

Московский физико-технический институт

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА
I ЗАДАНИЕ

Автор: Яфаров Руслан,
Б13-202

весна 2024

1. Виды сходимости случайных векторов

1. Пусть X_1, \dots, X_n — независимые одинаково распределенные случайные величины с распределением $Exp(\alpha)$, $\alpha > 0$.

Рассмотрим статистику $Y = \frac{1}{n} \sum_{i=1}^n X_i$. Найдите такие константы $a(\alpha)$ и $\sigma^2(\alpha) > 0$, что выполнено

$$\sqrt{n}(Y \sin Y - a(\alpha)) \xrightarrow{d} \mathcal{N}(0, \sigma^2(\alpha)), \text{ при } n \rightarrow \infty$$

Решение. По закону больших чисел $Y \xrightarrow{d} \frac{1}{\alpha}$. По ЦПТ $\sqrt{n}(Y - \frac{1}{\alpha}) \xrightarrow{d} \mathcal{N}(0, \frac{1}{\alpha^2})$. Воспользуемся теоремой 1.4 из С2. Положим $h(x) = x \sin x$, $b_n = \frac{1}{\sqrt{n}}$, $\xi_n = \sqrt{n}(Y - \frac{1}{\alpha})$, $a = \frac{1}{\alpha}$. Тогда получим, что $\frac{h(a+\xi_n b_n) - h(a)}{b_n} = \sqrt{n}(Y \sin Y - \frac{1}{\alpha} \sin \frac{1}{\alpha}) \xrightarrow{d} h'(a) \mathcal{N}(0, \frac{1}{\alpha^2}) \sim (\sin \frac{1}{\alpha} + \frac{1}{\alpha} \cos \frac{1}{\alpha}) \mathcal{N}(0, \frac{1}{\alpha^2}) \sim \mathcal{N}(0, (\frac{1}{\alpha} \sin \frac{1}{\alpha} + \frac{1}{\alpha^2} \cos \frac{1}{\alpha})^2) \Rightarrow a(\alpha) = \frac{1}{\alpha} \sin \frac{1}{\alpha}, \sigma(\alpha) = \frac{1}{\alpha} \sin \frac{1}{\alpha} + \frac{1}{\alpha^2} \cos \frac{1}{\alpha}$

2. Пусть $\{\xi_n\}_{n=1}^\infty, \{\eta_n\}_{n=1}^\infty, \{\zeta_n\}_{n=1}^\infty$ — последовательности случайных величин. Докажите, что если $\xi_n \xrightarrow{d} \xi, |\xi_n - \eta_n| \leq \zeta_n |\xi_n|, \zeta_n \xrightarrow{P} 0$, то $\eta_n \xrightarrow{d} \xi$

Решение. **TODO**

3. Задан набор независимых одинаково распределённых случайных величин X_1, \dots, X_n с распределением $\mathcal{N}(0, \sigma^2)$. Рассмотрим статистики $Y = \frac{1}{n} \sum_{i=1}^n |X_i|, Z = \frac{1}{n} \sum_{i=1}^n X_i^2, T = \sqrt{\frac{2}{\pi}} Z/Y$. Найдите предел сходимости по распределению выражения $\sqrt{n}(T - \sigma)$

Решение. Найдём необходимые моменты и ковариации:

$$\begin{aligned} \mathbb{E}|X_i| &= \int_{\mathbb{R}} |x| p_{X_i}(x) dx = 2 \int_0^{+\infty} x \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} dx = \frac{2}{\sqrt{2\pi\sigma^2}} \int_0^{+\infty} -\frac{\sigma^2}{x} x d(e^{-\frac{x^2}{2\sigma^2}}) = \sigma \sqrt{\frac{2}{\pi}} e^{-\frac{x^2}{2\sigma^2}} \Big|_0^{+\infty} \\ &= \sigma \sqrt{\frac{2}{\pi}} \end{aligned}$$

$$\mathbb{E}X_i^2 = \mathbb{D}X_i + (\mathbb{E}X_i)^2 = \sigma^2$$

$$\mathbb{E}|X_i|X_i^2 = 2 \int_0^{+\infty} x^3 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} dx = -2\sigma^2 \left(x^2 p_{X_i}(x) \Big|_0^{+\infty} - 2 \int_0^{+\infty} x p_{X_i}(x) dx \right) = 2\sigma^2 \mathbb{E}|X_i| = 2\sigma^3 \sqrt{\frac{2}{\pi}}$$

$$\mathbb{E}X_i^4 = -\sigma^2 \left(0 - \int_{\mathbb{R}} p_{X_i}(x) 3x^2 dx \right) = 3\sigma^2 \mathbb{E}X_i^2 = 3\sigma^4$$

$$\mathbb{D}|X_i| = \sigma^2 - \sigma^2 \frac{2}{\pi} = \sigma^2 \frac{\pi - 2}{\pi}, \mathbb{D}X_i^2 = \mathbb{E}X_i^4 - (\mathbb{E}X_i^2)^2 = 3\sigma^4 - \sigma^4 = 2\sigma^4$$

$$\text{cov}(|X_i|, X_i^2) = \mathbb{E}|X_i|X_i^2 - \mathbb{E}|X_i| \mathbb{E}X_i^2 = 2\sigma^3 \sqrt{\frac{2}{\pi}} - \sigma \sqrt{\frac{2}{\pi}} \sigma^2 = \sigma^3 \sqrt{\frac{2}{\pi}}$$

Положим $\eta_i = (|X_i|, X_i^2)^T$, тогда $\mathbb{D}\eta_i = \begin{pmatrix} \sigma^2 \frac{\pi - 2}{\pi} & \sigma^3 \sqrt{\frac{2}{\pi}} \\ \sigma^3 \sqrt{\frac{2}{\pi}} & 2\sigma^4 \end{pmatrix}$ Далее воспользуемся теоремой

