \geq a\}}$ .

**Definice 2.18** (Diskrétní rovnoměrné rozdělení). Náhodná veličina  $X$  má *diskrétní rovnoměrné rozdělení* na  $\{1, \dots, k\}$  právě tehdy, když

$$f_X(x) = \begin{cases} 1/k, & x = 1, \dots, k; \\ 0, & \text{jinak.} \end{cases}$$

**Definice 2.19** (Bernoulliho rozdělení). Náhodná veličina  $X$  má *Bernoulliho rozdělení* s parametrem  $p \in (0, 1)$  právě tehdy, když  $f_X(x) = p^x(1-p)^{1-x}$  pro  $x \in 0, 1$ . Zapisujeme  $X \sim \text{Alt}(p)$  nebo  $X \sim \text{Be}(p)$ . Tímto rozdělením modelujeme jevy, u kterých jsou pouze dva možné výsledky (úspěch/neúspěch, hod mincí).

**Definice 2.20** (Binomické rozdělení). Náhodná veličina  $X$  má *binomické rozdělení* s parametry  $n \in \mathbb{N}$  a  $p \in (0, 1)$  právě tehdy, když

$$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x} \chi_{\{x \in \{0, \dots, n\}\}}.$$

Zapisujeme  $X \sim \text{Bi}(n, p)$ . Používáme toto v případě sčítaně nezávislých<sup>2</sup> veličin s Bernoulliho rozdělením (počet úspěchů mezi  $n$  pokusy).

**Definice 2.21** (Geometrické rozdělení). Náhodná veličina  $X$  má *geometrické rozdělení* s parametrem  $p \in (0, 1)$  (zapisujeme  $X \sim \text{Geo}(p)$ ) právě tehdy, když

$$f_X(x) = p(1-p)^x$$

pro  $x \in \mathbb{N}_0$ . Taková náhodná veličina vyjadřuje počet neúspěchů před prvním úspěchem v posloupnosti nezávislých pokusů.

<sup>2</sup>Přesná definice nezávislých veličin bude uvedena později.

**Definice 2.22** (Negativně binomické rozdělení). Náhodná veličina  $X$  má *negativně binomické rozdělení* s parametry  $n \in \mathbb{N}$  a  $p \in (0, 1)$  (píšeme  $X \sim NB(n, p)$ ), jestliže platí

$$f_X(x) = \binom{n+x-1}{n-1} p^n (1-p)^x$$

pro  $x \in \mathbb{N}_0$ . Rozdělení vyjadřuje počet neúspěchů před  $n$ -tým úspěchem v posloupnosti nezávislých pokusů. Specifický případ  $NB(1, p) = Geo(p)$ .

**Definice 2.23** (Poissonovo rozdělení). Náhodná veličina  $X$  má *Poissonovo rozdělení* s parametrem  $\lambda > 0$  (zapisujeme  $X \sim Po(\lambda)$ ) právě tehdy, když

$$f_X(x) = e^{-\lambda} \frac{\lambda^x}{x!}$$

pro  $x \in \mathbb{N}_0$ . Jestliže  $X \sim Po(\lambda_X)$  a  $Y \sim Po(\lambda_Y)$  jsou nezávislé, potom  $X + Y \sim Po(\lambda_X + \lambda_Y)$ . Jestliže  $n \rightarrow \infty$  a  $np \rightarrow \lambda < \infty$ , potom  $Bi(n, p)$  konverguje k  $Po(\lambda)$ .

## 2.2 Absolutně spojité náhodné veličiny

**Definice 2.24** (Spojitě rovnoměrné rozdělení). Náhodná veličina  $X$  má *rovnoměrné rozdělení* na intervalu  $[a, b]$  právě tehdy, když  $f_X(x) = (b-a)^{-1} \chi_{\{x \in [a, b]\}}$ . Zapisujeme  $X \sim U(a, b)$  (uniform) nebo  $X \sim R(a, b)$  (rovnoměrné).

**Definice 2.25** (Normální rozdělení). Náhodná veličina  $X$  má *normální (Gaussovo) rozdělení* s parametry  $\mu \in \mathbb{R}$  a  $\sigma^2 > 0$  (zapisujeme  $X \sim N(\mu, \sigma^2)$ ) právě tehdy, když

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\}$$

pro  $x \in \mathbb{R}$ .

Toto rozdělení je enormně důležité, uvedeme si proto několik jeho vlastností. Nejprve, máme-li  $X \sim N(\mu, \sigma^2)$ , potom  $Z := (X - \mu)/\sigma \sim N(0, 1)$ . Tomuto rozdělení říkáme *standardní normální rozdělení*. Dále, máme-li dvě nezávislé normálně rozdělené veličiny  $X \sim N(\mu_X, \sigma_X^2)$ ,  $Y \sim N(\mu_Y, \sigma_Y^2)$ , potom  $X + Y \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$ .

Distribuční funkce  $N(0, 1)$  nejde vyjádřit analyticky, máme jen  $\Phi(x) := \int_{-\infty}^x \phi(t) dt$ , kde  $\phi(x) := \frac{1}{\sqrt{2\pi}} \exp -x^2/2$  je hustota standardního normálního rozdělení. Její hodnoty proto vyhledáváme v tabulkách, případně počítáme numericky.

**Příklad 2.26.** Předpokládejme, že  $X \sim N(3, 5)$ . Spočteme  $P[X \geq 1]$ . Dále spočtete  $q = F_X^{-1}(0.2)$ .

*Důkaz.* Počítáme přímo

$$P[X \geq 1] = 1 - P[X \leq 1] = 1 - P[Z \leq \frac{1-3}{\sqrt{5}}] \approx 1 - \Phi(-0.8944) = 0.81.$$

Dále z tabulek víme, že  $\Phi(-0.8416) = 0.2$ , potom

$$0.2 = P[X \leq q] = P[Z \leq \frac{q - \mu}{\sigma}] = \Phi[\frac{q - \mu}{\sigma}],$$

proto  $-0.8416 = \frac{q - \mu}{\sigma} = \frac{q - 3}{\sqrt{5}}$  a tedy  $q = 3 - 0.8416\sqrt{5} \approx 1.1181$ .  $\square$

**Definice 2.27** (Exponenciální rozdělení). Náhodná veličina  $X$  má *exponenciální rozdělení* s parametrem  $\lambda > 0$  (zapisujeme  $X \sim \text{Exp}(\lambda)$ ) právě tehdy, když

$$f_X(x) = \lambda e^{-\lambda x} \chi_{\{x > 0\}}.$$

**Definice 2.28** (Gamma rozdělení). Náhodná veličina  $X$  má *Gamma rozdělení* s parametry  $a, p > 0$  právě tehdy, když

$$f_X(x) = \frac{a^p}{\Gamma(p)} x^{p-1} e^{-ax} \chi_{\{x > 0\}},$$

kde  $\Gamma(p) = \int_0^\infty t^{p-1} e^{-t} dt$  je gamma funkce (spojité rozšíření faktoriálu). Zapisujeme  $X \sim \text{Gamma}(a, p)$  nebo  $X \sim \Gamma(a, p)$ . Exponenciální rozdělení  $\text{Exp}(a)$  je speciálním případem Gamma rozdělení s parametrem  $p = 1$ .

Opět máme součtový vzorec pro nezávislé veličiny  $X \sim \Gamma(a, p_X), Y \sim \Gamma(a, p_Y)$ , platí totiž  $X + Y \sim \Gamma(a, p_X + p_Y)$ .

**Definice 2.29** (Beta rozdělení). Náhodná veličina  $X$  má *Beta rozdělení* s parametry  $\alpha, \beta > 0$  právě tehdy, když

$$f_X(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1} \chi_{\{x \in (0,1)\}}.$$

Zapisujeme  $X \sim \text{Beta}(\alpha, \beta)$  nebo  $X \sim B(\alpha, \beta)$ . Všimněme si, že na rozdíl od předchozích rozdělení jde o rozdělení na kompaktu.

*konec 5. přednášky (3.3.2025)*

**Definice 2.30** ( $\chi^2$ -rozdělení). Náhodná veličina  $X$  má  $\chi^2$ -rozdělení s  $p$  stupni volnosti (zapisujeme  $X \sim \chi_p^2$ ) právě tehdy, když

$$f_X(x) = \frac{1}{\Gamma(p/2)2^{p/2}} x^{p/2-1} e^{-x/2} \chi_{\{x > 0\}}.$$

Máme-li soubor nezávislých náhodných veličin  $X_1, \dots, X_p \sim N(0, 1)$ , potom součet jejich druhých mocnin odpovídá  $\chi^2$ -rozdělení,  $\sum_{i=1}^p X_i^2 \sim \chi_p^2$ .

**Definice 2.31** (Studentovo  $t$ -rozdělení). Náhodná veličina  $X$  má Studentovo  $t$ -rozdělení s  $\nu$  stupni volnosti (zapisujeme  $X \sim t_\nu$ ) právě tehdy, když

$$f_X(x) = \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\pi\nu}} \frac{1}{(1+x^2/\nu)^{(\nu+1)/2}}.$$

**Definice 2.32** (Cauchyho rozdělení). Cauchyovo rozdělení je speciální případ  $t$ -rozdělení, když  $\nu = 1$ . Potom platí

$$f_X(x) = \frac{1}{\pi(1+x^2)}.$$

K zajímavým vlastnostem Cauchyova rozdělení patří například to, že nemá střední hodnotu (bude upřesněno později).

Přejdeme dále k vícerozměrným náhodným veličinám, jedním z jejich využití je například možnost formální definice pojmu nezávislosti několika náhodných veličin.

**Definice 2.33.** Necht  $(\Omega, \mathcal{A})$  je měřitelný prostor. *Náhodný vektor* je měřitelné zobrazení, které každému výsledku  $\omega$  přiřadí reálný  $d$ -rozměrný vektor  $\vec{X}(\omega)$ . To znamená, že

$$\vec{X} : \Omega \rightarrow \mathbb{R}^d \wedge \{\omega \in \Omega : \vec{X}(\omega) \leq \vec{x}\} \in \mathcal{A} \forall \vec{x} \in \mathbb{R}^d.$$

**Definice 2.34.** *Rozdělením náhodného vektoru*  $\vec{X} : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$  nazveme indukovanou pravděpodobnostní míru  $P_{\vec{X}}$  na  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$  definovanou jako

$$P_{\vec{X}}(B) := P[\{\omega \in \Omega : \vec{X}(\omega) \in B\}]$$

pro všechny  $B \in \mathcal{B}(\mathbb{R}^d)$ .

Již na první pohled je zřejmá analogie s jednorozměrnými náhodnými veličinami v tom, že  $P_{\vec{X}}$  je obraz míry  $P$  v zobrazení  $\vec{X}$ , kde se původní pravděpodobnostní prostor zobrazí na  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), P_{\vec{X}})$ . Platí, že pokud máme náhodný vektor  $\vec{X} = [X_1, \dots, X_d]^T$ , potom  $X_i$  je náhodná veličina pro všechna  $i \in \{1, \dots, d\}$  (důsledek definice, avšak platí i opačná implikace).

**Definice 2.35.** Necht  $(\Omega, \mathcal{A}, P)$  je pravděpodobnostní prostor. *Sdružená distribuční funkce* náhodného vektoru  $\vec{X}$  je funkce  $F_{\vec{X}} : \mathbb{R}^d \rightarrow [0, 1]$  definovaná jako

$$F_{\vec{X}}(\vec{x}) = P(\vec{X} \leq \vec{x}) = P\left(\bigcup_{l=1}^d \{X_l \leq x_l\}\right)$$

pro všechna  $\vec{x} = [x_1, \dots, x_d]^T \in \mathbb{R}^d$ .

Analogicky s náhodnými veličinami si zformulujeme tvrzení o základních vlastnostech sdružených distribučních funkcí.

**Věta 2.36** (Vlastnosti sdružené distribuční funkce). *Pokud je  $F$  sdružená distribuční funkce  $d$ -rozměrného náhodného vektoru  $\vec{X}$ , pak platí*

- (i)  $F$  je po složkách neklesající a zprava spojitá;
- (ii)  $\lim_{x_l \rightarrow -\infty} F(\vec{x}) = 0$  pro každé  $l = 1 \dots d$ ;



(iii)  $\lim_{x_l \rightarrow +\infty \forall l} F(\vec{x}) = 1$ .

*Důkaz.* Nejdříve dokážeme vlastnost (i). Fixujme  $x_1, \dots, x_{l-1}, x_{l+1}, \dots, x_d \in \mathbb{R}$  a definujeme funkci  $G(x) := F_{\vec{X}}(x_1, \dots, x_{l-1}, x, x_{l+1}, \dots, x_d)$ . Z monotonie pravděpodobnosti je  $G$  neklesající a nezáporná. Jelikož je  $G$  neklesající, nutně existuje limita  $\lim_{y \rightarrow x^+} G(y) \geq G(x)$ . Dokážeme, že dochází k rovnosti (čímž dokážeme spojitost zprava).

Z Heineovy věty plyne, že  $\lim_{y \rightarrow x^+} G(y) = \lim G(x + \frac{1}{n})$ . Označme  $B_n := (-\infty, x_1] \times \dots \times (-\infty, x_{l-1}] \times (-\infty, x + \frac{1}{n}) \times (-\infty, x_{l+1}] \times \dots \times (-\infty, x_d]$ . Potom máme, že  $B_n \searrow B := (-\infty, x_1] \times \dots \times (-\infty, x_{l-1}] \times (-\infty, x] \times (-\infty, x_{l+1}] \times \dots \times (-\infty, x_d]$ .

Dále s využitím věty o spojitosti míry (Věta 1.7) můžeme psát

$$\lim_{y \rightarrow x^+} G(y) = \lim_{n \rightarrow \infty} G(x + \frac{1}{n}) = \lim_{n \rightarrow \infty} P_{\vec{X}}(B_n) = P_{\vec{X}}\left(\bigcap_{n=1}^{\infty} B_n\right) = P_{\vec{X}}(B) = G(x),$$

čímž je ukončen důkaz vlastnosti (i).

K důkazu vlastnosti (ii) opět mějme pevná  $x_1, \dots, x_{l-1}, x_{l+1}, \dots, x_d \in \mathbb{R}$ . Opět uvažujme funkci  $G$  z předchozí části důkazu, která je neklesající a nezáporná. Proto musí existovat její limita  $\lim_{\vec{x} \rightarrow -\infty} G(\vec{x}) = \lim_{n \rightarrow \infty} G(-n)$  (opět plyne z Heineovy věty). Definujme  $C_n := (-\infty, x_1] \times \dots \times (-\infty, x_{l-1}] \times (-\infty, -n] \times (-\infty, x_{l+1}] \times \dots \times (-\infty, x_d]$ , potom platí že  $C_n \searrow \emptyset$ . Podobným argumentem jako posledně máme

$$\lim_{x_l \rightarrow -\infty} F(\vec{x}) = \lim_{x \rightarrow -\infty} G(x) = \lim_{n \rightarrow \infty} G(-n) = \lim_{n \rightarrow \infty} P_{\vec{X}}\left(\bigcap_{n=1}^{\infty} C_n\right) = P_{\vec{X}}(\emptyset) = 0,$$

čímž jsme dokázali vlastnost (ii).

Nakonec si uvědomíme, že podmínka z vlastnosti (iii) je ekvivalentní tomu, že  $\lim_{n \rightarrow \infty} \min\{x_l\} = \infty$ . Z již několikrát použité věty o spojitosti pravděpodobnosti máme, že  $1 \geq F_{\vec{X}}(\vec{x}) \geq F_{\vec{X}}(\min\{x_l\}[1, \dots, 1]^T)$ . Stačí tedy dokázat, že poslední uvedená limita je rovna  $\infty$ .

Položme  $H(x) := F_{\vec{X}}(x[1, \dots, 1]^T)$ . Z monotonie pravděpodobnosti máme, že funkce  $H$  je neklesající. Dále  $0 \leq H(x) \leq 1$ , tedy existuje  $\lim_{x \rightarrow \infty} H(x) = \lim_{n \rightarrow \infty} H(n)$  (jako limita posloupnosti). Položme  $D_n := (-\infty, n]^d$ . Opět z věty o spojitosti míry můžeme psát

$$\lim_{x_l \rightarrow +\infty \forall l} F(\vec{x}) = \lim_{n \rightarrow \infty} H(n) = \lim_{n \rightarrow \infty} P_{\vec{X}}(D_n) = P_{\vec{X}}(\mathbb{R}^d) = 1,$$

čímž jsme získali požadovanou rovnost.  $\square$

**Věta 2.37** (Marginální distribuční funkce). *Pokud je  $F_{\vec{X}}$  sdružená distribuční funkce náhodného vektoru  $\vec{X} = [X_1, \dots, X_d]^T$ , pak*

$$\lim_{x_d \rightarrow +\infty} F_{\vec{X}}(x_1, \dots, x_d) = F(x_1, \dots, x_{d-1}), \forall \vec{x} = [x_1, \dots, x_d]^T \in \mathbb{R}^d,$$

kde  $F$  je distribuční funkce náhodného podvektoru  $[X_1, \dots, X_{d-1}]^T$ .

*Důkaz.* Necht je dána libovolná posloupnost  $\{z_n\}_{n=1}^\infty$  taková, že  $\lim z_n = \infty$ . Dále označme  $B := \bigcap_{l=1}^{d-1} \{X_l \leq x_l\}$ ,  $B_n := B \cup \{X_d \leq z_n\}$  a  $D_n := \left(\bigcup_{m=n}^\infty B_m^C\right)^C$ . Platí  $D_n \subseteq B_n \subseteq B = \bigcup_{m=1}^\infty B_m$  a  $D_n \nearrow B$ . Ze spojitosti míry (Věta 1.7) máme, že  $\lim_{n \rightarrow \infty} P(D_n) = P(B)$  a z monotonie míry máme, že  $P(D_n) \leq P(B_n) \leq P(B)$ . Potom (viz věta o dvou strážnících) máme  $\lim_{n \rightarrow \infty} P(B_n) = P(B)$ . Nakonec z Heineovy věty máme, že  $\lim_{x_d \rightarrow \infty} P(B \cap \{X_d \leq x_d\}) = P(B)$ .  $\square$

Výše zmíněný limitní přechod můžeme opakovat vícekrát a “marginalizovat” až na jednorozměrný případ. Navíc, složky můžeme permutovat, tedy v kombinaci s touto větou můžeme “vyřadit” libovolnou složku.

**Definice 2.38** (Marginální rozdělení). Necht  $J \subseteq \{1, \dots, d\}$  a  $|J| = m$ . Potom *náhodný podvektor* definujeme jako  $\vec{Y} \equiv \{X_l\}_{l \in J} : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$  a *marginálním rozdělením* rozumíme indukovanou pravděpodobnostní míru  $P_{\vec{Y}}$  na prostoru  $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$ .

Ve speciálním případě  $J = \{1, \dots, m\}$  pak máme  $P_{\vec{Y}}(B) = P_{\vec{X}}(B \times \mathbb{R}^{d-m})$ . Pro  $|J| = 1$  celkem snadno vidíme, že se jedná o náhodnou veličinu.

V následujících definicích definujeme spojitě a diskrétní náhodné vektory podobně tomu, jak jsme to udělali u náhodných veličin.

**Definice 2.39.** Náhodný vektor  $\vec{X} = [X_1, \dots, X_d]^T$  nazveme *diskrétní*, jestliže existují nejvýše spočetná množina  $I \subseteq \mathbb{N}$  a posloupnosti  $\{\vec{x}_i\}_{i \in I}$  prvků  $\mathbb{R}^d$  a  $\{p_i\}_{i \in I}$  prvků intervalu  $(0, 1]$  takové, že platí

$$P_{\vec{X}} = \sum_{i \in I} p_i \delta_{\vec{x}_i} \text{ a } \sum_{i \in I} p_i = 1,$$

kde  $\delta_{\vec{u}}$  značí Diracovu míru v  $\vec{u} \in \mathbb{R}^d$ .

**Definice 2.40.** Náhodný vektor  $\vec{X} = [X_1, \dots, X_d]^T$  nazveme (*absolutně*) *spojitý*, jestliže  $P_{\vec{X}}$  je absolutně spojitá vůči  $d$ -rozměrné Lebesgueově míře  $\lambda^d$ .

V případě diskrétního náhodného vektoru pak můžeme explicitně uvést sdruženou distribuční funkci  $F_{\vec{X}}(\vec{x}) = \sum_{i \in I} p_i \chi_{[\vec{x}_i, \infty)}(\vec{x}) = \sum_{i \in I}^{\vec{x}_i \leq \vec{x}} p_i$ , kde relaci  $\leq$  uvažujeme po složkách (musí platit pro všechny složky zároveň).

