



Computergestützte Datenanalyse

BSc Psychologie SoSe 2021

Prof. Dr. Dirk Ostwald

(8) Anwendungsbeispiel

- Daten zweier Studien (“maps” und “flat”) des **Personality, Motivation, and Cognition Laboratory** mit $n = 330$ Versuchspersonen, Teil des **psychTools R Pakets**
- Fragebogenbasierte Messung von Affektvariablen vor und nach Anschauen eines Filmclips
- Filmclipbedingungen
 - (1) Frontline, Dokumentation über die Befreiung des Konzentrationslagers Bergen-Belsen
 - (2) Halloween, ein Horrorfilm
 - (3) National Geographic, eine Naturdokumentation über die Serengeti
 - (4) Parenthood, eine Komödie
- Messung von Tense Arousal (TA), Energetic Arousal (EA), Positive Affect (PA) und Negative Affect (NA) vor (1) und nach (2) Anschauen des Filmclips
- Erhebung von Daten mit fünf Skalen des Eysenck Persönlichkeitsinventar sowie State- und Trait-Anxiety Items. In der “maps” Studie außerdem Erhebung des BDI
- Für Details, siehe Rafaeli and Revelle (2006)

Der Affect Beispieldatensatz

```
install.packages("psychTools") # einmalige Installation des psychTools Pakets
library(psychTools)           # Laden des psychTools Pakets
?affect                       # Erläuterung des affect Datensatzes
data(affect)                  # Laden des affect Dataframes
View(affect)                  # Inspektion des affect Dataframes
```

#	Study	Film	ext	neur	imp	soc	lie	traitanx	state1	EA1	TA1	PA1	NA1	EA2	TA2	PA2	NA2	state2	MEQ	BDI	
1	maps		3	18.0	9.0	7.0	10.0	3.0	24.0	22.0	24.0	14.0	26.0	2.0	6.0	5.0	7.0	4.0	NA	NA	0.04761905
2	maps		3	16.0	12.0	5.0	8.0	1.0	41.0	40.0	9.0	13.0	10.0	4.0	4.0	14.0	5.0	5.0	NA	NA	0.33333333
3	maps		3	6.0	5.0	3.0	1.0	2.0	37.0	44.0	1.0	14.0	4.0	2.0	2.0	15.0	3.0	1.0	NA	NA	0.19047619
4	maps		3	12.0	15.0	4.0	6.0	3.0	54.0	40.0	5.0	15.0	1.0	0.0	4.0	15.0	0.0	2.0	NA	NA	0.38461538
5	maps		3	14.0	2.0	5.0	6.0	3.0	39.0	67.0	12.0	20.0	7.0	13.0	14.0	15.0	16.0	13.0	NA	NA	0.38095238
6	maps		1	6.0	15.0	2.0	4.0	5.0	51.0	38.0	9.0	14.0	5.0	1.0	7.0	12.0	2.0	2.0	NA	NA	0.23809524
7	maps		1	15.0	12.0	4.0	9.0	3.0	40.0	32.0	1.0	5.0	7.0	0.0	13.0	14.0	8.0	8.0	NA	NA	0.30769231
8	maps		2	18.0	10.0	7.0	9.0	2.0	32.0	41.0	17.0	11.0	10.0	1.0	19.0	15.0	16.0	0.0	NA	NA	0.00000000
9	maps		2	15.0	1.0	3.0	11.0	3.0	22.0	26.0	19.0	5.0	14.0	0.0	19.0	6.0	14.0	0.0	NA	NA	0.00000000
10	maps		2	8.0	10.0	2.0	5.0	2.0	35.0	31.0	15.0	8.0	7.0	0.0	28.0	19.0	11.0	2.0	NA	NA	0.33333333
11	maps		1	13.0	9.0	3.0	9.0	3.0	43.0	39.0	14.0	13.0	10.0	2.0	9.0	21.0	3.0	7.0	NA	NA	0.38095238
12	maps		4	14.0	1.0	3.0	12.0	6.0	33.0	25.0	24.0	15.0	23.0	2.0	27.0	11.0	29.0	0.0	NA	NA	0.14285714
13	maps		4	15.0	2.0	4.0	10.0	5.0	23.0	32.0	7.0	14.0	1.0	0.0	11.0	17.0	4.0	0.0	NA	NA	0.00000000
14	maps		4	19.0	3.0	7.0	11.0	0.0	23.0	23.0	21.0	13.0	20.0	1.0	23.0	18.0	21.0	2.0	NA	NA	0.00000000
15	maps		1	15.0	7.0	4.0	10.0	2.0	27.0	28.0	22.0	15.0	16.0	1.0	16.0	18.0	12.0	4.0	NA	NA	0.04761905
16	maps		1	11.0	13.0	6.0	5.0	7.0	45.0	28.0	2.0	0.0	5.0	0.0	23.0	26.0	21.0	9.0	NA	NA	0.19047619
17	maps		1	16.0	18.0	5.0	10.0	0.0	58.0	56.0	3.0	11.0	3.0	7.0	8.0	17.0	4.0	18.0	NA	NA	0.66666667
18	maps		1	17.0	11.0	6.0	11.0	4.0	39.0	44.0	19.0	18.0	23.0	1.0	21.0	20.0	21.0	6.0	NA	NA	0.33333333
19	maps		1	7.0	10.0	2.0	4.0	1.0	43.0	56.0	14.0	21.0	17.0	8.0	18.0	22.0	12.0	12.0	NA	NA	0.42857143
20	maps		1	13.0	12.0	4.0	8.0	4.0	38.0	35.0	9.0	6.0	1.0	0.0	11.0	10.0	8.0	1.0	NA	NA	0.00000000
21	maps		2	14.0	3.0	5.0	7.0	3.0	35.0	28.0	18.0	8.0	12.0	0.0	16.0	17.0	7.0	2.0	NA	NA	0.04761905
22	maps		2	11.0	10.0	3.0	7.0	1.0	37.0	47.0	11.0	12.0	5.0	1.0	16.0	22.0	12.0	4.0	NA	NA	0.52380952
23	maps		2	20.0	10.0	7.0	10.0	2.0	43.0	43.0	5.0	10.0	2.0	1.0	10.0	25.0	2.0	6.0	NA	NA	0.04761905

Tense Arousal

- *Activation-Deactivation Adjective Check List* nach Thayer (1986).
- Circa 25 Adjektive mit Vierpunkteskala
 - Wie gut beschreibt folgendes [Adjektiv] Ihre momentane Gefühlslage?
 - "I (1) do not feel, (2) cannot decide, (3) feel slightly, (4) definitely feel [Adjektiv]"
- Hohe TA Werte → Hohe Selbsteinschätzung bei *tense, clutched-up, fearful, jittery, intense, ...*
- Niedrige TA Werte → Hohe Selbsteinschätzung bei *calm, still, at-rest, quiet, placid, ...*

Wir wollen im Folgenden exemplarisch untersuchen, ob die Filmclipbedingung Einfluss auf die Tense Arousal Selbsteinschätzung hatte.

