орг: кр. 17 мая 18^{10}

Продолжаем тестирование гипотез.

$H_0: \theta \in \Xi_0$ $\quad \Xi_0 \cap \Xi_1 = \emptyset$

$H_1\ \theta \in \Xi_1$

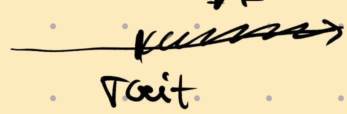
данные: $y_1 \ldots\ldots y_n$

типичный критерий:

шаг 1. $T = T(y_1 \ldots\ldots y_n)$ зав-т только от $y_1 \ldots\ldots y_n$ [и от значения пар-ра H_0/H_1]

Шаг 2 $\quad K$ - критич. область

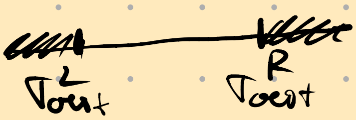
K



→ если $T \geq t_{crit}$, то H_0 отвергается [иначе нет]

→ Если $T \leq t_{crit}$, то H_0 не отвергается

→ если $\left[\begin{array}{l} T \geq t^{R}_{crit} \\ T \leq t^{L}_{crit} \end{array}\right.$, то H_0 отвергается



Как правило мы выбираем вероятность ошибки I рода

$\alpha = P(\text{отверг } H_0 \mid H_0 \text{ верна})$

верхнюю границу для ошибки I рода

α = уровень значимости $\geq$

$\geq P(\text{отв } H_0 \mid H_0 \text{ верна})$

двойственность: если пар-р θ - скалярный то можно переформулировать тестирование в CI.

Хи-квадрат тест Пирсона.

цель: проверить соответствие распределения качественной переменной заданному мультиномиальному.

Пример.

	караси	щуки	сомы
вер:	α	2α	$1-3\alpha$

α - неизв-ый пар-р

$\alpha \in [0; \frac{1}{3}]$

фактич. частоты	20	30	40

$H_0: \exists \alpha \begin{cases} p_{кар} = \alpha \\ p_{щуки} = 2\alpha \\ p_{сомы} = 1-3\alpha \end{cases}$ $\qquad H_1$: не H_0

Шаг 1 Состоят-ая оценка θ [оценка мет. макс./ оценка мет. мом-в]

$\hat{\theta}$

Шаг 2. есть ~~оценённые~~ вер-сти каждой категории

$\hat{p}_1, \hat{p}_2, \ldots, \hat{p}_k$

k - общее кол-во категорий.

оценённые вер-сти попадания в каждую категорию

Шаг 3

$$\chi^2 = \sum_{i=1}^{k} \frac{(N_i - N\hat{p}_i)^2}{N\hat{p}_i}$$

N - общее кол-во наб.

N_i - кол-во наблюдений попавших в i-ую категорию.

Если $\chi^2 > \chi^2_{crit}$, то H_0 отвергается [иначе - нет]

Теорема: При верной H_0 и $N \to +\infty$

$$\chi^2 \xrightarrow{dist} \chi^2_d$$

$$d = df_{UR} - df_R$$

df_{UR} – число своб-х пар-в в неогр. модели (H_0 или H_1)
df_R – число своб-х пар-в в огранич. модели (H_0)

практике: в каждую категорию попало ≥ 10

$$\alpha = P(\text{отв } H_0 \mid H_0 \text{ верна}) = 0.05$$

Шаг 1. оценка $\hat{\alpha}$. [метод макс правдоп]

