VEROVATNOĆA I STATISTIKA

Septembar 2005.

Rešenja zadataka

1. Neka je A događaj da je Aca pojeo tri palačinke.

   Poređajmo palačinke u niz. Broj načina na koje k dece može da podeli n palačinki pri uslovima zadatka odgovara broju mogućnosti razdvajanja n palačinki u nizu na k delova, odnosno umetanju $k-1$ "granice" u neke od $n-1$ praznina između palačinki, i iznosi $\binom{n-1}{k-1}$.

   Otuda je ukupan broj mogućih ishoda $\binom{19}{3}$. Povoljni ishodi za događaj A su oni kada je Aca uzeo svoje 3 palačinke, a preostalih 17 treba podeliti na 3 dela, što je moguće na $\binom{16}{2}$ načina.

   Tada je $P(A) = \frac{\binom{16}{2}}{\binom{19}{3}} = \frac{40}{323}$

2. Iz grafika zavisnosti Y od X i simetričnosti normalne $\mathcal{N}(0,1)$ raspodele dobija se da je za

   1) $y < -C$

   $$F_Y(y) = P\{Y \leq y\} = P\{X \geq -y\} = P\{X \leq y\} = F_X(y)$$

   2) $-C \leq y < C$

   $$\begin{aligned} F_Y(y) &= P\{Y \leq y\} = P\{-C < X \leq y\} + P\{X \geq C\} = P\{X \leq y\} - P\{X \leq -C\} + P\{X \geq C\} \\ &= P\{X \leq y\} - P\{X \geq C\} + P\{X \geq C\} = P\{X \leq y\} = F_X(y) \end{aligned}$$

   3) $y \geq C$

   $$F_Y(y) = P\{Y \leq y\} = P\{X \geq -y\} = P\{X \leq y\} = F_X(y)$$

   Kako je $F_Y(y) = F_X(y)$ za svako $y \in \mathbb{R}$, sledi da slučajna veličina Y ima takođe normalnu $\mathcal{N}(0,1)$ raspodelu.

   S obzirom da je $Z = X + Y = \begin{cases} 2X, & |X| < C \\ 0, & |X| \geq C \end{cases}$, zaključujemo da je $P\{Z = 0\} > 0$, pa slučajna veličina Z nije apsolutno neprekidnog tipa, te ne može imati normalnu raspodelu.

3. Neka je X_i broj dobijen na i-toj kockici, a $Y = \sum\limits_{i=1}^{100} X_i$ zbir dobijenih brojeva. Može se lako izračunati da je $EX_i = \frac{7}{2}$, $DX_i = \frac{35}{12}$, $EY = 350$ i $DY = \frac{875}{3}$. Tada je tražena verovatnoća:

   a) $P\{300 < \sum\limits_{i=1}^{100} X_i < 400\} = P\{300 < Y < 400\} = P\{300 - 350 < Y - EY < 400 - 350\} =$
   $= P\{-50 < Y - EY < 50\} = P\{|Y - EY| < 50\} = 1 - P\{|Y - EY| \geq 50\} \geq 1 - \frac{E(Y-EY)^2}{50^2} = 1 - \frac{DY}{50^2} = \frac{53}{60}$.

   b) $P\{300 < \sum\limits_{i=1}^{100} X_i < 400\} = P\{\frac{300-350}{\sqrt{\frac{875}{3}}} < \frac{\sum_{i=1}^{100} X_i - E\sum_{i=1}^{100} X_i}{\sqrt{D\sum_{i=1}^{100} X_i}} < \frac{400-350}{\sqrt{\frac{875}{3}}}\} \approx P\{-2.93 < X^* < 2.93\} = 0.997.$

4. Funkcija verodostojnosti je

$$L(x_1,...,x_n;\theta)=\prod_{i=1}^{n} f(x_i;\theta)=\left(\prod_{i=1}^{n} x_i\right)\cdot\theta^{2n}\cdot e^{-\theta\sum\limits_{i=1}^{n} x_i},$$

pa je

$$\ln L(x_1,...,x_n;\theta)=\sum_{i=1}^{n}\ln x_i+2n\ln\theta-\theta\sum_{i=1}^{n} x_i.$$

Rešavanjem jednačine verodostojnosti, $\frac{\partial}{\partial\theta}\ln L(x_1,...,x_n;\theta)=0$, dobija se ocena maksimalne verodostojnosti parametra θ

$$\hat\theta=\frac{2n}{\sum\limits_{i=1}^{n} X_i}.$$

Kako X_i ima gama raspodelu, to će zbog aditivnosti gama raspodele $Y=\sum\limits_{i=1}^{n} X_i$ takođe imati gama raspodelu sa gustinom $f(y)=\frac{y^{2n-1}}{\Gamma(2n)}\theta^{2n}e^{-\theta y}$.

