Семинар №4 ПМИ 239

Эффективные оценки

Листок

Определение 1. Пусть $\mathbf{X} = (X_1, \dots, X_n)$ – выборка из распределения $P_\theta(x)$, $\theta \in \Theta$ с обобщенной плотностью p_θ. Тогда *функцией правдоподобия* выборки $\mathbf{X}$ называется величина

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

Определение 2. Пусть $\Theta \subseteq \mathbb{R}$ – открытое множество. Семейство распределений P будем называть регулярным, если оно удовлетворяет следующим условиям:

1. Множество $A = \{x : p_\theta(x) > 0\}$ не зависит от параметра θ.
2. Для любой статистики $T(\mathbf{X})$ такой, что $E_\theta[T^2(\mathbf{X})] < +\infty$, законно дифференцирование по параметру под знаком интеграла:
$$\frac{\partial}{\partial\theta} E_\theta T(\mathbf{X}) = E_\theta T(\mathbf{X}) U_\theta(\mathbf{X}), \quad \text{где } U_\theta(\mathbf{X}) = \frac{\partial}{\partial\theta} \ln p_\theta(\mathbf{X}).$$
3. Информация Фишера выборки $\mathbf{X}$ положительна и конечна при всех $\theta \in \Theta : 0 < I_n(\theta) < \infty$.

Определение 3. Пусть $\mathbf{X} = (X_1, \dots, X_n)$ – выборка из распределения с обобщенной плотностью $p_\theta(x)$. Тогда *вкладом* выборки $\mathbf{X}$ называется

$$U_\theta(\mathbf{X}) = \frac{\partial}{\partial\theta} \ln p_\theta(\mathbf{X}).$$

Теорема 1. *Пусть выполнены условия регулярной модели. Тогда для любой несмещённой оценки $T(\mathbf{X})$ функции $\tau(\theta)$ справедливо неравенство Рао-Крамера:*

$$Var_\theta(T(\mathbf{X})) \geqslant \frac{(\tau'(\theta))^2}{nI_1(\theta)}. \tag{1}$$

Величина $I_n(\theta) = E_\theta[U_\theta(\mathbf{X})^2]$ называется ***информацией Фишера****.*

Теорема 2. *Неравенство (1) обращается в равенство тогда и только тогда, когда оценка $\hat{\theta}(\mathbf{X})$ является линейной функцией вклада выборки, то есть*

$$\hat{\theta}(\mathbf{X}) - \tau(\theta) = c(\theta) U_\theta(\mathbf{X}), \quad \text{где } c(\theta) = \frac{\tau'(\theta)}{nI_1(\theta)}.$$

Замечание 1. Если выполнены условия регулярной модели, то

$$I_n(\theta) = nI_1(\theta),$$

где $I_1(\theta)$ — информация Фишера одного наблюдения.

Также в регулярных семействах *информация Фишера* может быть посчитана

$$I_n(\theta) = -E\left[\frac{\partial}{\partial\theta} U_\theta(\mathbf{X})\right].$$

Определение 4. Оценка $\hat{\theta}(\mathbf{X})$ называется *эффективной*, если для неё достигается равенство в неравенстве Рао-Крамера.

Определение 5. Говорят, что распределение $\mathbb{P}_\theta$ принадлежит экспоненциальному семейству, если его обобщённая плотность представима в виде (каноническая параметризация с $\eta = \eta(\theta)$)

$$p_\eta(x) = e^{\eta T(x) - A(\eta)} h(x),$$

где T, A, h — некоторые функции, $\eta = \eta(\theta)$ — параметр распределения.

Замечание 2. Отметим, что в определении используется натуральная параметризация (η — натуральный параметр).

Теорема 3. *Пусть* $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ *— выборка из распределения* $\{\mathbb{P}_\theta : \theta \in \Theta \subseteq \mathbb{R}\}$. *Эффективная оценка* $\hat{\theta}(\mathbf{X})$ *для некоторой функции* $\tau(\theta)$ *существует тогда и только тогда, когда* $\mathbb{P}_\theta$ *принадлежит экспоненциальному семейству, причём*

$$\hat{\theta}(\mathbf{X}) = \frac{1}{n}\sum_{i=1}^{n} T(X_i), \quad \text{и} \quad \tau(\theta) = \frac{dA}{d\eta} \quad \left(\text{из курса матанализа } \frac{dA}{d\theta} = \frac{dA}{d\eta}\cdot\frac{d\eta}{d\theta}\right)$$

(последние два равенства следует понимать с точностью до линейного преобразования).

