A középérték és variancia azonosságának próbái: t-próba, F-próba

Háttérolvasmány:

Hipotézisállítás és -tesztelés: Reiczigel et al. 151–174. oldal.

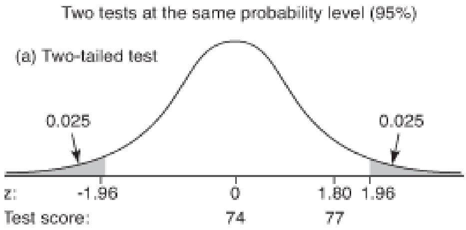
t-próba: Reiczigel et al. 194–202. oldal.

Hipotézisállítás

- A kísérlet alapjául szolgáló feltételezés: *a* minta egy adott szempont alapján más populációhoz tartozik, mint *b* minta.
- Ennek alapján megfogalmazzuk az alternatív vagy ellenhipotézist (H_1): egy bizonyos valószínűséggel állítható, hogy *b* minta átlaga nem ugyanahhoz a populációhoz tartozik, mint az *a* minta.
- Nullhipotézis (H_0): *a* minta és *b* minta egyazon populációhoz tartozik, azaz az átlaguk ugyanazon μ populációátlag körül szór.
- A nullhipotézist elutasíthatjuk, ha nagyon valószínűtlen, hogy a két minta ugyanahhoz a populációhoz tartozik.

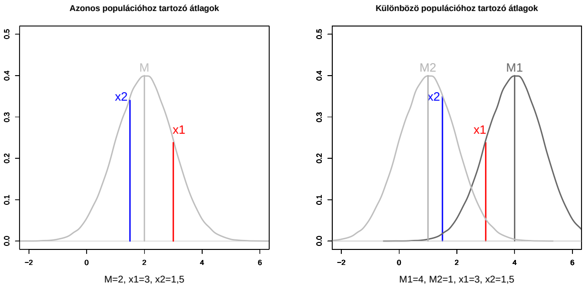
Hipotézis tesztelése $p = 95\%$-os megbízhatósággal

A humántudományokban a nullhipotézis valószínűségének szintjét 95%-nál húzzák meg. A 95%-os konfidenciaintervallumon kívül eső tartományt, a maradék 5%-ot α-ként definiálják.
Leggyakoribb nullhipotézis: a két minta NEM KÜLÖNBÖZIK.
Ehhez kétoldalú tesztek tartoznak, azaz α eloszlik a vizsgált minta alsó és felső két széle között, azaz mindkét oldalon 2,5%-ot fed le.
H_1: *a* nagy valószínűséggel **eltér** *b*-től.
H_0: *a* és *b* ugyanazon populáció része. Elutasítás: ha $\bar{x}$ a sűrűségfüggvény két szélén $\alpha/2$-be esik $\Rightarrow$ kétoldalú teszt.



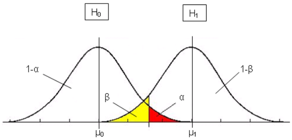
A nullhipotézisen alapuló eljárások hátulütői

- A 95%-os bizonyosság nem jelent teljes bizonyosságot!
- Vagyis: ha elutasítjuk a nullhipotézist, attól lehet igaz, az átlagok tartozhatnak egyazon populációhoz.
- És: lehet, hogy megtartjuk a nullhipotézist, pedig az átlagok két külön populációhoz tartoznak, csak átfednek egymással.



Hibatípusok

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

Összehasonlítás alapjai

- Átlagok,
- szórások,
- minta populációval $\leftrightarrow$ minta mintával,
- azonos varianciák $\leftrightarrow$ eltérő varianciák,
- független $\leftrightarrow$ párosított minták,
- parametrikus $\leftrightarrow$ ordinális vagy nem normális eloszlású minták.

Ha a populáció σ szórása ismert: átlagok z-eloszlás szerint szórnak μ körül.

A legtöbb esetben nem ismerjük a populáció szórását, ezért a mintaátlag szórását a Student-féle t eloszlással jellemezzük.

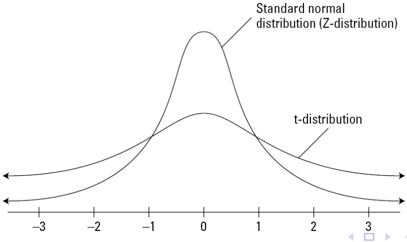
$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

$\sim$ normalizálás a z értékre, de σ helyett s.

t-eloszlás

Jellemzők:

- Szimmetrikus, átlaga 0, aszimptotikus, de nem normális eloszlású.
- Függ a minta méretétől, n-től.
- A t-eloszlás laposabb, mint $z \Rightarrow$ adott szignifikanciaszint határértékei messzebb esnek az átlagtól. Vagyis: egy adott átlag nagyobb valószínűséggel tartozik a vizsgált populációhoz, azaz a nullhipotézist nehezebb lesz elutasítani.
- $n = \infty$ esetén t eloszlás azonos z eloszlással.
- $n \geq 100$ esetén a különbség elhanyagolható, z-értékeket lehet használni.



