Normální rozdělení, Čebyševova nerovnost

23. – 27. 11. 2015

Normální rozdělení

1. Výška desetiletých chlapců v cm je náhodná veličina s normálním rozdělením $\mathsf{N}(136.1, 6.4^2)$.
   (a) S jakou pravděpodobností má náhodně vybraný desetiletý chlapec více než 150 cm?
   (b) S jakou pravděpodobností má desetiletý chlapec výšku v rozmezí 130 cm až 140 cm?
   (c) V obchodě prodávají konfekční velikosti oblečení na chlapce ve výškovém rozmezí 120 cm až 150 cm. S jakou pravděpodobností si náhodně vybraný desetiletý chlapec v obchodě nebude moci koupit oblečení, protože mu bude moc malé nebo moc velké?
   (d) Co říkají pravidlo 2 sigma a pravidlo 3 sigma pro výšku desetiletých chlapců?
   (e) Jakou výšku by musel mít Váš desetiletý bratr, aby patřil k 5 % nejvyšších v populaci?
   (f) Jakou výšku by musel mít Váš desetiletý bratr, aby patřil k 10 % nejnižších v populaci?

Čebyševova nerovnost

2. Náhodná veličina X má normální rozdělení $\mathsf{N}(\mu, \sigma^2)$. Odhadněte pravděpodobnosti

$$\mathsf{P}(|X-\mu| > 2\sigma) \quad \text{a} \quad \mathsf{P}(|X-\mu| > 3\sigma)$$

   pomocí Čebyševovy nerovnosti a proveďte srovnání s přesným výsledkem.

3. Hodíme n krát pravidelnou symetrickou mincí. Označíme ν_n relativní četnost líců, tj. $\nu_n = (\#\text{líců})/n$.
   (a) Spočítejte střední hodnotu a rozptyl ν_n.
   (b) K jakému číslu se bude ν_n blížit, budeme-li zvyšovat počet opakování n?
   (c) Jestliže provedeme $n = 100$ hodů, s jakou pravděpodobností dostaneme číslo, které je od 1/2 vzdálené o více než 0.1? (Řešte pomocí Čebyševovy nerovnosti).
   (d) Kolik musíme provést hodů, aby pravděpodobnost jevu $[|\nu_n - 1/2| \leq 0.1]$ byla alespoň 0.95? (Řešte pomocí Čebyševovy nerovnosti).

Tabulka distribuční funkce a kvantilové funkce $\mathsf{N}(0,1)$

x	0.000	0.100	0.200	0.300	0.400	0.500	0.600	0.700	0.800	0.900
$\Phi(x)$	0.500	0.540	0.579	0.618	0.655	0.691	0.726	0.758	0.788	0.816
x	0.100	0.200	0.300	0.400	0.500	0.600	0.700	0.800	0.900	1.000
$\Phi(x)$	0.540	0.579	0.618	0.655	0.691	0.726	0.758	0.788	0.816	0.841
x	1.100	1.200	1.300	1.400	1.500	1.600	1.700	1.800	1.900	2.000
$\Phi(x)$	0.864	0.885	0.903	0.919	0.933	0.945	0.955	0.964	0.971	0.977
x	2.100	2.200	2.300	2.400	2.500	2.600	2.700	2.800	2.900	3.000
$\Phi(x)$	0.982	0.986	0.989	0.992	0.994	0.995	0.997	0.997	0.998	0.999

α	0.5	0.8	0.9	0.95	0.975	0.99	0.995
q_α	0	0.842	1.282	1.645	1.960	2.326	2.576

Opakování z přednášky

Normální rozdělení. Normální rozdělení $\mathsf{N}(\mu, \sigma^2)$ má hustotu

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

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 $\mathsf{Var}\, X = \sigma^2$.
– Je-li $X \sim \mathsf{N}(\mu, \sigma^2)$ a $a, b \in \mathbb{R}$, $a \neq 0$, pak

$$aX + b \sim \mathsf{N}(a\mu + b, a^2\sigma^2),$$
$$\frac{X-\mu}{\sigma} \sim \mathsf{N}(0, 1).$$

– **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

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

Z hodnot distribuční funkce vyplývá, že pro náhodnou veličinu X s rozdělením $\mathsf{N}(\mu, \sigma^2)$ platí

$$\mathsf{P}[\mu - 2\sigma \leq X \leq \mu + 2\sigma] \approx 0.95$$
$$\mathsf{P}[\mu - 3\sigma \leq X \leq \mu + 3\sigma] \approx 0.997,$$

(odtud též tzv. pravidlo dvou sigma nebo pravidlo tří sigma).
– **Kvantily** q_α normálního rozdělení $\mathsf{N}(0, 1)$ jsou také uvedeny v tabulkách a platí $q_\alpha = -q_{1-\alpha}$
– Jsou-li X, Y nezávislé normálně rozdělené a $a, b \in \mathbb{R}$, pak $aX + bY$ má normální rozdělení (s příslušnými parametry).
– Normální rozdělení má v pravděpodobnosti zcela zásadní význam, jak ilustruje např. tzv. **centrální limitní věta**.

Čebyševova nerovnost: Je-li X náhodná veličina, pro kterou $0 < \mathsf{Var}\, X < \infty$, pak

$$\mathsf{P}(|X - \mathsf{E}X| > \varepsilon) \leq \frac{\mathsf{Var}\, X}{\varepsilon^2} \quad \text{pro všechna } \varepsilon > 0.$$