Увод у непараметарске тестове

- предности у односу на параметарске тестове: мање претпоставки о расподели узорка!

1. Тест знакова (sign test)
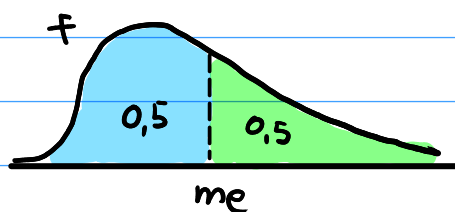
претпоставке: медијана расподеле me је 50% квантил

строга дефиниција: $m_e = \inf\{x : F(x) \geq \frac{1}{2}\} = F^{-1}(\frac{1}{2})$
ако је F строго растућа

објашњење: тачка са чије леве и десне стране оче-кујемо да се нађе приближно исти број тачака из псу.

видети значај: (за апс. непр. расподеле)

f - густина апс. непр. ф-је расподеле

ако EX < +∞!

специјално, ако је расподела симетрична: EX=me
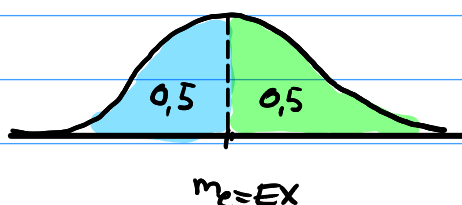
$$f(m_e+x)=f(m_e-x), \quad \forall x$$

зашто? $EX = \int_{-\infty}^{+\infty} x f(x)dx = \int_{-\infty}^{+\infty}(t+m_e)f(t+m_e)dt=$

$t=x-m_e$
$dt=dx$

$$= \int_{-\infty}^{+\infty} t f(t+m_e)dt + m_e \underbrace{\int_{-\infty}^{+\infty} f(t+m_e)dt}_{} = m_e$$

$$\underbrace{\phantom{\int_{-\infty}^{+\infty} t f(t+m_e)dt}}_{\overset{''}{0,}}$$

$t+m_e = x, \ dt = dx$

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

јер је t непарна, а

$f(t+m_e)$ парна ф-ја,

па им је производ непарна

ф-ја

Тестом знакова тестирамо хипотезе о меди-

јани <u>расподеле</u> из које долази ПСУ

↓

немамо додатне

претпоставке о

расподели

$H_0: \quad m_e = m_0 \qquad \text{vs} \qquad H_1: \ m_e \overset{>}{\underset{<}{\neq}} m_0$

Како да осмислимо тест статистику?

Ако је H_0 тачна, очекујемо да се приближно исти

број тачака из узорка нађе са обе стране m_0:

m_0
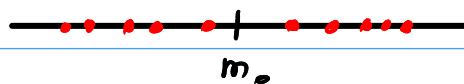
● – вредности из

узорка

Ако је $H_1: m_e > m_0$, каква слика иде у прилог томе?
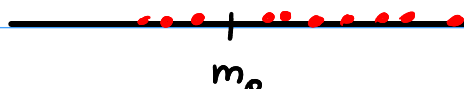
m_0

Дакле, бети број тачака из узорка са десне стране m_0 супериве на H_1

Како да направимо статистику која броји тачке бете од m_0?

$$T = \sum_{i=1}^{n} I\{X_i > m_0\}$$

при H_0: $P\{X_i > m_0\}$
$$= 1 - P\{X_i \leq m_0\} = \frac{1}{2}$$

При H_0: $I\{X_i > m_0\} \sim Ber\left(\frac{1}{2}\right)$, па тада $T \sim Bin\left(n, \frac{1}{2}\right)$

За $n > 10$ можемо е позвати на ЦГТ:

$$T^* = \frac{T - \frac{n}{2}}{\sqrt{\frac{n}{4}}} \underset{\text{при } H_0}{\sim} \mathcal{N}(0,1)$$

Дакле, за $H_1: m_e > m_0$ критична област те бити:

$$W = \{T \geq c_1\}$$

↓
по договору на овом
курсу ћемо за критичне
области узимати затворене
скупове;
када је тест статистика
при H_0 из апс. непр. расподеле,
за рачун нам није важно да
ли пишемо $>$ или $\geq$

За $n > 10$ користимо тест статистику T^*, па

$$\text{je} \quad W = \left\{ T \geq c_1 \right\} = \left\{ \frac{T - \frac{n}{2}}{\sqrt{\frac{n}{4}}} \geq \underbrace{\frac{c_1 - \frac{n}{2}}{\sqrt{\frac{n}{4}}}}_{\overset{..}{c}} \right\}$$

$$= \left\{ T^{*} \geq c \right\}$$

- Ako je H_1 : $m_e < m_o \Rightarrow W = \left\{ T^{*} \leq c \right\}$
- Ako je H_1 : $m_e \neq m_o \Rightarrow W = \left\{ |T^{*}| \geq c \right\}$