контрольная.

6 задач

про точечные оценки:
1. Найти оценку метод. макс. правд.
2. —— // —— моментов.
3. свойства оценок несмещ / состоят / эффект.

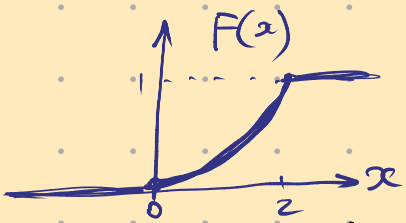
про CI:

4. построение асимп. CI с помощью ЦПТ / дельта-метод.
5. построение асимп CI для оценки макс правд.
6. классич CI (t, F, χ^2, N)

№5 построение асимп CI на базе оценки методом максим. правдоподобия

ф-ция распр. X_i (незав.)

$$F(x)=\begin{cases} x^a/2^a, & x\in[0;2] \\ 1, & x>2 \\ 0, & x<2 \end{cases}$$



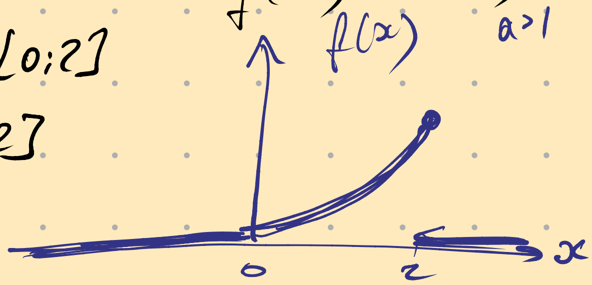
!! а) найдите точечную оценку $\hat{a}$ мет. макс. правд. по выборке из n набл.
б) оцените информацию Фишера
в) Постройте 95% асимпт интервал для a.

решение:

В некотор. случае $\max\limits_a f(x_1, \ldots, x_n)$

$$f(x)=F'(x)$$

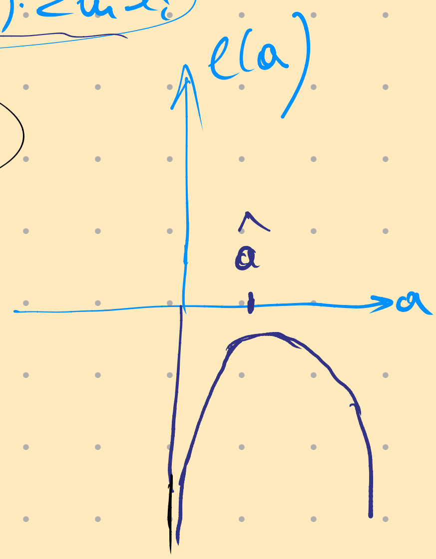
$$f(x)=\begin{cases} a x^{a-1}/2^a & x\in[0;2] \\ 0, & x\notin[0;2] \end{cases}$$



$$f(x_1, \ldots, x_n) = f(x_1)\cdot f(x_2)\cdot \ldots \cdot f(x_n) =$$
$$= \prod_{i=1}^{n} a x_i^{a-1}/2^a$$ ← функция правдоп.

$$l(a) = \ln f(x_1 \ldots x_n) =$$ ← лог функция правдоподобия
$$= (\ln a + (a-1)\ln x_i - a\ln 2)$$

$$= n\ln a - na\ln 2 + (a-1)\cdot\sum \ln x_i$$



$$\frac{\partial l}{\partial a} = \frac{n}{a} - n\ln 2 + \sum \ln x_i$$

$$\frac{\partial^2 l}{\partial a^2} = -\frac{n}{a^2} < 0$$

$$\frac{n}{\hat a} - n\ln 2 + \sum \ln x_i = 0.$$

$$\frac{n}{\hat a} = n\ln 2 - \sum \ln x_i$$

a) $\Downarrow$ $$\hat a = \frac{n}{n\ln 2 - \sum \ln x_i} = \frac{1}{\ln 2 - \frac{\sum \ln x_i}{n}} = \frac{1}{\ln 2 - \bar Y}$$

$\ln x_i = Y_i$

по ЦПТ: $\bar Y$ ас. норм.

