HW2

Smetanin Aleksandr

28 февраля 2020 г.

1. $X_1, \ldots, X_n$ - выборка из распределения Берннули с параметром p.
Тогда $X_1, X_1X_2, X_1(1-X_2)$ - несмещенные оценки для $p, p^2, p(1-p)$. Проверим это:
$EX_1 = p$,
$E(X_1 \cdot X_2) = EX_1 \cdot EX_2 = p \cdot p = p^2$,
$E(X_1(1-X_2)) = EX_1 \cdot E(1-X_2) = EX_1 \cdot (1 - EX_2) = p \cdot (1-p)$
Однако, такие оценки не являются состоятельными, так как нет стремеления $\theta^* \xrightarrow{\text{p}} \theta$ при $n \to \infty$, потому что оценка никак не зависит от n.

2. $F_n(y)$ - эмпирическая функкция распределения, которая строится по выборке из равномерного распределения на $[0, a], a > 1$. Для какого параметра $\theta = \theta(a)$ статистика $F_n(1)$ является несмещенной оценкой?

$$EF_n(1) = E(\frac{\#\{X_i < 1\}}{n}) = \frac{E(\#\{X_i < 1\})}{n} = \frac{1}{n}E\sum_{i=0}^{n} I(X_i < 1) = EI(X_i < 1) = \frac{1}{a}$$

Т.е. $F_n(1)$ - несмещенная оценка для $\theta(a) = a^{-1}$, проверим состоятельность, используя ЗБЧ:

$$F_n(1) = \frac{\#\{X_i < 1\}}{n} = \frac{1}{n}\sum_{i=0}^{n} I(X_i < 1) \xrightarrow{p} EI(X_i < 1) = \frac{1}{a}$$

3. Оценить параметры $\alpha > 0$ и $\beta \in \mathbb{R}$ по выборке из показательного распределения с плотностью $p_{\alpha,\beta}(y) = \alpha^{-1}e^{-(y-\beta)/\alpha}$ при $y > \beta$.

$$\begin{cases} m_1 = \int_\beta^{\inf} z\alpha^{-1}e^{-(z-\beta)/\alpha}dz \\ m_2 = \int_\beta^{\inf} z^2\alpha^{-1}e^{-(z-\beta)/\alpha}dz \end{cases} \quad \begin{cases} m_1 = \alpha + \beta \\ m_2 = 2\alpha^2 + 2\alpha\beta + \beta^2 \end{cases}$$

$$\begin{cases} \overline{X} = \alpha + \beta \\ \overline{X^2} = 2\alpha^2 + 2\alpha\beta + \beta^2 \end{cases} \quad \begin{cases} \alpha = \sqrt{\overline{X^2} - (\overline{X})^2} \\ \beta = \overline{X} - \sqrt{\overline{X^2} - (\overline{X})^2} \end{cases}$$

4. Найти оценку максимального правдоподобия параметра сдвига $\mu \in \mathbb{R}$ распределения Лапласа с плотностью $p_\mu(y) = \frac{e^{-|y-\mu|}}{2}$.

$$L(\theta) = \prod_{i=1}^{n} \frac{e^{-|X_i-\theta|}}{2}$$

$$l(\theta) = \log L(\theta) = \sum_{i=1}^{n} \log \frac{e^{-|X_i-\theta|}}{2} = -n * \log 2 - \sum_{i=1}^{n} |X_i - \theta|$$

Максимум функции $l(\theta)$ достигается в медиане выборки $X_1, \ldots, X_n$. т.е. $\hat{\theta}_n = X_{[\frac{n}{2}]}$.

5. Найти оценки максимального правдоподобия параметра $\theta > 0$, если распределение выборки имеет плотность

a. $\theta y^{\theta-1}$ при $y \in [0,1]$;

$$L(\theta) = \prod_{i=1}^{n} \theta X_i^{\theta-1}$$

$$l(\theta) = \log L(\theta) = n \log \theta + (\theta - 1) \sum_{i=1}^{n} \log X_i$$

$$l'(\theta) = \frac{n}{\theta} + \sum_{i=1}^{n} \log X_i = 0$$

$$\hat{\theta}_n = \frac{-n}{\sum_{i=1}^{n} \log X_i}$$

b. $\frac{2y}{\theta^2}$ при $y \in [0, \theta]$;

$$L(\theta) = \prod_{i=1}^{n} \frac{2X_i}{\theta^2}$$

$$l(\theta) = \log L(\theta) = n \log 2 - 2n \log \theta + \sum_{i=1}^{n} \log X_i$$

Из этой формулы видно, что максимум будет достигаться при $\log \theta \to 0$, однако из-за условия на плотности $y \in [0, \theta]$ лучшая оценка - $\hat{\theta}_n = \max_i(X_i)$.

c. $\frac{\theta e^{\frac{-\theta^2}{2y}}}{\sqrt{2\pi y^3}}$ при $y > 0$;

$$L(\theta) = \prod_{i=1}^{n} \frac{\theta e^{\frac{-\theta^2}{2X_i}}}{\sqrt{2\pi X_i^3}}$$

$$l(\theta) = \log L(\theta) = n \log \theta + \frac{1}{2} \sum_{i=1}^{n} (\frac{-\theta^2}{X_i} - \log(2\pi X_i^3))$$

$$l'(\theta) = \frac{n}{\theta} - \theta \sum_{i=1}^{n} X_i^{-1} = 0$$

$$\hat{\theta}_n = \sqrt{\frac{n}{\sum_{i=1}^{n} X_i^{-1}}}$$

d. $\frac{\theta (\log y)^{\theta-1}}{y}$ при $y \in [1, e]$;

$$L(\theta) = \prod_{i=1}^{n} \frac{\theta (\log X_i)^{\theta-1}}{X_i}$$

$$l(\theta) = \log L(\theta) = n \log \theta + (\theta - 1) \sum_{i=1}^{n} \log \log X_i - \sum_{i=1}^{n} \log X_i$$

$$l'(\theta) = \frac{n}{\theta} + \sum_{i=1}^{n} \log \log X_i = 0$$

$$\hat{\theta}_n = \frac{-n}{\sum_{i=1}^{n} \log \log X_i}$$

e. $\frac{e^{-|y|}}{2(1-e^{-\theta})}$ при $|y| \leq \theta$

$$L(\theta) = \prod_{i=1}^{n} \frac{e^{-|X_i|}}{2(1-e^{-\theta})} = \frac{e^{-\sum_{i=1}^{n}|X_i|}}{2^n(1-e^{-\theta})^n}$$

Так как мы ищем максимум $L(\theta)$, нам нужен положительный минимум $(1-e^{-\theta})$:

$$1 - e^{-\theta} \to 0$$

$$\theta \to 0$$

однако из-за условия на плотности $|y| \leq \theta$ лучшая оценка - $\hat{\theta}_n = \max_i(|X_i|)$.