1.4 для $\xi_n = \sum_{i=1}^n \eta_i$, $h(x, y) = \sqrt{\frac{2}{\pi}} \frac{y}{x}$, $b_n = 1/\sqrt{n}$, $a = \mathbb{E}\eta_1 = (\sigma\sqrt{\frac{2}{\pi}}, \sigma^2)^T$. Получим

$$\sqrt{n}(T - \sigma) \xrightarrow{d} (\nabla h|_a, \xi) \sim \mathcal{N}(0, \nabla h|_a^T \mathbb{D}\eta \nabla h|_a)$$

$$\nabla h = (-\sqrt{\frac{2}{\pi}} \frac{y}{x^2}, \sqrt{\frac{2}{\pi}} \frac{1}{x})^T \Rightarrow \nabla h|_a = (-\sqrt{\frac{\pi}{2}}, \frac{1}{\sigma})^T \Rightarrow \nabla h|_a^T \mathbb{D}\eta \nabla h|_a = \sigma^2 \left(\frac{\pi}{2} - 1 \right)$$

4. Пусть ξ, ξ_1, ξ_2, \dots - такие случайные величины, что $(\xi_n - \xi)^2 \xrightarrow{P} 0$ при $n \rightarrow \infty$. Показать, что $\xi_n^2 \xrightarrow{P} \xi^2$ при $n \rightarrow \infty$

Решение. $(\xi_n - \xi)^2 \xrightarrow{P} 0 \Leftrightarrow P(|\xi_n - \xi|^2 > \varepsilon^2) = P(|\xi_n - \xi| > \varepsilon) \rightarrow 0 \Rightarrow \xi_n \xrightarrow{P} \xi$. Далее пользуемся теоремой о наследовании сходимости для $h(x) = x^2$

5. Пусть ξ, ξ_1, ξ_2, \dots - случайные величины. Привести пример, когда

1. $\xi_n \xrightarrow{L_2} \xi$, $\xi_n \not\xrightarrow{\text{п.н.}} \xi$, $n \rightarrow \infty$

2. $\xi_n \xrightarrow{\text{п.н.}} \xi$, $\xi_n \not\xrightarrow{L_2} \xi$, $n \rightarrow \infty$

3. $\xi_n \xrightarrow{d} \xi$, $\xi_n \not\xrightarrow{P} \xi$, $n \rightarrow \infty$

Решение. 1. **TODO**

2. **TODO**

3. Пусть $\xi \sim \mathcal{N}(0, 1)$, $\xi_n = \xi$, тогда $\xi_n \xrightarrow{d} -\xi$, но $P(|\xi_n + \xi| > \varepsilon) = P(|\xi| > \varepsilon/2) \not\rightarrow 0$

6. Рассмотрим последовательность d -мерных случайных векторов $\bar{\xi}_n$. Доказать, что если при некотором $\bar{c} \in \mathbb{R}^d$ выполнено соотношение $\bar{\xi}_n \xrightarrow{d} \bar{c}$, то $\bar{\xi}_n \xrightarrow{P} \bar{c}$

Решение. $\bar{\xi}_n \xrightarrow{d} \bar{c} \Rightarrow \xi_n^i \xrightarrow{d} c^i \forall i$. Тогда, если докажем, что $\xi \in \mathbb{R}, \xi \xrightarrow{d} c \in \mathbb{R} \Rightarrow \xi \xrightarrow{P} c$, то утверждение будет доказано, т. к. покомпонентная сх-ть по в-ти влечет сх-ть вектора. Пусть $\xi_n \in \mathbb{R}, \xi_n \xrightarrow{d} c \in \mathbb{R}$. Тогда $P(|\xi_n - c| < \varepsilon) = P(c - \varepsilon < \xi_n < c + \varepsilon) = F_{\xi_n}(c + \varepsilon) - F_{\xi_n}(c - \varepsilon + 0) \geq F_{\xi_n}(c + \varepsilon) - F(c - \varepsilon/2) \rightarrow 1 \Rightarrow \xi_n \xrightarrow{P} c$

2. Статистики и оценки. Построение и сравнение оценок

7. Пусть X_1, \dots, X_n — выборка из распределения $R(0, \theta)$ (равномерного распределения на отрезке $[0, \theta]$). Проверьте на несмещенность, состоятельность и сильную состоятельность следующие оценки параметра θ : $2\bar{X}$, $\bar{X} + X_{(n)}/2$, $(n+1)X_{(1)}$, $X_{(1)} + X_{(n)}$, $\frac{n+1}{n}X_{(n)}$.

Решение.

1. $\mathbb{E}(2\bar{X}) = 2\mathbb{E}X_1 = 2\frac{\theta-0}{2} = \theta \Rightarrow$ оценка несмещенна.

По УЗБЧ $2\bar{X} \xrightarrow{\text{п. н.}} \theta \Rightarrow$ оценка сильно состоятельна.

2. Пусть $X_{(n)}/2 = \xi_n$. Найдем $p_{\xi_n} : F_{\xi_n}(x) = F_{X_1}^n(2x)$. $p_{\xi_n}(x) = \frac{d}{dx}F_{\xi_n}(x) = nF_{X_1}^{n-1}(2x)p_{X_1}(2x) \cdot 2$

$$2 = n2^n \frac{x^{n-1}}{\theta^n} I_{[0, \frac{\theta}{2}]}(x). \mathbb{E}\xi_n = \int_0^{\theta/2} xn2^n \frac{x^{n-1}}{\theta^n} dx = n2^n \frac{x^{n+1}}{\theta^n(n+1)} \Big|_0^{\theta/2} = \frac{n}{2n+2}\theta \Rightarrow \mathbb{E}(\bar{X} + X_{(n)}/2) = \frac{\theta}{2} \left(1 + \frac{n}{n+1}\right) \neq \theta \Rightarrow \text{оценка является смещенной}$$

$$F_{\xi_n}(x) = \begin{cases} 1, x \geq \theta/2, \\ \left(\frac{2x}{\theta}\right)^n, 0 < x < \theta/2 \\ 0, x \leq 0 \end{cases} \quad , \text{ тогда при } n \rightarrow \infty \text{ получаем } F(x) = \begin{cases} 1, x \geq \theta/2, \\ 0, x < \theta/2 \end{cases} \Rightarrow$$

