Tutorium Allgemeines Lineares Modell

BSc Psychologie SoSe 2022

11. Termin (9) T-Tests

Belinda Fleischmann

Selbstkontrollfragen

1. Erläutern Sie die Extremszenarien im Kontinuum von ALM Designs.
2. Erläutern Sie die Begriffe der faktoriellen und parametrischen ALM Designs.
3. Nennen Sie Beispiele für faktorielle, parametrische, und faktoriell-parametrische ALM Designs.
4. Erläutern Sie das Anwendungsszenario eines Einstichproben-T-Tests.
5. Erläutern Sie mögliche Hypothesenszenarien eines Einstichproben-T-Tests.
6. Geben Sie die Definition des Einstichproben-T-Test Modells wieder.
7. Geben Sie das Theorem zur Parameterschätzung im Einstichproben-T-Test Modell wieder.
8. Geben Sie das Theorem zur T-Teststatistik des Einstichproben-T-Tests wieder.
9. Geben Sie die Definition des zweiseitigen Einstichproben-T-Tests (mit ungerichteter Hypothese) wieder.
10. Skizzieren Sie die Testgütefunktion des zweiseitigen Einstichproben-T-Tests.
11. Geben Sie das Theorem zur Testumfangkontrolle im zweiseitigen Level-α_0-Einstichproben-T-Test wieder.
12. Erläutern Sie das praktische Vorgehen bei Durchführung eines zweiseitigen Level-α_0-Einstichproben-T-Tests.
13. Geben Sie die Definition des p-Wertes Werts für einen zweiseitigen Einstichproben-T-Test wieder.
14. Von welchen Werten hängt die Powerfunktion eines zweiseitigen Einstichproben-T-Tests ab?
15. Skizzieren Sie die Powerfunktion des zweiseitigen Einstichproben-T-Tests bei fester Stichprobengröße.
16. Skizzieren Sie die Powerfunktion des zweiseitigen Einstichproben-T-Tests bei festem Nichtzentralitätsparameter.
17. Betrachten Sie die PostBDI-PreBDI Differenzwertdaten der Waitlist Control Bedingung im Beispieldatensatz. Erstellen Sie ein Histogramm dieser Daten und evaluieren Sie Ihnen bekannte deskriptive Statistiken zu diesen Daten. Führen Sie einen zweiseitigen Einstichproben-T-Test mit Nullhypothesenparameter $\mu_0 = 0$ durch. Dokumentieren Sie Ihre Ergebnisse. Was folgern Sie aus den sich ergebenen Resultaten?

Selbstkontrollfragen

18. Erläutern Sie das Anwendungsszenario eines Zweistichproben-T-Tests.
19. Geben Sie die Definition des Zweistichproben-T-Test Modells wieder.
20. Geben Sie das Theorem zur Parameterschätzung im Zweistichproben-T-Test Modell wieder.
21. Geben Sie das Theorem zur T-Teststatistik des Zweistichproben-T-Tests wieder.
22. Erläutern Sie mögliche Hypothesenszenarien eines Zweistichproben-T-Tests.
23. Geben Sie die Definition des zweiseitigen Zweistichproben-T-Tests (mit ungerichteter Hypothese) wieder.
24. Skizzieren Sie die Testgütefunktion des zweiseitigen Zweistichproben-T-Tests.
25. Geben Sie das Theorem zur Testumfangkontrolle im zweiseitigen Level-α_0-Zweistichproben-T-Test wieder.
26. Erläutern Sie das praktische Vorgehen bei Durchführung eines zweiseitigen Level-α_0-Zweistichproben-T-Tests.
27. Geben Sie die Definition des p-Wertes Werts für einen zweiseitigen Zweistichproben-T-Test wieder.
28. Von welchen Werten hängt die Powerfunktion eines zweiseitigen Zweistichproben-T-Tests ab?
29. Skizzieren Sie die Powerfunktion des zweiseitigen Zweistichproben-T-Tests bei fester Stichprobengröße.
30. Skizzieren Sie die Powerfunktion des zweiseitigen Zweistichproben-T-Tests bei festem Nichtzentralitätsparameter.
31. Betrachten Sie die Daten zum Alter der Patient:innen in der Face-to-Face und Online Therapie Bedingung im Beispieldatensatz. Erstellen gruppenspezifische Histogramme dieser Daten und evaluieren Sie gruppenspezifische deskriptive Statistiken zu diesen Daten. Führen Sie einen zweiseitigen Zweistichproben-T-Test mit Nullhypothesenparameter $\mu_0 = 0$ durch. Dokumentieren Sie Ihre Ergebnisse. Was folgern Sie aus den sich ergebenen Resultaten?

