

HIPOTÉZISVIZSGÁLATOK, STATISZTIKAI PRÓBÁK

A sokaságot szeretnénk megismerni, de csak a minta áll rendelkezésünkre.

Az alapsokaságra vonatkozóan valamilyen feltevéssel élünk (pl. μ és/vagy σ értéke) és azt statisztikai próbával ellenőrizzük.

Jöhetnek-e az adatok olyan eloszlásból ...? Pl.: $\mu = \mu_0$

$$H_0 : \mu = \mu_0$$

nullhipotézis

$$H_1 : \mu \neq \mu_0$$

ellenhipotézis

***u*-próba**

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0$$

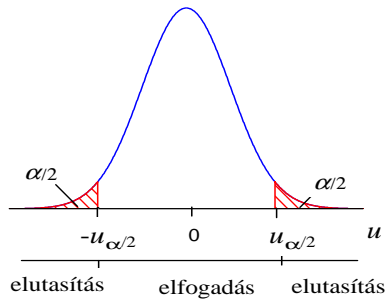
$$u = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

$$u_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

próbastatisztika

Ha H_0 igaz, $u_0 \sim u$

Ha u_0 olyan értékeket vesz föl, amelyeket u szokott, elfogadjuk



$$P(-u_{\alpha/2} < u_0 \leq u_{\alpha/2} | H_0) = 1 - \alpha$$

$$P\left(-u_{\alpha/2} < \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \leq u_{\alpha/2} | H_0\right) = 1 - \alpha$$

$$\mu_0 - u_{\alpha/2} \sigma / \sqrt{n} < \bar{x} < \mu_0 + u_{\alpha/2} \sigma / \sqrt{n}$$

$$\bar{x} - u_{\alpha/2} \sigma / \sqrt{n} < \mu_0 < \bar{x} + u_{\alpha/2} \sigma / \sqrt{n}$$

a konfidencia-intervallum tartalmazza a μ_0 értéket

u-próba

- kiszámítjuk a próbastatisztika aktuális értékét:

$$u_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \frac{\mu - \mu_0}{\sigma/\sqrt{n}}$$

értéke 0, ha H_0 igaz

u-eloszlású

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0, \text{ vagy } H_1 : \mu < \mu_0, \text{ vagy } H_1 : \mu > \mu_0.$$

- kijelöljük az elfogadási tartományt az előírt α szignifikanciaszinthez

Pl. $H_1 : \mu \neq \mu_0$ esetén

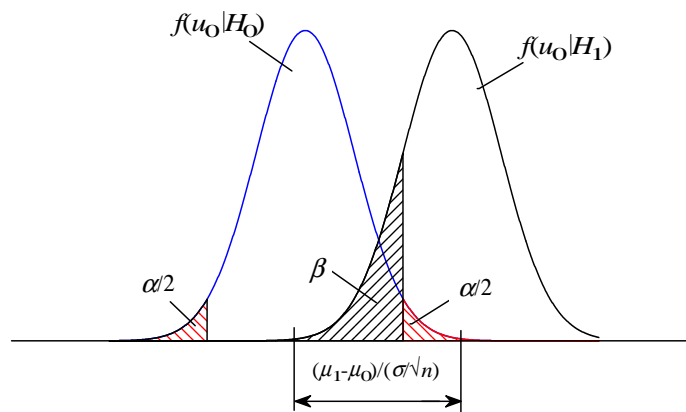
$$P\left(-u_{\alpha/2} < \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \leq u_{\alpha/2}\right) = P\left(\left|\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right| \leq u_{\alpha/2}\right) = 1 - \alpha$$

- megvizsgáljuk, hogy a próbastatisztika kiszámított értéke az elfogadási tartományban van-e
- ha igen, elfogadjuk a nullhipotézist

Első- és másodfajú hiba

nullhipotézis	döntés	
	a H_0 hipotézist	
	elfogadjuk	elutasítjuk
H_0 igaz	helyes döntés	elsőfajú hiba (α)
H_0 nem igaz	másodfajú hiba (β)	helyes döntés