$\xi_n \xrightarrow{d} \frac{\theta}{2} \Leftrightarrow \xi_n \xrightarrow{P} \frac{\theta}{2} \Rightarrow$ по теореме о наследовании сходимости получаем, что $\bar{X} + \xi_n \xrightarrow{P} \theta \Rightarrow$ оценка состоятельная

Докажем, что $X_{(n)} \xrightarrow{\text{п. н.}} \theta$. Заф. $\omega \in \Omega$, тогда последовательность $\{X_{(n)}(\omega)\}$ является неубывающей и ограниченной сверху $\Rightarrow \exists \xi(\omega) : \lim_{n \rightarrow \infty} X_{(n)}(\omega) = \xi(\omega) \Rightarrow$ оценка сильно состоятельная, но $X_{(n)} \xrightarrow{P} \theta \Rightarrow \xi(\omega) = \theta$. Аналогично доказывается, что $X_{(1)} \xrightarrow{\text{п. н.}} 0$. Тогда по теореме о наследовании сходимости оценка является сильно состоятельной

3. Найдем $p_{(n+1)X_{(1)}} : F_{X_{(1)}}(x) = 1 - (1 - F_{X_1}(x))^n \Rightarrow F_{(n+1)X_{(1)}}(x) = 1 - (1 - F_{X_1}(\frac{x}{n+1}))^n \Rightarrow$

$$p_{(n+1)X_{(1)}}(x) = n \left(1 - \frac{x}{(n+1)\theta}\right)^{n-1} \frac{1}{\theta} I_{[0, (n+1)\theta]}(x)$$

$$\mathbb{E}X_{(1)} = \int_0^\theta xn \left(1 - \frac{x}{\theta}\right)^{n-1} \frac{1}{\theta} dx = \theta n \int_0^1 t(1-t)^{n-1} dt = \theta n B(2, n) = \theta n \frac{\Gamma(2)\Gamma(n)}{\Gamma(n+2)} = \theta n \frac{1!(n-1)!}{(n+1)!} = \frac{\theta}{n+1} \Rightarrow \text{оценка } X_{(1)}(n+1) \text{ несмещенная.}$$

$$F_{(n+1)X_{(1)}}(x) = \begin{cases} 1, x \geq (n+1)\theta, \\ 1 - \left(1 - \frac{x}{(n+1)\theta}\right)^n, 0 < x < (n+1)\theta \\ 0, x \leq 0 \end{cases} \quad \text{при } n \rightarrow \infty \text{ получим } F(x) =$$

$$\begin{cases} 1 - e^{-\frac{x}{\theta}}, x \geq 0, \\ 0, x < 0 \end{cases} \Rightarrow \text{оценка не состоятельна и сл-но не сильно состоятельна}$$

4. $\mathbb{E}(X_{(1)} + X_{(n)}) = \frac{\theta}{n+1} + \frac{\theta n}{n+1} = \theta \Rightarrow$ оценка не смещена. Т. к. $F_{X_{(1)}} \rightarrow F(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases}$,

то $X_{(1)} \xrightarrow{P} 0 \Rightarrow$ по теореме о наследовании сходимости оценка является сильно состоятельной.

5. Оценка является несмещенной и сильно состоятельной (по модулю предыдущих выкладок это очев).

8. Пусть $\hat{\theta}_n(X)$ — асимптотически нормальная оценка параметра θ с асимптотической дисперсией $\sigma^2(\theta)$. Докажите, что тогда $\hat{\theta}_n(X)$ является состоятельной оценкой θ .

Решение. Пусть $\xi_n = \frac{1}{\sqrt{n}} \cdot \xi_n \xrightarrow{P} 0$ По теореме о наследовании сходимости $\xi_n \sqrt{n}(\hat{\theta}_n(X) - \theta) \xrightarrow{P} 0 \Rightarrow \hat{\theta}_n(X) \xrightarrow{P} \theta$

9. Пусть X_1, \dots, X_n - выборка из распределения с параметром σ^2 . Пусть, кроме того $D_{\sigma^2} X_1 = \sigma^2$. Докажите, что статистика $s = 1/n \sum_{i=1}^n (X_i - \bar{X})^2$ равна $\overline{X^2} - (\bar{X})^2$ и является состоятельной оценкой σ^2 . Является ли она несмещенной оценкой того же параметра?

Решение. $1/n \sum_{i=1}^n (X_i - \bar{X})^2 = 1/n \sum_{i=1}^n (X_i^2 - 2X_i\bar{X} + (\bar{X})^2) = 1/n \sum_{i=1}^n X_i^2 - 2\bar{X} 1/n \sum_{i=1}^n X_i - (\bar{X})^2 = \overline{X^2} - 2(\bar{X})^2 + (\bar{X})^2 = \overline{X^2} - (\bar{X})^2$. Подставляя в равенство вместо X_i $X_i - \mathbb{E}X_i$ мы получим, что данная оценка равна $\overline{(X - \mathbb{E}X_1)^2} - (\bar{X} - \mathbb{E}X_1)^2 \xrightarrow{п.н.} \mathbb{D}X_1$ по УЗБЧ. Но $\mathbb{E}s = \mathbb{D}X_1 - \mathbb{D}\bar{X} = \mathbb{D}X_1 - \frac{1}{n}\mathbb{D}X_1 \neq \mathbb{D}X_1 \Rightarrow$ оценка смещена

10. Пусть X_1, \dots, X_n — выборка из экспоненциального распределения с параметром θ . Покажите, что $\forall k \in \mathbb{N}$ статистика $\sqrt[k]{k!/\bar{X}^k}$ является асимптотически нормальной оценкой параметра θ . Найдите ее асимптотическую дисперсию.