Überblick - SKF 1 - Extremszenarien

1. Erläutern Sie die Extremszenarien im Kontinuum von ALM Designs.

Extremszenario (1) Die Erwartungswerte aller Datenvariablen sind identisch.

$$y_i \sim N(\mu, \sigma^2) \text{ u.i.v. für } i = 1, \dots, n \Leftrightarrow$$
$$y = X\beta + \varepsilon, X := 1_n \in \mathbb{R}^{n\times 1}, \beta := \mu \in \mathbb{R}, \varepsilon \sim N(0_n, \sigma^2 I_n) \quad (1)$$

$\Rightarrow$ Jegliche Datenvariabilität wird dem Fehlerterm zugeschrieben.

Extremszenario (2) Die Erwartungswerte aller Datenvariablen sind paarweise verschieden

$$y_i \sim N(\mu_i, \sigma^2) \text{ u.v. für } i = 1, \dots, n \Leftrightarrow$$
$$y = X\beta + \varepsilon \text{ mit } X := I_n \in \mathbb{R}^{n\times n}, \beta := (\mu_1, \dots, \mu_n)^T \in \mathbb{R}^n, \varepsilon \sim N(0_n, \sigma^2 I_n) \quad (2)$$

$\Rightarrow$ Jegliche Datenvariabilität wird dem Erwartungswertparameter zugeschrieben.

$\Rightarrow$ Es gilt $\hat{\beta} = (I_n^T I_n)^{-1} I_n^T y = y$ und $\hat{\sigma}^2 = \frac{(y - I_n y)^T (y - I_n y)}{n-p} = 0$.

Beide Extremszenarien sind wissenschaftlich nicht ergiebig, da sie keine theoriegeleitete systematische Abhängigkeit zwischen der UV und der AV repräsentieren. Die im weiteren Verlauf der Vorlesung betrachteten ALM Designs liegen zwischen den beiden Extremszenarien und repräsentieren verschiedene Formen der systematischen Abhängigkeit zwischen UV und AV.

Anmerkung:

- Mit Datenvariablen sind hier die Komponenten des Datenvektors y, also die y_i gemeint

Wiederholung: ALM Modelle u.i.v. und für $p = 2$ ausgeschrieben

ALM für u.i.v. Zufallsvektor mit Erwartungswertparameter $\mu \in \mathbb{R}$ und Varianzparameter σ^2. D.h., für jede Komponente des Datenvektors gilt $y_i \sim N(\mu, \sigma^2)$ für $i = 1, \dots, n$. Äquivalent dazu ($\Leftrightarrow$) in Matrixschreibweise

$$y \sim N(X\beta, \sigma^2 I_n) \text{ mit } X := 1_n \in \mathbb{R}^{n \times 1}, \beta := \mu \in \mathbb{R}^1, \varepsilon \sim N(0_n, \sigma^2 I_n).$$

$$y = X\beta + \varepsilon \Leftrightarrow \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \beta + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \mu + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \begin{pmatrix} \mu + \varepsilon_1 \\ \vdots \\ \mu + \varepsilon_n \end{pmatrix}$$

ALM für u.v. Zufallsvektor mit "individuellen" Erwartungswertparametern μ_i und Varianzparameter σ^2. D.h., für jede Komponente des Datenvektors gilt $y_i \sim N(\mu_i, \sigma^2)$ für $i = 1, \dots, n$. Für ein Beispiel mit p = 2 ($\beta = \begin{pmatrix} \beta_1 & \beta_2 \end{pmatrix}^T$) gilt äquivalent dazu in Matrixschreibweise

$$y \sim N(X\beta, \sigma^2 I_n) \text{ mit } X \in \mathbb{R}^{n \times 2}, \beta \in \mathbb{R}^2, \varepsilon \sim N(0_n, \sigma^2 I_n).$$

$$y = X\beta + \varepsilon \Leftrightarrow \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} \\ \vdots & \vdots \\ x_{n1} & x_{n2} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_5 \end{pmatrix} = \begin{pmatrix} x_{11}\beta_1 + x_{12}\beta_2 + \varepsilon_1 \\ \vdots \\ x_{n1}\beta_1 + x_{n2}\beta_2 + \varepsilon_n \end{pmatrix}$$