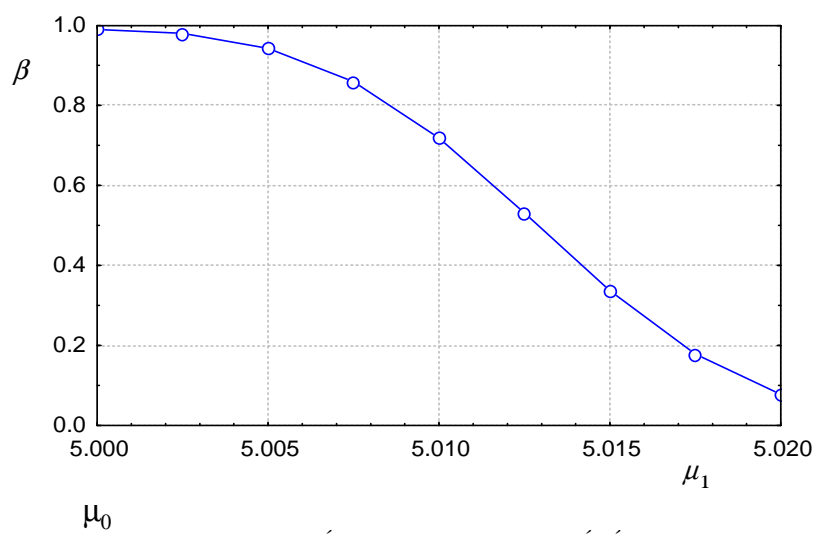
A másodfajú hiba valószínűsége



PARAMÉTERES STATISZTIKAI PRÓBÁK

7

Működési jelleggörbe (OC-görbe)



PARAMÉTERES STATISZTIKAI PRÓBÁK

8

7. példa

Táramérleggen négy ismételt tömegméréssel határoztuk meg egy tárgy tömegét. A 4 mérésből álló minta számtani középértéke 5.0125 g. Korábbi mérésekből tudjuk, hogy a mérés varianciája $\sigma^2 = 10^{-4} \text{ g}^2$. El kell döntenünk, hihető-e, hogy a várható érték (a tárgy valódi tömege) 5.0000 g.

$$H_0 : \mu = 5.0000, \quad H_1 : \mu \neq 5.0000$$

$$\bar{x} = 5.0125, \quad \sigma^2 = 10^{-4}, \quad n = 4, \quad \alpha = 0.05$$

$$u_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{5.0125 - 5.0000}{10^{-4} / 2} = 2.5$$

$$u_{\alpha/2} = 1.96$$

Az elfogadási tartomány: (-1.96; 1.96), a próbastatisztika aktuális értéke (2.5) ezen kívül van, így a H_0 hipotézist 0.05-os szignifikanciaszinten elvetjük.

Egymintás t -próba

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{\bar{x} - \mu}{s/\sqrt{n}} + \frac{\mu - \mu_0}{s/\sqrt{n}} = t + \frac{\mu - \mu_0}{s/\sqrt{n}}$$

$$P\left(-t_{\alpha/2} < \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \leq t_{\alpha/2}\right) = 1 - \alpha$$

8. példa

Egy analitikai módszer torzítatlanságának vizsgálatára 5 ismételt mérést végeztek. Az eredmények: 3.25, 3.27, 3.24, 3.26 és 3.24. Elfogadva, hogy az adatok közelítőleg normális eloszlásúak, ellenőrizzük 5%-os szignifikanciaszinten a torzítatlanság hipotézisét!

$$\bar{x} = 3.252$$

$$s = 0.013038$$

$$n = 5$$

$$t_0 = \frac{3.252 - 3.25}{0.013038/\sqrt{5}} = 0.342997$$

$$t_{\alpha/2}(4) = 2.776$$

Statisztikai próba és konfidencia-intervallum

Kétoldali eset

Elfogadási tartomány:

$$-t_{\alpha/2} < t_0 < t_{\alpha/2}$$

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Átrendezve

$$-t_{\alpha/2} s/\sqrt{n} < \bar{x} - \mu_0 \leq t_{\alpha/2} s/\sqrt{n}$$

$$\bar{x} - t_{\alpha/2} s/\sqrt{n} < \mu_0 \leq \bar{x} + t_{\alpha/2} s/\sqrt{n}$$

A μ várható érték $1-\alpha$ valószínűségű konfidencia-intervalluma

$$\bar{x} - t_{\alpha/2} s/\sqrt{n} < \mu \leq \bar{x} + t_{\alpha/2} s/\sqrt{n}$$

$$\bar{x} - t_{\alpha/2} s/\sqrt{n} < \mu_0 \leq \bar{x} + t_{\alpha/2} s/\sqrt{n}$$

Elfogadjuk a nullhipotézist ($\mu = \mu_0$), ha a konfidencia-intervallum tartalmazza a μ_0 feltételezett várható értéket.

