Zkouškový test z PMS

na vypracování máte 90 min

16.2.2017

.

T1 Jevy A, B, C jsou po dvou nezávislé a platí $P(A|B\cap C) = \frac{1}{2}$, $P(B|C) = 0.75$, $P(C|B) = \frac{2}{3}$, $P(A\cup B) = \frac{7}{8}$. Jsou jevy A, B, C nezávislé? Dokažte.

T2 Házím kostkou do té doby, než padne něco jiného než šestka.

a) Jaká je pravděpodobnost, že hodím kostkou alespoň 4x?

b) Jaká je střední hodnota počtu hodů?

c) Jaká je střední hodnota součtu hodů?

T3 Náhodná veličina X má rozdělení dané hustotou

$$f_X(u) = \begin{cases} e^{2u}\,, & u < 0\,, \\ cx^2\,, & u \in [1,2]\,, \\ 0\,, & \text{jinak}\,. \end{cases}$$

Určete konstantu c, distribuční funkci F_X a rozdělení veličiny $Y = g(X)$, kde funkce g je definována následovně:

$$g(x) = \begin{cases} 0\,, & x \le 0\,, \\ 1\,, & x > 0\,. \end{cases}$$

T4 V jisté firmě přichází lidé do práce rovnoměrně mezi 8:45 a 9:10. Určete minimální počet lidí, kteří musí do práce daný den přijít, aby se dalo s pravděpodobností 95% říci, že většina z nich přišla do práce včas, tj. do 9:00.

T5 Chyba měření hmotnosti součástek má normální rozdělení. Chyby se vyskytují nezávisle. Provedli jsme měření 7 součástek, z nichž po analýze vyplynuly hodnoty chyb v gramech:

$$1,\ -3,\ -1,\ 0.5,\ 1,\ -2,\ 1.$$

Testujte na hladině $\alpha = 0.01$, zda měření je zatíženo systematickou chybou.

Řešení

T1 Stačí ověřit $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$, neboť pro dvojice již podobná formule platí. Víme, že $P(B) = P(B|C) = 0.75$ a $P(C) = P(C|B) = \frac{2}{3}$, neboť jsou jevy A, B, C po dvou nezávislé. Ze stejného důvodu je dále

$$\frac{1}{2} = P(A|B \cap C) = \frac{P(A \cap B \cap C)}{P(B \cap C)} = \frac{P(A \cap B \cap C)}{P(B) \cdot P(C)},$$

z čehož plyne

$$P(A \cap B \cap C) = \frac{1}{2} P(B) \cdot P(C).$$

Protože $\frac{7}{8} = P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) - P(A) \cdot P(B)$ (jevy A, B jsou nezávislé), snadnými úpravami získáme $P(A) = \frac{1}{2}$, což nám dává $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$ a jevy A, B, C jsou tedy nezávislé.

T2

a) To je pravděpodobnost, že v prvních třech hodech hodím pokaždé šestku, tedy

$$P(\text{hodím alespoň 4x kostkou}) = \left(\frac{1}{6}\right)^3.$$

b) Jedná se o střední hodnotu geometrického rozdělení s parametrem $p = \frac{5}{6}$ (pravděpodobnost "úspěchu/ukončení hry"). Tedy průměrný počet hodů je $\frac{1}{p} = \frac{6}{5}$.

c) Označme průměrný součet hodů μ. Potom z principu hry plyne $\mu = \frac{1}{6}(1+2+3+4+5+6) + \frac{1}{6}\mu$. Tato rovnice má dvě řešení: $\mu = \frac{21}{5}$ a $\mu = \infty$. Je třeba dokázat, že druhá možnost nenastane. Hodnotu μ lze vyjádřit jako

$$\mu = \sum_{\substack{n=1 \\ 6 \nmid n}}^{\infty} \frac{n}{6^{[n/6]}} = \sum_{n=0}^{\infty} \frac{15 + 6n}{6^{n+1}},$$

kde $[x]$ označuje horní celou část reálného čísla x. Tato řada je konvergentní (stačí použít například podílové srovnávací kritérium nebo si uvědomit že je součtem konvergentní geometrické řady a derivace konvergentní geometrické řady), proto $\mu = \frac{21}{5}$.

