

Pravděpodobnost a statistika

Studijní materiál ke státnicím (M5)

Jakub Klesa

Pokryté předměty: NDMI002 (Diskrétní matematika), NMAI059 (Pravděpodobnost a statistika 1)

Obsah

Úvod	2
1 Pravděpodobnostní prostor, jevy, podmíněná pravděpodobnost, Bayes	2
1.1 Pravděpodobnostní prostor	2
1.2 Pravidla pro počítání s pravděpodobností	3
1.3 Podmíněná pravděpodobnost a nezávislost	3
1.4 Bayesův vzorec a věta o úplné pravděpodobnosti	4
2 Náhodné veličiny, střední hodnota, rozptyl, nerovnosti	5
2.1 Náhodné veličiny a distribuční funkce	5
2.2 Střední hodnota a její vlastnosti	6
2.3 Rozptyl a kovariance	8
3 Konkrétní rozdělení	9
3.1 Bernoulliho a binomické	9
3.2 Geometrické	9
3.3 Poissonovo	10
3.4 Exponenciální	10
3.5 Normální (Gaussovo)	11
4 Limitní věty: zákon velkých čísel a centrální limitní věta	12
4.1 Zákon velkých čísel	12
4.2 Centrální limitní věta	13
5 Bodové odhady	14
5.1 Vlastnosti odhadů	15
5.2 Metoda maximální věrohodnosti (MLE)	15
6 Intervalové odhady	16
6.1 CI pro střední hodnotu při známém σ	16
6.2 CI pro Bernoulliho parametr	17
7 Testování hypotéz	18
7.1 Základní rámec	18
7.2 p-hodnota	19
7.3 Příklad: test férovosti mince	19
7.4 Test dobré shody (chí-kvadrát) – nad rámec požadavků	20
Shrnutí	20

Úvod

Tento materiál pokrývá okruh **M5 – Pravděpodobnost a statistika** (předmět NMAI059 *Pravděpodobnost a statistika 1*, kombinatorické základy z NDMI002 *Diskrétní matematika*). Výklad stojí na skriptech R. Šámala k NMAI059 (verze LS 2024/25) a doplňkově na řešených cvičeních K. Krále, M. Mareše a R. Šámala (NMAI059, 2021), odkud čerpáme typové úlohy a tahákovou formu. Značení a hloubka odpovídají kurzu pro informatiky.

Věty uvádíme *se zněním*, doplněné intuicí; krátké důkazy, které pomáhají pochopení, jsou označeny jako *nepovinné*. Centrální limitní větu uvádíme bez důkazu (znění a aplikace), zákon velkých čísel se zněním a krátkým odvozením z Čebyševovy nerovnosti. Do výkladu jsou zařazeny obě řešené reálné otázky SZZ k okruhu (léto 2023: linearita střední hodnoty, léto 2025: centrální limitní věta); pohromadě jsou i v samostatném dokumentu *Příklady otázek*. Použité zdroje jsou uvedeny na konci.

Značení. Pravděpodobnostní prostor je trojice $(\Omega, \mathcal{F}, P)$, kde Ω je množina *elementárních jevů*, $\mathcal{F} \subseteq 2^\Omega$ je *prostor jevů* (σ -algebra) a $P: \mathcal{F} \rightarrow [0, 1]$ je *pravděpodobnost*. Jevy značíme $A, B, C \in \mathcal{F}$, doplněk jevu $A^c = \Omega \setminus A$. Náhodné veličiny značíme X, Y, Z (písmenem $X(\omega)$, ale skoro vždy píšeme jen X). Distribuční funkce: $F_X(x) = P(X \leq x)$. Diskrétní *pravděpodobnostní funkce* (PMF): $p_X(k) = P(X = k)$. Spojitá *hustota* (PDF): $f_X(x)$. Střední hodnota: $E(X)$, rozptyl: $\text{var}(X)$, směrodatná odchylka: $\sigma(X) = \sqrt{\text{var}(X)}$, kovariance: $\text{cov}(X, Y)$. Standardní rozdělení: $X \sim \text{Bern}(p)$, $\text{Bin}(n, p)$, $\text{Geom}(p)$, $\text{Pois}(\lambda)$, $\text{Exp}(\lambda)$, $N(\mu, \sigma^2)$. Symbolem Φ rozumíme distribuční funkci standardního normálního rozdělení $N(0, 1)$.

 **Video:** Playlist Essence of probability (neoficiální výběr videí kanálu 3Blue1Brown; série dává vizuální intuici k pravděpodobnosti; na konkrétní díly odkazujeme u jednotlivých témat níže).

1 Pravděpodobnostní prostor, jevy, podmíněná pravděpodobnost, Bayes

Začneme od modelu náhodnosti – pravděpodobnostního prostoru – a od základních pravidel pro počítání s P . Pak zavedeme *podmíněnou pravděpodobnost*, *nezávislost* a *Bayesův vzorec*, který umí „obrátit“ podmiňování.

1.1 Pravděpodobnostní prostor

Definice – Pravděpodobnostní prostor

Pravděpodobnostní prostor je trojice $(\Omega, \mathcal{F}, P)$, kde:

- (1) $\Omega \neq \emptyset$ je libovolná množina, prvky se nazývají *elementární jevy* (výsledky).
- (2) $\mathcal{F} \subseteq 2^\Omega$ je *prostor jevů* (σ -algebra): obsahuje $\emptyset$ a Ω , je uzavřený na doplněk a na spočetná sjednocení.
- (3) $P: \mathcal{F} \rightarrow [0, 1]$ je *pravděpodobnost*: $P(\Omega) = 1$ a pro každou posloupnost po dvou disjunktních jevů $A_1, A_2, \dots$ platí $P(\bigcup_i A_i) = \sum_i P(A_i)$ (*spočetná aditivita*).

 **Intuice.** Množina Ω popisuje, „co se může stát“: pro hod kostkou $\Omega = \{1, \dots, 6\}$, pro hod šipkou na terč Ω jednotkový kruh v rovině, pro nekonečnou posloupnost hodů mincí $\Omega = \{0, 1\}^{\mathbb{N}}$. Prostor jevů říká, „čemu umíme přiřadit pravděpodobnost“: pro konečné nebo spočetné Ω obvykle $\mathcal{F} = 2^\Omega$ (všechny podmnožiny). Pro $\Omega = \mathbb{R}$ se musíme omezit na *borelovskou σ -algebru* – ne každé podmnožině $\mathbb{R}$ dovedeme přiřadit délku konzistentně. Spočetná aditivita je hlavní axiom: dovoluje počítat s nekonečnými posloupnostmi jevů (např. sečíst $P(X = k)$ přes všechna $k \in \mathbb{N}$ u geometrického rozdělení nebo přejít k limitě v $\lim_{x \rightarrow \infty} F_X(x) = 1$), což by samotná *konečná* aditivita nezaručila.

i Poznámka. Klasický a geometrický prostor. Když je $|\Omega| < \infty$ a všechny výsledky jsou *stejně pravděpodobné* („symetrické“), říkáme prostoru *klasický* a $P(A) = |A|/|\Omega|$. Když $\Omega \subseteq \mathbb{R}^d$ má konečnou kladnou míru (délku, obsah, objem) a pravděpodobnost je úměrná míře, mluvíme o *geometrickém* prostoru a $P(A) = m(A)/m(\Omega)$.

1.2 Pravidla pro počítání s pravděpodobnostmi

Z axiomů snadno odvodíme základní pravidla, která neustále používáme:

★ Věta – Základní pravidla

V pravděpodobnostním prostoru $(\Omega, \mathcal{F}, P)$ pro libovolné $A, B \in \mathcal{F}$ platí:

- (1) $P(\emptyset) = 0$ a $P(A^c) = 1 - P(A)$.
- (2) Monotonie: $A \subseteq B \Rightarrow P(A) \leq P(B)$.
- (3) Sjednocení dvou: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
- (4) Subaditivita (Booleova nerovnost): $P(\bigcup_i A_i) \leq \sum_i P(A_i)$.

Důkaz (nepovinný – vše plyne z disjunktí aditivity).

- (1) $P(\emptyset) = 0$: ze spočetné aditivity pro $\emptyset = \emptyset \cup \emptyset \cup \dots$ je $P(\emptyset) = \sum_{i=1}^{\infty} P(\emptyset)$, což je konečné jen pokud $P(\emptyset) = 0$. Dále $\Omega = A \cup A^c$ je disjunktí rozklad, takže $1 = P(\Omega) = P(A) + P(A^c)$.
- (2) Pro $A \subseteq B$ rozkládáme $B = A \cup (B \setminus A)$ disjunktí, takže $P(B) = P(A) + P(B \setminus A) \geq P(A)$.
- (3) Rozkládáme $A \cup B = (A \setminus B) \cup (A \cap B) \cup (B \setminus A)$ disjunktí; sečtením $P(A) + P(B)$ přes tentýž rozklad dostaneme $P(A \cup B) + P(A \cap B)$.
- (4) *Zdisjunktění*: pro $B_i = A_i \setminus \bigcup_{j < i} A_j$ je $\bigcup_i A_i = \bigsqcup_i B_i$ a $P(B_i) \leq P(A_i)$ (monotonie); spočetnou aditivitou pak $P(\bigcup_i A_i) = \sum_i P(B_i) \leq \sum_i P(A_i)$.

□

★ Věta – Princip inkluze a exkluze

Pro jevy $A_1, \dots, A_n \in \mathcal{F}$

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{\emptyset \neq S \subseteq [n]} (-1)^{|S|+1} P\left(\bigcap_{i \in S} A_i\right).$$

i Intuice. Pro $n = 2$ je to pravidlo (3) výše. Obecný tvar přidává $\pm$ korekce na vícenásobná překrytí – analogie inkluze-exkluze pro velikosti množin (viz okruh M3). Detailní rozbor v M3.

1.3 Podmíněná pravděpodobnost a nezávislost

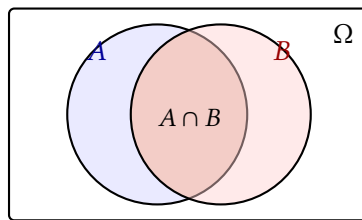
Když pozorujeme nějakou informaci B („vím, že nastalo B “), zúžíme prostor na B a přeškálujeme pravděpodobnosti tak, aby v něm znovu daly 1.

📖 Definice – Podmíněná pravděpodobnost

Pro jevy $A, B \in \mathcal{F}$ s $P(B) > 0$ definujeme

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

i Poznámka. Funkce $A \mapsto P(A | B)$ je opět pravděpodobnost na $(\Omega, \mathcal{F})$. Pro $P(B) = 0$ se podmíněná pravděpodobnost touto definicí nezavádí.



Podmínění jevem B „zúží“ výběrový prostor na B (jen červený kruh). $P(A | B)$ je podíl plochy průniku $A \cap B$ vůči ploše celého B , tj. jaká část B leží zároveň v A .

Definice – Nezávislost jevů

Jevy A, B jsou *nezávislé*, pokud $P(A \cap B) = P(A)P(B)$. Pro $P(B) > 0$ to je ekvivalentní s $P(A | B) = P(A)$. Konečná kolekce jevů $A_1, \dots, A_n$ je (*navzájem*) *nezávislá*, pokud pro každou podmnožinu $S \subseteq [n]$ platí $P(\bigcap_{i \in S} A_i) = \prod_{i \in S} P(A_i)$.

