t-Tests für unabhängige Stichproben

Ozgur Polat und Kathrin Einzmann

09 01 2020

## Datensatz: RepeatedPulse.csv

Var 1 = Pulsfrequenz (abhängig, intervallskaliert)

Var 2 = Tageszeit (morgens/abends, unabhängig nominalskaliert)

# 1. Hypothese:

H1: Es gibt einen Unterschied in der Herzfrequenz zwischen den beiden Zeitpunkten (morgens und abends).

H0: Es gibt keinen Unterschied in der Herzfrequenz zwischen den beiden Zeitpunkten (morgens und abends).

# 2. Voraussetzungen des t-Tests für abhängige Stichproben

• Die abhängige Variable ist intervallskaliert.

• Die unabhängige Variable ist nominalskaliert und hat zwei Ausprägungen. Unsere unabhängige Variable muss kategorial sein, daher nominalskaliert und muss zwei Ausprägungen haben. Die beiden Ausprägungen beziehen auf die beiden Gruppen, die wir vergleichen und sind oft, aber nicht zwangsläufig, Messzeitpunkte (z.B. Messzeitpunkt #1 verglichen mit Messzeitpunkt #2).

• Normalverteilung: Der gepaarte t-Test erwartet nicht, dass die Daten in den beiden Gruppen normalverteilt sein müssen, allerdings aber die Differenzen beider Gruppen. Genauer gesagt, bezieht sich diese Annahme auf die Residuen und nicht auf die Daten selbst.

• Abhängigkeit der Messungen: Der verbundene t-Test vergleicht zwei Gruppen. In der Regel wird dieselbe Person zu zwei unterschiedlichen Zeitpunkten vergleichen. Die Abhängigkeit muss vorher gegeben sein.

# 3. Grundlegende Konzepte: Was ist t-Test für abhängige Stichproben?

Der t-Test für abhängige Stichproben überprüft, ob die Mittelwerte zweier abhängiger/gepaarter Stichproben verschieden sind.

Die Fragestellung des t-Tests für abhängige Stichproben wird oft so verkürzt: “Unterscheiden sich die Mittelwerte von zwei abhängigen Stichproben?”

Möglichkeiten der Paarung gibt es:

• Messwiederholung (in unserem Fall): Die Messwerte stammen von der gleichen Person, zum Beispiel bei einer Messung vor einem Treatment und nach einem Treatment oder wenn verschiedene Treatments auf die gleiche Person angewendet werden und verglichen werden sollen.

• Natürliche Paare: Die Messwerte stammen von verschiedenen Personen, diese gehören aber zusammen, zum Beispiel Ehefrau – Ehemann, Psychologe – Patient, Anwalt – Klient, Eigentümer – Mieter oder Zwillinge.

• Matching: Die Messwerte stammen von verschiedenen Personen, die einander zugeordnet wurden, zum Beispiel aufgrund eines vergleichbaren Werts auf einer Drittvariablen (die nicht im Zentrum der Untersuchung steht).

# 4. Deskriptive Statistiken und Korrelation

### Normalverteilung

library(dplyr)

##   
## Attaching package: 'dplyr'

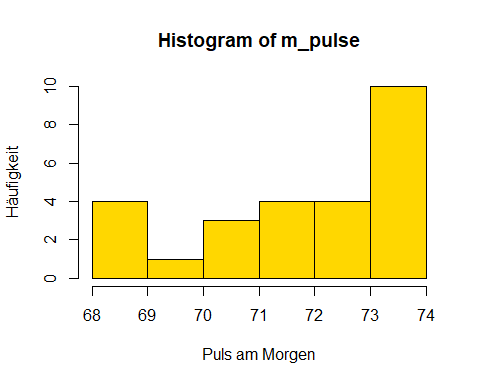
## The following object is masked from 'package:MASS':  
##   
## select

## The following object is masked from 'package:car':  
##   
## recode

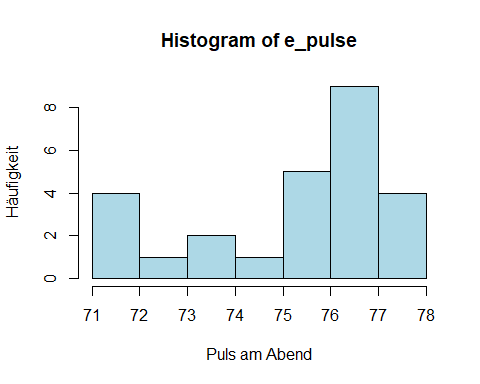
## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

