

Kevert modellek

Ismételt méréses varianciaanalízis

A nyelvészeti kísérletekben egy személytől szinte mindig többféle információt szokás begyűjteni → ismételt méréses módszerek.

Ismert módszer az ismételt méréses ANOVA, ahol a független változók közötti összefüggést egy *within-subjects* faktoron, azaz belső tényezőn belül vizsgáljuk: futóteljesítmény reggel, délben és este egyazon személynél mérve.

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 teljesítmény különbségeinek varianciája azonos a délben és este, valamint a reggel és este mért teljesítmények különbségeinek varianciájával.

Sok kutatási kérdésnél eleve nem várjuk a különbségek varianciájának azonosságát, például a kontrollfeltételként használt tényezőknél.

Szfericitást nem elváró alternatívák: ismételt méréses MANOVA (többváltozós ANOVA), kevert modellek.

A kevert modellek előnyei:

- ▶ egynél több belső tényező (pl. kísérleti személy és stimulus),
- ▶ ordinális adatok (pl. Likert-skála pontszámai),
- ▶ nem normális eloszlású adatok,
- ▶ üres cellák (pl. nincs minden faktorkombinációra adat, néhány kísérleti személy nem töltötte ki az utolsó oldalt a tesztalapon, stb.),
- ▶ nem kell cellaátlagokat számolni, mint az ismételt méréses (M)ANOVÁ-nál.

A kevert modellek hátrányai:

- ▶ Új, folyamatos fejlesztés alatt álló módszer.
- ▶ Kiszámú adatra nem megbízható (legalább 200 adatnak illik lennie).
- ▶ A modell nem tartalmaz szabadsági fokokat, ezért az eredmények nem feleltethetők meg egyértelműen p -értékeknek.
- ▶ A modell nem minden esetben konvergál, azaz bizonyos függvényekre nem ad semmilyen eredményt.
- ▶ Módszertani káosz a felhasználói oldalon.

Előny és hátrány: nem konzervatív eljárás, azaz nagyobb eséllyel talál szignifikáns különbséget, mint a klasszikus módszerek, pl. ANOVA.

Leírások:

Baayen, Harald (2008): Analyzing linguistic data. Cambridge: UP.
<http://www.sfs.uni-tuebingen.de/~hbaayen/publications/baayenCUPstats.pdf>

Field, Miles & Field (2012): Discovering statistics with R. London et al.: SAGE.

Winter, Bodo (2013): Linear models and linear mixed effect models in R with linguistic applications.
http://bodowinter.com/tutorial/bw_LME_tutorial2.pdf

Bodo Winter példája: alapfrekvencia az udvariasság függvényében, férfiaknál és nőknél.

Letölthető innen:

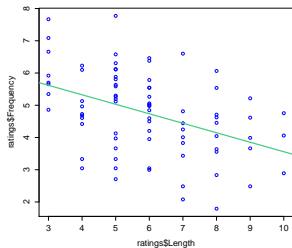
<http://bodowinter.com/tutorials.html>

dataset for tutorial 2

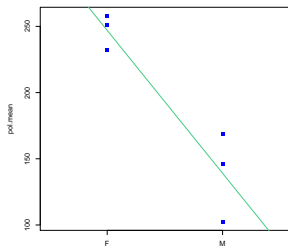
Betöltés legegyszerűbb `read.csv` függvénnyel, mert ott a vessző az alapértelmezett cellaelválasztó jel.

Objektum neve legyen *pol*.

Metszéspont jelentősége a lineáris modellekben: lineáris regresszió



Itt: a három női (F1-3) és három férfi (M1-3) beszélő átlagos alapfrekvencia-értékeire számolt regressziós egyenes.



