

```
#####
```

```
library(lattice)
library(ez)
library(lme4)
source(file.path(pfadu, "phoc.txt"))

ssb = read.table(file.path(pfadu, "ssb.txt"))
preasp = read.table(file.path(pfadu, "preasp.txt"))
glottal = read.table(file.path(pfadu, "glottal.txt"))
svot = read.table(file.path(pfadu, "svot.txt"))
dbwerte2 = read.table(file.path(pfadu, "dbwerte2.txt"))
alvtap = read.table(file.path(pfadu, "alvtap.txt"))
ice = read.table(file.path(pfadu, "ice.txt"))
h24 = read.table(file.path(pfadu, "h24.txt"))
vcv = read.table(file.path(pfadu, "vcv.txt"))
japan = read.table(file.path(pfadu, "japan2.txt"))

sig = function(k, m, add = T, ...)
{
  # Funktion um Sigmoiden auf Proportionen zu überlagern.
  curve(exp(m * x + k)/(1 + exp(m * x + k)), add = add, ...)
}
```

```
#####
```

```
# 1. Ich ziehe 12 Zahlen aus einem Hut mit Zahlen zwischen 20 und 80.
  Erstellen Sie ein 95% Konfidenzintervall für den Mittelwert der Zahlen.
  Was ist die Wahrscheinlichkeit, dass der Mittelwert zwischen 60 und 70
  liegt?
```

```
mu = mean(20:80); n = length(20:80);
s = sd(20:80) * sqrt((n-1)/n)
SE = s/sqrt(12)
unten = qnorm(0.025, mu, SE)
oben = qnorm(0.975, mu, SE)
pnorm(70, mu, SE) - pnorm(60, mu, SE)
```

```
# 2. Für die Daten in diesem Data-Frame:
dim(ssb)
```

```
# berechnen Sie (a) den Mittelwert und (b) die Standardabweichung der Variable
  F2 getrennt für jede Stufe der Kombination der Faktoren 'Wort' und
  'Alter' (d.h. 6 Mittelwerte sowie 6 Standardabweichungen jeweils für
  swoop.alt, used.alt, who'd.alt, swoop.jung, used.jung, who'd.jung).
names(ssb)
m = aggregate(F2 ~ interaction(Alter, Wort), mean, data = ssb)
s = aggregate(F2 ~ interaction(Alter, Wort), sd, data = ssb)
```

```
# 3. Für die Daten in diesem Data-Frame:
dim(preasp)
```

```
# erstellen Sie einen Density-plot der Vokaldauer (vdur) dividiert durch
  Verschlussdauer (clodur) für die drei Artikulationsstellen (cplace).
  Hier sollen die Density-Plots für /kk, pp, tt/ aufeinander überlagert
  werden. Modifizieren Sie diese Abbildung, indem Sie vdur/clodur nach
```

```

Region (region) aufteilen (also eine Abbildung für S, eine für C, und
eine für N jeweils mit überlagerten Densityplots für /kk, pp, tt/).
names(preasp)
ratio = with(preasp, vdur/clodur)
densityplot(~ratio, groups = cplace, auto.key=T, plot.points=F, ref=T, data =
  preasp)
densityplot(~ratio | region, groups = cplace, auto.key=T, plot.points=F,
  ref=T, data = preasp)

# 4. Die Daten in diesem Data-Frame:
dim(glottal)

# zeigen ob ein aspirierter (t) oder glottalisierter (tQ) Verschluss von einem
Erkennungssystem richtig (ja) erkannt wurde oder nicht (nein). Erstellen
Sie eine Abbildung, und führen Sie einen statistischen Test durch, um zu
prüfen, ob die Erkennung von dem Verschluss beeinflusst wurde.
head(glottal)
tab = with(glottal, table(verschluss, erkannt))
p = prop.table(tab, 1)
barchart(p, auto.key=T, horizontal=F)
g = glm(erkannt ~ verschluss, binomial, data = glottal)
summary(g)
# Die Erkennung wird nicht vom Verschluss signifikant beeinflusst, obwohl es
eine leichte Tendenz dazu gibt.

