

Математическая статистика.

Андрей Тищенко @AndrewTGk

2024/2025

Семинар 10 января

Задача 1

$x_1, \dots, x_n \sim F_\xi(x)$, найти функцию распределения для $X_{(n)}, X_{(1)}$
 $F_{X_{(n)}}(x) = P(X_{(n)} \leq x) = P(X_{(1)} \leq x, \dots, X_{(n)} \leq x) = P(X_1 \leq x, \dots, X_n \leq x) =$
 $= P(X_1 \leq x) \dots P(X_n \leq x) = (F_\xi(x))^n$
 $F_{X_{(1)}}(x) = P(X_{(1)} \leq x) = 1 - P(X_{(1)} > x) = 1 - P(X_1 > x, \dots, X_n > x) =$
 $= 1 - P(X_1 > x, \dots, X_n > x) = 1 - P(X_1 > x) \dots P(X_n > x) = 1 - (1 - F_\xi(x))^n$

Задача 2

$x_1, \dots, x_n \sim R(0, 1)$. Найти $EX_{(n)}, EX_{(1)}$.

$$F_{X_{(n)}}(x) = (F_\xi(x))^n$$
$$f_{X_{(n)}}(x) = (F_{X_{(n)}}(x))' = n(F_\xi(x))^{n-1} \cdot f_\xi(x)$$
$$F_\xi(x) = \begin{cases} 0, & x < 0 \\ x, & x \in [0, 1] \\ 1, & x > 1 \end{cases}$$

Подставим в предыдущее уравнение:

$$f_{X_{(n)}} = \begin{cases} 0, & x < 0 \\ nx^{n-1}, & x \in [0, 1] \\ 0, & x > 1 \end{cases}$$

$$EX_{(n)} = \int_{-\infty}^{+\infty} x f_{X_{(n)}}(x) dx = \int_0^1 x n x^{n-1} dx = n \int_0^1 x^n dx = \frac{n}{n+1}$$

Посчитаем для $X_{(1)}$:

$$F_{X_{(1)}}(x) = 1 - (1 - F_\xi(x))^n$$

$$f_{X_{(1)}}(x) = (F_{X_{(1)}}(x))' = n(1 - F_\xi(x))^{n-1} (F_\xi(x))' = n(1 - F_\xi(x))^{n-1} f_\xi(x) = \begin{cases} 0, & x < 0 \\ n(1-x)^{n-1}, & 0 \leq x \leq 1 \\ 0, & x > 1 \end{cases}$$

$$EX_{(1)} = \int_0^1 x n (1-x)^{n-1} dx = n \int_0^1 x (1-x)^{n-1} dx = \left\langle \begin{matrix} t = 1-x \\ x = 1-t \end{matrix} \right\rangle = -n \int_1^0 (1-t) t^{n-1} dt = n \int_0^1 (1-t) t^{n-1} dt =$$
$$= n \int_0^1 t^{n-1} dt - n \int_0^1 t^n dt = 1 - \frac{n}{n+1}$$

Задача 3

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$
$$E\bar{x} = E\left(\frac{1}{n} \sum_{i=1}^n x_i\right) = \frac{1}{n} \sum_{i=1}^n E(x_i) = Ex_i$$

$$\mathcal{D}(\bar{x}) = \mathcal{D}\left(\frac{1}{n} \sum_{i=1}^n x_i\right) = \frac{1}{n^2} \sum_{i=1}^n \mathcal{D}x_i = \frac{\mathcal{D}x_1}{n}$$

Посчитаем выборочную дисперсию:

$$S^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

$$ES^2 = E\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2\right) = \frac{1}{n} \sum_{i=1}^n E(x_i - \bar{x})^2 = \mathcal{D}(x_1 - \bar{x}) = \mathcal{D}(x_1) + \mathcal{D}(\bar{x}) - 2 \operatorname{cov}(x_1, \bar{x}) = \frac{(n+1)\mathcal{D}(x_1)}{n} - 2 \operatorname{cov}(x_1, \bar{x})$$

$$\operatorname{cov}(x_1, \bar{x}) = \operatorname{cov}\left(x_1, \frac{1}{n} \sum_{i=1}^n x_i\right) = \frac{1}{n} \operatorname{cov}(x_1, \sum_{i=1}^n x_i) = \frac{1}{n} \operatorname{cov}(x_1, x_1) = \frac{\mathcal{D}(x_1)}{n}$$

Тогда

$$ES^2 = \frac{(n+1)\mathcal{D}(x_1)}{n} - \frac{2\mathcal{D}(x_1)}{n} = \mathcal{D}(x_1) \left(1 - \frac{1}{n}\right)$$