 **Intuice.** Pozor: *po dvojicích nezávislé* $\neq$ *navzájem nezávislé*. Klasický protipříklad: dva hody férovou mincí; A = „v prvním hodu padl orel“, B = „ve druhém hodu padl orel“, C = „padly různé strany“. Každý z jevů má pravděpodobnost $1/2$ a každý průnik dvou z nich $1/4$, takže po dvojicích jsou nezávislé, ale $P(A \cap B \cap C) = 0 \neq 1/8 = P(A)P(B)P(C)$.

1.4 Bayesův vzorec a věta o úplné pravděpodobnosti

Často chceme „obrátit“ podmiňování: známe $P(\text{příznak} | \text{příčina})$ a hledáme $P(\text{příčina} | \text{příznak})$. To je Bayes.

 **Video:** 3Blue1Brown, díl Bayes theorem, the geometry of changing beliefs (*geometrický pohled na Bayese přes plochy jevů; chytře vysvětluje i lékařský test*).

★ Věta – Věta o úplné pravděpodobnosti

Nechť jevy $B_1, \dots, B_n$ tvoří rozklad Ω (po dvou disjunktní, $\bigcup_i B_i = \Omega$, všechny s $P(B_i) > 0$). Pak pro libovolný jev A

$$P(A) = \sum_{i=1}^n P(A | B_i) P(B_i).$$

★ Věta – Bayesův vzorec

Za stejných předpokladů a pro A s $P(A) > 0$ platí pro každé j

$$P(B_j | A) = \frac{P(A | B_j) P(B_j)}{\sum_{i=1}^n P(A | B_i) P(B_i)} = \frac{P(A | B_j) P(B_j)}{P(A)}.$$

 **Intuice.** Čítatel: „příznak a hypotéza B_j nastaly současně“, $P(A \cap B_j) = P(A | B_j)P(B_j)$. Jmenovatel: „příznak nastal vůbec“, rozložený přes všechny hypotézy. *Apriorní* pravděpodobnost $P(B_j)$ se přenásobí *věrohodností* $P(A | B_j)$ a normuje $P(A)$ – dostaneme *aposteriorní* pravděpodobnost $P(B_j | A)$.

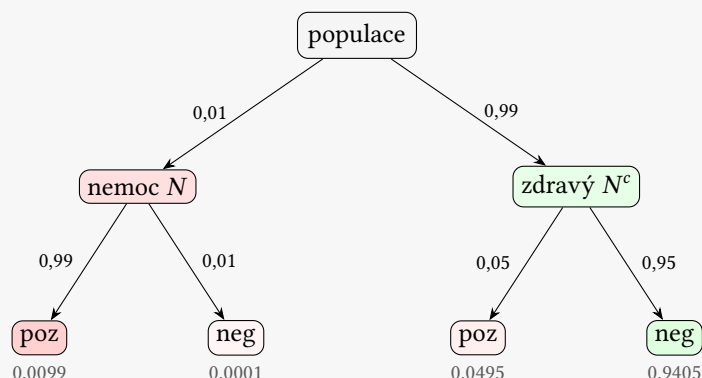
Příklad – Test nemoci

Zadání. V populaci má nemoc prevalenci $P(N) = 1\%$. Test má senzitivitu $P(\text{poz} | N) = 0,99$ a falešnou pozitivitu $P(\text{poz} | N^c) = 0,05$. Jaká je pravděpodobnost, že pacient s pozitivním výsledkem opravdu nemocí trpí?

Řešení. Bayes:

$$P(N | \text{poz}) = \frac{P(\text{poz} | N)P(N)}{P(\text{poz} | N)P(N) + P(\text{poz} | N^c)P(N^c)} = \frac{0,99 \cdot 0,01}{0,99 \cdot 0,01 + 0,05 \cdot 0,99} = \frac{0,0099}{0,0594} \approx \boxed{0,167}.$$

Tedy jen $\approx 16,7\%$ pozitivně testovaných je skutečně nemocných – nízká priora „zdědí“ většinu pozitivních výsledků z falešné pozitivity.



Bayesův strom pro test nemoci. První úroveň: *priorní* pravděpodobnosti $P(N) = 0,01$ vs. $P(N^c) = 0,99$. Druhá úroveň: *podmíněné* $P(\text{poz} | N) = 0,99$, $P(\text{poz} | N^c) = 0,05$. Hodnoty u listů jsou součiny podél cesty, tj. $P(N \cap \text{poz}) = 0,0099$ a $P(N^c \cap \text{poz}) = 0,0495$. Sečtením přes „poz“ listy: $P(\text{poz}) = 0,0594$. Bayes: $P(N | \text{poz}) = 0,0099/0,0594 \approx 0,167$.

Video: 3Blue1Brown, díl The medical test paradox, and redesigning Bayes' rule (proč i velmi přesný test dává při nízké prevalenci hodně falešně pozitivních – přesně efekt z příkladu výše; zavádí i tvar Bayese přes šance).

2 Náhodné veličiny, střední hodnota, rozptyl, nerovnosti

Náhodná veličina je číselná informace odvozená z náhodného experimentu – formálně funkce $X : \Omega \rightarrow \mathbb{R}$. Rozdělení n.v. popisujeme buď *distribuční funkcí* (univerzálně), nebo *pravděpodobnostní funkcí* (diskrétní případ), nebo *hustotou* (spojitý případ).

2.1 Náhodné veličiny a distribuční funkce

Definice – Náhodná veličina, distribuční funkce

Náhodná veličina (n.v.) na prostoru $(\Omega, \mathcal{F}, P)$ je zobrazení $X : \Omega \rightarrow \mathbb{R}$ takové, že pro každé $x \in \mathbb{R}$ je $\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$. Distribuční funkce je $F_X(x) = P(X \leq x)$.

★ Věta – Vlastnosti distribuční funkce

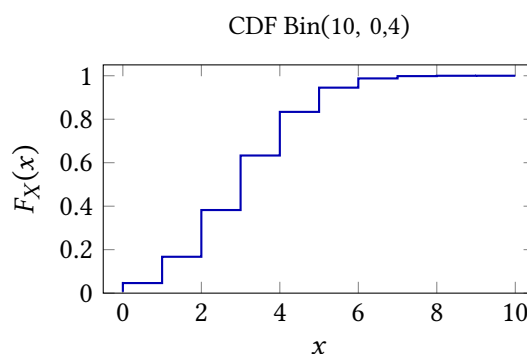
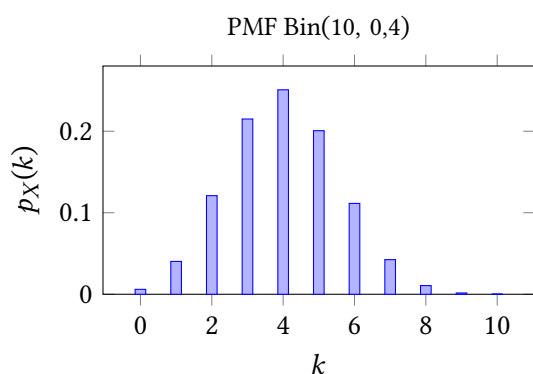
F_X je neklesající, zprava spojitá, $\lim_{x \rightarrow -\infty} F_X(x) = 0$, $\lim_{x \rightarrow +\infty} F_X(x) = 1$. Navíc $P(a < X \leq b) = F_X(b) - F_X(a)$ pro $a < b$.

Definice – Diskrétní a spojitá n.v.

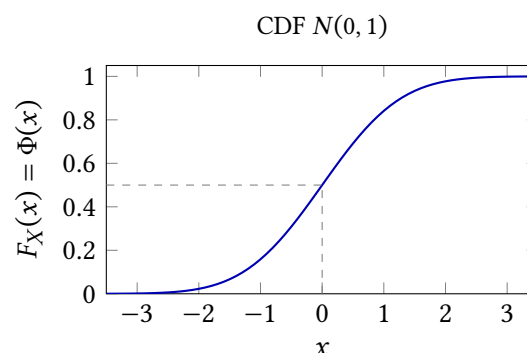
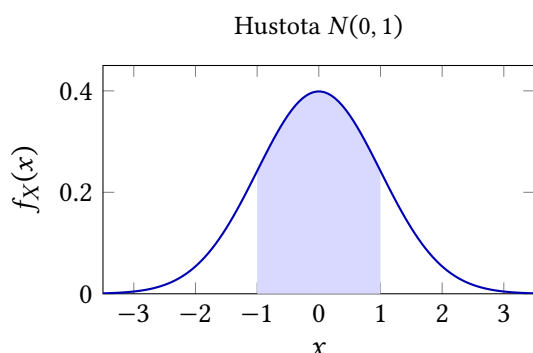
N.v. X je *diskrétní*, pokud nabývá nejvýše spočetně mnoha hodnot $x_1, x_2, \dots$. N.v. X je *spojitá*, pokud existuje nezáporná *hustota* f_X taková, že $F_X(x) = \int_{-\infty}^x f_X(t) dt$ pro všechna x . *Pravděpodobnostní funkce* diskrétní n.v. je $p_X(k) = P(X = k)$.

📺 **Video:** 3Blue1Brown, díl Why “probability of 0” does not mean “impossible” (proč u spojité veličiny $P(X = x) = 0$ a přesto má hustota smysl).

💡 **Intuice.** F_X je univerzální popis – funguje stejně pro diskrétní i spojité případ. Pro diskrétní je F_X schodovitá: skoky v bodech x_k o výšku $p_X(x_k)$. Pro spojitý je F_X hladká a $F'_X(x) = f_X(x)$ skoro všude. Hustota *není* pravděpodobnost: $f_X(x)$ může být > 1 ; pravděpodobnost je až $\int_a^b f_X(x) dx = P(a \leq X \leq b)$.



Diskrétní rozdělení Bin(10, 0,4). **Vlevo:** PMF $p_X(k) = \binom{10}{k} 0,4^k 0,6^{10-k}$, vrchol u $k = 4 = np$ (hodnota $\approx 0,251$). **Vpravo:** CDF je schodovitá; v každém celočíselném bodě skok o výšku $p_X(k)$, mezi nimi konstantní. $F_X(10) = 1$.



Standardní normální rozdělení $N(0,1)$. **Vlevo:** hustota $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$; modře vybarvená oblast $[-1,1]$ má míru $\Phi(1) - \Phi(-1) \approx 0,683$ (pravidlo 1σ). **Vpravo:** CDF Φ je hladká S-křivka, $\Phi(0) = 1/2$ (čárkované).

2.2 Střední hodnota a její vlastnosti

📖 Definice – Střední hodnota

Pro diskrétní n.v. X s hodnotami $\{x_i\}$ definujeme $E(X) = \sum_i x_i p_X(x_i)$, existuje-li suma absolutně. Pro spojitou n.v. s hustotou f_X $E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$, existuje-li integrál absolutně. Obecně pro $g: \mathbb{R} \rightarrow \mathbb{R}$ $E(g(X)) = \sum_i g(x_i) p_X(x_i)$ resp. $\int_{-\infty}^{\infty} g(x) f_X(x) dx$.

★ Věta – Linearita střední hodnoty

Pro libovolné n.v. X, Y (na též prostoru), které mají střední hodnotu, a konstanty $a, b \in \mathbb{R}$ platí

$$E(aX + bY) = a E(X) + b E(Y).$$

Speciálně $E(X + Y) = E(X) + E(Y)$ bez ohledu na nezávislost a $E(c) = c$ pro konstantní n.v.

