VEROVATNOĆA I STATISTIKA

Jun 2005.

Rešenja zadataka

1. Da bi se izvukle tačno tri kuglice iste boje, razlikujemo dva analogna slučaja, izvučene tri crvene i plava, odnosno tri plave i crvena. Izvlačenje tri crvene i jedne plave kuglice moguće je tako što izvučemo iz prve kutije dve crvene, a iz druge crvenu i plavu (plavu i crvenu) ili iz prve kutije crvenu i plavu (plavu i crvenu), a iz druge dve crvene. Verovatnoća tog događaja je onda

$$p_1^2 \cdot p_2(1-p_2) + p_1^2 \cdot (1-p_2)p_2 + p_1(1-p_1) \cdot p_2^2 + (1-p_1)p_1 \cdot p_2^2,$$

a analognog, tj. izvlačenje tri plave i jedne crvene je

$$(1-p_1)^2 \cdot (1-p_2)p_2 + (1-p_1)^2 \cdot p_2(1-p_2) + (1-p_1)p_1 \cdot (1-p_2)^2 + p_1(1-p_1) \cdot (1-p_2)^2.$$

Verovatnoća izvlačenja tačno tri kuglice iste boje onda je:

$$p_1^2 \cdot 2p_2(1-p_2) + 2p_1(1-p_1) \cdot p_2^2 + (1-p_1)^2 \cdot 2p_2(1-p_2) + 2p_1(1-p_1) \cdot (1-p_2)^2 = \frac{1}{2},$$

odnosno

$$(p_1 + p_2 - 2p_1p_2)(1 - p_1 - p_2 + 2p_1p_2) = \frac{1}{4}. \tag{1}$$

Zamenom $p_1 = \frac{1}{3}$ u (1), dobija se $(1+p_2)(2-p_2) = \frac{9}{4}$, odakle sledi $p_2 = \frac{1}{2}$, a zamenom $p_1 = \frac{1}{2}$ u (1), dobija se da je $\frac{1}{4} = \frac{1}{4}$, odakle sledi da p_2 može biti proizvoljno iz segmenta $[0,1]$.

2. Iz uslova $\int\int\limits_D f(x,y)dxdy = 1$, dobija se da je $C = \frac{1}{2}$.

Marginalna gustina za Y je

$$f_Y(y) = \int\limits_{y-2}^{2-y} \frac{1}{2}(x^2+y^2)dx = \frac{1}{6}(-8y^3 + 24y^2 - 24y + 16), \ \ 1 \le y \le 2,$$

pa je uslovna gustina za $X|Y$

$$f_{X|Y=y}(x) = \frac{f(x,y)}{f(y)} = \frac{3(x^2+y^2)}{-8y^3+24y^2-24y+16}, \ \ y-2 \le x \le 2-y, \ \ 1 \le y \le 2.$$

Marginalna gustina za X je

$$f_X(x) = \int\limits_1^{2-|x|} \frac{1}{2}(x^2+y^2)dy = \begin{cases} \int\limits_1^{x+2} \frac{1}{2}(x^2+y^2)dy, & -1 \le x < 0 \\ \int\limits_1^{2-x} \frac{1}{2}(x^2+y^2)dy, & 0 \le x \le 1 \end{cases} = \begin{cases} \frac{4x^3+9x^2+12x+7}{6}, & -1 \le x < 0 \\ \frac{-4x^3+9x^2-12x+7}{6}, & 0 \le x \le 1 \end{cases},$$

pa je uslovna gustina za $Y|X$

$$f_{Y|X=x}(y) = \frac{f(x,y)}{f(x)} = \begin{cases} \frac{3(x^2+y^2)}{4x^3+9x^2+12x+7}, & -1 \le x < 0 \\ \frac{3(x^2+y^2)}{-4x^3+9x^2-12x+7}, & 0 \le x \le 1 \end{cases}, \ \ 1 \le y \le 2-|x|.$$

3. Karakteristična funkcija slučajne veličine X_n je

$$\varphi_{X_n}(t) = E(e^{itX_n}) = \int_{-\infty}^{+\infty} e^{itx} f_{X_n}(x)dx = \int_{\frac{n+1}{n}}^{+\infty} e^{itx} e^{-(x-\frac{n+1}{n})}dx = \frac{e^{it\frac{n+1}{n}}}{1-it},$$

jer $|e^{(it-1)x}| = |e^{itx}||e^{-x}| = e^{-x} \underset{x\to+\infty}{\longrightarrow} 0$.

Kako za $(\forall t \in \mathbb{R})$ $\lim\limits_{n\to\infty} \varphi_{X_n}(t) = \frac{e^{it}}{1-it}$ i $\frac{e^{it}}{1-it}$ je neprekidna u nuli, a karakteristična funkcija slučajne veličine X je

$$\varphi_X(t) = E(e^{itX}) = \int_{-\infty}^{+\infty} e^{itx} f_X(x)dx = \int_{1}^{+\infty} e^{itx} e^{1-x}dx = \frac{e^{it}}{1-it},$$

zaključujemo da $X_n \xrightarrow[n\to\infty]{R} X$.