Lineáris regresszió: metszéspont azonos a kiindulási csoport (itt: Female) faktoronkénti átlagával, estimate a másik csoport(ok) átlagával.

```
h = lm(frequency~gender,pol)
summary(h)
```

	Estimate
(Intercept)	246.986
genderM	-108.110

Az Intercept értéke a default (nulladik) faktorszint átlaga, a genderM a M(ale) szint ehhez képest mért különbségét jelenti – ez egyben az egyenes meredeksége.

A kevert modellek minden egyes alanyra (= random hatásként definiált egységre) külön metszéspontot számolnak. A lineáris regresszióval szemben, ahol a hiba, vagyis az ϵ nem játszott szerepet a modellben, itt a fix és random hatások keverékéből áll össze a modell.

Lineáris regresszió: $\text{frequency} \sim \text{gender} + \epsilon$

Kevert modell: $\text{frequency} \sim \text{gender} + (1|\text{subject}) + \epsilon$

Itt a hiba az egyes személyeken belüli varianciára vonatkozik.

1: intercept számítása.

Több független változó és random hatás esetén:

$\text{frequency} \sim \text{gender} + \text{attitude} + (1|\text{subject}) + (1|\text{scenario}) + \epsilon$

Modell előnyei: (1) a cellánkénti varianciát is figyelembe veszi (szemben a cellaátlagot elváró (M)ANOVÁ-val), (2) több random hatás is integrálható egy modellbe (RM ANOVÁ-ban csak egy).

Eljárás

Fix hatások (*fixed effects*): független változók, megismételhető adatok (ha akarunk, többet is gyűjthetünk belőlük).

Random vagy véletlen hatások (*random effects*): belső tényezők (*within-subjects factors*), véletlenszerűen kiválasztott személyek/ítemek, nem megismételhetőek, mert más személy esetén más random hatás érvényesülhet, és újabb faktorszintek jelennek meg.

Eljárás: fix hatások szembeállítása, feltételezés: véletlen hatások varianciája ismeretlen.

Képlet:

függőváltozó \sim függetlenváltozó + randomhatás + ϵ

ahol ϵ az *error*, hiba terminus a nem ellenőrizhető varianciára.

lme4 csomag letöltése: `install.packages("lme4")`
betöltés: `library(lme4)`.

```
modell = lmer(fuggovaltozo ~ fuggetlentaltozo +  
(1|randomhatas), data=adatmatrix)
```

Random hatás nélkül nem működik a modell, hiszen a *mixed-effects* elnevezés a fix és random hatások keverékére vonatkozik.

```
pol.mod = lmer(frequency ~  
attitude+(1|subject)+(1|scenario), pol)
```

Random effects: hatásokon belüli variancia és szórás.

Residual: egyik random hatás által sem magyarázott variabilitás.
A random hatások és a reziduálisok átlaga 0 (normalizált, azaz centrált adatok).

Fixed effects: estimate of intercept: *informal* (ABC-ben első) kategória f0-átlaga. *Polite* kategóriáé ennél $-19,695$ Hertz-cel alacsonyabb \rightarrow meredekség (*slope*).
t-érték: átlag/standard hiba

```
pol.mod = lmer(frequency ~  
attitude+gender+(1|subject)+(1|scenario), pol)
```

Estimate of intercept: a nők f0-ja informális stílusban (az ábécében elől álló kategóriák).

Döntés a hipotézisről

Ha szignifikanciahatár $p = 0.05$, és a hipotézis kétoldalú (nem tudjuk, hogy a különbség pozitív vagy negatív lesz), a $p = 0.025$ valószínűségi szinthez tartozó t értékre van szükségünk.

De: honnan tudjuk a szabadsági fokot? Sehonnan ☹

Különböző megoldásokat használnak:

(1) Adott szabadsági fokhoz tartozó t -érték magasabb szabadsági fok esetén már alig változik. Megoldás: szabadsági fokot 60-nak (más szerint 100-nak) vesszük, itt $t = 2$. Tehát 2-nél nagyobb t -érték fölött szignifikánsnak tekintjük a különbséget.