	Караси	Щуки	Сомы
$[p_i(\alpha)]$	α	2α	$1-3\alpha$
факт. част $[N_i]$	20	30	40

$N=90$

$$\frac{(20+30+40)!}{20!\,30!\,40!}\cdot \alpha^{20}\cdot(2\alpha)^{30}\cdot(1-3\alpha)^{40} \to \max_{\alpha}$$

$$\ln(\ldots) + 20\ln\alpha + 30\ln(2\alpha) + 40\ln(1-3\alpha) \to \max_{\alpha}$$

$$\ell(\alpha) = \text{const} + 50\ln\alpha + 40\ln(1-3\alpha) \to \max_{\alpha}$$

$$\frac{\partial \ell}{\partial \alpha} = \frac{50}{\alpha} + \frac{40}{1-3\alpha}\cdot(-3)$$

$$\frac{50}{\hat{\alpha}} = \frac{3\cdot 40}{1-3\hat{\alpha}}$$

$$5(1-3\hat{\alpha}) = 12\hat{\alpha}$$

$$5-15\hat{\alpha} = 12\hat{\alpha}$$

$$\hat{\alpha} = \frac{5}{27}$$

$p_1(\alpha) = \alpha$ $\quad$ $p_2(\alpha) = 2\alpha$ $\quad$ $p_3(\alpha) = 1-3\alpha$

$\hat{p}_1(\hat{\alpha}) = \frac{5}{27}$ $\quad$ $\hat{p}_2 = \frac{10}{27}$ $\quad$ $\hat{p}_3 = \frac{12}{27}$

$$\chi^2 = \sum \frac{(N_i - N\hat{p}_i)^2}{N\hat{p}_i} = \frac{\left(20 - 90\cdot\frac{5}{27}\right)^2}{90\cdot\frac{5}{27}} + \frac{\left(30-90\cdot\frac{10}{27}\right)^2}{90\cdot\frac{10}{27}} +$$

$$+\frac{\left(40-90\cdot\frac{12}{27}\right)^2}{90\cdot\frac{12}{27}} = \ldots = 1$$

утверждение (т. Пифагора)

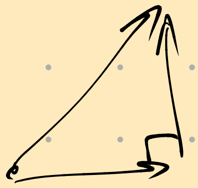
$$\sum\frac{(N_i - N\hat{p}_i)^2}{N\hat{p}_i} = \sum\frac{N_i^2}{N\hat{p}_i} - N$$

квадрат длины катета

кв. длины гип.

кв. длины катета

упр.



$$\chi^2 = \frac{20^2}{90\cdot 5/27} + \frac{30^2}{90\cdot\frac{10}{27}} + \frac{40^2}{90\cdot\frac{12}{27}} - 90 = 1$$

χ^2_{obs}

Считаем степени свободы !!