Intuice. Linearita je jeden z *nejsilnějších nástrojů* kombinatorické pravděpodobnosti: chceme-li $E(X)$ pro složitou veličinu, rozložíme ji na součet jednoduchých (často *indikátorů* jevů, I_A) a sečteme střední hodnoty. Indikátor má $E(I_A) = P(A)$, takže „počet úspěchů“ = součet indikátorů a $E(\text{počet úspěchů}) = \sum P(\text{úspěch}_i)$.

Definice – Nezávislost náhodných veličin

N.v. X, Y jsou *nezávislé*, pokud jsou jevy $\{X \leq x\}$ a $\{Y \leq y\}$ nezávislé pro všechna $x, y \in \mathbb{R}$, tj. $P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$. Pro diskrétní n.v. je to ekvivalentní s $P(X = x, Y = y) = P(X = x)P(Y = y)$ pro všechna x, y . Veličiny $X_1, \dots, X_n$ jsou nezávislé, pokud analogicky $P(X_1 \leq x_1, \dots, X_n \leq x_n) = \prod_i P(X_i \leq x_i)$ pro všechna $x_1, \dots, x_n$.

★ Věta – Střední hodnota součinu nezávislých n.v.

Pro *nezávislé* n.v. X, Y s konečnou střední hodnotou platí

$$E(XY) = E(X)E(Y).$$

Obecně pro nezávislé $X_1, \dots, X_n$ je $E(\prod_i X_i) = \prod_i E(X_i)$.

Poznámka. Bez předpokladu nezávislosti tvrzení obecně neplatí. Příklad: $X \in \{-1, +1\}$ stejně pravděpodobné, $Y = X$. Pak $E(X) = E(Y) = 0$, ale $E(XY) = E(X^2) = 1 \neq 0$.

★ Věta – Markovova nerovnost

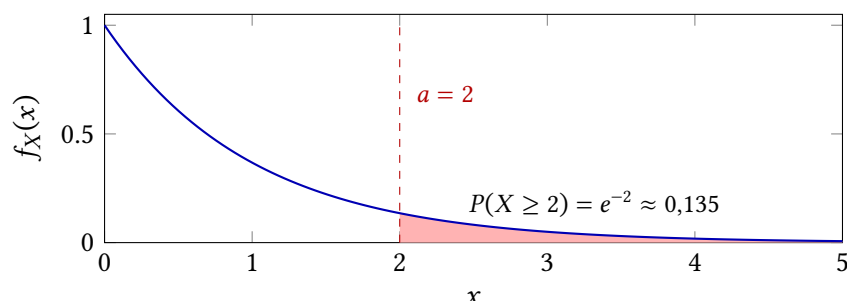
Pro nezápornou n.v. $X \geq 0$ a $a > 0$

$$P(X \geq a) \leq \frac{E(X)}{a}.$$

Důkaz (nepovinný – jeden řádek). Z $X \geq a \cdot I_{\{X \geq a\}}$ (na jevu $\{X \geq a\}$ je $X \geq a$, jinde $X \geq 0$) plyne $E(X) \geq a P(X \geq a)$. $\square$

Intuice. Markov dává *velmi hrubý* chvostový odhad pomocí jediného čísla $E(X)$ – bez znalosti rozptylu. Často nestačí (např. pro $a = 2E(X)$ dostaneme jen „nejvýše polovina“, což je málo), ale je univerzální. Když známe rozptyl, dostaneme silnější Čebyševovu nerovnost níže.

Markovova nerovnost: $P(X \geq a) \leq E(X)/a$ pro $X \sim \text{Exp}(1)$, $a = 2$



Markovova nerovnost pro $X \sim \text{Exp}(1)$, $E(X) = 1$, $a = 2$. Markovův odhad: $P(X \geq 2) \leq 1/2 = 0,5$. Skutečná hodnota: $P(X \geq 2) = \int_2^\infty e^{-x} dx = e^{-2} \approx 0,135$ (vybarvená plocha). Markov tedy nadhodnocuje téměř 4× – ukazuje, že je velmi hrubý.

2.3 Rozptyl a kovariance

Definice – Rozptyl, směrodatná odchylka, kovariance

Pro n.v. X s konečnou $E(X) = \mu$ definujeme

$$\text{var}(X) = E((X - \mu)^2) = E(X^2) - \mu^2, \quad \sigma(X) = \sqrt{\text{var}(X)}.$$

Pro dvojici n.v. X, Y

$$\text{cov}(X, Y) = E((X - E(X))(Y - E(Y))) = E(XY) - E(X)E(Y).$$

★ Věta – Vlastnosti rozptylu

- (1) $\text{var}(aX + b) = a^2 \text{var}(X)$ pro $a, b \in \mathbb{R}$ (posunutí neovlivní rozptyl, škálování čtvercově).
- (2) $\text{var}(X + Y) = \text{var}(X) + \text{var}(Y) + 2 \text{cov}(X, Y)$.
- (3) Obecně pro $X_1, \dots, X_n$ $\text{var}(\sum_i X_i) = \sum_i \text{var}(X_i) + 2 \sum_{i < j} \text{cov}(X_i, X_j)$.
- (4) Pro *nezávislé* n.v. je $\text{cov}(X_i, X_j) = 0$, takže $\text{var}(\sum_i X_i) = \sum_i \text{var}(X_i)$.

 **Intuice.** Bod (4) je nejčastěji používaný: pro průměr $\bar{X}_n = \frac{1}{n} \sum_i X_i$ nezávislých veličin se stejným rozptylem σ^2 máme $\text{var}(\bar{X}_n) = \sigma^2/n$. Rozptyl průměru tedy klesá jako $1/n$ a směrodatná odchylka jako $1/\sqrt{n}$. To je důvod, proč „více měření = lepší odhad“ – přesnost (směrodatná odchylka) se ale zlepšuje jen jako $1/\sqrt{n}$, ne jako $1/n$.

 **Poznámka.** $\text{cov}(X, Y) = 0$ *neimplikuje* nezávislost. Příklad: $X \sim N(0, 1)$, $Y = X^2$. Pak $\text{cov}(X, Y) = E(X^3) - E(X)E(X^2) = 0 - 0 = 0$, ale X a Y jsou *deterministicky závislé* (Y je funkcí X).

Z Markovovy nerovnosti aplikované na $(X - \mu)^2$ vyplyne přímo silnější chvostový odhad pomocí rozptylu. Čebyševovu nerovnost v požadavcích explicitně nenajdeme, ale tvoří *jediný krok* v důkazu zákona velkých čísel (sekce 4), proto ji zavádíme.

★ Věta – Čebyševova nerovnost

Pro n.v. X s konečnou střední hodnotou μ a rozptylem σ^2 a pro $k > 0$

$$P(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2}.$$

Důkaz (nepovinný – aplikace Markovovy). $P(|X - \mu| \geq k) = P((X - \mu)^2 \geq k^2) \leq E((X - \mu)^2)/k^2 = \sigma^2/k^2$. □

Linearitu střední hodnoty a práci s indikátory procvičuje následující reálná otázka SZZ.

Příklad – SZZ léto 2023 č. 4: Linearita střední hodnoty (hod kostkou ve skupině)

(Reálná otázka SZZ: léto 2023, č. 4 – Linearita střední hodnoty)

Zadání.

- (1) Uvedte větu o linearitě střední hodnoty (nedokazujte).
- (2) Ve skupině n lidí každý hodí férovou kostkou. Každá dvojice, která hodila stejné číslo, získá bod. Y = bod jedné předem vybrané dvojice. Určete $E(Y)$.
- (3) X = celkový počet bodů přes všechny dvojice. Určete $E(X)$.

Řešení.

- (1) Pro libovolné n.v. $X_1, \dots, X_n$ na též prostoru a konstanty $a_1, \dots, a_n, b \in \mathbb{R}$ platí $E(\sum_i a_i X_i + b) =$

$$\sum_i a_i E(X_i) + b.$$

- (2) $Y \in \{0, 1\}$ je indikátor jevu „vybraná dvojice hodila stejné číslo“. Tento jev nastává s pravděpodobností $P(Y = 1) = 6/36 = 1/6$ (z 36 stejně pravděpodobných výsledků dvou hodů má 6 shodu). Tedy $E(Y) = P(Y = 1) = \boxed{1/6}$.
- (3) $X = \sum_{\{i,j\}} Y_{ij}$, kde součet je přes všech $\binom{n}{2}$ dvojic a Y_{ij} je indikátor shody dvojice $\{i, j\}$. Z bodu (2) máme $E(Y_{ij}) = 1/6$ pro každou dvojici (veličiny Y_{ij} jsou sice po dvou nezávislé, ale jako celek ne: $Y_{12} = Y_{23} = 1$ vynutí $Y_{13} = 1$; linearita nezávislost nevyžaduje). Tedy

$$E(X) = \sum_{\{i,j\}} E(Y_{ij}) = \binom{n}{2} \cdot \frac{1}{6} = \boxed{\frac{n(n-1)}{12}}.$$

3 Konkrétní rozdělení

V tomto bloku zopakujeme šest standardních rozdělení: čtyři diskrétní (Bernoulliho, binomické, geometrické, Poissonovo) a dvě spojitá (exponenciální, normální). U každého uvedeme *kdy ho použít*, *PMF/hustotu*, *střední hodnotu a rozptyl* a *klíčové vztahy*.

3.1 Bernoulliho a binomické

Definice – Bern(p), Bin(n, p)

$X \sim \text{Bern}(p)$ pro $p \in [0, 1]$ je indikátor úspěchu s $P(X = 1) = p$, $P(X = 0) = 1 - p$. $X \sim \text{Bin}(n, p)$ pro $n \in \mathbb{N}$ je počet úspěchů v n *nezávislých* Bernoulliho pokusech, tedy $X = \sum_{i=1}^n X_i$, kde $X_i \stackrel{\text{iid}}{\sim} \text{Bern}(p)$. PMF:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad k = 0, 1, \dots, n.$$

★ Věta – Střední hodnota a rozptyl

$$\begin{aligned} E(\text{Bern}(p)) &= p, & \text{var}(\text{Bern}(p)) &= p(1 - p), \\ E(\text{Bin}(n, p)) &= np, & \text{var}(\text{Bin}(n, p)) &= np(1 - p). \end{aligned}$$

 **Intuice.** Vzorce stačí odvodit pro Bernoulli a sčítat: $E(X_i) = p$ (linearita), $\text{var}(X_i) = E(X_i^2) - p^2 = p - p^2 = p(1 - p)$. Pro Bin(n, p) pak $E = n \cdot p$ (linearita) a $\text{var} = n \cdot p(1 - p)$ (nezávislost).

 **Video:** 3Blue1Brown, díl Binomial distributions (vizualizace $\binom{n}{k} p^k (1 - p)^{n-k}$ a tvaru rozdělení).

3.2 Geometrické

Definice – Geom(p)

$X \sim \text{Geom}(p)$ pro $p \in (0, 1]$ je počet *nezávislých pokusů* potřebných k prvnímu úspěchu (počítáme i ten úspěšný). PMF:

$$P(X = k) = (1 - p)^{k-1} p, \quad k = 1, 2, 3, \dots$$

★ Věta – Střední hodnota a rozptyl geometrického

$$E(\text{Geom}(p)) = \frac{1}{p}, \quad \text{var}(\text{Geom}(p)) = \frac{1 - p}{p^2}.$$

💡 Intuice. Při pravděpodobnosti úspěchu p se na první úspěch průměrně čeká $1/p$ pokusů. Pro $p = 1/6$ (hod šestky) průměr 6. Geometrické má *paměťovou vlastnost* (memoryless): $P(X > m + k \mid X > m) = P(X > k)$ – dosavadní neúspěšné pokusy „nezvyšují šanci úspěchu příště“.

