Prvi kolokvij iz Statistike - 2008:04:11:08:30

U svim zadacima je $\Xi = \big(\xi_1, \cdots, \xi_n\big)$ prost slučajan uzorak opsega n izabran iz populacije s obilježjem, slučajnom varijablom ξ koja ima razdiobu ovisnu o parametru $\theta \in \Theta$.

1. Neka je $\xi \sim \mathcal{U}[0, \sqrt{\theta}\,]$, $\theta > 0$.

(a) Odrediti ML procjenitelja $\widehat{\theta}$ za θ na osnovu Ξ;
(b) normirati $\widehat{\theta}$ do nepristranog procjenitelja;
(c) pokazati da je takav procjenitelj konzistentan.

2. Neka je gustoća obilježja ξ zadana s

$$f_\theta(x) = 2\frac{x}{\theta}\,\chi_{[0,\theta]}(x) + 2\frac{1-\theta}{1-x}\,\chi_{(\theta,1]}(x) \qquad (0 < \theta < 1)\,.$$

(a) Naći procjenitelja $\widehat{\theta}$ metodom momenata za θ na osnovu Ξ;
(b) izračunati $\mathsf{D}\widehat{\theta}$.

3. Obilježje populacije $\xi \sim f_\alpha(x) = 2\alpha x(1-x^2)^{\alpha-1}\,\chi_{(0,1)}(x)$. Odrediti jednu dovoljnu statistiku za α na osnovu Ξ.

4. Kolika je Fisherova informacija uzorka Ξ, ako uzorak dolazi iz populacije s obilježjem

$$\xi \sim f_\theta(x) = 2\theta^2 x^3 e^{-\theta x^2}\,\chi_{[0,\infty)}(x) \qquad (\theta > 0)\,?$$

5. Što je *najefikasniji procjenitelj* $\widehat{\theta}_0$ nepoznatog parametra θ? Sav usput korišteni matematički aparat mora biti izveden i dokazan.

6. Iskazati Fisher – Neymanov teorem. "Dovoljno" je dokazati dovoljnost uvjeta!

7. (Nagradni zadatak). Daltonizam se pojavljuje s vjerojatnosti p kod muških i s vjerojatnosti p^2 kod ženskih osoba. Naći ML procjenu parametra p, ako od M muških je m, a kod F ženskih f daltonista u prostom slučajnom uzorku.

Svako točno riješeno pitanje u zadacima 1.–4. donosi 10 bodova, teorijska pitanja vrijede 15, a nagradni zadatak 20 (ako ga točno riješite, možete zabušavati ili sada, ili drugi put).
SRETAN I USPJEŠAN RAD!

Tibor Poganj

RJEŠENJA

1. (a) Nije teško uvidjeti da je funkcija vjerodostojnosti oblika

$$L_\theta(x) = \prod_{j=1}^{n} \frac{1}{\sqrt{\theta}}\, \chi_{(0,\sqrt{\theta})}(x_j) = \frac{1}{\theta^{n/2}}\, \chi\big\{0 < \xi_{(1)} \le \cdots \le \xi_{(n)} < \sqrt{\theta}\big\}\,,$$

gdje je $\chi\{A\}$ indikator slučajnog događaja A. Odatle je $\widehat{\theta} = \xi_{(n)}^2$.

(b) Gustoća n–te statistike poretka $\xi_{(n)}$ je

$$f_{(n)}(x) = \frac{nx^{n-1}}{\theta^{n/2}}\, \chi_{(0,\sqrt{\theta})}(x)\,.$$

Dakle, očekivanje procjenitelja $\widehat{\theta}$ je

$$\mathsf{E}\widehat{\theta} = \mathsf{E}\xi_{(n)}^2 = \frac{n}{\theta^{n/2}} \int_0^{\sqrt{\theta}} x^{n+1}\, \mathrm{d}x = \frac{n\theta}{n+2}\,;$$

$\widehat{\theta}$ očito nije nepristran, ali jeste

$$\widetilde{\theta} = \Big(1 + \frac{2}{n}\Big)\widehat{\theta} = \Big(1 + \frac{2}{n}\Big)\, \xi_{(n)}^2\,.$$

