10. cvičení z PSt — 20.–24.4.2026

Centrální limitní věta

1. 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).

(a) Pokud stáhneme 50 souborů, 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 na níže pro odhad.

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

MLE a nestrannost

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

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

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

(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: $\sum_{m=k}^{N} \binom{m}{k} = \binom{N+1}{k+1}$. Pak si připomeňte, 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.)**

4. Máme náhodný výběr $X_1, \ldots, X_n \sim N(\mu, 1)$ – data na vstupu $x_1, ..., x_n$ tedy pochází z normální distribuce s neznámým středem μ, ale známou směrodatnou odchylkou 1.

(a) Napište věrohodnostní (likelihood) funkci $p_\theta(x)$; později se může hodit pracovat s funkcí $\log p_\theta(x)$, které se říká log-likelihood.

(b) Derivací spočtěte $\hat{\mu}_{MLE}$, mělo by vyjít $\hat{\mu}_{MLE} = \frac{1}{n} \sum_{i=1}^{n} x_i$.

(c) Ověřte, že $\hat{\mu}_{MLE}$ je nestranný.

(d) Přesvědčte se, že kdyby směrodatná odchylka nebyla 1, ale byl to jakýkoliv (nám známý) parametr σ^2, $\hat{\mu}_{MLE}$ by vyšlo úplně stejně; volba $\sigma^2 = 1$ jen zjednodušuje algebru.

5. (Bonus.) Buďte $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$, oba parametry neznámé.

(a) Derivací ℓ podle μ a σ^2 ukažte, že MLE je $\hat{\mu}_{MLE} = \frac{1}{n} \sum_{i=1}^{n} x_i$ a $\hat{\sigma}^2_{MLE} = \frac{1}{n} \sum_{i=1}^{n} (X_i - \hat{\mu}_{MLE})^2$.

(b) Ověřte identitu $\sum (X_i - \hat{\mu}_{MLE})^2 = \sum (X_i - c)^2 - n(\hat{\mu}_{MLE} - c)^2$ pro každé c.

(c) S její pomocí ukažte, že $\mathbb{E}(\hat{\sigma}^2_{MLE}) = \frac{n-1}{n} \sigma^2$.

(d) Najděte korekci, aby byl estimátor nestranný: dostanete $s^2 = \frac{1}{n-1} \sum (X_i - \hat{\mu}_{MLE})^2$.

6. Najděte příklad nějakého nestranného estimátoru (třeba pro problémy z předchozích úloh), který je zjevně hodně špatný.