

Ismételt méréses varianciaanalízis

Ismételt méréses módszerek

Humán tudományok örök problémája: egy személytől általában nem egy, hanem többféle adatot gyűjtünk. Ennek elemzésére az egyszerű varianciaanalízis NEM alkalmas, mert ott alapfeltétel a minták függetlensége (mint a független mintás t -próba esetén).

A varianciaanalízis függő mintás megfelelője az **ismételt méréses** varianciaanalízis, angolul *repeated measures ANOVA*.

Fontos: az ismételt mérés nem arra vonatkozik, hogy egyazon beszélőtől többször vesszük fel ugyanazt az adatot (pl. mondatokat öt ismétléssel olvasnak fel), hanem hogy **egyazon személlyel** több szempontú mérést végzünk.

Például orvostudományban: egy bizonyos gyógyszer hatása kezelés előtt, a kezelés megkezdése után két héttel, egy hónappal stb.

Eljárás

Egy függő és egy vagy több független változó tesztelése, ahol az ismétlés **belső** tényezői (személy vagy tárgy, akiken/amiken az ismételt méréseket végeztük) közötti különbséget **véletlen** hatásnak, nem kívánatos varianciának tekintjük (*within-subjects factor*). Az alanyok tartozhatnak két vagy több különböző csoporthoz, (különböző nyelvek beszélői, egy növényfaj különböző fajtái stb.), ez a **köztes** tényező (*between-subjects factor*).

Feltételek:

- ▶ legalább öt alany (személy, növény, tárgy stb.),
- ▶ faktorkombinációnként egyetlen adat - azaz ha egyazon faktort egy alagnál több ismétléssel mértünk, ezeket a teszt futtatása előtt átlagolni kell minden egyes alanyra és cellára,
- ▶ kiegyensúlyozott dizájn, azaz ha az egyik faktor két szintjéhez két további faktor tartozik, akkor a másik faktornál is vizsgálni kell ugyanezt a két szintet akkor is, ha a kutatási kérdés szempontjából nem releváns.

Hátulütők

- ▶ R-ben nincs több faktor kombinációjára átlagoló beépített függvény,
- ▶ mivel átlagokkal számolunk, az egyes cellákon belüli varianciát nem tudjuk figyelembe venni,
- ▶ nem tudunk több *within-subject* tényezőt kombinálni (pl. résztvevő és többféle mondat egyazon kategóriából),
- ▶ csak akkor alkalmazható, ha a páronként összehasonlított faktorok különbségeinek varianciái azonosak (szfericitási feltétel),
- ▶ nincs post-hoc tesztje, csak t -próbák Bonferroni-korrekktúrával (konfidenciaszint/összes lehetséges kombináció száma).

Ezekre a kevert modellek jelentenek majd kiutat, amik cserében számos új problémát libbentenek fel.

Cellánkénti átlagok számítása

`anova.mean.r` nevű R-függvény letöltése innen:
phon.nytud.hu/mady/courses/statistics/materials

Szkript és függvény közötti különbség: függvényben létrehozott változók (R-objektumok) nem jelennek meg automatikusan a munkamemóriában. Szkript és függvény egyaránt betölthető a `source("eleresiutvonal")` paranccsal, a szkriptet közvetlenül be is lehet másolni egy szövegszerkesztőből az R-be (copy-paste).

A betöltött függvények és csomagok nevét a `search()` függvénnyel lehet lekérdezni.

Függvény első sora:

```
fuggvenynev = function(kotelezoargumentum1,  
kotelezoargumentum2, ...), ahol három pont további  
opcionális számú opcionális argumentumot jelöl.
```

Példa

Mondatvégi kétszótagú, /s/-re és /z/-re végződő szavakban megmértük a frikatíván belüli zöngés tartomány hosszát. Zöngésebbek-e a mondatvégi /z/-k, mint az /s/-ek?

zfin.RData, letöltés innen:

phon.nytud.hu/mady/courses/statistics/materials

```
zmean = anova.mean(zfin$cvoice,zfin$subj,zfin$voiced)
```

Kapott adatmátrix oszlopainak elnevezése:

```
names(zmean) = c("cvoice","subj","voiced")
```

Ismételt méréses varianciaanalízis függvénye

- ▶ Függő változó: mássalhangzó zöngességének tartama (cvoice).
- ▶ Független változó: zöngesség (voiced).
- ▶ Within-subjects factor: beszélő (subj).
- ▶ Between-subjects factor: nincs, mert a beszélők nincsenek csoportokra osztva.

```
h = summary(aov(cvoice ~ voiced +  
Error(subj/voiced), data=zmean))
```

Releváns p -érték: Error: subj:voiced sor alatt (ez jelzi az alanyok szerinti interakciót).

Ábrázolás:

```
interaction.plot(x-tengely, ismételt_mérés_alanya,  
paraméter)  
interaction.plot(zmean$voiced, zmean$subj, zmean$cvoice)
```

Több tényező

Többszörös varianciaanalízis képlete, pl. ha megelőző mássalhangzóra is kíváncsiak vagyunk:

```
summary(aov(cvoice ~ voiced*c1 +  
Error(subj/(voiced*c1)), data=zmean))
```

Ehhez a cellánkénti átlagokat újra kell számolni:

```
zmean = anova.mean(zfin$cvoice, zfin$subj,  
zfin$voiced, zfin$c1)
```


Eredmények

Értelmezés:

Error: subj:voiced zöngésségi tartamok beszélőnként, zöngésség függvényében (a p -érték változott, mert az átlagokat újraszámoltuk).

Error: subj:c1 zöngésségi tartamok beszélőnként, a megelőző mássalhangzó függvényében.

Error: subj:voiced:c1 zöngésségi tartamok beszélőnként, zöngésség és megelőző mássalhangzó interakciója, azaz befolyásolja-e a megelőző mássalhangzó a zöngésség hatását.

Köztes tényező (between-subjects factor)

Bodo Winter példája: alapfrekvencia az udvariasság függvényében (inf=informal, pol=polite), férfiaknál és nőknél.

Letölthető innen:

<https://bodo-winter.net/tutorials.html>
dataset for tutorial 2

Betöltés legegyszerűbb `read.csv` függvénnyel, mert ott a vessző az alapértelmezett cellaelválasztó jel. Mentés a `pol` objektumba.

Figyelem! A nem Mac-felhasználóknak gondjai lehetnek a betöltéssel a Mac-specifikus sortörésjelek miatt (ez a Windows és a Linux txt-formátuma között is jelenthet nehézséget).

Megoldás: megnyitás szövegmegjelenítővel (notepad++, gedit vagy más), és mentés a saját operációs rendszer kódolása szerinti formátumban. Vagy letöltés innen:

<https://phon.nytud.hu/mady/courses/statistics/materials/politeness.RData>

Értelmezzétek a kapott eredményeket!