Сегодня 18:30 кр.

→ Тестирование гипотез

Общий антураж

- набл.: $y_1, y_2 \ldots y_n$

простой:
– скалярные
– независимые

- знаем зак. распределения $f(y_1 \ldots y_n \mid \theta)$

простой
– θ - скаляр пар-р.

- θ не наблюдаем.

→ Байес. подход — задаём дополнительно априорное распр-ие θ $f(\theta)$.

→ частотн. подход. — нет доп. предп про θ; θ - неизв пар-р

Есть две гипотезы:
нулевая H_0 : $\theta \in \Xi_0$
альтернативная H_1 (H_A) : $\theta \in \Xi_1$

$$\Xi_0 \cap \Xi_1 = \varnothing$$

data $y_1 \ldots y_n$ → $\boxed{\text{тестирование}}$ →
- H_0 отвергается (H_0 is rejected)
- H_0 не отвергается (H_0 is not rejected)

(!) мы не говорим вердикт "H_0 верна"

(!) почему лучше не говорить „принимаем H_0"?

→ не отвергнуться может несколько противоречащих H_0^A и H_0^B.

Какая гипотеза нулевая? Какая альтернативная?

H_0: гипотеза, не требующая каких-либо действий (не выводим новое лек-во, не ост-м произв линию, не блокируем карту...)

H_1: гипотеза, треб-ая изменений привычного порядка действий.

решение \ истина	H_0	H_1
H_0 отвергнута	ош. I-го рода	☺
H_0 не отвергнута	☺	ош. II рода

(statistical test)

<u>def</u> <u>Статистический тест</u> $K \subseteq \mathbb{R}^n$

правило вида:

→ если $(y_1 \dots y_n) \in K$, то H_0 отвергается

→ если $(y_1 \dots y_n) \notin K$, то H_0 не отвергается.

K – критическая область (critical region)

$\boxed{\text{Как правило:}}$

Шаг 1. По вектору $(y_1 \dots y_n)$ считаем скалярную величину $T(y_1 \dots y_n)$ [тестовая статистика / test statistic]

<u>Шаг 2.</u>

$$\begin{cases} \text{если } T \geq T_{crit}, \text{ то } H_0 \text{ отвергается} \\ \text{если } T < T_{crit}, \text{ то } H_0 \text{ не отвергается} \end{cases}$$

<u>3 вар-та:</u>

$\begin{cases} \text{если } T \le T_{\text{отс}}, \text{ то } H_0 \text{ отвергается} \\ \text{если } T > T_{\text{отс}}, \text{ то } H_0 \text{ не отверг.} \end{cases}$

$\begin{cases} \text{если } T \le T^L_{\text{отс}} \text{ или } T \ge T^R_{\text{отс}}, \text{ то } H_0 \text{ отвергается} \\ \text{не отвергается иначе.} \end{cases}$

def Гипотеза → простая: множество Ξ состоит из одной точки

→ сложная: мн-во Ξ состоит более, чем из одной точки

$(H_0: p = \frac{1}{2})$

$(H_1: p \ne \frac{1}{2})$

---

def Размер теста (test size) с критической областью K

$$\alpha = \sup_{\theta \in \Xi_0} P(H_0 \text{ отвергнута} \mid \theta)$$

$$\alpha = \sup_{\theta \in \Xi_0} P((y_1 \dots y_n) \in K \mid \theta)$$

def. функция $\pi: \Theta \to P((y_1 \dots y_n) \in K \mid \theta)$ наз-ся функцией мощности (мощность теста) (power function)

def. уровень значимости (significance level) ! зависит от источника !

- синоним размеру теста
- α_0 - уровень значимости теста, если $\alpha_0 \ge$ размер теста

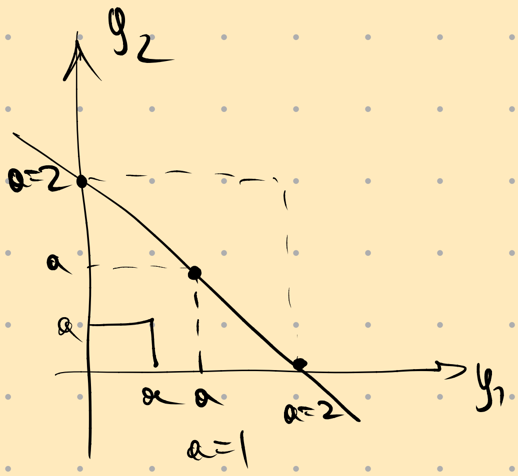
Упр. $y_1, \dots, y_n$ ~ незав Unif $[0; a]$ $\quad H_0: a \in [1; 1.5]$