📌 Poznámka. Pozor na konvenci: někteří autoři definují $\text{Geom}(p)$ jako *počet neúspěchů před prvním úspěchem* (hodnoty $0, 1, 2, \dots$), což posune E o 1 dolů. Šámalova skripta i my používáme „počet pokusů včetně úspěšného“, tj. $E = 1/p$.

3.3 Poissonovo

📖 Definice – $\text{Pois}(\lambda)$

$X \sim \text{Pois}(\lambda)$ pro $\lambda > 0$ je počet výskytů události v daném intervalu, kde *parametr* λ je průměrný (očekávaný) počet výskytů za tento interval. PMF

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

udává pravděpodobnost *právě* k výskytů; k je tedy nezáporné celé číslo, na jehož pravděpodobnost se ptáme. Faktor $e^{-\lambda}$ je normalizační konstanta zajišťující $\sum_{k \geq 0} P(X = k) = 1$ (plyne z $\sum_{k \geq 0} \lambda^k / k! = e^\lambda$).

★ Věta – Střední hodnota a rozptyl Poissonova

$$E(\text{Pois}(\lambda)) = \lambda, \quad \text{var}(\text{Pois}(\lambda)) = \lambda.$$

💡 Intuice. Poisson modeluje „počet vzácných nezávislých událostí v daném intervalu“: počet zákazníků za hodinu, počet překlepů na stránce, počet emailů za den. Vzniká jako limita $\text{Bin}(n, \lambda/n)$ pro $n \rightarrow \infty$ – proto pro „ n velké, p malé, np rozumné“ aproximujeme $\text{Bin}(n, p)$ Poissonem se $\lambda = np$.

3.4 Exponenciální

📖 Definice – $\text{Exp}(\lambda)$

$X \sim \text{Exp}(\lambda)$ pro $\lambda > 0$ má hustotu

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

a CDF $F_X(x) = 1 - e^{-\lambda x}$ pro $x \geq 0$.

★ Věta – Střední hodnota a rozptyl exponenciálního

$$E(\text{Exp}(\lambda)) = \frac{1}{\lambda}, \quad \text{var}(\text{Exp}(\lambda)) = \frac{1}{\lambda^2}.$$

💡 Intuice. Exponenciální je *spojitý protějšek geometrického*: doba čekání na první událost v Poissonově procesu s intenzitou λ . Také má *paměťovou vlastnost*: $P(X > s + t \mid X > s) = P(X > t)$ – mezi všemi spojitými rozděleními je tím *charakterizováno*.

3.5 Normální (Gaussovo)

Definice – $N(\mu, \sigma^2)$

$X \sim N(\mu, \sigma^2)$ pro $\mu \in \mathbb{R}$, $\sigma^2 > 0$ má hustotu

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in \mathbb{R}.$$

Standardní normální je $N(0, 1)$ s hustotou $\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ a CDF $\Phi(x) = \int_{-\infty}^x \varphi(t) dt$.

★ Věta – Střední hodnota, rozptyl a lineární transformace normálního

$$E(N(\mu, \sigma^2)) = \mu, \quad \text{var}(N(\mu, \sigma^2)) = \sigma^2.$$

Pokud $X \sim N(\mu, \sigma^2)$, pak pro $a, b \in \mathbb{R}$, $a \neq 0$

$$aX + b \sim N(a\mu + b, a^2\sigma^2).$$

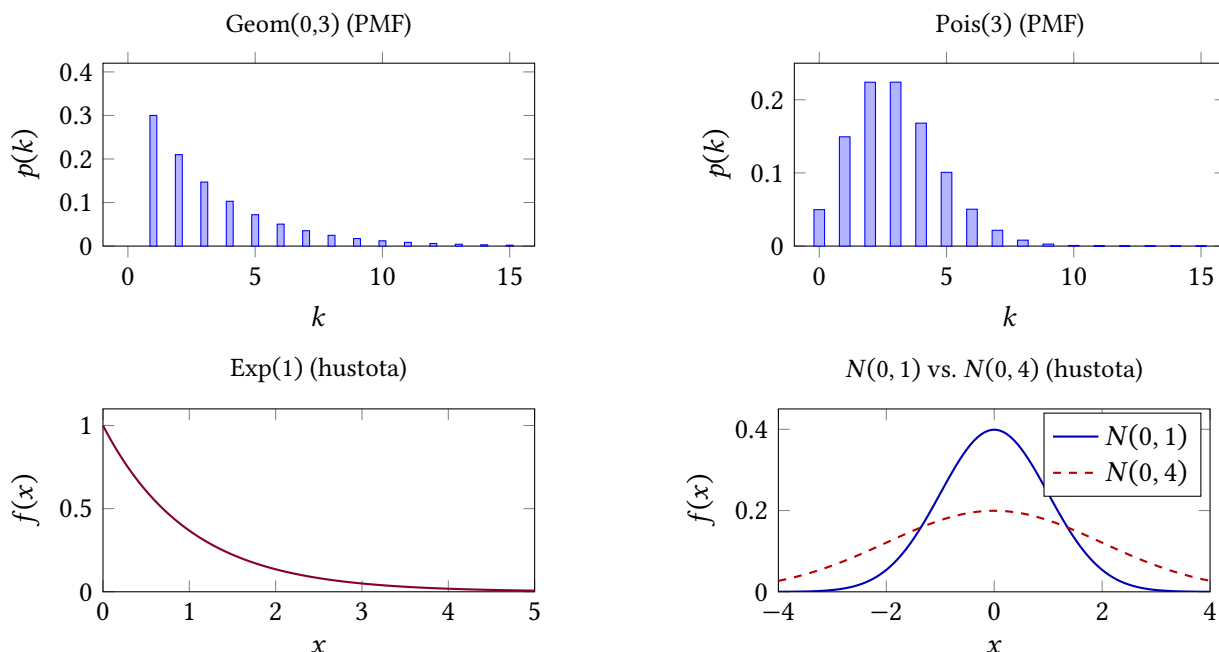
Speciálně $Z = (X - \mu)/\sigma \sim N(0, 1)$ – *standardizace*.

 **Intuice.** Normální rozdělení je všudypřítomné kvůli *centrální limitní větě* (níže) – součty mnoha nezávislých veličin se mu blíží. Pravidla $1\sigma/2\sigma/3\sigma$: $P(|X - \mu| < \sigma) \approx 0,683$, $P(|X - \mu| < 2\sigma) \approx 0,954$, $P(|X - \mu| < 3\sigma) \approx 0,997$. Hodnoty Φ se nevyjadřují v uzavřené formě a používá se *tabulka* nebo numerický výpočet.

 **Video:** 3Blue1Brown, díl Why π is in the normal distribution (odkud se v hustotě bere $\sqrt{2\pi}$ a proč je tvar právě $e^{-x^2/2}$).

Poznámka. Souhrnná tabulka.

Rozdělení	nosič	PMF / hustota	$E(X)$	$\text{var}(X)$
Bern(p)	$\{0, 1\}$	$p^k(1-p)^{1-k}$	p	$p(1-p)$
Bin(n, p)	$\{0, \dots, n\}$	$\binom{n}{k}p^k(1-p)^{n-k}$	np	$np(1-p)$
Geom(p)	$\{1, 2, \dots\}$	$(1-p)^{k-1}p$	$1/p$	$(1-p)/p^2$
Pois(λ)	$\{0, 1, 2, \dots\}$	$e^{-\lambda}\lambda^k/k!$	λ	λ
Exp(λ)	$[0, \infty)$	$\lambda e^{-\lambda x}$	$1/\lambda$	$1/\lambda^2$
$N(\mu, \sigma^2)$	$\mathbb{R}$	$\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$	μ	σ^2



Katalog standardních rozdělení. **Vlevo nahoře:** Geom(0,3) – exponenciálně klesající PMF, $E = 1/0,3 \approx 3,33$. **Vpravo nahoře:** Pois(3) – vrchol u $k = 2, 3$ (oba s $p \approx 0,224$), $E = \text{var} = 3$. **Vlevo dole:** Exp(1) – spojitě klesání, $E = 1$. **Vpravo dole:** $N(0, 1)$ vs. $N(0, 4)$ – širší rozdělení má větší rozptyl ($\sigma^2 = 4$ vs. $\sigma^2 = 1$); maximální výška $1/\sqrt{2\pi\sigma^2}$.

4 Limitní věty: zákon velkých čísel a centrální limitní věta

Dvě klíčové věty, které popisují chování průměru „velkého počtu“ nezávislých veličin. *Zákon velkých čísel* (ZVČ) říká *kam* průměr konverguje (k μ), *centrální limitní věta* (CLV) říká *jak rychle* (po normalizaci k $N(0, 1)$).

4.1 Zákon velkých čísel

★ Věta – Slabý zákon velkých čísel

Nechť $X_1, X_2, \dots$ jsou nezávislé stejně rozdělené n.v. (i.i.d.) s konečnou střední hodnotou μ a rozptylem σ^2 . Označme $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$. Pak pro každé $\varepsilon > 0$

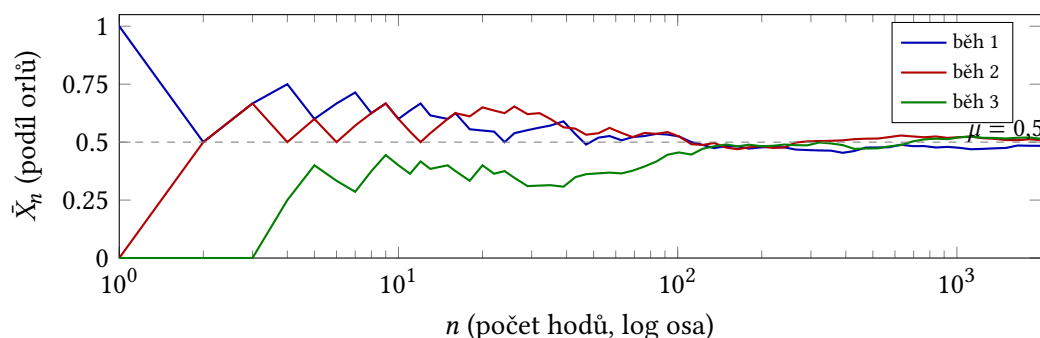
$$\lim_{n \rightarrow \infty} P(|\bar{X}_n - \mu| \geq \varepsilon) = 0.$$

Důkaz (nepovinný – aplikace Čebyševovy). $E(\bar{X}_n) = \mu$ (linearita), $\text{var}(\bar{X}_n) = \sigma^2/n$ (nezávislost). Čebyšev: $P(|\bar{X}_n - \mu| \geq \varepsilon) \leq \text{var}(\bar{X}_n)/\varepsilon^2 = \sigma^2/(n\varepsilon^2) \rightarrow 0$. $\square$

i Poznámka. Konvergence v pravděpodobnosti. Posloupnost n.v. Y_n konverguje v pravděpodobnosti k Y (píšeme $Y_n \xrightarrow{P} Y$), pokud pro každé $\varepsilon > 0$ platí $P(|Y_n - Y| \geq \varepsilon) \rightarrow 0$. Slabý ZVČ tedy říká právě $\bar{X}_n \xrightarrow{P} \mu$. Tento režim konvergence (slabší než po hodnotách) používáme dále v definici konzistence odhadu.