# 5. Für diese Daten:
dim(preasp)

# erstellen Sie eine Abbildung von Vokaldauer (vdur) auf der y-Achse als
Funktion von Verschlussdauer (clodur) auf der x-Achse getrennt für die
verschiedenen Wörter (Faktor word); also eine Abbildung mit 8 vdur-
clodur Plots.
head(preasp)
xyplot(vdur ~ clodur | word, data = preasp)

# 6. 10 Sprecher produzierten Wörter in der Laborsprache, der Lesesprache und
in der Spontansprache. Die jeweiligen Silbendauern dieser Sprecher waren
wie folgt:

# Spontan
c(12, 18, 15, 21, 19, 10, 18, 19, 23, 17)

# Lesesprache
c(15, 21, 16, 26, 23, 12, 17, 17, 27, 25)

# Laborsprache
c(18, 19, 15, 32, 22, 14, 21, 21, 30, 21)

# Prüfen Sie durch eine Abbildung und statistischen Test ob der Sprechstil
einen Einfluss auf die Silbendauer hat.
spontan = c(12, 18, 15, 21, 19, 10, 18, 19, 23, 17)
lese = c(15, 21, 16, 26, 23, 12, 17, 17, 27, 25)
labor = c(18, 19, 15, 32, 22, 14, 21, 21, 30, 21)
silbe = c(spontan, lese, labor)
stil.l = c(rep("spontan", length(spontan)), rep("lese", length(lese)),
  rep("labor", length(labor)))
vpn = rep(paste("S", 1:10, sep=""), 3)
stil.df = data.frame(silbe, S = stil.l, Vpn = vpn)
bwplot(silbe ~ S, data = stil.df)
ezANOVA(stil.df, .(silbe), .(Vpn), .(S))

```

```

# Die Silbendauer wurde signifikant von dem Sprechstil beeinflusst (F[2,18] =
  9.0, p < 0.01).

# 7. Für die Daten in diesem Data-Frame:
dim(svot)

# prüfen Sie mit einer Abbildung und einem statistischen Text, inwiefern VOT
  von Gender und Region beeinflusst wird. (Vpn enthält das
  Sprecherkürzel).
head(svot)
names(svot)
with(svot, table(Vpn, interaction(Region, G)))
bwplot(VOT ~ Region | G, data = svot)
ezANOVA(svot, .(VOT), .(Vpn), between = .(Region, G))
p = phoc(svot, .(VOT), .(Vpn), .(Region, G))
round(phsel(p$res), 3)
round(phsel(p$res, 2), 3)
# VOT wurde signifikant von der Region (F[2, 18] = 48.7, p < 0.001) und von
  Gender (F[1, 18] = 9.8, p < 0.001) beeinflusst, und es gab eine
  signifikante (F[2, 18] = 3.9, p < 0.05) Interaktion zwischen diesen
  Faktoren. Post-hoc Tests zeigten keinen Einfluss von Gender auf VOT. Sie
  zeigten auch signifikante Unterschiede zwischen Regionen A und C für
  Männer (p < 0.05) und für Frauen (p < 0.01) sowie zwischen Regionen A
  und B für Frauen (p < 0.01).

# 8. Für diese Daten:
dim(dbwerte2)

# prüfen Sie mit einer Abbildung und statistischem Test ,inwiefern dB aus der
  Dauer vorhergesagt werden kann.
head(dbwerte2)
names(dbwerte2)
plot(dB ~ Dauer, data = dbwerte2)
reg = lm(dB ~ Dauer, data = dbwerte2)
summary(reg)
abline(reg)
shapiro.test(resid(reg))
plot(resid(reg))
# dB kann aus der Dauer vorhergesagt werden (R^2 = .39, F[1, 98] = 61.2, p <
  0.001). Eventuell ist die Regression wegen Autokorrelation in den Daten
  nicht zulässig.

# 9. Für die Daten in diesem Data-Frame:
dim(alvtap)

# erstellen Sie eine Abbildung OHNE einen statistischen Test durchzuführen,
  inwiefern der Faktor K (ob ein Tap produziert wurde oder nicht) vom
  Alter und/oder vom Herkunftsland (Her) beeinflusst wird.
head(alvtap)
tab = with(alvtap, table(Alter, Her, K))
p = prop.table(tab, 1:2)
barchart(p, auto.key=T, horizontal = F)

# 10. Diese Daten:
dim(ice)

```

```

#aus dem Isländischen zeigen Dauerwerte für die Aspiration (Faktor Type)
  entweder vor dem Verschluss (Stufe p = Präaspiration) oder nach dem
  Verschluss (Stufe q = Postaspiration). Die Daten sind von 8 Sprechern
  (Faktor Sprecher) produziert worden. Erstellen Sie eine Abbildung und
  führen Sie einen statistischen Test durch, um zu prüfen, inwiefern die
  Dauer von der Aspiration beeinflusst wird.