$\beta = 2\alpha$ $\gamma = 1-3\alpha$

ограниченная модель: H_0: | α | 2α | $1-3\alpha$ |

$df_R = 1$

неогранич. модель H_0 или H_1 | p_1 | p_2 | p_3 |

$1-p_1-p_2$

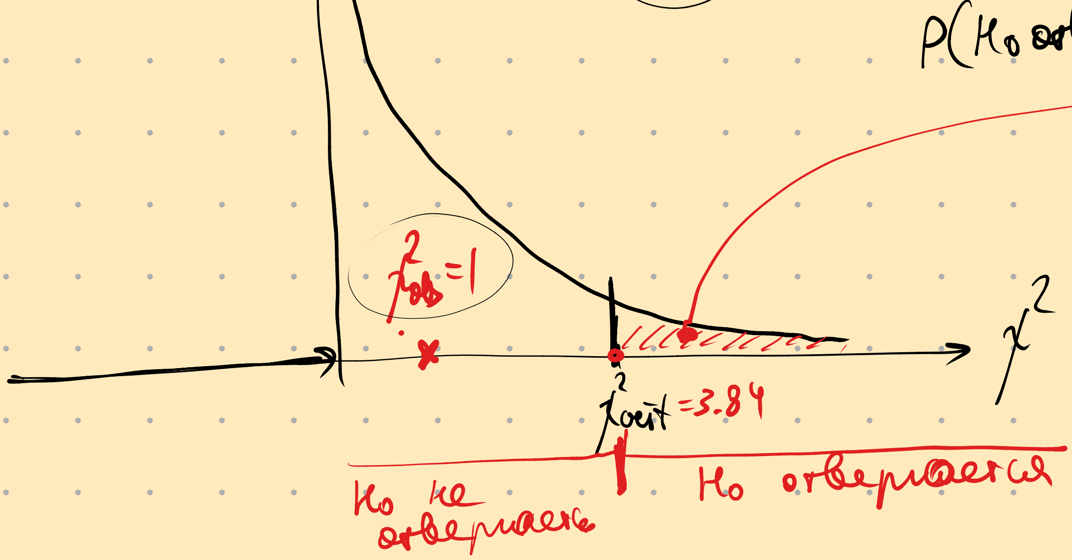
$p_1+p_2+p_3=1$

$df_{UR} = 2$

χ^2_d $\quad d = df_{UR} - df_R = 2-1 = 1$ ← число степеней свободы

ф. плотн. χ^2_1

$P(H_0 \text{ отвер} \mid H_0 \text{ верна}) = 0.05$



$$\chi^2_{crit} = 3.84$$

критер: если $\chi^2 \geq 3.84$, то H_0 отвергается
$\chi^2 < 3.84$, то H_0 не отвергается

$\chi^2 = 1 \Rightarrow$ H_0 не отвергается.

---

p- значение [p-value]

Взгляд алгоритмический

$T = T(y_1 \dots y_n)$

Теорема: если H_0 и $n \to \infty$ то $T \to Law$

Три вар-та
- если $T \geq T_{crit}$ то H_0 отверг.
- если $T \leq T_{crit}$ то H_0 отверг
- если $\begin{cases} T \geq T^{R}_{crit} \\ T \leq T_{Lcrit} \end{cases}$ то H_0 отверг.

переделываем критерий так.

$p\text{-value} = p\text{-value}(y_1 \dots y_n)$

если H_0 и $n \to \infty$ то $p\text{-value} \to U[0;1]$

если $p\text{-value} < \alpha$, то H_0 отвергается.
[иначе, то не отвергается]

$\alpha = P(\text{отверг } H_0 \mid H_0 \text{ верна})$

Опр. „пороговое" (устаревш.) p-value – пороговый уровень значимости (α)

если $p\text{-value} > \alpha$, то H_0 не отвергается
если $p\text{-value} \leq \alpha$, то H_0 отвергается

Опр. „смысловое" [для $K = [T_{crit}; +\infty)$]

$p\text{-value} = P(T_{new} \geq T_{obs} \mid H_0, T_{obs})$

предполагаем, что H_0 верна

фиксируем полученное знач-ие статистики T_{obs}.

Вер-ть того, что в новом эксперименте (с тем же числом наблюдений) тестовая статистика превысит полученную.

<u>продолжение упр-ия.</u>

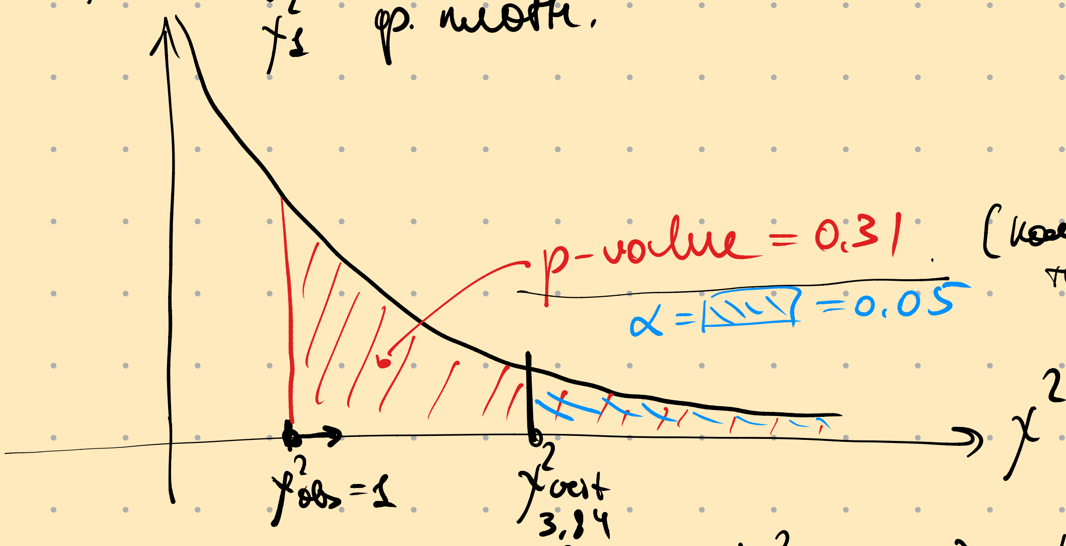
- посчитайте p-value в данной задаче.