Pro spojitě náhodné vektory si uvědomíme, že sdružená distribuční funkce má derivaci  $\frac{\partial^d}{\partial x_1 \dots \partial x_d} F_{\vec{X}}$  s.v. vzhledem k  $\lambda^d$  a platí následující vztah pro sdruženou hustotu

$$f_{\vec{X}}(\vec{x}) = \frac{\partial^d}{\partial x_1 \dots \partial x_d} F_{\vec{X}}.$$

Tento vztah platí  $\lambda^d$ -skoro všude a navíc v námi zkoumaných příkladech je  $F_{\vec{X}}$  dostatečně hladká, tedy nezáleží na pořadí derivací. Potom také můžeme z hustoty spočítat distribuční funkci pomocí vztahu

$$F_{\vec{X}}(\vec{x}) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_d} f_{\vec{X}}(t_1, \dots, t_d) dt_d \dots dt_1,$$

pro všechna  $\vec{x} = [x_1, \dots, x_d]^T \in \mathbb{R}^d$ . Díky Fubiniově větě opět nezáleží na pořadí integrálů.

*konec 6. přednášky (4.3.2025)*

**Věta 2.41** (Hustota vzhledem k součinové referenční míře). *Nechť  $P_{\vec{X}}$  je rozdělení náhodného vektoru  $\vec{X} = [X_1, \dots, X_d]^T$  a nechť existují  $\sigma$ -konečné míry  $\mu_l, l \in \{1, \dots, d\}$  na  $\mathbb{R}$  takové, že pro jejich součin platí  $P_{\vec{X}} \ll \mu_1 \otimes \dots \otimes \mu_d$ . Potom  $P_{X_l} \ll \mu_l$  pro všechny složky  $l \in \{1, \dots, d\}$ . Dále pak existují nezáporné měřitelné funkce (hustoty)  $f_{\vec{X}} : \mathbb{R}^d \rightarrow [0, \infty)$  a  $f_{X_l} : \mathbb{R} \rightarrow [0, \infty)$  pro  $l \in \{1, \dots, d\}$  takové, že*

$$P_{\vec{X}}(\times_{l=1}^d (-\infty, x_l]) = F_{\vec{X}}(\vec{x}) = \int_{\times_{l=1}^d (-\infty, x_l]} f_{\vec{X}}(\vec{t}) d(\mu_1 \otimes \dots \otimes \mu_d)$$

pro všechna  $\vec{x} = [x_1, \dots, x_d]^T$ . Pro borelovskou množinu  $B \in \mathcal{B}(\mathbb{R}^d)$  navíc platí

$$P_{\vec{X}}(B) = \int_B f_{\vec{X}}(\vec{t}) d(\mu_1 \otimes \dots \otimes \mu_d).$$

Potom také  $F_{X_l}(x_l) = \int_{-\infty}^{x_l} f_{X_l}(t) d\mu_l$  pro všechna  $x_l \in \mathbb{R}$  a všechny složky  $l \in \{1, \dots, d\}$ , kde

$$f_{X_l}(t) = \int_{\mathbb{R}^{d-1}} f_{\vec{X}}(t_1, \dots, t_d) d(\mu_1 \otimes \dots \otimes \mu_{l-1} \otimes \mu_{l+1} \otimes \dots \otimes \mu_d)$$

platí  $\mu_l$ -skoro všude.

*Důkaz.* Existence hustot plyne přímo z Radonovy-Nikodymovy věty. Zbytek plyne z Fubiniho věty (předpoklady splněny díky Radonově-Nikodymově větě a faktu, že pravděpodobnostní prostor je vždy normalizovaný).  $\square$

Poznamenejme, že předpoklad existence příslušných měr je automaticky splněn v případě diskrétních nebo absolutně spojitých náhodných vektorů.

**Příklad 2.42.** Mějme absolutně spojitý náhodný vektor  $[X, Y]^T$ . Potom existuje jeho sdružená distribuční funkce  $F_{[X, Y]^T}(x, y)$ . Chceme-li dostat jednorozměrnou distribuční funkci  $F_X(x)$ , s použitím Věty 2.37 dostáváme  $F_X(x) = \lim_{y \rightarrow \infty} F_{[X, Y]^T}(x, y)$ . Potom jeho hustotu dostaneme, zderivováním  $f_X(x) = F'_X(x)$ . Navíc z předchozí věty (Věta 2.41) máme, že existuje sdružená hustota  $f_{[X, Y]^T}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{[X, Y]^T}(x, y)$ . Pro získání jednorozměrné hustoty  $f_X$  pak už jen stačí integrovat podle  $y$  přes celou reálnou osu.

Nechť  $\vec{X}$  je diskrétní náhodný vektor a  $\nu$  čítací míra na  $\{\vec{x}_i\}_{i \in I} \subset \mathbb{R}^d$ , pak hustotu tohoto vektoru vzhledem k čítací míře  $\nu$  nazýváme *pravděpodobnostní funkcí* diskrétního mnohorozměrného rozdělení  $\vec{X}$ .

**Příklad 2.43.** Uvažujme dvojrozměrný náhodný vektor  $[X, Y]^T$ . Pro přehlednost uvedeme i řádkové/sloupkové součty (jde o marginální hustoty).

	$Y = 0$	$Y = 1$	
$X = 0$	1/9	2/9	1/3
$X = 1$	2/9	4/9	2/3
	1/3	2/3	1

Potom přímým výpočtem můžeme získat například  $f_{[X,Y]^T}(1,1) = P(X = 1, Y = 1) = 4/9$ .

**Příklad 2.44.** Necht dvourozměrný náhodný vektor  $[X, Y]^T$  je rovnoměrně rozdělen na jednotkovém čtverci. Pak z Věty 2.41 okamžitě vychází  $f_{[X,Y]^T}(x, y) = \chi_{\{(x,y) \in [0,1]^2\}}$ . Určíme  $P[X < 1/2, Y < 1/2]$ . Tato událost  $B := \{X < 1/2, Y < 1/2\}$  odpovídá podmnožině jednotkového čtverce, tedy zintegrováním přes tuto podmnožinu nepočítáme nic jiného než plošný obsah této množiny. Z Fubiniovy věty tedy dostáváme  $P(B) = 1/4$ .

Uvedeme užitečný důsledek předchozí věty pouze v případě dvourozměrných vektorů, snadno se však dají rozšířit i pro případ vícerozměrných vektorů.

**Důsledek 2.45** (Marginální rozdělení pro dvourozměrné náhodné vektory). *Pokud diskrétní náhodný vektor  $[X, Y]^T$  má sdruženou pravděpodobnostní funkci  $f_{[X,Y]^T}$ , pak marginální pravděpodobnostní funkce pro  $X$  je*

$$f_X(x) = P[X = x] = \sum_y P(X = x, Y = y) = \sum_y f_{[X,Y]^T}(x, y).$$

*Pokud spojitý náhodný vektor  $[X, Y]^T$  má sdruženou pravděpodobnostní funkci  $f_{[X,Y]^T}$ , pak marginální hustota pro  $X$  je*

$$f_X(x) = \int f_{[X,Y]^T}(x, y) dy.$$

Věnujme opět pozornost pojmu nezávislosti náhodných veličin. Všimněme si, že platí následující vlastnost

$$\lim_{x_j \rightarrow \infty \forall j \in \{1, \dots, d\} \setminus l} F_{\vec{X}}(\vec{x}) = \lim_{x_j \rightarrow \infty \forall j \in \{1, \dots, d\} \setminus l} P[X_1 \leq x_1, \dots, X_d \leq x_d] = P[X_l \leq x_l] =: F_{X_l}(x_l).$$

**Definice 2.46.** Náhodné veličiny  $X_1, \dots, X_d$  jsou *nezávislé*, pokud  $F_{\vec{X}}(\vec{x}) = \prod_{l=1}^d F_{X_l}(x_l)$  pro každý vektor  $\vec{x} = [x_1, \dots, x_d]^T \in \mathbb{R}^d$ .

Analogicky s předchozí definicí definujeme nezávislost náhodných vektorů. V literatuře se vyskytuje i jiná, ekvivaletní, definice nezávislosti, kterou uvedeme později.

**Definice 2.47.** Náhodné vektory  $\vec{X}_1, \dots, \vec{X}_d$  jsou *nezávislé*, pokud

$$F_{\vec{X}}(\vec{x}) = \prod_{l=1}^d F_{\vec{X}_l}(\vec{x}_l)$$

pro každý “nad-vektor”  $\vec{x} = [\vec{x}_1^T, \dots, \vec{x}_d^T]^T \in \mathbb{R}^{\sum_{l=1}^d d_l}$  kde  $\vec{X} = [\vec{X}_1^T, \dots, \vec{X}_d^T]^T$  a  $\vec{X}_l$  jsou  $d_l$ -rozměrné náhodné vektory pro všechna  $l \in \{1, \dots, d\}$ .

Dalším důležitým pojmem je takzvaný nosič náhodné veličiny. Rozumíme tím v zásadě množinu, kde náhodná veličina “žije”. Uvedeme zde definici pro diskrétní a spojitě náhodné veličiny. Tyto pojmy se budou hodit pro vymezení prostoru, přes který poté budeme integrovat.

**Definice 2.48.** *Nosičem* diskrétní náhodné veličiny  $X$  nazýváme následující množinu  $S(X) = \{x \in \mathbb{R} : P[X = x] > 0\}$ . *Nosičem* spojitě náhodné veličiny  $Y$  rozumíme množinu  $S(Y) = \{y \in \mathbb{R} : f_Y(y) > 0\}$ . Obdobně definujeme i nosič náhodného vektoru (cvičení).

**Věta 2.49** (Ekvivalentní charakterizace nezávislosti). *Nechť sdružená pravděpodobnostní funkce diskrétního náhodného vektoru  $\vec{X} = [X_1, \dots, X_d]^T$  je  $f_{\vec{X}}(\vec{x}) = P[\vec{X} = \vec{x}]$ . Pak platí, že náhodné veličiny  $\{X_1, \dots, X_d\}$  jsou nezávislé právě tehdy, když*

$$P[\vec{X} = \vec{x}] = \prod_{l=1}^d P[X_l = x_l]$$

pro všechny vektory  $\vec{x} = [x_1, \dots, x_d]^T \in \times_{l=1}^d S(X_l)$ .

*Nechť sdružená pravděpodobnostní funkce spojitěho náhodného vektoru  $\vec{X} = [X_1, \dots, X_d]^T$  je  $f_{\vec{X}}(\vec{x})$ . Pak platí, že  $\{X_1, \dots, X_d\}$  jsou nezávislé právě tehdy, když*

$$P[\vec{X} = \vec{x}] = \prod_{l=1}^d f_{X_l}(x_l)$$

pro  $\lambda^d$ -skoro všechny vektory  $\vec{x} = [x_1, \dots, x_d]^T \in \times_{l=1}^d S(X_l)$ .

*Důkaz.* Dokážeme oba případy (spojitý a diskrétní) najednou tak, že budeme uvažovat příslušnou referenční součinovou míru.

Nejdříve dokážeme implikaci  $\Rightarrow$ . Uvažujme vektor  $\vec{x} = [x_1, \dots, x_d]^T$ . Potom z definice nezávislosti a linearity integrálu dostáváme

$$F_{\vec{X}}(\vec{x}) = \prod_{l=1}^d F_{X_l}(x_l) = \prod_{l=1}^d \int_{-\infty}^{x_l} f_{x_l}(t_l) d\mu_l = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_d} \prod_{l=1}^d f_{x_l}(t_l) d\mu_d \cdots d\mu_1,$$

kde druhá rovnost plyne z věty o hustotě vzhledem k součinové referenční míře (Věta 2.41) s mírou  $\lambda^d$ , případně sčítací mírou  $\nu$  na  $\mathbb{R}^d$ . Dále díky Fubiniově větě můžeme pokračovat v úpravách

$$\int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_d} \prod_{l=1}^d f_{x_l}(t_l) d\mu_d \cdots d\mu_1 = \int_{(-\infty, \vec{x}]} \prod_{l=1}^d f_{x_l}(t_l) d(\mu_1 \otimes \cdots \otimes \mu_d).$$

Pak už ale nutně musí platit  $f_{\vec{X}}(t_1, \dots, t_d) = \prod_{l=1}^d f_{x_l}(t_l)$ .

Implikace  $\Leftarrow$  se dokáže obráceným postupem (cvičení). □

**Důsledek 2.50.** Předpokládejme, že  $S([X, Y]^T) = S(X) \times S(Y)$ . Pokud pro sdruženou pravděpodobnostní funkci diskrétního náhodného vektoru  $[X, Y]^T$  platí  $P[X = x, Y = y] = g(x)h(y)$  pro nějaké měřitelné funkce  $g, h$  a pro všechna  $[x, y]^T \in S([X, Y]^T)$ , pak  $X$  a  $Y$  jsou nezávislé.

Obdobně, pokud pro sdruženou pravděpodobnostní funkci spojitého náhodného vektoru  $[X, Y]^T$  platí  $f_{[X, Y]^T}(x, y) = g(x)h(y)$  pro nějaké měřitelné funkce  $g, h$  a pro  $\lambda^2$ -skoro všechna  $[x, y]^T \in S([X, Y]^T)$ , pak  $X$  a  $Y$  jsou nezávislé.

Poznámka:  $g$  a  $h$  nutně nemusí být hustoty nebo pravděpodobnostní funkce, ale normovaná funkce  $\tilde{g} := \frac{g}{\int g}$  již ano.

*Důkaz.* Definujme  $\tilde{g}$  a  $\tilde{h}$  jako v poznámce. Potom musí existovat verze  $\tilde{g}, \tilde{h} \geq 0$  měřitelná. Zbytek dostaneme z předchozí věty.  $\square$

konec 7. přednášky (10.3.2025)

**Věta 2.51** (Alternativní (ekvivalentní) definice nezávislosti). Náhodné veličiny  $X_1, \dots, X_d$  jsou nezávislé právě tehdy, když

$$P_{\vec{X}} = \otimes_{l=1}^d P_{X_l}.$$

*Důkaz.* Začneme implikací zprava doleva ( $\Leftarrow$ ). Jestliže pro všechny množiny  $B_l \in \mathcal{B}(\mathbb{R}), l = 1, \dots, d$  platí

$$P_{\vec{X}}(\times_{l=1}^d B_l) = \prod_{l=1}^d P_{X_l}(B_l),$$

pak vezmeme  $B_l = (-\infty, x_l]$  pro fixní  $\vec{x} = [x_1, \dots, x_d]$  generátory borelovské  $\sigma$ -algebry a můžeme počítat

$$F_{\vec{X}}(\vec{x}) = P_{\vec{X}}(\times_{l=1}^d (-\infty, x_l]) = \prod_{l=1}^d P_{X_l}((-\infty, x_l]) = \prod_{l=1}^d F_{X_l}(x_l),$$

což je přímo definice nezávislosti.

K důkazu opačné implikace (předpokládáme platnost  $F_{\vec{X}}(\vec{x}) = \prod_{l=1}^d F_{X_l}(x_l)$  pro všechna  $\vec{x} \in \mathbb{R}^d$ ) použijeme Dynkinův systém

$$D = \{ \times_{l=1}^d (-\infty, x_l], x_l \in \mathbb{R}, \forall l \in \{1, \dots, d\} \}.$$

Tento systém je uzavřený na průniky a generuje celou borelovskou  $\sigma$ -algebru  $\mathcal{B}(\mathbb{R}^d)$ . Jelikož  $P_{\vec{X}}(\mathbb{R}^d) = 1$ , dostáváme z věty o jednoznačnosti míry rovnost obou měr ( $P_{\vec{X}}$  a  $\otimes_{l=1}^d P_{X_l}$ ) na celé  $\mathcal{B}(\mathbb{R}^d)$ .  $\square$

Vrátíme se opět k podmíněnosti, tentokrát budeme zkoumat podmíněnost náhodných veličin. Motivačním příkladem budiž zjištění průměrné mzdy občana, který vystudoval MatFyz na základě znalosti průměrné mzdy všech občanů ČR. Opět se jedná o zjednodušenou definici, ta obecnější bude uvedena v pokročilejších kurzech.

**Definice 2.52.** Pro diskrétní náhodný vektor  $[X, Y]^T$  definujeme *podmíněnou pravděpodobnostní funkci*  $X$  za podmínky  $Y = y$  vztahem

$$f_{X|Y}(x|y) \equiv P[X = x|Y = y] := \frac{P[X = x, Y = y]}{P[Y = y]} \equiv \frac{f_{[X,Y]^T}(x, y)}{f_Y(y)},$$

pokud  $P[Y = y] > 0$ .

Pro spojitý náhodný vektor  $[X, Y]^T$  je *podmíněná hustota*  $X$  za podmínky  $Y = y$

$$f_{X|Y}(x|y) := \frac{f_{(X,Y)}(x, y)}{f_Y(y)},$$

pokud  $f_Y(y) > 0$ .

Je třeba si uvědomit, že podmíněná pravděpodobnostní funkce (hustota) jsou funkce argumentu  $x$  s parametrem  $y$ . V případech nepokrytých touto definicí můžeme definovat podmíněnou pravděpodobnost libovolně. Několik užitečných vzorců pro výpočet podmíněné pravděpodobnosti, pro diskrétní vektor platí  $P[X \in A|Y = y] = \sum_{x \in A} P[X = x|Y = y]$  a pro spojitý vektor platí  $P[X \in A|Y = y] = \int_A f_{X|Y}(x|y)dy$ .

V dalších kapitolách budeme pracovat s mnohorozměrným normálním rozdělením, je proto vhodné si ho zadefinovat už teď. V obecném případě nestačí zadefinovat chování po složkách, je třeba nějakým způsobem zahrnout i vztahy mezi jednotlivými složkami.

**Definice 2.53.** Náhodný vektor  $\vec{X} = [X_1, \dots, X_d]^T$  má *d-rozměrné normální rozdělení* s parametry  $\mu \in \mathbb{R}^d$  a  $\Sigma \in \mathbb{R}^{d \times d}$  (značíme  $X \sim N_d(\mu, \Sigma)$ ), pokud existuje  $k$ -rozměrný náhodný vektor  $\vec{Y} = [Y_1, \dots, Y_k]^T$  a matice  $\mathbb{A} \in \mathbb{R}^{d \times k}$  takové, že

- (i)  $\{Y_1, \dots, Y_k\}$  jsou nezávislé;
- (ii)  $Y_j \sim N(0, 1)$  pro všechny složky  $j \in \{1, \dots, k\}$ ;
- (iii)  $\mathbb{A}\mathbb{A}^T = \Sigma$ ;
- (iv)  $\vec{X} = \mathbb{A}\vec{Y} + \mu$ .

Takto komplikovaná definice je potřeba, neboť matice  $\Sigma$  nutně nemusí mít jednoznačně určenou “druhou odmocninu”, proto musíme použít nějakou druhou odmocninu, která bude případně dávat nižší hodnotu. Takto definovaná matice  $\mathbb{A}$  a vektor  $\mu$  jsou deterministické (nenáhodné). Pozornost si zaslouží tzv. standardní  $d$ -rozměrné normální rozdělení  $N_d(\vec{0}, I_d)$ .

Na závěr se budeme chvíli věnovat transformacím náhodných veličin. V obecném případě je možné formalizovat tuto představu pomocí věty o substituci z TMI, avšak pro naše účely postačí uvést jen několik speciálních případů.

Mějme diskrétní náhodnou veličinu  $X$  a ne nutně monotónní funkci  $t : \mathbb{R} \rightarrow \mathbb{R}$ , pro kterou platí  $Y := t(X)$ . Chceme odvodit pravděpodobnostní funkci

$P[Y = y]$ . Můžeme psát

$$P[Y = y] = P[t(X) = y] = P[X \in t^{-1}(y)] = \sum_{t(x)=y} P[X = x].$$

Dále mějme spojitou náhodnou veličinu  $X$ , známe její hustotu  $f_X(x)$ . Cílem je spočítat hustotu  $f_Y(y)$ , kde  $Y = t(X)$ . Pro každé  $y$  můžeme nalézt množinu  $\mathcal{T}_y = \{x : t(x) \leq y\}$ . Poté můžeme spočítat distribuční funkci rozdělení  $Y$ .

$$F_Y(y) = P[Y \leq y] = P[t(X) \leq Y] = P[\omega : t(X(\omega)) \leq Y] = \int_{\mathcal{T}(y)} f_X(x) dx,$$

hustotu poté můžeme získat pouhým zderivováním distribuční funkce  $F_Y$ .

