11. cvičení z PSt — 28.4.–2.5.2025

Aplikace nerovností a Centrální Limitní Věty

1. Házíme kostkou, za 1 a 2 dostaneme bod. Označme X počet bodů, které dostaneme po n (nezávislých) hodech. Odhadněte pravděpodobnost, že $X \geq n/2$.

(a) Pomocí Markovovy nerovnosti.

(b) Pomocí Čebyševovy nerovnosti.

(c) Pro konkrétní n, jak lze tuto hodnotu určitě přesně?

2. Statistik chce odhadnout průměrnou výšku h (v metrech) lidí v nějaké populaci, pomocí n nezávislých vzorků $X_1, \dots, X_n$, které vybíráme uniformně náhodně se všech možných lidí. Pro odhad použije výběrový průměr $\bar{X}_n = (X_1 + \cdots + X_n)/n$. Odhaduje, že směrodatná odchylka jednoho výběru je nejvýše 1 metr.

(a) Jak velké n má volit, aby směrodatná odchylka $\bar{X}_n$ byla nejvýše 1 cm?

(b) Pro jaké n zajistí Čebyševova nerovnost, že $\bar{X}_n$ se liší od h nejvýše o 5 cm s pravděpodobností alespoň 99 %? (Neboli $P(|\bar{X}_n - h| \leq 5) \geq 0.99$.)

(c) Statistik si všimne, že všichni měření lidé mají výšku v intervalu $(1.4, 2.1)$. Jak má upravit odhad směrodatné odchylky? Jak se změní odpovědi na předchozí otázky?

3. Odhadněte $\binom{100}{30}$ pomocí CLV. Nápověda: použijte CLV pro odhad $P(29.5 < X < 30.5)$ pro vhodnou n.v. X. Na druhou stranu pro $P(X = 30)$ máme vzorec $\binom{100}{30}/2^{100}$ z binomického rozdělení. Alternativně, můžete použít Moivre–Laplaceho větu. Pokud máte po ruce vhodný stroj, vyčíslete.

Soupis vzorečků

- **Markovova nerovnost:** $P(X \geq a\mathbb{E}(X)) \leq \frac{1}{a}$ pro $X \geq 0$.
- **Čebyševova nerovnost:** $P(|X - \mathbb{E}(X)| \geq t\sigma_X) \leq \frac{1}{t^2}$.
- **Centrální limitní věta:** Nechť $X_1, \dots, X_n$ jsou stejně rozdělené n.n.v. se střední hodnotou μ a rozptylem σ^2. Označme $Y_n = ((X_1 + \cdots + X_n) - n\mu)/(\sqrt{n} \cdot \sigma)$. Pak $Y_n \xrightarrow{d} N(0,1)$. Neboli

$$\lim_{n \to \infty} F_{Y_n}(x) = \Phi(x) \quad \text{pro každé } x \in \mathbb{R}.$$

- **De Moivre–Laplaceho věta** říká o něco silnější věc: v okolí pn je i pravděpodobnostní funkce $Bin(n,p)$ dobře aproximována hustotou $N(np, np(1-p))$, tedy vhodně přeškálovanou Gaussovou funkcí φ. Přesněji:

$$\binom{n}{k} p^k (1-p)^{n-k} \simeq \frac{1}{\sqrt{2\pi np(1-p)}} e^{-\frac{(k-np)^2}{2np(1-p)}}$$

  pro k blízké pn. Ještě přesněji: pokud pro nějaké c je $|k - pn| < c\sqrt{np(1-p)}$ a n se blíží nekonečnu, tak poměr dvou výrazů nahoře se blíží k jedné.