Решение.

$$\psi_{X_1}(t) = \mathbb{E}e^{itX_1} = \int_0^{+\infty} \theta e^{-\theta x} e^{itx} dx = \int_0^{+\infty} \theta e^{x(it-\theta)} dx = -\frac{\theta}{it-\theta} = \left(1 - \frac{it}{\theta}\right)^{-1}$$

$$\frac{d^k}{dt^k} \psi_{X_1}(t) = \frac{i^k}{\theta^k} k! \left(1 - \frac{it}{\theta}\right)^{-k-1} \Rightarrow \mathbb{E}X_1^k = \frac{k!}{\theta^k}$$

Пусть $\xi_n = \sqrt{n} \left(\bar{X}^k - \frac{k!}{\theta^k}\right)$. Тогда по ЦПТ $\xi_n \xrightarrow{d} \mathcal{N}(0, \mathbb{D}X^k)$, $\mathbb{D}X^k = \frac{(2k)!}{\theta^{2k}} - \frac{(k!)^2}{\theta^{2k}}$. Пользуясь теоремой 1.4 из С2 для $h(x) = \sqrt[k]{k!/x}$, $b_n = 1/\sqrt{n}$ и $a = \frac{k!}{\theta^k}$. $h'(x) = (\sqrt[k]{k!} x^{-\frac{1}{k}})' = -\frac{1}{k} \sqrt[k]{\frac{k!}{x^{k+1}}} \Rightarrow h'(a) = -\frac{\theta^{k+1}}{k!k} \Rightarrow \sqrt{n} \left(\sqrt[k]{k!/\bar{X}^k} - \theta\right) \xrightarrow{d} \mathcal{N}(0, \sigma^2(\theta))$, где $\sigma^2(\theta) = \mathbb{D}X^k h'(a)^2 = \theta^2 \frac{(2k)! - (k!)^2}{(k!k)^2}$

11. Пусть X_1, \dots, X_n - выборка из распределения с плотностью

$$p_{\alpha, \beta}(x) = \frac{1}{\alpha} e^{(\beta-x)/\alpha} I_{[\beta, +\infty)}(x)$$

где $\theta = (\alpha, \beta)$ — двумерный параметр. Найдите для θ оценку максимального правдоподобия. Докажите, что полученная для α оценка $\hat{\alpha}_n$ является асимптотически нормальной, и найдите ее асимптотическую дисперсию.

Решение. Пусть $\theta = (\alpha, \beta)$

$$\mathcal{L}(x, \theta) = \frac{1}{\alpha^n} e^{\sum_{i=1}^n \frac{\beta-x_i}{\alpha}} I_{[\beta, +\infty)}(x_1, \dots, x_n) = \frac{1}{\alpha^n} e^{\sum_{i=1}^n \frac{\beta-x_i}{\alpha}} I_{[\beta, +\infty)}(\min(x_1, \dots, x_n))$$

Для того, чтобы произведение было не 0, должно выполняться $\beta \leq \min(x_1, \dots, x_n)$, в то же время $\mathcal{L}(x, \theta) \xrightarrow{\beta} \max_{\beta} \Leftrightarrow \sum_{i=1}^n \beta - x_i \xrightarrow{\beta} \max \Rightarrow \hat{\beta} = X_{(1)}$. $l(x, \theta) = \sum_{i=1}^n \frac{\beta-x_i}{\alpha} - n \ln \alpha \Rightarrow \hat{\alpha}_n = \sum_{i=1}^n \frac{X_i - \beta}{n} = \bar{X} - X_{(1)}$ в силу того, что функция $f(x) = \frac{c}{x} - \ln x$, $c \leq 0$ имеет глобальный максимум в т. $x = -c$

TODO Доказать асимптотическую нормальность

12. Найдите оценку максимального правдоподобия для параметра сдвига в распределении Коши, т.е. плотность равна

$$p_{\theta}(x) = \frac{1}{\pi(1 + (x - \theta)^2)}$$

если выборка состоит из а) одного наблюдения, б) двух наблюдений (т.е. $n = 1, 2$).

Решение.

а $\pi \mathcal{L}(x, \theta) = \frac{1}{1+(x-\theta)^2} \Rightarrow \hat{\theta} = X_1$

б $l(x, \theta) = -\ln(1 + (x_1 - \theta)^2) - \ln(1 + (x_2 - \theta)^2) - 2 \ln \pi \Rightarrow \frac{dl}{d\theta} = \frac{2(x_1+x_2-2\theta)(\theta^2-\theta x_1-\theta x_2+x_1 x_2+1)}{(1+(x_1-\theta)^2)(1+(x_2-\theta)^2)}.$

Получаем 2 случая:

(а) $|x_1 - x_2| \leq 2$

Тогда уравнение правдоподобия имеет единственный корень $\frac{x_1+x_2}{2}$. Легко заметить, что в этой точке достигается максимум всей функции

(б) $|x_1 - x_2| > 2$. Тогда уравнение имеет 3 корня и максимум достигается в одной из точек $\frac{x_1+x_2}{2} + \sqrt{\left(\frac{x_1-x_2}{2}\right)^2 - 1}, \frac{x_1+x_2}{2} - \sqrt{\left(\frac{x_1-x_2}{2}\right)^2 - 1}$. Легко убедиться, что значение функции правдоподобия совпадают на них

Тогда оценка максимального правдоподобия $\hat{\theta} = \begin{cases} \frac{X_1+X_2}{2}, |X_1 - X_2| \leq 2 \\ \frac{X_1+X_2}{2} + \sqrt{\left(\frac{X_1-X_2}{2}\right)^2 - 1}, |X_1 - X_2| > 2 \end{cases}$

13. Пусть $X_1 \sim R(0, \theta)$. Найдите несмещённую оценку параметра $1/\theta$.

Решение. Для $n = 1$ $\mathbb{E}g(X_1) = \int_0^\theta \frac{1}{\theta} g(x) dx = \frac{1}{\theta} \Rightarrow \forall \theta \int_0^\theta g(x) dx = 1$, что невозможно.

