

Analýza a vizualizácia dát

26. februára 2024

- ▶ rozcvičky na konci hodiny:
 - 11.3.2024
 - 8.4.2024
 - 29.4.2024
 - 13.5.2024 - opravná / náhradná

náhodné premenné, rozdelenia pravdepodobnosti

- ▶ diskkrétne
- ▶ spojité
- ▶ ani diskkrétne, ani spojité

funkcie spojené s rozdeleniami

- ▶ hustota, pravdepodobnostná fcia, kumulatívne fcie - *ROZDELENIE.DIST()*
- ▶ kvantilová fcia (inverzná k distribučnej) - *ROZDELENIE.INV()*

cvičenie

- ▶ vykreslite pravdepodobnostné funkcie binomického rozdelenia s rôznymi hodnotami parametrov
- ▶ demonštrujte vizuálne De Moivreovu - Laplaceovu vetu
- ▶ vykreslite kumulatívnu distribučnú funkciu $bin(10; 0,2)$
- ▶ upravte grafy tak, aby boli v prezentovateľnej podobe
- ▶ uložte vytvorené grafy

cvičenie

- ▶ vykreslite hustoty χ^2 -rozdelenia s rôznymi hodnotami parametrov
- ▶ demonštrujte vizuálne fakt, že jeho limitným rozdelením je normálne rozdelenie
- ▶ vykreslite kvantilovú funkciu χ^2 -rozdelenia

cvičenie

- vykreslite hustotu dvojrozmerného normálneho rozdelenia

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \right).$$

$$c = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$$

$$f(x, y) = c \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[\left(\frac{x-\mu_1}{\sigma_1} \right)^2 + \left(\frac{y-\mu_2}{\sigma_2} \right)^2 - 2\rho \left(\frac{x-\mu_1}{\sigma_1} \right) \left(\frac{y-\mu_2}{\sigma_2} \right) \right] \right\}$$

samostatné cvičenie

- ▶ vykreslite pravdepodobnostnú funkciu Poissonovho rozdelenia pre rôzne hodnoty parametrov v jednom grafe
- ▶ vizuálne demonštrujte, že Poissonovo rozdelenie je limitným rozdelením binomického (a normálne limitným rozdelením Poissonovho)
- ▶ vykreslite hustotu normálneho rozdelenia (výška žien, mužov) s parametrami
 - $\mu = 167$; $sd = 5,5$
 - $\mu = 181$; $sd = 6,4$
- ▶ vykreslite hustoty Studentovho t-rozdelenia s rôznymi hodnotami parametrov a demonštrujte vizuálne fakt, že jeho limitným rozdelením je normálne rozdelenie
- ▶ upravte grafy tak, aby boli v prezentovateľnej podobe
- ▶ uložte vytvorené grafy

pozor

- ▶ pmf (probability mass function) alebo cdf (cumulative distribution function)
- ▶ nie pre každý typ grafu sú rovnaké možnosti (scatter plot a line chart majú iné možnosti, hoci vedia vyzeráť rovnako)
- ▶ či sa správne zobrazuje to, čo sa má
- ▶ prezentácia grafu