Пример 1. *Для биномиального распределения* $Bin(n, p)$*:*

- $\eta = \ln\frac{p}{1-p}$,
- $A(\eta) = n\ln(1 + e^\eta)$,
- $\tau(p) = np$, *и* $\hat{p} = \frac{1}{n}\sum X_i = \overline{\mathbf{X}}$ *— эффективная оценка.*

Задачи

$\boxed{1}$ Посчитайте информацию Фишера и найдите эффективную оценку параметра θ, если дана выборка $X_1, \ldots, X_n$ из:

(a) $\mathrm{Ber}(\theta)$;

(b) $\mathcal{N}(\theta, \sigma^2)$ с известным $\sigma^2 > 0$;

(c) $\mathcal{N}(\mu, \theta)$ с известным $\mu, \theta > 0$.

$\boxed{\mathbf{2}}$ Пусть $X_1, \ldots, X_n$ — выборка из экспоненциального распределения $\mathrm{Exp}(\theta)$ с параметром $\theta > 0$. Для какой функции $\tau(\theta)$ существует эффективная оценка?

$\boxed{\mathbf{3}}$ Пусть $\mathbf{X} = (X_1, \ldots, X_n)$ — выборка из экспоненциального распределения $\mathrm{Exp}(\theta)$ с параметром $\theta > 0$. Далее, пусть $X_{(1)} = \min_{1 \leqslant i \leqslant n} X_i$. Вычислите информацию Фишера статистик $X_{(1)}$ и $\sum_{i=1}^n X_i$. Сравните с информацией Фишера выборки, посчитанной в предыдущей задаче.

Подсказка: Сумма независимых случайных величин с экспоненциальным распределением $\mathrm{Exp}(\theta)$ имеет гамма-распределение $\Gamma(n, \theta)$.

$\boxed{\mathbf{4}}$ Пусть дана выборка $X_1, \ldots, X_n \sim U([0, \theta])$ из равномерного распределения. Покажите, что оценка $\frac{n+1}{n} X_{(n)}$ является сверх-эффективной для параметра θ, то есть её дисперсия меньше, чем нижняя граница в неравенстве Рао-Крамера. Объясните, почему для оценки $\frac{n+1}{n} X_{(n)}$ не выполняется неравенство Рао-Крамера.

$\boxed{\mathbf{5}}$ В каждом из следующих случаев установите, относится ли соответствующее распределение к экспоненциальному семейству. В случае положительного ответа определите функции $\tau(\theta)$, для которых существует эффективная оценка, и установите вид этой оценки:

(a) $\mathcal{N}(\theta, \sigma^2)$, где $\sigma^2 > 0$ известно:

$$p_\theta(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\theta)^2}{2\sigma^2}}, \quad x \in \mathbb{R};$$

(b) $\mathcal{N}(\mu, \theta)$, где $\mu \in \mathbb{R}$ известно:

$$p_\theta(x) = \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{(x-\mu)^2}{2\theta}}, \quad x \in \mathbb{R};$$

(c) $\mathrm{Exp}(\theta), \theta > 0$:

$$p_\theta(x) = \theta e^{-\theta x} \mathbb{I}\{x > 0\};$$

(d) $\Gamma(\alpha, \theta)$, где $\alpha > 0$ известно:

$$p_\theta(x) = \frac{\theta^\alpha x^{\alpha-1}}{\Gamma(\alpha)} e^{-\theta x} \mathbb{I}\{x > 0\};$$

(e) Биномиальное распределение $\mathrm{Bin}(k, \theta)$, $\theta \in (0, 1]$:

$$p_\theta(x) = C_k^x \theta^x (1-\theta)^{k-x}, \quad x \in \{0, 1, \ldots, k\};$$

(f) $\mathrm{Pois}(\theta)$:

$$p_\theta(x) = e^{-\theta} \frac{\theta^x}{x!}, \quad \theta > 0, \quad x \in \{0, 1, \ldots\}.$$

Решение задач

Задача 1

(a) Бернуллиевское распределение Ber(θ)