Dále uvažujme případ (dvourozměrného) diskrétního náhodného vektoru  $[X, Y]^T$  a transformace  $Z = t(X, Y)$ . Ze znalosti diskrétního rozdělení vektoru  $[X, Y]$  chceme spočítat  $P[Z = z]$ . Můžeme psát

$$\begin{aligned} P[Z = z] &= P[t(X, Y) = z] = P[\omega : t([X, Y]^T(\omega)) = z] = \\ &= P[[X, Y]^T \in t^{-1}(z)] = \sum_{t(x,y)=z} P[X = x, Y = y]. \end{aligned}$$

Nakonec mějme spojitý náhodný vektor  $[X, Y]^T$ , pro nějž známe hustotu  $f_{[X,Y]^T}(x, y)$  a chceme získat hustotu  $f_Z(z)$ , jestliže náhodná veličina  $Z$  je definována  $Z := t(X, Y)$ . V analogii se spojitou náhodnou veličinou nalezneme pro každé  $z \in \mathbb{R}$  množinu  $\mathcal{T}(z) = \{[x, y] : t(x, y) \leq z\}$ . Opět spočteme distribuční funkci  $F_Z$  pomocí následujících kroků

$$\begin{aligned} F_Z(z) &= P[Z \leq z] = P[t(X, Y) \leq z] = P[\omega : t([X, Y]^T(\omega)) \leq z] = \\ &= \iint_{\mathcal{T}(z)} f_{[X,Y]^T}(x, y) dx dy, \end{aligned}$$

hustotu  $f_Z$  opět můžeme získat pomocí zderivování funkce  $F_Z$ .



### 3 Střední hodnota

V této kapitole se budeme věnovat pojmu střední hodnoty, laicky řečeno, kolem jaké hodnoty se nachází naše rozdělení. Nejedná se ani o průměr ani o prostřední, případně nejčastější hodnotu, tyto pojmy zadefinujeme později a ve statistice mají svůj vlastní význam odlišný od střední hodnoty.

**Definice 3.1.** *Střední hodnota* náhodné veličiny  $X$  je reálné číslo

$$\mathbb{E}[X] = \mathbb{E}X = \int X dP \equiv \int X(\omega) dP(\omega),$$

pokud pravá strana existuje.

Tato definice je velmi teoretická, k praktickému výpočtu se hodí následující věta, kde převedeme integrál na výpočet pomocí obrazu pravděpodobnostní míry.

**Věta 3.2.** *Střední hodnota náhodné veličiny  $X$  je  $\mathbb{E}X = \int x dP_X(x)$ , pokud pravá strana existuje.*

*Důkaz.* Z věty o přenosu integrace (2.5) při volbě  $g = Id$  a  $(\mathbb{M}, \mathcal{M}) = (\mathbb{R}, \mathcal{B})$  dostáváme požadované tvrzení.  $\square$

Z Radon-Nikodymovy věty ihned plyne následující pozorování

**Pozorování 3.3.** *Střední hodnota veličiny  $X$  je*

$$\mathbb{E}X = \begin{cases} \int_{-\infty}^{\infty} x f_X(x) dx, & X \text{ spojitá}; \\ \sum_{x \in S(X)} x P[X = x], & X \text{ diskrétní}. \end{cases}$$

Střední hodnota nemusí existovat vždy, jeden z takových případů uvedeme v následujícím příkladu.

**Příklad 3.4.** Pokud  $X \sim \text{Cauchy}$  (Definice 2.32), pak  $\mathbb{E}X$  neexistuje. Pomocí integrování per partes můžeme počítat

$$\int_0^{\infty} \frac{x}{\pi(1+x^2)} dx = [x \arctan(x)]_0^{\infty} - \int_0^{\infty} \arctan(x) dx = \infty.$$

Dostali jsme, že pro integrál přes celou reálnou přímku není definován výraz  $\infty - \infty$ .

Uvažujme teď transformaci  $Y = t(X)$ . Následující věta nám umožní počítat střední hodnotu transformované náhodné veličiny.

**Věta 3.5** (Pravidlo líného statistika). *Buď  $t : \mathbb{R} \rightarrow \mathbb{R}$  měřitelná funkce a necht  $Y = t(X)$ , kde  $X$  je nějaká náhodná veličina. Pak*

$$\mathbb{E}Y = \int t(x) dP_X(x),$$

*pokud pravá strana existuje.*

*Důkaz.* Z Věty 2.5 dostáváme

$$\mathbb{E}Y = \mathbb{E}[t(X)] = \int_{\mathbb{R}} t(X(\omega))dP(\omega) = \int_{\mathbb{R}} t(x)dP_X(x).$$

□

Poznamenejme si explicitní vzorce pro transformaci spojitých a diskrétní náhodných veličin, které jsou přímým důsledkem předchozí věty:

**Důsledek 3.6.** *Mějme náhodné veličiny  $X$  a  $Y$  takové, že platí  $Y = t(X)$  pro nějakou transformaci  $t : \mathbb{R} \rightarrow \mathbb{R}$ . Má-li  $X$  diskrétní rozdělení, potom*

$$\mathbb{E}Y = \sum_{x \in S(X)} t(x)P[X = x].$$

*Je-li  $X$  spojitá, potom platí*

$$\mathbb{E}Y = \int_{\mathbb{R}} t(x)f_X(x)dx.$$

Přímé využití pravidla lineárního statistika si uvedeme v definici a aplikacích následujícího pojmu, který jistým způsobem umožňuje charakterizovat chování rozdělení.

**Definice 3.7.** Pro reálné číslo  $k$  definujeme  $k$ -tý *moment* náhodné veličiny  $X$  jako  $\mathbb{E}[X^k]$  za předpokladu, že  $\mathbb{E}[|X|^k] < \infty$ . Dále definujeme  $k$ -tý *absolutní moment* jako  $\mathbb{E}[|X|^k]$ , pokud existuje.

V praxi se nejčastěji setkáme s momenty jen pro  $k$  přirozené, pokud nebude řečeno jinak, všechny momenty budou mít přirozený parametr.

**Věta 3.8.** *Pokud existuje  $k$ -tý moment, pak existuje  $l$ -tý moment pro jakékoli  $l \in \{1, \dots, k\}$ .*

*Důkaz.* Potřebujeme ukázat, že  $\mathbb{E}[|X|^l] < \infty$ . Můžeme počítat

$$\begin{aligned} \mathbb{E}[|X|^l] &= \int_{\mathbb{R}} |x|^l dP_X(x) = \int_{|x| \leq 1} |x|^l dP_X(x) + \int_{|x| > 1} |x|^l dP_X(x) \leq \\ &\int_{|x| \leq 1} dP_X(x) + \int_{|x| > 1} |x|^k dP_X(x) \leq \int_{\mathbb{R}} dP_X(x) + \int_{\mathbb{R}} |x|^k dP_X(x). \end{aligned}$$

Dostáváme  $1 + \mathbb{E}[|X|^k] < \infty$ , čímž je důkaz ukončen. □

Některá zobrazení mají jen několik prvních momentů a žádné vyšší momenty neexistují, jako například Studentovo  $t$ -rozdělení (Definice 2.31) s  $\nu = 3$  stupni volnosti.

**Příklad 3.9.** Pro  $X \sim t_3$  platí  $\mathbb{E}X = 0$ ,  $\mathbb{E}X^2 = 2$  ale  $\mathbb{E}|X|^3 = \infty$ . (cvičení, použijte per partes)

Definujeme dále  $\mathcal{L}^p$  prostory náhodných veličin, které jsou podobné  $L^p$  prostorům z teorie míry a funkcionální analýzy.

**Definice 3.10.** Pro reálné číslo  $p$  (v praxi se vystačíme pouze s případem  $p \geq 1$ ) definujeme prostor  $\mathcal{L}^p$  tak, že náhodná veličina  $X \in \mathcal{L}^p$ , jestliže  $\mathbb{E}[|X|^p] < \infty$ .

Ukážeme si pár základních vlastností prostoru  $\mathcal{L}^1$ , které se mohou hodit při praktických aplikacích.

**Věta 3.11** (Základní vlastnosti prostoru  $\mathcal{L}^1$ ). *Nechť jsou dány  $X_1, \dots, X_d \in \mathcal{L}^1$  a  $a_1, \dots, a_d$  jsou konstanty, pak platí linearita ve smyslu*

$$\mathbb{E} \left( \sum_{l=1}^d a_l X_l \right) = \sum_{l=1}^d a_l \mathbb{E} X_l.$$

*Dále mějme  $X_1, \dots, X_d \in \mathcal{L}^1$  nezávislé náhodné veličiny, potom platí*

$$\mathbb{E} \left( \prod_{l=1}^d X_l \right) = \prod_{l=1}^d \mathbb{E} X_l.$$

*Důkaz.* Linearita plyne z věty o přenosu integrace (Věta 2.5) a linearitě Lebesgueova integrálu.

Dokážeme druhou vlastnost. Nejprve ukážeme, že hledaná střední hodnota je dobře definovaná. Uvažujme posloupnost funkcí  $\{g_n : \mathbb{R}^d \rightarrow \mathbb{R}\}$  definovaných jako  $g_n(\vec{x}) = \prod_{l=1}^d |x_l| \chi_{\{|x_l| \leq n\}}$ . Pak pro každé  $n \in \mathbb{N}$  je  $g_n(\vec{X})$  omezená a existuje její první moment  $\mathbb{E}[g_n(\vec{X})] \in \mathbb{R}$ . Díky nezávislosti můžeme psát

$$\mathbb{E}[g_n(\vec{X})] = \int_{\mathbb{R}^d} \prod_{l=1}^d |x_l| \chi_{\{|x_l| \leq n\}} d(\otimes_{l=1}^d P_{X_l}),$$

odkud z Fubiniovy věty a následně linearitě integrálu plyne

$$= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \prod_{l=1}^d |x_l| \chi_{\{|x_l| \leq n\}} dP_{X_1} \cdots dP_{X_d} = \prod_{l=1}^d \mathbb{E}[|X_l| \chi_{\{|X_l| \leq n\}}] \leq \prod_{l=1}^d \mathbb{E}[|X_l|].$$

Platí, že funkce  $g_n(\vec{x})$  jsou nezáporné a  $g_n(\vec{x}) \uparrow \prod_{l=1}^d |x_l|$  na celém  $\mathbb{R}^d$ . Tudíž z Leviho věty plyne, že  $\mathbb{E}[g_n(\vec{X})] \uparrow \mathbb{E}[\prod_{l=1}^d |X_l|]$ . Potom ale nutně  $\mathbb{E} \left| \prod_{l=1}^d X_l \right| \leq \prod_{l=1}^d \mathbb{E}|X_l| < \infty$ , tedy příslušný první moment existuje.

Dále můžeme počítat

$$\begin{aligned} \mathbb{E} \left[ \prod_{l=1}^d X_l \right] &= \int_{\mathbb{R}^d} \prod_{l=1}^d x_l dP_{\vec{X}} = \int_{\mathbb{R}} \prod_{l=1}^d x_l d(\otimes_{l=1}^d P_{X_l}) = \\ &= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \prod_{l=1}^d x_l dP_{X_1} \cdots dP_{X_d} = \prod_{l=1}^d \int_{\mathbb{R}} x_l dP_{X_l} = \prod_{l=1}^d \mathbb{E} X_l, \end{aligned}$$

kde druhá rovnost plyne z nezávislosti náhodných veličin  $X_1, \dots, X_d$ , třetí z Fubiniovy věty a předposlední z linearitě integrálu.  $\square$

konec 8. přednášky (11.3.2025)

Teď definujeme další neplnohodnotnou (jinými slovy, neurčuje danou náhodnou veličinu, případně její rozdělení jednoznačně) charakteristiku. (Poznámka: příkladem plnohodnotné charakteristiky je distribuční funkce rozdělení)

**Definice 3.12.** Rozptyl náhodné veličiny  $X$  je definován jako

$$\text{Var } X = \mathbb{E}(X - \mathbb{E}X)^2,$$

za předpokladu, že pravá strana je dobře definovaná. Pak *směrodatná odchylka* též náhodné veličiny je definovaná je

$$\text{sd}(X) = \sqrt{\text{Var } X}.$$

Uvědomme si, že některé volby charakterizování variability rozdělení nejsou vhodné, například na první pohled logické " $\mathbb{E}(X - \mathbb{E}X)$ " je nulová všude, kde je definovaná.

**Věta 3.13** (Vlastnosti rozptylu). *Za předpokladu, že uvažované druhé momenty jsou konečné, potom*

$$\text{Var } X = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \geq 0.$$

*Pokud  $a, b \in \mathbb{R}$ , pak*

$$\text{Var}(aX + b) = a^2 \text{Var } X,$$

*jinými slovy, rozptyl se chová jako kvadratická forma. Pokud  $X_1, \dots, X_2$  jsou nezávislé a  $a_1, \dots, a_d$  jsou reálné konstanty, pak*

$$\text{Var} \left( \sum_{l=1}^d a_l X_l \right) = \sum_{l=1}^d a_l^2 \text{Var } X_l.$$

*Důkaz.* Dokážeme první vlastnost. Máme

$$\text{Var } X = \mathbb{E}[(X - \mathbb{E}X)^2] = \mathbb{E}[X^2 - 2X(\mathbb{E}X) + (\mathbb{E}X)^2] = \mathbb{E}X^2 - 2(\mathbb{E}X)^2 + (\mathbb{E}X)^2,$$

kde předposlední rovnost plyne z linearity střední hodnoty a faktu, že  $E[c] = c$  pro konstantu  $c$ . Nezápornost plyne z toho, že počítáme střední hodnotu nezáporné náhodné veličiny.

Dále pro druhou vlastnost píšme

$$\text{Var}(aX + b) = \mathbb{E}[(aX + b - \mathbb{E}(aX + b))^2] = \mathbb{E}[a^2 (X - \mathbb{E}X)^2],$$

kde druhá rovnost plyne z linearity střední hodnoty a rovnosti  $b = \mathbb{E}b$ , dále můžeme psát

$$\mathbb{E}[a^2 (X - \mathbb{E}X)^2] = a^2 \mathbb{E}(X - \mathbb{E}X)^2 = a^2 \text{Var } X.$$

K důkazu poslední vlastnosti začneme opět rozepsáním definice

$$\begin{aligned}\text{Var}\left(\sum_{l=1}^d a_l X_l\right) &= \mathbb{E}\left[\sum_{l=1}^d a_l X_l - \mathbb{E}\left(\sum_{l=1}^d a_l X_l\right)\right]^2 = \mathbb{E}\left[\sum_{l=1}^d a_l (X_l - \mathbb{E}X_l)\right]^2 = \\ &= \mathbb{E}\left[\sum_{l=1}^d a_l^2 (X_l - \mathbb{E}X_l)^2 + 2 \sum_{1 \leq j < l \leq d} a_j a_l (X_j - \mathbb{E}X_j)(X_l - \mathbb{E}X_l)\right] = \\ &= \sum_{l=1}^d a_l^2 \mathbb{E}(X_l - \mathbb{E}X_l)^2 + 2 \sum_{1 \leq j < l \leq d} a_j a_l \mathbb{E}(X_j - \mathbb{E}X_j)(X_l - \mathbb{E}X_l) = \sum_{l=1}^d a_l^2 \text{Var} X_l,\end{aligned}$$

kde poslední rovnost plyne z toho, že v případě  $j \neq l$  máme díky nezávislosti

$$\mathbb{E}(X_j - \mathbb{E}X_j)(X_l - \mathbb{E}X_l) = \mathbb{E}[X_j X_l] - \mathbb{E}[X_l]\mathbb{E}[X_j] = 0.$$

□

Dalším pojmem, kterému se budeme věnovat, je kovariance a korelace, které charakterizují lineární vztah mezi dvěma náhodnými veličinami.

**Definice 3.14.** Kovariance mezi  $X$  a  $Y$  je definována jako

$$\text{Cov}(X, Y) = \mathbb{E}((X - \mathbb{E}X)(Y - \mathbb{E}Y)).$$

Pokud  $\text{Var}(X) \text{Var}(Y) > 0$ , pak definujeme *korelaci* mezi  $X$  a  $Y$  vztahem

$$\rho_{X,Y} \equiv \text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}}.$$

Je třeba si dávat pozor, že tyto pojmy necharakterizují libovolnou souvislost mezi náhodnými veličinami, ale pouze lineární. Navíc, korelace nemusí způsobovat kauzalitu (spotřeba čokolády v dané zemi sice koreluje s počtem Nobelových laureátů, ale nemůžeme zvýšit počet laureátů tím, že zvýšíme spotřebu čokolády).

Dále si všimneme, že z pravidla líného statistika (Věta 3.5) okamžitě plynou následující vztahy.

**Důsledek 3.15.** Pro  $X, Y$  spojité platí  $\mathbb{E}XY = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{(X,Y)}(x, y) dx dy$ .  
Pro  $X, Y$  diskrétní platí  $\mathbb{E}XY = \sum_{x \in S(X), y \in S(Y)} xy P[X = x, Y = y]$ .

Dále si zformulujeme několik vlastností kovariance a korelace.

**Věta 3.16.** Pro náhodné veličiny  $X$  a  $Y$  platí následující tvrzení (jsou-li příslušné matematické objekty dobře definovány).

- (i)  $\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}X\mathbb{E}Y$ ;
- (ii)  $-1 \leq \text{Corr}(X, Y) \leq 1$ ;

(iii)  $|\text{Corr}(X, Y)| = 1 \Leftrightarrow Y = aX + b$  s pravděpodobností 1 pro nějaké hodnoty  $a, b \in \mathbb{R}$ .

(iv) Pro nezávislé  $X$  a  $Y$  platí  $\text{Cov}(X, Y) = 0$ . Pozor: opačná implikace nemusí platit (stačí vzít  $X \sim U(-1, 1), Y = X^2$ ).

*Důkaz.* Budeme postupovat postupně, k důkazu první vlastnosti použijeme následující výpočet:

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}X)(Y - \mathbb{E}Y)] = \mathbb{E}(XY) - \mathbb{E}X\mathbb{E}Y,$$

kde druhá rovnost se získá roznásobením závorek analogicky s důkazem předchozí věty. Z tohoto okamžitě plyne vlastnost (iv), neboť nezávislost  $X$  a  $Y$  implikuje, že  $\mathbb{E}(XY) = \mathbb{E}X\mathbb{E}Y$ .

K důkazu vlastnosti (ii) použijeme Cauchyovu-Schwarzovu nerovnost (Věta 4.4), kterou si zde dokážeme. Definujeme funkci  $g(a) := \mathbb{E}(aX - Y)^2$ , potom

$$0 \leq \mathbb{E}(aX - Y)^2 = \mathbb{E}(a^2X^2 - 2aXY + Y^2) = a^2\mathbb{E}X^2 - 2a\mathbb{E}XY + \mathbb{E}Y^2.$$

Funkci  $g(a)$  můžeme zderivovat, dostáváme

$$g'(a) = 2a\mathbb{E}X^2 - 2\mathbb{E}XY,$$

svého minima tedy funkce  $g$  nabývá v bodě  $\frac{\mathbb{E}XY}{\mathbb{E}X^2}$  (bez újmy na obecnosti  $\mathbb{E}X^2 \neq 0$ , v opačném případě máme  $X = 0$  skoro jistě, z čehož vlastnosti z věty plynou triviálně). Dosadíme tuto hodnotu do předpisu funkce  $g(a)$  a dostáváme.

$$g\left(\frac{\mathbb{E}XY}{\mathbb{E}X^2}\right) = \frac{(\mathbb{E}XY)^2}{\mathbb{E}X^2} - 2\frac{(\mathbb{E}XY)^2}{\mathbb{E}X^2} + \mathbb{E}Y^2 \geq 0.$$

Z toho již plyne, že  $(\mathbb{E}XY)^2 \leq (\mathbb{E}X^2)(\mathbb{E}Y^2)$  (dokázali jsme Cauchy-Schwarz!), z čehož plyne požadované tvrzení.

