Математическая статистика. ДЗ 12.

Прохоров Юрий, 776

Задача 1

Доказать, что если функцию правдоподобия $L(\mathbf{x}\,|\,\theta)$ можно представить в виде

$$\frac{d}{d\theta}\ln L(\mathbf{x}\,|\,\theta) = a(\theta)\big(T(\mathbf{x}) - \theta\big), \tag{$*$}$$

то

(a) оценка максимального правдоподобия имеет вид

$$\widehat{\theta}^{\text{ОМП}}(\mathbf{X}) = T(\mathbf{X})$$

(b) существует эффективная оценка

$$\widehat{\theta}^{\text{эфф}}(\mathbf{X}) = T(\mathbf{X})$$

(c) коэффициент $a(\theta)$ равен информации по Фишеру

$$a(\theta) = I_n(\theta)$$

Решение:

(a) Зафиксируем конкретное значение выборки $\mathbf{x}$. Тогда оценка максимального правдоподобия $\widehat{\theta}^{\text{ОМП}}$ находится из условия:

$$L(\mathbf{x}\,|\,\theta) \longrightarrow \max_\theta \qquad \Longleftrightarrow \qquad \ln L(\mathbf{x}\,|\,\theta) \longrightarrow \max_\theta \qquad \Longleftrightarrow \qquad \frac{d}{d\theta}\ln L(\mathbf{x}\,|\,\theta)\Big|_{\theta=\widehat{\theta}^{\text{ОМП}}} = 0$$

Тогда из условия $(*)$ в точке $\theta = \widehat{\theta}^{\text{ОМП}}$:

$$\frac{d}{d\theta}\ln L(\mathbf{x}\,|\,\theta)\Big|_{\theta=\widehat{\theta}^{\text{ОМП}}} = a\big(\widehat{\theta}^{\text{ОМП}}\big)\big(T(\mathbf{x}) - \widehat{\theta}^{\text{ОМП}}\big) = 0$$

Отсюда следует, что

$$\widehat{\theta}^{\text{ОМП}} = T(\mathbf{x})$$

(b) Пусть $T(\mathbf{X})$ — некоторая несмещенная оценка параметра θ.

При доказательстве неравенства Рао-Крамера используется неравенство Коши-Буняковского для случайных величин $\frac{d}{d\theta}\ln L(\mathbf{X}\,|\,\theta)$ и $\big(T(\mathbf{X}) - \theta\big)$. Из теории вероятностей известно, что оно обращается в равенство тогда и только тогда, эти случайные величины линейно зависимы, то есть когда выполнено уравнение $(*)$.

У нас условие $(*)$ выполнено, значит, неравенство Рао-Крамера обращается в равенство и $T(\mathbf{X})$ — эффективная оценка.

(c) Посчитаем в уравнении $(*)$ дисперсию левой и правой частей. Левая часть:

$$\mathbb{V}\left[\frac{d}{d\theta}\ln L(\mathbf{X}\,|\,\theta)\right] = \mathbb{E}\left[\left(\frac{d}{d\theta}\ln L(\mathbf{X}\,|\,\theta) - \underbrace{\mathbb{E}\left[\frac{d}{d\theta}\ln L(\mathbf{X}\,|\,\theta)\right]}_{0}\right)^2\right] = \mathbb{E}\left[\left(\frac{d}{d\theta}\ln L(\mathbf{X}\,|\,\theta)\right)^2\right] = I_n(\theta)$$

Покажем, что матожидание в равенстве выше равно 0:

$$\int L(\mathbf{x}\,|\,\theta)\,d\mathbf{x} = 1 \qquad \Longrightarrow \qquad 0 = \frac{d}{d\theta}\int L(\mathbf{x}\,|\,\theta)\,d\mathbf{x} = \int \frac{d}{d\theta}L\,d\mathbf{x} = \int \frac{d\ln L}{d\theta}L\,d\mathbf{x} = \mathbb{E}\left[\frac{d}{d\theta}\ln L(\mathbf{X}\,|\,\theta)\right]$$

Левая часть, с учетом того, что $T(\mathbf{X})$ — несмещенная эффективная оценка:

$$\mathbb{V}\left[a(\theta)\cdot\big(T(\mathbf{X})-\theta\big)\right]=a^2(\theta)\cdot\mathbb{V}\big[T(\mathbf{X})\big]=a^2(\theta)\cdot I_n^{-1}(\theta)$$

Отсюда получаем, что

$$I_n(\theta)=a^2(\theta)\cdot I_n^{-1}(\theta)\qquad\Longrightarrow\qquad a(\theta)=I_n(\theta)$$

Задача 2

Пусть $\mathbf{X}=(X_1,\ldots,X_n)$, $X\sim\operatorname{Exp}(\theta)$.