по дельта-методу: $\hat a$ ас. норм.

б) инфо Фишера интуит:

сколько (в уе/в попугаях) инфо о неизв. a содержится в выборке.

$I = n \cdot$ кол-во инфо об $\boxed{a}$ от одного наблюдения

формально

a – скаляр

$$I = Var\left(\frac{\partial l}{\partial a}\right) = -E(H) = -E\left(\frac{\partial l}{\partial a}\right)$$

опр.

так считать легче как правило

$$\frac{\partial^2 l}{\partial a^2} = -\frac{n}{a^2}$$

...нет ничего сверх-го!

n, a, константа [n - изв-ая, a - неизв.]

$x_1, x_2 \dots x_n$, $\hat a$ — СВ. [какие]

$$I = -E\left(-\frac{n}{a^2}\right) = \frac{n}{a^2} \;!!$$

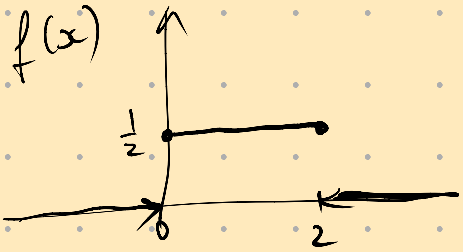
оценка инфо Фишера $\hat I = \frac{n}{\hat a^2}$, где $\hat a = \frac{n}{n\ln 2 - \sum \ln x_i}$

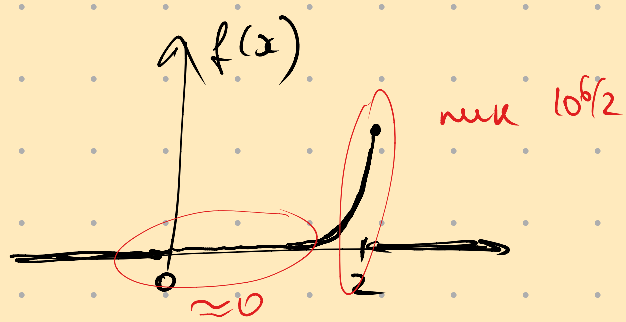
$$\boxed{\hat I = \frac{(n\ln 2 - \sum \ln x_i)^2}{n}}$$

почему форма зав-сть от a интуитивна?

a – малое, $a=1$

a – большое, $a = 1000000$





$$f(x) = \begin{cases} a x^{a-1}/2^a, & x \in [0;2] \\ 0, & x \notin [0;2] \end{cases}$$

выборка: 0.7, 1.2, 0.3, 1.4...

выборка: 1.9999, 1.9998, 1.9999, 1.9998...

интуиция: при малых a выборка больше несёт инфо об a.

основная роль инфо Фишера

$$\lim_{n\to\infty} Var(\hat a)\cdot I_F = I$$

identity matrix (ед. матрица)

Fisher info

$\hat a$ метод. максим. правд.

$$\hat{Var}(\hat a) = \hat I_F^{-1} \qquad \hat{Var}(\hat a) = \frac{n}{(n\ln 2 - \sum \ln x_i)^2}$$

оценили дисперсию оценки!

б) 95% asy CI для a
99% asy CI для a

$\hat{a}$ асимп. норм-но.

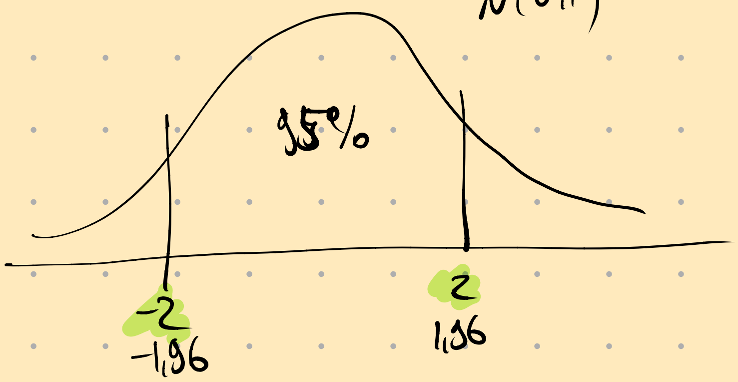
$$\frac{\hat{a}-a}{\sqrt{\hat{Var}(\hat{a})}} \xrightarrow[n\to\infty]{dist} N(0;1)$$