💡 Intuice. ZVČ formalizuje „empirický průměr se ustálí na střední hodnotě“. Pokud házíme férovou mincí, podíl orlů (=empirický průměr indikátorů) se blíží k $1/2$. Pozor: *neříká*, že počet orlů minus $n/2$ se zmenšuje; ten naopak (typicky) roste jako $\sqrt{n}$.

Zákon velkých čísel: tři trajektorie $\bar{X}_n$ pro fair mince



ZVČ ilustračně. Tři nezávislé simulace házení férovou mincí (běh 1, běh 2, běh 3). Pro malé n silně kolísají (např. při $n = 10$ podíl orlů snadno 0,2 nebo 0,8), s rostoucím n se všechny tři blíží k $\mu = 0,5$ (čárkovaná čára). Při $n = 2000$: $\bar{X} = 0,484, 0,509, 0,515$.

4.2 Centrální limitní věta

★ Věta – Centrální limitní věta, CLV

Nechť $X_1, X_2, \dots$ jsou i.i.d. n.v. s konečnou střední hodnotou μ a rozptylem $\sigma^2 > 0$. Označme standardizovaný průměr

$$Y_n = \frac{(X_1 + \dots + X_n) - n\mu}{\sqrt{n}\sigma} = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}.$$

Pak Y_n konverguje v distribuci k $N(0, 1)$, značeno $Y_n \xrightarrow{d} N(0, 1)$: pro každé $x \in \mathbb{R}$

$$\lim_{n \rightarrow \infty} P(Y_n \leq x) = \Phi(x).$$

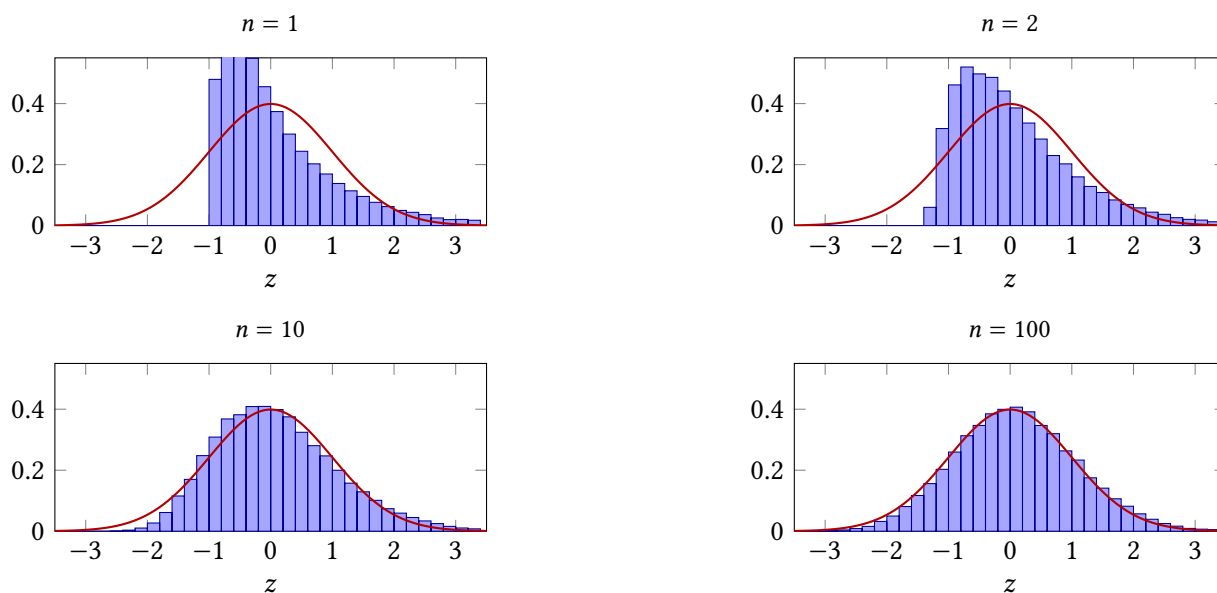
📺 **Video:** 3Blue1Brown, díl But what is the Central Limit Theorem? (vizuálně proč se standardizované součty i.i.d. veličin blíží $N(0, 1)$; Galtonova deska, role μ a σ).

📺 **Video:** 3Blue1Brown, díl Convolutions – why $X + Y$ in probability is a beautiful mess (rozdělení součtu nezávislých veličin je konvoluce jejich rozdělení; opakovaná konvoluce hladí libovolné rozdělení do zvonu – mechanismus za CLV, viz panel $n = 2$ níže).

📌 **Poznámka. Konvergence v distribuci** $Y_n \xrightarrow{d} Y$ znamená, že distribuční funkce F_{Y_n} bodově konverguje k F_Y v každém bodě, kde je F_Y spojitá. Pro $Y \sim N(0, 1)$ je $F_Y = \Phi$ spojitá všude, takže dostáváme bodovou konvergenci $\lim_n F_{Y_n}(x) = \Phi(x)$ ve všech $x \in \mathbb{R}$. Není to konvergence náhodných veličin po hodnotách – konvergují jen jejich rozdělení.

💡 **Intuice.** Klíčová vlastnost: *rozdělení původních X_i je libovolné* (jen s konečným rozptylem). Po sečtení a standardizaci dostaneme univerzálně $N(0, 1)$. To je důvod, proč je normální rozdělení tak časté: cokoliv vzniká součtem mnoha drobných nezávislých příspěvků (měřicí chyby, doba běhu, výška populace) má rozdělení blízké normálnímu. *Rychlost* konvergence závisí na původním rozdělení – pro symetrická a hladká stačí $n \approx 20$ – 30 , pro silně vychýlená (např. exponenciální) typicky $n \approx 100$.

📌 **Poznámka. Aplikační forma.** Pokud sčítáme $S_n = X_1 + \dots + X_n$, pak $S_n \approx N(n\mu, n\sigma^2)$ pro velká n , neboli $P(S_n \leq s) \approx \Phi((s - n\mu)/(\sqrt{n}\sigma))$. Tohle je tvar, který používáme při výpočtech.



CLV ilustračně. Histogramy standardizovaného průměru $Y_n = \sqrt{n}(\bar{X}_n - 1)$ pro $X_i \sim \text{Exp}(1)$ ($\mu = \sigma = 1$); $5 \cdot 10^4$ simulací pro každé n . $n = 1$: silně vychýlené (Exp). $n = 2$: stále zešíkmené. $n = 10$: téměř symetrické, ale vrchol posunut. $n = 100$: prakticky neodlišitelné od $N(0, 1)$ (červená křivka = teoretická hustota ϕ).

Znění CLV a její aplikaci procvičuje následující reálná otázka SZZ.

Příklad – SZZ léto 2025 č. 4: Centrální limitní věta (doba běhu programu)

(Reálná otázka SZZ: léto 2025, č. 4 – Centrální limitní věta)

Zadání.

- (1) Uveďte znění Centrální limitní věty.
- (2) Program zpracovává jeden vstup náhodnou dobou se střední hodnotou 100 ms a směrodatnou odchylkou 20 ms. Odhadněte pomocí CLV pravděpodobnost, že 100 vstupů zabere více než 10,3 s. Použijte tabulku Φ : $\Phi(1,5) \approx 0,93$.
- (3) Jaké předpoklady jsme udělali a jak jsou realistické?

Řešení.

- (1) (Viz znění výše.) $X_1, X_2, \dots$ i.i.d. s $E(X_i) = \mu$ a $\text{var}(X_i) = \sigma^2$; pak $Y_n = ((X_1 + \dots + X_n) - n\mu)/(\sqrt{n}\sigma) \xrightarrow{d} N(0, 1)$, tedy $\lim_{n \rightarrow \infty} F_{Y_n}(x) = \Phi(x)$ pro každé x .
- (2) Označme X_i dobu zpracování i -tého vstupu (v sekundách). $\mu = 0,1$, $\sigma = 0,02$, $n = 100$. Hledáme

$$P(S_{100} > 10,3) = P\left(Y_{100} > \frac{10,3 - 100 \cdot 0,1}{\sqrt{100} \cdot 0,02}\right) = P\left(Y_{100} > \frac{0,3}{0,2}\right) = P(Y_{100} > 1,5).$$

Podle CLV $P(Y_{100} \leq 1,5) \approx \Phi(1,5) \approx 0,93$, takže

$$P(S_{100} > 10,3) \approx 1 - \Phi(1,5) \approx 1 - 0,93 = \boxed{0,07} \quad (\approx 7\%).$$

- (3) **Předpoklady CLV:** (a) doby jsou *nezávislé*; (b) jsou *stejně rozdělené* s konečnou μ a σ^2 . *Konkrétní tvar rozdělení nás nezajímá* – to je síla CLV. **Realistické?** Nezávislost je často porušena: pokud je počítač přetížen, všechny vstupy se zpomalí společně. Stejně rozdělení je rozumné, pokud vstupy nemají systematicky různou náročnost. Náhrada limity konečným $n = 100$ je u CLV obvykle bezpečná, pokud původní rozdělení není extrémně zešíkmené.

5 Bodové odhady

Předpokládáme *model*: pozorování $X_1, \dots, X_n$ jsou i.i.d. z rozdělení F_θ , kde θ je neznámý *parametr* (např. střední hodnota, rozptyl, parametr Bernoulliho). Z dat chceme zrekonstruovat hodnotu θ , kterou „nevi-

díme“ přímo, jen skrze náhodná pozorování.

Definice – Náhodný výběr, statistika, bodový odhad

Náhodný výběr rozsahu n je n -tice i.i.d. veličin $X_1, \dots, X_n$ se společným rozdělením F_θ . *Statistika* je libovolná (měřitelná) funkce výběru $T = T(X_1, \dots, X_n)$; jako funkce náhodných veličin je sama náhodnou veličinou. *Bodový odhad* parametru θ je statistika $\hat{\theta}_n = \hat{\theta}_n(X_1, \dots, X_n)$, jejíž hodnotu používáme jako odhad neznámého θ .

 **Intuice.** Statistika smí záviset *jen* na datech, ne na neznámém θ – jinak by ji nešlo z pozorování spočítat. Dobrý bodový odhad je takový, který je „co nejbliž“ skutečné θ ; co přesně „blízko“ znamená, upřesňují vlastnosti níže (nestrannost, konzistence, MSE).

5.1 Vlastnosti odhadů

Definice – Nestrannost, konzistence, MSE

Nechť $\hat{\theta}_n$ je odhad parametru θ .

- (1) $\hat{\theta}_n$ je *nestranný* (unbiased), pokud $E_\theta(\hat{\theta}_n) = \theta$ pro každé θ . *Vychýlení* (bias) je $E_\theta(\hat{\theta}_n) - \theta$.
- (2) $\hat{\theta}_n$ je (*slabě*) *konzistentní*, pokud $\hat{\theta}_n \xrightarrow{P} \theta$ pro $n \rightarrow \infty$, tj. $P(|\hat{\theta}_n - \theta| \geq \varepsilon) \rightarrow 0$ pro každé $\varepsilon > 0$.
- (3) *Střední kvadratická chyba* (MSE): $MSE(\hat{\theta}_n) = E_\theta((\hat{\theta}_n - \theta)^2) = \text{var}_\theta(\hat{\theta}_n) + (\text{bias}(\hat{\theta}_n))^2$.