Несмещённая выборочная дисперсия (её математическое ожидание равняется дисперсии x_1):

$$\tilde{S}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

Посчитаем дисперсию S^2 :

$$\begin{aligned} \mathcal{D}\left(x_1 - \frac{1}{n} \sum_{i=1}^n x_i\right) &= \mathcal{D}\left(\frac{(n-1)x_1}{n}\right) + \mathcal{D}\left(\frac{1}{n} \sum_{i=2}^n x_i\right) = \frac{(n-1)^2}{n^2} \mathcal{D}(x_1) + \frac{n-1}{n^2} \mathcal{D}(x_1) = \\ &= \mathcal{D}(x_1) \left(\frac{(n-1)(n-1+1)}{n^2}\right) = \mathcal{D}(x_1) \frac{n-1}{n} \end{aligned}$$

Семинар 17 января.

$$T(x_1, x_2, \dots, x_n) = \sqrt{\frac{\pi}{2}} \frac{1}{n} \sum_{i=1}^n |x_i - m|, x_i \sim N(m, \theta^2)$$

$$ET(x_1, x_2, \dots, x_n) = \sqrt{\frac{\pi}{2}} \frac{1}{n} \sum_{i=1}^n E|x_i - m| = \sqrt{\frac{\pi}{2}} E|x_1 - m| = \sqrt{\frac{\pi}{2}} \int_{-\infty}^{+\infty} |x - m| \frac{1}{\sqrt{2\pi}\theta} e^{-\frac{(x-m)^2}{2\theta^2}} dx$$

Заменим $\frac{x-m}{\theta}$ на y

$$\frac{\theta}{2} \int_{-\infty}^{+\infty} |y| \cdot e^{-\frac{y^2}{2}} dy = \theta \int_0^{+\infty} y \cdot e^{-\frac{y^2}{2}} dy = \theta(1 - 0) = \theta$$

$$\frac{1}{n} \sum_{i=1}^n \sqrt{\frac{\pi}{2}} |x_i - m| \xrightarrow[n \rightarrow +\infty]{\text{п. н.}} E \sqrt{\frac{\pi}{2}} |x_i - m|$$

Задача

$$X = (X_1, \dots, X_n), X_i \sim R(0, \theta)$$

$$\hat{\theta} = X_{(n)}, \text{ доказать } \lim_{n \rightarrow \infty} EX_{(n)} = \theta$$

$$F_{X_{(n)}}(x) = (F_{X_i}(x))^n = \left(\frac{x}{\theta}\right)^n$$

$$f_{X_{(n)}}(x) = \frac{dF_{X_{(n)}}}{dx} = \frac{nx^{n-1}}{\theta^n}$$

$$EX_{(n)} = \int_0^\theta \frac{nx^n}{\theta^n} dx = \frac{nx^{n+1}}{(n+1)\theta^n} \Big|_0^\theta = \frac{n}{n+1} \theta \xrightarrow[n \rightarrow \infty]{} \theta. \text{ То есть смещённая, но асимптотически несмещённая.}$$

Докажем состоятельность, хотим:

$$\forall \varepsilon > 0 \quad P(|\hat{\theta} - \theta| < \varepsilon) \xrightarrow[n \rightarrow \infty]{} 1$$

$$P(-\varepsilon < X_{(n)} - \theta < \varepsilon) = F_{X_{(n)}}(\varepsilon + \theta) - F_{X_{(n)}}(\theta - \varepsilon) = 1 - \left(\frac{\theta - \varepsilon}{\theta}\right)^n \xrightarrow[n \rightarrow \infty]{} 1$$

Задача

$$I_n(\theta) = E \left(\frac{\delta \ln f(x, \theta)}{\delta \theta} \right)^2, \quad I_n(\theta) = n I_1(\theta), \quad x_1, \dots, x_n \sim N(\theta, \sigma^2).$$

$$f(x, \theta) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\theta)^2}{2\sigma^2}}$$

$$\ln f(x, \theta) = \ln \left(\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\theta)^2}{2\sigma^2}} \right) = -\frac{(x-\theta)^2}{2\sigma^2} + \ln \frac{1}{\sqrt{2\pi}\sigma}$$

$$\frac{\delta \ln f(x, \theta)}{\delta \theta} = -\frac{2(x-\theta)}{2\sigma^2} \cdot (-1) = \frac{x-\theta}{\sigma^2}$$

$$E \left(\frac{x-\theta}{\sigma^2} \right)^2 = \frac{1}{\sigma^4} E(x-\theta)^2 = \frac{1}{\sigma^4} \sigma^2 = \frac{1}{\sigma^2} = I_1(\theta)$$

$$\mathcal{D}\hat{\theta} \geq \frac{1}{n I_1(\theta)} = \frac{\sigma^2}{n} = \mathcal{D}\bar{x}$$