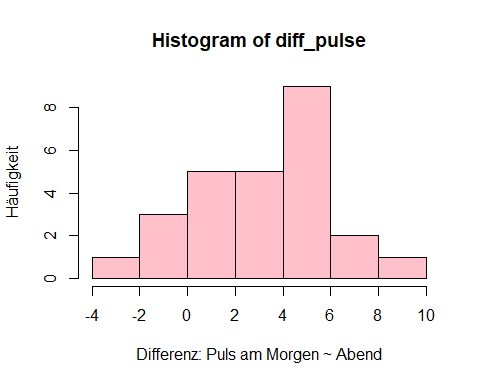
morning<- PULSE%>% filter(Time=="morning")  
evening <- PULSE%>% filter(Time=="evening")  
  
m\_pulse <- (filter(PULSE, Time=="morning")$Pulse)  
e\_pulse <- (filter(PULSE, Time=="evening")$Pulse)  
  
diff\_pulse <- e\_pulse - m\_pulse  
  
hist(m\_pulse, xlab = "Puls am Morgen", ylab = "Häufigkeit", col="gold", nclass = 5)



hist(e\_pulse, xlab = "Puls am Abend", ylab = "Häufigkeit", col="lightblue", nclass = 5)



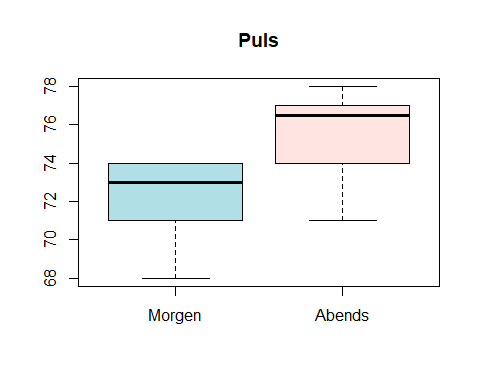
hist(diff\_pulse, xlab = "Differenz: Puls am Morgen ~ Abend", ylab = "Häufigkeit", col = "pink", nclass = 5)



Ergebnis: Die beiden Stichproben sind nicht normalverteilt. Es soll aber geprüft werden, ob die Verteilung der Differenzen zwischen den beiden Messwerten normal ist. Dies trifft zu, die Differenzen sind normalverteilt.

### Boxplots

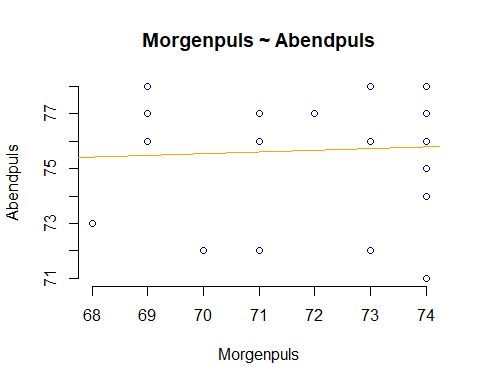
boxplot(morning$Pulse,evening$Pulse,names=c("Morgen","Abends"), col=c('powderblue', 'mistyrose'),main="Puls")



Der Boxplot zeigt, dass sowohl der Rechnungsbetrag als auch das Trinkgeld einige Ausreißer nach oben haben.

<–scatterplot(m\_pulse,e\_pulse, main = “Morgenpuls ~ Abendpuls”,xlab = “Morgenpuls”, ylab = “Abendpuls”, pch = 21, frame = FALSE, col=‘darkblue’,regLine=TRUE)

plot(m\_pulse,e\_pulse, main = "Morgenpuls ~ Abendpuls",xlab = "Morgenpuls", ylab = "Abendpuls",  
pch = 21, frame = FALSE, col='darkblue')  
abline(lm( e\_pulse ~ m\_pulse, data = PULSE), col = "orange")



Es zeigt sich augenscheinlich eine schwach positive lineare Abhängigkeit zwischen den Werten zu den beiden Tageszeiten.

### Korrelation

cor <- cor.test(x=morning$Pulse, y= evening$Pulse, method = "pearson")  
cor

##   
## Pearson's product-moment correlation  
##   
## data: morning$Pulse and evening$Pulse  
## t = 0.27975, df = 24, p-value = 0.7821  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## -0.3378009 0.4347618  
## sample estimates:  
## cor   
## 0.0570108

names(cor)

## [1] "statistic" "parameter" "p.value" "estimate" "null.value"   
## [6] "alternative" "method" "data.name" "conf.int"

r <- cor$estimate  
r

## cor   
## 0.0570108

cor$p.value

## [1] 0.7820664

Ergebnis der Korrelationsanalyse: Die Werte für den Puls zu den beiden Messzeitpunkten korrelieren für ein Fehlerniveau alpha<=0.05 mit p=0.7821 nicht signifikant (r= 0.057, n=26). Die fehlenden Korrelationswerte sind Versuchsdesign bedingt, da sie nicht sauber miteinander verbunden ist.