 **Intuice.** Nestrannost říká „v průměru přes opakované experimenty trefíme správně“. Konzistence: „s rostoucím počtem dat se k θ blížíme“. MSE: kompromis mezi vychýlením a rozptylem; někdy se vyplatí mít trochu vychýlený odhad, pokud má výrazně menší rozptyl (typický příklad: *regularizace* v ML).

★ Věta – Výběrový průměr a výběrový rozptyl

Pro i.i.d. vzorek $X_1, \dots, X_n$ z rozdělení s $\mu = E(X_i)$ a $\sigma^2 = \text{var}(X_i)$:

- (1) *Výběrový průměr* $\bar{X}_n = \frac{1}{n} \sum_i X_i$ je nestranný odhad μ : $E(\bar{X}_n) = \mu$. Je i konzistentní (ZVČ).
- (2) $\text{var}(\bar{X}_n) = \sigma^2/n$.
- (3) *Výběrový rozptyl* $S_n^2 = \frac{1}{n-1} \sum_i (X_i - \bar{X}_n)^2$ je nestranný odhad σ^2 (jmenovatel $n-1$, ne $n!$).

 **Poznámka.** Jmenovatel $n-1$ ve výběrovém rozptylu se nazývá *Besselova korekce*. „Naivní“ verze s jmenovatelem n podhodnocuje σ^2 , protože ve výpočtu nahrazujeme neznámé μ odhadem $\bar{X}_n$, což „ubere stupeň volnosti“. $E(\frac{1}{n} \sum_i (X_i - \bar{X}_n)^2) = \frac{n-1}{n} \sigma^2$, kdežto $E(S_n^2) = \sigma^2$.

5.2 Metoda maximální věrohodnosti (MLE)

Požadavky jmenují „alespoň jednu metodu pro tvorbu bodových odhadů“. Volíme *metodu maximální věrohodnosti* (MLE), vedle metody momentů jednu ze dvou metod ve Šámalových skriptech a v praxi nej-používanější. Princip: zvolíme $\hat{\theta}$ tak, aby *maximalizovala* pravděpodobnost (resp. hustotu) pozorovaných dat při tomto θ .

Definice – Věrohodnostní funkce, MLE

Pro i.i.d. vzorek $x_1, \dots, x_n$ a parametrickou rodinu $\{f_\theta\}$ (PMF nebo hustota) definujeme *věrohodnostní funkci*

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

Odhad metodou maximální věrohodnosti (MLE) je

$$\hat{\theta}_{\text{MLE}} = \arg \max_{\theta} L(\theta; x_1, \dots, x_n) = \arg \max_{\theta} \ell(\theta; x_1, \dots, x_n),$$

kde $\ell = \log L = \sum_i \log f_{\theta}(x_i)$ je *log-věrohodnost*.

Intuice. Logaritmování mění součin na součet – snazší derivování. Maximalizujeme ℓ obvykle přes $\partial \ell / \partial \theta = 0$ a kontrolou znaménka druhé derivace. MLE má řadu hezkých teoretických vlastností: za rozumných podmínek je *konzistentní*, *asymptoticky nestranný* a *asymptoticky normální* (s rozptylem daným „Fisherovou informací“).

Příklad – MLE pro Bernoulliho parametr

Zadání. Provedeme n nezávislých Bernoulliho pokusů s neznámou pravděpodobností úspěchu p a pozorujeme k úspěchů. Najděte MLE pro p .

Řešení. Pozorování $x_1, \dots, x_n \in \{0, 1\}$, $k = \sum_i x_i$. Věrohodnost $L(p) = p^k(1-p)^{n-k}$. Log-věrohodnost $\ell(p) = k \log p + (n-k) \log(1-p)$. Derivace:

$$\frac{d\ell}{dp} = \frac{k}{p} - \frac{n-k}{1-p} = 0 \implies k(1-p) = (n-k)p \implies \boxed{\hat{p}_{\text{MLE}} = \frac{k}{n}}.$$

Pro $0 < k < n$ je druhá derivace $-k/p^2 - (n-k)/(1-p)^2 < 0$, takže jde o maximum. (Pro $k = 0$ je $L(p) = (1-p)^n$ klesající a maximum je v $p = 0$, pro $k = n$ je $L(p) = p^n$ rostoucí a maximum je v $p = 1$; vzorec k/n tedy platí i tehdy.) MLE je výběrový průměr indikátorů, což je intuitivní: nejlepší odhad pravděpodobnosti úspěchu je relativní četnost úspěchů.

Příklad – MLE pro Poissonovo λ

Zadání. Pozorujeme $x_1, \dots, x_n \in \mathbb{N}_0$ z $\text{Pois}(\lambda)$. Najděte MLE.

Řešení. $L(\lambda) = \prod_i e^{-\lambda} \lambda^{x_i} / x_i!$, $\ell(\lambda) = -n\lambda + (\sum_i x_i) \log \lambda - \sum_i \log(x_i!)$. Derivace $\partial \ell / \partial \lambda = -n + (\sum_i x_i) / \lambda = 0$, takže $\hat{\lambda}_{\text{MLE}} = \bar{x}_n$ (druhá derivace $-\sum_i x_i / \lambda^2 < 0$, pokud $\sum_i x_i > 0$). Opět výběrový průměr, protože $E_{\lambda}(X) = \lambda$. Obecně ale MLE průměrem být nemusí: např. pro rovnoměrné rozdělení na $[0, \theta]$ je $\hat{\theta}_{\text{MLE}} = \max_i x_i$.

6 Intervalové odhady

Bodový odhad dá jedno číslo a neříká, jak je spolehlivý. Intervalový odhad místo jednoho čísla dá náhodný interval $[\hat{\theta}_L, \hat{\theta}_U]$, který neznámé θ pokryje s pravděpodobností (alespoň) $1 - \alpha$. Požadavky chtějí metodu založenou na *aproximaci normálním rozdělením* (pomocí CLV).

Definice – Konfidenční interval

Konfidenční interval (CI) pro parametr θ s *hladinou spolehlivosti* $1 - \alpha \in (0, 1)$ je dvojice statistik $\hat{\theta}_L < \hat{\theta}_U$ (závisí na $X_1, \dots, X_n$) taková, že pro každé θ

$$P_{\theta}(\hat{\theta}_L \leq \theta \leq \hat{\theta}_U) \geq 1 - \alpha.$$

6.1 CI pro střední hodnotu při známém σ

Pro i.i.d. vzorek z rozdělení s rozptylem σ^2 máme z CLV $\bar{X}_n \approx N(\mu, \sigma^2/n)$. Standardizací $Z_n = (\bar{X}_n - \mu) \sqrt{n}/\sigma \approx N(0, 1)$. Hledáme symetrický interval kolem μ s pokrytím $1 - \alpha$. Protože $N(0, 1)$ je symetrické kolem nuly, rozdělíme „nepokrytou“ pravděpodobnost α rovnoměrně do obou ocasů po $\alpha/2$. Označme

$z_{\alpha/2}$ kvantil $N(0, 1)$ úrovně $1 - \alpha/2$ (tj. $\Phi(z_{\alpha/2}) = 1 - \alpha/2$); pak $P(|Z| \leq z_{\alpha/2}) = 1 - \alpha$ pro $Z \sim N(0, 1)$. Aplikováno na náš Z_n :

$$P(|Z_n| \leq z_{\alpha/2}) \approx 1 - \alpha \iff P\left(\bar{X}_n - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X}_n + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right) \approx 1 - \alpha.$$

★ Věta – CI pro střední hodnotu, σ známé

Při i.i.d. vzorku s konečným rozptylem σ^2 je

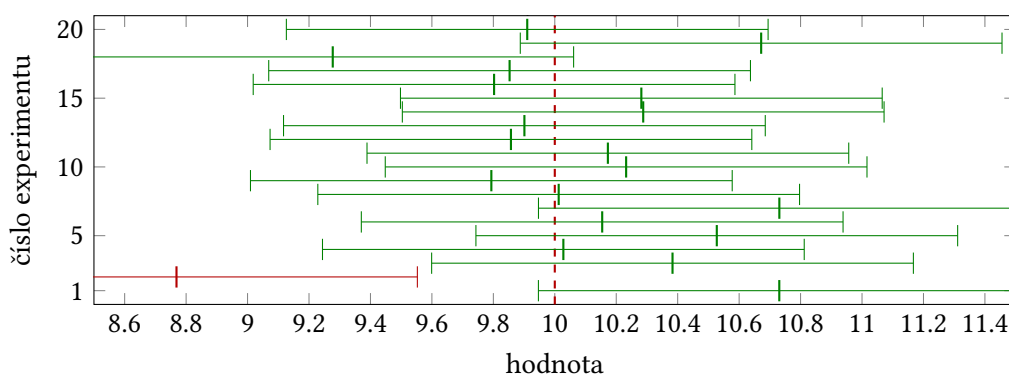
$$\bar{X}_n \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

přibližný $(1 - \alpha)$ -konfidenční interval pro μ . Pro $\alpha = 0,05$ je $z_{0,025} \approx 1,96$.

💡 Intuice. Šířka CI = $2z_{\alpha/2}\sigma/\sqrt{n}$. Klesá jako $1/\sqrt{n}$, tedy abychom zúžili interval 10×, potřebujeme 100× víc dat. Závisí na *spolehlivosti* $1 - \alpha$ (vyšší spolehlivost → širší interval) a na σ (vyšší šum → širší interval).

📌 Poznámka. Interpretace 95 % CI. Náhodný je interval (jeho meze závisí na vzorku), *pevné* je neznámé μ . Význam „95 %“: kdybychom experiment opakovali mnohokrát, ≈ 95 % takto vypočtených intervalů by obsahovalo skutečné μ . Nelze říct „s pravděpodobností 95 % leží μ v tomto konkrétním intervalu“ – po výpočtu už není nic náhodné.

Konfidenční intervaly: 20 simulací z $N(10, 4)$, $n = 25$, 95 % CI



20 nezávislých 95 % konfidenčních intervalů $\bar{X}_n \pm 1,96 \cdot 2/\sqrt{25} = \bar{X}_n \pm 0,784$ ze vzorků o velikosti $n = 25$ z $N(\mu = 10, \sigma^2 = 4)$. **Zelené** intervaly obsahují $\mu = 10$, **červený** (č. 2) ho mine. Pokrytí $19/20 = 95$ % tu náhodou přesně odpovídá teoretické hladině (v jiném běhu simulace by mohlo vyjít např. 18/20 nebo 20/20). Šířka všech intervalů je stejná (σ známé), liší se jen poloha středu $\bar{X}_n$.

6.2 CI pro Bernoulliho parametr

V praxi často chceme CI pro *podíl* p (preferenční volby, podíl chybných výrobků). Pro $X_i \sim \text{Bern}(p)$ je $\bar{X}_n = k/n$ bodový odhad a $\text{var}(X_i) = p(1 - p)$. Z CLV:

$$\frac{k/n - p}{\sqrt{p(1 - p)/n}} \approx N(0, 1).$$

Po nahrazení neznámého $p(1 - p)$ odhadem $\hat{p}(1 - \hat{p})$ (Slutského věta) máme přibližný CI

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}.$$

Poznámka. Tato aproximace funguje dobře, pokud $n\hat{p} \geq 5$ a $n(1 - \hat{p}) \geq 5$ (cca „aspoň 5 úspěchů a 5 neúspěchů“). Pro malé vzorky existují přesnější metody (Wilson, Clopper-Pearson), které nejsou v rozsahu okruhu.

