

T -próbák és nem-parametrikus megfelelőik

Ismétlés a Statisztika I kurzusról

Gyakori összehasonlítások, amiket statisztikai módszerekkel végzünk:

- ▶ valamilyen szempontból különböző csoportok összevetése,
- ▶ eltérő faktorok hatása ugyanarra az egyedre,
- ▶ adott csoport összevetése egy ismert irányadó értékkel.

Példa: nyelvi teljesítményt mérő teszt eredménye különböző anyanyelvű csoportoknál, vagy egyazon csoportban különböző időpontokban, vagy egy konkrét csoport teljesítménye az ismert átlaghoz képest.

Klasszikus eset: kétmintás független t -próba.

Háttér: gyakran feltételezik, hogy vannak könnyebben és nehezebben megtanulható nyelvek. A kínait például sokan nehéznek tartják. Egyes nyelvészek szerint nincsenek könnyű és nehéz nyelvek, inkább az számít, hogy az idegen nyelv szerkezete hasonlít-e az anyanyelvhez vagy más korábban megtanult idegen nyelvekhez.

Kísérlet: két éve kínaiul tanuló, angol és vietnámi nyelvtanulók teljesítményét hasonlítják össze egy nyelvi teszten (maximum: 100 pont). Adatok:

`phon.nytud.hu/mady/courses/statistics/materials/chinese_test.csv`

Nullhipotézis: a két anyanyelvi csoport teljesítménye azonos, nem különbözik szignifikánsan.

Kutatási hipotézis: a vietnámi csoport az anyanyelvük közelsége miatt két év alatt magasabb szintre jut el, jobb lesz a teszteredménye.

Adatok elemzése

A .csv sima szöveges formátum (comma-separated values). A letöltött fájlt szövegfeldolgozó szoftverben lehet megnézni (Jegyzettömb, Notepad++, Linuxon `less` parancs stb.).

A beolvasáshoz fontos kérdések: (1) az első sorban adatok vagy oszlopnevek vannak, (2) mi választja el az oszlopokat egymástól.

```
chinese = read.csv("chinese_test.csv")
```

?read.csv vagy help(read.csv) megmutatja, hogy az oszlopnevek jelenléte és a vessző mint elválasztó a függvény paramétereinek default beállítása.

Ábrázolás dobozdiagrammal:

```
boxplot(score~L1, data = Chinese)
```

t -próba futtatása

```
t.test(score~L1, Chinese)
```

Lényeges információk: t abszolút értéke és iránya, p -érték, csoportok szerinti átlagok. Ha t **negatív**, akkor az **első** (kiindulási) csoport átlaga alacsonyabb, mint a másodiké.

Ha a két csoport között varianciahomogenitás áll fenn, a p -érték pontosabb.

Ellenőrzés: `var.equal()`.

A varianciák azonossága a nullhipotézis. Ha $p > 0,05$, a varianciák azonosságát feltételezzük.

Módosított függvény:

```
t.test(score~L1, Chinese, var.equal = T)
```

Ábrázoljuk a vietnámi és angol anyanyelvű tanulók eredményeit sűrűségfüggvényen.

Szükséges lépések:

- ▶ logikai vektor létrehozása az egyik csoport (pl. vietnámi) adatsoraira,
- ▶ minimum és maximum pontszám definiálása az x tengelyen,
- ▶ y tengely minimum és maximum értékének rögzítése,
- ▶ vietnámi csoport értékeinek ábrázolása zöld színnel,
- ▶ ugyanabban az ábrában az angol csoport értékeinek ábrázolása kékkel.

A függvények a `ttest_wilcoxon_R.txt` szöveges fájlban találhatóak a félév anyagainak könyvtárában.

A nyelvi teszten a résztvevők átlagosan 50 pontot érnek el.

Igaz-e, hogy a vietnámi tanulók a két nyelv hasonló szerkezete miatt átlagon felül teljesítenek?

Eljárás: **egymintás t -próba**. Itt nem két csoportot hasonlítunk össze, hanem egy csoport pontszámait az átlaggal. A vietnámi csoportra már létrehoztuk a viet logikai vektort, ezt használjuk most.

```
t.test(Chinese$score[viet], mu=50)
```

Fontos, hogy itt más a `t.test()` függvény szintaxisa, mint a korábbi dián. Itt egyetlen mintát (50 vietnámi L1-hez tartozó pontszámot) adunk meg, és a populáció átlagával, μ -vel vetjük össze az átlagukat.

Az angol anyanyelvű tanulók csoportja azt állítja, hogy a teszt olyan résztvevőknek kedvez, akik az ázsiai oktatási rendszerben tanulnak. Szerintük a kínai tudásuk szintje nem marad el a vietnámiakétól, ezért egy másik tesztet akarnak kitölteni.

Eljárás: páros t -próba alkalmazása. Nem a két csoport eredményeit vetjük össze, hanem azt vizsgáljuk, hogy az egyes résztvevők pontszámának különbsége szignifikánsan eltér-e nullától.

A két teszt adatai a `chinese_english_test.csv` fájlban találhatóak, a "test1" és "test2" oszlopban. Töltsük be az adatmátrixot a `Chinese.english` változóba.

```
t.test(Chinese.english$test1, Chinese.english$test2,  
paired = T)
```

Tényleg szignifikánsan jobb teljesítményt érnek el a hallgatók a másik teszten?

Tegyük fel, hogy a csoportok teljesítményét nem egy 0 és 100 pont közötti skálán adják meg, hanem öt fokozatban: "beginner (1)", "developing (2)", "intermediate (3)", "advanced (4)", "excellent(5)".

Ez a skála nem parametrikus, mert nem igaz rá, hogy az értékek közötti távolságok azonosak (ekvidisztánsak) lennének.

Az összehasonlítást nem-parametrikus próbákkal végezhetjük el.

- ▶ Két független minta: Mann-Whitney próba.
- ▶ Egy minta ismert **medián** mellett: egymintás Wilcoxon-próba.
- ▶ Páros összehasonlítás: páros Wilcoxon-próba.

A tesztek függvénye egyaránt `wilcoxcon.test()`.

Ellenőrizzük nemparametrikus teszttel, hogy a kategóriákban összefoglalt eredmények alapján szignifikáns-e a különbség a két csoport eredményei között.

Eredmények: `chinese_test_categories.csv`

```
wilcox.test(category~L1, Chinese.cat)
```

A teszt medián eredménye három pont, azaz "intermediate" minősítés.

Hogyan viszonyul ehhez a vietnámi és az angol anyanyelvű csoportok eredménye?