

A gyakorló feladatok megoldásai (6.alkalom)

1. Két-szemponos variancia analízist célszerű javasolni, amennyiben a tippelésekre vonatkozó változó normál eloszlású és csoportonként a variancia a populáció szintjén azonosnak tekinthető.

Tegyük fel, hogy az adatbázist a C meghajtóra mentetted! Egy lehetséges megoldás a következő:

```
a=read.table("C:\\idoerzek.txt", sep="," , header=T)
summary(aov(a[,1]~as.factor(a[,2])*as.factor(a[,3])))
```

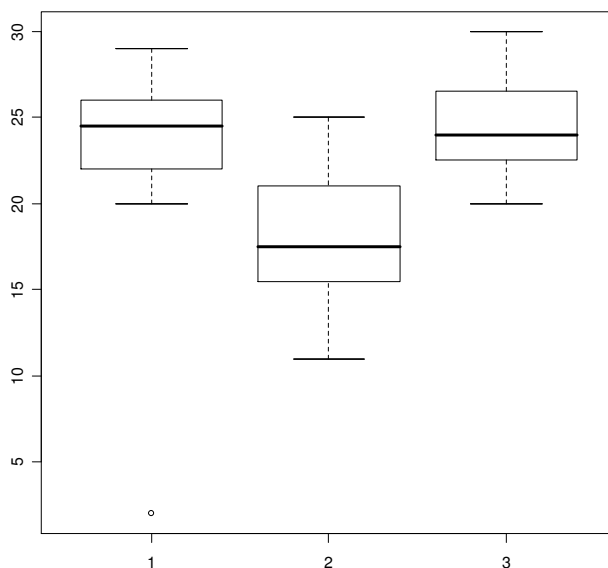
Az interakciós tag nem szignifikáns ($F=2,2$, $df=(6,36)$, $p<0,062$), az a napszak nem befolyásolja, hogy hogyan hatnak a különböző befolyásoltságot okozó szerek. És fordítva: A befolyásoltságot okozó szerek nem hatnak a napszak hatására a becslési feladatban.

A napszak hatása nem szignifikáns, azaz a napszak nem hatott szignifikánsan a becslés jóságára ($F=0.41$, $df=(3,36)$, $p<0.75$).

A befolyásoltságot okozó szer mibenléte szignifikáns hatású ($F=10.3$, $df=(2,36)$, $p<0,001$), azaz van legalább két olyan szer, melyek használata esetén a becslések a populáció szintjén is eltérő jóságúak. Bővebb értelmezés csak páronkénti vizsgálat után lehetséges. A páronkénti vizsgálatok közül leginkább a Tukey-próbát alkalmaznám, mert négy csoportunk is van, illetve minden csoportba 12 fő esik.

```
TukeyHSD(aov(a[,1]~ as.factor(a[,3])))
```

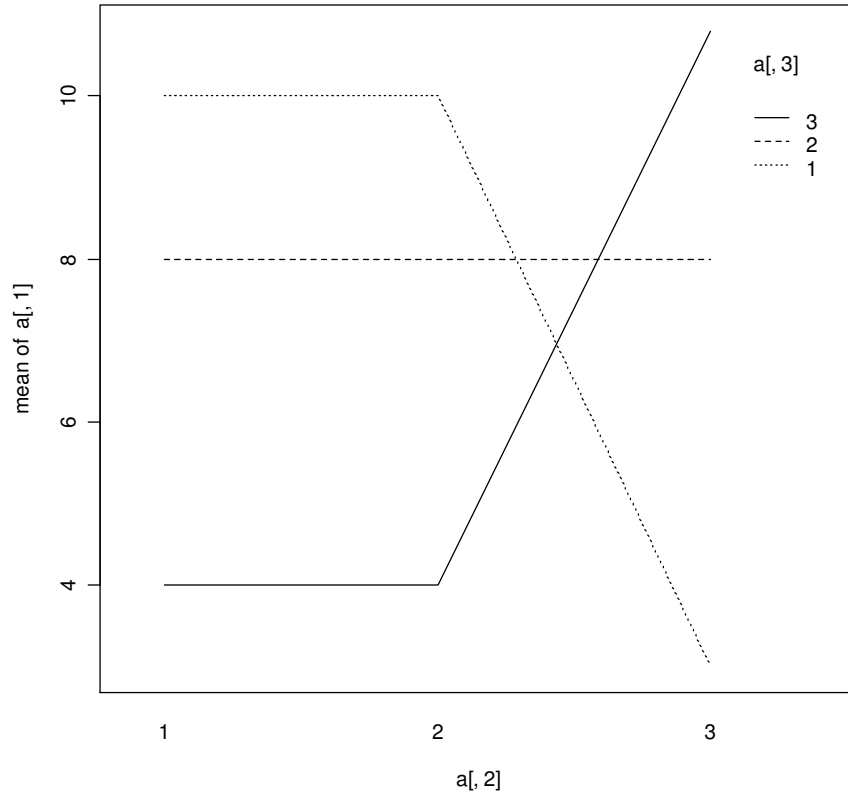
Az eredmények alapján a kettessel jelölt szer hatása különbözik mindkét másik szer hatásától.



És az ábrából látszik, hogy az aszpirin hatása alatt álló személyek átlaga jóval kisebb, mint a másik két csoporté, azaz sokkal jobban becsülnek meg 20mp-et.

2.

```
- a=read.table("C:\\kata\\adatbazisok\\juttalom.txt", sep=" ", header=T)
  interaction.plot(a[,2],a[,3],a[,1])
```



-Tegyük fel, hogy az időtartamokat tartalmazó változó normál eloszlású!
(Gyakorlatban tesztelnünk kell!) Vizsgáljuk meg a csoportonkénti varianciák
azonosságát!

```
library(car)
```

```
levene.test(a[,1], a[,2])
```

A jutalom által képzett csoportok esetén feltételezhetjük a varianciák azonosságát a
populációban ($F=0,465$, $df=(2,42)$, $p<0,63$)

```
levene.test(a[,1], a[,3])
```

A különböző osztályba járó gyermekek esetén is feltételezhetjük az adatok alapján,
hogy az időtartamok varianciája azonos a populáció szintjén ($F=2$, $df=(2,42)$, $p<0,14$).
Azaz használhatjuk a kérdés megválaszolására a kétszemponos varianciaanalízist.

```
- summary(aov(a[,1]~as.factor(a[,2])*as.factor(a[,3])))
```

Szignifikáns az interakciós tag ($F=20$, $df=(4,36)$, $p<0,001$), azaz különbözőképpen
motiválja a különböző korú gyermekeket egy-egy jutalom. Az ábra alapján is láthatjuk,
hogy a negyedikesek kb. egyenlő mértékben próbálkoznak a három jutalom
bármelyikének reményében. Az elsősök akkor a legmotiváltabbak, ha édességet ígérnek
nekik és akkor a legkevésbé, ha pénzt. A hetedikesek viszont éppen ellenkezőleg, a
pénz inkább motiválja őket, mint az édesség.