Zápočtová písemka z PST 30.4.2018

T1 V populaci trpí 3% jistou chorobou A. Vyštření na chorobu A ukáže výsledek správně v 70% případů, pokud člověk chorobou A trpí a v 90% případů, pokud chorobou A netrpí.

a) Náhodně zvolíme člověka a vyšetříme jej na chorobu A. Jaká je pravděpodobnost, že test vyjde pozitivní?

b) Test u daného člověka vyšel pozitivní. Protože léčba je narozdíl od vyšetření zatěžující, chceme se přesvědčit, že je člověk opravdu nemocný a vyšetříme jej tedy znovu. Test 2 vyjde pozitvně. Jaká je pravděpodobnost, že chorobou A netrpí?

T2 Hodíme 4x šestistěnnou kostkou a 2x mincí. Rub platí jako bod, líc ne. Náhodná veličina X udává počet bodů, kolik jsme dohromady hodili. Určete $\mathbb{E}X$, $\mathrm{D}\,X$ a $\mathbb{E}X^2$.

T3 Je-li ze sta výrobků průměrně 6 vadných, určete praveděpodobnost, že z tisíce výrobků jich bude vadných mezi 60 a 68.

Řešení

T1 Označme jev, že náhodně vybraný člověk z populace trpí chorobou A jako jev A a to, že má pozitivní resp. negativní test jako jev T_+, resp. T_-. Označme také T_{++} jako jev, že náhodně vybraný člověk, který je podroben dvojímu testování vyjdou oba testy pozitivně.

a) Dle věty o úplné pravděpodobnosti máme

$$P(T_+) = P(T_+|A)P(A) + P(T_+|\overline{A})P(\overline{A}) = 0.7 \cdot 0.03 + 0.1 \cdot 0.97 = 0.118,$$

kde $\overline{A}$ značí doplněk jevu A.

b) Chceme získat podmíněnou pravděpodobnost $P(\overline{A}|T_{++})$. Podle Bayesova vzorce dostáváme

$$P(\overline{A}|T_{++}) = \frac{P(T_{++}|\overline{A})P(\overline{A})}{P(T_{++})}.$$

Za předpokladu, že daná osoba trpí chorobou A, je pravděpodobnost, že jí druhé testování vyjde pozitivní, stejná jako u prvního testovaní, konkrétně $P(T_+|A)$. Obdobně tomu je v případě, pokud daná osoba chorobou A netrpí, tedy pravděpodobnost, že výsledek kteréhokoliv testování jí vyjde pozitivně, je rovna $P(T_{++}|\overline{A})$. Proto máme $P(T_{++}|\overline{A}) = P(T_+|\overline{A})^2$ a podobně opět podle věty o úplné pravděpodobnosti

$$\begin{aligned} P(T_{++}) &= P(T_{++}|A)P(A) + P(T_{++}|\overline{A})P(\overline{A}) = P(T_+|A)^2P(A) + P(T_+|\overline{A})^2P(\overline{A}) \\ &= (0.7)^2 \cdot 0.03 + (0.1)^2 \cdot 0.97 = 0.0244. \end{aligned}$$

Dosazením do původního vzorce dostáváme

$$P(\overline{A}|T_{++}) = \frac{P(T_+|\overline{A})^2P(\overline{A})}{P(T_{++})} = \frac{(0.1)^2 \cdot 0.97}{0.0244} = 0.398.$$

T2 Označme X_i $i \in \{1, 2, 3, 4\}$ výsledek i-tého hodu kostkou a X_5, resp. X_6 počet bodů získaných prvním, resp. druhým hodem mincí. Potom máme

$$X = \sum_{i=1}^{6} X_i.$$

Platí $\mathbb{E}X = \sum_{i=1}^{6} \mathbb{E}X_i = 4 \cdot 3.5 + 2 \cdot 0.5 = 15$, neboť $\mathbb{E}X_1 = ... = \mathbb{E}X_4 = 3.5$ a $\mathbb{E}X_5 = \mathbb{E}X_6 = 0.5$ (je snadno vidět, lze ověřit výpočtem). Neboť jsou veličiny $X_1, ..., X_6$ nezávislé, rozptyl X dostaneme jako součet rozptylů veličin X_i. Máme

$$\mathbb{E}X_1^2 = \sum_{a \in \mathbb{R}} a^2 P(X_1 = a) = \frac{1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2}{6} = \frac{91}{6},$$

tedy $\mathrm{D}\, X_1 = \mathbb{E}X_1^2 - (\mathbb{E}X_1)^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}$. Obdobně

$$\mathbb{E}X_5^2 = \sum_{a \in \mathbb{R}} a^2 P(X_5 = a) = \frac{0^2 + 1^2}{2} = \frac{1}{2}$$

a $\mathrm{D}\, X_5 = \mathbb{E}X_5^2 - (\mathbb{E}X_5)^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$. Celkově tedy máme

$$\mathrm{D}\, X = \sum_{i=1}^{6} \mathrm{D}\, X_i = 4 \cdot \frac{35}{12} + 2 \cdot \frac{1}{4} = \frac{73}{6},$$

$$\mathbb{E}X^2 = \mathrm{D}\, X + (\mathbb{E}X)^2 = \frac{73}{6} + (15)^2 = 237.166.$$

T3 Označme X_i, $i \in \{1, ...1000\}$ veličinu, která nabývá hodnoty 0, pokud i-tý výrobek není vadný a 1 pokud je vadný. Veličiny X_i mají alternativní (nula-jedničkové) rozdělení s parametrem $p = 0.06$. Ptáme se na pravděpodobnost $P(60 \leq \sum_{i=1}^{1000} X_i \leq 68)$. Víme, že je $\mathbb{E}X_i = p = 0.06$ a $\mathrm{D}\, X_i = p(1-p) = 0.0564$ pro všechna $i \in \{1, ..., 1000\}$. Chceme-li použít centrální limitní větu, musíme předpokládat, že vady na výrobcích se vyskytují nezávisle. Potom má veličina

$$U = \frac{\sum_{i=1}^{1000} X_i - 1000 \cdot 0.06}{\sqrt{1000 \cdot 0.0594}} = \frac{\sum_{i=1}^{1000} X_i - 60}{7.707}$$

přibližně rozdělení $\mathrm{N}(0,1)$ a je tedy

$$P\left(60 \leq \sum_{i=1}^{1000} X_i \leq 68\right) = P\left(0 \leq U \leq \frac{8}{\sqrt{56.4}}\right) = \Phi(1.065) - \Phi(0) = 0.85 - \frac{1}{2} = 35\%.$$