Rešitve izpita iz verjetnosti in statistike z dne 6. 9. 2010

Matematika – UNI-BOL

1. Računajmo:

$$
\begin{aligned}
P(Y > 1) &= \int_1^\infty e^{-y}\,\mathrm{d}y = \frac{1}{e}\,,\\
P(X > Y, Y > 1) &= \int_1^\infty \int_y^\infty e^{-x}\,e^{-y}\,\mathrm{d}x\,\mathrm{d}y = \int_1^\infty e^{-2y}\,\mathrm{d}y = \frac{1}{2e^2}\,,\\
P(X > Y \mid Y > 1) &= \frac{P(X > Y, Y > 1)}{(Y > 1)} = \frac{1}{2e}\,.
\end{aligned}
$$

2. Označimo z A_i dogodek, da i-ti otrok ne dobi nobene žoge, in pišimo:

$$
S = X_1 + X_2 + \cdots + X_n\,,\ \text{kjer je}\ X_i = \mathbf{1}_{A_i} = \begin{cases} 1 & ;\ i\text{-ti otrok ne dobi nobene žoge} \\ 0 & ;\ \text{sicer} \end{cases}
$$

Iz $E(X_i) = P(A_i) = \left(1 - \frac{1}{n-1}\right)^{n-1}$ dobimo $E(S) = n\left(1 - \frac{1}{n-1}\right)^{n-1}$. Za izračun disperzije pa pišimo:

$$
\begin{aligned}
D(S) &= E(S^2) - \big(E(S)\big)^2,\\
E(S^2) &= \sum_{i=1}^n E(X_i^2) + \sum_{\substack{1\le i,j\le n\\ i\ne j}} E(X_iX_j) =\\
&= \sum_{i=1}^n P(A_i) + \sum_{\substack{1\le i,j\le n\\ i\ne j}} P(A_i \cap A_j) =\\
&= n\left(1 - \frac{1}{n-1}\right)^{n-1} + n(n-1)\left(1 - \frac{1}{n-1}\right)^2\left(1 - \frac{2}{n-1}\right)^{n-2}.
\end{aligned}
$$

Torej je:

$$
\begin{aligned}
D(S) = n\left(1 - \frac{1}{n-1}\right)^{n-1} &+ n(n-1)\left(1 - \frac{1}{n-1}\right)^2\left(1 - \frac{2}{n-1}\right)^{n-2} -\\
&- n^2\left(1 - \frac{1}{n-1}\right)^{2n-2}.
\end{aligned}
$$

Opomba. Zanimivo je raziskati asimptotično obnašanje količin, ki jih računamo, ko gre n proti neskončno. Če z $a_n \sim b_n$ označimo, da je $\lim_{n\to\infty} a_n/b_n = 1$, se da s sredstvi iz elementarne analize dokazati, da je:

$$
E(S) \sim n/e \quad \text{in} \quad D(S) \sim n(e^{-1} - 2e^{-2})\,.
$$

Z bolj sofisticiranimi sredstvi iz verjetnostnega računa pa se da raziskati tudi asimptotično obnašanje cele porazdelitve slučajne spremenljivke S, in sicer se izkaže, da slučajne spremenljivke:

$$\frac{S - n/e}{\sqrt{n}}$$

šibko konvergirajo proti normalni porazdelitvi z matematičnim upanjem nič in disperzijo $e^{-1} - 2e^{-2}$.

3. Ker lahko zapišemo:

$$p(x) = \frac{a}{x}\, e^{a \ln x}$$

in za ustrezni vzorec:

$$p_{X_1,\ldots X_n}(x_1, \ldots x_n) = \frac{a^n}{x_1 x_2 \cdots x_n}\, e^{a(\ln x_1 + \cdots + \ln x_n)}\,,$$

gre za eksponentno družino, kjer ima naravni parametrični prostor neprazno notranjost. Torej je nepristranska cenilka za $1/a$ z najmanjšo možno disperzijo, če obstaja, funkcija minimalne zadostne statistike $\ln X_1 + \cdots + \ln X_n$. Iz:

$$E(\ln X_1) = a \int_0^1 x^{a-1} \ln x \,\mathrm{d}x = x^a \ln x \Big|_0^1 - \int_0^1 x^{a-1}\,\mathrm{d}x = -\frac{1}{a}$$

dobimo $E(\ln X_1 + \cdots + \ln X_n) = -n/a$. Iskana cenilka je torej:

$$-\frac{\ln X_1 + \ln X_2 + \cdots + \ln X_n}{n}\,.$$

To cenilko dobimo tudi po metodi največjega verjetja, a je potrebno preveriti, da je nepristranska.

4. Vzorčno povprečje: $\bar{X} = 24˙95$.
Popravljeni vzorčni standardni odklon: $S_{\mathrm{p}} \doteq 3˙776$.
Kvantila porazdelitve hi kvadrat pri $df = 9$: $\chi^2_{0.025} \doteq 2˙700$, $\chi^2_{0.975} \doteq 19˙02$.
Interval zaupanja (ustrezno zaokrožen): $2˙59 < \sigma < 6˙90$.