Для $n \geq 2$ пусть $\hat{\theta} = \frac{1}{4\sqrt{X_1 X_2}}$. Тогда

$$\mathbb{E}\hat{\theta} = \int_{[0,\theta]^2} \frac{1}{\theta^2} \frac{1}{2\sqrt{x_1}} \frac{1}{2\sqrt{x_2}} dx_1 dx_2 = \frac{1}{\theta}$$

14. Найдите несмещённую оценку λ^3 по выборке X_1, \dots, X_n из распределения $Pois(\lambda)$.

Решение. $\varphi_{X_1}(t) = \mathbb{E}e^{tX_1} = e^{\lambda(e^t-1)} \Rightarrow \frac{d}{dt}\varphi_{X_1}(t) = \lambda e^{t+\lambda(e^t-1)}, \frac{d^2}{dt^2}\varphi_{X_1}(t) = \lambda e^{t+\lambda(e^t-1)} + \lambda^2 e^{2t+\lambda(e^t-1)}, \frac{d^3}{dt^3}\varphi_{X_1}(t) = \lambda e^{t+\lambda(e^t-1)} + \lambda^2 e^{2t+\lambda(e^t-1)} + \lambda^2(2 + \lambda e^t)e^{2t+\lambda(e^t-1)} \Rightarrow \mathbb{E}X_1 = \lambda, \mathbb{E}X_1^2 = \lambda + \lambda^2, \mathbb{E}X_1^3 = \lambda^3 + 3\lambda^2 + \lambda \Rightarrow \hat{\lambda} = X_1^3 - 3X_1^2 + 2X_1$

15. Пусть X_1, \dots, X_n — выборка из равномерного распределения на отрезке $[0, \theta]$. Сравните следующие оценки параметра θ в равномерном подходе с квадратичной функцией потерь: $\theta : 2\bar{X}, (n+1)X_{(1)}, \frac{n+1}{n}X_{(n)}$.

Решение. По задаче 7 все оценки несмещены \Rightarrow нужно сравнить дисперсии этих случайных величин.

$$1. \mathbb{D}(2\bar{X}) = \frac{4}{n}\mathbb{D}X_1 = \frac{\theta^2}{3n}.$$

$$2. \mathbb{D}(n+1)X_{(1)} = (n+1)^2\mathbb{D}X_{(1)} \cdot \mathbb{E}X_{(1)}^2 = \int_0^\theta x^2 n \left(1 - \frac{x}{\theta}\right)^{n-1} \frac{1}{\theta} dx = n\theta^2 \int_0^1 t^2 (1-t)^{n-1} dt = \\ \theta^2 n B(3, n) = \theta^2 n \frac{2!(n-1)!}{(n+2)!} = \frac{2\theta^2}{(n+1)(n+2)} \Rightarrow \mathbb{D}X_1 = \frac{2\theta^2}{(n+1)(n+2)} - \frac{\theta^2}{(n+1)^2} = \frac{n\theta^2}{(n+1)^2(n+2)} \Rightarrow \\ \mathbb{D}(n+1)X_{(1)} = \theta^2 \left(1 - \frac{2}{n+2}\right)$$

$$3. \mathbb{E}X_{(n)}^2 = \int_0^\theta nx^2 \frac{x^{n-1}}{\theta^n} dx = \theta^2 \frac{n}{n+2} \Rightarrow \mathbb{D}X_{(n)} = \theta^2 n \left(\frac{1}{(n+2)} - \frac{n}{(n+1)^2} \right) = \theta^2 \frac{n}{(n+1)^2(n+2)} \Rightarrow \\ \mathbb{D} \frac{n+1}{n} X_{(n)} = \frac{\theta^2}{(n+1)(n+2)}$$

Таким образом (асимптотически), $3_{\text{оц}}$ лучше $1_{\text{оц}}$ лучше $2_{\text{оц}}$

16. Пусть $\theta_1^*(X)$ и $\theta_2^*(X)$ — две наилучшие в среднеквадратичном подходе оценки параметра θ в классе всех оценок с од- ним и тем же математическим ожиданием $\tau(\theta)$. Докажите, что тогда для любого θ они совпадают почти наверное.

Решение. $\mathbb{E}(\theta_1^*(X) - \theta)^2 = \mathbb{E}(\theta_2^*(X) - \theta)^2$, при этом $\mathbb{E}\theta_1^*(X) = \mathbb{E}\theta_2^*(X) \Rightarrow \mathbb{E}(\theta_1^*(X))^2 = \mathbb{E}(\theta_2^*(X))^2$. Предположим, $\theta_1^*(X)$ и $\theta_2^*(X)$ не совпадают почти наверное. Тогда $\mathbb{E}(\theta_1^*(X) - \theta_2^*(X))^2 > 0 \Rightarrow \mathbb{E}\theta_1^*(X)\theta_2^*(X) < \mathbb{E}(\theta_1^*(X))^2$. Тогда

$$\mathbb{E}(\theta_1^*(X) - \theta)^2 - \mathbb{E} \left(\frac{\theta_1^*(X) + \theta_2^*(X)}{2} - \theta \right)^2 = \mathbb{E}(\theta_1^*(X))^2 - \frac{\mathbb{E}(\theta_1^*(X))^2}{2} + \frac{\mathbb{E}\theta_1^*(X)\theta_2^*(X)}{2} = \\ \frac{1}{2}(\mathbb{E}(\theta_1^*(X))^2 - \mathbb{E}\theta_1^*(X)\theta_2^*(X)) > 0 \Rightarrow$$

оценка $\frac{\theta_1^*(X) + \theta_2^*(X)}{2}$ лучше в данном классе.

17. Пусть X_1, \dots, X_n - выборка из нормального распределения с параметрами (a, σ^2) . Найдите эффективную оценку а) параметра a , если σ известно; б) параметра σ^2 , если a известно. Вычислите информацию Фишера одного наблюдения в обоих случаях

Решение.

18. Пусть X_1, \dots, X_n — выборка из логистического распределения со сдвигом θ , имеющего плотность

$$p_\theta(x) = \frac{\exp\{\theta - x\}}{(1 + \exp\{\theta - x\})^2}$$

Найдите информацию Фишера $i(\theta)$ одного наблюдения в этой модели.