with(ice, table(Sprecher, Type))
ice.m = aggregate(Dauer ~ Type * Sprecher, mean, data = ice)
bwplot(Dauer ~ Type, data = ice.m)
ezANOVA(ice.m, .(Dauer), .(Sprecher), .(Type))
# Die Dauer wurde signifikant vom Verschluss beeinflusst (F[1, 8] = 6.5, p <
  0.05)
# Oder mit Mixed-Model
ice.lmer = lmer(Dauer ~ Type + (1+Type|Sprecher), data = ice)
anova(ice.lmer)
ohne = update(ice.lmer, ~ . -Type)
anova(ice.lmer, ohne)
# Die Dauer wurde signifikant vom Verschluss beeinflusst ( $\chi^2[1] = 5.3$ , p <
  0.05)

# 11.
# In diesen Daten:
dim(h24)

# wurden Vokale (Faktor V) Versuchspersonen präsentiert. Der Faktor Urteil
  zeigt, ob die Versuchspersonen die Vokale richtig (1) oder falsch (0)
  identifiziert hatten. Prüfen Sie durch eine Abbildung und statistischen
  Test ob das Urteil von dem Vokal (Faktor V) beeinflusst wurde.
tab = with(h24, table(V, Urteil))
p = prop.table(tab, 1)
barchart(p, horizontal=F, auto.key=T)
g = glm(Urteil ~ V, binomial, data = h24)
ohne = update(g, ~ . -V)
anova(ohne, g, test="Chisq")
# Das Urteil wird vom Vokal beeinflusst ( $\chi^2[2] = 15.3$ , p < 0.001).

# 12. Diese Daten:
dim(vcv)

# zeigen Reaktionszeiten (RT) von verschiedenen Versuchspersonen (Subject) auf
  3 Vokale (Faktor Vowel) in verschiedenen davor (Left) und danach (Right)
  kommenden Konsonanten-Kontexten. Prüfen Sie durch eine Abbildung und
  statistischen Test, ob die Reaktionszeiten vom Vokal (Vowel) und/oder
  Sprachgruppe (Faktor Lang) beeinflusst wurden.
with(vcv, table(Subject, interaction(Vowel, Lang)))
vcv.m = aggregate(RT ~ Vowel * Lang * Subject, mean, data = vcv)

bwplot(RT ~ Vowel | Lang, data = vcv.m)
ezANOVA(vcv.m, .(RT), .(Subject), .(Vowel), between = .(Lang))
# Die Daten zeigen einen signifikanten Einfluss vom Vokal (F[2,32] = 9.0, p <
  0.001) aber keine Einflüsse von der Sprachgruppe und es gab auch keine
  Interaktion zwischen diesen Faktoren.

# 13.
# In diesen Daten:
dim(japan)

```

wurde ein Kontinuum erstellt, indem die Grundfrequenz in einem japanischen Satz in 11 Schritten (Faktor Stim) herabgestuft wurde. Eine Versuchsperson musste beurteilen, ob der Satz eher wie eine Aussage (Aus) oder Erstaunen (Ers) klingt (Faktor Urteil). Prüfen Sie durch eine Abbildung und statistischen Test, inwiefern das Urteil von der Grundfrequenz (Faktor Stim) beeinflusst wurde. Zu welchem Stimulus-Wert kommt der Umkipppunkt zwischen Aussage und Erstaunen vor? (Sie können eine Sigmoid mit der Funktion $\text{sig}(k, m)$ überlagern, wo k und m der Intercept und Steigung der Sigmoid sind).

```
levels(japan$Urteil)
P = japan$Urteil == "Ers"
Q = !P
j.m = aggregate(cbind(P, Q) ~ Stim, sum, data = japan)
p = with(j.m, P/(P+Q))
j.m = cbind(j.m, p)
plot(p ~ Stim, data = j.m)
japan.glm = glm(Urteil ~ Stim, binomial, data = japan)
k = coef(japan.glm)[1]
m = coef(japan.glm)[2]
sig(k, m, col = "blue")
u = -k/m
abline(v=u, lty=2)
summary(japan.glm)
# Das Urteil (ob eine Aussage oder eine Frage) wurde signifikant vom Stimulus
  beeinflusst (z = 7.4, p < 0.001).
```