

```

library(lattice)
source(file.path(pfadu, "proben.R"))
form = read.table(file.path(pfadu, "bet.txt"))
e.df = read.table(file.path(pfadu, "e.txt"))

#####
### 1. SE (Standard Error) und Konfidenzintervall
#####
# (Siehe Vorlesung, Normalverteilung)

##### SE
# Der SE ist die Populationsstandardabweichung von Mittelwerten.
# z.B. Wir werfen 12 Würfeln zusammen, und berechnen den Mittelwert
# der Zahlen.
# Der SE setzt sich zusammen aus  $\sigma/\sqrt{k}$ 
#  $\sigma$  ist die Populationsstandardabweichung =  $\text{sd}(1:6) * \text{sqrt}(5/6)$ 
#  $k$  ist die Anzahl der Würfel = 12
SE =  $\text{sd}(1:6) * \text{sqrt}(5/6) / \text{sqrt}(12)$ 

# Bedeutung: wenn wir 12 Würfeln zusammenwerfen,
# davon den Mittelwert berechnen,
# diesen Vorgang unendlich viel Mal wiederholen,
# sodass wir unendlich viele Mittelwerte hätten,
# dann davon die Standardabweichung berechnen,
# dann wäre diese Standardabweichung genau
#  $\text{SE} = \text{sd}(1:6) * \text{sqrt}(5/6) / \text{sqrt}(12)$ 
# Wir müssten ziemlich nah an diesen SE mit z.B.
# 50000 Mittelwerten kommen:
o =  $\text{proben}(k=12, N = 50000)$ 
sd(o)

##### Konfidenzintervall
# Wir wollen zwei Werte a und b auf eine solche Weise berechnen,
# sodass der Mittelwert zwischen a und b
# mit einer Wahrscheinlichkeit von 0.95 liegt.
# Hier benötigen wir den Populations-Mittelwert:
#
mu =  $\text{mean}(1:6)$ 
# und SE
SE =  $\text{sd}(1:6) * \text{sqrt}(5/6) / \text{sqrt}(12)$ 
# a und b
a =  $\text{qnorm}(0.025, mu, SE)$ 
b =  $\text{qnorm}(0.975, mu, SE)$ 
# d.h. wenn wir 12 Würfeln zusammen werfen,
# davon den Mittelwert berechnen,
# dann fällt der Mittelwert zwischen
# (a) 2.533725 und (b) 4.466275
# mit einer Wahrscheinlichkeit von 0.95
# (= ein Mittelwert von weniger
# als 2.533725 oder mehr als 4.466275 wird meistens nur
# in 5/100 Fällen vorkommen).
# Prüfen: 12 Würfeln werfen, davon den Mittelwert berechnen, diesen
# Vorgang 100 Mal wiederholen (daher 100 Mittelwerte)
m =  $\text{proben}(k = 12, N = 100)$ 
# Wieviele dieser Mittelwerte sind weniger
# als 2.486562 oder größer als 4.513438?
sum(m < 2.533725 | m > 4.466275)

#####
### 2. Konfidenzintervall einer Stichprobe

```

```
#####
# Hier sind 12 Dauerwerte von einem /a:/ Vokal
d = c(119, 111, 105, 130, 133, 122, 124, 129, 95, 100, 109, 111)

# Wir wollen aufgrund der Stichprobe ein Konfidenzintervall
# für den Dauer-Mittelwert von /a:/ erstellen (= dieser Mittelwert
# fällt zwischen a und b mit einer Wahrscheinlichkeit von 95%).
# Zu diesem Zweck können
# i. (mu, SE) aufgrund
# der Stichprobe wie folgt eingeschätzt werden, angenommen dass es
# sich um eine randomisiert ausgewählte Stichprobe handelt:
mu = mean(d)
SE = sd(d)/sqrt(12)

# ii. das Konfidenzintervall wird nicht mit
# der Normalverteilung sondern durch die t-Verteilung
# mit einer gewissen Anzahl von Freiheitsgraden erstellt.
#
# Die t-Verteilung ist der Normalverteilung sehr ähnlich,
# und nähert sich der Normalverteilung
# je höher die Anzahl der Freiheitsgrade
# z.B. t-Verteilung
# mit 3, 6, 20 Freiheitsgrade auf die Normalverteilung überlagern:
curve(dnorm(x, 0, 1), xlim = c(-4, 4), ylab = "Wahrscheinlichkeitsdichte")
# t-Verteilung mit 3 df
curve(dt(x, 3), add = T, col = "red")
# t-Verteilung mit 6 df
curve(dt(x, 6), add = T, col = "blue")
# t-Verteilung mit 20 df
curve(dt(x, 20), add = T, col = "green")

# Die Anzahl der Freiheitsgrade = die Anzahl der Stichproben minus 1
df = 11
# Das Konfidenzintervall für die obige Stichprobe ist:
a = mu + SE * qt(0.025, df)
b = mu - SE * qt(0.025, df)