Решение.

$$u_\theta(x) = \frac{1 - e^{\theta-x}}{1 + e^{\theta-x}} \Rightarrow i(\theta) = \int_{\mathbb{R}} \frac{(1 - e^{\theta-x})^2}{(1 + e^{\theta-x})^2} \frac{e^{\theta-x}}{(1 + e^{\theta-x})^2} dx = \int_0^{+\infty} \frac{(1-t)^2}{(1+t)^4} dt =$$

$$\int_1^{+\infty} \frac{(2-u)^2}{u^4} du = \int_1^{+\infty} \frac{4 - 4u + u^2}{u^4} du = \left(4\frac{u^{-3}}{-3} - 4\frac{u^{-2}}{-2} + \frac{u^{-1}}{-1} \right) \Big|_1^{+\infty} = - \left(-\frac{4}{3} + 2 - 1 \right) = \frac{1}{3}$$

19. С помощью метода моментов построить оценку параметра θ для следующих распределений: а) $Bern(\theta)$, б) $Pois(\theta)$, в) $\mathcal{N}(\theta, 1)$, г) $\exp(\theta)$. Является ли данная оценка:

1. несмещенной?
2. состоятельной?
3. сильно состоятельной?
4. асимптотически нормальной?

Решение. В первых 3 случаях оценка $\hat{\theta} = \bar{X}$ является несмещенной, состоятельной и сильно состоятельной по УЗБЧ. По ЦПТ также она является асимптотически нормальной с асимптотической дисперсией равной дисперсии этой случайной величины, то есть

а $\mathbb{D}X_1 = \theta(1 - \theta)$

б $\mathbb{D}X_1 = \theta$

с $\mathbb{D}X_1 = 1$

д $\mathbb{E}X_1 = 1/\theta \Rightarrow \hat{\theta} = \frac{1}{\bar{X}}$. По методу моментов оценка сильно состоятельная. По задаче 10 оценка асимптотически нормальная с ас-й дисперсией θ^2 . Найдем $p_{\bar{X}}(x)$. В силу независимости $p_{X_1+X_2}(x) = \int_{\mathbb{R}} p_{X_1}(x-t)p_{X_2}(t)dt = \int_{\mathbb{R}} \theta e^{-\theta(x-t)} I_{x-t \geq 0} \theta e^{-\theta t} I_{t \geq 0} dt = \int_0^x \theta^2 e^{-\theta x} = \theta^2 x e^{-\theta x} I_{x \geq 0}$. По индукции получаем, что $p_{X_1+\dots+X_n}(x) = \theta^n \frac{x^{n-1}}{(n-1)!} e^{-\theta x} I_{x \geq 0}$. $p_{\xi/n}(x) = \frac{d}{dx} F_{\xi/n}(x) = \frac{d}{dx} F_{\xi}(nx) = np_{\xi}(nx) \Rightarrow p_{\bar{X}}(x) = n^n \theta^n \frac{x^{n-1}}{(n-1)!} e^{-\theta nx} I_{x \geq 0}$.

$$\mathbb{E}1/\bar{X} = \int_0^{+\infty} n^n \theta^n \frac{x^{n-2}}{(n-1)!} e^{-\theta nx} dx = -\frac{n^{n-1} \theta^{n-1}}{n-1} \left(\frac{x^{n-2}}{(n-2)!} e^{-\theta nx} \Big|_0^{+\infty} - \int_0^{+\infty} \frac{x^{n-3}}{(n-3)!} e^{-\theta nx} dx \right) =$$

$$\frac{n^{n-1} \theta^{n-1}}{n-1} \int_0^{+\infty} \frac{x^{n-3}}{(n-3)!} e^{-\theta nx} dx = \dots = \frac{(n\theta)^{n-(n-2)}}{n-1} \int_0^{+\infty} e^{-\theta nx} dx = \frac{n\theta}{n-1} \neq \theta \Rightarrow \text{оценка смещена}$$

20. Решить предыдущую задачу используя вместо метода моментов метод максимального правдоподобия.

Решение. Пусть $\sum_{i=1}^n x_i = a$

а $L(\theta) = \theta^a (1 - \theta)^{n-a} \Rightarrow l(\theta) = a \ln \theta + (n - a) \ln(1 - \theta) \Rightarrow \frac{d}{d\theta} l(\theta) = \frac{a}{\theta} - \frac{n-a}{1-\theta}$, причем $\frac{d^2}{d\theta^2} l(\theta) = \frac{-a}{\theta^2} - \frac{n-a}{(1-\theta)^2} < 0 \forall \theta \Rightarrow \theta = a/n$ точка строгого глобального максимума $\Rightarrow \hat{\theta} = \bar{X}$

б $L(\theta) = \frac{\theta^a e^{-n\theta}}{x_1! x_2! \dots x_n!} \rightarrow \max_{\theta} \Leftrightarrow H(\theta) = \theta^a e^{-n\theta} \rightarrow \max_{\theta} .h(\theta) = a \ln \theta - n\theta \Rightarrow \frac{d}{d\theta} h(\theta) = \frac{a}{\theta} - n \Rightarrow \hat{\theta} = \bar{X}$

с $L(\theta) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{\sum_{i=1}^n (x_i - \theta)^2}{2}} \rightarrow \max_{\theta} \Leftrightarrow \sum_{i=1}^n (x_i - \theta)^2 \rightarrow \min_{\theta} \Rightarrow \hat{\theta} = \bar{X}$

$$d L(\theta) = \theta^n e^{-\theta a} I_{x(1) \geq 0}. \text{ Опустим индикатор, он ни на что не влияет. } l(\theta) = n \ln \theta - \theta a \Rightarrow \\ \frac{d}{d\theta} l(\theta) = \frac{n}{\theta} - a \Rightarrow \hat{\theta} = \frac{1}{\bar{X}}$$

Во всех случаях получили те же оценки

21. Рассмотрим распределения Коши с плотностью $p_\theta(x) = \frac{1}{\pi(1+(x-\theta)^2)}$. С помощью выборочной медианы построить асимптотически нормальную оценку для θ^2 и найти ее асимптотическую дисперсию.