Информация Фишера

Функция правдоподобия для выборки $X_1, \dots, X_n$:

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

Логарифмическая функция правдоподобия:

$$\ln p_\theta(\mathbf{X}) = \sum_{i=1}^{n} X_i \ln\theta + (n - \sum_{i=1}^{n} X_i)\ln(1-\theta).$$

Первая производная по θ:

$$U_\theta(\mathbf{X}) = \frac{\partial}{\partial\theta}\ln p_\theta(\mathbf{X}) = \frac{\sum_{i=1}^{n} X_i}{\theta} - \frac{n - \sum_{i=1}^{n} X_i}{1-\theta}.$$

Вторая производная:

$$\frac{\partial^2}{\partial\theta^2}\ln p_\theta(\mathbf{X}) = -\frac{\sum_{i=1}^{n} X_i}{\theta^2} - \frac{n - \sum_{i=1}^{n} X_i}{(1-\theta)^2}.$$

Информация Фишера для одного наблюдения:

$$I_1(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2}\ln p_\theta(\mathbf{X})\right] = \frac{1}{\theta} + \frac{1}{1-\theta} = \frac{1}{\theta(1-\theta)}.$$

Для выборки размера n:

$$I_n(\theta) = nI_1(\theta) = \frac{n}{\theta(1-\theta)}.$$

Эффективная оценка

Несмещенная оценка $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Дисперсия оценки:

$$\mathrm{Var}(\hat{\theta}) = \frac{\theta(1-\theta)}{n}.$$

Нижняя граница Рао-Крамера:

$$\frac{(\tau'(\theta))^2}{nI_1(\theta)} = \frac{1}{n \cdot \frac{1}{\theta(1-\theta)}} = \frac{\theta(1-\theta)}{n}.$$

Равенство достигается, следовательно, $\hat{\theta}$ эффективна.

(b) Нормальное распределение $\mathcal{N}(\theta, \sigma^2)$, σ^2 известно

Информация Фишера

Логарифмическая функция правдоподобия:

$$\ln p_\theta(\mathbf{X}) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i - \theta)^2.$$

Первая производная по θ:

$$U_\theta(\mathbf{X}) = \frac{\partial}{\partial\theta}\ln p_\theta(\mathbf{X}) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \theta).$$

Вторая производная:

$$\frac{\partial^2}{\partial\theta^2}\ln p_\theta(\mathbf{X}) = -\frac{n}{\sigma^2}.$$

Информация Фишера:

$$I_n(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2}\ln p_\theta(\mathbf{X})\right] = \frac{n}{\sigma^2}.$$

Эффективная оценка

Оценка $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Дисперсия:

$$\mathrm{Var}(\hat{\theta}) = \frac{\sigma^2}{n}.$$

Нижняя граница Рао-Крамера:

$$\frac{1}{n \cdot \frac{1}{\sigma^2}} = \frac{\sigma^2}{n}.$$

Равенство выполняется, оценка эффективна.

(c) Нормальное распределение $\mathcal{N}(\mu, \theta)$, μ известно

Информация Фишера

Логарифмическая функция правдоподобия:

$$\ln p_\theta(\mathbf{X}) = -\frac{n}{2}\ln(2\pi\theta) - \frac{1}{2\theta}\sum_{i=1}^{n}(X_i - \mu)^2.$$

Первая производная по θ:

$$U_\theta(\mathbf{X}) = -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^{n}(X_i - \mu)^2.$$

Вторая производная:

$$\frac{\partial^2}{\partial\theta^2}\ln p_\theta(\mathbf{X}) = \frac{n}{2\theta^2} - \frac{1}{\theta^3}\sum_{i=1}^{n}(X_i - \mu)^2.$$

Информация Фишера:

$$I_n(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2}\ln p_\theta(\mathbf{X})\right] = \frac{n}{2\theta^2}.$$

Эффективная оценка Оценка $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$. Дисперсия:

$$\mathrm{Var}(\hat{\theta}) = \frac{2\theta^2}{n}.$$

Нижняя граница Рао-Крамера:

$$\frac{1}{n \cdot \frac{1}{2\theta^2}} = \frac{2\theta^2}{n}.$$

Равенство достигается, оценка эффективна.