# Bedeutung: aufgrund dieser Stichprobe fällt der Dauermittelwert
# für /a/ zwischen 107.8 ms und 123.5 ms mit einer
# Wahrscheinlichkeit von 0.95.

#####
### 3. Konfidenzintervall für den Unterschied zwischen 2 Stichproben
#####

# Die Dauerwerte wenn 12 Sprecher einen betonten /a/ produzierten sind wie
# folgt:
bet = c(119, 111, 105, 130, 133, 122, 124, 129, 95, 100, 109, 111)
# Die Dauerwerte wenn die selben 12 Sprecher einen unbetonten /a/ produzierten
# sind:
un = c(110, 95, 108, 80, 120, 110, 120, 95, 72, 83, 90, 95)

# Wir wollen einen 95% Konfidenzintervall erstellen für den
# Unterschied zwischen betont und unbetont
mu = mean(bet - un)
SE = sd(bet - un)/sqrt(12)
# 95% Konfidenzintervall
mu + SE * qt(0.025, 11)
mu + SE * qt(0.975, 11)

# Das 95% Konfidenzintervall für den Unterschied zwischen den
```

```

# Mittelwerten ist  $8.7 \text{ ms} \leq \mu \leq 26.3 \text{ ms}$ 
# Bedeutung: der Unterschied zwischen
# dem Dauer-Mittelwert von einem betontem /a/
# und dem Dauer-Mittelwert von einem unbetonten /a/
# liegt zwischen
#  $8.7 \text{ ms}$  und  $26.3 \text{ ms}$  mit einer Wahrscheinlichkeit von  $0.95$ .

#####
### 4. Prüfen, ob sich Mittelwerte signifikant unterscheiden
#####
#  $H_0$  (die Null Hypothese).
# Betonung hat keinen Einfluss auf die Dauer.
# Bedeutung: der Unterschied zwischen den Mittelwerten ist Null ( $0$ ).

#  $H_1$  (die Alternativ-Hypothese). Betonung beeinflusst die Dauer.
# Bedeutung: der Mittelwert-Unterschied weicht von Null ab.

#  $\alpha$ -Wert (alpha-Wert)
#  $0.05$  ist der hier gewählte, sogenannte  $\alpha$ -Wert (alpha-Wert)
# bei dem wir  $H_0$  verwerfen (Ein Konfidenzintervall von  $0.95$ 
# bedeutet einen  $\alpha$ -Wert von  $1 - 0.95 = 0.05$ ).

# Prüfen: wenn  $0$  außerhalb des Konfidenzintervalls
#  $8.7 \text{ ms} \leq \mu \leq 26.3 \text{ ms}$  fällt, verwerfen wir  $H_0$ 
# und akzeptierten  $H_1$ .

# Schlussfolgerung für das obige Beispiel:
# Wir verwerfen  $H_0$  und akzeptieren  $H_1$ .

# Berichten:
#
# Die Dauer wird signifikant von der Betonung beeinflusst ( $p < 0.05$ )
# Bedeutung (i):
# die Wahrscheinlichkeit, dass die Dauer NICHT von der
# Betonung beeinflusst wird, ist weniger als  $0.05$  (weniger als  $5\%$ ).
# Bedeutung (ii): die Wahrscheinlichkeit, dass der Unterschied
# zwischen dem Dauer-Mittelwert von betontem /a/ und
# dem Dauer-Mittelwert von unbetontem /a/ Null sein könnte, ist
# weniger als  $0.05$ .

#####
### 5. Der t-test
#####
# Wir bekommen dieselbe Auswertung einfacher durch den sogenannten t-test:
# Manuell noch einmal:
mu = mean(bet - un)
SE = sd(bet - un)/sqrt(12)
# 95% Konfidenzintervall
mu + SE * qt(0.025, 11)
mu + SE * qt(0.975, 11)
# mit dem t-test
#
t.test(bet - un)
# Identifizieren:
# mu (17.5)
#
# Freiheitsgrade (11)
#
# 95% Konfidenzintervall (8.739938 26.260062)

```

```

#
# t = 4.3969 (der t-Statistik).
# Der Abstand zwischen mu und 0 (Null) in SE-Einheiten.
(mu - 0)/SE
#
# p-value = 0.001069:
# Bedeutung (i): Die Wahrscheinlichkeit dass
# der Unterschied zwischen den Mittelwerten Null sein könnte.
# Bedeutung (ii): Die Wahrscheinlichkeit, dass H0 zutrifft.
# Bedeutung (iii): die Wahrscheinlichkeit, dass der Unterschied zwischen
# den Mittelwerten außerhalb des Konfidenzintervalls fällt, auch:
(1- pt(4.396914, 11)) * 2