4. Kako je

$$EA_n = \frac{1}{6n} E\sum_{i=1}^{n} I\{X_i \le 3\} = \frac{1}{6n}\sum_{i=1}^{n} E\,I\{X_i \le 3\} = \frac{1}{6n}\sum_{i=1}^{n} P\{X_i \le 3\} = \frac{1}{6n}\sum_{i=1}^{n} 6\theta = \theta,$$

$$EB_n = \frac{6}{22} - \frac{1}{22}E\overline{X_n} = \frac{6}{22} - \frac{1}{22}\frac{1}{n}\sum_{i=1}^{n} EX_i = \frac{6}{22} - \frac{1}{22}\frac{1}{n}\sum_{i=1}^{n}(6-22\theta) = \theta.$$

Kako su ocene nepristrasne, onda je

$$\begin{aligned}
E(A_n-\theta)^2 &= DA_n = \frac{1}{36n^2} D\sum_{i=1}^{n} I\{X_i \le 3\} = \frac{1}{36n^2}\sum_{i=1}^{n} DI\{X_i \le 3\} \\
&= \frac{1}{36n^2}\sum_{i=1}^{n} P\{X_i \le 3\}(1-P\{X_i \le 3\}) = \frac{1}{36n^2}\sum_{i=1}^{n} 6\theta(1-6\theta) \\
&= \frac{\theta}{6n} - \frac{\theta^2}{n}, \\
E(B_n-\theta)^2 &= DB_n = \frac{1}{22^2} D\overline{X_n} = \frac{1}{484}\frac{1}{n^2} D\sum_{i=1}^{n} X_i = \frac{1}{484}\frac{1}{n^2}\sum_{i=1}^{n} DX_i \\
&= \frac{1}{484}\frac{1}{n^2}\sum_{i=1}^{n}(84\theta - 484\theta^2) = \frac{84\theta}{484n} - \frac{\theta^2}{n}.
\end{aligned}$$

Upoređivanjem ova dva izraza dobija se $E(A_n-\theta)^2 < E(B_n-\theta)^2$ za svako n, pa sledi da je ocena A_n bolja u srednje kvadratnom smislu.

5. Kako su elementi uzorka nezavisne slučajne veličine sa normalnom $\mathcal{N}(m, \sigma^2)$ raspodelom, to $X_{2i} - X_{2i-1}$ ima normalnu $\mathcal{N}(0, 2\sigma^2)$ raspodelu, odnosno $\frac{X_{2i}-X_{2i-1}}{\sigma\sqrt{2}}$ ima normalnu $\mathcal{N}(0,1)$ raspodelu, pa uzimajući u obzir nezavisnost slučajnih veličina $X_{2i} - X_{2i-1}$, $i = 1, 2, ..., 20$, slučajna veličina $Y = \sum_{i=1}^{20} \left(\frac{X_{2i}-X_{2i-1}}{\sqrt{2}\sigma}\right)^2$ ima χ^2_{20} raspodelu.

   Konstantu C dobijamo iz uslova da je nivo značajnosti testa 0.1, odnosno

$$P_{H_0}\{\sum_{i=1}^{20}(X_{2i} - X_{2i-1})^2 \leq C\} = P_{H_0}\{\sum_{i=1}^{20}\left(\frac{X_{2i} - X_{2i-1}}{\sqrt{2}}\right)^2 \leq \frac{C}{2}\} = 0.1$$

$$P_{H_0}\{\sum_{i=1}^{20}\left(\frac{X_{2i} - X_{2i-1}}{\sqrt{2}}\right)^2 > \frac{C}{2}\} = 0.9$$

   Odavde sledi da je $C = 2 \cdot \chi^2_{20;\,0.9} = 24.886$.

   Verovatnoća greške druge vrste

$$\beta = P_{H_1}\{\sum_{i=1}^{20}(X_{2i} - X_{2i-1})^2 > 24.886\} = P_{H_1}\{Y > 24.886\} \approx 0.2.$$

6. Kako je disperzija vremena rastvaranja tablete jednaka 40, to znači da treba da testiramo hipotezu da vreme rastvaranja ima χ^2_{20} raspodelu.

   S obzirom da je domen za χ^2 raspodelu $[0, +\infty)$, dodavanjem još jedne klase dobijamo sledeću tabelu:

	[0,12.44]	(12.44,14.58]	(14.58,19.34]	(19.34,28.41]	$(28.41, +\infty)$
M_k	12	8	35	45	0
p_k	0.1	0.1	0.3	0.4	0.1
np_k	10	10	30	40	10
$\frac{(M_k-np_k)^2}{np_k}$	$\frac{4}{10}$	$\frac{4}{10}$	$\frac{25}{30}$	$\frac{25}{40}$	10

   Verovatnoće p_k računaju se iz tablice za χ^2_{20} raspodelu, na primer

$$p_2 = P_{H_0}\{X \in (12.44, 14.58]\} = P_{H_0}\{X > 12.44\} - P_{H_0}\{X > 14.58\} = 0.9 - 0.8 = 0.1.$$

   Odavde je $\chi^2_0 = \sum_{k=1}^{5} \frac{(M_k-np_k)^2}{np_k} = 12.258$. Kako je kritična oblast $W = \{\chi^2_0 \geq C\}$, gde je $C = \chi^2_{5-1;\,0.01} = 13.277$, zaključujemo da uzorak ne pripada kritičnoj oblasti, pa nema razloga da odbacimo hipotezu da vreme rastvaranja ima χ^2 raspodelu.