Kevert modellek

Ismételt méréses varianciaanalízis

A nyelvészeti kísérletekben egy személytől szinte mindig többféle információt szokás begyűjteni $\rightarrow$ 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, itt: random hatás (pl. kísérleti személy és stimulus).
- Ordinális adatok (pl. Likert-skála pontszámai), ha parametrikusan értelmezhetőek.
- Nem normális eloszlású adatok, ha a reziduálisok eloszlása normális.
- Üres cellák (pl. nincs minden faktorkombinációra adat, néhány kísérleti személy nem töltötte ki az utolsó oldalt a tesztlapon, stb.).
- Legfontosabb: nem kell cellaátlagokat számolni, így az egyénen belüli varianciát is figyelembe veszi.

A kevert modellek hátrányai

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

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

A konzervatív eljárásokba be van építve az 1-es típusú (α) hiba kiküszöbölése. Ez egynél több páronkénti összehasonlításnál fontos.

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 (2014): A very basic tutorial for performing linear mixed effects analyses (Tutorial 2) https://bodo-winter.net/tutorials.html

Winter, Bodo (2020). Statistics for linguists: An introduction using R. New York & London: Routledge.

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

Letölthető innen:
https://bodo-winter.net/tutorials.html

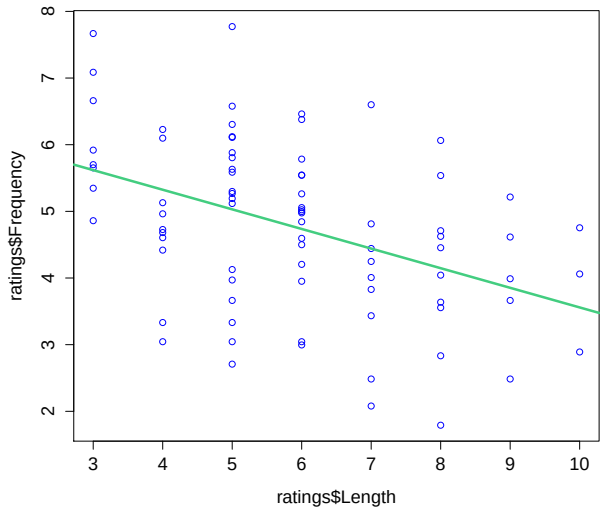
dataset for tutorial 2

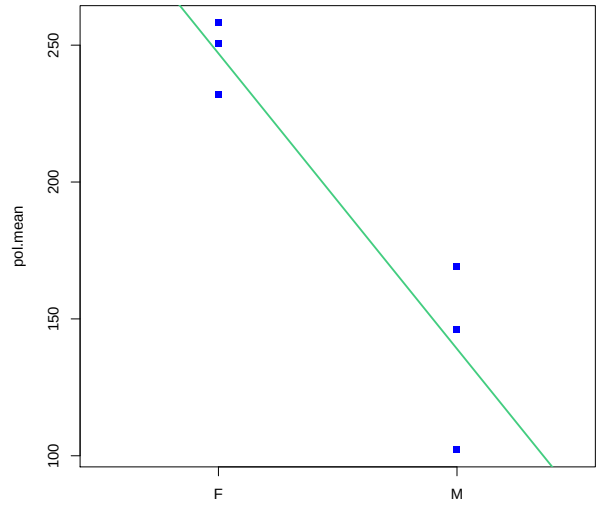
Betöltés legegyszerűbb `read.csv` függvénnyel, mert ott a vessző az alapértelmezett cellaelválasztó jel.

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 gondot).

Megoldás: megnyitás szövegmegjelenítővel (notepad++, gedit vagy más), és mentés a ti OP-rendszeretek kódolása szerinti formátumban. Vagy letöltés innen:
https://phon.nytud.hu/mady/courses/statistics/materials/politeness.RData

Metszéspont jelentősége a lineáris modellekben: lineáris regresszió





Lineáris regresszió: metszéspont azonos a faktoronkénti átlaggal, 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, az ábécé szerint első) faktorszint átlaga, a genderM a M(ale) szint ehhez képest mért különbségét jelenti. A nők (genderF) átlagos alapfrekvenciájához képest a férfiaké $\sim$ 108 Hertz-cel alacsonyabb. Ez megegyezik a két csoport átlagai közötti egyenes meredekségével.

A kevert modellekben minden egyes random hatásként definiált egységre külön metszéspontot lehet számolni (de nem muszáj). A lineáris regresszióval szemben, ahol 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. Ha (0|subject): csak meredekség.

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.
Random vagy véletlen hatások (*random effects*): belső tényezők (*within-subject factors*).
Eljárás: fix hatások szembeállítása, feltételezés: véletlen hatások varianciája ismeretlen.

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

képlet:
függővváltozó $\sim$ függetlenváltozó $+$ randomhatás $+ \epsilon$
ahol ϵ az *error*, hiba terminus a nem ellenőrizhető varianciára.

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

Random hatás nélkül nem működik a modell! Hiszen azért kevert, mert van benne fix és random hatás.

```
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.

Fixed effects: estimate of intercept: *informal* (ABC-ben első) kategória f0-átlaga. *Polite* kategóriáé ennél –19,695 Hertz-cel alacsonyabb → 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 elöl 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 ☹

Sokféle megoldást használnak. Két egyszerű eljárás:

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

Valószínűségek szimulálása Monte-Carlo modellel vagy `Anova` függvénnyel a `car` csomagból.

