

Mintaméret meghatározása

Mintaméret meghatározása standard hiba alapján

Standard hiba:

$$se = \frac{s}{\sqrt{n}}$$

Ez a populáció feltételezett eloszlása, amit a minta szórása alapján határozunk meg.

Probléma: a szórást ismernünk kell már az adatgyűjtés megkezdése előtt. Becslés vagy korábbi kísérletek, vagy saját előkísérlet eredménye alapján.

Példa: két kilós krumpliszsákok valós tömegét szeretnénk megsaccolni. A standard hiba értéke legyen $SE = 0,2$ kg.

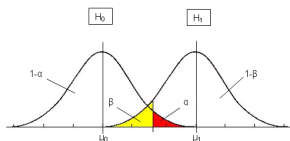
$$\frac{s}{\sqrt{n}} = 0,2$$

azaz

$$\left(\frac{s}{0,2}\right)^2 = n$$

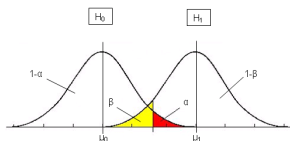
Hipotézis hibás elutasítása

1. α -hiba (első fajta, elsőfajú hiba): elutasítjuk H_0 -t, mert az átlag a megadott konfidenciaintervallumon kívül esik $\rightarrow \alpha$ része (piros tartomány).
2. β -hiba (második fajta, másodfajú hiba): megtartjuk H_0 -t, holott az átlag más populációhoz tartozik (sárga tartomány).



Hipotézis hibás elutasítása

1. α -hiba (első fajta, elsőfajú hiba): elutasítjuk H_0 -t, mert az átlag a megadott konfidenciaintervallumon kívül esik $\rightarrow \alpha$ része (piros tartomány).
2. β -hiba (második fajta, másodfajú hiba): megtartjuk H_0 -t, holott az átlag más populációhoz tartozik (sárga tartomány).



	H_0 -t megtartjuk	H_0 -t elvetjük
H_0 igaz	helyes döntés	α -hiba (álpozitív)
H_1 igaz	β -hiba (álnegatív)	helyes döntés

A próba ereje

A próba ereje (*power*): H_1 fennállása esetén H_0 -t – helyesen – elvetjük, azaz $1 - \beta$. Erő: helyes döntés valószínűsége H_1 igazságának esetén.

DE: a β -hiba valószínűsége függ (1) a konfidenciahatártól és (2) a két átlag különbségétől. Nagyon nagy minta esetén szinte bármilyen különbséget szignifikánsnak akarunk kimutatni.

Pl. 1000 véletlen egész szám 1 és 100, valamint 2 és 101 között. t -próba nem szignifikáns, de 10000 ugyanilyen számra igen.

Minta erejének meghatározása

Az adott elemszám mellett elég nagy-e a próba ereje?

```
power.t.test()
```

delta: szakmailag már relevánsnak tekinthető különbség

sd: szórás

sig.level: szignifikanciaszint

type: kétmintás t-próba

alternative: egy- vagy kétoldali teszt

Kimenetben output: $\text{power} = 0.93$ ← a kívánt szignifikáns különbség elszalasztásának lehetősége a minta ereje $1 - \text{power}$, azaz 6,7%. A jelenlegi teszt ereje jó, tehát fölösleges a mintaméretet növelni.

```
analóg: power.anova.test()
```

Mintaméret meghatározása adott erő mellett

Tudjuk, hogy mekkora javulást szeretnénk elérni. A próba erejét 0,9-es értékűnek tervezzük.

`power.prop.test()`

p1: első mintában adott előfordulások aránya

p2: második mintában adott előfordulások aránya

sig.level: szignifikanciaszint

power: elvárt próbaerő

alternative: egy- vagy kétoldali

n ismeretében reális-e ennyi adatközlővel kísérletet lefuttatni?