Решение. По теореме о выборочной медиане $\hat{\mu}$ является асимптотически нормальной оценкой параметра $z_{\frac{1}{2}} = \theta$ с асимптотической дисперсией $\frac{1}{4p_\theta(z_{\frac{1}{2}})} = \frac{\pi^2}{4}$. Ф-я $f(x) = x^2$ является дифференцируемой, по теореме о наследовании асимптотической нормальности получаем, что $\hat{\mu}^2$ является асимптотически нормальной оценкой параметра θ^2 с асимптотической дисперсией $(2\theta)^2 \frac{\pi^2}{4} = (\pi\theta)^2$

22. Пусть X_1, \dots, X_n — выборка из распределения: а) $Bern(\theta)$, б) $Pois(\theta)$, в) $\mathcal{N}(\theta, 1)$, г) $\exp(\theta)$. Для каких функций существует эффективная оценка? Найти соответствующую эффективную оценку и количество информации (фишеровской), содежащейся в одном сообщении.

Решение. Для всех распределений $\Theta = \mathbb{R}$, носитель не зависит от θ и все функции распределений “хорошие”, то есть выполняются первые 3 пункта условий регулярности.

$$\text{а } p_{X_1}(x) = \theta I_1(x) + (1 - \theta) I_0(x) \Rightarrow u_\theta(x) = \frac{I_1(x) - I_0(x)}{p_{X_1}(x)} \Rightarrow i(\theta) = \frac{1}{\theta} + \frac{1}{1-\theta} = \frac{1}{\theta(1-\theta)}$$

TODO

3. Достаточные статистики. Полные статистики. Оптимальные оценки.

23. Приведите пример такого параметрического семейства распределений P и нетривиальной неполной достаточной статистики $S(X_1, \dots, X_n)$, где X_1, \dots, X_n — выборка из неизвестного распределения $P \in \mathcal{P}$, что размерность статистики S равна 1.

Решение. Рассмотрим $X_1 \sim Bern(\theta)$. Правдоподобие выборки равно $L(\theta) = \theta^{\sum_{i=1}^n x_i} (1 - \theta)^{n - \sum_{i=1}^n x_i} \Rightarrow$ по критерию факторизации $S(X) = \sum_{i=1}^n X_i$ является достаточной статистикой.

24. Пусть X_1, \dots, X_n — выборка из нормального распределения с параметрами (a, σ^2) , $a \in \mathbb{R}$, $\sigma > 0$. Найдите оптимальную оценку параметра $\theta = (a, \sigma^2)$.

Решение. $L(\theta) = \frac{1}{(2\pi)^{n/2}\sigma^n} e^{-\frac{\sum_{i=1}^n (x_i - a)^2}{2\sigma^2}} \Rightarrow T(X) = (\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$ является полной достаточной статистикой по теореме об экспоненциальных семействах. Решая уравнение несмещённости, получаем, что $S(X) = (\bar{X}, \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2)$ оптимальная оценка.

25. Пусть X_1, \dots, X_n — выборка из нормального распределения с параметрами $(0, \theta^2)$. Найдите оптимальную оценку для θ .

Решение. Так как нормальное распределение из экспоненциального семейства, $T(X) = \sum_{i=1}^n X_i^2$ является полной достаточной статистикой. Найдём распределение $T(X)$.

$$\frac{T(X)}{\theta^2} \sim \chi^2(n), p_{\theta^2 \chi^2} = \frac{d}{dx} F_{\chi^2}(x/\theta^2) = 1/\theta^2 p_{\chi^2}(x/\theta^2)$$

$$\mathbb{E}\varphi(T) = \frac{1}{2^{n/2}\Gamma(n/2)} \int_0^{+\infty} \varphi(x) \left(\frac{x}{\theta^2}\right)^{n/2-1} e^{-\frac{x}{2\theta^2}} dx = \frac{1}{2^{n/2}\Gamma(n/2)} \int_0^{+\infty} \varphi(t\theta^2) t^{n/2-1} e^{-\frac{t}{2}} dt \Rightarrow$$

$$\text{положив } \varphi(x) = \sqrt{x} \text{ получаем } \frac{\theta}{2^{n/2}\Gamma(n/2)} \int_0^{+\infty} t^{(n+1)/2-1} e^{-\frac{t}{2}} dt = \frac{\theta 2^{(n+1)/2}}{2^{n/2}\Gamma(n/2)} \int_0^{+\infty} u^{(n+1)/2-1} e^{-u} du$$

$$= \frac{\theta \sqrt{2}\Gamma(\frac{n+1}{2})}{\Gamma(n/2)}. \text{ Окончательно получаем } S(X) = \frac{\Gamma(\frac{n}{2})}{\sqrt{2}\Gamma(\frac{n+1}{2})} \sqrt{\sum_{i=1}^n X_i^2}$$

26. Пусть X_1, \dots, X_n — выборка из пуассоновского распределения с параметром $\theta > 0$. Найдите $\mathbb{E}\left(X_1^2 \middle| \sum_{i=1}^n X_i\right)$.

Решение. $\sum_{i=1}^n X_i$ полная достаточная статистика для параметра θ . $\mathbb{E}X_1^2 = \mathbb{D}X_1 + (\mathbb{E}X_1)^2 = \theta + \theta^2 \Rightarrow$ матожидание равно оптимальной оценке $\theta + \theta^2$. $T(X) = \sum_{i=1}^n X_i \sim \text{Pois}(n\theta) \Rightarrow \mathbb{E}T = n\theta, \mathbb{E}T^2 = n\theta + n^2\theta^2 \Rightarrow S(T) = \frac{T^2 - T}{n^2} + \frac{T}{n}$

27. С помощью критерия факторизации найти достаточную статистику для следующего семейства распределений: а) $\text{Bern}(\theta)$, б) $\text{Pois}(\theta)$, в) $\mathcal{N}(\theta, 1)$, г) $\exp(\theta)$. Проверить, является ли полученная статистика полной.