Ответ

(a) $I_n(\theta) = \dfrac{n}{\theta(1-\theta)}$, эффективная оценка $\hat{\theta} = \dfrac{1}{n}\sum X_i$. (b) $I_n(\theta) = \dfrac{n}{\sigma^2}$, эффективная оценка $\hat{\theta} = \overline{X}$. (c) $I_n(\theta) = \dfrac{n}{2\theta^2}$, эффективная оценка $\hat{\theta} = \dfrac{1}{n}\sum(X_i - \mu)^2$.

Задача 2

Проверка принадлежности к экспоненциальному семейству

Плотность распределения $\mathrm{Exp}(\theta)$ задается как:

$$p_\theta(x) = \theta e^{-\theta x} \quad \text{для } x \geqslant 0.$$

Преобразуем её к канонической форме экспоненциального семейства:

$$p_\theta(x) = e^{-\theta x + \ln\theta}.$$

Сравнивая с общим видом $p_\eta(x) = e^{\eta T(x) - A(\eta)}h(x)$, получаем:

$$\eta = -\theta, \quad T(x) = x, \quad A(\eta) = -\ln(-\eta), \quad h(x) = 1 \quad \text{для } x \geqslant 0.$$

Таким образом, распределение принадлежит экспоненциальному семейству.

Определение функции $\tau(\theta)$

Согласно Теореме 2, эффективная оценка существует для функции:

$$\tau(\theta) = \frac{A'(\theta)}{\eta'(\theta)},$$

где $\eta = -\theta$. Вычислим производные:

$$\eta'(\theta) = -1,$$

$$A(\eta) = -\ln(-\eta) \implies A'(\eta) = \frac{1}{\eta}.$$

Переходя к переменной θ:

$$A'(\theta) = \frac{1}{-\theta} = -\frac{1}{\theta}.$$

Подставляем в формулу для $\tau(\theta)$:

$$\tau(\theta) = \frac{-\frac{1}{\theta}}{-1} = \frac{1}{\theta}.$$

Эффективная оценка

Согласно теореме, эффективная оценка имеет вид:

$$\hat{\tau}(\mathbf{X}) = \frac{1}{n}\sum_{i=1}^{n} T(X_i) = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Проверка несмещенности

Математическое ожидание X_i для $\mathrm{Exp}(\theta)$:

$$E[X_i] = \frac{1}{\theta}.$$

Следовательно:

$$E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n}\cdot n\cdot\frac{1}{\theta} = \frac{1}{\theta}.$$

Оценка несмещенная.

Проверка эффективности через неравенство Рао-Крамера

Логарифм плотности:

$$\ln p_\theta(x) = \ln\theta - \theta x.$$

Первая производная по θ:

$$\frac{\partial}{\partial\theta}\ln p_\theta(x) = \frac{1}{\theta} - x.$$

Вторая производная:

$$\frac{\partial^2}{\partial\theta^2}\ln p_\theta(x) = -\frac{1}{\theta^2}.$$

Информация Фишера для одного наблюдения:

$$I_1(\theta) = -E\left[-\frac{1}{\theta^2}\right] = \frac{1}{\theta^2}.$$

Для выборки размера n:

$$I_n(\theta) = n\cdot\frac{1}{\theta^2} = \frac{n}{\theta^2}.$$

Нижняя граница Рао-Крамера для $\tau(\theta) = \frac{1}{\theta}$:

$$\frac{(\tau'(\theta))^2}{nI_1(\theta)} = \frac{\left(-\frac{1}{\theta^2}\right)^2}{n \cdot \frac{1}{\theta^2}} = \frac{\frac{1}{\theta^4}}{\frac{n}{\theta^2}} = \frac{1}{n\theta^2}.$$

Дисперсия оценки $\hat{\tau}(\mathbf{X}) = \overline{X}$:

$$\mathrm{Var}(\overline{X}) = \frac{\mathrm{Var}(X_i)}{n} = \frac{\frac{1}{\theta^2}}{n} = \frac{1}{n\theta^2}.$$

Равенство достигается, следовательно, оценка эффективна.

Ответ

Для функции $\tau(\theta) = \dfrac{1}{\theta}$ существует эффективная оценка $\hat{\tau}(\mathbf{X}) = \dfrac{1}{n}\sum_{i=1}^n X_i$.