Vlastnost (iii) budeme dokazovat po implikacích. Nejdříve předpokládejme, že  $Y = aX + b$  pro nějaká  $a, b \in \mathbb{R}$ . Potom máme

$$\begin{aligned} \text{Cov}(X, Y) &= \text{Cov}(X, aX + b) = \mathbb{E}[X(aX + b)] - \mathbb{E}X\mathbb{E}(aX + b) = \\ &= a\mathbb{E}X^2 + b\mathbb{E}X - a(\mathbb{E}X)^2 - b\mathbb{E}X = a \text{Var } X. \end{aligned}$$

a můžeme psát

$$|\text{Corr}(X, Y)| = \frac{|\text{Cov}(X, Y)|}{\sqrt{\text{Var } X \text{Var } Y}} = \frac{|a \text{Var } X|}{\sqrt{\text{Var } X a^2 \text{Var } X}} = 1.$$

Nakonec, k důkazu poslední implikace si uvědomíme, že rovnost nastává v případě  $|\text{Cor}(X, Y)| = \sqrt{\text{Var } X \text{Var } Y}$ . To nastane právě tehdy, když

$$[\mathbb{E}(X - \mathbb{E}X)(Y - \mathbb{E}Y)]^2 = [\mathbb{E}(X - \mathbb{E}X)^2][\mathbb{E}(Y - \mathbb{E}Y)^2].$$

Položme  $\tilde{X} = X - \mathbb{E}X$  a  $\tilde{Y} = Y - \mathbb{E}Y$ . Předchozí výraz pak bude mít tvar

$$[\mathbb{E}\tilde{X}\tilde{Y}]^2 = \mathbb{E}\tilde{X}^2\mathbb{E}\tilde{Y}^2.$$

Dosadíme  $a = \frac{\mathbb{E}\tilde{X}\tilde{Y}}{\mathbb{E}\tilde{X}^2}$  do  $g(a)$  z důkazu vlastnosti (ii), dostáváme (všimněte si, že jde o bod, kde funkce  $g$  nabývá svého minima)  $0 = \mathbb{E} \left[ \frac{\mathbb{E}\tilde{X}\tilde{Y}}{\mathbb{E}\tilde{X}^2} \tilde{X} - \tilde{Y} \right]^2$ , a tedy musí platit  $P[a\tilde{X} - \tilde{Y} = 0] = 1$ . Pak s pravděpodobností 1 musí platit  $aX - a\mathbb{E}X + \mathbb{E}Y = Y$ , což jsme chtěli dokázat (stačí vzít  $b = -a\mathbb{E}X - \mathbb{E}Y$ ).  $\square$

Jednoduchým důsledkem tohoto tvrzení (plyne z důkazu poslední vlastnosti z Věty 3.13) je následující tvrzení umožňující počítat rozptyl součtu ne nutně nezávislých veličin.

**Důsledek 3.17** (Rozptyl součtu). *Pokud  $X_1, \dots, X_d \in \mathcal{L}^2$  a  $a_1, \dots, a_d$  jsou reálné konstanty, potom*

$$\text{Var} \left( \sum_{l=1}^d a_l X_l \right) = \sum_{l=1}^d a_l^2 \text{Var} X_l + 2 \sum_{1 \leq j < l \leq d} a_j a_l \text{Cov}(X_j, X_l).$$

Pro vícerozměrné náhodné vektory můžeme definovat obdobné pojmy jako pro náhodné veličiny.

**Definice 3.18.** *Střední hodnotu náhodného vektoru  $\vec{X} = [X_1, \dots, X_d]^T$  definujeme předpisem*

$$\mathbb{E}\vec{X} = [\mathbb{E}X_1, \dots, \mathbb{E}X_d]^T.$$

*Varianční-kovarianční matice náhodného vektoru  $\vec{X} = [X_1, \dots, X_d]^T$  je definována jako*

$$\text{Var} \vec{X} = \begin{bmatrix} \text{Var} X_1 & \text{Cov}(X_1, X_2) & \cdots & \text{Cov}(X_1, X_d) \\ \text{Cov}(X_2, X_1) & \text{Var} X_2 & \cdots & \text{Cov}(X_2, X_d) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_d, X_1) & \text{Cov}(X_d, X_2) & \cdots & \text{Var}(X_d) \end{bmatrix}.$$

Všimneme si, že platí  $\text{Cov}(X, X) = \text{Var} X$  a  $\text{Cov}(X, Y) = \text{Cov}(Y, X)$ . Z toho plyne, že takto definovaná kovarianční matice je symetrická a navíc  $\text{Var} \vec{X} [\text{Cov}(X_i, X_j)]_{i,j=1\dots d}$ .

*konec 9. přednášky (17.3.2025)*

Budeme pokračovat základními vlastnostmi variančních matic, které se chovají podobně rozptylu jednorozměrné náhodné veličiny.

**Věta 3.19** (Vlastnosti varianční matice). *Máme-li náhodný vektor  $\vec{X}$  a reálné vektory  $\vec{a}, \vec{b}$  takové, že následující výrazy mají smysl, potom*

$$\mathbb{E}(\vec{a}^T \vec{X} + \vec{b}) = \vec{a}^T \mathbb{E}\vec{X} + \vec{b}, \text{Var}(\vec{a}^T \vec{X} + \vec{b}) = \vec{a}^T (\text{Var} \vec{X}) \vec{a}.$$

*Máme-li náhodný vektor  $\vec{X}$  a  $\vec{A}, \vec{B}$  jsou reálné matice, pak*

$$\mathbb{E}(A\vec{X} + B) = A\mathbb{E}\vec{X} + B, \text{Var}(A\vec{X} + B) = A(\text{Var} \vec{X})A^T.$$

*Důkaz.* Z definice násobení matic a linearity operátoru  $\mathbb{E}$ .  $\square$

**Definice 3.20.** Pro danou náhodnou veličinu  $X$  definujeme *momentovou vytvořující funkci* (MGF) vztahem

$$\psi_X(t) = \mathbb{E}[\exp(tX)] = \int_{\mathbb{R}} e^{tx} dP_X(x)$$

pro  $t \in \mathbb{R}$ , pokud pravá strana existuje. Speciální případ  $\psi_X(-t)$  nazýváme *Laplaceovou transformací*  $X$ .

**Věta 3.21** (Vlastnosti MGF). *Platí následující vlastnosti MGF:*

- (i) *Existuje-li  $\varepsilon > 0$  takové, že na  $(-\varepsilon, \varepsilon)$  existuje  $\psi_X(t)$ , potom  $\psi_X^{(m)}(0) = \mathbb{E}X^m$ ,  $m \in \mathbb{N}_0$ ;*
- (ii) *Pokud  $Y = aX + b$ , pak  $\psi_Y(t) = e^{bt}\psi_X(at)$ ;*
- (iii) *Pokud  $X_1, \dots, X_d$  jsou nezávislé a  $Y = \sum_{l=1}^d X_l$ , pak platí  $\psi_Y(t) = \prod_{l=1}^d \psi_{X_l}(t)$ .*

*Důkaz.* Dokážeme první vlastnost. Příklad  $m = 0$  je triviální, nechť tedy máme  $m > 0$ . Nejdříve budeme uvažovat případ  $m = 1$  a chceme použít větu o konvergentní majorantě. Nechť tedy

$$g(x) := e^{\frac{-\varepsilon x}{2}} + e^{\frac{\varepsilon x}{2}},$$

potom platí  $\exp tx \leq g(x)$  pro všechna  $t \in [-\varepsilon/2, \varepsilon/2]$  a libovolné  $x \in \mathbb{R}$ . Dále z předpokladu máme, že

$$\int_{\mathbb{R}} g(x) dP_X(x) = \psi_X(-\varepsilon/2) + \psi_X(\varepsilon/2) < +\infty.$$

Dostáváme, že  $g$  je hledaná konvergentní majoranta. Z věty o konvergentní majorantě tedy můžeme provést záměnu integrálu a derivace.

$$\frac{d}{dt}\psi(t) = \frac{d}{dt} \int_{\mathbb{R}} e^{tx} dP_X(x) = \int_{\mathbb{R}} x e^{tx} dP_X(x) \stackrel{t=0}{=} \int_{\mathbb{R}} x dP_X(x) = \mathbb{E}X^1.$$

Zbytek se dokáže indukcí s použitím podobné majoranty, v  $m$ -tém kroku dostaneme  $\int_{\mathbb{R}} x^m dP_X(x) =: \mathbb{E}X^m$ .

Druhou vlastnost dokážeme přímým rozepsáním definice

$$\begin{aligned} \psi_Y(t) &= \psi_{aX+b}(t) = \mathbb{E}[\exp(taX + tb)] = \mathbb{E}[\exp\{atX\}e^{tb}] = \\ &= e^{tb}\mathbb{E}[\exp(atX)] = e^{tb}\psi_X(at). \end{aligned}$$

Nakonec, poslední vlastnost se dokáže následně

$$\psi_Y(t) = \psi_{\sum_{l=1}^d X_l}(t) = \mathbb{E}[\exp\{t \sum_{l=1}^d X_l\}] = \mathbb{E}\left[\prod_{l=1}^d e^{tX_l}\right].$$



Dále využijeme nezávislost (která se přenáší i na veličiny transformované stejnou měřitelnou funkcí) a dostaneme

$$\mathbb{E} \left[ \prod_{l=1}^d e^{tX_l} \right] = \prod_{l=1}^d \mathbb{E}(e^{tX_l}) = \prod_{l=1}^d \psi_{X_l}(t).$$

□

Poznámka: pokud  $\psi_X(t) = \psi_Y(t)$  pro všechna  $t$  v nějakém otevřeném intervalu kolem 0, pak  $X$  a  $Y$  se rovnají v distribuci.

**Definice 3.22.** Pro danou náhodnou veličinu  $X$  definujeme *charakteristickou funkci* (CF) vztahem

$$\varphi_X(t) = \mathbb{E}[\exp(itX)] = \int_{\mathbb{R}} e^{itx} dP_X(x)$$

pro  $t \in \mathbb{R}$ .

Na rozdíl od momentové vytvořující funkce takto definovaná charakteristická funkce je dobře definovaná pro všechna  $t \in \mathbb{R}$ . Opět máme speciální název pro vyhodnocení charakteristické funkce v bodě  $-t$ , říkáme tomu *Fourierova transformace*. Z definice exponenciály z komplexní analýzy okamžitě dostáváme vyjádření  $\phi_X(t) = \mathbb{E} \cos(tX) + i\mathbb{E} \sin(tX)$ .

**Věta 3.23** (Vlastnosti CF). *Platí následující vlastnosti CF:*

- (i)  $\varphi_X$  existuje pro jakékoli rozdělení  $X$ ;
- (ii)  $\varphi_X(0) = 1$ ;
- (iii)  $|\varphi_X(t)| \leq 1$  pro všechna  $t \in \mathbb{R}$ ;
- (iv)  $\varphi_X$  je stejnoměrně spojitá, tedy  $\forall \varepsilon > 0 \exists \delta > 0 : |\varphi_X(t) - \varphi_X(s)| \leq \varepsilon$  kdykoli  $|t - s| \leq \delta$ ;
- (v)  $\varphi_{aX+b}(t) = e^{ibt} \varphi_X(at)$  pro všechna  $t, a, b \in \mathbb{R}$ ;
- (vi)  $\varphi_{-X}(t) = \bar{\varphi}_X(t)$  (komplexně sdružená funkce);
- (vii)  $\varphi_X(t) \in \mathbb{R} \forall t \in \mathbb{R}$  právě tehdy když rozdělení je symetrické kolem bodu  $t = 0$ .
- (viii) Jsou-li  $X, Y$  nezávislé, potom  $\varphi_{X+Y}(t) = \varphi_X(t) \varphi_Y(t)$ .

*Důkaz.* Budeme dokazovat vlastnosti postupně

- (i) Víme, že pro všechna  $x$  a všechna  $t$  platí  $|e^{itx}|^2 = \sin^2(tx) + \cos^2(tx) = 1$ . Pak  $\mathbb{E}|e^{itx}|^2 = 1$ , z Jensenovy nerovnosti (bude později) máme, že  $\mathbb{E}|e^{itx}| \leq \sqrt{\mathbb{E}|e^{itx}|^2} = 1$  a tedy  $e^{itx}$  je integrovatelná.
- (ii) Přímým dosazením dostaneme  $\int_{\mathbb{R}} dP_X(x) = 1$ .

(iii) Viz důkaz vlastnosti (i).

(iv) Položme  $h := s - t$ , potom

$$\begin{aligned} |\varphi_X(t) - \varphi_X(s)| &= |\mathbb{E}[e^{itX}] - \mathbb{E}[e^{i(t+h)X}]| \leq \\ &\mathbb{E}[|e^{itX}(1 - e^{ihX})|] \leq \\ &\mathbb{E}[|e^{itX}| \cdot |1 - e^{ihX}|] \leq \mathbb{E}[|e^{ihX} - 1|]. \end{aligned}$$

Víme, že  $e^{ihX} - 1 \rightarrow 0$  když  $h \rightarrow 0$  a zároveň  $|e^{ihX} - 1| \leq 2$ . Máme tedy konvergentní majorantu. Dle Lebesgueovy věty tedy platí  $\lim_{h \rightarrow 0} \mathbb{E}[e^{ihX} - 1] = 0$ . Z toho již plyne stejnoměrná spojitost.

(v) Z definice dostáváme

$$\varphi_{aX+b}(t) = \mathbb{E}[e^{it(aX+b)}] = e^{ibt} \mathbb{E}[e^{itaX}] = e^{ibt} \varphi_X(at).$$

(vi) Využijeme přepisu do goniometrického tvaru (viz poznámka před touto větou), dostaneme

$$\varphi_{-X}(t) = \mathbb{E}[\cos(-tX) + i \sin(-tX)] = \mathbb{E}[\cos(tX) - i \sin(tX)] = \bar{\varphi}_X(t).$$

(vii) Necht' nejdříve  $X$  má rozdělení symetrické kolem 0, potom  $X \stackrel{d}{=} -X$ , z čehož máme  $\varphi_{-X}(t) = \varphi_X(t)$ . Aplikací již dokázané vlastnosti (v) máme, že  $\varphi_X = \bar{\varphi}_X(t)$ , tedy  $\varphi_X(t) \in \mathbb{R}$ . Opačná implikace se dokáže stejným postupem v opačném pořadí.

(viii) Rozepsání definice

$$\varphi_{X+Y}(t) = \mathbb{E}[e^{it(X+Y)}] = \mathbb{E}[e^{itX} e^{itY}],$$

dále díky nezávislosti dostáváme

$$\mathbb{E}[e^{itX} e^{itY}] = \mathbb{E}[e^{itX}] \mathbb{E}[e^{itY}] = \varphi_X(t) \varphi_Y(t).$$

□

Následující věta nám umožňuje jednoznačně popisovat rozdělení jak podle distribuční funkce, tak i podle charakteristické funkce.

**Věta 3.24** (Leviho inverzní formule pro charakteristickou funkci). *Pro jakékoli rozdělení  $X$  a libovolné  $a < b$  platí*

$$\lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-T}^T \frac{e^{-ita} - e^{-itb}}{it} \varphi_X(t) dt = P[a < X < b] + \frac{P[X = a] + P[X = b]}{2}.$$

*Důkaz.* Mějme  $T \in \mathbb{R}$  a  $a < b$ , potom

$$\begin{aligned} \frac{1}{2\pi} \int_{-T}^T \frac{e^{ita} - e^{itb}}{it} \varphi_X(t) dt &= \frac{1}{2\pi} \int_{-T}^T \frac{e^{ita} - e^{itb}}{it} \int_{\mathbb{R}} e^{itx} dP_X = \\ \int_{-T}^T \int_{-\infty}^{\infty} \frac{e^{it(x-a)} - e^{it(x-b)}}{2\pi it} dP_X dt &\stackrel{\text{Fubini}}{=} \int_{-\infty}^{\infty} \int_{-T}^T \frac{e^{it(x-a)} - e^{it(x-b)}}{2\pi it} dt dP_X. \end{aligned}$$

Všimneme si, že pro každou konstantu  $c \in \mathbb{R}$  platí  $\int_{-T}^T \frac{e^{itc}}{2it} dt = \int_0^T \frac{\sin(tc)}{t} dt$  a tento integrál se navíc rovná  $\frac{\pi}{2} \operatorname{sgn}(c)$ . Potom platí

$$\begin{aligned} \frac{1}{2\pi} \int_{-T}^T \frac{e^{ita} - e^{itb}}{it} \varphi_X(t) dt &= \\ \int_{-\infty}^{\infty} \frac{1}{\pi} \left[ \int_0^T \frac{\sin(t(x-a))}{t} dt - \int_0^T \frac{\sin(t(x-b))}{t} dt \right] dP_X. \end{aligned}$$

Když pošleme  $T$  do nekonečna, dostaneme následující hodnoty

$$\frac{1}{\pi} \int_0^T \frac{\sin(t(x-a))}{t} dt \xrightarrow{T \rightarrow \infty} \begin{cases} -\frac{1}{2}, & x < a, \\ \frac{1}{2}, & x > a, \\ 0, & x = a. \end{cases}$$

Potom

$$\frac{1}{\pi} \left[ \int_0^T \frac{\sin(t(x-a))}{t} dt - \int_0^T \frac{\sin(t(x-b))}{t} dt \right] \xrightarrow{T \rightarrow \infty} \begin{cases} \frac{1}{2}, & x = a, b \\ 1, & a < x < b, \\ 0, & \text{jinak.} \end{cases}$$

Dosazením do předchozího vzorce a užitím Lebesgueovy věty o konvergentní majorantě dostáváme

$$\frac{1}{2\pi} \int_{-T}^T \frac{e^{ita} - e^{itb}}{it} \varphi_X(t) dt = P[a < X < b] + \frac{P[X = a] + P[X = b]}{2}.$$

□

Z předchozí věty okamžitě plyne následující důsledek.

**Důsledek 3.25** (Jednoznačná charakterizace rozdělení). *Platí  $\varphi_X = \varphi_Y \Leftrightarrow X \stackrel{d}{=} Y$ .*

Nakonec definujeme charakteristickou funkci pro náhodné vektory. Obdobným způsobem pro ní můžeme dokázat vlastnosti, které jsme již dokázali pro jednorozměrné náhodné veličiny.

**Definice 3.26.** *Charakteristická funkce* náhodného vektoru  $\vec{X} = [X_1, \dots, X_d]^T$  je definována vztahem

$$\varphi_{\vec{X}}^{\vec{t}} = \mathbb{E}[e^{it^T \vec{X}}] = \int_{\mathbb{R}} e^{it^T \vec{X}} dP_{\vec{X}}$$

pro  $\vec{t} \in \mathbb{R}^d$ .

*konec 10. přednášky (18.3.2025)*

## 4 Stochastické nerovnosti

V této kapitole budeme studovat užitečné nerovnosti, které budeme moci aplikovat pro odhady některých statistických veličin. Začneme odhady pro hodnoty pravděpodobnosti samotné (tzv. nerovnosti Markovovského typu)

**Věta 4.1** (Markovova nerovnost). *Nechť  $X$  je nezáporná náhodná veličina a předpokládejme, že  $\mathbb{E}X$  existuje. Potom pro každé  $\varepsilon > 0$  platí*

$$P[X \geq \varepsilon] \leq \frac{\mathbb{E}X}{\varepsilon}.$$

*Důkaz.* Z předpokladu máme, že  $X \geq 0$ , tedy  $P[X \geq 0] = 1$  a  $P[X < 0] = 1 - 1 = 0$ . Z toho plyne, že

$$\int_{\{\omega \in \Omega: X(\omega) < 0\}} dP(\omega) = 0.$$

Potom pro  $\varepsilon > 0$  máme, že pravděpodobnost

$$\begin{aligned} P[X \geq \varepsilon] &= \int_{\{\omega \in \Omega: X(\omega) \geq \varepsilon\}} dP(\omega) \leq \int_{\{\omega \in \Omega: X(\omega) \geq \varepsilon\}} \frac{X(\omega)}{\varepsilon} dP(\omega) \leq \\ &\int_{\Omega} \frac{X(\omega)}{\varepsilon} dP(\omega) = \frac{1}{\varepsilon} \int_{\Omega} X(\omega) dP(\omega) = \frac{\mathbb{E}X}{\varepsilon}, \end{aligned}$$

kde první nerovnost platí díky tomu, že  $X(\omega) > \varepsilon$  a tedy  $\frac{X(\omega)}{\varepsilon} > 1$ .  $\square$

**Důsledek 4.2** (Zobecněná Markovova nerovnost). *Nechť  $X$  je nezáporná náhodná veličina a předpokládejme, že  $\mathbb{E}X^r$  existuje pro nějaké  $r > 0$ . Potom pro každé  $\varepsilon > 0$  platí*

$$P[X \geq \varepsilon] \leq \frac{\mathbb{E}X^r}{\varepsilon^r}.$$

*Důkaz.* Nechť  $Y := X^r$ ,  $\tilde{\varepsilon} = \varepsilon^r$ , poté použijeme předchozí větu pro  $P[Y \leq \tilde{\varepsilon}]$ .  $\square$

**Věta 4.3** (Čebyševova nerovnost). *Nechť  $X$  je náhodná veličina a předpokládejme, že  $\mathbb{E}[X]$  existuje, potom pro každé  $\varepsilon > 0$  platí*

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

*Důkaz.* Položme  $Y = |X - \mathbb{E}X| \geq 0$  a  $\tilde{\varepsilon} = \varepsilon^2$ , potom stačí aplikovat Větu 4.1.  $\square$

Dále si uvedeme několik nerovností, které přímo poskytují odhad pro střední hodnotu.

