2. kolokvij iz OVS (1. 6. 2010, rešitve)

1. Približno 8% ljudi v Sloveniji je levičarjev. Ljubljana ima 280000 prebivalcev. Oceni verjetnost, da je v Ljubljani med 22500 in 23000 levičarjev (predpostavi, da prebivalci Ljubljane predstavljajo naključni vzorec prebivalcev Slovenije).

   Rešitev: Predpostavljamo, da je vsak prebivalec Ljubljane 'indikatorska slučajna spremenljivka' s porazdelitveno shemo

   $$I_k \sim \begin{pmatrix} 0 & 1 \\ 0.92 & 0.08 \end{pmatrix}.$$

   Število levičarjev v Ljubljani je potem $X = I_1 + I_2 + \cdots + I_{280000}$. Vemo, da je $E(I_k) = p = 0.08$ in $D(I_k) = p(1-p) = 0.08 \cdot 0.92 = 0.0736$, zato je $E(X) = 280000 \cdot E(I_k) = 22400$ in $D(X) = 280000 \cdot D(I_k) = 20608$, saj so I_k neodvisne. Po CLI je $X \dot{\sim} N(\mu, \sigma)$, za $\mu = E(X) = 22400$ in $\sigma = \sqrt{D(X)} = 143.555$.

   Torej: $P(22500 \leq X \leq 23000) = P(X \leq 23000) - P(X \leq 22500) \doteq \Phi\left(\frac{23000-\mu}{\sigma}\right) - \Phi\left(\frac{22500-\mu}{\sigma}\right) = \Phi(4.18) - \Phi(0.697) = 0.243$.

2. Stehtali smo vzorec 2000-glave populacije belih miši in dobili naslednje rezultate (v gramih).

   29.3 29.7 30.1 28.9 30.3 29.4 29.9 31.1 29.3 29.5 30.2

   a) Izračunaj vzorčno povprečje in popravljeni vzorčni standardni odklon.

   b) Izračunaj interval zaupanja za standardni odklon pri 95% stopnji zaupanja.

   Rešitev: a) Vzorčno povprečje $\overline{X} = 29.79$, popravljen vzorčni standardni odklon $\hat{S} = 0.614$.
   b) Interval zaupanja za σ pri stopnji zaupanja $\gamma = 0.95$ je $\left[\frac{\sqrt{n-1}\cdot\hat{S}}{\sqrt{\chi^2_{0.975}(n-1)}}, \frac{\sqrt{n-1}\cdot\hat{S}}{\sqrt{\chi^2_{0.025}(n-1)}}\right]$. V našem primeru je $n = 11$ in dobimo interval zaupanja $[0.429, 1.08]$.

3. Predpostavljajmo, da je reakcijski čas voznikov porazdeljen normalno s povprečjem 1.2 s in standardnim odklonom 0.2 s.

   a) Izračunaj verjetnost, da pri izbranem enostavnem naključnem vzorcu velikosti 50 voznikov izračunamo vzorčno povprečje, manjše od 1.1 s.

   b) Oceni delež voznikov, ki imajo reakcijski čas daljši od 1.4 s.

   Rešitev: a) Za vzorec 50 voznikov je vzorčno povprečje $\overline{X}$ porazdeljeno kot $\overline{X} \sim N(1.2, \frac{0.2}{\sqrt{50}})$. Sledi: $P(\overline{X} \leq 1.1) = \Phi\left(\frac{1.1-1.2}{0.2/\sqrt{50}}\right) = \Phi(-3.536) = 1 - \Phi(3.536) = 0.0002$ ali 0.2%.
   b) $P(\text{reakcijski čas} \geq 1.4) = 1 - P(\text{reakcijski čas} \leq 1.4) = 1 - \Phi\left(\frac{1.4-1.2}{0.2}\right) = 1 - \Phi(1) = 0.1587$, tj. skoraj 16% voznikov.

4. Pri igri rulete je verjetnost, da se kroglica ustavi na črnem polju $\frac{18}{37}$, na rdečem polju prav tako $\frac{18}{37}$ in na zelenem polju $\frac{1}{37}$. Če je ta pogoj izpolnjen, igro imenujemo pravična ruleta. Igralnica je zelo zainteresirana, da je ruleta prevična, saj lahko le tako prepreči, da bi se kakšen pozorno opazujoč igralec okoristil z načrtnimi stavami. Zato v igralnici beležijo izide naključno izbranih iger. V bazi imamo 74 iger, od katerih se je pri 37 kroglica ustavila na rdečem, pri 32 na črnem in pri 5 na zelenem polju. Ali lahko s stopnjo značilnosti 0.05 trdiš, da njihova ruleta ni pravična?

   Rešitev: Iz podatkov dobimo tako tabelo:

Barva polja	črna	rdeča	zelena	
Frekvenca	32	37	5	74
Pričakovana f.	36	36	2	74

   Testiramo hipotezo $H_0 : X \sim \begin{pmatrix} \text{črna} & \text{rdeča} & \text{zelena} \\ \frac{18}{37} & \frac{18}{37} & \frac{1}{37} \end{pmatrix}$ pri stopnji značilnosti 0.05. Naredimo χ^2-test:

   $$\chi^2 = \frac{(32-36)^2}{36} + \frac{(37-36)^2}{36} + \frac{(5-2)^2}{2} = 4.972.$$

   Poleg tega je $\chi^2_{1-0.05}(3-1) = \chi^2_{0.95}(2) = 5.991$. Ker je $4.972 = \chi^2 < \chi^2_{0.95}(2) = 5.991$ hipoteze ne zavrnemo in *ne moremo trditi, da ruleta ni pravična.*