```
Anova(pol.mod)
```

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

A modell együtthatói

Az itt használt modell az úgynevezett random konstans (*random intercept*) modell. Azt feltételezi, hogy a random hatásokon belül a metszéspontok (konstans, intercept) eltérnek, viszont a másik kondícióhoz képest a különbség azonos.

`coef(modell)`
Az egyes random hatásokhoz tartozó elemek metszéspontja (beszélő és szcenárió).

`ranef(modell)`
Az átlagtól való eltérések a random hatások elemeiként.

(2) Modellek összevetése restricted 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.

Felállítjuk a modellt úgy, hogy csak a nem szerepel fix hatásként, és összevetjük a nem + stílusra felállított modellel.
Vagyis: lemérem a vízzel és elemlámpával futott időmet, majd eldobom az elemlámpát, és így is felfutok. Ha így lassabb vagyok, a lámpa volt a ludas. (Tegyük fel, hogy nem fáradok.)

```
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:
`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, ha fix hatásokhoz tesztelünk, REML=F. Van, aki szerint fordítva. Tartsunk kéznél referenciákat. Ebben a példában Winter a fix hatás tesztelésére REML=F beállítást használ.

(3) Valószínűségek szimulálása Monte-Carlo modellel vagy `Anova` függvénnyel a `car` csomagból.

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

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

Modell együtthatóinak elemzése

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

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

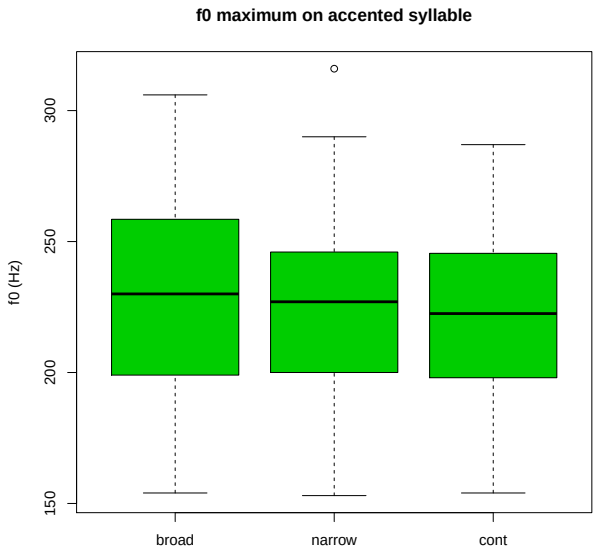
`attitudepol` értékei mindig negatívak, vagyis udvarias stílusban minden személynél és minden item esetén alacsonyabb az f0. Stílus hatásának szignifikanciája ellenőrizhető a modellek összehasonlításával.

Ellenőrizhető az interaction.plot() függvénnyel.

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

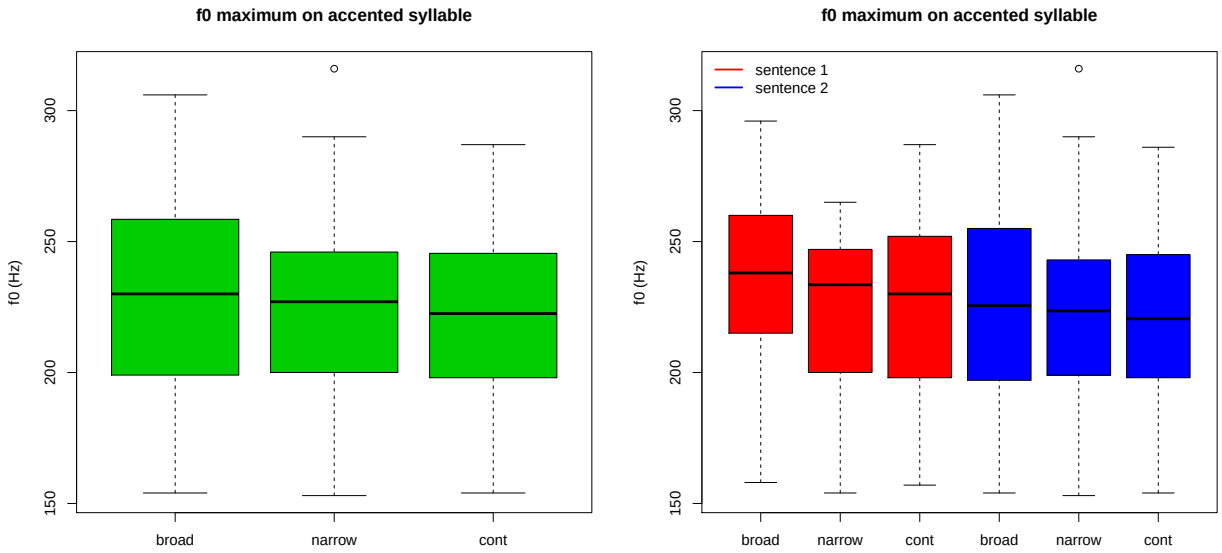
Példa

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



Példa

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



Az első mondatban jól látszik a tendencia, a másodikban nem $\leftrightarrow$ 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 lineáris kevert modellt random intercept és random slope modell alkalmazása esetén. Milyen különbségeket találunk?