

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 Ω 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 zadanými 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 Ω , 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 Ω , které budeme říkat σ -algebra.

Definice 1.3. Necht $\Omega \neq \emptyset$ je množina a $\mathcal{A} \subset 2^\Omega$ soubor jejích podmnožin. Této množině \mathcal{A} říkáme σ -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 (Ω, \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 (Ω, \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 (Ω, \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 (Ω, \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 Ω 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. Házáme 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 Ω 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 Ω 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 , $P(B|A_3) = 0.01$ ¹. 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. □

¹Tyto 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 (Ω, \mathcal{A}) je měřitelný prostor. *Náhodná veličina* je měřitelné zobrazení, které přiřazuje každému výsledku ω 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 ω . 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 (Ω, \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 μ, ν dvě σ -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 μ -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 λ ,
- Čítačí 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 μ , kde μ je σ -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 μ .

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 σ -konečné míře μ , 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 (Ω, \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 λ , 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 aditivity 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 (v prázdné množině), 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 μ_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 λ , $P_{X_{ds}}$ (diskrétní spojitá) je absolutně spojitá vůči číselní míře μ na nějaké spočetné podmnožině \mathbb{R} a nakonec $P_{X_{sg}}$ (singulární) není absolutně spojitá vůči λ ani ji nelze napsat jako spočetnou kombinaci Diracových měr δ_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 δ_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 λ .

Pro spojitou náhodnou veličinu X vždy existuje hustota f_X (nezáporná a jednoznačná až na množinu λ -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

Příklad 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 $F_X(x) = \chi_{\{x \geq a\}}$.

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

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

Příklad 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).

Příklad 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 nezávislých² veličin s Bernoulliho rozdělením (počet úspěchů mezi n pokusy).

Příklad 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.

²Přesná definice nezávislých veličin bude uvedena později.

Příklad 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 $NB(1, p) = Geo(p)$.

Příklad 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

Příklad 2.24 (Spojité 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).

Příklad 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

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

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

Příklad 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 $X + Y \sim \Gamma(a, p_X + p_Y)$.

Příklad 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)