χ^2 -próba a variancia vizsgálatára

$$H_0: \sigma^2 \leq \sigma_0^2 \quad H_1: \sigma^2 > \sigma_0^2$$

Ha H_0 igaz, akkor a következő kifejezés χ^2 -eloszlású,
szabadsági foka: $\nu = n - 1$

$$\chi_0^2 = \frac{s^2(n-1)}{\sigma_0^2}, \quad P\left(\frac{s^2(n-1)}{\sigma_0^2} \leq \chi_\alpha^2\right) = 1 - \alpha$$

9. példa

A 17. példa adatai alapján ellenőrizzük $\alpha=0.05$ -os szignifikanciaszinten, hogy elfogadható-e az az állítás, mely szerint a mérési módszer varianciája (σ^2) legfeljebb 10^{-4} (%)².

$$s = 0.013038$$

$$s^2 = 1.700 \cdot 10^{-4}$$

$$\chi_0^2 = \frac{0.00017 \cdot 4}{10^{-4}} = 6.8$$

$$\chi_0^2 = \frac{\sigma^2}{\sigma^2} \cdot \frac{0.00017 \cdot 4}{10^{-4}} = \frac{\sigma^2}{10^{-4}} \cdot \frac{0.00017 \cdot 4}{\sigma^2} = 6.8$$

$$\chi_{0.05}^2(4) = 9.488, \text{ felső határ}$$

> 1, ha H_1 igaz

$$H_0: \sigma^2 \leq 10^{-4}$$

$$H_1: \sigma^2 > 10^{-4}$$

χ^2

Két szórásnégyzet összehasonlítása (F-próba)

$$H_0: \sigma_1^2 = \sigma_2^2$$

A próbastatisztika: $F_0 = \frac{s_1^2}{s_2^2}; (n_1 - 1, n_2 - 1)$

Egyik oldali ellenhipotézis esetén: $H_1: \sigma_1^2 > \sigma_2^2$

Akkor utasítjuk el a nullhipotézist, ha

$$s_1^2 / s_2^2 > F_\alpha$$

Kétoldali ellenhipotézis esetén: $H_1: \sigma_1^2 \neq \sigma_2^2$

Akkor utasítjuk el a nullhipotézist, ha

$$\frac{s_1^2}{s_2^2} < F_{1-\alpha/2} \quad \text{vagy} \quad \frac{s_1^2}{s_2^2} < F_{\alpha/2}$$

$$s_1^2 / s_2^2 \geq 1$$

**elég az elfogadási tartomány
felső határát ellenőrizni**

95 %-os egyoldali szint = a 90 %-os kétoldali szintnek

Kétmintás t -próba

Adott a két független minta elemszáma (n_1 és n_2),
és szórásnégyzetük (s_1^2 és s_2^2).

Tételezzük fel, hogy a két sokaság varianciája
megegyezik. (Ezt F -próbával ellenőrizni kell!)

$$d = \bar{x}_1 - \bar{x}_2 \quad E(d) = \mu_1 - \mu_2$$

$$Var(d) = Var(\bar{x}_1 - \bar{x}_2) = \sigma^2 / n_1 + \sigma^2 / n_2$$

$$s_d^2 = s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)$$

$$s^2 = \frac{1}{n_1 + n_2 - 2} \left[s_1^2 (n_1 - 1) + s_2^2 (n_2 - 1) \right]$$

A következő kifejezés t -eloszlású

$$t = \frac{d - E(d)}{s_d} = \frac{d - E(d)}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \quad \nu = n_1 + n_2 - 2$$

$$H_0: \mu_1 = \mu_2, \text{ ekkor } E(d) = 0$$

A próbat statisztika:

$$t_0 = \frac{d - 0}{s_d} = \frac{d}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \quad \nu = (n_1 - 1) + (n_2 - 1)$$

A $\sigma_1^2 = \sigma_2^2$ feltevést F -próbával ellenőrizzük