$$\boxed{p\text{-value} = P(\chi^2_{new} \geq 1 \mid H_0) =}$$

значение χ^2_{obs} в нашем эксперименте

(при H_0 $\chi^2_{new} \sim \chi^2_1$)

χ^2_1 ф. плотн.



<u>Старая формулировка</u> если $\chi^2_{obs} \in K = [\chi^2_{crit}; +\infty)$, то H_0 отв.

если $\chi^2_{obs} \geq \underset{3{,}84}{\chi^2_{crit}}$, то <u>$H_0$ отверг.</u> [сравниваем числа на оси]

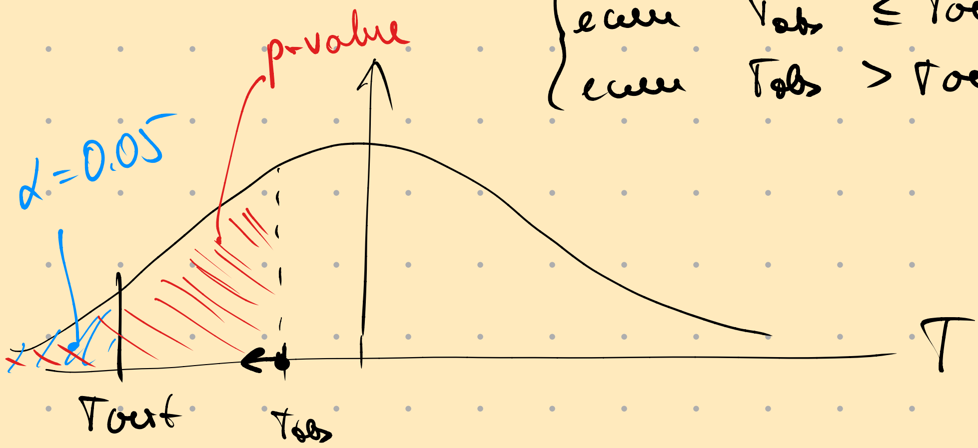
<u>формулировка через p-value</u>

$$\begin{cases} \text{если } p\text{-value} < \overset{0{,}05}{\alpha}, \text{ то } H_0 \text{ отверг.} \\ \text{если } p\text{-value} \geq \alpha, \text{ то } H_0 \text{ не отверг.} \end{cases}$$

[сравнение площадей]

левая крит. часть : $K = (-\infty; T_{crit}]$

$$\begin{cases} \text{если } T_{obs} \le T_{crit}, \text{ то } H_0 \text{ отверг.} \\ \text{если } T_{obs} > T_{crit} \text{ то } H_0 \text{ не отвер.} \end{cases}$$



$$\boxed{\text{p-value} = P(T_{new} \le T_{obs} \mid H_0, T_{obs})}$$

$\boxed{H_0 \text{ отверг}} \iff \boxed{T_{obs} < T_{crit}} \iff \boxed{\text{p-value} < \alpha}$

$\boxed{H_0 \text{ не отверг}} \iff \boxed{T_{obs} \ge T_{crit}} \iff \boxed{\text{p-value} \ge \alpha}$

инт. p-value — это вер-сть получить более экстремальное значение статистики, чем получено в нашем эксперименте при верной H_0.