Anwendungsbeispiel

- Exploration und Deskription
- T-Tests
- Einfaktorielle Varianzanalyse
- Zweifaktorielle Varianzanalyse
- Übungen und Selbstkontrollaufgaben

Anwendungsbeispiel

- **Exploration und Deskription**
- T-Tests
- Einfaktorielle Varianzanalyse
- Zweifaktorielle Varianzanalyse
- Übungen und Selbstkontrollaufgaben

Visualisierung der Rohdatenverteilungen

```
# Histogrammparameter
TA_min      = min(min(affect$TA1),           # kleinster TA Wert
                  min(affect$TA2))
TA_max      = max(max(affect$TA1),           # kleinster TA Wert
                  max(affect$TA2))

h           = 1                               # gewünschte Klassenbreite
b_0         = TA_min                          # b_0
b_k         = TA_max                          # b_k
k           = ceiling((b_k - b_0)/h)          # Anzahl der Klassen
b           = seq(b_0, b_k, by = h)           # Klassen [b_{j-1}, b_j[
ylimits     = c(0, .2)                       # y-Achsenlimits
xlimits     = c(0, 31)                       # x-Achsenlimits
film        = c("Frontline",                # Filmbedingungen
                "Halloween",
                "Geographic",
                "Parenthood")
```

Visualisierung der Rohdatenverteilungen (fortgeführt)

```
# Abbildungsparameter
par(
  mfcol      = c(2,4),           # für Details siehe ?par
  family     = "sans",         # 2 x 4 Panelstruktur
  pty       = "m",             # Serif-freier Fonttyp
  bty       = "l",             # Maximale Abbildungsregion
  lwd       = 1,               # L förmige Box
  las       = 1,               # Liniendicke
  xaxs     = "i",              # Horizontale Achsenbeschriftung
  yaxs     = "i",              # x-Achse bei y = 0
  font.main = 1,               # y-Achse bei x = 0
  cex      = .6,               # Non-Bold Titel
  cex.main = 1.1)              # Textvergrößerungsfaktor
                                # Titeltextrvergrößerungsfaktor

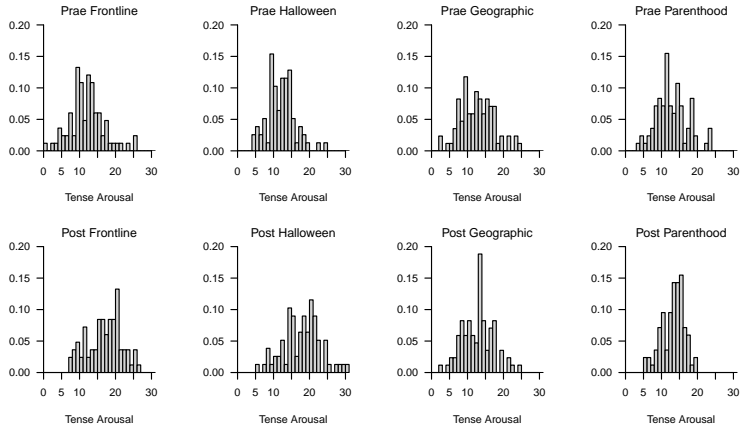
# Iteration über Filmclipbedingungen
for(i in 1:4){

  # Prä-Filmclipexposition TA Werte
  hist(
    affect$TA1[affect$Film == i], # TA Werte von Filmclipbedingung i
    breaks  = b,                  # Histogrammklassen
    freq    = F,                  # normierte relative Häufigkeit
    xlim    = xlimits,            # x-Achsenlimits
    ylim    = ylimits,            # y-Achsenlimits
    xlab    = "Tense Arousal",    # x-Achsenbeschriftung
    ylab    = "",                 # y-Achsenbeschriftung
    main    = paste("Prae", film[i]) # Titelbeschriftung

  # Post-Filmclipexposition TA Werte
  hist(
    affect$TA2[affect$Film == i], # TA Werte von Filmclipbedingung i
    breaks  = b,                  # Histogrammklassen
    freq    = F,                  # normierte relative Häufigkeit
    xlim    = xlimits,            # x-Achsenlimits
    ylim    = ylimits,            # y-Achsenlimits
    xlab    = "Tense Arousal",    # x-Achsenbeschriftung
    ylab    = "",                 # y-Achsenbeschriftung
    main    = paste("Post", film[i]) # Titelbeschriftung
  )
}
```

Visualisierung der Rohdatenverteilungen (fortgeführt)

- Normalisierte relative Häufigkeiten der Tense Arousal Selbsteinschätzungen



- Visuelle Inspektion legt Erhöhung der TA Werte nach Frontline und Halloween nahe

Evaluation der Post-Prae-Filmexposition Tense Arousal Wertdifferenz (DTA)

```
data(affect) # Laden des affect Datensatzes  
affect$DTA = affect$TA2 - affect$TA1 # Post-TA minus Prae-TA Werte
```

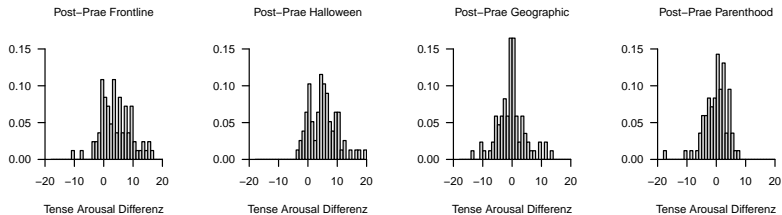
Visualisierung DTA (fortgeführt)

```
# Histogrammparameter
h          = 1                                # gewünschte Klassenbreite
b_0       = min(affect$DTA)                  # b_0
b_k       = max(affect$DTA)                  # b_k
k         = ceiling((b_k - b_0)/h)           # Anzahl der Klassen
b         = seq(b_0, b_k, by = h)            # Klassen [b_{j-1}, b_j[
ylimits   = c(0, .18)                        # y-Achsenlimits
xlimits   = c(-20, 20)                       # x-Achsenlimits
film      = c("Frontline", "Halloween",     # Filmclipbedingungen
              "Geographic", "Parenthood")

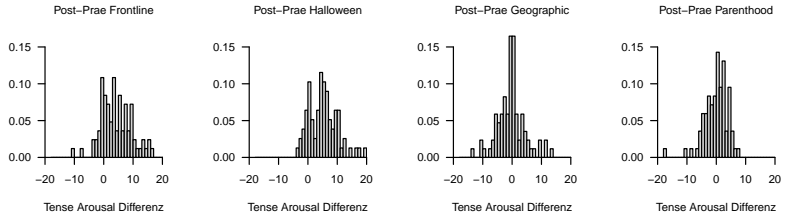
# Abbildungsparameter
par(
  mfcol    = c(1, 4),                        # für Details siehe ?par
  family   = "sans",                         # 1 x 4 Panelstruktur
  pt       = "m",                            # Serif-freier Fonttyp
  bty      = "n",                            # Maximale Abbildungsregion
  las      = 1,                              # L förmige Box
  xaxs     = "i",                            # Horizontale Achsenbeschriftung
  yaxs     = "i",                            # x-Achse bei y = 0
  font.main = 1,                             # y-Achse bei x = 0
  cex      = .6,                             # Non-Bold Titel
  cex.main = 1)                             # Textvergrößerungsfaktor
                                          # Titeltextvergrößerungsfaktor

# Iteration über Filmclipbedingungen
for(i in 1:4){
  hist(
    affect$DTA[affect$Film == i],             # TA Werte von Filmclipbedingung i
    breaks  = b,                             # Histogrammklassen
    freq    = F,                             # normierte relative Häufigkeit
    xlim    = xlimits,                       # x-Achsenlimits
    ylim    = ylimits,                       # y-Achsenlimits
    xlab    = "Tense Arousal Differenz",     # x-Achsenbeschriftung
    ylab    = "",                            # y-Achsenbeschriftung
    main    = paste("Post-Præ", film[i]))    # Titelbeschriftung
}
```

Visualisierung DTA (fortgeführt)



Visualisierung DTA (fortgeführt)



- Visuelle Inspektion legt Erhöhung der TA Werte nach Frontline und Halloween nahe
- Visuelle Inspektion legt leichtes Sinken der TA Werte nach Parenthood nahe

Zusammenfassung und Visualisierung der DTA Werte

```
# Summary Statistics Dataframe Initialisierung
library(psychTools)
data(affect)
affect$DTA = affect$TA2 - affect$TA1
con = 4
STA = data.frame(
  n = rep(NA,con),
  Max = rep(NA,con),
  Min = rep(NA,con),
  Median = rep(NA,con),
  Mean = rep(NA,con),
  Var = rep(NA,con),
  Std = rep(NA,con),
  Sem = rep(NA,con),
  row.names = c("Frontline",
                "Halloween",
                "Geographic",
                "Parenthood"))

# Datensatzladen
# Datensatzladen
# Post-TA minus Prae-TA Werte
# Anzahl Filmclipbedingungen
# Dataframeerzeugung
# Stichprobengrößen
# Maxima
# Minima
# Medians
# Means
# Variances
# Standarddeviations
# Standarderrors
# Filmclipbedingungen

# Iterationen über Filmclipbedingungen
for(i in 1:con){
  data = affect$DTA[affect$Film == i]
  STA$n[i] = length(data)
  STA$Max[i] = max(data)
  STA$Min[i] = min(data)
  STA$Median[i] = median(data)
  STA$Mean[i] = mean(data)
  STA$Var[i] = var(data)
  STA$Std[i] = sd(data)
  STA$Sem[i] = sd(data)/sqrt(length(data))
}
print.AsIs(STA)
```

	n	Max	Min	Median	Mean	Var	Std	Sem
Frontline	83	26	-10	4.00	4.989	32.5	5.70	0.626
Halloween	78	20	-3	5.25	5.712	23.0	4.80	0.543
Geographic	85	14	-13	0.00	0.140	22.7	4.77	0.517
Parenthood	84	8	-18	1.00	0.396	16.3	4.04	0.441