Tada je

$$E\hat\theta=2nE\frac{1}{Y}=2n\int\limits_0^{+\infty}\frac{1}{y}\frac{y^{2n-1}}{\Gamma(2n)}\theta^{2n}e^{-\theta y}dy=\frac{2n}{\Gamma(2n)}\theta\cdot\Gamma(2n-1)=\frac{2n}{2n-1}\theta,$$

pa $\hat\theta$ nije nepristrasna ocena parametra θ.

Koristeći Čebišovljevu nejednakost dobijamo da za svako $\varepsilon>0$

$$\begin{aligned}P\{|\hat\theta-\theta|\geq\varepsilon\} &= P\{|\frac{2n}{Y}-\theta|\geq\varepsilon\}\leq\frac{E(\frac{2n}{Y}-\theta)^2}{\varepsilon^2}=\frac{4n^2E\frac{1}{Y^2}-4n\theta E\frac{1}{Y}+\theta^2}{\varepsilon^2}\\ &= \frac{4n^2\frac{\theta^2}{(2n-1)(2n-2)}-4n\theta\frac{\theta}{2n-1}+\theta^2}{\varepsilon^2}=\frac{[\frac{4n^2}{(2n-1)(2n-2)}-\frac{4n}{2n-1}+1]\theta^2}{\varepsilon^2}\underset{n\to+\infty}{\longrightarrow}0,\end{aligned}$$

pa zaključujemo da je ocena $\hat\theta$ postojana.

5. Kad X_i ima normalnu $\mathcal{N}(m,\sigma^2)$ raspodelu, tada $\frac{\overline{X_n}-m}{\sigma}\sqrt{n}$ ima normalnu $\mathcal{N}(0,1)$, a $\frac{n\overline{S_n}^2}{\sigma^2}$ ima χ^2_{n-1} raspodelu. Zbog toga je:

$$\begin{aligned}P\{0<\overline{X_{25}}<6\quad,\quad 216.744<\overline{S_{25}}^2<280.044\} &= P\{0<\overline{X_{25}}<6\}\cdot P\{216.744<\overline{S_{25}}^2<280.044\}=\\ &= P\{\frac{0-3}{\sqrt{300}}\sqrt{25}<\frac{\overline{X_{25}}-3}{\sqrt{300}}\sqrt{25}<\frac{6-3}{\sqrt{300}}\sqrt{25}\}\cdot P\{\frac{25\cdot 216.744}{\sqrt{300}}<\frac{25\cdot\overline{S_{25}}^2}{\sqrt{300}}<\frac{25\cdot 280.044}{\sqrt{300}}\}\\ &= P\{-0.87<X^*<0.87\}\cdot P\{18.062<\chi^2_{24}<23.337\}=0.6157\cdot(0.8-0.5)=0.18471.\end{aligned}$$

6. Neka je l registrovani broj pojavljivanja pisma u 100 bacanja. Formiramo pomoćnu tabelu

	pismo	glava
M_k	l	$100-l$
p_k	$\frac{1}{2}$	$\frac{1}{2}$
np_k	50	50

pa je test statistika $\chi^2_0=\sum\limits_{k=1}^{2}\frac{(M_k-np_k)^2}{np_k}=\frac{(l-50)^2}{50}+\frac{(100-l-50)^2}{50}=\frac{(l-50)^2}{25}$.

Kritična oblast je $W=\{\chi^2_0\geq\chi^2_{2-1;0.05}\}=\{\chi^2_0\geq\chi^2_{1;0.05}\}=\{\chi^2_0\geq 3.841\}$, pa hipotezu ne odbacujemo ako je $\frac{(l-50)^2}{25}<3.841$, odnosno $40.20077<l<59.79923$. Dakle, najmanji broj pojavljivanja pisma pri kome nema razloga za odbacivanje hipoteze je 41, a najveći 59.