Szabadsági fokok

Szabadsági fok, *degree of freedom, df*: a szabadon változtatható elemek száma, ami mellett a minta egy adott tulajdonsága változatlan marad. A t-eloszlás meredekségét a szabadsági fokok száma határozza meg.

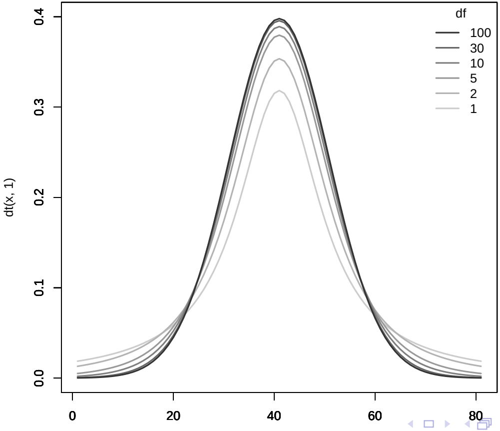
Pl. egy $n = 5$ elemű minta átlaga $\bar{x} = 10$. Hány elem változtatható szabadon a mintaátlag változatlansága mellett?

Négy, hiszen az ötödik elemet úgy kell kiválasztani, hogy a minta átlaga 10 maradjon, tehát csak négy elem változtatható szabadon.

Tehát $df = n - 1$.

t-eloszlás és szabadsági fokok

A *t*-eloszlás lapossága függ a szabadsági fokoktól. Minél nagyobb a szabadsági fok, annál közelebb esik a kritikus érték (= szignifikanciahatár, konfidenciaintervallum szélső értéke) az átlaghoz. Vagyis: meredekebb görbe esetén egy bizonyos átlag nagyobb valószínűséggel fog kívül esni a szignifikanciaküszöbön.



Egymintás Student-féle t-próba

- Feltétel: normális eloszlású változó.
- Alkalmazás: populáció vagy nagyszámú referenciaminta átlaga ismert. Ilyen például az IQ $= 100$, aminek 100-ban van rögzítve az átlaga.
- Eljárás: ha $t_{minta} > t_{1-\alpha(n-1)} \Rightarrow H_0$ elvetése. Vagyis: kiszámítjuk a minta t-értékét, és összevetjük az adott mintához tartozó t-eloszlás küszöbértékével. Ha a minta t-értéke abszolút értékben nagyobb, akkor elvetjük a nullhipotézist.

Példa

A Kincskereső óvodába 60 okos és ügyes gyerek jár. Átlagos IQ-juk 108, a szórás 10. Okosabbak-e az oda járó gyerekek az átlagnál?

$$t_{minta} = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} = \frac{108-100}{10/\sqrt{60}} = \frac{8}{1{,}29} = 6,2$$

Kritikus értékhez tartozó t meghatározása ($p = 1 - \alpha = 0,95$): adott kvantilishez (0,95) tartozó t-érték 59-es szabadsági fok mellett:

`qt(p,df)`, itt: `qt(0.975,59)` $\rightarrow$ 2,000995

vagyis egy 59-es szabadsági fokú mintához kétoldalú teszt esetén, ahol α az eloszlás két szélén levő 2,5%, nagyjából $t = 2$ a szignifikanciaküszöb.
Ennél a minta t értéke jóval nagyobb, ezért H_0-t elutasítjuk.

A Kincskereső óvodába járó gyerekek IQ-ja tehát az átlagtól szignifikánsan különbözik. Mivel t értéke pozitív, a minta átlagértéke magasabb, mint a rögzített IQ-átlagé.

Kétmintás független t-próba

- Két minta alapján két ismeretlen μ értéket hasonlítunk össze.
- Minták kiválasztása egymástól független (pl. spanyol óvodások és cseh óvodások).
- Feltétel: normális eloszlás, azonos varianciák.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s}{n_1} + \frac{s}{n_2}}}$$

ahol a szórás mindkét mintában azonos: a közös variancia becslése a mintánkénti szórásokból.

DE: a szórások azonosságát ritkán állíthatjuk biztosan!