Visualisierung der DTA Mittelwerte und Standardabweichungen

```
# Abbildungsparameter
library(latex2exp)
par(
  family      = "sans",
  pty        = "m",
  bty        = "l",
  lwd        = 1,
  las        = 1,
  font.main  = 1,
  cex        = 1,
  cex.main   = 1.2)

# TeX Annotations
# für Details siehe ?par
# Serif-freier Fonttyp
# Maximale Abbildungsregion
# L förmige Boz
# Linienstärke
# Horizontale Achsenbeschriftung
# Non-Bold Titel
# Textvergrößerungsfaktor
# Titeltextrvergrößerungsfaktor

# Balkendiagramm
x = barplot(
  STA$Mean,
  ylim = c(-5,15),
  names.arg = c("Frontline",
                "Halloween",
                "Geographic",
                "Parenthood"),
  col = "gray90",
  main = TeX("DTA Mittelwerte  $\pm$  STD"))

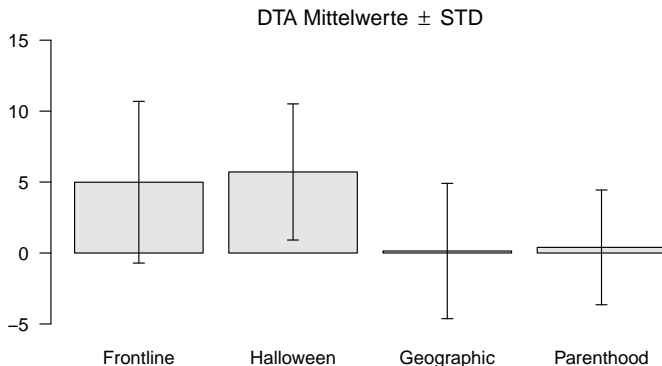
# Ausgabe der x-Ordinaten (?barplot für Details)
# y-Werte (DTA Mittelwerte)
# y-Achsenlimits
# x-Achsenbeschriftung

# Balkenfarbe
# Titel

# Fehlerbalken
arrows(
  x0 = x,
  y0 = STA$Mean - STA$Std,
  x1 = x,
  y1 = STA$Mean + STA$Std,
  code = 3,
  angle = 90,
  length = 0.05)

# für Details siehe ?arrows
# arrow start x-ordinate
# arrow start y-ordinate
# arrow end x-ordinate
# arrow end y-ordinate
# Pfeilspitzen beiderseits
# Pfeilspitzenwinkel -> Linie
# Linielänge
```

Visualisierung der DTA Mittelwerte und Standardabweichungen

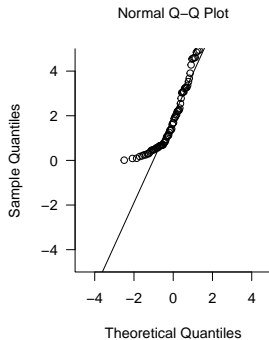
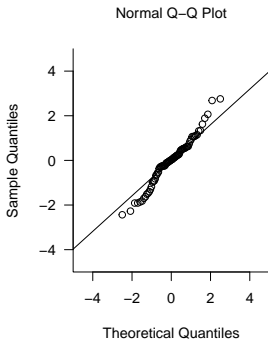


Visuelle Inspektion der Normalität durch Quantil-Quantil-Plots

```
# Abbildungsparameter
par(
  mfcol      = c(1,2),
  family     = "sans",
  pty       = "m",
  bty       = "l",
  lwd       = 1,
  las       = 1,
  xaxs     = "i",
  yaxs     = "i",
  font.main = 1,
  cex      = .8,
  cex.main = .8)
x_norm = rnorm(80)
x_chisq = rchisq(80,2)
qqnorm(x_norm)
qqline(x_norm)
qqnorm(x_chisq)
qqline(x_chisq)

#für Details siehe ?par
# 1 x 4 Panelstruktur
# Serif-freier Fonttyp
# Maximale Abbildungsregion
# L förmige Box
# Liniendicke
# Horizontale Achsenbeschriftung
# x-Achse bei y = 0
# y-Achse bei x = 0
# Non-Bold Titel
# Textvergrößerungsfaktor
# Titeltextrvergrößerungsfaktor
# Normalverteilungsrealisation
# \chi^2(2) Realisation
# Q-Q-Plot
# Q-Q-Plotlinie
# Q-Q-Plot
# Q-Q-Plotlinie
```

Visuelle Inspektion der Normalität durch Quantil-Quantil-Plots



Visuelle Inspektion der Normalität durch Quantil-Quantil-Plots

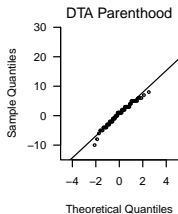
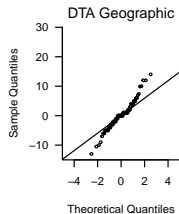
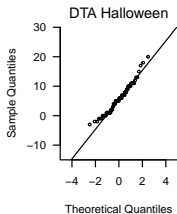
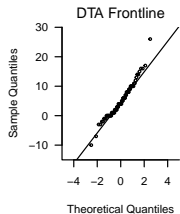
```
# TA Differenzwerte
data(affect)
affect$DTA = affect$TA2 - affect$TA1

# Datensatzladen
# Post-TA minus Prae-TA Werte

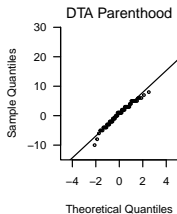
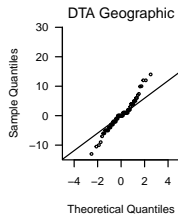
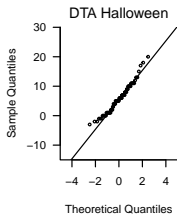
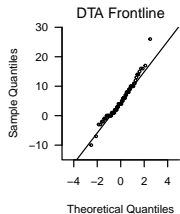
# Abbildungsparameter
par(
  mfcol      = c(1,4),
  family     = "sans",
  pty       = "m",
  bty       = "l",
  lwd       = 1,
  las       = 1,
  xaxs     = "i",
  yaxs     = "i",
  font.main = 1,
  cex      = .6,
  cex.main = 1.2)
# Details siehe ?par
# 1 x 4Panelstruktur
# Serif-freier Fonttyp
# Maximale Abbildungsregion
# L förmige Box
# Liniendicke
# Horizontale Achsenbeschriftung
# x-Achse bei y = 0
# y-Achse bei x = 0
# Non-Bold Titel
# Textvergrößerungsfaktor
# Titeltextvergrößerungsfaktor
# Filmbedingungen
film = c("Frontline",
        "Halloween",
        "Geographic",
        "Parenthood")

# Q-Q-Plots
for(i in 1:4){
  qqnorm(
    affect$DTA[affect$Film == i],
    cex = .5,
    ylim = c(-15,30),
    xlim = c(-5,5),
    main = paste("DTA", film[i]))
  qqline(affect$DTA[affect$Film == i])
}
```

Visuelle Inspektion der Normalität durch Quantil-Quantil-Plots



Visuelle Inspektion der Normalität durch Quantil-Quantil-Plots



- Normalität bis auf Geographic im Wesentlichen bestätigt.
- Normalverteilungsannahme nicht ungerechtfertigt

95%-Konfidenzintervalle für die DTA Erwartungwertparameter bei $DTA[i] \sim N(\mu, \sigma^2)$
⇒ Konfidenzintervall für μ bei $N(\mu, \sigma^2)$ mit $\sigma^2 > 0$ unbekannt