library(abind, pos=24)  
library(e1071, pos=25)  
numSummary(PULSE[,"Pulse", drop=FALSE], groups=PULSE$Time, statistics=c("mean", "sd", "IQR", "quantiles", "skewness", "kurtosis"), quantiles=c(0,.25,.5,.75,1), type="2")

## mean sd IQR skewness kurtosis 0% 25% 50% 75% 100%  
## evening 75.69231 2.131089 2.75 -0.9601058 -0.3104955 71 74.25 76.5 77.00 78  
## morning 72.23077 1.924738 3.00 -0.8291676 -0.5120082 68 71.00 73.0 74.00 74  
## noon 74.15385 2.556440 2.75 0.2191578 -0.4039852 69 72.25 74.0 75.00 79  
## one 75.38462 1.444352 2.50 -0.3114562 -0.2354335 72 74.25 75.0 76.75 78  
## Pulse:n  
## evening 26  
## morning 26  
## noon 26  
## one 26

Ergebnis von Korrelationsanalyse: Die Werte für den Puls zu den beiden Messzeitpunkten korrelieren für ein Fehlerniveau alpha<=0.05 mit p=0.7821 nicht signifikant (r= 0.057, n=26).

Erfüllte Voraussetzungen:

• Die abhängige Variable Puls ist intervallskaliert.

• Die unabhängige Variable Tageszeit ist nominalskaliert und hat zwei Ausprägungen (morgens/abends).

• Normalverteilung: Der gepaarte t-Test erwartet nicht, dass die Daten in den beiden Gruppen normalverteilt sein müssen, allerdings aber die Differenzen beider Gruppen. Dies ist gegeben (siehe Histogramm der Differenzen)

• Abhängigkeit der Messungen: Der verbundene t-Test vergleicht hier die zwei Zeitpunkte morgens und abends. Die Abhängigkeit ist daher gegeben.

# 5. Ergebnisse des t-Tests für abhängige Stichproben

t\_test <- with(PULSE, (t.test(m\_pulse, e\_pulse, alternative='two.sided', conf.level=.95,   
 paired=TRUE)))  
  
t\_test

##   
## Paired t-test  
##   
## data: m\_pulse and e\_pulse  
## t = -6.3286, df = 25, p-value = 1.267e-06  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -4.588036 -2.335041  
## sample estimates:  
## mean of the differences   
## -3.461538

Die Teststatistik beträgt t = -6.3286 und der zugehörige Signifikanzwert p = .000. Damit ist der Unterschied signifikant: Die Mittelwerte des Pulses zu den beiden Messzeitpunkten (Morgens und Abends) unterscheiden sich (t = -6.3286, p = .000, n = 26).

# 6. Berechnung der Effektstärk

Um die Bedeutsamkeit eines Ergebnisses zu beurteilen, werden Effektstärken berechnet. Im Beispiel ist der Mittelwertsunterschied zwar signifikant, doch es stellt sich die Frage, ob der Unterschied gross genug ist, um ihn als bedeutend einzustufen.

Es gibt verschiedene Arten die Effektstärke zu messen. Zu den bekanntesten zählen die Effektstärke von Cohen (d) und der Korrelationskoeffizient (r) von Pearson. Der Korrelationskoeffizient eignet sich sehr gut, da die Effektstärke dabei immer zwischen 0 (kein Effekt) und 1 (maximaler Effekt) liegt. Wenn sich jedoch die Gruppen hinsichtlich ihrer Grösse stark unterscheiden, wird empfohlen, d von Cohen zu wählen, da r durch die Grössenunterschiede verzerrt werden kann.

t\_value <- t\_test$statistic  
t\_value

## t   
## -6.328619

df <- t\_test$parameter  
df

## df   
## 25

R <- sqrt(t\_value^2/(t\_value^2+df))  
R

## t   
## 0.7846583

Zur Beurteilung der Grösse des Effektes dient die Einteilung von Cohen (1992):

r = .10 entspricht einem schwachen Effekt r = .30 entspricht einem mittleren Effekt r = .50 entspricht einem starken Effekt

Damit entspricht eine Effektstärke von 0,78 einem starken Effekt.

# 7. Eine Aussage

Es zeigt sich, dass die Herzfrequenz gemessen zu zwei Zeitpunkten (morgens und abends) sich statistisch signifikant unterscheiden (t = -6.3286, p = .000, n = 26). Der Morgenpuls fällt niedriger aus (Mittelwert = 72.2351, SD = 1.92) als der Abendpuls (M = 75.69, SD = 2.13). Die Effektstärke nach Cohen (1992) liegt bei r = 0,785 und entspricht damit einem starken Effekt.H0 wird somit abgelehnt und H1 angenommen.