Задача 3

Информация Фишера для статистики $X_{(1)} = \min_{1\leqslant i\leqslant n} X_i$

Распределение $X_{(1)}$ для выборки $X_1, \ldots, X_n \sim \mathrm{Exp}(\theta)$ имеет плотность:

$$p_{X_{(1)}}(x) = n\theta e^{-n\theta x} \quad \text{для } x \geqslant 0.$$

Логарифм плотности

$$\ln p_{X_{(1)}}(x) = \ln(n\theta) - n\theta x.$$

Первая производная по θ

$$\frac{\partial}{\partial\theta} \ln p_{X_{(1)}}(x) = \frac{n}{n\theta} - nx = \frac{1}{\theta} - nx.$$

Вторая производная по θ

$$\frac{\partial^2}{\partial\theta^2} \ln p_{X_{(1)}}(x) = -\frac{1}{\theta^2}.$$

Информация Фишера для $X_{(1)}$

$$I_{X_{(1)}}(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2} \ln p_{X_{(1)}}(x)\right] = -E\left[-\frac{1}{\theta^2}\right] = \frac{1}{\theta^2}.$$

—

Информация Фишера для статистики $\sum_{i=1}^n X_i$

Сумма $\sum_{i=1}^n X_i$ имеет гамма-распределение $\mathrm{Gamma}(n, \theta)$ с плотностью:

$$p_{\sum X_i}(x) = \frac{\theta^n x^{n-1} e^{-\theta x}}{\Gamma(n)} \quad \text{для } x \geqslant 0.$$

Логарифм плотности

$$\ln p_{\sum X_i}(x) = n\ln\theta + (n-1)\ln x - \theta x - \ln\Gamma(n).$$

Первая производная по θ

$$\frac{\partial}{\partial\theta}\ln p_{\sum X_i}(x) = \frac{n}{\theta} - x.$$

Вторая производная по θ

$$\frac{\partial^2}{\partial\theta^2}\ln p_{\sum X_i}(x) = -\frac{n}{\theta^2}.$$

Информация Фишера для $\sum X_i$

$$I_{\sum X_i}(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2}\ln p_{\sum X_i}(x)\right] = -E\left[-\frac{n}{\theta^2}\right] = \frac{n}{\theta^2}.$$

Сравнение с информацией Фишера выборки
Из предыдущей задачи информация Фишера для всей выборки $X_1, \ldots, X_n$:

$$I_n(\theta) = \frac{n}{\theta^2}.$$

- **Для $X_{(1)}$**
$$I_{X_{(1)}}(\theta) = \frac{1}{\theta^2} \quad \text{(значительно меньше } I_n(\theta)\text{)}.$$

- **Для $\sum X_i$**
$$I_{\sum X_i}(\theta) = \frac{n}{\theta^2} \quad \text{(совпадает с } I_n(\theta)\text{)}.$$

Ответ

Информация Фишера:

- Для $X_{(1)}$: $I_{X_{(1)}}(\theta) = \dfrac{1}{\theta^2}$.

- Для $\sum X_i$: $I_{\sum X_i}(\theta) = \dfrac{n}{\theta^2}$.

Сравнение: $I_{\sum X_i}(\theta) = I_n(\theta)$, $I_{X_{(1)}}(\theta) < I_n(\theta)$.

Задача 4

Распределение статистики $X_{(n)} = \max_{1\leqslant i\leqslant n} X_i$
Плотность распределения $X_{(n)}$:

$$p_{X_{(n)}}(x) = \frac{nx^{n-1}}{\theta^n} \quad \text{для } 0 \leqslant x \leqslant \theta.$$

Математическое ожидание $X_{(n)}$

$$E[X_{(n)}] = \int_0^\theta x \cdot \frac{nx^{n-1}}{\theta^n}dx = \frac{n}{\theta^n}\int_0^\theta x^n dx = \frac{n}{\theta^n}\cdot\frac{\theta^{n+1}}{n+1} = \frac{n}{n+1}\theta.$$

Дисперсия $X_{(n)}$

Сначала найдем $E[X_{(n)}^2]$:

$$E[X_{(n)}^2] = \int_0^\theta x^2 \cdot \frac{nx^{n-1}}{\theta^n} dx = \frac{n}{\theta^n} \int_0^\theta x^{n+1} dx = \frac{n}{\theta^n} \cdot \frac{\theta^{n+2}}{n+2} = \frac{n}{n+2}\theta^2.$$