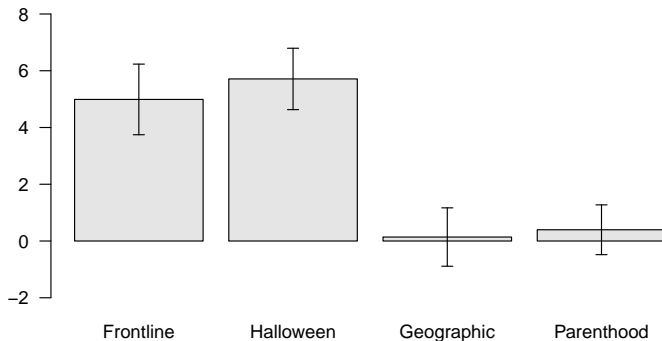
```
# KI Berechnung basierend auf dataframe STA
delta      = 0.95                                # Konfidenzlevel
CI_delta   = STA$Std/sqrt(STA$n)*qt((1+delta)/2,STA$n-1) # Konfidenzlevel Delta
STA$CI_L   = STA$Mean - CI_delta                # Untere Konfidenzlevelgrenze
STA$CI_U   = STA$Mean + CI_delta                # Obere Konfidenzlevelgrenze

# Ausgabe
print.AsIs(STA)
```

	n	Max	Min	Median	Mean	Var	Std	Sem	CI_L	CI_U
Frontline	83	26	-10	4.00	4.989	32.5	5.70	0.626	3.745	6.23
Halloween	78	20	-3	5.25	5.712	23.0	4.80	0.543	4.630	6.79
Geographic	85	14	-13	0.00	0.140	22.7	4.77	0.517	-0.889	1.17
Parenthood	84	8	-18	1.00	0.396	16.3	4.04	0.441	-0.481	1.27

95%-Konfidenzintervalle für die DTA Erwartungswertparameter

Erwartungswertparameterschätzer und Konfidenzintervalle



Anwendungsbeispiel

- Exploration und Deskription
- **T-Tests**
- Einfaktorielle Varianzanalyse
- Zweifaktorielle Varianzanalyse
- Übungen und Selbstkontrollaufgaben

Einstichproben-T-Test mit $H_0 : \mu = \mu_0$ und $H_1 : \mu \neq \mu_0$

Anwendungsszenario

- Eine Stichprobe experimenteller Einheiten
- Annahme der unabhängigen und identischen Normalverteilung $N(\mu, \sigma^2)$
- μ und σ^2 unbekannt
- Quantifizieren der Unsicherheit beim inferentiellen Vergleich von μ mit einem μ_0 beabsichtigt.

Anwendungsbeispiel

- Analyse der TA Werte vor Filmclipexposition "Halloween" in der "maps" Studie
 - $\mu \neq \mu_0 := 12 ? \Rightarrow$ Vom normalen TA-Ruhewert abweichende Proband:innengruppe?

Datenauswahl

```
library(psychTools)           # Datensatzladen
data(affect)                  # Datensatzladen
film = 2                       # Filmbedingung
x = affect$TA1[affect$Study == "maps" &
               affect$Film == film] # Prae-Filmexpositionsdaten in "maps" Studie
                                   # ... für Filmbedingung "film"
```

Einstichproben-T-Test mit $H_0 : \mu = \mu_0$ und $H_1 : \mu \neq \mu_0$

```
# Manueller Einstichproben-T-Test
n      = length(x)                # Stichprobengröße
mu_0   = 12                      # Nullhypothese H_0
alpha_0 = 0.05                   # Signifikanzniveau
k_alpha_0 = qt(1 - (alpha_0/2), n-1) # kritischer Wert
x_bar  = mean(x)                 # Stichprobenmittel
s_n    = sd(x)                   # Stichprobenstandardabweichung
t      = sqrt(n)*(x_bar - mu_0)/s_n # Teststatistik
if(abs(t) >= k_alpha_0){         # Test 1_{|T(X)| >= k_alpha_0}
  phi = 1                         # Ablehnen von H_0
} else {
  phi = 0                         # Nicht Ablehnen von H_0
}
pval   = 2*(1 - pt(abs(t), n-1)) # p-Wert
cat("\nfg      = ", n-1,          # Ausgabe
    "\nx_bar   = ", x_bar,
    "\ns_n     = ", s_n,
    "\nt       = ", t,
    "\nalpha_0  = ", alpha_0,
    "\nk_alpha_0 = ", k_alpha_0,
    "\nphi      = ", phi,
    "\np-Wert   = ", pval)
```

```
fg      = 36
x_bar   = 12.5
s_n     = 4.3
t       = TRUE
alpha_0 = 0.05
k_alpha_0 = 2.03
phi     = 0
p-Wert  = 0.52
```

Einstichproben-T-Test mit $H_0 : \mu = \mu_0$ und $H_1 : \mu \neq \mu_0$

```
# Automatischer Einstichproben-T-Test
varphi = t.test(
  x,
  alternative = c("two.sided"),
  mu = 12,
  conf.level = 1-alpha_0)
# ?t.test für Details
# Datensatz
# H_1: \mu \neq \mu_0
# \mu_0 (sic!)
# \delta = 1 - \alpha_0 (sic!)

# Ausgabe
print(varphi)
```

One Sample t-test

```
data: x
t = 0.7, df = 36, p-value = 0.5
alternative hypothesis: true mean is not equal to 12
95 percent confidence interval:
 11.0 13.9
sample estimates:
mean of x
 12.5
```

```
# Genauere Ausgabe t
paste(varphi[1])
```

```
[1] "c(t = 0.650169688685817)"
```

```
# Genauere Ausgabe p
paste(varphi[3])
```

```
[1] "0.519710450116824"
```

Anwendungsszenario

- Eine Stichprobe experimenteller Einheiten
- Annahme der unabhängigen und identischen Normalverteilung $N(\mu, \sigma^2)$
- μ und σ^2 unbekannt
- Quantifizieren der Unsicherheit beim inferentiellen Vergleich von μ und σ^2 beabsichtigt.

Anwendungsbeispiel

- Analyse der TA Werte vor Filmclipexposition "Parenthood" in der "maps" Studie
 - $\mu > \mu_0 := 12? \Rightarrow$ Proband:innengruppe mit höheren Werten als TA-Ruhewert?

Datenauswahl

```
library(psychTools) # Datensatzladen
data(affect) # Datensatzladen
film = 4 # Filmbedingung
x = affect$TA1[affect$Study == "maps" & # Prae-Filmexpositionsdaten in "maps" Studie
              affect$Film == film] # ... für Filmbedingung "film"
```

Einstichproben-T-Test mit $H_0 : \mu \leq \mu_0$ und $H_1 : \mu > \mu_0$

```
# Manueller Einstichproben-T-Test
n      = length(x)
mu_0   = 12
alpha_0 = 0.05
k_alpha_0 = qt(1 - alpha_0, n-1)
x_bar  = mean(x)
s_n    = sd(x)
t      = sqrt(n)*((x_bar - mu_0)/s_n)
if(t >= k_alpha_0){
  phi = 1
} else {
  phi = 0
}
pval   = 1 - pt(t, n-1)
cat("\nfg      = ", n-1,
    "\nx_bar    = ", x_bar,
    "\ns_n     = ", s_n,
    "\nt       = ", t,
    "\nalpha_0  = ", alpha_0,
    "\nk_alpha_0 = ", k_alpha_0,
    "\nphi     = ", phi,
    "\np-Wert   = ", pval)

# Stichprobengröße
# Nullhypothese H_0
# Signifikanzniveau
# kritischer Wert
# Stichprobenmittel
# Stichprobenstandardabweichung
# Teststatistik
# Test 1_{T >= k_alpha_0}
# Ablehnen von H_0

# Nicht Ablehnen von H_0

# p-Wert
# Ausgabe
```

```
fg      = 37
x_bar   = 14.2
s_n     = 4.69
t       = 2.94
alpha_0 = 0.05
k_alpha_0 = 1.69
phi     = 1
p-Wert  = 0.00283
```