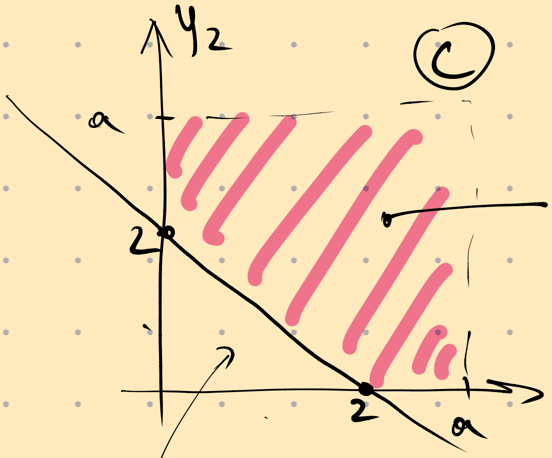
$H_1: a \in [2, 3]$

тест: если $y_1+y_2 \geq 2$, то H_0 отвергается
если $y_1+y_2 < 2$ то H_0 не отвергается.

а) $\pi(a)$ – ф-ция мощности
б) размер теста $= \sup_{a\in[1;1.5]} \pi(a) = 2\cdot\frac{0.5^2}{1.5^2} = \frac{2}{9}$
в) $P(\text{ош I рода} \mid a=1)$? $=0$
г) $P(\text{ош II рода} \mid a=2)$? $= 1-\frac{1}{2} = \frac{1}{2}$

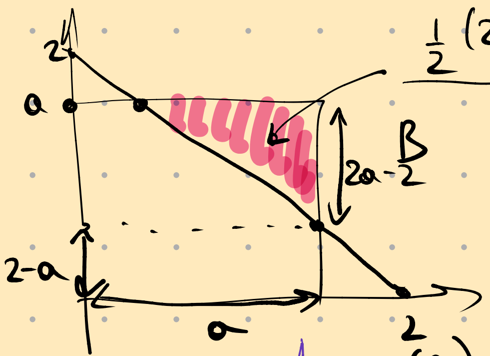
$$\pi(a) = P(y_1+y_2 \geq 2 \mid a)$$



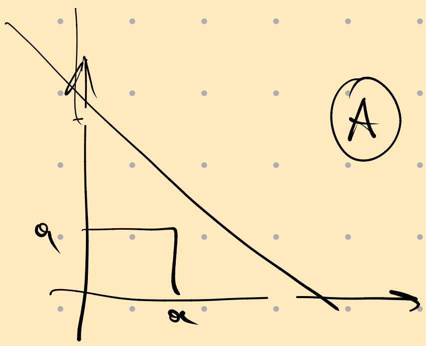


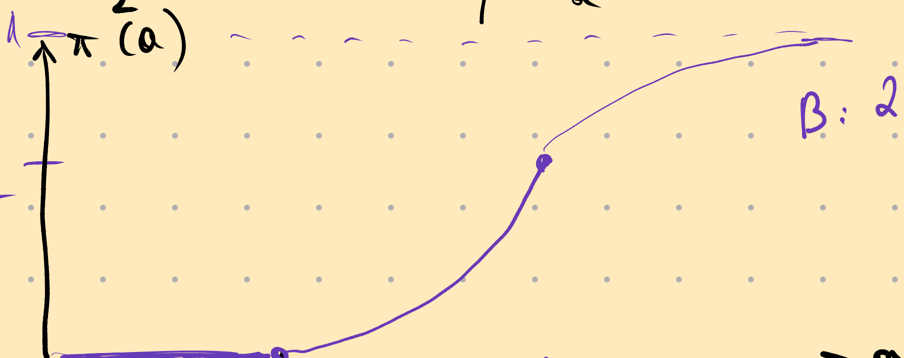
$$\frac{\frac{1}{2}\cdot 2\cdot 2}{a^2} = \frac{2}{a^2}$$

$$\pi(a) = P(\text{отверг } H_0 \mid a) = \begin{cases} 0 & a < 1 & A \\ \frac{2(a-1)^2}{a^2} & a \in [1;2] & B \\ 1-\frac{2}{a^2} & a > 2 & C \end{cases}$$



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





B: $2\cdot\left(1-\frac{1}{a}\right)^2$

Тестирование гипотез и построение довер-го интервала.

Упр. $X_1, X_2 \dots X_{10} \sim N(\mu; \sigma^2)$ незав $n=10$

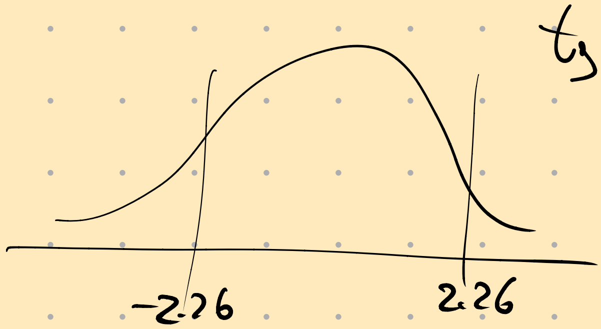
$\bar{X} = 20$ $\hat{\sigma}^2 = \frac{\sum (X_i - \bar{X})^2}{n-1} = 100$

а) 95% CI для μ.

