Limitní věty II.

19.4. a 23.4. 2018

1. Nechť má náhodná veličina X normální rozdělení $\mathsf{N}(1,4)$. Pomocí tabulek určete $\mathsf{P}(X<1)$, $\mathsf{P}(X>5)$ a $\mathsf{P}(|X|<2)$.

2. Nechť $X_1, X_2, \dots$ jsou nezávislé náhodné veličiny s normálním rozdělením $\mathsf{N}(\mu, \sigma^2)$.
   (a) Určete, s jakou pravděpodobností leží X_1 v intervalu $(\mu - 2\sigma, \mu + 2\sigma)$.
   (b) Jaké je rozdělení veličin $S_n = \sum_{i=1}^n X_i$, $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ a
   $$Z_n = \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma}?$$
   (c) Nechť $\mu = 1$ a $\sigma^2 = 4$. Jak velké n je třeba zvolit, abychom měli zaručeno, že $\overline{X}_n$ bude kladné číslo s pravděpodobností alespoň 0.99?

3. Jaká je pravděpodobnost, že při 10 000 hodech symetrickou mincí padne rub více než 4900 krát? Nalezněte nejprve přesné řešení (výraz) a následně řešte pomocí CLV.

4. Na server má přístup 100 uživatelů. Z dřívějších zkušeností víme, že uživatel má na serveru průměrně 120MB dat, směrodatná odchylka množství dat je 40 MB. Jak velký diskový prostor potřebujeme, aby s pravděpodobností větší než 99% nedošlo k jeho nedostatku?

5. Nechť ν_n značí relativní četnost líců v n hodech mincí (mince je symetrická a hody provádíme nezávisle). Zjistěte, kolik musíme provést hodů, aby pravděpodobnost jevu $\big[|\nu_n - 1/2| \le 0.01\big]$ byla alespoň 0.95? Řešte
   (a) pomocí Čebyševovy nerovnosti,
   (b) pomocí CLV.

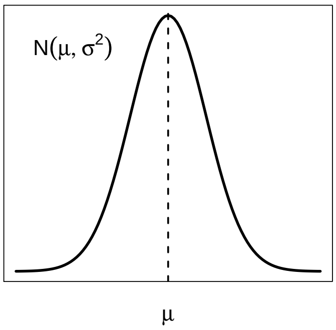
6. Životnost jedné žárovky má exponenciální rozdělení se střední hodnotou 10 hodin. Jakmile se jedna žárovka porouchá, nahradíme ji ihned další. Kolik máme zakoupit žárovek, abychom měli jistotu, že budeme moci svítit alespoň 600 hodin s pravděpodobností alespoň 95%?

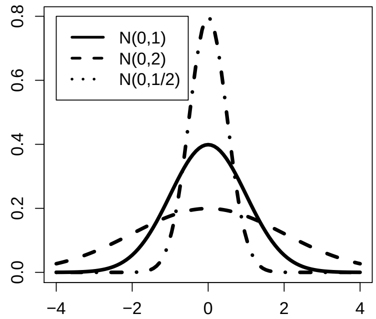
7. Pojišťovna má pojištěno 1 000 osob stejného věku. Pravděpodobnost úmrtí v daném roce je u každého pojištěného 0,01. Pojištěnci platí roční pojistné 1 200 Kč a v případě úmrtí je oprávněné osobě vyplaceno 80 000 Kč.
   (a) Jaký je v daném roce očekávaný zisk (resp. ztráta) pojišťovny?
   (b) Jaká je pravděpodobnost, že pojišťovna utrpí v daném roce ztrátu?
   (c) Jaký zisk je ředitel pojištovny schopen garantovat představenstvu s pravděpodobností 90 %?
   (d) Kolik by musela mít pojišťovna klientů (při stávajícím nastavení výše plateb), aby s pravděpodobností alespoň 99 % vydělala více než 10 000 Kč ročně?

Normální rozdělení. Normální rozdělení $\mathsf{N}(\mu, \sigma^2)$ je spojité rozdělení s hustotou

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

kde $\mu \in \mathbb{R}$ a $\sigma^2 > 0$ jsou parametry. Je-li $\mu = 0$ a $\sigma^2 = 1$, tj. $f(x) = \frac{1}{\sqrt{2\pi}} \exp\{-x^2/2\}$, pak se toto rozdělení nazývá **normované** normální rozdělení a značí se $\mathsf{N}(0,1)$.





- Je-li $X \sim \mathsf{N}(\mu, \sigma^2)$, pak $\mathsf{E}X = \mu$ a $\mathsf{Var}\, X = \sigma^2$.
  Dále $aX + b \sim \mathsf{N}(a\mu + b, a^2\sigma^2)$ pro všechna $a, b \in \mathbb{R}$, $a \neq 0$. Speciálně, $\dfrac{X-\mu}{\sigma} \sim \mathsf{N}(0,1)$.
- Jsou-li X, Y nezávislé normálně rozdělené a $a, b \in \mathbb{R}$, pak $aX + bY$ má normální rozdělení (s příslušnými parametry).
- **Distribuční funkce** rozdělení $\mathsf{N}(0,1)$ se značí jako Φ, tj. $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\{-t^2/2\} dt$. Tento určitý integrál je možné spočítat jen **numericky**, a proto hodnoty funkce Φ nalezneme **v tabulkách** (nebo dostaneme pomocí vhodného softwaru). Ze symetrie platí

$$\Phi(x) + \Phi(-x) = 1 \qquad \text{pro všechna } x \in \mathbb{R}.$$

- Hodnoty $\Phi^{-1}(\alpha) = q_\alpha$ (tzv. kvantily $\mathsf{N}(0,1)$) jsou také uvedeny v tabulkách a platí $q_\alpha = -q_{1-\alpha}$.

Centrální limitní věta (CLV): Nechť $\{X_n\}$ je posloupnost nezávislých stejně rozdělených náhodných veličin s $0 < \mathsf{Var}\, X_1 < \infty$. Pak

$$\mathsf{P}\left(\frac{\sum_{k=1}^n X_k - n\mathsf{E}X_1}{\sqrt{n\,\mathsf{Var}\, X_1}} \leq x\right) \xrightarrow{n\to\infty} \Phi(x), \quad x \in \mathbb{R},$$

neboli ekvivalentně

$$\mathsf{P}\left(\sqrt{n} \cdot \frac{\overline{X}_n - \mathsf{E}X_1}{\sqrt{\mathsf{Var}\, X_1}} \leq x\right) \xrightarrow{n\to\infty} \Phi(x), \quad x \in \mathbb{R},$$

kde Φ je distribuční funkci normálního rozdělení $\mathsf{N}(0,1)$.
Zkráceně píšeme

$$Z_n := \frac{\sum_{k=1}^n X_k - n\mathsf{E}X_1}{\sqrt{n\mathsf{Var}\, X_1}} = \sqrt{n} \cdot \frac{\overline{X}_n - \mathsf{E}X_1}{\sqrt{\mathsf{Var}\, X_1}} \overset{asympt.}{\sim} \mathsf{N}(0,1)$$

a říkáme, že Z_n má asymptoticky normální rozdělení $\mathsf{N}(0,1)$. CLV nám tedy říká, že distribuční funkce veličiny Z_n se při $n \to \infty$ blíží k Φ, tedy $\{Z_n\}$ konvergují v distribuci k náhodné veličině s $\mathsf{N}(0,1)$ rozdělením. Pro n dost velké tedy lze uvažovat

$$\mathsf{P}(Z_n \leq x) \doteq \Phi(x).$$