Einstichproben-T-Test mit $H_0 : \mu \leq \mu_0$ und $H_1 : \mu > \mu_0$

```
# Automatischer Einstichproben-T-Test
varphi = t.test(
  x,
  alternative = c("greater"),
  mu = 12,
  conf.level = 1-alpha_0)

# ?t.test für Details
# Datensatz
# H_1: \mu > \mu_0
# \mu_0 (sic!)
# \delta = 1 - \alpha_0 (sic!)

# Ausgabe
print(varphi)
```

One Sample t-test

```
data: x
t = 3, df = 37, p-value = 0.003
alternative hypothesis: true mean is greater than 12
95 percent confidence interval:
 13 Inf
sample estimates:
mean of x
 14.2
```

```
# Genauere Ausgabe t
paste(varphi[1])
```

```
[1] "c(t = 2.93821956016834)"
```

```
# Genauere Ausgabe p
paste(varphi[3])
```

```
[1] "0.00282732752593068"
```

Anwendungsszenario

- Zwei Stichproben experimenteller Einheiten
- Annahme der unabhängiger und identischen Normalverteilungen $N(\mu_1, \sigma_1^2)$ und $N(\mu_2, \sigma_2^2)$
- $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ unbekannt
- Quantifizieren der Unsicherheit beim inferentiellen Vergleich von μ_1 und μ_2 beabsichtigt.

Anwendungsbeispiel

- Analyse der TA Werte nach Filmclipexpositionen "Frontline" und "Halloween" in "maps" Studie
 - $\mu_1 \neq \mu_2 ? \Rightarrow$ Induzieren "Frontline" und "Halloween" unterschiedliche TA Level?

Datenauswahl

```
library(psychTools)                                # Datensatzladen
data(affect)                                       # Datensatzladen
x_1 = affect$TA2[affect$Study == "maps" &         # Post-Filmexpositionsdaten in "maps" Studie
               affect$Film == 1]                 # ... für Filmbedingung "Frontline"
x_2 = affect$TA2[affect$Study == "maps" &         # Post-Filmexpositionsdaten in "maps" Studie
               affect$Film == 2]                 # ... für Filmbedingung "Halloween"
```

Zweistichproben-T-Test bei unabhängigen Stichproben unter Annahme identischer Varianzen

```
# Manueller Zweistichproben-T-Test
n_1      = length(x_1)           # Stichprobengröße n_1
n_2      = length(x_2)           # Stichprobengröße n_2
alpha_0  = 0.05                  # Signifikanzniveau
k_alpha_0 = qt(1 - (alpha_0/2), n_1+n_2-2) # kritischer Wert
x_bar_1  = mean(x_1)             # x_bar_1
x_bar_2  = mean(x_2)             # x_bar_2
s_12     = sqrt((sum((x_1-x_bar_1)^2)+sum((x_2-x_bar_2)^2))/ # gepoolte Standardabweichung s_12
              (n_1+n_2-2))
t        = sqrt((n_1*n_2)/(n_1+n_2))*((x_bar_1-x_bar_2)/s_12) # Zweistichproben-T-Teststatistik
if(abs(t) >= k_alpha_0){        # Test 1_{|T(X)| >= k_alpha_0}
  phi = 1                        # Ablehnen von H_0
} else {
  phi = 0                        # Nicht Ablehnen von H_0
}
pval     = 2*(1 - pt(abs(t), n_1+n_2-2)) # p-Wert
cat("\nx_bar_1 = ", x_bar_1,        # Ausgabe
    "\nx_bar_2 = ", x_bar_2,
    "\nfg      = ", n_1 + n_2 - 2,
    "\nalpha_0  = ", alpha_0,
    "\nk_alpha_0 = ", k_alpha_0,
    "\nt       = ", t,
    "\nphi     = ", phi,
    "\np-Wert  = ", pval)
```

```
x_bar_1 = 18.9
x_bar_2 = 17.9
fg      = 77
alpha_0 = 0.05
k_alpha_0 = 1.99
t       = 0.828
phi     = 0
p-Wert  = 0.41
```

Zweistichproben-T-Test bei unabhängigen Stichproben unter Annahme identischer Varianzen

```
# Automatischer Zweistichproben-T-Test
varphi = t.test(                                # ?t.test für Details
  x_1,                                          # Datensatz x_1
  x_2,                                          # Datensatz x_2
  var.equal = TRUE,                            # \sigma_1^2 = \sigma_2^2
  alternative = c("two.sided"),                # H_1: \mu_1 \neq \mu_2
  conf.level = 1-alpha_0)                     # \delta = 1 - \alpha_0 (sic!)

# Ausgabe
print(varphi)
```

Two Sample t-test

```
data: x_1 and x_2
t = 0.8, df = 77, p-value = 0.4
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.38  3.35
sample estimates:
mean of x mean of y
  18.9    17.9
```

```
# Genauere Ausgabe t
paste(varphi[1])
```

```
[1] "c(t = 0.828099919448388)"
```

```
# Genauere Ausgabe p
paste(varphi[3])
```

```
[1] "0.410173261831605"
```

Zweistichproben-T-Test bei unabhängigen Stichproben ohne weitere Varianzannahmen

```
# Automatischer Zweistichproben-T-Test
varphi = t.test(
  x_1,
  x_2,
  var.equal = FALSE,
  alternative = c("two.sided"),
  conf.level = 1-alpha_0)
# ?t.test für Details
# Datensatz x_1
# Datensatz x_2
# Behrens-Fisher Problem
# H_1: \mu_1 \neq \mu_2
# \delta = 1 - \alpha_0 (sic!)

# Ausgabe
print(varphi)
```

Welch Two Sample t-test

```
data: x_1 and x_2
t = 0.8, df = 73, p-value = 0.4
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.40  3.37
sample estimates:
mean of x mean of y
  18.9    17.9
```

```
# Genauere Ausgabe t
paste(varphi[1])
```

```
[1] "c(t = 0.821929786001458)"
# Genauere Ausgabe p
paste(varphi[3])
```

```
[1] "0.413805439616455"
```

Anwendungsszenario

- **Zwei Messungen** an einer Gruppe experimenteller Einheiten.
- Annahme "abhängiger" Normalverteilungen $N(\mu_1, \sigma_1^2)$ und $N(\mu_2, \sigma_2^2)$
- $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ unbekannt.
- Quantifizieren der Unsicherheit beim inferentiellen Vergleich von μ_1 mit μ_2 beabsichtigt.

Anwendungsbeispiele

- Tense Arousal Datenanalyse bei einer Filmclipbedingung
 - Erste Messung vor Filmclipexposition, zweite Messung nach Filmclipexposition
 - $\mu_1 \neq \mu_2$? \Leftrightarrow Hat Filmclipexposition einen Einfluss auf TA?

Evaluation der Post-Præ-Filmexposition Tense Arousal Wertdifferenz (DTA)

```
library(psychTools)           # psychTools library
data(affect)                  # Laden des affect Datensatzes
affect$DTA = affect$TA2 - affect$TA1  # Post-TA minus Prae-TA Werte
```

Zweistichproben-T-Test bei abhängigen Stichproben

95%-Konfidenzintervalle für die DTA Erwartungswertparameter

Erwartungswertparameterschätzer und Konfidenzintervalle



Zweistichproben-T-Test bei abhängigen Stichproben

Filmclipbedingung "Halloween"

```
x      = affect$DTA[affect$Study == "maps" &
              affect$Film == 2]
n      = length(x)
mu_0   = 0
alpha_0 = 0.05
k_alpha_0 = qt(1 - alpha_0/2, n-1)
x_bar  = mean(x)
s_n    = sd(x)
t      = sqrt(n)*((x_bar - mu_0)/s_n)
if(t >= k_alpha_0){
  phi = 1
} else {
  phi = 0
}
pval   = 2*(1-pt(abs(t), n-1))
cat("\nfg      = ", n-1,
    "\n x_bar   = ", x_bar,
    "\n s_n     = ", s_n,
    "\n t       = ", t,
    "\n alpha_0 = ", alpha_0,
    "\n k_alpha_0 = ", k_alpha_0,
    "\n phi     = ", phi,
    "\n p-Wert  = ", pval)
```

