№7 М.С. Некрасов Артём 216

*1. Пусть $X_1, \dots, X_n$ – выборка из дискретного равномерного распределения на множестве $\{1, \dots, \theta\}$, $\theta \in N$.
Найдите полную достаточную статистику в данной модели и оптимальную оценку параметра θ.*

Подсказка. При проверке статистики на полноту и решении уравнения несмещенности помните, что равенства вида

$$\mathbb{E}_\theta f(S(X)) = 0;$$
$$\mathbb{E}_\theta f(S(X)) = \theta$$

должны быть выполнены сразу для всех θ. В данной задаче полезно начать решать эти равенства «по индукции», начиная с $\theta = 1$.

Бред сумашедшего.

$$p_\theta(X) = \prod_{i=1}^{n} \frac{1}{\theta} \cdot I\{1 \le X_i \le \theta\} = \frac{1}{\theta^n} I\{X_{(1)} \ge 1, X_{(n)} \le \theta\} = \frac{1}{\theta^n} \cdot I\{X_{(1)} \ge 1\} \cdot I\{X_{(n)} \le \theta\} \quad \Rightarrow \quad X_{(n)} - \text{достаточная статистика.}$$

i. $\mathbb{E}_\theta f(X_{(n)}) = \sum_{i=1}^{\theta} f(i) P(X_{(n)} = i) = 0 \quad \forall \theta$

$\theta = 1$: $f(1) \cdot P(X_{(n)} = 1) = 0 \;\Rightarrow\; f(1) \cdot 1 = 0 \;\Rightarrow\; f(1) = 0$

$\theta = 2$: $\underbrace{f(1) \cdot P(X_{(n)} = 1)}_{0} + f(2) \cdot \underbrace{P(X_{(n)} = 2)}_{>0} = 0 \;\Rightarrow\; f(2) = 0$

$\Rightarrow$ аналогично $f(k) = 0 \;\; \forall k \in \mathbb{N} \;\Rightarrow\; f \equiv 0$ на $\mathbb{N}$

ii. $\mathbb{E}X_1 = \sum_{k=1}^{\theta} k \cdot \frac{1}{\theta} = \frac{1}{\theta} \cdot \frac{\theta(\theta+1)}{2} = \frac{\theta+1}{2} \;\Rightarrow\;$ оценка $\hat\theta = 2X_1 - 1$ является несмещенной.

По теореме $\mathbb{E}(\hat\theta | s(X))$ – несмещенная оценка θ ($s(X) = X_{(n)}$ – достаточная), так как $X_{(n)}$ – полная и $\mathbb{E}(\hat\theta | s(X))$ – функция от $s(X)$, то $\mathbb{E}(\hat\theta | s(X))$ – оптимальная.

$$\mathbb{E}(\hat\theta | X_{(n)}) = \mathbb{E}(2X_1 - 1 | X_{(n)}) = 2\mathbb{E}(X_1 | X_{(n)}) - 1$$

$$\mathbb{E}(X_1 | X_{(n)} = s) = \sum_{k=1}^{\theta} k \cdot P(X_1 = k | X_{(n)} = s)$$

$$P(X_1 = k | X_{(n)} = s) = \frac{P(X_1 = k \cap X_{(n)} = s)}{P(X_{(n)} = s)} = \frac{P(X_1 = k \cap X_{(n)} = s)}{P(X_{(n)} \le s) - P(X_{(n)} \le s - 1)} \circleddash$$

$$k = s \circleddash \frac{P(X_1 = s \cap X_{(n)} = s)}{\left(\frac{s}{\theta}\right)^n - \left(\frac{s-1}{\theta}\right)^n} = \frac{P(X_1 = s \cap X_2 \le \cdots s)}{\left(\frac{s}{\theta}\right)^n - \left(\frac{s-1}{\theta}\right)^n} = \frac{s^{n-1}}{s^n - (s-1)^n}$$