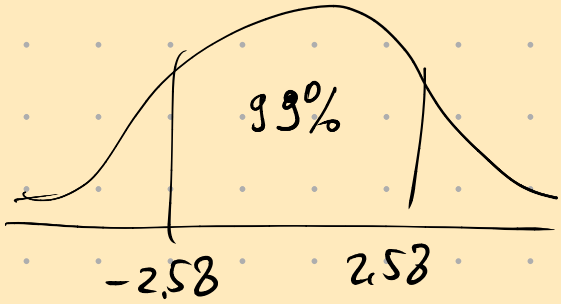
$$\underbrace{se(\hat{a})} = \sqrt{\hat{Var}(\hat{a})}$$

standard error
станд. ошибка

$$\frac{\hat{a}-a}{se(\hat{a})} \xrightarrow{dist} N(0;1)$$

$N(0;1)$





95% asy $[\hat{a} - 1{,}96 \cdot se(\hat{a}); \hat{a} + 1{,}96\, se(\hat{a})]$

99% asy $[\hat{a} - 2{,}58\, se(\hat{a}); \hat{a} + 2{,}58\, se(\hat{a})]$

$n = 1000$
$\sum \ln x_i = 400$

$$\hat{a} = \frac{n}{n\ln 2 - \sum \ln x_i} = \frac{1000}{1000\ln 2 - 400} = 3.4$$

$$\hat{I} = \frac{n}{\hat{a}^2} = \frac{1000}{3{,}4^2} = 85.9$$

$$\hat{Var}(\hat{a}) = \hat{I}^{-1} = 1/85.9 \approx 0.0116$$

$$se(\hat{a}) = \sqrt{0.0116} \approx 0.11$$

95% asy $[3.4 - 2\cdot 0.11; 3.4 + 2\cdot 0.11]$

$[3{,}18; 3{,}62]$

№6 → CI для $\mu = E(X_i)$
→ CI для $p = P(X_i = 1)$

→ CI для $\mu_a - \mu_b$
→ CI для $p_a - p_b$
→ CI для σ_a^2 / σ_b^2

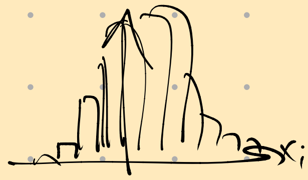
$\sigma^2 = Var(X)$

→ CI для $\sigma^2 = Var(X_i)$

суекорий !

если: $X_1, X_2, \ldots, X_n \sim$ незав $N(\mu; \sigma^2)$

на проспише

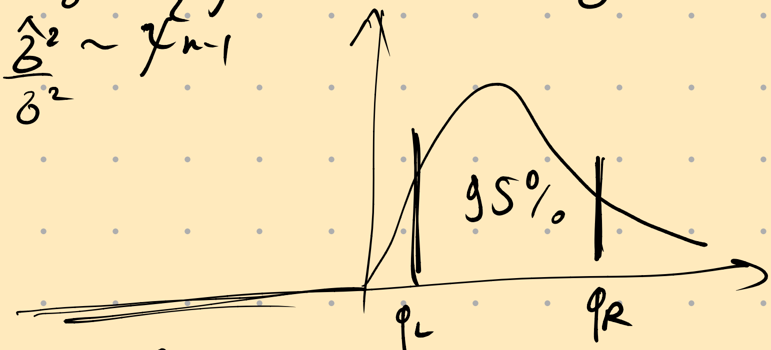


Теоремка! то:

$$\sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{\sigma^2} \sim \chi^2_{n-1}$$

$$\hat{\sigma}^2 = \frac{\sum (X_i - \bar{X})^2}{n-1}$$

$$(n-1) \cdot \frac{\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-1}$$



$$P\left(q_L \le \frac{(n-1) \cdot \hat{\sigma}^2}{\sigma^2} \le q_R\right) = 0{,}95$$

