Rešitve izpita iz verjetnosti in statistike z dne 23. 8. 2013

Matematika – univerzitetni študij

1. a) Slučajna spremenljivka N lahko zavzame celoštevilske vrednosti od 2 do $2n-2k+2$. Za i izmed teh števil se dogodek $\{N = i\}$ zgodi, če je med prvimi $i-1$ nogavicami natanko ena karirasta in če je i-ta karirasta. To pomeni, da je med zadnjimi $2n-i$ nogavicami natanko $2k-2$ karirastih. Sledi:

$$P(N = i) = \frac{(i-1)\binom{2n-i}{2k-2}}{\binom{2n}{2k}}\,.$$

b) Za $n = 4$ in $k = 2$ dobimo:

$$N \sim \begin{pmatrix} 2 & 3 & 4 & 5 & 6 \\ \frac{15}{70} & \frac{20}{70} & \frac{18}{70} & \frac{12}{70} & \frac{5}{70} \end{pmatrix},$$

od koder izračunamo:

$$E(N) = \frac{18}{5}\,, \quad E(N^2) = \frac{72}{5}\,, \quad D(N) = \frac{36}{25}\,.$$

2. Ker gre za 100 neodvisnih seštevancev, je smiselno uporabiti približek na podlagi centralnega limitnega izreka. Velja:

$$E(X_iY_i) = E(X_i)\,E(Y_i) = 3\,, \quad E(S) = 300\,,$$
$$E(X_i^2) = D(X_i) + \big(E(X_i)\big)^2 = 9\,, \quad E(Y_i^2) = D(Y_i) + \big(E(Y_i)\big)^2 = 10\,,$$
$$E(X_i^2Y_i^2) = E(X_i^2)\,E(Y_i^2) = 90\,, \quad D(X_iY_i) = 81\,, \quad D(S) = 8100\,,$$

torej je približno:

$$P(S < 240) \approx \Phi\left(\frac{240-300}{90}\right) + \frac{1}{2} = \frac{1}{2} - \Phi\left(\frac{2}{3}\right) \approx 0{\cdot}252\,.$$

3. Označimo ustrezni aritmetični sredini z $\mu_{\check{\mathrm{Z}}}$ in μ_{M} ter ustrezna popravljena standardna odklona z $S_{\mathrm{p},\check{\mathrm{Z}}}$ in $S_{\mathrm{p},\mathrm{M}}$. Naj bo v vzorcu n moških in n žensk. Ničelno hipotezo zavrnemo v korist alternativne, če je:

$$\frac{\mu_{\mathrm{M}} - \mu_{\check{\mathrm{Z}}}}{\sqrt{\dfrac{S_{\mathrm{p},\mathrm{M}}^2 + S_{\mathrm{p},\check{\mathrm{Z}}}^2}{2}}}\sqrt{\frac{n}{2}} > t_{0.99}(2n-1)$$

ali ekvivalentno:

$$\frac{t_{0.99}(2n-1)}{\sqrt{n}} < \frac{\mu_{\mathrm{M}} - \mu_{\check{\mathrm{Z}}}}{\sqrt{S_{\mathrm{p},\mathrm{M}}^2 + S_{\mathrm{p},\check{\mathrm{Z}}}^2}} \doteq 0{\cdot}18034\,,$$

kar velja za $n \geq 169$ (če kvantil Studentove porazdelitve aproksimiramo s kvantilom standardne normalne porazdelitve, pa ustrezno neenakost dobimo že za $n \geq 162$).

4. Označimo z R razdaljo od izhodišča do najbližje točke (če točk sploh ni, pa naj bo $R = \infty$). Najlažje je računati komplemenarno kumulativno porazdelitveno funkcijo, torej $P(R > r)$. To je verjetnost dogodka, da v krogu okoli izhodišča s polmerom r ni nobene točke.
Kvadrat oddaljenosti posamezne točke od izhodišča ima porazdelitev hi kvadrat z dvema prostostnima stopnjama, ki ima gostoto:

$$g(x) = \begin{cases} \frac{1}{2}\, e^{-x/2} & ; x > 0 \\ 0 & ; \text{sicer} \end{cases} .$$

Verjetnost, da je posamezna točka od izhodišča oddaljena za več kot r, je enaka:

$$\int_{r^2}^{\infty} g(x)\,\mathrm{d}x = e^{-r^2/2} .$$

Od tod sledi:

$$P(R > r \mid N) = e^{-Nr^2/2}$$

in:

$$P(R > r) = \sum_{n=0}^{\infty} \frac{\lambda^n\, e^{-\lambda}}{n!}\, e^{-nr^2/2} = e^{\lambda(e^{-r^2/2}-1)} .$$

Kumulativna porazdelitvena funkcija je torej enaka:

$$F_R(r) = \begin{cases} 0 & ; r \le 0 \\ 1 - e^{\lambda(e^{-r^2/2}-1)} & ; r \ge 0 \end{cases} .$$

Opazimo, da njena limita, ko gre r proti neskončno, ni 1, temveč $1 - e^{-\lambda}$. Ta limita je verjetnost, da je slučajna spremenljivka R končna ali, ekvivalentno, da je v ravnini vsaj ena točka.

4P. Najprej določimo kumulativno porazdelitveno funkcijo $F_Y(y) = P(Y < y)$. Ker lahko Y zavzame vrednosti od 0 do 4, je za $y \le 0$ očitno $F_Y(y) = 0$, za $y \ge 4$ pa $F_Y(y) = 1$. Za $0 \le y \le 1$ velja:

$$F_Y(y) = P(X^2 < y) = P\left(\sqrt{y} < X < \sqrt{y}\right) = \frac{2}{3}\sqrt{y} ,$$

za $1 \le y \le 2$ pa velja:

$$F_Y(y) = P(X^2 < y) = P\left(\sqrt{y} < X < \sqrt{y}\right) = P\left(1 < X < \sqrt{y}\right) = \frac{\sqrt{y}+1}{3} .$$