Семинар 24 января

Задача 4 ДЗ

$$\hat{K}_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y}) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X} + Ex_1 - Ex_1)(y_i - \bar{Y} + Ey_1 - Ey_1)$$

$$E\hat{K}_{xy} = E \frac{1}{n} \sum_{i=1}^n ((x_i - Ex_1) - (\bar{X} - Ex_1)) ((y_i - Ey_1) - (\bar{Y} - Ey_1)) =$$

$$= E((x_i - Ex_1) - (\bar{X} - Ex_1)) \cdot ((y_i - Ey_1) - (\bar{Y} - Ey_1)) = E((x_1 - Ex_1)(y_1 - Ey_1) + (x_1 - Ex_1)(\bar{Y} - Ey_1) +$$

$$= \text{cov}(x, y) - \frac{1}{n} \text{cov}(x, y) - \frac{1}{n} \text{cov}(x, y) + \frac{1}{n} \text{cov}(x, y)$$

Задача 5 ДЗ

Решал у доски, всем gl.

Задача 1

$X_1, \dots, X_n \sim \Pi(\theta)$. Проверить, что оценка $\hat{\theta} = \bar{X}$ является R-эффективной.

$$E\hat{\theta} = E \frac{1}{n} \sum_{i=1}^n x_i = Ex_1 = \theta$$

$$\mathcal{D} \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n} \theta$$

$$P(\xi = x_1) = \frac{e^{-\theta} \theta^{x_1}}{x_1!}. \text{ Логарифмируем:}$$

$$\ln \frac{e^{-\theta} \theta^{x_1}}{x_1!} = -\theta + x_1 \ln \theta - \ln x_1!$$

Возьмём частную производную:

$$\frac{\delta(-\theta + x_1 \ln \theta - \ln x_1!)}{\delta \theta} = -1 + \frac{x_1}{\theta}$$

Возьмём матожидание квадрата этой величины:

$$E(-1 + \frac{x_1}{\theta})^2 = \frac{1}{\theta^2} E(x_1 - \theta)^2 = \frac{\mathcal{D}x_1}{\theta^2} = \frac{1}{\theta} \Rightarrow I_n(\theta) = \frac{n}{\theta}$$

Попробуем самостоятельно подогнать оценку:

$$U(x, \theta) = \sum_{i=1}^n -1 + \frac{x_i}{\theta} = \frac{1}{\theta} \sum_{i=1}^n (x_i - \theta) = \frac{1}{\theta} (-n\theta + \sum_{i=1}^n \frac{x_i}{n}) = \frac{n}{\theta} (\sum_{i=1}^n (\frac{x_i}{n}) - \theta)$$

$$\hat{\theta} - \theta = a(\theta)U(x, \theta) \Rightarrow a(\theta) = \frac{\theta}{n}, \quad \hat{\theta} = \sum_{i=1}^n \frac{x_i}{n}$$

ДЗ

Задача 1

$$X_1, \dots, X_n \sim N(\theta, \sigma^2) \Rightarrow \forall i = \overline{1, n} \quad f(x_i, \theta) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\theta)^2}{2\sigma^2}}$$

$$\ln f(x_i, \theta) = \ln \frac{1}{\sqrt{2\pi}\sigma} - \frac{(x-\theta)^2}{2\sigma^2} = \ln \frac{1}{\sqrt{2\pi}\theta} - \frac{x^2}{2\sigma^2} + \frac{\theta x}{\sigma^2} - \frac{\theta^2}{2\sigma^2} \Rightarrow \frac{\delta}{\delta \theta} f(x_i, \theta) = \frac{x}{\sigma} - \frac{\theta}{\sigma^2}$$

$$U(x, \theta) = \sum_{i=1}^n \left(\frac{x_i}{\sigma} - \frac{\theta}{\sigma^2} \right)$$

По критерию эффективности хотим:

$$\hat{\theta} - \theta = \alpha(x)U(x, \theta)$$

Преобразуем: $U(x, \theta) = \left(\sum_{i=1}^n \frac{x_i}{\sigma} \right) - \frac{n\theta}{\sigma^2} \Rightarrow \underbrace{\frac{\sigma^2}{n}}_{\alpha(\sigma)} U(x, \theta) = \underbrace{\left(\frac{1}{n} \sum_{i=1}^n \sigma x_i \right)}_{\hat{\theta}} - \theta$

Задача 2

$$X_1, \dots, X_n \sim N(m, \theta) \Rightarrow f(x_i, \theta) = \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{(x-m)^2}{2\theta}}$$

$$\ln f(x, \theta) = \ln \frac{1}{\sqrt{2\pi}} - \frac{1}{2} \ln \theta - \frac{(x-m)^2}{2\theta} \Rightarrow \frac{\delta}{\delta \theta} f(x, \theta) = -\frac{1}{2\theta} + \frac{(x-m)^2}{2\theta^2}$$

