Rešitve kolokvija iz verjetnosti in statistike z dne 6. 6. 2013

Matematika – univerzitetni študij

1. Najprej izračunajmo:

$$E(X_i) = a\,, \qquad D(X_i) = E\big[(X_i - a)^2\big] = \frac{a^2+2}{3}\,.$$

Če z $\bar{X}$ označimo aritmetično sredino, torej velja $E(\bar{X}) = a$ in $D(\bar{X}) = (a^2+2)/300$. Po centralnem limitnem izreku je približno $\bar{X} \sim \mathrm{N}\big(a, \sqrt{(a^2+2)/300}\big)$, torej je:

$$P\left(\bar{X} > \frac{11a}{10}\right) \approx \frac{1}{2} - \Phi\left(\frac{\frac{a}{10}}{\sqrt{\frac{a^2+2}{300}}}\right) = \frac{1}{2} - \Phi\left(\frac{a\sqrt{3}}{\sqrt{a^2+2}}\right).$$

Torej mora biti približno $\Phi\big(a\sqrt{3}/\sqrt{a^2+2}\big) = 0^{\cdot}45$ oziroma približno $a\sqrt{3}/\sqrt{a^2+2} = z_{0.95} \doteq 1^{\cdot}645$, kar je res pri $a = z_{0.95}\sqrt{2/(3 - z_{0.95}^2)} \doteq 4^{\cdot}3$.

2. Iz gostote dobimo logaritem verjetja:

$$\ln L = n \ln \frac{\sin(\pi a)}{\pi} + (a-1)\sum_{i=1}^n \ln X_i - a \sum_{i=1}^n \ln(1 - X_i)$$

Ta funkcija gre proti minus nekončno, ko gre a proti 0 ali proti 1. Torej bo maksimum zagotovo dosežen tam, kjer je odvod:

$$\frac{\mathrm{d}\ln L}{\mathrm{d}a} = n\pi\,\mathrm{ctg}(\pi a) + \sum_{i=1}^n \ln \frac{X_i}{1 - X_i}$$

enak nič, to pa se zgodi, ko je a enak:

$$\begin{aligned}
\hat{a} &= \frac{1}{\pi}\,\mathrm{arcctg}\left(-\frac{1}{n\pi}\sum_{i=1}^n \ln\frac{X_i}{1-X_i}\right) = \\
&= \frac{1}{\pi}\,\mathrm{arcctg}\left(\frac{1}{n\pi}\sum_{i=1}^n \ln\frac{1-X_i}{X_i}\right) = \\
&= \frac{1}{2} + \frac{1}{\pi}\,\mathrm{arctg}\left(\frac{1}{n\pi}\sum_{i=1}^n \ln\frac{X_i}{1-X_i}\right).
\end{aligned}$$

Izražava:

$$\hat{a} = \frac{1}{\pi}\,\mathrm{arctg}\left(-\frac{n\pi}{\sum_{i=1}^n \ln\frac{X_i}{1-X_i}}\right) = \frac{1}{\pi}\,\mathrm{arctg}\left(\frac{n\pi}{\sum_{i=1}^n \ln\frac{1-X_i}{X_i}}\right)$$

pa ne da vedno pravilnega rezultata, saj je tedaj $-\frac{1}{2} < a < \frac{1}{2}$, medtem ko mora biti v našem statističnem modelu $0 < a < 1$.

3. a) Imamo podatke:

$$m = 17\,, \quad \bar{X} = 45\text{˙}76\,, \quad S_{\mathrm{p},X} = 15\text{˙}61\,, n = 12\,, \quad \bar{Y} = 31\text{˙}17\,, \quad S_{\mathrm{p},Y} = 20\text{˙}31\,.$$

Popravljeni vzorčni standardni odklon za vse študente znaša:

$$S_{\mathrm{p}} = \sqrt{\frac{(m-1)S_{\mathrm{p},X}^2 + (n-1)S_{\mathrm{p},Y}^2}{m+n-2}} \doteq 17\text{˙}68\,.$$

b) Po primerjavi testne statistike $T \doteq 2\text{˙}19$ s kritičnima vrednostma $t_{0.975}(27) \doteq 2\text{˙}05$ in $t_{0.995}(27) \doteq 2\text{˙}77$ dobimo, da so odstopanja statistično značilna, niso pa zelo značilna.

4. Nastavimo neznano funkcijo h, za katero bo veljalo $E\big[h(X)\big] = \lambda\, e^{\lambda}$. Iz Poissonove porazdelitve dobimo:

$$E\big[h(X)\big] = \sum_{k=0}^{\infty} h(k)\, \frac{\lambda^k\, e^{-\lambda}}{k!}\,,$$

Torej mora veljati:

$$\sum_{k=0}^{\infty} \frac{h(k)}{k!}\, \lambda^k = \lambda\, e^{2\lambda} = \sum_{k=1}^{\infty} \frac{2^{k-1}}{(k-1)!}\, \lambda^k$$

in to velja za $h(k) = k\, 2^{k-1}$. To pomeni, da je $X\, 2^{X-1}$ nepristranska cenilka za $\lambda\, e^{\lambda}$.

4P. S prvim momentom:

$$m_1 = E(X) = \frac{1}{2a} \int_{-a}^{a} x \, \mathrm{d}x = 0$$

si ne moremo nič pomagati, zato pa si lahko pomagamo z drugim:

$$m_2 = E(X^2) = \frac{1}{2a} \int_{-a}^{a} x^2 \, \mathrm{d}x = \frac{a^2}{3}\,.$$

Cenilka za a je torej:

$$\hat{a} = \sqrt{3m_2} = \sqrt{\frac{3}{n}(X_1^2 + X_2^2 + \cdots + X_n^2)}\,.$$