7 Testování hypotéz

Místo odhadu konkrétní hodnoty zde rozhodujeme *ano/ne* otázku: data podporují, nebo zamítají, nějakou hypotézu? Klasický příklad: „je tato mince férová?“, „je nový lék účinnější než placebo?“

7.1 Základní rámec

Definice – Nulová a alternativní hypotéza, testová statistika

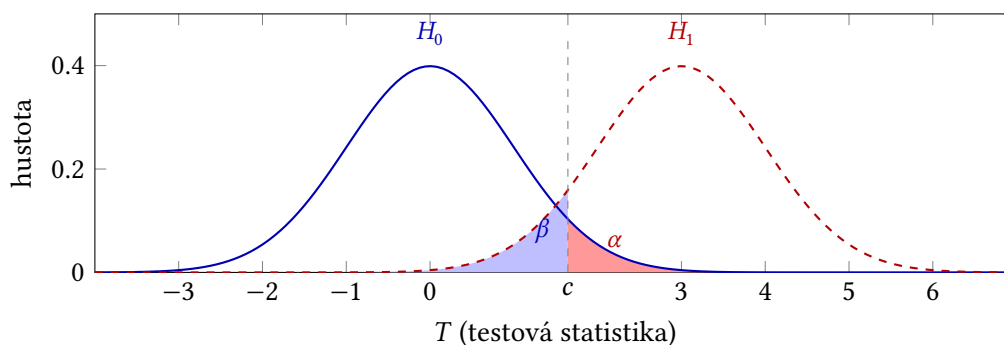
- *Nulová hypotéza* H_0 : „defaultní, konzervativní“ tvrzení o parametru („nic nového se neděje“). Např. $H_0: p = 1/2$ pro férovou minci.
- *Alternativní hypotéza* H_1 : tvrzení, které platí, pokud zamítneme H_0 .
- *Testová statistika* $T = T(X_1, \dots, X_n)$: funkce dat, jejíž rozdělení za *platnosti* H_0 známe.
- *Kritický obor* W : množina hodnot T , pro které H_0 zamítáme.

Definice – Chyby 1. a 2. druhu, hladina významnosti

- *Chyba 1. druhu* („falešné zamítnutí“): zamítneme H_0 , ačkoliv ve skutečnosti platí. Její pravděpodobnost značíme $\alpha = P(T \in W | H_0)$ a nazýváme *hladina významnosti* (signifikance).
- *Chyba 2. druhu* („falešné přijetí“): nezamítneme H_0 , ačkoliv ve skutečnosti neplatí (platí H_1). Její pravděpodobnost značíme $\beta = P(T \notin W | H_1)$.
- *Síla testu* (power): $1 - \beta = P(T \in W | H_1)$ – pravděpodobnost správného zamítnutí.

Intuice. α si volíme (typicky $\alpha = 0,05$ nebo $0,01$) a podle ní stanovíme kritický obor. β vyplývá z volby α , velikosti vzorku n a skutečné polohy v H_1 . *Vztah:* při fixním n se snížením α zvyšuje β (užší kritický obor $\rightarrow$ méně falešných zamítnutí, ale i méně skutečných objevů). Jediný způsob, jak snížit *obojí* současně, je zvětšit vzorek.

Chyby 1. a 2. druhu: $H_0: T \sim N(0, 1)$, $H_1: T \sim N(3, 1)$



Chyby 1. a 2. druhu pro jednostranný test. $H_0: T \sim N(0, 1)$ (modrý zvonec), $H_1: T \sim N(3, 1)$ (červený, posunutý). Kritická hodnota $c = z_{0,05} \approx 1,645$ (čárkovaná čára), kritický obor $W = [c, \infty)$. Červená oblast pod H_0 vpravo od c má míru $\alpha = 0,05$ (chyba 1. druhu). Modrá oblast pod H_1 vlevo od c má míru $\beta = \Phi(c - 3) = \Phi(-1,355) \approx 0,088$ (chyba 2. druhu). Síla testu $1 - \beta \approx 0,912$.

7.2 p-hodnota

Definice – p-hodnota

p-hodnota pozorování $t^{\text{obs}} = T(x_1, \dots, x_n)$ je pravděpodobnost (za H_0), že bychom dostali výsledek *minimálně tak extrémní* jako pozorovaný. Pro jednostranný test ve směru velkých hodnot: $p = P(T \geq t^{\text{obs}} \mid H_0)$.

 **Intuice.** Rozhodovací pravidlo: $p \leq \alpha \Rightarrow$ zamítáme H_0 . Stejně jako porovnání $t^{\text{obs}} \in W$, ale p-hodnota dává navíc *kvantitativní* informaci – „jak silně“ data H_0 porušují. Časté *nedorozumění*: p-hodnota *není* pravděpodobnost, že H_0 platí. Je to pravděpodobnost dat alespoň tak extrémních jako pozorovaná, za předpokladu, že H_0 platí.

7.3 Příklad: test férovosti mince

Příklad – Test férovosti

Zadání. Hodili jsme mincí $n = 100$ krát a padlo 65 orlů. Je mince férová ($p = 1/2$) na hladině $\alpha = 0,05$?

Řešení. $H_0 : p = 1/2$, $H_1 : p \neq 1/2$ (oboustranný test).

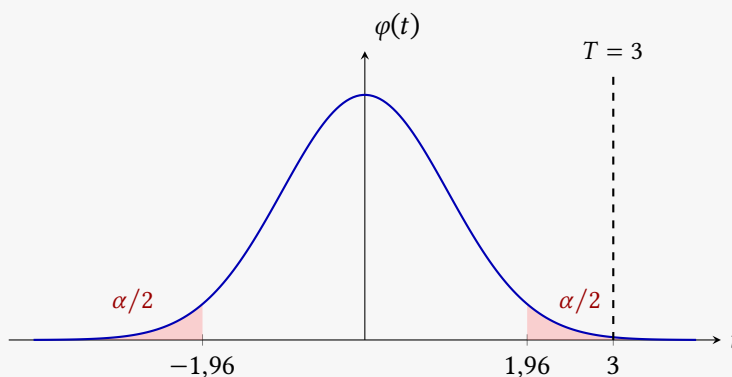
Testová statistika. Označme K počet orlů ve $n = 100$ hodech. Za H_0 je $K \sim \text{Bin}(100, 1/2)$ s $E(K) = np_0 = 50$ a $\text{var}(K) = np_0(1 - p_0) = 25$. Standardizace přes CLV dává

$$T = \frac{K - np_0}{\sqrt{np_0(1 - p_0)}} = \frac{65 - 50}{\sqrt{25}} = 3, \quad T \approx N(0, 1) \text{ za } H_0.$$

Kritický obor. Pro oboustranný test $W = \{|T| \geq z_{\alpha/2}\} = \{|T| \geq 1,96\}$. Pozorujeme $|T| = 3 \in W$, takže H_0 **zamítáme**.

p-hodnota. $p = P(|N(0, 1)| \geq 3) = 2(1 - \Phi(3)) \approx 2 \cdot 0,00135 = 0,0027$. To je *výrazně* menší než 0,05, takže silné zamítnutí.

Závěr. Na hladině $\alpha = 0,05$ (i 0,01) data nepodporují hypotézu férové mince.



Kritický obor oboustranného testu při $\alpha = 0,05$ jsou oba ocasy za $\pm 1,96$ (vybarvené, dohromady míra 0,05). Pozorovaná hodnota $T = 3$ leží daleko v pravém ocasu ($3 > 1,96$), takže H_0 (férová mince) na hladině 5 % **zamítáme**.

 **Poznámka. Obecný z-test pro střední hodnotu.** Test férovosti je speciální případ. Pro i.i.d. výběr se známým rozptylem σ^2 a $H_0 : \mu = \mu_0$ použijeme statistiku $T = (\bar{X}_n - \mu_0)\sqrt{n}/\sigma$, která je za H_0 podle CLV přibližně $N(0, 1)$. Oboustranný kritický obor je $\{|T| \geq z_{\alpha/2}\}$, jednostranný (proti $H_1 : \mu > \mu_0$) $\{T \geq z_\alpha\}$. Test je tedy „obrácený“ konfidenční interval: H_0 na hladině α zamítneme oboustranně právě tehdy, když μ_0 neleží uvnitř $(1 - \alpha)$ -CI $\bar{X}_n \pm z_{\alpha/2}\sigma/\sqrt{n}$.

7.4 Test dobré shody (chí-kvadrát) – nad rámec požadavků

i Poznámka. Následující věta je *nad rámec požadavků* okruhu (testování hypotéz vyžaduje jen základní rámec, chyby a hladinu). Uvádíme ji jako konkrétní příklad testu z přednášky NMAI059; statistické testy včetně chí-kvadrát testu dále rozvíjí okruh Strojové učení.

★ Věta – Chí-kvadrát test dobré shody

Mějme rozdělení do K kategorií s teoretickými pravděpodobnostmi $p_1, \dots, p_K$ za H_0 ($\sum p_i = 1$). Pozorovali jsme četnosti $N_1, \dots, N_K$ z $n = \sum N_i$ pokusů. Statistika

$$\chi^2 = \sum_{i=1}^K \frac{(N_i - np_i)^2}{np_i}$$

má za H_0 pro velká n rozdělení χ^2_{K-1} (chí-kvadrát s $K - 1$ stupni volnosti). Zamítáme H_0 , pokud $\chi^2 \geq \chi^2_{K-1, 1-\alpha}$ (kvantil úrovně $1 - \alpha$). Aproximace se obvykle považuje za dostatečnou, když všechny očekávané četnosti $np_i \geq 5$.

💡 Intuice. Každý člen $(N_i - np_i)^2 / (np_i)$ srovnává pozorovanou a očekávanou četnost v kategorii i . Velký rozdíl v některé kategorii vyrobí velký příspěvek. Stupně volnosti jsou $K - 1$ (ne K), protože jeden vztah $\sum N_i = n$ ubírá stupeň volnosti.

Shrnutí

i Poznámka. Klíčové vlastnosti a vzorce (přehled k SZZ). Telegrafický výčet klíčových definic, vět a vzorců okruhu M5. Obě reálné SZZ otázky (léto 2023: linearita E ; léto 2025: CLV + aplikace) jsou v bloku Střední hodnota a Limitní věty.

Pravděpodobnostní prostor a základní pravidla.

- Trojice $(\Omega, \mathcal{F}, P)$: Ω – elementární jevy, $\mathcal{F}$ – σ -algebra jevů, $P : \mathcal{F} \rightarrow [0, 1]$ – pravděpodobnost se **spočetnou aditivitou** ($P(\bigcup A_i) = \sum P(A_i)$ pro disjunktní A_i) a $P(\Omega) = 1$.
- Klasický prostor: $P(A) = |A|/|\Omega|$ (symetrický, konečný Ω); geometrický: $P(A) = m(A)/m(\Omega)$.
- Odvozená pravidla: $P(\emptyset) = 0$; $P(A^c) = 1 - P(A)$; monotonie $A \subseteq B \Rightarrow P(A) \leq P(B)$; $P(A \cup B) = P(A) + P(B) - P(A \cap B)$; subaditivita $P(\bigcup A_i) \leq \sum P(A_i)$.
- **Inkluze-exkluze**: $P(\bigcup_{i=1}^n A_i) = \sum_{S \neq \emptyset} (-1)^{|S|+1} P(\bigcap_{i \in S} A_i)$.

