Centrální limitní věta a opakování

12.12.2018

1. 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%?

2. Nechť $X_1, \ldots, X_n$ je náhodný výběr z rovnoměrného rozdělení na intervalu $[\theta - 1, \theta + 1]$, kde $\theta \in \mathbb{R}$ je nějaký neznámý parametr.
   (a) Nalezněte odhad $\widehat{\theta}_n$ parametru θ momentovou metodou a vyšetřete vlastnosti takového odhadu.
   (b) Kolik alespoň pozorování musíme do odhadu zapojit, aby pravděpodobnost, že je náš odhad $\widehat{\theta}_n$ od skutečné hodnoty θ o dál než 0,1, byla menší než 5 %.

3. Zajímá nás odhad pravděpodobnosti p výskytu krátkozrakosti u náhodně vybraného obyvatele ČR ve věku 18–40 let. Navrhněte experiment, pomocí kterého získáme takový odhad, a zjistěte, kolik osob musíme do experimentu zapojit, aby byl náš odhad natolik přesný, že je od skutečné hodnoty p vzdálen o méně než $0,01$ s pravděpodobností větší než 90 %.

4. Mezi studenty zapsanými na předmět Pravděpodobnost a matematická statistika se vyskytují dívky a chlapci v poměru 1:4. Dále budeme předpokládat, že jedna čtvrtina studentů píše levou rukou. Označme jako X identifikátor toho, zda je daná osoba chlapec, Y identifikátor, zda daná osoba píše levou rukou, a p pravděpodobnost, že je daná osoba chlapec píšící levou rukou.
   (a) Zapište sdružené rozdělení náhodného vektoru $(X, Y)^\top$. Jaká konkrétní hodnota p odpovídá situaci, kdy by byly veličiny X a Y nezávislé?
   (b) Za předpokladu nezávislosti X a Y spočtěte (přibližně) s jakou pravděpodobností bude mezi 36 studenty na zkoušce více než 6 chlapců píšících levou rukou.
   (c) Nyní předpokládejte, že je $\mathsf{Corr}\,(X, Y) = 1/\sqrt{3}$. Určete hodnotu p a podmíněnou pravděpodobnost, že osoba píše levou rukou, víme-li, že je to chlapec.
   (d) Pro situaci z (c) spočtěte pravděpodobnost z (b).
   (e) Zopakujte výpočet pro $\mathsf{Corr}\,(X, Y) = -1/\sqrt{3}$.

5. Doba výpočtu (v sekundách) jedné úlohy s náhodným vstupem je náhodná veličina s rozdělením s hustotou $f(x) = 2/x^3$ pro $x \geq 1$ a $f(x) = 0$ jinak. Budeme spouštět úlohy sériově (jednu za druhou) a zajímá nás, s jakou pravděpodobností bude celková doba výpočtu 100 úloh delší než 250 sekund. Můžeme využít centrální limitní větu?

Opakování

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

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

neboli ekvivalentně

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

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 F_n veličiny Z_n se při $n \to \infty$ blíží k Φ. Pro n dost velké tedy lze uvažovat

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

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
$\Phi^{-1}(\alpha) = q_\alpha$	0	0.842	1.282	1.645	1.960	2.326	2.576

Platí

$$\Phi(x) + \Phi(-x) = 1, \quad q_\alpha = -q_{1-\alpha}.$$