

# Ismételt méréses multifaktoriális varianciaanalízis (*repeated measures MANOVA*)

# Varianciaanalízis

Adott egy parametrikus függő változó és egy vagy több kategoriális független változó. Pl. függő változó: vércukorszint, kategoriális (= nominális) változók: (1) kezelés (inzulin, placebo), (2) napszak (reggel, délben, este),

Vizsgált kérdések: (1) van-e különbség a csoportok átlaga között ( $t$ -próba általánosítása), (2) van-e hatása a vizsgált tényezőnek (regressziószámítás: magyarázó változók hatása a függő változóra).

Előfeltétel: minta normális eloszlása és csoportokon belüli varianciák homogenitása (nagyjából megegyezése). Tesztek: Kolmogorov-Smirnov-próba, Levene-próba stb.

## Varianciaanalízis típusai: egyváltozós próbák

- ▶ **Egytényezős varianciaanalízis (one-way ANOVA):** egy függő változó (vércukorszint), egy független változó (kezelés): 200 kísérleti személy (ksz) vércukorszintje, a fele inzulint kap, a másik fele placebót.
- ▶ **Kéttényezős varianciaanalízis (two-way ANOVA):** egy függő változó (vércukorszint), két független változó (kezelés, napszak): 600 kísérleti személy, 300 inzulint kap, 300 placebót, ezek közül csoportonként 100 ksz értékeit reggel mérik, 100-ét délben, 100-ét este.
- ▶ **Többtényezős varianciaanalízis ( $n$ -way ANOVA):** egy függő változó (vércukorszint),  $n$  független változó (vércukorszint, napszak, nem, korosztály, bőrszín stb.). Minden csoportban 100-100 **eltérő** személy vércukorszintjét mérik.

## Varianciaanalízis típusai: egyváltozós, ismételt méréses tesztek

Nem tudunk/akarunk *faktorkombináció* \* 100 kísérleti személyt felhajtani, ezért a kísérleti személyeket reggel, délben és este megvizsgáljuk  $\Rightarrow$  kéttényezős dizájnban 600 helyett elég 200 ksz.

- ▶ **Belső tényező (within-subjects factor):** *egy személyen belül* napszakonként több mérés.
- ▶ **Köztes tényező (between-subjects factor):** csoportok tagjainak összehasonlítása, azaz inzulinnal, ill. placebóval kezelt csoport.

További előfeltétel: szfericitás, azaz a feltételek függetlensége. Bármely két feltétel közötti összefüggésnek azonosnak kell lennie bármely másik két feltétel közötti összefüggéssel, pl. reggel és délben mért vércukorszintek különbségeinek varianciája azonos délben és este, valamint este és reggel mért vércukorszintek különbségeinek varianciájával.)

# Ismételt méréses többváltozós varianciaanalízis

Teszt Mauchly-próbával, korrekció Greenhouse-Geisser-próba segítségével (SPSS-felhasználók előnyben).

Elegánsabb megoldás: *többváltozós* varianciaanalízis. Itt az egyik függő változó továbbra is a vércukorszint, a másik pedig a különböző szintek közötti kovariancia.

R-ben szükséglet: `car` csomag, plusz két további szkript (`anova.mean.r` és `Anova.prepare.r`), letölthető innen: [clara.nytud.hu/~mady/courses/statistics/materials](http://clara.nytud.hu/~mady/courses/statistics/materials)

Emlékeztetőül: az ismételt mérés **nem azt jelenti, hogy egyazon embertől** tízszer veszünk vért reggel, egyazon beszélővel ötször mondatjuk el ugyanazt a mondatot stb. Az ilyen ismétlés nem kerül bele az ismételt méréses varianciaanalízisbe, ehelyett az ismétlések átlagát vesszük.

## 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:

[clara.nytud.hu/~mady/courses/statistics/materials](http://clara.nytud.hu/~mady/courses/statistics/materials)

Egyazon beszélő azonos feltételen belüli ismételt felolvasásainak átlaga:

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

Kapott adatmátrix oszlopainak elnevezése:

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

# Ismételt méréses ANOVA

```
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)
```

# Ismételt mérés MANOVA

R-kód

```
library(car)
code = c("d", "s", "w")
d = zmean[,c("cvoice", "subj", "voiced")]
d.t = Anova.prepare(d, code)
d.lm = lm(d.t$d~1)
Anova(d.lm, idata=d.t$w, idesign=~voiced)
```

code: *d*: függő változó (dependent variable), *s*: ismételt mérés tárgya (subject), *w*: belső tényező (within-subjects factor), *b*: csoportok közötti tényező (between-subjects factor).

Az `Anova.prepare` függvény **kizárólag** a *d*, *s*, *w* és *b* változókat tudja értelmezni! Csak a *w* változóból lehet több, ha több tényezőnk van.



# Többtényezős MANOVA

Többtényezős ismételt mérésű többváltozós varianciaanalízis képlete, ha nincs *between subjects factor*, pl. ha megelőző mássalhangzóra is kíváncsiak vagyunk.

Először a cellánkénti átlagokat újra kell számolni:

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

```
code = c("d", "s", "w", "w")  
d = zmean.c[,c("cvoice", "subj", "voiced", "c1")]  
d.t = Anova.prepare(d, code)  
d.lm = lm(d.t$d~1)  
Anova(d.lm, idata=d.t$w, idesign=~voiced*c1)
```

Megegyeznek az eredmények az ismételt mérésű ANOVA által kiadottal?

## Ismételt méréses MANOVA, több csoport

Hogyan hat a tág, szűk és kontrasztív fókusz a fókuszban levő szó hangsúlyos szótagjának tartamára, és a megelőző topik hangsúlyos szótagjának tartamára? Hatással van-e az életkor a hangsúlyos szótag tartamára (age változó)?

Letöltés: `accdur.RData`

hangsúlyos szótag tartama: `dur` változó, milliszekundumokban.

Először egy mérések átlagát tartalmazó táblázatot kell létrehozni az `anova.mean` függvénnnyel (`accmean`).

```
code = c("d", "s", "w", "b")
d = accmean[,c("dur", "subj", "focus", "age")]
d.t = Anova.prepare(d, code)
d.lm = lm(d.t$d~factor(d.t$b))
Anova(d.lm, idata=d.t$w, idesign=~focus)
```

Probléma, mint RM-ANOVA esetén: nincs post-hoc tesztje.  
Helyette: páronkénti összehasonlítás.

Itt probléma: a figyelmen kívül hagyott faktor szintjét törölni kell a

```
levels(valtozo)[n] = NA
```

függvénnyel.

Ehhez érdemes a minket érdeklő sorokat egy új data.frame-be írni, és ebben törölni a felesleges faktorszintet, így az eredeti, minden adatot tartalmazó adatmátrixunk változatlan marad.

Kódot lásd az rm\_manova.txt fájlban.