Тогда:

$$\mathrm{Var}(X_{(n)}) = E[X_{(n)}^2] - (E[X_{(n)}])^2 = \frac{n}{n+2}\theta^2 - \left(\frac{n}{n+1}\theta\right)^2 = \frac{n\theta^2}{(n+1)^2(n+2)}.$$

Теперь покажем, что $\hat{\theta} = \frac{n+1}{n} X_{(n)}$ будет несмещенной оценкой:

$$E[\hat{\theta}] = \frac{n+1}{n} E[X_{(n)}] = \frac{n+1}{n} \cdot \frac{n}{n+1}\theta = \theta.$$

Для дальнейших выкладок сразу посчитаем дисперсию оценки

$$\mathrm{Var}(\hat{\theta}) = \left(\frac{n+1}{n}\right)^2 \mathrm{Var}(X_{(n)}) = \left(\frac{n+1}{n}\right)^2 \cdot \frac{n\theta^2}{(n+1)^2(n+2)} = \frac{\theta^2}{n(n+2)}.$$

Нижняя граница Рао-Крамера

Для регулярных моделей нижняя граница вычисляется как:

$$\frac{(\tau'(\theta))^2}{nI_1(\theta)}.$$

Однако модель $U([0,\theta])$ **не является регулярной**, так как носитель распределения $\{x : p_\theta(x) > 0\} = [0,\theta]$ зависит от параметра θ. Это нарушает первое условие регулярности.

Вычисление информации Фишера

Логарифмическая функция правдоподобия:

$$\ln p_\theta(\mathbf{X}) = -n \ln \theta.$$

Первая производная по θ:

$$U_\theta(\mathbf{X}) = \frac{\partial}{\partial \theta} \ln p_\theta(\mathbf{X}) = -\frac{n}{\theta}.$$

Вторая производная:

$$\frac{\partial^2}{\partial \theta^2} \ln p_\theta(\mathbf{X}) = \frac{n}{\theta^2}.$$

Информация Фишера:

$$I_n(\theta) = -E\left[\frac{\partial^2}{\partial \theta^2} \ln p_\theta(\mathbf{X})\right] = -E\left[\frac{n}{\theta^2}\right] = \frac{n}{\theta^2}.$$

Нижняя граница Рао-Крамера для несмещенной оценки $\tau(\theta) = \theta$:

$$\frac{(1)^2}{n \cdot \frac{1}{\theta^2}} = \frac{\theta^2}{n}.$$

Вспомним, что $\hat{\theta}$

$$\mathrm{Var}(\hat{\theta}) = \frac{\theta^2}{n(n+2)} < \frac{\theta^2}{n} = I_n(\theta),$$

то есть оценка $\hat{\theta}$ является сверх-эффективной.

Почему неравенство Рао-Крамера не выполняется? Модель $U([0,\theta])$ нерегулярна: носитель распределения зависит от θ. Отметим, что керавенство Рао-Крамера применимо только в регулярных моделях, где выполнены все условия (независимость носителя от θ, дифференцируемость под знаком интеграла).

Ответ

Оценка $\hat{\theta} = \frac{n+1}{n} X_{(n)}$ имеет дисперсию $\mathrm{Var}(\hat{\theta}) = \dfrac{\theta^2}{n(n+2)}$, которая меньше формальной нижней границы Рао-Крамера $\dfrac{\theta^2}{n}$. Неравенство Рао-Крамера не выполняется, так как модель $U([0,\theta])$ нерегулярна (носитель зависит от θ).

Задача 5

(a) Нормальное распределение $\mathcal{N}(\theta, \sigma^2)$, σ^2 известно

Проверка принадлежности к экспоненциальному семейству:

Плотность распределения:

$$p_\theta(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\theta)^2}{2\sigma^2}}, \quad x \in \mathbb{R}.$$

Преобразуем к канонической форме:

$$p_\theta(x) = \underbrace{e^{\frac{\theta x}{\sigma^2} - \frac{\theta^2}{2\sigma^2}}}_{\text{Экспоненциальный член}} \cdot \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}}_{h(x)}.$$

Сравнение с общим видом $p_\eta(x) = e^{\eta T(x) - A(\eta)} h(x)$ даёт:

$$\eta = \frac{\theta}{\sigma^2}, \quad T(x) = x, \quad A(\eta) = \frac{\sigma^2 \eta^2}{2}, \quad h(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}.$$

Вывод: Да, распределение принадлежит экспоненциальному семейству.