$$k < s \circleddash \frac{P(X_1 = k \cap X_{(n)} = s)}{\left(\frac{s}{\theta}\right)^n - \left(\frac{s-1}{\theta}\right)^n} = \frac{P(X_1 = k) \cdot P(\max_{[2,n]} X_i = s)}{\left(\frac{s}{\theta}\right)^n - \left(\frac{s-1}{\theta}\right)^n} = \frac{\frac{1}{\theta}\left(P(\max_{[2,n]} X_i \le s) - P(\max_{[2,n]} X_i \le s - 1)\right)}{\left(\frac{s}{\theta}\right)^n - \left(\frac{s-1}{\theta}\right)^n} = \frac{\frac{1}{\theta}\left(\left(\frac{s}{\theta}\right)^{n-1} - \left(\frac{s-1}{\theta}\right)^{n-1}\right)}{\left(\frac{s}{\theta}\right)^n - \left(\frac{s-1}{\theta}\right)^n} =$$

$$= \frac{s^{n-1} - (s-1)^{n-1}}{s^n - (s-1)^n}$$

$$\Rightarrow \mathbb{E}(X_1 | X_{(n)} = s) = \frac{s^{n-1} - (s-1)^{n-1}}{s^n - (s-1)^n} \sum_{k=1}^{s-1} k + s \cdot \frac{s^{n-1}}{s^n - (s-1)^n} + 0 \cdot \sum_{k=s+1}^{\theta} k = \frac{s(s-1)}{2} \cdot \frac{s^{n-1} - (s-1)^{n-1}}{s^n - (s-1)^n} + \frac{s^n}{s^n - (s-1)^n} =$$

$$= \frac{1}{2} \cdot \frac{s^n(s-1) - (s-1)^n s}{s^n - (s-1)^n} + \frac{s^n}{s^n - (s-1)^n}$$

$$\Rightarrow \mathbb{E}(\hat\theta | X_{(n)} = s) = 2\mathbb{E}(X_1 | X_{(n)} = s) - 1 = \frac{s^n(s-1) - (s-1)^n s + s^n + (s-1)^n}{s^n - (s-1)^n} = \frac{s^n(s-1+1) - (s-1)^n(s-1)}{s^n - (s-1)^n} = \frac{s^{n+1} - (s-1)^{n+1}}{s^n - (s-1)^n}$$

$\Rightarrow$ оптимальная оценка: $\boxed{\dfrac{X_{(n)}^{n+1} - (X_{(n)} - 1)^{n+1}}{X_{(n)}^n - (X_{(n)} - 1)^n}}$

*2. Пусть $X_1, \dots, X_n$ – выборка из экспоненциального распределения с параметром θ. Априорное распределение θ есть $\Gamma(\alpha, \beta)$.
Найдите байесовскую оценку параметра θ и проверьте ее на состоятельность.*

Байесовская оценка: $\hat\theta_\pi(X) = \widetilde{\mathbb{E}}[\theta | X]$, $\quad \hat\theta_\pi(X) = \int_\Theta t \cdot p(t|X)dt$, $\quad$ где $\quad p(\theta|X) = \dfrac{p_\theta(X)\pi(t)}{\int_\Theta p_t(X)\pi(t)dt}$

Решение.

$$p_\theta(X) = \prod_{i=1}^{n} \theta \cdot \exp\{-\theta X_i\} \cdot I\{X_i > 0\} = \theta^n \exp\left\{-\theta \cdot \sum_{i=1}^{n} X_i\right\} \cdot I\{X_i > 0\}$$

Плотность $\Gamma(\alpha, \beta)$: $\quad \pi(\theta) = \dfrac{x^{\alpha-1} \cdot \exp\{-\theta\beta\}}{\Gamma(\alpha)} \cdot \beta^\alpha \cdot I\{\theta > 0\}$

Посчитаем $p(\theta|X)$.