Welch-próba

Mint a kétmintás független t-próba, de nem feltételezzük a varianciák egyenlőségét.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

A gyakorlatban többnyire a Welch-próbát alkalmazzuk, ha nem akarunk a varianciák homogenitásának (azonosságának) tesztelésével foglalkozni.

Példa

Eltér-e az alábbi mintában a nőstény és hím borjak, bikák és üszők születéskor mért testtömege?

bika (kg)	46	37	39	37	33	48	35		
üsző (kg)	27	37	35	41	35	34	43	38	40

Vektor létrehozása az R-ben:

```
bika = c(46,37,39,37,33,48,35)
uszo = c(27,37,35,41,35,34,43,38,40)
```

Példa

Normális eloszlásúak-e a minták?
`shapiro.test(bika), shapiro.test(uszo)`

Eredmény értelmezése: ha NEM igaz, hogy a minták normális eloszlásúak, p kisebb, mint 0,05. Ha p az adott szignifikanciaszintnél nagyobb, elfogadjuk a normális eloszlás feltételezését.

Két minta összehasonlítása t-próbával:
`t.test(bika,uszo)`
A függvény alapértelmezett paraméterei: kétoldalú teszt (alternative=two.sided), varianciák nem egyenlőek (var.equal=FALSE).

$p > 0,05 \Rightarrow$ különbség nem szignifikáns.

Kétmintás páros t-próba

A minta egyazon elem vagy összetartozó elemek kétszeri megfigyeléséből áll. Vagyis egyazon elemen kétféle dolgot mérünk, pl. egy személy vérnyomása gyógyszer szedése előtt és hat héttel később.

Feltétel: egy elem két értékének különbsége normális eloszlású.

$$t = \frac{\bar{d}}{s_d/\sqrt{n}}$$

ahol $\bar{d}$ a különbségek átlaga, s_d a különbségek becsült szórása, n a párok száma (tehát az elemszám, nem a mérések száma).

A kétmintás páros t-próba elve hasonlít az egymintás t-próbához. Itt nem egyetlen minta átlagát hasonlítjuk össze egy ismert populációéval, mint az IQ-nál, hanem azt teszteljük, hogy az elemek között mért különbségek szignifikánsan eltérnek-e 0-tól.

Példa

`ratings` adatmátrix.
Növények és állatok nevének gyakorisága és ismertsége (Frequency, meanFamiliarity) egy 1-től 9-ig terjedő skálán megadva.
Különbözik-e a gyakorisági mutató és az ismertség páronként?

Normális eloszlás tesztelése:

```
shapiro.test(ratings$Frequency)
shapiro.test(ratings$meanFamiliarity)
```

Páros t-próba:

```
t.test(ratings$Frequency,ratings$meanFamiliarity,
paired=T)
```

$p \ll 0,001$, tehát a megkérdezettek az állatok és növények méretét szignifikánsan nagyobbra becslik egy adott skálán, mint a súlyukat.

Varianciára vonatkozó próbák

Tesztek és feltételeik (legalább) két minta esetén:

- *F*-**próba:** mindkét mintában normális eloszlás, független minták. R: `var.test()`.
- **Levene-próba:** közelítő próba, de normális eloszlás hiányában is használható, több mintára is. R: `leveneTest` a `car` könyvtárban.
- **Bartlett-próba:** normális eloszlás, páros mintákra is használható. R: `bartlett.test()`.

Csomag telepítése: `install.packages("car"), library(car)`

Feladat 1

Töltsük le a trans.RData fájlt innen:
clara.nytud.hu/~mady/courses/statistics/materials/trans.RData
Letöltés `load("konyvtar/trans.RData")` függvénnyel (NEM `read.table()`).

A mátrixban angol, ill. portugál, kb. 1500 szavas szövegek hossza van megadva, majd a másik nyelre való lefordítás utáni hosszuk.

A. Ellenőrizzük, azonos-e az angol és portugál szövegek szószámának varianciája. Ehhez először ellenőriznünk kell, hogy az angol és a portugál szövegek hosszának eloszlása normális-e.

Milyen teszteket kell elvégeznünk?

Feladat 2

A. Hasonlítsuk össze a bikák és üszők születési súlyának varianciáit a fenti három teszttel. Milyen különbségeket látunk az eredményekben?

Teszteljük t-próbával:

B. Igaz-e, hogy a ratings mátrixban szereplő állatnevek gyakorisága alacsonyabb mértékű, mint a növényeké? És az ismertségük?

C. Milyen értéket vesz fel p, ha a növények és állatok súly- és méretbecslését külön-külön vizsgáljuk?

Teszteljük minden esetben, hogy az adatok normális eloszlásúak-e, és hogy a varianciák homogének (egyenlőek)-e.