**Věta 4.4** (Cauchy-Schwarzova nerovnost). *Pokud mají náhodné veličiny  $X$  a  $Y$  konečné rozptyly, potom*

$$|\mathbb{E}XY| \leq \sqrt{\mathbb{E}X^2 \mathbb{E}Y^2} \text{ a } |\text{Cov}(X, Y)| \leq \sqrt{\text{Var } X \text{Var } Y}.$$

*Důkaz.* Plyne z důkazu vlastnosti (ii) ve Větě 3.16 o vlastnostech kovariance a korelace.  $\square$

**Věta 4.5** (Jensenova nerovnost). *Pokud je  $g$  konvexní, pak  $\mathbb{E}g(X) \geq g(\mathbb{E}X)$ . Dále, pokud je  $g$  konkávní  $\mathbb{E}g(X) \leq g(\mathbb{E}X)$ .*

*Důkaz.* Necht máme  $t(x) = a + tx$  tečna k funkci  $g$  v bodě  $\mathbb{E}X$ . Pokud  $g$  je konvexní, pak  $t(x) \leq g(x)$  pro všechna  $x \in \mathbb{R}$ . Také platí  $t(\mathbb{E}X) = g(\mathbb{E}X)$ . Potom  $E[g(X)] \geq E[t(X)] = E[a + bX] = a + b\mathbb{E}X = t(\mathbb{E}X) = g(\mathbb{E}X)$ . Pro konkávní  $g$  se tvrzení dokáže analogicky.  $\square$

## 5 Stochastické konvergence

V této kapitole budeme studovat druhy konvergence v pravděpodobnostních prostorech, které jsou často jiné, neboť náš prostor je vždy normovaný na 1.

**Definice 5.1.** Necht  $X_1, X_2, \dots$  je posloupnost náhodných veličin a necht  $X$  je jiná náhodná veličina. Necht  $F_n$  označuje distribuční funkci  $X_n$  a necht  $F$  označuje distribuční funkci  $X$ . Potom  $X_n$  konverguje k  $X$  v pravděpodobnosti (předpokládáme, že  $X_i, X$  všechny “žijí” na stejném pravděpodobnostním prostoru), značíme  $X_n \xrightarrow[n \rightarrow \infty]{P}$ , pokud pro každé  $\varepsilon > 0$ ,

$$P[|X_n - X| > \varepsilon] \xrightarrow{n \rightarrow \infty} 0.$$

Dále  $X_n$  konverguje k  $X$  v distribuci, značíme  $X_n \xrightarrow[n \rightarrow \infty]{D} X$ , pokud

$$\lim_{n \rightarrow \infty} F_n(x) = F(x)$$

pro všechna  $x$  kde je  $F$  spojitá.

$X_n$  konverguje k  $X$  v  $L_p$  pro  $p \geq 1$ , značíme  $X_n \xrightarrow[n \rightarrow \infty]{L_p} X$ , pokud

$$\mathbb{E}|X_n - X|^p \xrightarrow{n \rightarrow \infty} 0.$$

$X_n$  konverguje k  $X$  skoro jistě, značíme  $X_n \xrightarrow[n \rightarrow \infty]{P-s.j.} X$ , pokud

$$P[\lim_{n \rightarrow \infty} X_n = X] \equiv P[\omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)] = 1.$$

**Věta 5.2** (Implikace mezi typy konvergence). *Platí následující implikace*

$$(i) \quad X_n \xrightarrow[n \rightarrow \infty]{P-s.j.} X \implies X_n \xrightarrow[n \rightarrow \infty]{P} X;$$

$$(ii) \quad \text{pro } p \geq 1 \text{ platí } X_n \xrightarrow[n \rightarrow \infty]{L_p} X \implies X_n \xrightarrow[n \rightarrow \infty]{P} X;$$

$$(iii) \quad \text{pro } p \geq q \geq 1 \text{ platí } X_n \xrightarrow[n \rightarrow \infty]{L_p} X \implies X_n \xrightarrow[n \rightarrow \infty]{L_q} X;$$

$$(iv) \quad X_n \xrightarrow[n \rightarrow \infty]{P} X \implies X_n \xrightarrow[n \rightarrow \infty]{D} X;$$

$$(v) \quad \text{Pokud } X_n \xrightarrow[n \rightarrow \infty]{D} X \text{ a } P[X = c] = 1 \text{ pro nějaké } c \in \mathbb{R}, \text{ pak } X_n \xrightarrow[n \rightarrow \infty]{P} X.$$

konec 11. přednášky (24.3.2025)

*Důkaz.* Budeme dokazovat postupně každou implikaci.

(i) Mějme  $\varepsilon > 0$ . Pro  $n \in \mathbb{N}$  definujeme náhodné události

$$A_n := \{\omega \in \Omega : \exists m \geq n : |X_m(\omega) - X(\omega)| \geq \varepsilon\};$$

$$B_n := \{\omega \in \Omega : |X_n(\omega) - X(\omega)| \geq \varepsilon\}.$$

Chceme ukázat, že  $P(B_n) \rightarrow 0$ . Víme, že  $A_n \supseteq B_n$ , tedy díky monotonii pravděpodobnosti stačí ukázat, že  $P(A_n) \rightarrow 0$ . Reálná posloupnost  $\{X_m(\omega)\}_m$  je konvergentní, jestliže existuje přirozené číslo  $N$  takové, že pro žádné  $m \geq N$  neplatí  $|X_m(\omega) - X(\omega)| \geq \varepsilon$ . Potom  $\lim A_n$  je jev, že posloupnost  $\{X_m(\omega)\}_m$  diverguje. Ale dle předpokladu  $X_n$  konverguje skoro jistě, tedy  $P(\lim A_n) = 0$ . Jelikož  $A_1 \supset A_2 \supset \dots$ , z věty o spojitosti míry (Věta 1.7) dostáváme  $P(\lim A_n) = \lim P(A_n) = 0$ , čímž jsme dostali požadovanou konvergenci v pravděpodobnosti.

(ii) Nechť  $X_n \xrightarrow{L_p} X$ . Podle Markovovy nerovnosti (Věta 4.1) platí

$$P(|X_n - X| \geq \varepsilon) = P(|X_n - X|^p \geq \varepsilon^p) \leq \frac{\mathbb{E}|X_n - X|^p}{\varepsilon^p} \rightarrow 0$$

pro  $n \rightarrow \infty$ , neboť se jedná o posloupnost čísel a chování čitatele vyplývá z předpokladu konvergence v  $L_p$ .

(iii) Nechť  $X_n \xrightarrow{L_p} X$  a  $p \geq q \geq 1$ . Dle Jensenovy nerovnosti pro konvexní funkci (Věta 4.5)  $g(x) := x^{p/q}$  pro  $x \geq 0$  dostáváme  $g(\mathbb{E}[|X_n - X|^q]) \leq \mathbb{E}[|X_n - X|^{q \cdot \frac{p}{q}}] = \mathbb{E}|X_n - X|^p \rightarrow 0$ , kde limitní přechod plyne z předpokladu konvergence v  $L_p$ , tedy jsme přímo ukázali konvergenci v  $L_q$ .

(iv) Nechť  $X_n \xrightarrow{P} X$ . Zvolme  $\varepsilon > 0$  libovolně. Nechť  $x \in \mathbb{R}$  je libovolný bod, v němž je limitní distribuční funkce  $F$  spojitá. Potom můžeme psát

$$\begin{aligned} F_n(x) &= P(X_n \leq x) = P(X_n \leq x, X \leq x + \varepsilon) + P(X_n \leq x, X > x + \varepsilon) \leq \\ &= P(X \leq x + \varepsilon) + P(|X_n - X| > \varepsilon) = F(x + \varepsilon) + P(|X_n - X| > \varepsilon). \end{aligned}$$

Taktéž dostaneme

$$\begin{aligned} F(x - \varepsilon) &= P(X \leq x - \varepsilon) = P(X \leq x - \varepsilon, X_n \leq X) + \\ &+ P(X \leq x - \varepsilon, X_n > \varepsilon) \leq F_n(x) + P(|X_n - X| > \varepsilon). \end{aligned}$$

Potom  $F(x - \varepsilon) - P(|X_n - X| > \varepsilon) \leq F_n(x) \leq F(x + \varepsilon) + P(|X_n - X| > \varepsilon)$ . Jelikož  $\varepsilon > 0$  bylo voleno libovolně a  $F$  je spojitá v  $x$ , tedy  $\lim_{n \rightarrow \infty} F_n(x) = F(x)$ .

(v) Nechť  $X_n \xrightarrow{D} X$  a nechť  $P[X = c] = 1$  pro nějaké  $c \in \mathbb{R}$ . Zvolme libovolné  $\varepsilon > 0$  a můžeme počítat

$$\begin{aligned} P(|X_n - X| \geq \varepsilon) &= P(X_n \leq c - \varepsilon) + P(X_n \geq c + \varepsilon) \leq \\ &= P(X_n \leq c - \varepsilon) + P(X_n > c + \frac{\varepsilon}{2}) \leq \\ &= F_n(c - \varepsilon) + 1 - F_n(c + \frac{\varepsilon}{2}) \xrightarrow{n \rightarrow \infty} 0 + 1 - 1 = 0, \end{aligned}$$

čímž jsme dokončili důkaz této věty.

□

Uvedeme si několik protipříkladů, na kterých si ukážeme, že implikace opačné k právě uvedeným nemusí platit.

**Příklad 5.3.** Ukážeme, že konvergence v pravděpodobnosti neimplikuje konvergenci skoro jistě. Mějme prostor  $(\Omega = [0, 1], \mathcal{A} = \mathcal{B}(\Omega), P = \lambda)$ . Každé přirozené číslo můžeme jednoznačně zapsat ve tvaru  $2^n + m$ , kde  $m \in \{0, 1, \dots, 2^n - 1\}$  a definovat

$$X_{2^n+m}(\omega) = \chi_{\{\omega \in (m2^{-n}, (m+1)2^{-n}]\}}, \omega \in [0, 1].$$

Například, protože  $33 = 2^5 + 1$ , dostaneme  $X_{33}(\omega) = \chi_{\{\omega \in (2^{-5}, 2^{-4}]\}}$ . Pak pro každé  $\varepsilon \in (0, 1)$  dostaneme  $P[|X_{2^n+m}| > \varepsilon] = 2^{-n} \rightarrow 0$  pro  $n \rightarrow \infty$ . Tedy  $X_n \xrightarrow{P} 0$ . Avšak pro každé  $\omega \in (0, 1]$ ,  $X_j(\omega) = 1$  a  $X_j(\omega) = 0$  pro nekonečně mnoho různých  $j$  a tedy posloupnost  $X_n$  nekonverguje skoro jistě.

**Příklad 5.4.** Ukážeme, že konvergence v pravděpodobnosti neimplikuje konvergenci v  $L_p$ . Mějme prostor  $(\Omega = [0, 1], \mathcal{A} = \mathcal{B}(\Omega), P = \lambda)$ . Každé přirozené číslo můžeme jednoznačně zapsat ve tvaru  $2^n + m$ , kde  $m \in \{0, 1, \dots, 2^n - 1\}$  a definovat

$$X_{2^n+m}(\omega) = 2^n \chi_{\{\omega \in ((m-1)2^{-n}, m2^{-n}]\}}, \omega \in [0, 1].$$

Pak opět pro každé  $\varepsilon \in (0, 1)$  dostaneme  $P[|X_{2^n+m}| > \varepsilon] = 2^{-n} \rightarrow 0$ . Tedy  $X_n \xrightarrow{P} 0$ . Nicméně,  $\mathbb{E}|X_{2^n+m} - 0| = 2^n P[X_{2^n+m} = 2^n] = 2^n 2^{-n} = 1$  a tedy posloupnost nekonverguje v  $L_1$ , tedy to nemůže konvergovat ani v vyšších  $L_p, p > 1$ .

**Příklad 5.5.** Ukážeme, že konvergence v  $L_q$  neimplikuje konvergenci v  $L_p$  pro  $p > q \geq 1$ . Mějme prostor  $(\Omega = [0, 1], \mathcal{A} = \mathcal{B}(\Omega), P = \lambda)$ . Každé přirozené číslo můžeme jednoznačně zapsat ve tvaru  $2^n + m$ , kde  $m \in \{0, 1, \dots, 2^n - 1\}$  a definovat

$$X_{2^n+m}(\omega) = 2^{n/2} \chi_{\{\omega \in ((m-1)2^{-n}, m2^{-n}]\}}, \omega \in [0, 1].$$

Pak  $\mathbb{E}|X_{2^n+m} - 0| = 2^{n/2} P[X_{2^n+m} = 2^{n/2}] = 2^{n/2} 2^{-n} = 2^{-n/2}$  a tedy posloupnost konverguje v  $L_1$ . Nicméně, pro  $p = 2$  máme  $\mathbb{E}|X_{2^n+m} - 0|^2 = 2^{2(n/2)} 2^{-n} = 1$  a tedy posloupnost nekonverguje v  $L_2$ .

**Příklad 5.6.** Ukážeme, že konvergence v distribuci neimplikuje konvergenci v pravděpodobnosti. Nechť  $X \sim N(0, 1)$  a  $X_n := -X, n \in \mathbb{N}$ . Tedy  $X_n \sim N(0, 1)$  pro každé  $n \in \mathbb{N}$ . Tedy triviálně  $\lim F_n(x) = F(x)$  pro všechna  $x \in \mathbb{R}$ . Tedy  $X_n \xrightarrow{D} X$ .

Nicméně,  $P[|X_n - X| > \varepsilon] = P[|2X| > \varepsilon] = P[|X| > \varepsilon/2] \neq 0$  (nezávislé na  $n$ ), tedy posloupnost  $X_n$  nekonverguje v pravděpodobnosti.

**Věta 5.7** (o spojitém zobrazení (CMT)). *Nechť  $\vec{X}, \vec{X}_1, \vec{X}_2, \dots$  jsou  $d$ -rozměrné náhodné vektory a  $g : \mathbb{R}^d \rightarrow \mathbb{R}^m$  je spojitá v každém bodě množiny  $C$  takové, že  $P[\vec{X} \in C] = 1$ . Potom platí*



- $\vec{X}_n \xrightarrow{P-s.j.} \vec{X} \implies g(\vec{X}_n) \xrightarrow{P-s.j.} g(\vec{X});$
- $\vec{X}_n \xrightarrow{P} \vec{X} \implies g(\vec{X}_n) \xrightarrow{P} g(\vec{X});$
- $\vec{X}_n \xrightarrow{D} \vec{X} \implies g(\vec{X}_n) \xrightarrow{D} g(\vec{X}).$

*Důkaz.* Dokážeme pouze první dvě vlastnosti. Nejdříve, nechť  $X_n$  konverguje skoro jistě, potom ze spojitosti  $g$  máme, že vlastnost  $\lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)$  implikuje  $\lim_{n \rightarrow \infty} g(X_n(\omega)) = g(X(\omega))$ . Potom  $P[\omega \in \Omega : \lim_{n \rightarrow \infty} g(X_n(\omega)) = g(X(\omega))] = P[\omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)] = 1$ , přičemž poslední rovnost plyne z definice konvergence skoro jistě.

K důkazu druhé vlastnosti zvolme  $\varepsilon > 0$ . Potom uvažujeme pro libovolné  $\delta > 0$  množinu

$$B_\delta := \{x \in \mathbb{R} : \exists y \in \mathbb{R} : |x - y| < \delta \wedge |g(x) - g(y)| \geq \varepsilon\}.$$

*konec 12. přednášky (25.3.2025)*

Zřejmě  $B_\delta \rightarrow \emptyset$  pro  $\delta \rightarrow 0^+$ . Potom můžeme psát

$$P(|g(\vec{X}_n) - g(\vec{X})| \geq \varepsilon) = P(|g(\vec{X}_n) - g(\vec{X})| \geq \varepsilon \cap |\vec{X}_n - \vec{X}| \geq \delta) +$$

$$P(|g(\vec{X}_n) - g(\vec{X})| \geq \varepsilon \cap |\vec{X}_n - \vec{X}| < \delta) \leq P(|\vec{X}_n - \vec{X}| \geq \delta) + P(\vec{X} \in B_\delta).$$

Jelikož  $\delta$  bylo voleno libovolně, platí  $P(\vec{X} \in B_\delta) \xrightarrow{\delta \rightarrow 0} 0$  a  $P(|\vec{X}_n - \vec{X}| \geq \delta) \xrightarrow{n \rightarrow \infty} 0$ , čímž jsme dokázali konvergenci v pravděpodobnosti.  $\square$

Poznámka: z  $\vec{X}_n \xrightarrow{L_p} \vec{X}$  nutně neplyne  $g(\vec{X}_n) \xrightarrow{L_p} g(\vec{X})$ .

Dalším významným tvrzením teorie pravděpodobnosti je takzvaná Slutského věta (v anglické literatuře se také používá název Cramer-Slutského věta).

**Věta 5.8** (Slutského věta). *Pokud  $X_n \xrightarrow{D} X$  a  $Y_n \xrightarrow{P} c \in \mathbb{R}$ , pak  $X_n + Y_n \xrightarrow{D} X + c$  a  $X_n Y_n \xrightarrow{D} cX$ .*

*Důkaz.* Dokážeme pouze výrok pro součet, část pro součin se dokáže analogicky.

Mějme  $x \in \mathbb{R}$  bod, v němž je spojitá distribuční funkce veličiny  $X + c$ . Potom  $x - c$  je nutně bodem spojitosti distribuční funkce  $F_X$ . Zvolme  $\eta > 0$ . Potom existuje  $\varepsilon_0 > 0$  takové, že  $|F_X(x - c) - F_X(x - c - \varepsilon)| < \frac{\eta}{3}$  pro každé  $|\varepsilon| < \varepsilon_0$ .

Jelikož  $F_X$  je distribuční funkce, má nejvýše spočetně mnoho bodů nespojitosti. Dokážeme tuto vlastnost. Nechť tedy  $D$  je množina bodů nespojitosti  $F_X$ . Pak  $\forall y \in D : F_X(y_-) < F_X(y^+)$ , tedy existuje racionální číslo  $q_y \in \mathbb{Q}$  takové, že  $F_X(y_-) < q_y < F_X(y^+)$ . Jelikož  $F_X$  je neklesající, pak  $y \neq z \in D \implies q_y \neq q_z$ . Tedy zobrazení  $y \mapsto q_y$  je prosté.

Z toho máme, že v každém okolí bodu  $x - c$  můžeme nalézt vlevo i vpravo od  $x - c$  nějaký bod, v němž je  $F_X$  spojitá. Z definice spojitosti existuje  $0 < \varepsilon < \varepsilon_0$  takové, že  $F_X$  je spojitá v  $x - c + \varepsilon$  i v  $x - c - \varepsilon$ . Potom

$$P(X_n + Y_n \leq x) = P(X_n + Y_n \leq x \cap |Y_n - c| < \varepsilon) + P(X_n + Y_n \leq x \cap |Y_n - c| \geq \varepsilon) \leq$$

$$P(X_n + c \leq x + \varepsilon \cap |Y_n - c| < \varepsilon) + P(|Y_n - c| \geq \varepsilon) \leq \\ P(X_n \leq x - c + \varepsilon) + P(|Y_n - c| \geq \varepsilon).$$

Jelikož  $F_X$  je spojitá v bodě  $x - c + \varepsilon$  a  $Y_n \xrightarrow{P} c$ , existuje  $n_1 \in \mathbb{N}$  takové, že pro všechna  $n \geq n_1$  platí  $P(X_n \leq x - c + \varepsilon) \leq P(X_n \leq x - c) + \frac{\eta}{3}$  a zároveň  $P(|Y_n - c| \geq \varepsilon) \leq \frac{\eta}{3}$ . Dohromady pro  $n \geq n_1$  máme

$$P(X_n + Y_n \leq x) \leq P(X \leq x - c + \varepsilon) + \frac{2}{3}\eta \leq P(X \leq x - c) + \eta.$$

Opačná nerovnost se dokáže analogicky. Pro  $n \geq \max\{n_1, n_2\}$  dohromady máme

$$P(X + c \leq x) - \eta \leq P(X_n + Y_n \leq x) \leq P(X + c \leq x) + \eta.$$

Jelikož  $\eta > 0$  bylo voleno libovolně, věta je dokázána.  $\square$

Z této věty okamžitě plyne následující důsledek (neboť konvergence v pravděpodobnosti implikuje konvergenci v distribuci, viz Věta 5.2, vlastnosti (iv) a (v)).