Post minus Prae Filmezpositionen TA Differenzen
... für Filmbedingung Halloween
Stichprobengröße
Nullhypothese $\mu_1 = \mu_2 \Leftrightarrow \mu = 0$
Signifikanzniveau
kritischer Wert
Stichprobenmittel
Stichprobenstandardabweichung
Teststatistik
Test $1_{\{T \geq k_{\alpha_0}\}}$
Ablehnen von H_0

Nicht Ablehnen von H_0

p-Wert
Ausgabe

```
fg      = 36
x_bar   = 5.49
s_n     = 5.05
t       = 6.6
alpha_0 = 0.05
k_alpha_0 = 2.03
phi     = 1
p-Wert  = 1.09e-07
```

Zweistichproben-T-Test bei abhängigen Stichproben

Filmclipbedingung "Parenthood"

```
x      = affect$DTA[affect$Study == "maps" &
              affect$Film == 4]
n      = length(x)
mu_0   = 0
alpha_0 = 0.05
k_alpha_0 = qt(1 - alpha_0/2, n-1)
x_bar  = mean(x)
s_n    = sd(x)
t      = sqrt(n)*((x_bar - mu_0)/s_n)
if(t >= k_alpha_0){
  phi = 1
} else {
  phi = 0
}
pval   = 2*(1-pt(abs(t), n-1))
cat("\nfg      = ", n-1,
    "\n\bar{x}   = ", x_bar,
    "\ns_n     = ", s_n,
    "\nt      = ", t,
    "\n\alpha_0 = ", alpha_0,
    "\nk_alpha_0 = ", k_alpha_0,
    "\nphi     = ", phi,
    "\np-Wert  = ", pval)
```

Post minus Prae Filmezpositionen TA Differenzen
... für Filmbedingung Halloween
Stichprobengröße
Nullhypothese $\mu_1 = \mu_2 \Leftrightarrow \mu = 0$
Signifikanzniveau
kritischer Wert
Stichprobenmittel
Stichprobenstandardabweichung
Teststatistik
Test $1_{\{T \geq k_{\alpha_0}\}}$
Ablehnen von H_0

Nicht Ablehnen von H_0

p-Wert
Ausgabe

```
fg      = 37
x_bar   = 0.211
s_n     = 3.74
t       = 0.347
alpha_0 = 0.05
k_alpha_0 = 2.03
phi     = 0
p-Wert  = 0.73
```

Anwendungsbeispiel

- Exploration und Deskription
- T-Tests
- **Einfaktorielle Varianzanalyse**
- Zweifaktorielle Varianzanalyse
- Übungen und Selbstkontrollaufgaben

Anwendungsszenario

- **Zwei oder mehr Stichproben** (im Folgenden oft **Gruppen** genannt) experimenteller Einheiten.
- Annahme der unabhängigen und identischen Normalverteilung $N(\mu_i, \sigma^2)$ der Daten.
- μ_i und σ^2 unbekannt.
- Absicht des inferentiellen Testens der Nullhypothese identischer Gruppenerwartungswerte.

Tense Arousal Differenzwertanalyse der vier Filmclipbedingungen im Affect Datensatz

- Inferentielle Evidenz für Gruppenerwartungswertunterschiede?
- Evidenz für unterschiedliche Tense Arousal Induktion durch unterschiedliche Filmclips?

Erwartungswertparameterschätzer und Konfidenzintervalle



Einfaktorielle Varianzanalyse

Datenaufbereitung

```
L      = list(                                     # Liste der DTA Gruppenwerte
  x_1j = affect$DTA[affect$Film == 1],
  x_2j = affect$DTA[affect$Film == 2],
  x_3j = affect$DTA[affect$Film == 3],
  x_4j = affect$DTA[affect$Film == 4])

m      = c(                                       # Vektor der Gruppengrößen
  length(L$x_1j),
  length(L$x_2j),
  length(L$x_3j),
  length(L$x_4j))

m_max  = max(m)                                  # Maximale Gruppengröße
X      = data.frame(                             # Dataframeerzeugung mit NA Padding
  x_1j = c(L$x_1j, rep(NA, m_max-length(L$x_1j))),
  x_2j = c(L$x_2j, rep(NA, m_max-length(L$x_2j))),
  x_3j = c(L$x_3j, rep(NA, m_max-length(L$x_3j))),
  x_4j = c(L$x_4j, rep(NA, m_max-length(L$x_4j))))
```

Dataframezeilen ab Zeile 78

	x_1j	x_2j	x_3j	x_4j
78	4	3	-4.0	-1
79	1	NA	0.0	-18
80	11	NA	9.9	2
81	4	NA	1.0	1
82	0	NA	5.0	-5
83	7	NA	-3.0	2
84	NA	NA	1.0	-3
85	NA	NA	0.0	NA

Manuelle Berechnung bei fehlenden Daten (NAs)

```
p      = ncol(X)                # Gruppenanzahl
m      = colSums(!is.na(X))     # Gruppengrößen bei gleichen Gruppengrößen
n      = sum(m)                 # Gesamtstichprobengröße
x_bar  = mean(as.matrix(X), na.rm = TRUE) # Gesamtmittelwert
x_bar_i = colMeans(as.matrix(X), na.rm = TRUE) # Gruppenmittelwerte
alpha_0 = 0.05                 # Signifikanzlevel
k_alpha_0 = qf(1 - alpha_0, p-1, n-p) # kritischer Wert
SQT    = sum((as.vector(as.matrix(X)) - x_bar)^2, na.rm = TRUE) # Total Sum of Squares
SQB    = sum(m*(x_bar_i - x_bar)^2, na.rm = TRUE ) # Between-Group Sum of Squares
SQW    = sum((t(as.matrix(X)) - as.vector(x_bar_i))^2, na.rm = TRUE ) # Within-Group Sum of Squares
MSB    = SQB/(p-1)             # Between-Group Mean Squares
MSW    = SQW/(n-p)            # Within-Group Mean Squares
f_tilde = MSB/MSW             # Realisierung der F-Teststatistik
if(f_tilde >= k_alpha_0){     # \tilde{f} >= k_alpha_0
  phi = 1                      # Testwert
} else {
  phi = 0                      # Testwert
}
p_value = 1 - pf(f_tilde, p-1, n-p) # p-Wert
cat("Einfaktorielle Varianzanalyse",
    "\nGruppenanzahl      = ", p,
    "\nGruppengrößen       = ", m,
    "\nGesamtstichprobengröße = ", n,
    "\nGesamtmittelwert     = ", x_bar,
    "\nGruppenmittelwerte   = ", x_bar_i,
    "\nalpha_0              = ", alpha_0,
    "\nk_alpha_0           = ", k_alpha_0,
    "\nSQT                  = ", SQT,
    "\nSQB                  = ", SQB,
    "\nSQW                  = ", SQW,
    "\nMSB                  = ", MSB,
    "\nMSW                  = ", MSW,
    "\nf_tilde              = ", f_tilde,
    "\nphi                  = ", phi,
    "\np-Wert               = ", p_value,
    "\n\n")
```

Einfaktorielle Varianzanalyse

Manuelle Berechnung bei fehlenden Daten (NAs)

Einfaktorielle Varianzanalyse

Gruppenanzahl = 4
Gruppengrößen = 83 78 85 84
Gesamtstichprobengröße = 330
Gesamtmittelwert = 2.74
Gruppenmittelwerte = 4.99 5.71 0.14 0.396
alpha_0 = 0.05
k_alpha_0 = 2.63
SQT = 9847
SQB = 2145
SQW = 7702
MSB = 715
MSW = 23.6
f_tilde = 30.3
phi = 1
p-Wert = 0

Automatische Berechnung bei fehlenden Daten (NAs)

```
XAOV = data.frame(dta = c(X$x_1j, X$x_2j, X$x_3j, X$x_4j),  
                 group = as.factor(c(rep(1,length(X$x_1j)),  
                                   rep(2,length(X$x_2j)),  
                                   rep(3,length(X$x_3j)),  
                                   rep(4,length(X$x_4j)))))  
  
res.aov = aov(dta ~ group, data = XAOV)  
summary(res.aov)
```

```
          Df Sum Sq Mean Sq F value Pr(>F)  
group      3   2145     715   30.3 <2e-16 ***  
Residuals 326   7702      24  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
10 observations deleted due to missingness
```

- Die Nullhypothese identischer Gruppenerwartungswertparameter kann abgelehnt werden.
- Die Filmclipbedingungen induzieren recht sicher unterschiedliche TA Werte.