Применим критерий эффективности:

$$\begin{aligned} U(x, \theta) &= \sum_{i=1}^n \left(\frac{(x-m)^2}{2\theta^2} - \frac{1}{2\theta} \right) = \sum_{i=1}^n \left(\frac{(x-m)^2 - \theta}{2\theta^2} \right) = \frac{1}{2\theta^2} \sum_{i=1}^n ((x-m)^2 - \theta) = \\ &= \frac{1}{2\theta^2} \left(\sum_{i=1}^n ((x-m)^2) - n\theta \right) = \frac{n}{2\theta^2} \left(\frac{1}{n} \sum_{i=1}^n ((x-m)^2) - \theta \right) \Rightarrow \underbrace{\frac{2\theta^2}{n}}_{\alpha(\theta)} U(x, \theta) = \underbrace{\left(\frac{1}{n} \sum_{i=1}^n (x-m)^2 \right)}_{\hat{\theta}} - \theta \end{aligned}$$

Задача 3

$X_1, \dots, X_n \sim G(\theta) \Rightarrow Ex = \frac{1}{\theta}$. Проверить оценку $\hat{\theta} = \frac{1}{\bar{X}}$ на несмещённость.

Хотим $E\hat{\theta} = \theta$. Попробуем по определению:

$$E\hat{\theta} = E \frac{n}{\sum_{i=1}^n x_i} = nE \frac{1}{\sum_{i=1}^n x_i}?$$

Попробуем решить через функцию правдоподобия:

$$L(x_1, \dots, x_n, \theta) = \prod_{i=1}^n P(\xi = x_i, \theta) = \prod_{i=1}^n (1-\theta)^{x_i-1} \theta \approx f(x, \theta)$$

$$\ln f(x_i, \theta) = \ln ((1-\theta)^{x_i-1} \theta) = (x_i-1) \ln(1-\theta) + \ln \theta$$

$$\frac{\delta}{\delta \theta} \ln f(x, \theta) = \frac{1}{\theta} - \frac{x_i-1}{1-\theta} = \frac{1-\theta-\theta x_i+\theta}{\theta-\theta^2} = \frac{1-\theta x_i}{\theta-\theta^2}$$

Применим критерий эффективности:

$$U(x, \theta) = \sum_{i=1}^n \frac{1-\theta x_i}{\theta-\theta^2} = \frac{1}{\theta-\theta^2} \left(n - \theta \sum_{i=1}^n x_i \right) = \frac{n}{\theta-\theta^2} \left(1 - \frac{\theta}{n} \sum_{i=1}^n x_i \right) = \frac{n\bar{X}}{\theta-\theta^2} \left(\frac{1}{\bar{X}} - \theta \right)$$

Значит $\frac{1}{\bar{X}}$ является R-эффективной, то есть несмещённой.

Задача 4

$X_1, \dots, X_n \sim Bi(k, \theta)$. Показать, что $\hat{\theta} = \frac{\bar{X}}{k}$ R-эффективная.

Посчитаем функцию правдоподобия:

$$L(x_1, \dots, x_n, \theta) = \prod_{i=1}^n P(\xi = x_i, \theta) = \prod_{i=1}^n C_n^k \theta^{x_i} \cdot (1-\theta)^{k-x_i} \approx f(x_i, \theta)$$

$$\ln f(x_i, \theta) = \ln \frac{n!}{k!(n-k)!} + x_i \ln \theta + (k-x_i) \ln(1-\theta)$$

$$\frac{\delta}{\delta\theta} \ln f(x_i, \theta) = \frac{x_i}{\theta} + \frac{x_i - k}{1 - \theta} = \frac{x_i - \theta x_i + \theta x_i - \theta k}{\theta - \theta^2} = \frac{x_i - \theta k}{\theta - \theta^2}$$

Применим критерий эффективности:

$$U(x, \theta) = \sum_{i=1}^n \frac{x_i - \theta k}{\theta - \theta^2} = \frac{1}{\theta - \theta^2} \left(-n\theta k + \sum_{i=1}^n x_i \right) = \frac{nk}{\theta - \theta^2} \left(\frac{1}{nk} \sum_{i=1}^n (x_i) - \theta \right) = \frac{nk}{\theta - \theta^2} \left(\frac{\bar{X}}{k} - \theta \right)$$

Получается, что $\frac{\bar{X}}{k}$ является R-эффективной (тут также выполнена 5 задача, так как я применил критерий эффективности в доказательстве).