Числитель: $\quad p_\theta(X)\pi(\theta) \;=\; \theta^n \exp\left\{-\theta \cdot \sum_{i=1}^{n} X_i\right\} \cdot I\{X_i > 0\} \;\cdot\; \dfrac{x^{\alpha-1} \cdot \exp\{-\theta\beta\}}{\Gamma(\alpha)} \cdot \beta^\alpha \cdot I\{\theta > 0\} =$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \theta^{\alpha-1}\theta^n \cdot \exp\{-\theta\beta\} \cdot \exp\left\{-\theta \sum X_i\right\} \cdot I\{\theta > 0, X_i > 0\} = \underbrace{\frac{\beta^\alpha}{\Gamma(\alpha)}}_{const} \cdot \theta^{n+\alpha-1} \exp\left\{-\theta\beta - \theta \sum X_i\right\} \cdot I\{\theta > 0, X_i > 0\} =$$

$$= \text{const} \cdot \theta^{n+\alpha-1} \exp\left\{-\theta\left(\beta + \sum X_i\right)\right\}$$

$\begin{cases} n + \alpha = a \\ \beta + \sum X_i = b \end{cases} \Rightarrow \quad \Gamma\left(n + \alpha, \beta + \sum X_i\right)$ – апостериорное распределение $p(\theta|X)$.

Тогда $\hat\theta_\pi(X) = \mathbb{E}\left[\Gamma\left(n + \alpha, \beta + \sum X_i\right)\right] \Rightarrow \boxed{\hat\theta_\pi(X) = \dfrac{n + \alpha}{\beta + \sum X_i}}$

Проверим состоятельность:

$$\frac{n+\alpha}{\beta + \sum X_i} = \frac{n+\alpha}{\beta + \bar X n} = \frac{n}{\beta + \bar X n} + \frac{\alpha}{\beta + \bar X n} = \frac{n}{\bar X\left(\frac{\beta}{\bar X} + n\right)} + \frac{\alpha}{\bar X\left(\frac{\beta}{\bar X} + n\right)} = \frac{n}{\bar X n\left(\frac{\beta}{\bar X n} + 1\right)} + \frac{\alpha}{\bar X n\left(\frac{\beta}{\bar X n} + 1\right)} = \underbrace{\frac{1}{\bar X}}_{\theta, \text{т.к.} \bar X \to \frac{1}{\theta}} \cdot \underbrace{\frac{1}{\frac{\beta}{\bar X n} + 1}}_{1 \text{ при } n \to \infty} + \underbrace{\frac{\alpha}{\bar X n\left(\frac{\beta}{\bar X n} + 1\right)}}_{0 \text{ при } n \to \infty} = \theta$$

$\boxed{\hat\theta_\pi(X) \xrightarrow{\text{a.s.}} \theta \;\Rightarrow\; \text{по УЗБЧ оценка состоятельная.}}$

3. Пусть $X_1, \dots, X_n$ – выборка из распределения Бернулли с параметром θ. Найдите байесовскую оценку θ, если априорное распределение θ есть

a) равномерное распределение на отрезке $[0,1]$;

Решение.

i. $p_\theta(X) = \prod_{i=1}^{n} \theta^{X_i}(1-\theta)^{1-X_i} = \theta^{\sum X_i}(1-\theta)^{n - \sum X_i}$

ii. $\pi(\theta) = 1 \cdot I\{0 \le \theta \le 1\}$

iii. $p_\theta(X)\pi(\theta) = \theta^{\sum X_i}(1-\theta)^{n - \sum X_i} \cdot I\{0 \le \theta \le 1\}$