Anwendungsbeispiel

- Exploration und Deskription
- T-Tests
- Einfaktorielle Varianzanalyse
- **Zweifaktorielle Varianzanalyse**
- Übungen und Selbstkontrollaufgaben

Zweifaktorielle Varianzanalyse

Anwendungsszenario

Zweifaktorielle Versuchspläne mit crossed design (cf. Modul A.2)

- Eine univariate Messvariable bestimmt an individuellen experimentellen Einheiten.
- Zwei diskrete experimentelle Variablen, die mindestens zweistufig sind.
- Die experimentellen Variablen werden *Faktoren* genannt.
- Die Stufen der Faktoren werden auch *Faktorlevel* genannt.
- Jedes Level eines Faktors wird mit allen Level des anderen Faktors kombiniert

TA Score Pre-Post-Differenzwertanalyse für zwei Filmclips und zwei Studien

- Faktor Filmclip (A) mit Level Halloween (1) und Geographic (2)
- Faktor Studien (B) mit Level Flat (1) und Maps (2)
- 2×2 faktorieller Versuchsplan mit crossed design $\Rightarrow 2 \times 2$ ANOVA
- Inferentielle Evidenz für einen Haupteffekt der Filmclipbedingung?
- Inferentielle Evidenz für einen Haupteffekt der Studie?
- Inferentielle Evidenz für eine Interaktion = Änderung der Filmclipeffekte zwischen den Studien?

Zweifaktorielle Varianzanalyse

Datenaufbereitung

```
L      = list(
  A1B1 = affect$DTA[affect$Film == 2 & affect$Study == "maps"], # Halloween , Maps
  A1B2 = affect$DTA[affect$Film == 2 & affect$Study == "flat"], # Halloween , Flat
  A2B1 = affect$DTA[affect$Film == 3 & affect$Study == "maps"], # Geographic, Maps
  A2B2 = affect$DTA[affect$Film == 3 & affect$Study == "flat"]) # Geographic, Flat

m      = c(
  length(L$A1B1), # Vektor der Gruppengrößen
  length(L$A1B2),
  length(L$A2B1),
  length(L$A2B2))

m_min  = min(m) # minimale Gruppengröße
X      = data.frame(
  A1B1 = L$A1B1[1:m_min],
  A1B2 = L$A1B2[1:m_min],
  A2B1 = L$A2B1[1:m_min],
  A2B2 = L$A2B2[1:m_min])
```

Dataframezeilen ab Zeile 35

	A1B1	A1B2	A2B1	A2B2
35	5	6	-1	-4.0
36	2	1	-5	0.0
37	2	10	12	9.9

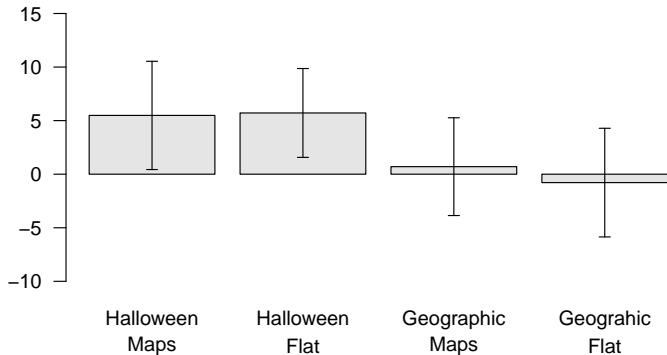
Zweifaktorielle Varianzanalyse

Darstellung der Gruppenmittelwerte und Gruppenstandardabweichungen durch Balkendiagramm

```
dev.new()
library(latex2exp)
par(
  family = "sans",
  pty     = "m",
  bty     = "l",
  lwd     = 1,
  las     = 1,
  font.main = 1,
  cex     = 1.1,
  cex.main = 1.2)
groupmeans = colMeans(X, na.rm = TRUE)
groupstds  = apply(X,2,sd)
x = barplot(
  groupmeans,
  ylim       = c(-10,15),
  col        = "gray90",
  names.arg  = c("Halloween" , "Halloween",
                 "Geographic" , "Geographic"))
text(x[1],-16, "Maps" , xpd = T, cex = 1)
text(x[2],-16, "Flat" , xpd = T, cex = 1)
text(x[3],-16, "Maps" , xpd = T, cex = 1)
text(x[4],-16, "Flat" , xpd = T, cex = 1)
arrows(
  x0      = x,
  y0      = groupmeans - groupstds,
  x1      = x,
  y1      = groupmeans + groupstds,
  code    = 3,
  angle   = 90,
  length  = 0.05)
dev.copy2pdf(
  file    = file.path(getwd(), "08_Abbildungen", "cda_08_aov_2_barplot.pdf"),
  width   = 7,
  height  = 4)
# Figure Initialisierung
# TeX Annotations
# für Details siehe ?par
# Serif-freier Fonttyp
# Maximale Abbildungsregion
# L förmige Boz
# Liniendicke
# Horizontale Achsenbeschriftung
# Non-Bold Titel
# Textvergrößerungsfaktor
# Titeltextvergrößerungsfaktor
# Gruppenmittelwerte
# Gruppenmittelwerte
# Ausgabe der x-Ordinaten (?barplot für Details)
# Labels
# für Details siehe ?arrows
# arrow start x-ordinate
# arrow start y-ordinate
# arrow end x-ordinate
# arrow end y-ordinate
# Pfeilspitzen beiderseits
# Pfeilspitzenwinkel -> Linie
# Linienlänge
```

Zweifaktorielle Varianzanalyse

Darstellung der Gruppenmittelwerte und Gruppenstandardabweichungen durch Balkendiagramm



Zweifaktorielle Varianzanalyse

Darstellung der Gruppenmittelwerte und Gruppenstandardabweichungen durch Liniendiagramm

```
dev.new()
library(latex2exp)
x          = 1:2
groupmeans = matrix(colMeans(X, na.rm = TRUE), nrow = 2)
groupstds  = matrix(apply(X, 2, sd, na.rm = TRUE), nrow = 2)
cols       = c("gray30", "gray70")
lwds       = c(4,2)
graphics.off()
fdir       = file.path(getwd(), "08_Abbildungen")
dev.new()
par(
  family = "sans",
  pty    = "s",
  bty    = "l",
  lwd    = 1,
  las    = 1,
  mgp    = c(2,1,0),
  xaxs   = "i",
  yaxs   = "i",
  font.main = 1,
  cex    = 1.5,
  cex.main = 1)
matplot(
  x,
  groupmeans,
  type = "b",
  pch  = 21,
  xlim = c(.5,2.5),
  ylim = c(-10,15),
  lty  = 1,
  col  = cols,
  lwd  = lwds,
  xlab = "",
  ylab = "",
  xaxt = "n")
```

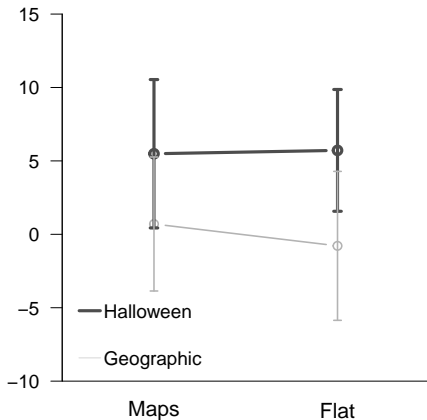
Zweifaktorielle Varianzanalyse

Darstellung der Gruppenmittelwerte und Gruppenstandardabweichungen durch Liniendiagramm

```
# Fehlerbalken und Annotationen
for(i in 1:2){
  arrows(
    x0      = x,
    y0      = grupmeans[,i] - groupstds[,i],
    x1      = x,
    y1      = grupmeans[,i] + groupstds[,i],
    col     = cols[[i]],
    lwd     = lwds[[i]],
    code    = 3,
    angle   = 90,
    length  = 0.05)
}
legend(
  .4,
  -2,
  c("Halloween", "Geographic"),
  lty     = 1,
  col     = c("gray30", "gray90"),
  lwd     = lwds,
  bty     = "n",
  cex     = 1,
  x.intersp = .1,
  seg.len = 0.6,
)
text(1,-12, "Maps" , xpd = T, cex = 1.1)
text(2,-12, "Flat" , xpd = T, cex = 1.1)
dev.copy2pdf(
  file     = file.path(getwd(), "08_Abbildungen", "cda_08_aov_2_lineplot.pdf"),
  width    = 7,
  height   = 7)
```