б) $T = \frac{\bar{X} - \mu}{\sqrt{\frac{\hat{\sigma}^2}{n}}}$ Найдите критическую область вида $\{|T| \geq T_{crit}\}$ такую, что $P(\text{ошибки I рода} \mid \mu = 30) = 0.05$.

в) проведём тест: $H_0: \mu = 30$, $H_1: \mu \neq 30$ на ур-не значимости 5%

а) $\frac{\bar{X} - \mu}{\sqrt{\frac{\hat{\sigma}^2}{n}}} \sim t_{n-1}$



$$-t_{crit} \leq \frac{\bar{X} - \mu}{\sqrt{\frac{\hat{\sigma}^2}{n}}} \leq t_{crit}$$

↳ решаем нер-во отн-но μ

$$\mu \in \left[\bar{X} - t_{crit} \cdot \sqrt{\frac{\hat{\sigma}^2}{n}};\ \bar{X} + t_{crit} \cdot \sqrt{\frac{\hat{\sigma}^2}{n}}\right]$$

$$\mu \in \left[20 - 2.26 \cdot \sqrt{\frac{100}{10}};\ 20 + 2.26 \cdot \sqrt{\frac{100}{10}}\right]$$

$$\mu \in [12.85;\ 27.15]$$

б) $P(|T| \geq T_{crit} \mid \mu = 30) \stackrel{?}{=} 0.05$

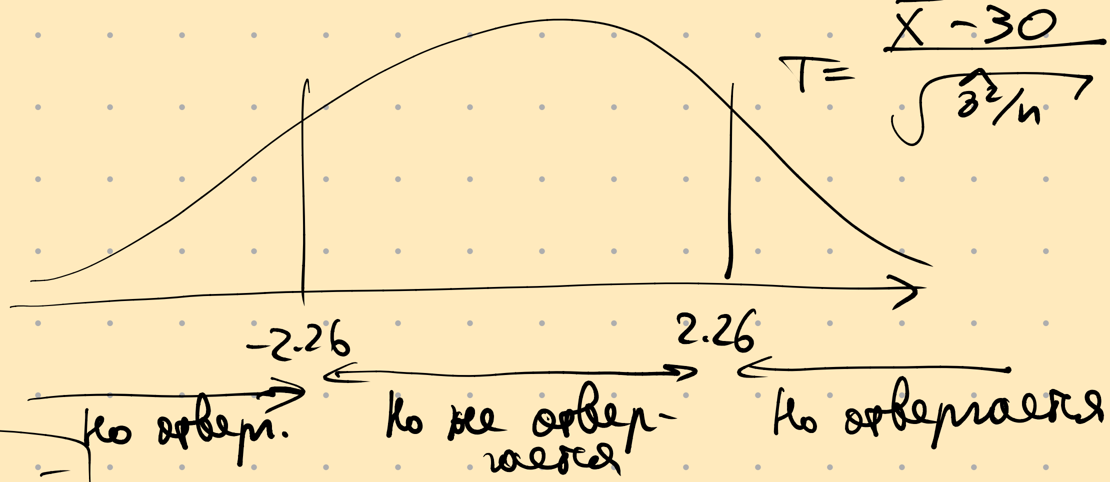
$$P\left(\left|\frac{\bar{X}-\mu}{\sqrt{\hat{\sigma}^2/n}}\right| \geq t_{crit} \;\middle|\; \mu=30\right)=0.05$$

t_g

$$P(|t_g| \geq t_{crit})=0.05$$

t_g

$$T=\frac{\bar{X}-30}{\sqrt{\hat{\sigma}^2/n}}$$



б) **Критерий**

(даёт вер-сть ошибки I рода в 5% при $\mu=30$)

Если $T=\dfrac{\bar{X}-30}{\sqrt{\hat{\sigma}^2/n}}$ $|T| \geq 2.26$, то H_0 отвергается

если T $|T| < 2.26$, то H_0 не отвергается

в) применим критерий !!

$$T=\frac{20-30}{\sqrt{100/10}}=\frac{-10}{\sqrt{10}}=-3.16$$

вывод: H_0 отвергается

Мораль: де-факто проверка гипотезы

оказалась бы самой построению CI.

$30 \in CI \xrightarrow{\text{нет}}$ H_0 отверг.

$30 \in CI \xrightarrow{\text{да}}$ H_0 не отверг.

$$T = \frac{\bar{X} - 30}{\sqrt{\hat{\sigma}^2 / n}}$$

$|T| \geq T_{crit} \to$ H_0 отверг.

$|T| < T_{crit} \to$ H_0 не отверг.

Мораль:
если пар-р скалярный, то построение CI гораздо нагляднее :)