МАТЕМАТИЧКА СТАТИСТИКА - Решење колоквијума

1. Нека је $\mathbf{X}$ узорак из фамилије расподела $\begin{pmatrix} -1 & 2 & 3 \\ 1-3\theta & \theta & 2\theta \end{pmatrix}$, $0 < \theta < \frac{1}{3}$.

а) Написати узорачку расподелу у облику производа користећи индикаторе одговарајућих догађаја и затим показати да дата фамилија расподела припада експоненцијалним фамилијама расподела.

б) Наћи минималну довољну статистику $T(\mathbf{X})$ параметра θ. Да ли је она комплетна?

в) Наћи непристрасну оцену параметра θ која је функција $T(\mathbf{X})$. Да ли њена дисперзија достиже доњу границу Рао-Крамера?

Решење.

а) Пошто је

$$p(x;\theta) = 0, x \notin \{-1,2,3\},$$

и

$$\begin{aligned} p(x;\theta) &= (1-3\theta)^{I\{x=-1\}}\theta^{I\{x=2\}}(2\theta)^{I\{x=3\}} \\ &= e^{I\{x=-1\}\ln(1-3\theta)+I\{x=2\}\ln(\theta)+I\{x=3\}(2\theta)} \\ &= e^{I\{x=-1\}\ln(1-3\theta)+(1-I\{x=-1\})\ln\theta+I\{x=3\}\ln 2} \\ &= e^{I\{x=-1\}\ln(\frac{1-3\theta}{\theta})+\ln\theta+I\{x=3\}\ln 2}, x \in \{-1,2,3\}, \end{aligned}$$

имамо да је $A(\theta) = \ln(\frac{1-3\theta}{\theta}), K(x) = I\{x=-1\}, S(x) = I\{x=3\}\ln 2, q(\theta) = \ln\theta$. Пошто скуп вредности за X, $S = \{-1,2,3\}$ не зависи од θ, на основу овог, случајна величина X је из експоненцијалне фамилије расподела.

б) Довољна статистика за θ, за узорак обима n, је $\sum_{i=1}^n K(X_i) = \sum_{i=1}^n I\{X_i = -1\}$. Како је димензија ове статистике 1, она мора бити и минимална. Да бисмо доказали комплетност ове статистике довољно је показати регуларност експоненцијане фамилије. Можемо користити једноставније услове регуларности. Носач расподеле не зависи од параметра и параметарски простор $\Theta = (0,1/3)$ садржи отворени интервал (и сам јесте отворен интервал). Дакле, $T(\mathbf{X}) = \sum_{i=1}^n I\{X_i = -1\}$ је комплетна довољна статистика за θ.

в) $E(T(\mathbf{X})) = n(1-3\theta)$, па је непристрасна оцена за θ статистика $U(\mathbf{X}) = \frac{1}{3}(1 - \frac{T(\mathbf{X})}{n})$. Како је ова непристрасна оцена линеарна функција довољне статистике која фигурише у изразу експоненцијалне фамилије расподела, она достиже доњу границу Рао-Крамера.

Ефикасност ове оцене могла је да се утврди и рачунањем и упоређивањем дисперзије ове оцене и доње границе дисперзија Рао-Крамера. У том случају добија се да је $DU(\mathbf{X}) = G = \frac{\theta(1-3\theta)}{3n}$. Услови регуларности важе, и нису се морали проверавати јер су они слабији од услова регуларности за експоненцијалну фамилију.

2. Нека је $\mathbf{X}$ узорак из фамилије Паретових расподела с густином $f(x;\theta) = \theta(1+x)^{-(1+\theta)}$, $x \geq 0, \theta > 2$.

а) Израчунати математичко очекивање $m(\theta)$ и дисперзију $\sigma^2(\theta)$ елемента узорка X_i.

б) Нека је $\bar{X}_n = \frac{1}{n}\sum_{k=1}^n X_k$. Испитати конвергенцију у расподели низа $\sqrt{n}(\bar{X}_n - m(\theta))$, коришћењем централне граничне теореме или утицајне криве.

в) Наћи $\hat{\theta}_n$, оцену параметра θ методом момената.

г) Наћи асимптотску расподелу $\sqrt{n}(\hat{\theta}_n - \theta)$ кад $n \to \infty$ користећи резултат под б) и делта метод.

Решење.

а) $m(\theta) = \frac{1}{\theta-1}$, $\sigma^2(\theta) = \frac{\theta}{(\theta-1)^2(\theta-2)}$ за $\theta > 2$.

б) Први начин: Како је $\sigma^2(\theta)$ коначно за свако $\theta > 2$, услови централне граничне теореме су испуњени, па важи

$$\sqrt{n}(\bar{X}_n - m(\theta)) \xrightarrow{D} Z, \text{ где је } Z : \mathcal{N}(0, \sigma^2(\theta)). \tag{1}$$

Други начин: Пошто је $m(\theta)$ први моменат расподеле то је утицајна крива $\phi(x;\theta) = x - m(\theta)$, па је $\int_{-\infty}^{\infty} \phi^2(x;\theta) f(x;\theta) dx = DX = \sigma^2(\theta) < \infty$, те важи (1).

в) Решавањем једначине $m(\theta) = \bar{X}_n$ добијамо $\hat{\theta}_n = \frac{1}{\bar{X}_n} + 1$.

г) Пошто је

$$\begin{aligned}\sqrt{n}(\hat{\theta}_n - \theta) &= \sqrt{n}\Big(\frac{1}{\bar{X}_n} + 1 - \frac{1}{m(\theta)} - 1\Big) \\ &= \sqrt{n}\Big(\frac{1}{\bar{X}_n} - \frac{1}{m(\theta)}\Big),\end{aligned}$$

а функција $g(m(\theta)) = 1/m(\theta)$ диференцијабилна за свако $m(\theta) \in (0,1)$ и $g'(m(\theta)) = -\frac{1}{(m(\theta))^2} \neq 0$, тражена гранична расподела добија се директно применом делта метода и резултата под б):

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{D} (g'(m(\theta))^2 Z,$$

па је гранична расподела нормална $\mathcal{N}\left(0, \frac{\sigma^2(\theta)}{(m(\theta))^4}\right)$, односно $\mathcal{N}\left(0, \frac{\theta(\theta-1)^2}{\theta-2}\right)$.