# Ergebnis berichten.
# Wir wählen immer drei  $\alpha$ -Werte aus:  $p < 0.05$ ,  $p < 0.01$ ,  $p < 0.001$ 
# und wählen den  $\alpha$ -Wert, der am nächsten über dem p-Wert liegt
# Hier wählen wir  $p < 0.01$ , da der p-Wert (0.001069) über 0.001 aber unter
# 0.01 liegt.
# Berichten
# Entweder:
# Betonung hatte einen signifikanten Einfluss auf die
# Dauer (t[11] = 4.4,  $p < 0.01$ )
# Oder:
# Die Betonung wurde signifikant von der Dauer beeinflusst
# (t[11] = 4.4,  $p < 0.01$ )
#
# Sollte  $p > 0.05$  sein, dann ist das Ergebnis nicht signifikant
# (Wir verwerfen nicht  $H_0$ ), und schreiben:
# Die Betonung hatte keinen signifikanten Einfluss auf die Dauer
# oder
# Die Dauer wurde nicht signifikant von der Dauer beeinflusst.

#####
### 6. Gepaarte t-tests
#####
# Das obige Beispiel ist ein gepaarter t-test, weil eine
# Differenz - ob betont oder unbetont - pro Paar berechnet wird
# (daher 12 Paare in dem obigen Beispiel).
# Ein gepaarter t-test kommt in der Phonetik meistens
# vor, wenn Stichproben-Paare pro Versuchsperson verglichen werden.
#
# Die Reaktionszeitmessungen (ms)
# für 10 Versuchspersonen wenn sie ein
# Wort in dem linken und rechten ohr wahrgenommen haben, waren wie folgt:
links = c(66, 62, 49, 65, 59, 59, 58, 58, 65, 57)
rechts = c(48, 37, 51, 61, 66, 48, 46, 40, 65, 51)
# Hat das Ohr einen Einfluss auf die Reaktionszeiten?
# = Wird die Reaktionszeit (abhängige Variable) von
# dem Ohr (unabhängiger Faktor mit zwei Stufen, links/rechts)
# beeinflusst?
#
# Dies ist gepaart (wir vergleichen Paare von 2 Stichproben pro Vpn).

# Vorgang für einen gepaarten Test
# 1. Unterschiede zwischen Paaren berechnen
unt = links - rechts
# 2. Abbildung (boxplot) der Unterschiede pro Versuchsperson
# Wenn 0 außerhalb vom Rechteck fällt, wird das Ergebnis
# wahrscheinlich signifikant sein
bwplot(~ unt)

```

```

# 3. t-test der Unterschiede
t.test(unt)
#
# Berichten:
# Die Reaktionszeiten wurden vom Ohr signifikant beeinflusst
# (t[9] = 2.7, p < 0.05)
# Übrigens bekommen wir das gleiche Ergebnis in die andere Richtung
# aber mit einem Minus-Zeichen
t.test(rechts - links)
# Im Bericht wird das Minus-Zeichen ignoriert, daher:
# Die Reaktionszeiten wurden vom Ohr signifikant beeinflusst
# (t[9] = 2.7, p < 0.05)

# Zweites Beispiel: wenn (wie üblich)
# die Werte in einem Data-Frame stecken.

# 12 Versuchspersonen produzierten jeweils einen betonten
# unbetonten /i/. Unterscheiden sich der betonte und unbetonte /i/
# in F2?
dim(form)
head(form)
form

# Die Frage immer umstellen:
# Wurde F2 (abhängige Variable) von der Betonung
# (unabhängiger Faktor mit 2 Stufen: betont/unbetont) beeinflusst?
# Der Test ist gepaart: es gibt ein paar von Werten pro Versuchsperson.

# Vorgang genau wie vorher aber die Befehle
# für 1 und 2 sind etwas anders
# 1. Unterschiede pro Paar (hier Versuchsperson) berechnen
d = aggregate(F2 ~ Vpn, diff, data = form)
# 2. Boxplot der Unterschiede
bwplot(~F2, data = d)
# 3. t.test
t.test(d$F2)
# F2 wurde signifikant von der Betonung beeinflusst
# (t[11] = 4.4, p < 0.01).

#####
### 7. Ungepaarte t-tests
#####
# Wenn nicht Paare von Stichproben sondern
# zwei Gruppen verglichen werden.
# In der Phonetik: meistens wenn zwei verschiedene
# Sprecher-Gruppen (männlich/weiblich; Bayern/Hessen; englisch/deutsch)
# verglichen werden.
# Vorgang: wie oben aber
# die Stichproben werden nicht paarweise voneinander
# subtrahiert (da sie nicht gepaart sind).
#
# Unterscheiden sich deutsch und englisch in F2 von /e/?
# = Wird F2 (abhängige Variable) von der Sprache (unabhängige Variable
# mit 2 Stufen: englisch/deutsch) beeinflusst?
head(e.df)
dim(e.df)
# 1. Boxplot und/oder Densityplot der Unterschiede
bwplot(F2 ~ Sprache, data = e.df)
densityplot(~ F2, groups = Sprache, data = e.df, auto.key=T, plot.points=F,

```

```
      ref=T)
# Hier prüfen wir, ob signifikante Unterschiede
# zwischen den Mittelwerten der beiden Gruppen vorliegen
t.test(F2 ~ Sprache, data = e.df)
# F2 wurde signifikant (t[21.1] = 2.3, p < 0.05) von der Sprache beeinflusst.
# (NB: in einem nicht gepaarten t-Test bekommt man fast immer
# Bruchteile von Freiheitsgraden).
```