Получили что-то очень похожее на плотность бета-распределения: $\dfrac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1}$ – с точность до *const*

$$\begin{cases} \sum X_i = \alpha - 1 \\ n - \sum X_i = \beta - 1 \end{cases} \Rightarrow \begin{cases} \alpha = \sum X_i + 1 \\ \beta = n - \sum X_i + 1 \end{cases} \Rightarrow \quad B\left(\sum X_i + 1, n - \sum X_i + 1\right)$$

$$\boxed{\hat\theta_\pi(X) = \frac{\sum X_i + 1}{(\sum X_i + 1) + (n - \sum X_i + 1)} = \frac{\sum X_i + 1}{n + 2}}$$

b) двухточечное распределение со значениями $\frac{1}{4}$ и $\frac{3}{4}$ с вероятностями $\pi\left(\frac{1}{4}\right) = \pi\left(\frac{3}{4}\right) = \frac{1}{2}$.

Решение.

$\pi(\theta) = \dfrac{1}{b-a} I\{0 \le \theta \le 1\}$ – плотность для равномерного.

$$\pi(\theta) = \frac{1}{2} I\left\{\theta = \frac{1}{4}\right\} + \frac{1}{2} I\left\{\theta = \frac{3}{4}\right\}$$

$$p_\theta(X)\pi(\theta) = \theta^{\sum X_i}(1-\theta)^{n - \sum X_i}\left(\frac{1}{2} I\left\{\theta = \frac{1}{4}\right\} + \frac{1}{2} I\left\{\theta = \frac{3}{4}\right\}\right)$$

$$p_\theta(X)\pi(\theta) = \begin{cases} \frac{1}{2}\left(\frac{1}{4}\right)^{\sum X_i} \cdot \left(\frac{3}{4}\right)^{n - \sum X_i}, & \theta = \frac{1}{4} \\ \frac{1}{2}\left(\frac{3}{4}\right)^{\sum X_i} \cdot \left(\frac{1}{4}\right)^{n - \sum X_i}, & \theta = \frac{3}{4} \end{cases}$$

$$\pi\left(\theta = \frac{1}{4} \Big| X\right) = \frac{p_{\frac{1}{4}}(X)\pi\left(\frac{1}{4}\right)}{p_{\frac{1}{4}}(X)\pi\left(\frac{1}{4}\right) + p_{\frac{3}{4}}(X)\pi\left(\frac{3}{4}\right)} = \frac{\frac{1}{2}\left(\frac{1}{4}\right)^{\sum X_i} \cdot \left(\frac{3}{4}\right)^{n - \sum X_i}}{\frac{1}{2}\left(\frac{1}{4}\right)^{\sum X_i} \cdot \left(\frac{3}{4}\right)^{n - \sum X_i} + \frac{1}{2}\left(\frac{3}{4}\right)^{\sum X_i} \cdot \left(\frac{1}{4}\right)^{n - \sum X_i}} = \frac{\left(\frac{1}{4}\right)^{\sum X_i} \cdot \left(\frac{3}{4}\right)^{n - \sum X_i}}{\left(\frac{1}{4}\right)^{\sum X_i} \cdot \left(\frac{3}{4}\right)^{n - \sum X_i} + \left(\frac{3}{4}\right)^{\sum X_i} \cdot \left(\frac{1}{4}\right)^{n - \sum X_i}} =$$

$$= \frac{3^n}{3^n + 3^{2\sum X_i}}$$

$$\pi\left(\theta = \frac{3}{4} \Big| X\right) = 1 - \pi\left(\theta = \frac{1}{4} \Big| X\right) = 1 - \frac{3^n}{3^n + 3^{2\sum X_i}} = \frac{3^{2\sum X_i}}{3^n + 3^{2\sum X_i}}$$

$$\boxed{\hat\theta_\pi(X) = \frac{1}{4} \cdot \frac{3^n}{3^n + 3^{2\sum X_i}} + \frac{3}{4} \cdot \frac{3^{2\sum X_i}}{3^n + 3^{2\sum X_i}} = \frac{3 \cdot 3^{2\sum X_i} + 3^n}{4(3^n + 3^{2\sum X_i})}}$$