Podmíněná pravděpodobnost, nezávislost, Bayes.

- $P(A | B) = P(A \cap B) / P(B)$ pro $P(B) > 0$; multiplikační pravidlo $P(A \cap B) = P(A | B)P(B)$.
- **Nezávislost**: A, B nezávislé $\iff P(A \cap B) = P(A)P(B)$ ($\iff P(A | B) = P(A)$ pro $P(B) > 0$). Kolekce $A_1, \dots, A_n$ navzájem nezávislá $\iff$ každá podmnožina splňuje podmínku součinu – po dvojicích nestačí!
- **Věta o úplné pravděpodobnosti**: rozklad $B_1, \dots, B_n$ prostoru Ω : $P(A) = \sum_i P(A | B_i)P(B_i)$.
- **Bayesův vzorec**: $P(B_j | A) = P(A | B_j)P(B_j) / \sum_i P(A | B_i)P(B_i)$ – „obráť“ podmiňování (apriorní $\times$ věrohodnost, normalizováno).

Náhodné veličiny a jejich popis.

- N.v. $X : \Omega \rightarrow \mathbb{R}$; **distribuční funkce** $F_X(x) = P(X \leq x)$: neklesající, zprava spojitá, $F_X(-\infty) = 0$, $F_X(+\infty) = 1$; $P(a < X \leq b) = F_X(b) - F_X(a)$.
- **Diskrétní**: PMF $p_X(k) = P(X = k)$; F_X je schodovitá, $\sum_k p_X(k) = 1$.
- **Spojitá**: hustota $f_X \geq 0$ s $F_X(x) = \int_{-\infty}^x f_X(t) dt$; $P(a \leq X \leq b) = \int_a^b f_X$; f_X *není* pravděpodobnost (může být > 1).

Střední hodnota a rozptyl.

- $E(X) = \sum_k k p_X(k)$ (diskrétní) resp. $\int x f_X(x) dx$ (spojitá), existuje-li absolutně.

- **Linearita střední hodnoty** (SZZ léto 2023): $E(\sum_i a_i X_i + b) = \sum_i a_i E(X_i) + b$ – bez ohledu na nezávislost. Indikátor: $E(I_A) = P(A)$, takže počítání přes indikátory funguje i pro závislé n.v.
- **Nezávislost n.v.:** $P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$ pro všechna x, y (diskrétní: $P(X = x, Y = y) = P(X = x)P(Y = y)$).
- **Střední hodnota součinu nezávislých:** $E(XY) = E(X)E(Y)$ (bez nezávislosti obecně neplatí; z $E(XY) = E(X)E(Y)$ naopak nezávislost neplyne).
- **Markovova nerovnost** (pro $X \geq 0, a > 0$): $P(X \geq a) \leq E(X)/a$.
- **Rozptyl:** $\text{var}(X) = E((X - \mu)^2) = E(X^2) - E(X)^2$; $\sigma(X) = \sqrt{\text{var}(X)}$; $\text{var}(aX + b) = a^2 \text{var}(X)$.
- **Rozptyl součtu:** $\text{var}(X + Y) = \text{var}(X) + \text{var}(Y) + 2 \text{cov}(X, Y)$; pro nezávislé $\text{cov} = 0$, takže $\text{var}(\sum_i X_i) = \sum_i \text{var}(X_i)$ a $\text{var}(\bar{X}_n) = \sigma^2/n$. Pozor: $\text{cov} = 0 \nRightarrow$ nezávislost.
- **Čebyševova nerovnost:** $P(|X - \mu| \geq k) \leq \sigma^2/k^2$ (z Markova na $(X - \mu)^2$).

Konkrétní rozdělení (E a var).

- Bern(p): $\{0, 1\}$, PMF $p^k(1-p)^{1-k}$; $E = p$, $\text{var} = p(1-p)$.
- Bin(n, p): počet úspěchů v n nezávislých Bernoulli pokusech, PMF $\binom{n}{k} p^k(1-p)^{n-k}$; $E = np$, $\text{var} = np(1-p)$. Aproximace Poissonem pro n velké, p malé, $\lambda = np$ rozumné.
- Geom(p): počet pokusů do prvního úspěchu (vč. úspěšného), PMF $(1-p)^{k-1}p$; $E = 1/p$, $\text{var} = (1-p)/p^2$. Paměťová vlastnost: $P(X > m+k | X > m) = P(X > k)$.
- Pois(λ): počet vzácných událostí, PMF $e^{-\lambda} \lambda^k / k!$; $E = \lambda$, $\text{var} = \lambda$.
- Exp(λ): hustota $\lambda e^{-\lambda x}$ ($x \geq 0$), CDF $1 - e^{-\lambda x}$; $E = 1/\lambda$, $\text{var} = 1/\lambda^2$. Paměťová vlastnost (spojitý protějšek Geom).
- $N(\mu, \sigma^2)$: hustota $\frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{(x-\mu)^2}{2\sigma^2})$; $E = \mu$, $\text{var} = \sigma^2$. Standardizace: $(X - \mu)/\sigma \sim N(0, 1)$; CDF Φ (tabulkovaná). Lineární transformace: $aX + b \sim N(a\mu + b, a^2\sigma^2)$. Pravidla $1\sigma/2\sigma/3\sigma$: 68 %/95 %/99,7 %.

Limitní věty.

- **Zákon velkých čísel (ZVČ):** $X_1, X_2, \dots$ i.i.d. s $E(X_i) = \mu$, $\text{var}(X_i) = \sigma^2 \Rightarrow \bar{X}_n \xrightarrow{P} \mu$, tj. $P(|\bar{X}_n - \mu| \geq \varepsilon) \rightarrow 0$ pro každé $\varepsilon > 0$. Důkaz: Čebyšev + $\text{var}(\bar{X}_n) = \sigma^2/n \rightarrow 0$.
- **Centrální limitní věta (CLV)** (SZZ léto 2025): $X_1, X_2, \dots$ i.i.d. s $\mu, \sigma^2 > 0 \Rightarrow Y_n = \frac{S_n - n\mu}{\sqrt{n}\sigma} \xrightarrow{d} N(0, 1)$, tj. $\lim_{n \rightarrow \infty} P(Y_n \leq x) = \Phi(x)$ pro všechna x . Rozdělení X_i je libovolné!
- **Aplikační forma CLV:** $S_n = X_1 + \dots + X_n \approx N(n\mu, n\sigma^2)$, tedy $P(S_n \leq s) \approx \Phi((s - n\mu)/(\sqrt{n}\sigma))$. Předpoklady: nezávislost + stejné rozdělení + konečné μ, σ^2 . Konkrétní tvar rozdělení nezáleží.

Bodové odhady.

- **Nestrannost:** $E(\hat{\theta}_n) = \theta$; **konzistence:** $\hat{\theta}_n \xrightarrow{P} \theta$; **MSE** = $\text{var}(\hat{\theta}_n) + \text{bias}^2$.
- **Výběrový průměr** $\bar{X}_n$: nestranný + konzistentní odhad μ (ze ZVČ), $\text{var}(\bar{X}_n) = \sigma^2/n$.
- **Výběrový rozptyl** $S_n^2 = \frac{1}{n-1} \sum (X_i - \bar{X}_n)^2$: nestranný odhad σ^2 (Besselova korekce – jmenovatel $n-1$, ne n !).
- **MLE:** $\hat{\theta} = \arg \max_{\theta} \ell(\theta)$, kde $\ell(\theta) = \sum_{i=1}^n \log f_{\theta}(x_i)$ (log-věrohodnost). Pro Bern(p): $\hat{p} = k/n$; pro Pois(λ): $\hat{\lambda} = \bar{x}_n$. MLE je (za podmínek regularity) konzistentní a asymptoticky normální.

Intervalové odhady (z CLV).

- **CI pro μ (σ znám):** $\bar{X}_n \pm z_{\alpha/2} \sigma / \sqrt{n}$, kde $\Phi(z_{\alpha/2}) = 1 - \alpha/2$; pro $\alpha = 0,05$: $z_{0,025} \approx 1,96$.
- **CI pro Bernoulliho podíl:** $\hat{p} \pm z_{\alpha/2} \sqrt{\hat{p}(1-\hat{p})/n}$ (funguje pro $n\hat{p} \geq 5, n(1-\hat{p}) \geq 5$).
- Šířka CI $\propto 1/\sqrt{n}$ (pro 10× užší CI potřeba 100× víc dat); vyšší spolehlivost $1 - \alpha \uparrow \Rightarrow$ širší interval.
- Interpretace: *interval* je náhodný; $1 - \alpha$ je podíl intervalů přes opakované experimenty, které θ obsahují – ne pravděpodobnost pro konkrétní interval.

Testování hypotéz.

- **Rámec:** H_0 (nulová, konzervativní) vs. H_1 (alternativní); testová statistika T se daným rozdělením za H_0 ; kritický obor W – množina hodnot T , pro které H_0 zamítáme.
- **Chyba 1. druhu** (falešné zamítnutí): $\alpha = P(T \in W | H_0)$ – *hladina významnosti*; typicky $\alpha = 0,05$ nebo $0,01$.

- **Chyba 2. druhu** (falešné přijetí): $\beta = P(T \notin W \mid H_1)$; **síla testu** $1 - \beta$. Snížení α (při fixním n) zvyšuje β – lze snížit obojí jen zvětšením n .
- **p-hodnota**: $p = P(T \geq t^{\text{obs}} \mid H_0)$ (jednostranný); zamítáme H_0 při $p \leq \alpha$. p-hodnota *není* pravděpodobnost, že H_0 platí.
- **Postup z-testu pro střední hodnotu**: statistika $T = ((\bar{X}_n - \mu_0)\sqrt{n})/\sigma$, za H_0 je $T \approx N(0, 1)$; kritický obor oboustranný $\{|T| \geq z_{\alpha/2}\}$.
- **Chí-kvadrát test dobré shody** (nad rámec): $\chi^2 = \sum (N_i - np_i)^2 / (np_i) \sim \chi_{K-1}^2$ za H_0 .

Zdroje

- R. Šámal: *NMAI059 Pravděpodobnost a statistika 1*, skripta k přednášce, MFF UK, verze 14. 6. 2025 (primární zdroj výkladu: pravděpodobnostní prostor, náhodné veličiny, rozdělení, limitní věty, odhady a testování hypotéz).
<https://iuuk.mff.cuni.cz/~samal/vyuka/2425/PSt1/skripta.pdf>
- K. Král, M. Mareš, R. Šámal: *Řešená cvičení: NMAI059 Pravděpodobnost a statistika 1*, MFF UK, 2021 (sbírka řešených úloh ke cvičením; typové úlohy).
- Video odkazovaná v textu: 3Blue1Brown, *Bayes theorem, the geometry of changing beliefs; The medical test paradox, and redesigning Bayes' rule; Binomial distributions a Why “probability of 0” does not mean “impossible”* (série *Probabilities of probabilities*); *Why π is in the normal distribution; But what is the Central Limit Theorem?*; *Convolutions: Why $X+Y$ in probability is a beautiful mess* (YouTube, odkazy u příslušných témat).
- MFF UK: požadavky ke státní závěrečné zkoušce (Informatika, Bc.) a zadání otázek z minulých termínů.
<https://www.mff.cuni.cz/cs/studenti/bakalarske-studium/statni-zaverecne-zkousky/bakalarske-statni-zkousky-studijniho-programu-informatika>