Функция $\tau(\theta)$ и эффективная оценка:

Согласно теореме 2:

$$\tau(\theta) = \frac{A'(\eta)}{\eta'(\theta)}, \quad \text{где } \eta = \frac{\theta}{\sigma^2}.$$

Вычислим производные:

$$\begin{aligned} A(\eta) &= \frac{\sigma^2\eta^2}{2} \implies A'(\eta) = \sigma^2\eta = \theta, \\ \eta'(\theta) &= \frac{1}{\sigma^2}. \end{aligned}$$

Тогда:

$$\tau(\theta) = \frac{\theta}{\frac{1}{\sigma^2}} = \theta\sigma^2.$$

Корректная параметризация показывает, что $\tau(\theta) = \theta$. Эффективная оценка:

$$\hat{\tau}(X) = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

(b) Нормальное распределение $\mathcal{N}(\mu, \theta)$, μ известно

Проверка принадлежности к экспоненциальному семейству:

Плотность распределения:

$$p_\theta(x) = \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{(x-\mu)^2}{2\theta}}, \quad x \in \mathbb{R}.$$

Преобразуем к канонической форме:

$$p_\theta(x) = \underbrace{e^{-\frac{(x-\mu)^2}{2\theta} - \frac{1}{2}\ln(2\pi\theta)}}_{\text{Экспоненциальный член}}.$$

Сравнение с общим видом даёт:

$$\eta = -\frac{1}{2\theta}, \quad T(x) = (x-\mu)^2, \quad A(\eta) = -\frac{1}{2}\ln(-2\eta) - \frac{1}{2}\ln(2\pi), \quad h(x) = 1.$$

Вывод: Да, распределение принадлежит экспоненциальному семейству.

Функция $\tau(\theta)$ и эффективная оценка:

Вычислим производные:

$$\begin{aligned} A(\eta) &= -\frac{1}{2}\ln(-2\eta) - \frac{1}{2}\ln(2\pi) \implies A'(\eta) = \frac{1}{2\eta}, \\ \eta'(\theta) &= \frac{1}{2\theta^2}. \end{aligned}$$

Тогда:

$$\tau(\theta) = \frac{\frac{1}{2\eta}}{\frac{1}{2\theta^2}} = \theta.$$

Эффективная оценка:

$$\hat{\tau}(\mathbf{X}) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2.$$

(c) Экспоненциальное распределение $\mathrm{Exp}(\theta)$

Проверка принадлежности к экспоненциальному семейству:

Плотность распределения:

$$p_\theta(x) = \theta e^{-\theta x}, \quad x > 0.$$

Преобразуем к канонической форме:

$$p_\theta(x) = \underbrace{e^{-\theta x + \ln\theta}}_{\text{Экспоненциальный член}} \cdot \underbrace{1}_{h(x)}.$$

Сравнение с общим видом даёт:

$$\eta = -\theta, \quad T(x) = x, \quad A(\eta) = -\ln(-\eta), \quad h(x) = 1.$$

Вывод: Да, распределение принадлежит экспоненциальному семейству.

Функция $\tau(\theta)$ и эффективная оценка:

Вычислим производные:

$$A(\eta) = -\ln(-\eta) \implies A'(\eta) = \frac{1}{\eta},$$
$$\eta'(\theta) = -1.$$

Тогда:

$$\tau(\theta) = \frac{\frac{1}{-\theta}}{-1} = \frac{1}{\theta}.$$

Эффективная оценка:

$$\hat{\tau}(\mathbf{X}) = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

(d) Гамма-распределение $\Gamma(\alpha, \theta)$, $\alpha > 0$ известно

Проверка принадлежности к экспоненциальному семейству:

Плотность распределения:

$$p_\theta(x) = \frac{\theta^\alpha x^{\alpha-1}}{\Gamma(\alpha)} e^{-\theta x}, \quad x > 0.$$

Преобразуем к канонической форме:

$$p_\theta(x) = \underbrace{e^{-\theta x + \alpha\ln\theta}}_{\text{Экспоненциальный член}} \cdot \underbrace{\frac{x^{\alpha-1}}{\Gamma(\alpha)}}_{h(x)}.$$

Сравнение с общим видом $p_\eta(x) = e^{\eta T(x) - A(\eta)} h(x)$ даёт:

$$\eta = -\theta, \quad T(x) = x, \quad A(\eta) = -\alpha\ln(-\eta), \quad h(x) = \frac{x^{\alpha-1}}{\Gamma(\alpha)}.$$

Вывод: Да, распределение принадлежит экспоненциальному семейству.