(c) Dovoljno je pokazati da $\mathsf{D}\widetilde{\theta} \to 0$, kada $n \to \infty$, budući da je $\widetilde{\theta}$ nepristrani procjenitelj (to slijedi iz nejednakosti Чебышева). A to je ispunjeno zbog

$$\lim_{n\to\infty} \mathsf{D}\widetilde{\theta} = \lim_{n\to\infty} \Big(1 + \frac{2}{n}\Big)^2 \mathsf{D}\xi_{(n)}^2 = \lim_{n\to\infty} \frac{4\theta^2}{n(n+4)} = 0\,.$$

2. (a) Traženi procjenitelj je $\widehat{\theta} = 3\,\overline{X}_n - 1$, jer je očekivanje obilježja

$$\mathsf{E}\xi = \frac{2}{\theta}\int_0^\theta x^2 \mathrm{d}x + \frac{2}{1-\theta}\int_0^\infty x(1-x)\mathrm{d}x = \frac{1+\theta}{3}\,.$$

(b) Kako je $\mathsf{E}\xi^2 = (\theta^2 + \theta + 1)/6$, lako slijedi

$$\mathsf{D}\widehat{\theta} = 9\,\mathsf{D}\overline{X}_n = \frac{9}{n^2}\,\mathsf{D}\sum_{\ell=1}^{n} \xi_\ell = \frac{9}{n^2}\sum_{\ell=1}^{n} \mathsf{D}\xi_\ell = \frac{9}{n}\,\mathsf{D}\xi = \frac{1}{2n}\,(\theta^2 - \theta + 1)\,.$$

3. Funkciju vjerodostojnosti možemo napisati u obliku

$$L_\alpha(x) = 2^n\alpha^n \prod_{j=1} x_j(1 - x_j^2)^{\alpha-1} = \underbrace{\alpha^n\Big(\prod_{j=1}^{n}(1 - x_j^2)\Big)^{\alpha-1}}_{K_1(U;\alpha)} \cdot \underbrace{\prod_{\ell=1}^{n} 2x_\ell}_{K_2(x)}\,.$$

Na osnovu faktorizacijskog teorema statistika $U = \prod_{j=1}^{n}(1 - x_j^2)$ je dovoljna za α.

4. Po definiciji je

$$\mathcal{I}_1(\theta) = \mathsf{E}\Big(\frac{\partial}{\partial\theta}\ln f_\theta(\xi)\Big)^2 = \mathsf{E}\Big(\frac{2}{\theta} - \xi\Big)^2 = 2\theta^2\int_0^\infty \Big(\frac{2}{\theta} - x\Big)^2 x^3 e^{-\theta x^2}\,\mathrm{d}x\,.$$

Uvodimo supstituciju $\theta x^2 = t$, što integral svodi na linearnu kombinaciju tri Gamma funkcije:

$$\mathcal{I}(\theta) = n\mathcal{I}_1(\theta) = n\Big\{\frac{4}{\theta^2}\Gamma(2) - \frac{4}{\theta^2}\Gamma(3) + \frac{1}{\theta^2}\Gamma(4)\Big\} = \frac{2n}{\theta^2}\,.$$

5. To je zapravo nejednakost Rao–Cramér.

6. Dokaz je u bilježnici.

7. Funkcija vjerodostojnosti je

$$L_p(x) = \binom{M}{m}p^m(1-p)^{M-m}\binom{F}{f}p^{2f}(1-p^2)^{F-f}\,.$$

Sada se lako nalazi jednadžba vjerodostojnosti:

$$\frac{\partial}{\partial p}\ln L_p(x) = 0 \quad\Leftrightarrow\quad (M+2F)p^2 + (M-m)p - (m+2f) = 0 \tag{$\star$}$$

odnosno

$$\widehat{p} = \frac{\sqrt{(M-m)^2 + 4(M+2F)(m+2f)} - (M-m)}{2(M+2F)}\,.$$

Drugo rješenje ($\star$) je negativno, dakle ne dolazi u obzir kao procjenitelj.