```
library(lattice)
library(ez)
source(file.path(pfadu, "phoc.txt"))
blang  = read.table(file.path(pfadu, "blang.txt"))
v.df = read.table(file.path(pfadu,"vokal.txt"))
dg = read.table(file.path(pfadu, "dg.txt"))


######################################################################
# 1. gepaarter t-test und within-subjects ANOVA
######################################################################
head(blang); dim(blang)

# Jede Stufe des unabhängigen within-Faktors (Betonung)
# wird einmal pro Vpn belegt
with(blang, table(Vpn, Betonung))

# Differenz-Berechnung
d = aggregate(F2 ~ Vpn, diff, data = blang)
# Boxplot
bwplot(d$F2)
# gepaarter t-Test
t.test(d$F2)
# F2 wurde von der Betonung beeinflusst (t[11] = 4.5, p < 0.01)
# Lösung mit ANOVA
ezANOVA(blang, .(F2), .(Vpn), .(Betonung))
# 2 Betonung   1  11 18.95986 0.001147148     * 0.4113659
# Identifizieren: die F-Statistik = 18.95986; der p-Wert: 0.001147148
# die Freiheitsgrade im Zähler und im Nenner: 1  11
# F2 wurde von der Betonung beeinflusst
# (F[1,11] = 19.0, p < 0.01)
# Beobachten: die F-Statistik ist die t-Statistik hoch 2:
4.3543^2
sqrt(18.95986)

######################################################################
# 2. ungepaarter t-test und between-subjects ANOVA
######################################################################
head(v.df); dim(v.df)
# Die Stufen des unabhängigen between-Faktors
# werden von unterschiedlichen Vpn belegt
with(v.df, table(Vpn, Sprache))
# d.h. jede Reihe ist [1,0] oder [0,1]

# Boxplot oder density plot
bwplot(F2 ~ Sprache, data = v.df)
densityplot(~F2, groups=Sprache, data = v.df, auto.key=T, plot.points=F, ref=T)

# ungepaarter t-Test
t.test(F2 ~ Sprache, data = v.df)
# F2 wurde von der Sprache beinflusst
# (t[11.8] = 2.7, p < 0.05)
# Lösung mit ANOVA
```

```
ezANOVA(v.df, .(F2), .(Vpn), between = .(Sprache))
#
# F2 wurde von der Sprache beeinflusst
# F[1,18] = 7.2, p < 0.05)
#
#
# (Dies nur nebenbei)
####################################################
# Die Bedeutung von dem Levene-Test
# Nur wenn p > 0.05 ist der ANOVA berechtigt

# d.h. ein ANOVA mit between-Faktor wird unter
# der Annahme durchgeführt, dass sich die
# Varianzen in den Stufen nicht  signifikant unterscheiden
# Der ungepaarte t-Test kann geändert werden,
# um diese Annahme zu berücksichtigen mit var.equal=T
#
t.test(F2 ~ Sprache, data = v.df, var.equal=T)
# t-Statistik: 2.688
# 2.688^2
# gibt 7.225344, die F-Statistik in dem between-Faktor ANOVA
#############################################################
# 3. Zwei Faktoren und Interaktionen
######################################################################
head(dg); dim(dg)
#
# Wenn nur ein 1 Wert pro Reihe belegt ist,
# dann sind beide Faktoren 'between'
with(dg, table(Vpn, interaction(Region, Gen)))

# 2 Faktoren, 3 Fragen:
# Wird F2 von Gender beeinflusst? Von der Region?
bwplot(F2 ~ Gen | Region, data = dg, layout=c(3,1))
densityplot(~F2 | Region, groups = Gen, auto.key=T, plot.points=F,
data = dg, layout=c(3,1))

# Liegt eine Interaktion zwischen Gender und Region
# vor? Bedeutung: sind die Unterschiede zwischen A, B, C
# ähnlich für männlich und weiblich? (Oder umgekehrt:
# sind die Unterschiede zwischen männlich und weiblich
# ähnlich in den Regionen A, B, C?). Wenn nein, dann
# liegt eine Interaktion vor.
# Man kann teilweise eine Interaktion auch graphisch darstellen,
# indem die Medianwerte von dem Boxplot miteinander
# verbunden werden:
# Hier sind die Medianwerte pro Stufe
dg.m = aggregate(F2 ~ Gen * Region, median, data = dg)
# Hier werden die Medianwerte getrennt für männlich
# und weiblich verbunden
xyplot(F2 ~ Region, groups = Gen, data = dg.m, type="b", auto.key=T)
# Wenn die Linien mehr oder weniger parallel zueinander
# verlaufen, liegt keine Interaktion vor. Je mehr sie davon
# abweichen, umso wahrscheinlicher wird die
```

```
# Interaktion signifikant sein.
ezANOVA(dg, .(F2), .(Vpn), between =.(Region, Gen))
# F2 wurde von der Region (F[2,54] = 119.6, p < 0.001)
# und vom Geschlecht (F[1,54] = 106.2, p < 0.001)
# signifikant beeinflusst und es gab eine signifikante
# Interaktion (F[2,54] = 12.1, p < 0.001) zwischen
# diesen beiden Faktoren. Post-hoc Bonferroni-korrigierte
# t-Tests zeigten...

###################### Post-hoc t-Tests
## (Wenn eine signifikante Interaktion vorliegt).
## Post-hoc t-Tests vergleichen alle möglichen Stufen-Paare
## d.h. es gibt A, B, C sowie m, w daher Kombinationen von
## A-m mit A-w, A-m mit B-m, A-m mit B-w...
## Es gibt 15 mögliche Paare
## (errechnet sich as 6!/(4!2!) = 30/2 = 15)
## 6 weil es 3 * 2 = 6 Stufen gibt
## 6! bedeutet: 6 * 5 * 4 * 3 * 2 * 1
## Zusätzlich werden die Wahrscheinlichkeiten
## mit der Anzahl der Tests multipliziert
## (d.h. es ist 15 Mal so schwierig, Signifikanzen
## zu bekommen). z.B. ein Wert von 0.04
## ist nach dieser Berechnung 0.04 * 15 = 0.6
## daher nicht signifikant
vok.ph = phoc(dg, .(F2), .(Vpn), .(Region, Gen))
## Wir benötigen aber nicht alle 15 Tests
## sondern nur diejenigen wenn ein Faktor variiert
## und die anderen konstant sind, daher:
# Faktor 1 variiert, Faktor 2 ist konstant
# Bedeutung: gibt es Unterschiede zwischen den
# Dialekten (Faktor 1 variiert) getrennt für Männer
# und Frauen?
p1 = phsel(vok.ph$res, 1)
round(p1, 3)
# Faktor 2 variiert, Faktor 1 ist konstant
# Bedeutung: gib es Unterschiede zwischen männlich/weiblich
# (Faktor 2 variiert) getrennt in den Dialekten?
p2 = phsel(vok.ph$res, 2)
round(p2, 3)
# Post-hoc Bonferroni-korrigierte
# t-Tests zeigten signifikante Unterschiede
# zwischen den Dialekten A und C
# (Für Männer: p < 0.001; Für Frauen p < 0.001)
# sowie zwischen
# den Dialekten B und C
# (Für Männer: p < 0.001; Für Frauen p < 0.001)
# Der Unterschied
# zwischen den Dialekten A und B war nicht signifikant,
# weder in Männern noch in Frauen.
# Schließlich gab es signifikante Unterschiede
# zwischen Männern und Frauen in Dialekt A (p < 0.001)
# und in Dialekt B (p < 0.001) aber nicht in Dialekt C.
```