

Statisztika1

Paraméteres eljárások,
Normalitásvizsgálat
Egymintás t-próba

Felhasznált irodalmak

- Balázs Katalin órai diasorai
- Field, A., Miles, J., & Field, Z. (2012).
Discovering Statistics Using R. SAGE
Publications.

Paraméteres eljárások

- Becsülik a populáció egy paraméterét
- Alkalmazásuknak számos feltétele van (paraméterek és a változó eloszlása)
- Csak normál eloszlású, folytonos változók esetén alkalmazhatók!

Normalitásvizsgálat

- Grafikus exploratív technikák, pl. hisztogram (szimmetria, egycsúcsúság)
- Normalitás tesztelése, pl. ferdeségi együttható, csúcsossági együttható,
- Amennyiben a változó nagy valószínűséggel nem normális elosztású, robusztus, nemparaméteres eljárások alkalmazása javasolt!

$$z_{\text{skewness}} = \frac{S - 0}{SE_{\text{skewness}}}$$

$$z_{\text{kurtosis}} = \frac{K - 0}{SE_{\text{kurtosis}}}$$

- `library(moments)`
- `agostino.test(adat$kor)`
- `anscombe.test(adat$kor)`

Shapiro-Wilk teszt

- 1965, Samuel Sanford Shapiro és Martin Wilk
- H_0 : Normál elosztású populációból származik a minta
 - Tehát ha $p < 0,05$ akkor a minta NEM normál eloszlású populációból származik
- Legmegbízhatóbbnak tartják
- `shapiro.test(adat$kor)`

Lilliefors teszt

- Hubert Lilliefors
- Kolmogorov-Szmirnov teszten alapul
- H_0 : Normál elosztású populációból származik a minta
 - Tehát ha $p < 0,05$ akkor a minta NEM normál eloszlású populációból származik

```
library(nortest)
```

```
lillie.test(adat$kor)
```


Student-féle t-eloszlás

- Egy ismert várhatóértékű (μ) és szórású (σ) normális eloszlásból vett véletlenszerű minta empirikus várható értékével számított u paraméter:

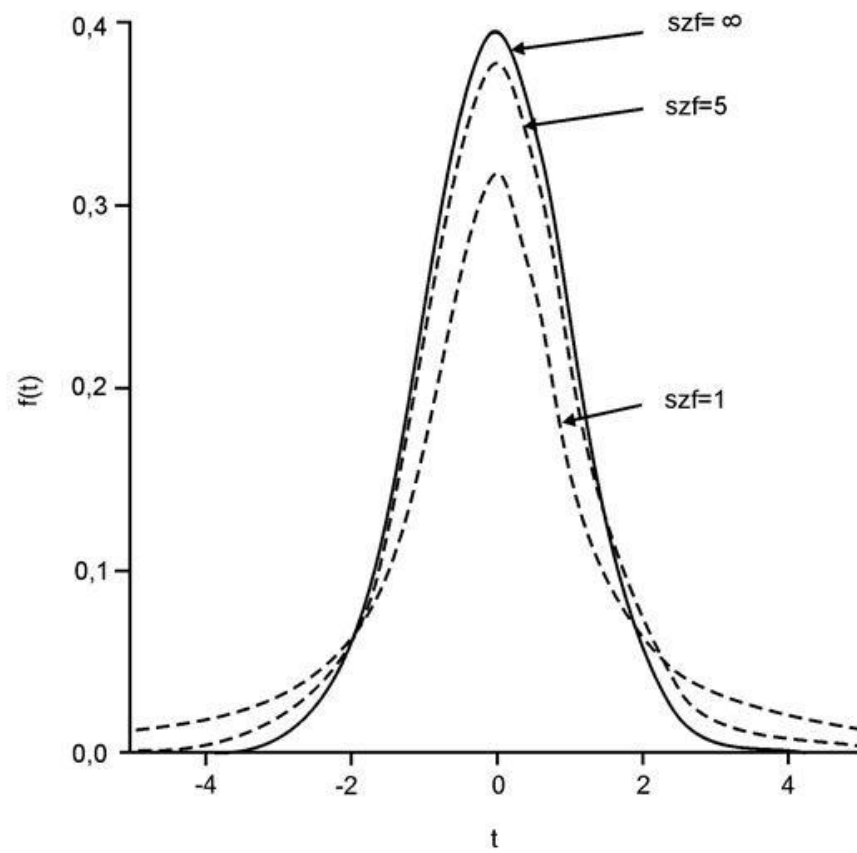
$$u = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

- standard normális eloszlást követ.

- Ha nem ismert a szórás?
- W. S. Gossett statisztikus és „Student” álnéven, 1908

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





Egymintás t-próba

- Adott populációátlag lehetséges-e a a minta alapján?
- Azt ellenőrzi, hogy a változó középértéke különbözik-e egy adott próbaértéktől?

$$t = \frac{\bar{X} - \mu}{s / \sqrt{n}}$$

- H_0 : A populáció átlaga egyenlő egy feltételezett értékkel

- `t.test(adat$iq, mu=120)`