(2) Szabadsági fok meghatározása a megfigyelések száma alapján. Kétoldali teszt esetén: $2 * (1 - pt(abs(t), n - 2))$ ahol a $pt()$ függvénnyel kiszámoljuk az adott fix hatásra kapott t -értékhez tartozó p -értéket az elemszám–fix hatás paraméterszámát.

(3) Modellek összevetése restricted/relativised/residual maximum likelihood (REML) alapján.

Fix és random hatásokra egyaránt alkalmazható. Összehasonlítás alapja: egyszerűbb modell. Ha a bonyolultabb modell más becsléshez vezet, mint az egyszerűbb, akkor a plusz faktornak van hatása.

Példa Winter nyomán: felfutok egy hegyre adott idő alatt. Van nálam két liter víz és egy elemlámpa. Felfutok egy hegyre ezek nélkül, és látom, hogy így gyorsabb vagyok. Tesztelni akarom, hogy a vizesüveg vagy az elemlámpa miatt voltam lassabb.

Vagyis: lemérem a vízzel és elemlámpával futott időmet, majd eldobom az elemlámpát, és így is felfutok. Mindhárom eredményre számolok egy modellt, és összevetem, hogy különböznek-e. Ha így lassabb vagyok, a lámpa volt a ludas. (Tegyük fel, hogy nem fáradok.)

Nem (gender) és stílus (attitude) fix hatások befolyásának tesztelése:

Felállítjuk a modellt úgy, hogy csak a nem szerepel fix hatásként, és összevetjük a nem + stílusra felállított modellel.

```
pol.null = lmer(frequency ~  
gender+(1|subject)+(1|scenario),pol, REML=F)  
pol.mod = lmer(frequency ~  
attitude+gender+(1|subject)+(1|scenario),pol,REML=F)
```

A két modell összehasonlítása az egyszerűbbtől a bonyolultabb modell felé haladva:

```
anova(pol.null, pol.mod)
```

χ^2 -érték és hozzá tartozó p -érték.

Másik irányadó érték: AIC (Akaike's Information Criterion). Ha a két AIC-érték között kettőnél nagyobb a különbség, akkor a modellek szignifikánsan különböznek, vagyis a komplexebb modell jobb becslést ad.

Interakció tesztelése: interakció nélküli és interakciós modell összehasonlítása: `attitude+gender` és `attitude*gender`.

REML-ről alkotott vélemény sajnos nem egységes ☹. Van, akiknél random hatásokhoz REML=T javasolt, fix hatásokhoz REML=F. Van, aki szerint fordítva. Tartsunk kéznél referenciákat. Ebben a példában Winter a fix hatás tesztelésére REML=F beállítást használ.

(4) Valószínűségek szimulálása Anova függvénnyel a car csomagból.

```
Anova(pol.null)
```

Kimenet: varianciaanalízishez hasonló táblázat (summary(aov)).

Modell együttthatóinak elemzése

```
pol.mod = lmer(frequency attitude+gender+(1|subject)+  
(1|scenario),pol,REML=F)  
coef(pol.mod) vagy ranef(pol.mod)$subject (RANDOM  
EEffects)
```

Metszéspont minden személyre és minden scénárióra (itemre) különböző, de a meredekségek egyformák, vagyis azt feltételezzük, hogy a stílus hatása minden személyre és minden itemre azonos – *random intercept model*. Pedig feltehetően nem.

Helyette: *random slope model*

```
pol.mod =  
lmer(frequency attitude+gender+(1+attitude|subject)  
+(1+attitude|scenario),pol,REML=F)
```

Azaz: a modell különböző default-értékekből (intercept) és a stílus függvényében különböző válaszadási tendenciákból indul ki mindkét random hatás esetén.

`attitude` pol értékei mindig negatívak, vagyis udvarias stílusban minden személynél és minden item esetén alacsonyabb az f_0 . Stílus hatásának szignifikanciája ellenőrizhető a modellek összehasonlításával.