**Důsledek 5.9.** *Pokud  $x_n \xrightarrow{P} a \in \mathbb{R}$  a  $Y_n \xrightarrow{P} b \in \mathbb{R}$ , potom  $X_n + Y_n \xrightarrow{P} a + b$  a  $X_n Y_n \xrightarrow{P} ab$ .*

Ještě budeme potřebovat následující větu, která ekvivalentně charakterizuje konvergenci v distribuci, její důkaz je však aktuálně nad naše schopnosti.

**Věta 5.10** (Lévyho věta o spojitosti). *Platí  $\vec{X}_n \xrightarrow{D} \vec{X} \Leftrightarrow \varphi_{\vec{X}_n}(\vec{t}) \rightarrow \varphi_{\vec{X}}(\vec{t})$  pro všechna  $\vec{t} \in \mathbb{R}^d$ .*

Ukážeme si explicitní vyjádření charakteristické funkce normálního rozdělení.

**Příklad 5.11.** Necht  $X \sim N(\mu, \sigma^2)$ . Potom platí

$$\varphi_X(t) = \exp \left\{ i\mu t - \frac{1}{2}\sigma^2 t^2 \right\}.$$

*Důkaz.* Nejdříve uvažujme  $X \sim N(0, 1)$  standardní normální rozdělení. Potom

$$\varphi_X(t) = \mathbb{E}[\exp\{itx\}] = \mathbb{E}[\cos(tx)] + i\mathbb{E}[\sin(tx)] = \\ \int_{-\infty}^{\infty} \cos(tx) dP_X + i \int_{-\infty}^{\infty} \sin(tx) dP_X = \mathbb{E}[\cos(tX)],$$

kde poslední rovnost plyne z faktu, že  $\sin$  je lichá funkce, a tedy příslušný integrál je nulový. Zderivujeme  $\varphi_X(t)$  s použitím majoranty  $x$ .

$$\frac{d}{dt} \varphi_X(t) = - \int_{-\infty}^{\infty} x \sin(tx) dP_X = - \int_{-\infty}^{\infty} x \sin(tx) \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{x^2}{2} \right\} dx =$$

$$\stackrel{p.p.}{=} [x \sin(tx) f_X(x)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} t \cos(tx) f_X(x) dx = -t \mathbb{E}[\cos(tX)].$$

Dostali jsme, že  $\frac{d}{dt} \varphi_X(t) = -t \varphi_X(t)$ . Tato diferenciální rovnice s počáteční podmínkou  $\varphi_X(0) = \cos 0 = 1$  má jediné řešení  $\varphi_X(t) = \exp \left\{ -\frac{t^2}{2} \right\}$ .

Potom pro  $Y = \mu + \sigma X$  ( $Y \sim N(\mu, \sigma^2)$ ) dosadíme a aplikujeme výsledek pro standardní rozdělení, dostaneme

$$\begin{aligned} \varphi_Y(t) &= \mathbb{E}[\exp\{itY\}] = E[\exp\{it(\mu + \sigma X)\}] = \\ &= \exp\{it\mu\} \cdot \mathbb{E}[\exp\{i(t\sigma)X\}] = \exp \left\{ it\mu - \frac{1}{2} \sigma^2 t^2 \right\}. \end{aligned}$$

□

*konec 13. přednášky (31.3.2025)*

**Definice 5.12.** *Posloupností nezávislých náhodných veličin* rozumíme posloupnost náhodných veličin  $\{X_n\}_{n \in \mathbb{N}}$  takovou, že pro každou  $J \subset \mathbb{N}$ ,  $|J| < \infty$  a všechna  $\{x_j\}_{j \in J}$  platí

$$F_{\{X_j\}_{j \in J}}(\{x_j\}_{j \in J}) = \prod_{j \in J} F_{X_j}(x_j).$$

**Definice 5.13.**  $\{X_n\}_{n \in \mathbb{N}}$  je *posloupnost nezávislých stejně rozdělených náhodných veličin (IID)*, jestliže to je posloupnost nezávislých náhodných veličin majících stejnou distribuční funkci.

Tyto definice lze celkem snadno přepsat i pro náhodné vektory (cvičení).

**Věta 5.14** (Slabý zákon velkých čísel). *Jestliže  $\{X_n\}_{n \in \mathbb{N}}$  je IID posloupnost náhodných veličin a  $\mathbb{E}|X_1| < \infty$ , pak*

$$\bar{X}_n := \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow[n \rightarrow \infty]{P} \mathbb{E}X_1.$$

*Důkaz.* Rozvineme komplexní exponenciálu do Taylorova rozvoje do prvního řádu. Platí  $e^{ity} = 1 + ity + o(t)$  pro všechna  $y \in \mathbb{R}$ . Potom  $\mathbb{E}[e^{itY}] = 1 + it\mathbb{E}Y + o(t)$ . Dále můžeme počítat

$$\begin{aligned} \varphi_{\bar{X}_n}(t) &= \varphi_{\frac{1}{n} \sum X_i}(t) = \varphi_{\sum \frac{X_i}{n}}(t) = \left( \varphi_{\frac{X_1}{n}}(t) \right)^n = \left( \varphi_{X_1} \left( \frac{t}{n} \right) \right)^n = \\ &= \left( \mathbb{E}[e^{\frac{itX_1}{n}}] \right)^n = \left( 1 + \frac{it}{n} \mathbb{E}X_1 + o \left( \frac{t}{n} \right) \right)^n. \end{aligned}$$

Pro  $n \rightarrow \infty$  dále máme

$$\lim_{n \rightarrow \infty} \varphi_{\bar{X}_n}(t) = \lim_{n \rightarrow \infty} \left( 1 + \frac{it}{n} \mathbb{E}X_1 + o \left( \frac{t}{n} \right) \right)^n = \lim_{n \rightarrow \infty} (\mathbb{E}[e^{it\mathbb{E}X_1/n}])^n =$$

$$\lim_{n \rightarrow \infty} (e^{it\mathbb{E}X_1/n})^n = e^{it\mathbb{E}X_1} = \varphi_{\mathbb{E}X_1}(t).$$

Z Lévyho věty (Věta 5.10) tedy dostáváme konvergenci  $\bar{X}_n \xrightarrow{D} \mathbb{E}X_1$  v distribuci. Dále využijeme faktu, že jestliže posloupnost konverguje v distribuci ke konstantě, potom máme i konvergenci v pravděpodobnosti, čímž je důkaz ukončen.  $\square$

Jestliže konvergenci v pravděpodobnosti nahradíme konvergencí skoro jistě, dostaneme silný zákon velkých čísel. Jeho důkaz je však nad rámec tohoto kurzu. Kdybychom předpokládali, že  $\text{Var } X_1 < \infty$ , pak tvrzení plyne okamžitě z Čebyševovy nerovnosti (Věta 4.3).

**Věta 5.15** (Centrální limitní věta). *Pokud  $\{X_n\}_{n \in \mathbb{N}}$  je posloupnost IID náhodných veličin s  $\mathbb{E}X_1^2 < \infty$  a  $\text{Var } X_1 > 0$ , pak*

$$Z_n := \sqrt{n} \frac{\bar{X}_n - \mathbb{E}X_1}{\sqrt{\text{Var } X_1}} \xrightarrow{D} Z \sim N(0, 1).$$

Jinými slovy, pro všechna  $x \in \mathbb{R}$ ,

$$\lim_{n \rightarrow \infty} P[Z_n \leq x] = \Phi(x) \equiv \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt.$$

Zkráceně tedy máme  $Z_n \xrightarrow[n \rightarrow \infty]{D} N(0, 1)$ .

*Důkaz.* Tentokrát budeme potřebovat rozvoj exponenciály do druhého řádu. Je tedy  $e^{ity} = 1 + ity + \frac{(ity)^2}{2} + o(t^2)$ . Potom také  $\mathbb{E}[e^{itY}] = 1 + it\mathbb{E}Y - \frac{t^2}{2}\mathbb{E}Y^2 + o(t^2)$ .

Definujme  $Y_n := \frac{(\bar{X}_n - \mathbb{E}X_1)}{\sqrt{\text{Var } X_1}}$ . Pak  $\mathbb{E}Y_n = 0$ ,  $\text{Var } Y_n = 1$  a platí

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i = \sqrt{n} \left( \frac{1}{n} \sum_{i=1}^n Y_i \right) = \sqrt{n} \bar{Y}_n = \sqrt{n} \frac{\bar{X}_n - \mathbb{E}X_1}{\sqrt{\text{Var } X_1}} =: Z_n$$

pro každé  $n \in \mathbb{N}$ . Dále můžeme počítat

$$\begin{aligned} \varphi_{Z_n}(t) &= \varphi_{\sum_{i=1}^n \frac{1}{\sqrt{n}} Y_i}(t) = \left( \varphi_{\frac{Y_1}{\sqrt{n}}}(t) \right)^n = \left( \varphi_{Y_1} \frac{t}{\sqrt{n}} \right)^n = \\ &= \left( 1 + \frac{it}{\sqrt{n}} \mathbb{E}Y_1 - \frac{t^2}{2n} \mathbb{E}Y_1^2 + o\left(\frac{t^2}{n}\right) \right)^n = \left( 1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right) \right)^n, \end{aligned}$$

neboť  $\mathbb{E}Y_1 = 0$  a  $\mathbb{E}Y_1^2 = \text{Var } Y_1 = 1$ . Jelikož  $e^y = \lim_{n \rightarrow \infty} \left( 1 + \frac{y}{n} \right)^n$ , dostáváme

$$\lim_{n \rightarrow \infty} \varphi_{Z_n}^t = \lim_{n \rightarrow \infty} \left( 1 - \frac{t^2}{2n} \right)^n = e^{-\frac{1}{2}t^2} = \varphi_Z(t),$$

kde  $Z \sim N(0, 1)$ . Z Lévyho věty (Věta 5.10) je konvergence charakteristických funkcí ekvivalentní konvergenci v distribuci, čímž je důkaz ukončen.  $\square$

Následující věta nám umožní zformulovat mnohorozměrnou verzi centrální limitní věty.

**Věta 5.16** (Cramér-Woldova věta). *Platí  $\vec{X}_n \xrightarrow{D} X$  právě tehdy, když pro všechna  $\vec{t} \in \mathbb{R}^d$  platí  $\vec{t}^T \vec{X}_n \xrightarrow{D} \vec{t}^T X$ .*

*Důkaz.* Důkaz plyne okamžitě z Věty 5.10 a spojitosti lineárních zobrazení.  $\square$

**Věta 5.17** (Mnohorozměrná CLT). *Pokud  $\{\vec{X}_n\}_{n \in \mathbb{N}}$  je IID posloupnost d-rozměrných náhodných vektorů s pozitivně definitní varianční-kovarianční maticí  $\text{Var } \vec{X}_1$ , pak*

$$\sqrt{n}(\vec{X}_n - \mathbb{E}\vec{X}_1) \xrightarrow{D} N_d(\vec{0}, \text{Var } \vec{X}_1), n \rightarrow \infty.$$

**Věta 5.18** (Delta metoda). *Pokud  $\sqrt{n}(Y_n - \mu) \xrightarrow{D} N(0, \sigma^2)$  a  $g$  je spojitě diferencovatelná v okolí  $\mu$  taková, že  $g'(\mu) \neq 0$ , pak*

$$\sqrt{n}(g(Y_n) - g(\mu)) \xrightarrow{D} N(0, (g'(\mu))^2 \sigma^2).$$

*Důkaz.* Podle věty o střední hodnotě dostáváme  $g(Y_n) = g(\mu) + g'(\tilde{\mu})(Y_n - \mu)$ , kde  $\tilde{\mu}$  leží mezi  $Y_n$  a  $\mu$ . Když  $\sqrt{n}(Y_n - \mu) \xrightarrow{D} N(0, \sigma^2)$ , potom  $Y_n - \mu \xrightarrow{D} 0$ . Pak tedy i  $Y_n - \mu \xrightarrow{P} 0$ . Jelikož  $|\tilde{\mu} - \mu| \leq |Y_n - \mu|$ , musí platit, že  $\tilde{\mu} - \mu \xrightarrow{P} 0$ . Dle věty o spojitém zobrazení (Věta 5.7) dostáváme  $g'(\tilde{\mu}) \xrightarrow{P} g'(\mu)$  (předpokládáme spojitost  $g'$ ).

Dále můžeme dosadit  $\sqrt{n}[g(Y_n) - g(\mu)] = g'(\tilde{\mu})\sqrt{n}(Y_n - \mu)$ . Nakonec, ze Slutského věty (Věta 5.8) máme (využili jsme předpoklad, že  $\sqrt{n}(Y_n - \mu) \xrightarrow{D} N(0, \sigma^2)$ )

$$\sqrt{n}[g(Y_n) - g(\mu)] \xrightarrow{D} N(0, [g'(\mu)]^2 \sigma^2).$$

$\square$

Obdobně se dá zformulovat i mnohorozměrná delta-metoda (za předpokladu nenulovosti jakobiánu  $g$ ).

*konec 14. přednášky (1.4.2025)*

## 6 Statistické učení

V této kapitole se budeme věnovat základům matematické statistiky, což je obor, který bude středobodem naší pozornosti po celý zbytek semestru. Začneme formalizací pojmů týkajících se opakovaného provádění experimentu a charakterizací statistických modelů.

**Definice 6.1.** Pokud jsou  $X_1, \dots, X_n$  nezávislé a každá má stejné marginální rozdělení a distribuční funkci  $F$ , říkáme, že  $X_1, \dots, X_n$  jsou IID (nezávislé a stejně rozdělené) a píšeme

$$X_1, \dots, X_n \stackrel{IID}{\sim} F.$$

Takové  $X_1, \dots, X_n$  nazýváme *náhodný výběr* velikosti  $n$  z  $F$ .

Obecně si představujeme měřitelná zobrazení  $X_1, \dots, X_n$ . V praxi však většinou dostaneme pouze reálná čísla  $X_i(\omega)$  pro jedno konkrétní  $\omega \in \Omega$ . Možná rozdělení těchto náhodných veličin budeme modelovat pomocí takzvaných parametrických modelů, tedy množin  $\mathcal{F}$  rozdělení, jež se dají parametrizovat konečným počtem parametrů.

**Příklad 6.2** (Normální model).

$$\mathcal{F} = \left\{ f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(x - \mu)^2}{2\sigma^2} \right\}, \mu \in \mathbb{R}, \sigma^2 > 0 \right\}.$$

Taková data pochází z normálního rozdělení se dvěma parametry  $\mu$  a  $\sigma^2$ .

Všechny parametrické modely můžeme obecně zapsat ve tvaru

$$\mathcal{F} = \{ f(\cdot; \vec{\theta}) : \vec{\theta} \in \vec{\Theta} \subseteq \mathbb{R}^d \}.$$

V dalším textu budeme využívat následující značení:

$$P_\theta[X \in A] := \int_A f(x; \theta) dx;$$

$$\mathbb{E}_\theta[g(X)] := \int_{\mathbb{R}} g(x) f(x; \theta) dx.$$

Velkou skupinou modelů jsou také neparametrické modely, který nemůžeme parametrizovat konečným počtem parametrů. Například, celou funkci hustoty můžeme považovat za nekonečnědimenzionální prostor. Uvedeme si jeden příklad takového neparametrického modelu.

**Příklad 6.3** (Model Sobolevova prostoru).

$$\mathcal{F} = \left\{ f : \int_{\mathbb{R}} (f''(x))^2 dx < \infty \right\}.$$

Data pochází z rozdělení s nepříliš “vlnitou” hustotou.

**Definice 6.4.** Bodový odhad  $\hat{\theta}_n$  parametru  $\theta$  je měřitelná funkce  $t$  náhodných veličin  $X_1, \dots, X_n$ :

$$\hat{\theta}_n = t(X_1, \dots, X_n).$$

V této definici předpokládáme, že  $\theta$  je pevné ale neznámé reálné číslo (vektor). Avšak získaný odhad  $\hat{\theta}_n$  je sice náhodná veličina, ale umíme ji přesně charakterizovat.

**Definice 6.5.** Odhad  $\hat{\theta}_n$  je *nestranný*, pokud  $\mathbb{E}[\hat{\theta}_n] = \theta$  pro všechna  $n \in \mathbb{N}$ . *Vychýlení* odhadu definujeme jako  $\text{bias}(\hat{\theta}_n) := \mathbb{E}[\hat{\theta}_n] - \theta$ . Odhad je *konzistentní*, jestliže  $\hat{\theta}_n \xrightarrow{P} \theta$  pro  $n \rightarrow \infty$ .

V dnešní době je díky vývoji výpočetní techniky nestrannost více upozadována, větší důraz proto klademe na konzistenci modelu.

**Definice 6.6.** Rozdělení odhadu  $\hat{\theta}_n$  nazýváme *výběrové rozdělení*. Standardní odchylku  $\hat{\theta}_n$  nazýváme *standardní chyba*  $\text{se}(\hat{\theta}_n) = \sqrt{\text{Var} \hat{\theta}_n}$ .

V těchto případech je standardní chyba se neznámá veličina (parametr), ale obvykle ji můžeme odhadnout. Takovou odhadnutou standardní chybu značíme  $\hat{\text{se}}$ .

**Příklad 6.7.** Mějme Bernoulliho náhodný výběr  $X_1, \dots, X_n \stackrel{IID}{\sim} \text{Be}(p)$  a parametr  $p \in (0, 1)$ . Potom můžeme uvažovat odhad

$$\hat{p}_n := \frac{1}{n} \sum_{i=1}^n X_i$$

a z toho získáme odhad standardní chyby (díky nezávislosti a stejné rozdělenosti máme  $\text{Var}(\hat{p}_n) = \frac{p(1-p)}{n}$ ). Jelikož přesná hodnota  $p$  je neznámá, musíme tento parametr také odhadnout, proto

$$\hat{\text{se}}(\hat{p}_n) := \sqrt{\frac{\hat{p}_n(1 - \hat{p}_n)}{n}}.$$

**Definice 6.8.** Kvalitu bodového odhadu můžeme posuzovat pomocí *střední kvadratické chyby*

$$\text{MSE}(\hat{\theta}_n) := \mathbb{E}_\theta[\hat{\theta}_n - \theta]^2.$$

Mějme na paměti, že  $\mathbb{E}_\theta$  se v případě nezávislých a stejně rozdělených  $X_i$  vztahuje k očekávané hodnotě vzhledem k rozdělení

$$f(x_1, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i; \theta).$$

**Věta 6.9** (Rozklad střední kvadratické chyby). *Mějme odhad  $\hat{\theta}_n$ . Pak vždy platí  $\text{MSE}(\hat{\theta}_n) = \text{bias}^2(\hat{\theta}_n) + \text{Var}(\hat{\theta}_n)$ .*

*Důkaz.* Rozepsáním definice dostáváme

$$\begin{aligned}\text{MSE}[\hat{\theta}_n] &= \mathbb{E}_\theta[\hat{\theta}_n - \theta]^2 = \mathbb{E}_\theta \left\{ [\hat{\theta}_n - \mathbb{E}\hat{\theta}_n + \mathbb{E}\hat{\theta}_n - \theta]^2 \right\} = \\ &= \mathbb{E}_\theta \left\{ [\hat{\theta}_n - \mathbb{E}\hat{\theta}_n]^2 - 2[\hat{\theta}_n - \mathbb{E}\hat{\theta}_n][\mathbb{E}\hat{\theta}_n - \theta] + [\mathbb{E}\hat{\theta}_n - \theta]^2 \right\} = \\ &= \text{Var}(\hat{\theta}_n) + \text{bias}^2(\hat{\theta}_n),\end{aligned}$$

kde poslední rovnost plyne z toho, že druhý sčítanec je nulový, což plyne z linearity střední hodnoty.  $\square$

**Věta 6.10** (Postačující podmínka pro konzistenci). *Nechť platí  $\text{bias}(\hat{\theta}_n) \rightarrow 0$  a  $\text{Var}(\hat{\theta}_n) \rightarrow 0$ . Potom platí  $\hat{\theta}_n$  je konzistentní.*

*Důkaz.* Z Věty 6.9 dostáváme, že  $\text{MSE}(\hat{\theta}_n) = \mathbb{E}_\theta[\hat{\theta}_n - \theta]^2 \rightarrow 0$ . Z definice  $L_2$  konvergence dostáváme, že  $\hat{\theta}_n \xrightarrow{L_2} \theta$ . Zbytek dostáváme z faktu, že  $L_2$  konvergence implikuje konvergenci v pravděpodobnosti.  $\square$

**Příklad 6.11.** Mějme stejnou situaci jako v Příkladu 6.7. Jelikož náš odhad je nestranný ( $\mathbb{E}(\hat{p}_n) = p$ ) a  $\text{Var}(\hat{p}_n) = \frac{p(1-p)}{n} \rightarrow 0$  pro  $n \rightarrow \infty$ , dostáváme díky Věte 6.10, že  $\hat{p}_n \xrightarrow{P} p$ .