Функция $\tau(\theta)$ и эффективная оценка:

Согласно теореме 2:

$$\tau(\theta) = \frac{A'(\eta)}{\eta'(\theta)}, \quad \text{где } \eta = -\theta.$$

Вычислим производные:

$$A(\eta) = -\alpha \ln(-\eta) \implies A'(\eta) = -\frac{\alpha}{\eta},$$
$$\eta'(\theta) = -1.$$

Тогда:

$$\tau(\theta) = \frac{-\frac{\alpha}{-\theta}}{-1} = \frac{\alpha}{\theta}.$$

Эффективная оценка:

$$\hat{\tau}(\mathbf{X}) = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

(e) Биномиальное распределение $\mathbf{Bin}(k, \theta)$

Проверка принадлежности к экспоненциальному семейству:

Плотность распределения:

$$p_\theta(x) = \binom{k}{x}\theta^x(1-\theta)^{k-x}, \quad x \in \{0, 1, \ldots, k\}.$$

Преобразуем к канонической форме:

$$p_\theta(x) = \underbrace{e^{x\ln\left(\frac{\theta}{1-\theta}\right)+k\ln(1-\theta)}}_{\text{Экспоненциальный член}} \cdot \underbrace{\binom{k}{x}}_{h(x)}.$$

Сравнение с общим видом даёт:

$$\eta = \ln\left(\frac{\theta}{1-\theta}\right), \quad T(x) = x, \quad A(\eta) = k\ln(1+e^\eta), \quad h(x) = \binom{k}{x}.$$

Вывод: Да, распределение принадлежит экспоненциальному семейству.

Функция $\tau(\theta)$ и эффективная оценка:

Вычислим производные:

$$A(\eta) = k\ln(1+e^\eta) \implies A'(\eta) = \frac{ke^\eta}{1+e^\eta},$$
$$\eta'(\theta) = \frac{1}{\theta(1-\theta)}.$$

Тогда:

$$\tau(\theta) = \frac{\frac{ke^\eta}{1+e^\eta}}{\frac{1}{\theta(1-\theta)}} = k\theta.$$

Эффективная оценка:

$$\hat{\tau}(\mathbf{X}) = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

То есть для $\tau(\theta)$ эффективная оценка $\hat{\tau}(\mathbf{X})$ будет выборочное среднее $\overline{\mathbf{X}}$.

(f) Пуассоновское распределение Pois(θ)

Проверка принадлежности к экспоненциальному семейству:

Плотность распределения:

$$p_\theta(x) = e^{-\theta}\frac{\theta^x}{x!}, \quad x \in \{0, 1, \dots\}.$$

Преобразуем к канонической форме:

$$p_\theta(x) = \underbrace{e^{x\ln\theta-\theta}}_{\text{Экспоненциальный член}} \cdot \underbrace{\frac{1}{x!}}_{h(x)}.$$

Сравнение с общим видом даёт:

$$\eta = \ln\theta, \quad T(x) = x, \quad A(\eta) = e^\eta = \theta, \quad h(x) = \frac{1}{x!}.$$

Вывод: Да, распределение принадлежит экспоненциальному семейству.

Функция $\tau(\theta)$ и эффективная оценка:

Вычислим производные:

$$A(\eta) = e^\eta \implies A'(\eta) = e^\eta,$$
$$\eta'(\theta) = \frac{1}{\theta}.$$

Тогда:

$$\tau(\theta) = \frac{e^\eta}{\frac{1}{\theta}} = \theta.$$

Эффективная оценка:

$$\hat{\tau}(\mathbf{X}) = \frac{1}{n}\sum_{i=1}^{n} X_i.$$