Ellenőrizhető az `interaction.plot()` függvényvel.

Alapvető módszertani probléma: a random intercept modellek antikonzervatívok, vagyis sok esetben mutatnak szignifikáns különbséget ott is, ahol nincsenek!

Adattisztogatás

Egyes vélemények szerint a nyilvánvalóan hibás adatpontokat ki lehet zárni az elemzésből, pl. véletlenül rossz gombnyomás, irreálisan hosszú vagy rövid (< 200 ms) reakcióidő (ld. Baayen 2008: 243ff.).

Egyénekenkénti adatpontok megszámolása:

`qqmath(~frequency|subject, pol)`

Nagyjából lineáris tendencia: normális eloszlás. Szélső értékek: kilógó pontok.

Egyes ajánlások szerint az alany vagy az item átlagától legalább 2 vagy 3 szigma távolságra eső pontokat is ki lehet hagyni az elemzésből. Baayen szerint szükségtelen.

Ismétlések hatása

Rátanulás vagy fáradás megjelenítése:

```
xyplot(frequency~scenario|subject,pol)
```

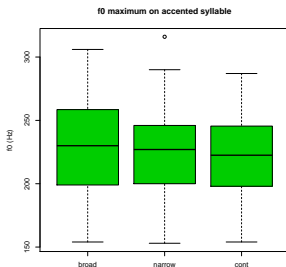
Változik-e az f_0 a kísérlet során egyazon kondícióon belül?

```
z = pol$attitude=="pol"
```

```
xyplot(frequency~scenario|subject,pol[z,])
```

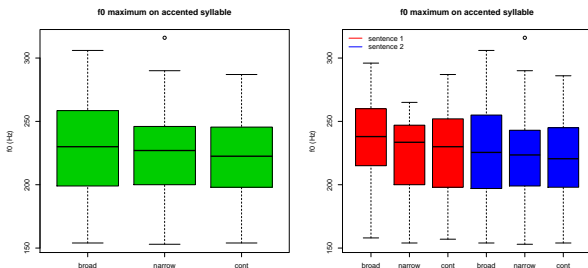
Gyakorló feladat

Alapfrekvencia maximum értéke fókuszbán levő szó hangsúlyos szótagán, két mondatban, hét beszélő felolvasásában.



Gyakorló feladat

Alapfrekvencia maximum értéke fókuszban levő szó hangsúlyos szótagán, két mondatban, hét beszélő felolvasásában.



Az első mondatban jól látszik a tendencia, a másodikban nem
↔ mondat is random hatás, azaz belső tényező.

Fenti példa: focusacc.Rdata

```
acc.lmer = lmer(f0max ~ focus + (1|subj) + (1|sent),  
data=adatmatrix)
```

Függő változó: f0 maximum, fixed effect: fókusztípus, random effect: beszélő és mondat.

Eredmények

Fixed effects:

	t value
focuscontrastive	-2.422
focusnarrow	-1.825

Faktorszintek automatikus sorrendezése alfanumerikusan, tehát sorrend: broad < contrastive < narrow. A kapott t értékek a broad vs. contrastive, broad vs. narrow összehasonlításra vonatkoznak.

Ha contrastive vs. narrow összehasonlításra vagyunk kíváncsiak:

```
focusacc$focus = relevel(focusacc$focus,  
"contrastive")
```

Ekkor "contrastive" kerül az első helyre, összehasonlítás ehhez képest.

```
focus = factor(focus,levels=c(" broad", " narrow", " contr"))  
vagy  
focus = factor(focus, levels = levels(focus)[c(2, 1, 3)])
```

További feladat az `accdur.RData` fájl újraelemzése: találjuk meg a maximálisan szükséges komplexitású kevert modellt a random intercept és a random slope modell alkalmazása esetén. Milyen különbségeket találunk?