Решение. Пусть $\sum_{i=1}^n x_i = a, T(X) = \sum_{i=1}^n X_i$

а $L(\theta) = \theta^a (1-\theta)^{n-a} \Rightarrow T(X)$ — достаточная статистика. $T \sim \text{Bin}(n, \theta) \mathbb{E}\varphi(T) = \sum_{k=1}^n \varphi(k) C_n^k \theta^k (1-\theta)^{n-k} = 0 \forall \theta \Leftrightarrow \sum_{k=1}^n \varphi(k) C_n^k z^k = 0 \forall z, \text{ где } z = \frac{\theta}{1-\theta} \Rightarrow \forall k \varphi(k) = 0 \Rightarrow \text{статистика полная.}$

б $L(\theta) = \frac{\theta^a e^{-n\theta}}{x_1! x_2! \dots x_n!} \Rightarrow T(X)$ — достаточная статистика. $T(X) \sim \text{Pois}(n\theta) \Rightarrow \mathbb{E}\varphi(T) = \sum_{k=0}^{\infty} \varphi(k) e^{-n\theta} \frac{(n\theta)^k}{k!} = 0 \forall \theta \frac{d^m}{d\theta^m} e^{n\theta} \mathbb{E}\varphi(T) \big|_{\theta=0} = \sum_{k=m}^{\infty} \varphi(k) n^k \frac{\theta^{k-m}}{(k-m)!} \big|_{\theta=0} = n^m \varphi(m) = 0 \Rightarrow \varphi(m) = 0 \Rightarrow \text{статистика является полной.}$

в $L(\theta) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{\sum_{i=1}^n (x_i - \theta)^2}{2}} = \frac{1}{(2\pi)^{n/2}} e^{-\frac{\sum_{i=1}^n x_i^2}{2}} e^{\theta a - \frac{n\theta^2}{2}} \Rightarrow T(X)$ достаточная статистика. $T(X) \sim \mathcal{N}(\theta, 1/n)$

$$\mathbb{E}\varphi(T) = \int_{\mathbb{R}} \sqrt{\frac{n}{2\pi}} \varphi(x) e^{-\frac{n}{2}(x-\theta)^2} dx$$

TODO

d $L(\theta) = \theta^n e^{-\theta a} I_{x_{(1)} \geq 0} \Rightarrow T(X)$ - достаточная.

28. Построить оптимальную оценку функции $\tau(\theta) = 5\theta^2 + 3\theta + 7$ для $Bern(\theta)$.

Решение. При $n = 1$ построить несмещенную оценку для θ^2 невозможно, далее полагаем $n \neq 1$. По предыдущей задаче известно, что статистика $T(X) = \sum_{i=1}^n X_i$ является полной достаточной. $T(X) \sim Bin(n, \theta) \Rightarrow \mathbb{E}T(X) = n\theta, \mathbb{D}T(X) = n(\theta - \theta^2) \Rightarrow \mathbb{E}T^2(X) = n\theta - n\theta^2 + n^2\theta^2 \Rightarrow \theta^2 = \frac{\mathbb{E}T^2 - \mathbb{E}T}{n^2 - n} \Rightarrow \varphi(T) = 5\frac{T^2 - T}{n^2 - n} + 3\frac{T}{n} + 7$

29. Построить оптимальную оценку функции $\tau(\theta) = \sqrt{\theta}$ для $\exp(\theta)$

Решение. Так как семейство экспоненциальное с правдоподобием $L(\theta) = e^{-\theta \sum_{i=1}^n x_i + n \ln \theta}$, то оценка $T(X) = \sum_{i=1}^n X_i$ является полной достаточной по теореме об экспоненциальном распределении. Решим уравнение несмещенности. По задаче 19 известно, что $p_T(x) = \theta^n \frac{x^{n-1}}{(n-1)!} e^{-\theta x} I_{x \geq 0}$. Положим $\varphi(x) = \frac{1}{\sqrt{x}}$. Тогда

$$\begin{aligned} \mathbb{E}\varphi(T(X)) &= \frac{\theta^n}{(n-1)!} \int_0^{+\infty} x^{n-3/2} e^{-\theta x} dx = -\frac{\theta^{n-1}}{(n-1)!} \left(x^{n-3/2} e^{-\theta x} \Big|_0^{+\infty} - (n-3/2) \int_0^{+\infty} x^{n-5/2} e^{-\theta x} dx \right) \\ &= \frac{\theta^{n-1}}{(n-1)!} (n-3/2) \int_0^{+\infty} x^{n-5/2} e^{-\theta x} dx = \dots = \frac{\theta^2}{(n-1)!} \prod_{i=1}^{n-2} (n-i-1/2) \int_0^{+\infty} \sqrt{x} e^{-\theta x} dx \\ \int_0^{+\infty} \sqrt{x} e^{-\theta x} dx &= \int_0^{+\infty} 2t^2 e^{-\theta t^2} dt = -\frac{1}{\theta} \int_0^{+\infty} t d(e^{-\theta t^2}) = \frac{1}{\theta} \int_0^{+\infty} e^{-\theta t^2} dt = \frac{\sqrt{\pi}}{2} \frac{1}{\theta \sqrt{\theta}} \Rightarrow \\ \mathbb{E}\varphi(T(X)) &= \frac{\sqrt{\pi} \prod_{i=1}^{n-2} (n-i-1/2)}{2(n-1)!} \sqrt{\theta} \end{aligned}$$

Отсюда положив $\psi(T(X)) = \frac{2(n-1)!}{\sqrt{\pi} \prod_{i=1}^{n-2} (n-i-1/2)} \frac{1}{\sqrt{T(X)}} = \frac{2(n-1)!}{\sqrt{\pi} \prod_{i=1}^{n-2} (n-i-1/2)} \frac{1}{\sqrt{\sum_{i=1}^n X_i}}$ мы получим оптимальную оценку.