**Definice 6.12.** Odhad  $\hat{\theta}_n$  parametru  $\theta$  se nazývá *asymptoticky standardně normální*, jestliže pro  $n \rightarrow \infty$  platí

$$\frac{\hat{\theta}_n - \theta}{\text{se}(\hat{\theta}_n)} \xrightarrow{D} N(0, 1).$$

**Definice 6.13.**  $(1-\alpha)$ -*interval spolehlivosti* (konfidenční interval) pro parametr  $\theta$  je interval  $C_n = (a, b)$ , kde  $a = a(X_1, \dots, X_n)$  a  $b = b(X_1, \dots, X_n)$  jsou měřitelné funkce dat takové, že pro všechna  $\theta \in \Theta$

$$P_\theta[\theta \in C_n] = 1 - \alpha.$$

*Asymptotický* (přibližný)  $(1-\alpha)$ -*interval spolehlivosti* pro parametr  $\theta$  je interval  $C_n$  takový, že pro všechna  $\theta \in \Theta$

$$\lim_{n \rightarrow \infty} P_\theta[\theta \in C_n] = 1 - \alpha.$$

Tato definice říká, že interval  $C_n$  zachytí  $\theta$  s pravděpodobností (přibližně)  $1 - \alpha$ . Tento parametr nazýváme *pokrytí* intervalu spolehlivosti (CI). Interval spolehlivosti je náhodná veličina, i přestože  $\theta$  je pevné deterministické. Pro vícerozměrné prostory uvažujeme kouli/elipsoid spolehlivosti (ale toto rozšíření je komplikovanější, protože na  $\mathbb{R}^d$ ,  $d > 1$  neexistuje vhodné uspořádání).

*konec 15. přednášky (8.4.2025)*



**Příklad 6.14** (Interval spolehlivosti a normalita). Mějme stejný setup jako v Příkladu 6.2, tedy  $X_1, \dots, X_n \stackrel{IID}{\sim} N(\mu, \sigma^2)$  a parametr  $\mu \in \mathbb{R}$  je *neznámý* a má být *odhadnut* (bodový odhad a konfidenční interval),  $\sigma^2 > 0$  je předpokládáno *známé*.

Nezávislost a stejná rozdělenost nám poskytuje možnost pracovat s mnoho-rozměrným vektorem  $\vec{X} := (X_1, \dots, X_n)^T \sim ((\mu, \dots, \mu)^T, \sigma^2 I_n)$ .

Uvažujme bodový odhad  $\hat{\mu}_n = \frac{1}{n} \sum_{i=1}^n X_i \equiv \bar{X}_n$ . Tento odhad je konzistentní, což plyne ze zákona velkých čísel 5.14. Z linearity normálních rozdělení dostáváme  $\sqrt{n}(\hat{\mu}_n - \mu)/\sigma \sim N(0, 1)$ .

Nechť  $u_\beta := \Phi^{-1}(\beta)$  je  $\beta$ -kvantil ( $\beta \in (0, 1)$ ) standardního normálního rozdělení a necht  $Y = \sqrt{n} \frac{\bar{X}_n - \mu}{\sqrt{\sigma^2}}$ . Potom

$$P[-u_{1-\alpha/2} \leq Y \leq u_{1-\alpha/2}] = P[Y \leq u_{1-\alpha/2}] - P[Y \leq -u_{1-\alpha/2}] = 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1 - \alpha$$

pro všechna  $n \in \mathbb{N}$  a  $\mu \in \mathbb{R}$ . Ve druhé rovnosti jsme použili vlastnost  $u_\beta = -u_{1-\beta}$ , která se snadno ověří přímým dosazením do definice distribuční funkce normálního rozdělení.

Jednoduchými algebraickými úpravami získáme nerovnost pro  $\mu$  (zapišeme to rovnou ve tvaru intervalu):

$$\mu \in \left( \bar{X}_n - u_{1-\alpha/2} \sqrt{\frac{\sigma^2}{n}}, \bar{X}_n + u_{1-\alpha/2} \sqrt{\frac{\sigma^2}{n}} \right).$$

Tento interval je  $1 - \alpha$  interval spolehlivosti pro hodnotu  $\mu$ . Tedy s pravděpodobností  $1 - \alpha$  leží hodnota  $\mu$  v tomto intervalu.

Pro výpočet také můžeme použít centrální limitní větu, v tomto případě dostaneme stejný interval spolehlivosti (s poznámkou, že jde o přibližný, tedy asymptotický interval spolehlivosti).

Zkoumejme délku získaného intervalu spolehlivosti. Z předchozího příkladu máme délku  $2u_{1-\alpha/2} \sqrt{\frac{\sigma^2}{n}}$ . Poznamenejme si, že s klesajícím  $\alpha$  (povolená tolerance) roste délka intervalu. Taktéž roste délka intervalu s rostoucím rozptylem  $\sigma^2$  a klesajícím rozsahem výběru  $n$ . Mějme danou délku  $d$ . Kolik pozorování potřebujeme, abychom získali interval spolehlivosti užší než  $d$ ? Vychází

$$n \geq \left\lceil \frac{4u_{1-\alpha/2}^2 \sigma^2}{d^2} \right\rceil + 1.$$

**Věta 6.15** (Interval spolehlivosti založený na normalitě). *Předpokládejme, že  $\hat{\theta}_n$  je asymptoticky standardně normální odhad parametru  $\theta$  a  $\widehat{\text{se}}(\hat{\theta}_n)$  je konzistentní odhad  $\text{se}(\theta_n)$ , tj.  $\widehat{\text{se}}(\hat{\theta}_n) - \text{se}(\hat{\theta}_n) \xrightarrow{P} 0$ . Necht  $u_{1-\alpha/2}$  je  $(1 - \alpha/2)$ -kvantil standardního normálního rozdělení a*

$$C_n = \left( \hat{\theta}_n - u_{1-\alpha/2} \widehat{\text{se}}(\hat{\theta}_n), \hat{\theta}_n + u_{1-\alpha/2} \widehat{\text{se}}(\hat{\theta}_n) \right).$$

Pak

$$P_\theta[\theta \in C_n] \xrightarrow{n \rightarrow \infty} 1 - \alpha.$$

*Důkaz.* Z definice asymptoticky standardně normálního odhadu máme

$$\frac{\hat{\theta}_n - \theta}{\text{se}(\hat{\theta}_n)} \xrightarrow{D} N(0, 1)$$

a máme konzistentní odhad standardní chyby odhadu

$$\widehat{\text{se}}(\hat{\theta}_n H) - \text{se}(\hat{\theta}_n) \xrightarrow{P} 0.$$

Ze Slutského věty (Věta 5.8) dostáváme  $Y := \frac{\hat{\theta}_n - \theta}{\widehat{\text{se}}(\hat{\theta}_n)} \xrightarrow{D} N(0, 1)$ . Potom již platí

$$\lim_{n \rightarrow \infty} P[-u_{1-\alpha/2} \leq Y \leq u_{1-\alpha/2}] = \Phi(u_{1-\alpha/2}) - \Phi(u_{\alpha/2}) = 1 - \alpha.$$

□

Neformálně zapisujeme  $\theta_n \approx N(\theta, \widehat{\text{se}}(\hat{\theta}_n))$ . Přibližně platí pro 95%-intervaly spolehlivosti  $\alpha = 0.05$  a  $u_{0.975} \approx 1.96 \approx 2$  vedoucí k explicitnímu intervalu spolehlivosti  $\hat{\theta}_n \pm 2\widehat{\text{se}}(\hat{\theta}_n)$ .

**Příklad 6.16.** Pokračujeme v Příkladu 6.11. Již jsme spočítali  $\hat{p}_n \xrightarrow{P} p$ ,  $\text{se}(\hat{p}_n) = \sqrt{p(1-p)/n}$  a  $\widehat{\text{se}}(\hat{p}_n) := \sqrt{\hat{p}_n(1-\hat{p}_n)/n}$ . Ze Slutského věty pak máme  $\widehat{\text{se}}(\hat{p}_n) - \text{se}(\hat{p}_n) \xrightarrow{P} 0$ .

Z centrální limitní věty dostáváme, že  $\frac{\hat{p}_n - p}{\widehat{\text{se}}(\hat{p}_n)} \xrightarrow{D} N(0, 1)$ . Dále opětovným použitím Slutského věty získáme  $\frac{\hat{p}_n - p}{\widehat{\text{se}}(\hat{p}_n)} \xrightarrow{D} N(0, 1)$ . Tedy z předchozí věty

$$\hat{p}_n \pm u_{1-\alpha/2} \sqrt{\hat{p}_n(1-\hat{p}_n)/n}$$

je asymptotický (přibližný)  $(1 - \alpha)$ -interval spolehlivosti pro  $p$ .

## 7 Statistické funkcionály

Nechť  $X_1, \dots, X_n$  je IID náhodný výběr z  $F$  s rozsahem výběru  $n$ . Chceme odhadnout  $F$  jejím empirickým protějškem.

**Definice 7.1** (ECDF). Pro  $x \in \mathbb{R}$  definujeme

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n \chi_{\{X_i \leq x\}}.$$

Takto definovaná empirická distribuční funkce splňuje všechny vlastnosti normální distribuční funkce a přiřazuje váhu  $\frac{1}{n}$  každému pozorování  $X_i$ . Takéž ECDF můžeme definovat jako relativní četnost  $X$  menších nebo rovných pevnému  $x$ , to znamená  $\frac{1}{n} |\{X_i \leq x\}|$ .

*konec 16. přednášky (14.4.2025)*

**Věta 7.2** (Bodové vlastnosti ECDF). Pro libovolné pevné  $x \in \mathbb{R}$ ,

- (i)  $\mathbb{E} [\hat{F}_n(x)] = F(x);$
- (ii)  $\text{Var} [\hat{F}_n(x)] = \frac{F(x)(1-F(x))}{n};$
- (iii)  $\text{MSE} (\hat{F}_n(x)) = \frac{F(x)(1-F(x))}{n} \rightarrow 0$  pro  $n \rightarrow \infty;$
- (iv)  $\hat{F}_n(X) \xrightarrow{P} F(x)$  pro  $n \rightarrow \infty.$

*Důkaz.* Platí  $\mathbb{E} [\hat{F}_n(x)] = \mathbb{E} [\frac{1}{n} \sum_{i=1}^n \chi_{\{X_i \leq x\}}] = \frac{1}{n} \sum_{i=1}^n P[X_i \leq x] = F(x)$ . Tím jsme dokázali první vlastnost.

Dále platí  $\text{Var} [\hat{F}_n(x)] = \text{Var} [\frac{1}{n} \sum_{i=1}^n \chi_{\{X_i \leq x\}}] = \frac{1}{n^2} \sum_{i=1}^n \text{Var} [\chi_{\{X_i \leq x\}}] = \frac{1}{n^2} \sum_{i=1}^n [\mathbb{E} \chi^2 - (\mathbb{E} \chi)^2] = \frac{1}{n^2} \sum_{i=1}^n (F(x) - F(x)^2) = \frac{F(x)(1-F(x))}{n}$ , čímž jsme dokázali druhou vlastnost.

K důkazu třetí rovnosti si uvědomíme, že díky již dokázané vlastnosti (i) je  $\text{bias}(\hat{F}_n(x)) = 0$  a tedy  $\text{MSE} (\hat{F}_n(x)) = \text{Var} [\hat{F}_n(x)]$ .

Nakonec, díky zákonu velkých čísel (Věta 5.14) máme

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n \chi_{\{X_i \leq x\}} \xrightarrow{P} \mathbb{E} [\chi_{\{X_1 \leq x\}}] = F(x).$$

□

**Definice 7.3.** *Funkcionál* je zobrazení  $T : \mathcal{F} \rightarrow \mathbb{R}$ , kde  $\mathcal{F}$  je nějaká množina funkcí.

**Definice 7.4.** *Statistický funkcionál* je zobrazení  $T$ , které přiřadí rozdělení  $P_X$  reálné číslo.

Můžeme také definovat vektorové funkcionály, stačí obor hodnot nahradit  $\mathbb{R}^d$ . Uvedeme si několik příkladů statistických funkcionálů.

**Příklad 7.5.** Následující operátory jsou statistické funkcionály:

- střední hodnota  $\mu = \mathbb{E}X = \int x dP_X(x)$ ;
- rozptyl  $\sigma^2 = \text{Var } X = \int (x - \mu)^2 dP_X(x)$ ;
- medián  $F^{-1}(1/2) \equiv \inf\{x : P_X((-\infty, x]) > 1/2\}$ .

**Definice 7.6.** Pokud  $T(P_X) = \int r(x) dP_X(x)$  pro nějakou měřitelnou funkci  $r$ , pak  $T$  nazýváme *lineární statistický funkcionál*.

Motivací této definice je fakt, že takto definovaný funkcionál  $T$  je lineární ve svých argumentech, jinými slovy,

$$T(aP_X + bP_Y + c) = aT(P_X) + bT(P_Y) + c$$

pro  $a, b, c \in \mathbb{R}$ .

Z předchozího příkladu dostaneme, že střední hodnota a rozptyl jsou lineární a medián není (neexistuje vhodná měřitelná funkce  $r$ ).

**Definice 7.7.** Necht  $X_1, \dots, X_n$  je náhodný výběr z  $F$  s rozsahem výběru  $n$ , kde  $X_i : (\Omega, \mathcal{A}, P) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$  pro  $i = 1, \dots, n$ . Pak

$$\hat{P}_n(B) := \frac{1}{n} \sum_{i=1}^n \chi_{\{X_i \in B\}} \equiv \frac{1}{n} \sum_{i=1}^n \delta_{X_i}(B)$$

pro  $B \in \mathcal{B}(\mathbb{R})$  se nazývá *empirická pravděpodobnostní míra*.

Právě definovaný objekt je *náhodná pravděpodobnostní míra*, které má diskrétní rovnoměrné pravděpodobnostní rozdělení (součet Diracových měr) na náhodných bodech  $X_1, \dots, X_n$ , kde každý tento bod má váhu  $\frac{1}{n}$ .

**Definice 7.8.** *Plug-in odhad* neznámého parametru  $\theta = T(P_X)$  je  $\hat{\theta}_n := T(\hat{P}_n)$ .

Myšlenkou definice plug-in odhadu je nahrazení neznámé pravděpodobnostní míry jejím odhadem.

**Příklad 7.9.** Platí  $\hat{F}_n(x) = \hat{P}_n((-\infty, x])$  pro  $x \in \mathbb{R}$ .

**Definice 7.10.** *Empirický (plug-in) odhad* pro lineární statistický funkcionál  $T(P_X) = \int r(x) dP_X(x)$  je

$$T(\hat{P}_n) = \int r(x) d\hat{P}_n(x).$$

**Věta 7.11** (Výpočet plug-in odhadu pro lineární statistický funkcionál). *Pro empirický odhad lineárního statistického funkcionálu  $T(P_X) = \int r(x) dP_X(x)$  platí*

$$T(\hat{P}_n) = \frac{1}{n} \sum_{i=1}^n r(X_i).$$

*Důkaz.* Necht  $\omega \in \Omega$  je dáno. Z definice empirického odhadu lineárního statistického funkcionálu dostáváme

$$\begin{aligned} T(\hat{P}_n)(\omega) &= \int_{\mathbb{R}} r(x) d\hat{P}_n(\omega)(x) = \int_{\mathbb{R}} r(x) d\left(\frac{1}{n} \sum_{i=1}^n \delta_{X_i(\omega)}(x)\right) = \\ &= \frac{1}{n} \sum_{i=1}^n \int_{\mathbb{R}} r(x) d\delta_{X_i(\omega)}(x) = \frac{1}{n} \sum_{i=1}^n r(X_i(\omega)). \end{aligned}$$

□

**Příklad 7.12.** Spočteme empirickou střední hodnotu. Máme  $\mu = T(P_X) = \int x dP_X(x)$  a tedy díky předchozí větě

$$\hat{\mu}_n = \int x d\hat{P}_n(x) = \bar{X}_n.$$

**Příklad 7.13.** Spočteme empirický rozptyl. Z definice rozptylu máme

$$\text{Var } X = \sigma^2 = T(P_X) = \int (x - \mu)^2 dP_X(x) = \int x^2 dP_X(x) - \left(\int x dP_X(x)\right)^2.$$

Potom

$$\begin{aligned} \hat{\sigma}_n^2 &= \int x^2 d\hat{P}_n(x) - \left(\int x d\hat{P}_n(x)\right)^2 = \\ &= \frac{1}{n} \sum_{i=1}^n X_i^2 - \left(\frac{1}{n} \sum_{i=1}^n X_i\right)^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2. \end{aligned}$$

**Příklad 7.14.** Spočteme empirickou korelaci. Necht tedy  $Z = [X, Y]^T$  a necht  $\rho = T(P_{[X, Y]^T})$  označuje příslušnou korelaci. Můžeme psát

$$T(P_{[X, Y]^T}) = a(T_1(P_{[X, Y]^T}), T_2(P_{[X, Y]^T}), T_3(P_{[X, Y]^T}), T_4(P_{[X, Y]^T}), T_5(P_{[X, Y]^T})),$$

kde

$$\begin{aligned} T_1(P_{[X, Y]^T}) &= \int x dP_{[X, Y]^T}(x, y), \\ T_2(P_{[X, Y]^T}) &= \int y dP_{[X, Y]^T}(x, y), \\ T_3(P_{[X, Y]^T}) &= \int xy dP_{[X, Y]^T}(x, y), \\ T_4(P_{[X, Y]^T}) &= \int x^2 dP_{[X, Y]^T}(x, y), \\ T_5(P_{[X, Y]^T}) &= \int y^2 dP_{[X, Y]^T}(x, y) \end{aligned}$$

a zároveň  $a(t_1, t_2, t_3, t_4, t_5) = \frac{t_3 - t_1 t_2}{\sqrt{(t_4 - t_1^2)(t_5 - t_2^2)}}$ . Dosazením se snadno ověří, že tímto jsme opravdu získali vzorec pro daný funkcionál. Nahrazením distribuční funkce jejím empirickým protějškem nakonec dostáváme

$$\hat{\rho} = \frac{\sum_i (X_i - \bar{X}_n)(Y_i - \bar{Y}_n)}{\sqrt{\sum_i (X_i - \bar{X}_n)^2 \sum_j (Y_j - \bar{Y}_n)^2}}.$$

Tuto veličinu nazýváme *výběrovou korelací*.

**Definice 7.15.** Připomínka: pro  $p \in (0, 1)$  definujeme *p-tý kvantil* jako  $T(F) = F^{-1}(p) = \inf\{x : F(x) > p\}$ .

Nyní definujeme

$$T(\hat{F}_n) = \hat{F}_n^{-1}(p) = \inf\{x : \hat{F}_n(x) > p\}$$

a tento objekt nazýváme *p-tý výběrový kvantil*. Obdobně definujeme *výběrový medián* jako  $\hat{F}_n^{-1}(1/2)$ . Navíc mezikvartilové rozpětí  $\tilde{T}(F) = F^{-1}(3/4) - F^{-1}(1/4)$  lze odhadnout pomocí *výběrového mezikvartilového rozpětí*  $\tilde{T}(\hat{F}_n) = \hat{F}_n^{-1}(3/4) - \hat{F}_n^{-1}(1/4)$ .

## 8 Parametrická inference

V této kapitole se budeme věnovat problému odhadování parametru tak, aby získané rozdělení co nejvhodněji pasovalo na experimentální data. Hlavním objektem zkoumání bude rodina parametrických modelů

$$\mathcal{F} := \{f(\cdot, \vec{\theta}) : \vec{\theta} \in \vec{\Theta} \subseteq \mathbb{R}^d\},$$

kde  $\vec{\Theta}$  je parametrický prostor a  $\vec{\theta} = (\theta_1, \dots, \theta_d)$  je parametr.