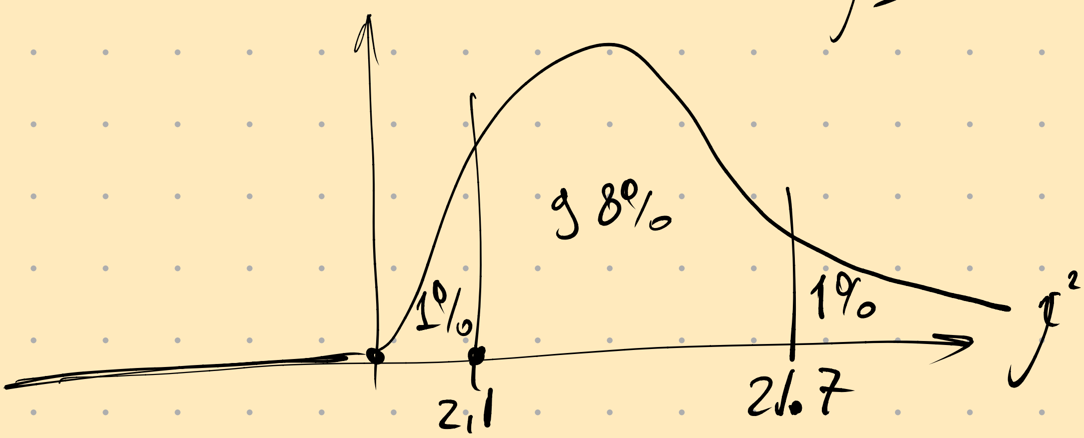
$$P\left(\frac{(n-1) \cdot \hat{\sigma}^2}{q_R} \le \sigma^2 \le \frac{(n-1) \cdot \hat{\sigma}^2}{q_L}\right) = 0{,}95$$

$$\frac{\sum (X_i - \bar{X})^2}{q_R} \le \sigma^2 \le \frac{\sum (X_i - \bar{X})^2}{q_L}$$

Пример. $n = 10$ $\quad \sum X_i^2 = 1000$ $\quad \sum X_i = 20$ $\quad X_i \sim$ нез $N(\mu; \sigma^2)$

постройте 98% CI для σ^2.

χ^2_9

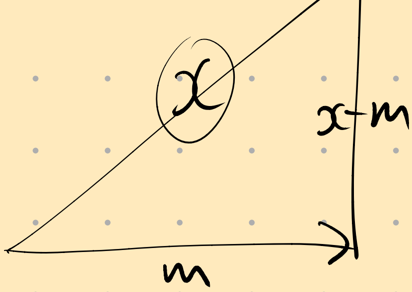


$$CI: \left[\frac{\sum (X_i - \bar{X})^2}{21.7}; \frac{\sum (X_i - \bar{X})^2}{2.1}\right]$$

$$\sum (X_i - \bar{X})^2 = \ldots = \sum X_i^2 - n \cdot \bar{X}^2 = 1000 - 10 \cdot \left(\frac{20}{10}\right)^2 = 1000 - 400 = 600.$$

$$\|x - m\|^2 = \|x\|^2 - \|m\|^2$$

$(x_1$



$$x = \begin{pmatrix} \vdots \\ x_n \end{pmatrix}$$

98%

$$CI: \left[\frac{600}{21.7} ; \frac{600}{2.1} \right]$$

$$m = \begin{pmatrix} \bar{x} \\ \bar{x} \\ \vdots \\ \bar{x} \end{pmatrix}$$

$$x - m = \begin{pmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ \vdots \\ x_n - \bar{x} \end{pmatrix}$$