(a) Вычислить информацию по Фишеру $I_n(\theta)$.

(b) Вычислить матожидание и дисперсию для

$$\widehat{\theta}(\mathbf{X})=\frac{n-1}{\sum_{i=1}^n X_i}$$

(c) Показать, что $\widehat{\theta}(\mathbf{X})$ не является эффективной оценкой.

Решение:

(a) Обозначим для простоты

$$S_n=S_n(\mathbf{X})=\sum_{i=1}^n X_i$$

Функция правдоподобия:

$$L(\mathbf{X}\,|\,\theta)=\prod_{i=1}^n\theta e^{-\theta X_i}=\theta^n e^{-\theta S_n}$$

$$\ln L(\mathbf{X}\,|\,\theta)=n\ln\theta-\theta S_n,\qquad\frac{d^2}{d\theta^2}\ln L(\mathbf{X}\,|\,\theta)=-\frac{n}{\theta^2}$$

Информация по Фишеру:

$$I_n(\theta)=-\mathbb{E}\left[\frac{d^2}{d\theta^2}\ln L(\mathbf{X}\,|\,\theta)\right]=\frac{n}{\theta^2}$$

(b) Для подсчета матожидания будем использовать тот факт, что

$$2\theta\sum_{i=1}^n X_i=2\theta S_n\sim\chi^2(2n),\qquad f_{\chi^2(2n)}(x)=\frac{x^{n-1}e^{-x/2}}{2^n(n-1)!},\quad x\geq 0$$

Тогда

$$\mathbb{E}\,\widehat{\theta}=2\theta(n-1)\cdot\mathbb{E}\left[\frac{1}{2\theta S_n}\right]=2\theta(n-1)\int\limits_0^{+\infty}\frac{1}{x}\cdot\frac{x^{n-1}e^{-x/2}}{2^n(n-1)!}\,dx=\frac{2\theta(n-1)}{2(n-1)}\underbrace{\int\limits_0^{+\infty}\underbrace{\frac{x^{n-2}e^{-x/2}}{2^{n-1}(n-2)!}}_{f_{\chi^2(2(n-1))}}\,dx}_{1}=\theta,$$

то есть $\widehat{\theta}(\mathbf{X})$ — несмещенная оценка.

$$\mathbb{V}\,\widehat{\theta}=\mathbb{E}\,\widehat{\theta}^2-\big(\mathbb{E}\,\widehat{\theta}\big)^2=\mathbb{E}\left[\frac{(n-1)^2}{S_n^2}\right]-\theta^2=4\theta^2(n-1)^2\mathbb{E}\left[\frac{1}{(2\theta S_n)^2}\right]-\theta^2=$$

$$=4\theta^2(n-1)^2\int\limits_0^{+\infty}\frac{1}{x^2}\cdot\frac{x^{n-1}e^{-x/2}}{2^n(n-1)!}\,dx-\theta^2=\frac{4\theta^2(n-1)^2}{4(n-1)(n-2)}\underbrace{\int\limits_0^{+\infty}f_{\chi^2(2(n-2))}(x)\,dx}_{1}-\theta^2=$$

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

(c) Данное семейство распределений является регулярным, поэтому можно записать неравенство Рао-Крамера:

$$\mathbb{V}\,\widehat{\theta} \geq \frac{1}{I_n^{-1}} = \frac{\theta^2}{n}$$

У нас равенства нет, поэтому оценка не является эффективной.
Можно было доказать неэффективность этой оценки, показав, что $\frac{d}{d\theta}\ln L(\mathbf{X}\,|\,\theta)$ и $\left(\widehat{\theta}(\mathbf{X}) - \theta\right)$ не пропорциональны:

$$\frac{d}{d\theta}\ln L(\mathbf{X}\,|\,\theta) = \frac{n}{\theta} - S_n, \qquad \widehat{\theta} - \theta = \frac{n-1}{S_n} - \theta$$