Je zřejmé, že ne každý model je schopen pokrýt všechna možná rozdělení vyskytující se v přírodě. Musíme proto aproximovat a umět dobře odhadnout, kdy máme “dost dobrý” odhad. Budeme se zajímat o odhad nějaké funkce  $T(\vec{\theta})$ . Například pro  $X_i \sim N(\mu, \sigma^2)$ , pokud naším parametrem zájmu je  $\mu$ , stačí volit  $T(\vec{\theta}) = \mu$  a  $\sigma^2$  se potom nazývá *nežádoucí/rušivý parametr*.

**Příklad 8.1.** Připomeňme si, že náhodná veličina  $X$  má rozdělení  $\Gamma(a, p)$ , jestliže

$$f_X(x; a, p) = \frac{a^p}{\Gamma(p)} x^{p-1} \exp\{-ax\}.$$

kde  $a, p > 0$  a

$$\Gamma(p) = \int_0^\infty y^{p-1} e^{-y} dy.$$

Parametrem ve smyslu úvodu je tedy vektor  $\vec{\theta} = (a, p)$ . Chceme-li spočítat průměrnou délku života (což je jedna z věcí, k modelování kterých se používá Gamma rozdělení), dostáváme

$$T(a, p) = \mathbb{E}_{\vec{\theta}} X = \int_0^\infty \frac{a^p x^p}{\Gamma(p)} e^{-ax} dx = \frac{1}{a\Gamma(p)} \int_0^\infty y^p e^{-y} dy = \frac{\Gamma(p+1)}{a\Gamma(p)} = \frac{p}{a}.$$

V dalším textu uvažujme náhodný výběr  $X_1, \dots, X_n \stackrel{IID}{\sim} F \in \mathcal{F}$ .

**Příklad 8.2.** Uvažujme  $\mathcal{F} = \{F(\mu) : \mathbb{E}_{F(\mu)} = \mu \wedge |\mu| < \infty\}$  rodinu modelů s konečnou střední hodnotou. Potom  $\bar{X}_n$  je konzistentní a nestranný odhad  $\mu$  a  $X_1$  je nestranný, ale ne konzistentní odhad  $\mu$ .

Dále uvažujme  $\mathcal{F} = \{F(\sigma^2) : \text{Var}_{F(\sigma^2)} = \sigma^2 < \infty\}$  rodinu modelů s konečným rozptylem. Potom  $\hat{\sigma}_n^2 = n^{-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$  je konzistentní, ale ne nestranný odhad  $\sigma^2$  a  $S_n^2 = (n-1)^{-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$  je konzistentní a nestranný odhad  $\sigma^2$ .

*konec 17. přednášky (15.4.2025)*

## A Ukázková zápočtová písemka – Varianta A

**Příklad A.1.** V šesti urnách máme v každé 10 míčků. V jedné urně je osm černých, ve dvou urnách je po pěti černých a ve třech urnách je po  $k$  černých. Zbylé míčky jsou bílé.

- (a) Nechť  $k = 3$ . Z náhodně vybrané urny jsem s vracením vytáhl dva míčky, oba bílé. S jakou pravděpodobností bylo v urně pět černých míčků?  
*Nechť  $V_i$  je událost, že byla vybrána urna s  $i$  bílými míčky,  $H$  je událost, že oba vytážené míčky jsou bílé. Z Bayesovy věty (Věta 1.15) máme, že  $P(V_5|H) = P(H|V_5) \frac{P(V_5)}{P(H)}$ . Dále máme, že  $P(H|V_5) = \left(\frac{1}{2}\right)^2 = \frac{1}{4}$  a  $P(V_5) = \frac{2}{6} = \frac{1}{3}$ . Nakonec spočteme*

$$P(H) = \sum_i P(H|V_i)P(V_i) = \frac{1}{6} \cdot \left(\frac{1}{5}\right)^2 + \frac{2}{6} \cdot \left(\frac{1}{2}\right)^2 + \frac{3}{6} \cdot \left(\frac{7}{10}\right)^2 = \frac{67}{200}.$$

*Dosazením do výše uvedeného vzorce dostáváme  $P(V_5|H) = \frac{1}{4} \cdot \frac{1}{3} \cdot \frac{200}{67} = \frac{50}{201}$ .*

- (b) Kolik musí být  $k$ , aby pravděpodobnost vytážení černého míčku z náhodně vybrané urny byla  $\frac{1}{2}$ .  
*Využijeme větu o úplné pravděpodobnosti (Věta 1.14). Nechť  $B, W$  jsou jevy, že jsme vytáhli černý/bílý míček a  $A_i$  reprezentuje jev, že byla vybrána  $i$ -tá urna. Platí*

$$\begin{aligned} \frac{1}{2} = P[B] &= \sum_{i=1}^n P[B|A_i]P[A_i] = \underbrace{P[B|A_1]A_1}_{\frac{8}{10} \cdot \frac{1}{6} = \frac{2}{15}} + \underbrace{2 \cdot P[B|A_2]A_2}_{2 \cdot \frac{5}{10} \cdot \frac{1}{6} = \frac{1}{6}} + \\ &\underbrace{3 \cdot P[B|A_4]P[A_4]}_{3 \cdot \frac{k}{10} \cdot \frac{1}{6} = \frac{k}{20}} = \frac{2}{15} + \frac{1}{6} + \frac{k}{20} = \frac{6+k}{20}. \end{aligned}$$

*Vyřešením této lineární rovnice dostáváme, že  $k = 4$ .*

- (c) Nechť  $k = 2$ . Náhodně vybereme jednu urnu, kterou vynecháme, ze zbylých náhodně vytáhneme po jednom míčku. Jaká je pravděpodobnost, že všechny vytážené míčky jsou bílé?  
*Nechť  $V_i$  je událost, že byla vyřazena  $i$ -tá urna. Potom nechť  $W_i$  je událost, že byl vytážen bílý míček z  $i$ -té urny. Jevy  $W_i$  jsou navzájem nezávislé a zároveň jsou nezávislé jevy  $W_j$  a  $V_i$  pro  $i \neq j$ . Označme hledanou pravděpodobnost  $P$ , můžeme psát*

$$\begin{aligned} P &= \sum_{i=1}^6 P(V_i) \left( \prod_{j \neq i} P(W_j) \right) = \frac{1}{6} \underbrace{\left( \left(\frac{5}{10}\right)^2 \cdot \left(\frac{8}{10}\right)^3 \right)}_{\frac{8}{375}} + \\ &\underbrace{\frac{2}{6} \left( \frac{2}{10} \cdot \frac{5}{10} \cdot \left(\frac{8}{10}\right)^3 \right)}_{\frac{32}{1875}} + \underbrace{\frac{3}{6} \left( \frac{2}{10} \cdot \left(\frac{5}{10}\right)^2 \cdot \left(\frac{8}{10}\right)^2 \right)}_{\frac{2}{125}} = \frac{34}{625}. \end{aligned}$$



**Příklad A.2.** Náhodná veličina  $Y$  má spojité rozdělení s hustotou

$$f_Y(y) = \begin{cases} 3e^{-3y}, & y \geq 0; \\ 0, & \text{jinak.} \end{cases}$$

Definujeme  $U := \lceil Y \rceil$  a  $V := \lceil Y \rceil - Y$  (tedy horní celou a frakcionální část  $Y$ ).

- (a) Určete rozdělení náhodné veličiny  $U$ .

*Zřejmě půjde o diskrétní rozdělení, kde pro  $u \geq 1$  budeme mít*

$$P[U = u] = P[u-1 < Y < u] = \int_{u-1}^u 3e^{-3y} dy = e^{-3u}(e^3 - 1).$$

- (b) Určete rozdělení náhodné veličiny  $V$ .

*Tentokrát půjde o spojité rozdělení, spočteme jeho distribuční funkci. Pro  $v \in (0, 1)$  máme*

$$F_V(v) = P[V < v] = \sum_{t=1}^{\infty} \int_{t-v}^t 3e^{-3y} dy = \sum_{t=1}^{\infty} e^{-3t}(e^{3v} - 1) = \frac{e^{3v} - 1}{e^3 - 1}.$$

*Pro  $v \leq 0$  je  $F_V$  nulová, pro  $v \geq 1$  máme  $F_V(v) = 1$ .*

- (c) Spočítejte  $\mathbb{E}(Ye^{-Y} - 1)$ .

*Využijeme pravidlo líného statistika (Věta 3.5) s transformací  $t(y) = ye^{-y}$ , dostaneme*

$$\mathbb{E}(Ye^{-Y}) = \int_0^{\infty} (ye^{-y} \cdot 3e^{-3y}) = \int_0^{\infty} 3ye^{-4y} = \frac{3}{16}.$$

*S využitím linearity střední hodnoty dostaneme  $\mathbb{E}(Ye^{-Y} - 1) = \mathbb{E}(Ye^{-Y}) - 1 = -\frac{13}{16}$ .*

**Příklad A.3.** Náhodný vektor  $(X, Y)^T$  má spojité rozdělení s hustotou

$$f(x, y) = \begin{cases} ay, & -1 < x < -1, 0 < y \leq x^2; \\ 0, & \text{jinak,} \end{cases}$$

kde  $a \in \mathbb{R}$  je vhodná konstanta.

- (a) Určete  $a \in \mathbb{R}$ .

*Platí, že integrál hustoty přes celý prostor je roven 1. Můžeme tedy psát*

$$1 = \int_{-1}^1 \int_0^{x^2} ay dy dx = \frac{a}{5}.$$

*Tedy  $a = 5$ .*

- (b) Určete marginální rozdělení a střední hodnotu náhodné veličiny  $X$ .  
Platí pro  $|x| < 1$  (jinde je hustota nulová)

$$f_X(x) = \int_{\mathbb{R}} f(x, y) dy = \int_0^{x^2} 5y dy = \frac{5x^4}{2},$$

dále spočteme

$$\mathbb{E}[X] = \int_{-1}^1 x f_X(x) dx = \int_{-1}^1 \frac{5x^5}{2} = 0.$$

- (c) Určete  $P(0 < X < \sqrt{2Y})$ .

Nalezneme průsečík grafů funkcí  $y = x^2$  a  $x = \sqrt{2y}$ . Máme  $y = x^2 = \frac{1}{2}x^2$ , tedy grafy těchto funkcí se protnou pouze počátku a jinde je hodnota druhé funkce ostře menší než hodnota první. Požadovaná podmínka je splněna pro hodnoty  $(x, y) \in \mathbb{R}^2$ , které se nachází mezi těmito dvěma parabolami. Můžeme tedy integrovat

$$P(0 < X < \sqrt{2Y}) = \int_0^1 \int_{\frac{1}{2}x^2}^{x^2} 5y dy dx = \frac{3}{8}.$$

- (d) Spočítejte kovarianci veličin  $X$  a  $Z$ , kde  $Z = XY$ . Dle známého vzorce máme  $\text{Cov}(X, XY) = \mathbb{E}(X \cdot XY) - \mathbb{E}X \mathbb{E}XY = \mathbb{E}[X^2Y] - \mathbb{E}X \mathbb{E}XY$ . Opět s využitím pravidla lineární statistika můžeme integrovat

$$\text{Cov}(X, XY) = \int_{-1}^1 \int_0^{x^2} x^2 y \cdot 5y dy dx - 0 = \frac{10}{27}.$$

- (e) Rozhodněte, zda jsou veličiny  $X$  a  $Y$  nezávislé a své rozhodnutí zdůvodněte.  
Náhodné veličiny  $X$  a  $Y$  nejsou nezávislé, neboť

$$0 = P[0 < x < \frac{1}{2}, y > \frac{1}{4}] \neq P[0 < x < \frac{1}{2}]P[y > \frac{1}{4}],$$

jelikož zjevně ani jeden z činitelů na pravé straně není nulový.

## B Ukázková zápočtová písemka – Varianta B

**Příklad B.1.** Máme šest truhel a v každé z nich je jedna stříbrná mince. Do jedné truhly vložíme tři zlaté mince, do dvou truhel dvě zlaté mince a do zbylých tří truhel po jedné zlaté minci (dohromady jsme tak dodali deset zlatých mincí).

- (a) Z náhodně vybrané truhly dvakrát táhneme s vracením. Jaká je pravděpodobnost, že vytáhneme jednu zlatou a jednu stříbrnou minci.  
Nechť  $H$  je událost, že jsme vytáhli jednu zlatou a jednu stříbrnou minci z dané truhly a  $V_i$  jsou události, že jsme náhodně vybrali  $i$ -tou truhlu. Potom  $P(H) = 2^{\frac{p-1}{p}} \cdot \frac{1}{p}$  ( $p$  je počet mincí v dané truhle, chceme jednu zlatou

a jednu stříbrnou a nezáleží na pořadí) a  $P(V_i) = \frac{1}{6}$ . Potom ze zákona úplné pravděpodobnosti dostáváme

$$P(H) = \sum_{i=1}^6 P(V_i)P(H|V_i) = \frac{1}{6} \cdot 2 \cdot \frac{1}{4} \cdot \frac{3}{4} + \frac{2}{6} \cdot 2 \cdot \frac{1}{3} \cdot \frac{2}{3} + \frac{3}{6} \cdot 2 \cdot \frac{1}{2} \cdot \frac{1}{2} =$$

$$\frac{1}{16} + \frac{4}{27} + \frac{1}{4} = \frac{199}{432}.$$

- (b) Z náhodně vybrané truhly jsme dvakrát táhli s vracením a vytáhli jednu zlato a jednu stříbrnou minci (bez ohledu na pořadí). S jakou pravděpodobností to byla truhla s alespoň dvěma zlatými mincemi?

*Z Bayesovy věty (Věta 1.15) máme (použijeme stejnou notaci jako v předchozím příkladu)  $P(V_i|H) = P(H|V_i) \frac{P(V_i)}{P(H)}$ . Chceme spočítat*

$$\sum_{i=1}^3 P(V_i|H) = \sum_{i=1}^3 P(H|V_i) \frac{P(V_i)}{P(H)} = \frac{1}{P(H)} (P(H|V_1)P(V_1) +$$

$$2P(H|V_2)P(V_2)) = \frac{432}{199} \left( 2 \cdot \frac{1}{4} \cdot \frac{3}{4} \cdot \frac{1}{6} + 2 \cdot 2 \cdot \frac{1}{3} \cdot \frac{2}{3} \cdot \frac{1}{6} \right) = \frac{91}{199}.$$

- (c) Náhodně vybereme tři truhly a z každé náhodně vytáhneme po jedné minci. S jakou pravděpodobností bude alespoň jedna z nich zlatá?

*Spočteme nejdříve pravděpodobnost opačného jevu, tedy události, že budou vytaženy tři stříbrné mince. Označme  $V_{(i,j,k)}$  pravděpodobnost, že bylo náhodně zvoleno  $i$  truhel se třemi zlatými mincemi,  $j$  truhel se dvěma a  $k$  truhel s jednou zlatou mincí. Pravděpodobnost tohoto jevu se spočte pomocí vzorce*

$$P(V_{(i,j,k)}) = \frac{\binom{1}{i} \binom{2}{j} \binom{3}{k}}{\binom{6}{3}} = \frac{\binom{2}{j} \binom{3}{k}}{20}.$$

*Dále označme  $S$  jev, že byly vytaženy 3 stříbrné mince. Potom*

$$P(S) = \sum_{i,j,k} P(V_{(i,j,k)}) \frac{1}{4^i} \cdot \frac{1}{3^j} \cdot \frac{1}{2^k}.$$

*Možné hodnoty  $i, j, k$  a příslušné pravděpodobnosti  $P(V_{(i,j,k)})$  zapíšeme do tabulky.*

$i$	$j$	$k$	$P(V_{(i,j,k)})$	sčítanec v $P(S)$
1	2	0	1/20	1/720
1	1	1	6/20	1/80
1	0	2	3/20	3/320
0	0	3	1/20	1/160
0	2	1	3/20	1/120
0	1	2	6/20	1/40

Potom  $P(S) = \frac{181}{2880}$  a tedy hledaná pravděpodobnost vytažení alespoň jedné zlaté mince je  $P(S^C) = 1 - P(S) = \frac{2699}{2880}$ .

**Příklad B.2.** Náhodná veličina  $X$  má spojité rozdělení s hustotou

$$f(x) = \begin{cases} 2e^{-2x}, & x \geq 0; \\ 0, & \text{jinak.} \end{cases}$$

Definujme  $U := \lfloor X \rfloor$  a  $V := X - \lfloor X \rfloor$  (jinými slovy spodní celá část a frakcionální část  $X$ ).

(a) Určete rozdělení náhodné veličiny  $U$ .

*Jedná se o diskrétní náhodnou veličinu s pravděpodobnostní funkcí*

$$P[U = u] = \int_u^{u+1} f(x) dx = [-e^{-2x}]_u^{u+1} = e^{-2u}(1 - e^{-2})$$

pro  $u \geq 0$ .

(b) Určete rozdělení náhodné veličiny  $V$ .

*Tentokrát máme spojitou náhodnou veličinu, spočteme její distribuční funkci. Nechť  $v \in (0, 1)$ , potom*

$$\begin{aligned} F_V(v) = P[V \leq v] &= \sum_{u=0}^{\infty} \int_0^v f(u+t) dt = \sum_{u=0}^{\infty} \int_0^v 2e^{-2u} e^{-2t} dt = \\ &= \sum_{u=0}^{\infty} e^{-2u}(1 - e^{-2v}) = \frac{1 - e^{-2v}}{1 - e^{-2}}. \end{aligned}$$

Pro  $v \leq 0$  máme  $F_V(v) = 0$ , pro  $v \geq 1$  máme  $F_V(v) = 1$ .

(c) Spočtete  $\mathbb{E}(Xe^{-X} - 1)$ .

*S využitím pravidla líného statistika (Věta 3.5) dostáváme (transformace  $t(x) = xe^{-x} - 1$ )*

$$\mathbb{E}t(X) = \int_0^{\infty} t(x)f(x)dx = \int_0^{\infty} (xe^{-x} - 1)2e^{-2x}dx = \frac{2}{9} - 1 = -\frac{7}{9}.$$

**Příklad B.3.** Náhodný vektor  $(X, Y)^T$  má spojité rozdělení s hustotou

$$f(x, y) = \begin{cases} 2x, & 0 < x < 1, -x^2 < y < x^2; \\ 0, & \text{jinak.} \end{cases}$$

(a) Určete  $P(0 < 2Y < X^2)$ .

*Počítáme obsah útvaru mezi osou  $x$  a křivkou  $y = \frac{1}{2}x^2$  s hustotou  $f(x, y)$ . Máme integrál*

$$\int_0^1 \int_0^{\frac{1}{2}x^2} 2xdydx = \int_0^1 x^3 dx = \frac{1}{4}.$$

- (b) Určete marginální rozdělení a střední hodnotu veličiny  $X$ .  
*Hustotu  $f_X$  získáme zintegrováním hustoty  $f(x, y)$  přes všechny možné hodnoty  $y$ , tedy*

$$f_X(x) = \int_{-x^2}^{x^2} 2x dy = 4x^3.$$

*Dále platí*

$$\mathbb{E}X = \int_0^1 x f_X(x) dx = \int_0^1 4x^4 dx = \frac{4}{5}.$$

- (c) Spočtěte kovarianci veličin  $X$  a  $W$ , kde  $W = X^2Y$ .  
*Platí, že  $\text{Cov}(X, X^2Y) = E[X^3Y] - E[X]E[X^2Y]$ , tedy potřebujeme dopočítat chybějící střední hodnoty*

$$E[X^3Y] = \int_0^1 \int_{-x^2}^{x^2} x^3 y \cdot 2x dy dx = 0;$$

$$E[X^2Y] = \int_0^1 \int_{-x^2}^{x^2} x^2 y \cdot 2x dy dx = 0,$$

*(v obou případech je integrand lichý v proměnné  $y$ , a tedy integrál přes interval  $[-a, a]$  je roven nule). Vychází  $\text{Cov}(X, X^2Y) = 0 - 0 = 0$ .*

- (d) Rozhodněte, zda jsou veličiny  $X$  a  $Y$  nezávislé a své rozhodnutí zdůvodněte.  
*Veličiny  $X$  a  $Y$  nejsou nezávislé, neboť  $P[0 < X < \frac{1}{2}, Y > \frac{1}{4}] = 0 \neq P[0 < X < \frac{1}{2}]P[Y > \frac{1}{4}]$  (součin zřejmě nenulových hodnot).*