Zweifaktorielle Varianzanalyse

Darstellung der Gruppenmittelwerte und Gruppenstandardabweichungen durch Liniendiagramm



Zweifaktorielle Varianzanalyse

Manuelle Berechnung

```
# Analyseparameter
p           = 2
q           = 2
m           = m_min
n           = p*q*m
alpha_0     = 0.05

# Mittelwerte
x_bar      = mean(as.matrix(X))
x_bar_i    = colMeans(cbind(as.matrix(c(X$A1B1, X$A1B2)),
                             as.matrix(c(X$A2B1, X$A2B2))))
x_bar_j    = colMeans(cbind(as.matrix(c(X$A1B1, X$A2B1)),
                             as.matrix(c(X$A1B2, X$A2B2))))
x_bar_ij   = colMeans(X)

# Parameterschätzer
mu_hat     = x_bar
alpha_hat_i = x_bar_i - x_bar
beta_hat_j  = x_bar_j - x_bar
gamma_hat_ij = c(x_bar_ij[1] - (mu_hat + alpha_hat_i[1] + beta_hat_j[1]),
                 x_bar_ij[2] - (mu_hat + alpha_hat_i[1] + beta_hat_j[2]),
                 x_bar_ij[3] - (mu_hat + alpha_hat_i[2] + beta_hat_j[1]),
                 x_bar_ij[4] - (mu_hat + alpha_hat_i[2] + beta_hat_j[2]))
sigsqr_hat = sum((t(as.matrix(X)) - as.vector(x_bar_ij))^2)/(p*q*(m-1))

# Sum of squares
SQT        = mean((as.matrix(X) - x_bar)^2)
SQA        = q*m*sum(alpha_hat_i^2)
SQB        = p*m*sum(beta_hat_j^2)
SQG        = m*sum(gamma_hat_ij^2)
SQR        = p*q*(m-1)*sigsqr_hat

# Anzahl Level Faktor A
# Anzahl Level Faktor B
# Anzahl Datenpunkte pro Zelle
# Gesamtstichprobengröße
# Signifikanzlevel

# Gesamtmittelwert \bar{X}
# Faktor A Mittelwerte \bar{X}_i
# Faktor B Mittelwerte \bar{X}_j
# Gruppenmittelwerte \bar{X}_{ij}

# \mu Schätzer
# \alpha_i Schätzer
# \beta_j Schätzer
# \gamma_ij Schätzer

# \sigma^2 Schätzers

# Total SOS
# Faktor A SOS
# Faktor B SOS
# Interaktions SOS
# Residual SOS
```

Zweifaktorielle Varianzanalyse

Manuelle Berechnung

```
# Mean Sum of Squares
MSA      = SQA/(p-1)                # Mean Square Faktor A
MSB      = SQB/(q-1)                # Mean Square Faktor B
MSG      = SQG/((p-1)*(q-1))        # Mean Square Interaktion
MSR      = SQR/(p*q*(m-1))          # Mean Square Residuals

# Teststatistiken
F_A      = MSA/MSR                  # Faktor A F-Teststatistik
F_B      = MSB/MSR                  # Faktor B F-Teststatistik
F_AB     = MSG/MSR                  # Interaktion F-Teststatistik

# kritische Werte
k_A_alpha_0 = qf(1-alpha_0, p-1, p*q*(m-1)) # k^A_\alpha_0
k_B_alpha_0 = qf(1-alpha_0, q-1, p*q*(m-1)) # k^B_\alpha_0
k_AB_alpha_0 = qf(1-alpha_0, (p-1)*(q-1), p*q*(m-1)) # k^AB_\alpha_0

# Tests
phi       = c(F_A, F_B, F_AB) >= c(k_A_alpha_0, k_B_alpha_0, k_AB_alpha_0) # simultaner Test

# p-Werte
p_values  = 1-c(pf(F_A, p-1, p*q*(m-1)),
                pf(F_B, q-1, p*q*(m-1)),
                pf(F_AB, (p-1)*(q-1), p*q*(m-1)))
```

Manuelle Berechnung

Zweifaktorielle Varianzanalyse

Anzahl Level Faktor A = 2
Anzahl Level Faktor B = 2
Gruppengrößen = 37
Gesamtstichprobengröße = 148
alpha_0 = 0.05

Haupteffekt Faktor A

SQA = 1178
MSA = 1178
F_A = 52.8
k_A_alpha_0 = 3.91
phi_A = 1
p_A-Wert = 2.16e-11

Haupteffekt Faktor B

SQB = 14.7
MSB = 14.7
F_B = 0.657
k_B_alpha_0 = 3.91
phi_B = 0
p_B-Wert = 0.419

Interaktion A x B

SQG = 27.3
MSG = 27.3
F_AxB = 1.22
k_AxB_alpha_0 = 3.91
phi_AxB = 0
p_AxB-Wert = 0.27

Zweifaktorielle Varianzanalyse

Automatische Berechnung

```
XAOV = data.frame(X = c(X$A1B1, X$A1B2, X$A2B1, X$A2B2) , # Reformatierung des Datensatzes
                 A = as.factor(c(rep(1,2*m), rep(2,2*m))) , # Faktor A Level Labels
                 B = as.factor(c(rep(1,m), rep(2,m), rep(1,m), rep(2,m)))) # Faktor B Level Labels
res.aov = aov(X ~ A + B + A:B, data = XAOV) # Benutzung von R's aov Funktion
summary(res.aov)
```

```
          Df Sum Sq Mean Sq F value Pr(>F)
A           1   1178     1178   52.79 2.2e-11 ***
B           1     15       15    0.66  0.42
A:B         1     27       27    1.22  0.27
Residuals 144   3214        22
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Die Nullhypothese keines Filmcliphaupteffektes kann abgelehnt werden.

- Die Filmclips induzieren recht sicher unterschiedliche TA Erwartungswertparameter.

Die Nullhypothese keines Studienhaupteffektes kann nicht abgelehnt werden.

- Die Studien induzieren recht sicher keine unterschiedlichen TA Erwartungswertparameter.

Die Nullhypothese keiner Interaktion kann nicht abgelehnt werden.

- Die durch die Filmclips induzierten TA Erwartungswertparameter unterscheiden sich zwischen den Studien vermutlich nicht.

Anwendungsbeispiel

- Exploration und Deskription
- T-Tests
- Einfaktorielle Varianzanalyse
- Zweifaktorielle Varianzanalyse
- **Übungen und Selbstkontrollaufgaben**

1. Dokumentieren Sie die in dieser Einheit eingeführten R Befehle in einem R Skript.
2. Nennen Sie eine R Funktion zur visuellen Inspektion der Normalität von Daten.
3. Nennen Sie eine R Funktion zur Durchführung von Einstichproben-T-Tests.
4. Nennen Sie eine R Funktion zur Durchführung von Zweistichproben-T-Tests.
5. Nennen Sie eine R Funktion zur Durchführung einfaktorieller Varianzanalysen.
6. Nennen Sie eine R Funktion zur Durchführung zweifaktorieller Varianzanalysen.

References

- Rafaeli, Eshkol, and William Revelle. 2006. "A Premature Consensus: Are Happiness and Sadness Truly Opposite Affects?" *Motivation and Emotion* 30 (1): 1–12. <https://doi.org/10.1007/s11031-006-9004-2>.
- Thayer, Robert E. 1986. "Activation-Deactivation Adjective Check List: Current Overview and Structural Analysis." *Psychological Reports* 58 (2): 607–14. <https://doi.org/10.2466/pr0.1986.58.2.607>.