T3 Pro hustotu f_X musí platit:

$$1 = \int_{-\infty}^{\infty} f_X(u)\, \mathrm{d}u = \int_{-\infty}^{0} e^{2u}\, \mathrm{d}u + \int_{1}^{2} cu^2\, \mathrm{d}u = \left[\frac{e^{2u}}{2}\right]_{-\infty}^{0} + \left[c\frac{u^3}{3}\right]_{1}^{2} = \frac{1}{2} + c\frac{7}{3}.$$

Z toho dostáváme $c = \frac{3}{14}$. Distribuční funkci získáme jednoduše integrací

$$F_X(u) = \int_{-\infty}^{u} f_X(x)\, \mathrm{d}x = \begin{cases} \frac{e^{2u}}{2}\,, & u < 0\,, \\ \frac{1}{2}\,, & u \in [0,1)\,, \\ \frac{6+u^3}{14}\,, & u \in [1,2)\,, \\ 1\,, & u \geq 1\,. \end{cases}$$

Rozdělení veličiny Y je alternativní s parametrem $\frac{1}{2}$, neboť funkce g nabývá pouze hodnot 0 a 1 a platí $P(Y = 0) = P(g(X) = 0) = P(X \leq 0) = \frac{1}{2}$ a $P(Y = 1) = P(X > 0) = \frac{1}{2}$.

T4 Označme X_i, $i \in \{1, ..., n\}$ náhodnou veličinu, která nabývá jedničky, pokud i-tý pracovník přišel do práce včas a nuly, pokud pozdě. Abychom k výpočtu mohli použít centrální limitní větu, je třeba předpokládat, že příchody zaměstnanců do práce jsou nezávislé (což v praxi většinou nejsou! - jedou spolu autem, přijíždějí stejnou tramvají...). Za tohoto předpokladu hledáme minimální rozsah výběru $X_1, ..., X_n$, aby platilo $P\left(\overline{X} > \frac{1}{2}\right) \geq 95\%$, tj. $P\left(\overline{X} \leq \frac{1}{2}\right) < 0.05$. Jest pro každé $i \in \{1, ..., n\}$ pravda $P(X_i = 1) = \frac{15}{25} = 0.6 =: p$. Aproximací pomocí CLV dostáváme

$$P\left(\overline{X} \leq \frac{1}{2}\right) = P\left(\frac{\overline{X} - p}{\sqrt{(1-p)p}}\sqrt{n} \leq \frac{1/2 - p}{\sqrt{(1-p)p}}\sqrt{n}\right) = \Phi\left(\frac{1/2 - p}{\sqrt{(1-p)p}}\sqrt{n}\right) = \Phi\left(\frac{-1}{2\sqrt{6}}\sqrt{n}\right) < 0.05,$$

z čehož (složením poslední rovnice s kvantilovou funkcí $\mathrm{N}(0,1)$) plyne

$$\frac{-1}{2\sqrt{6}}\sqrt{n} < -1.65$$
$$n > 65.34. \Rightarrow n \geq 66.$$

T5 Úlohu řešíme jednovýběrovým T-testem. Označíme-li náhodný výběr $X_1, ..., X_7,\ X_i \sim \mathrm{N}(\mu, \sigma^2)$ pro $i = 1, ..., 7$, potom testujeme hypotézu $H_0:\ \mu = 0$ proti alternativě $H_1:\ \mu \neq 0$. Spočteme hodnoty $\overline{X} = \frac{-5}{14}$ a $S_n^2 = \frac{1}{6}\left(\sum_{i=1}^{7} X_i^2 - 7\overline{X}^2\right) = \frac{1}{6}\left(17.25 - \frac{6.25}{7}\right) \doteq 2.73$. Za platnosti nulové hypotézy má testovací statistika

$$T = \frac{\overline{X} - 0}{S_n}\sqrt{n} = \frac{-5/14}{1.65}\sqrt{7} = -0.57$$

Studentovo t-rozdělení o 6 stupních volnosti. Porovnáním hodnoty $|T| = 0.57$ s kvantilem $q_{t_6}(0.99) = 3.143$ dostáváme $|T| < q_{t_6}(0.99)$, což znamená, že na hladině spolehlivosti 0.01 nelze tvrdit, že je měření zatíženo systematickou chybou.