10. cvičení z PSt — 23. 4. 2026

Věta 1. *Nechť* $X_1, \dots, X_n$ *jsou stejně rozdělené n.n.v. se střední hodnotou* μ *a rozptylem* σ^2. *Nechť* $Y_n = \frac{(X_1+\ldots+X_n)-n\cdot\mu}{\sigma\cdot\sqrt{n}}$. *Pak platí*

$$\overline{X}_n \xrightarrow{s.j.} \mu \qquad \textit{(Silný zákon velkých čísel)}$$

$$(\forall \varepsilon > 0) \qquad P(|\overline{X}_n - \mu| > \varepsilon) \le \frac{\sigma^2}{n\varepsilon^2} \qquad \textit{(Slabý zákon velkých čísel)}$$

$$(\forall x \in \mathbb{R}) \qquad \lim_{n\to\infty} F_{Y_n}(x) = \Phi(x) \qquad \textit{(Centrální limitní věta)}$$

x	-2.5	-2.0	-1.5	-1.0	-0.5	0.0	0.5	1.0	1.5	2.0	2.5
$\Phi(x)$	0.01	0.02	0.07	0.16	0.31	0.5	0.69	0.84	0.93	0.98	0.99

Zákony velkých čísel

1. Počítání obsahu kruhu náhodným samplováním. Vygenerujeme náhodný bod ve čtverci (obě souřadnice budou mít rozdělení $U(0,1)$). Označíme X_i indikátor jevu „i-tý bod leží v kruhu s poloměrem 1 a středem v počátku“.

(a) Určete $\mathbb{E}(X_i)$, $\mathrm{var}(X_i)$.
(b) Položte $\overline{X}_n = (X_1 + \cdots + X_n)/n$. Určete $\mathbb{E}(\overline{X}_n)$, $\mathrm{var}(\overline{X}_n)$.
(c) Rozmyslete si, co říká slabý a silný zákon velkých čísel o aproximaci π pomocí X_n?
(d) Pro jaké n čekáte, že dostaneme výsledek správně na jedno desetinné místo? Na dvě, tři, …?
(e) Jiný výpočet: $Y_i = \sqrt{1 - U_i^2}$, kde $U_i \sim U(0,1)$. Jaké je $\mathrm{var}(Y_i)$ a $\mathrm{var}(\overline{Y}_n)$? *(Hint dole)*[1]
(f) Která metoda je přesnější?

Centrální limitní věta

2. Připomeňme, že standardizací n.v. X myslíme $\mathrm{stand}(X) = (X - \mathbb{E}(X))/\sigma_X$.

(a) Ukažte, že $\mathrm{stand}(X)$ má střední hodnotu 0 a rozptyl 1.
(b) Ukažte, že Y_n v CLV je rovna $\mathrm{stand}(\overline{X}_n)$ a také $\mathrm{stand}(X_1 + \cdots + X_n)$.

3. Měříme rychlost stahování souborů z cloudového úložiště. Každý čas stahování jednoho souboru je náhodná veličina s průměrem $\mu = 5$ minut a standardní odchylkou $\sigma = 2$ minuty. Předpokládejme, že časy stahování jednotlivých souborů jsou na sobě nezávislé, stahování probíhá jedno po druhém (tj. vždy se stahuje jen jeden soubor, hned po jeho dokončení začneme stahovat další). Stahujeme celkem 50 souborů.

(a) Jaká je přibližná pravděpodobnost, že celková doba stahování přesáhne 270 minut?
(b) Jaká je přibližná pravděpodobnost, že průměrná doba stahování na soubor je kratší než 4,5 minuty?

Použijte Centrální limitní větu. Napište přesnou formuli pomocí funkce Φ a použijte tabulku pro odhad.

4. Označme $S = \sum_{k=0}^{30} \binom{100}{k}$. Označme dále $X = \sum_{i=1}^{100} X_i$, kde X_i je 0 nebo 1, obojí s pravděpodobností 1/2 a veličiny X_1, …, X_n jsou nezávislé. Je tedy $X \sim \mathrm{Bin}(100, 1/2)$.

(a) Vyjádřete S pomocí distribuční funkce F_X.
(b) Použijte CLV na odhad této pravděpodobnosti.
(c) Případně vyčíslete S vhodným softwarem a srovnejte.

Kvantilová funkce

5. Kvantilovou funkci jsme definovali předpisem $Q_X(p) = F_X^{-1}(p)$.

(a) Jaký je obor hodnot Q_X? Kdy dává definice smysl?
(b) Rozmyslete si, jak je rozumné definovat Q_X pro diskrétní n.v. (a v čem je vlastně problém).

6*. Ukažte pro libovolnou n.v. X a $U \sim U(0,1)$ že $Q_X(U)$ má stejné rozdělení jako X. *(Hint dole)*[2]

[1] Uvědomte si, že $\mathbb{E}(Y_i)$ je obsah čtvrtkruhu, tedy $\pi/4$.
[2] Ukažte, že $Q_X(U)$ i X mají stejnou distribuční funkci.

MLE a nestrannost

7. Připomeňte si, co je to věrohodnost, estimátor, bias estimátoru.

8. Němci vyrábějí tanky s pořadovými čísly $1, \dots, N$ pro neznámé N. Ukořistíme k z nich a vidíme pořadová čísla $X_1, \dots, X_k$, tj. rovnoměrně náhodnou k-prvkovou podmnožinu $\{1, \dots, N\}$. Nechť $M = \max(X_1, \dots, X_k)$.

(a) Ukažte, že $P(M = m) = \binom{m-1}{k-1} / \binom{N}{k}$ pro $m \in \{k, \dots, N\}$.

(b) Připomeňte si, že M je MLE pro N (ukazovali jsme na přednášce).

(c) Spočtěte $\mathbb{E}(M) = \frac{k(N+1)}{k+1}$. Může se hodit hockey-stick identity: $\sum_{m=k}^{N} \binom{m}{k} = \binom{N+1}{k+1}$.

(d) Jak z toho plyne nestranný odhad $\hat{N}_{unbiased} = \frac{k+1}{k} M - 1$?

(Pozor: na přednášce bylo místo -1 napsáno $-\frac{k+1}{k}$ — to byla chyba.)

K procvičení

9. Víme, že průměrný počet bodů z písemky byl 40 (ze 100). Odhadněte odsud podíl studentů s alespoň 80 body. Vylepšete odhad, pokud víte, že směrodatná odchylka počtu bodů je 10.

10. Chceme odhadnout, zda naše mince (a způsob jak s ní házíme) je spravedlivá. Pokud ze sta hodů padne orel více než 55-krát, řekneme, že spravedlivá není. Jaká je pravděpodobnost, že se zmýlíme?

11. Vzpomeňte si na větu z přednášky. Nechť $U \sim U(0, 1)$. Jak vyrobíte náhodnou veličinu

(a) s rozdělením $U(a, b)$?

(b) s rozdělením $N(0, 1)$? (Využijte funkce Φ jako "black box".)

(c) s uniformním rozdělením na množině $\{1, 2, \dots, 6\}$?

12. Odhadněte $\binom{100}{30}$ pomocí CLV. *(Hint dole)*[3]

[3]Použijte CLV pro odhad $P(29.5 < X < 30.5)$ pro vhodnou n.v. X.