Extremszenarien aus SKF 1 ausgeschrieben

Extremszenario (1) Die Erwartungswerte aller Datenvariablen sind identisch $\Rightarrow$ wir haben nur ein μ für alle Datenvariablen. $\Rightarrow$ Jegliche Datenvariabilität wird dem Fehlerterm zugeschrieben

$$y \sim N(X\beta, \sigma^2 I_n) \text{ mit } X := 1_n \in \mathbb{R}^{n\times 1}, \beta := \mu \in \mathbb{R}^1, \varepsilon \sim N(0_n, \sigma^2 I_n).$$

$$y = X\beta + \varepsilon \Leftrightarrow \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \beta + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \mu + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \begin{pmatrix} \mu + \varepsilon_1 \\ \mu + \varepsilon_2 \\ \vdots \\ \mu + \varepsilon_n \end{pmatrix}$$

Extremszenario (2) Die Erwartungswerte aller Datenvariablen sind paarweise verschieden $\Rightarrow$ jede Datenvariable y_i hat ein "individuelles" μ_i. $\Rightarrow$ Jegliche Datenvariabilität wird dem Erwartungswertparameter zugeschrieben. $\Rightarrow$ Es gilt $\hat{\beta} = y$ und $\hat{\sigma}^2 = 0$.

$$y \sim N(X\beta, \sigma^2 I_n) \text{ mit } X := I_n \in \mathbb{R}^{n\times n}, \beta := (\mu_1, \ldots, \mu_n)^T \in \mathbb{R}^2, \varepsilon \sim N(0_n, \sigma^2 I_n).$$

$$y = X\beta + \varepsilon \Leftrightarrow \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & \ldots & 0 \\ 0 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & 1 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 1\mu_1 + 0\mu_2 + \ldots + 0\mu_n + 0 \\ 0\mu_1 + 1\mu_2 + \ldots + 0\mu_n + 0 \\ \vdots \\ 0\mu_1 + 0\mu_2 + \ldots + 1\mu_n + 0 \end{pmatrix}$$

2. Erläutern Sie die Begriffe der faktoriellen und parametrischen ALM Designs.

Faktorielle ALM Designs

- Designmatrizen mit 1en und 0en, manchmal -1en.
- Betaparameter repräsentieren Gruppenerwartungswerte.
- Betaparameterschätzer repräsentieren Gruppenstichprobenmittel.

Parametrische ALM Designs

- Designmatrizen besitzen Spalten mit kontinuierlichen reellen Werten.
- Die Designmatrixsspalten werden *Regressoren*, *Prädiktoren*, oder *Kovariaten* genannt.
- Betaparameter repräsentieren Steigungsparameter.
- Betaparameterschätzer ergeben sich als normalisierte Regressor-Daten Kovarianzen.
- Es besteht ein enger Bezug zur Theorie der Korrelation.

Faktoriell-parametrische ALM Designs

- Designmatrizen mit mehreren faktoriellen und parametrischen Werten.
- Die parametrischen Regressoren werden oft als kontrollierte Kovariaten betrachtet.

3. Nennen Sie Beispiele für faktorielle, parametrische, und faktoriell-parametrische ALM Designs.

Faktorielle ALM Designs

- T-Tests, Einfaktorielle Varianzanalyse, Mehrfaktorielle Varianzanalyse
- Anwendungsbeispiel: Gruppenvergleich Treatment vs. No-Treatment

Parametrische ALM Designs

- Einfache lineare Regression, Multiple lineare Regression
- Anwendungsbeispiel: Effekt von Medikamentendosierung und/oder Anzahl Therapiestunden auf Symptomreduktion

Faktoriell-parametrische ALM Designs

- Kovarianzanalyse
- Anwendungsbeispiel: Kombination Treatment vs. No-Treatment und Medikamentendosierung

4. Erläutern Sie das Anwendungsszenario eines Einstichproben-T-Tests.

Im Anwendungsszenario eines Einstichproben-T-Tests betrachten wir **eine Gruppe** (Stichprobe) randomisierter experimenteller Einheiten. Wir nehmen an, dass die Datenpunkte unabhängig und identisch normalverteilte Realisierungen von ZVen ($y_i \sim N(\mu, \sigma^2)$) sind, wobei μ und σ^2 unbekannt sind.

Wir sind daran interessiert die *Unsicherheit*, die mit dem inferentiellen Vergleich von μ und μ_0 verbunden ist, im Sinne eines Hypothesentests zu *quantifizieren*.

5. Erläutern Sie mögliche Hypothesenszenarien eines Einstichproben-T-Tests.

Einfache Nullhypothese, einfache Alternativhypothese $H_0 : \mu = \mu_0, H_1 : \mu = \mu_1$

- Theoretisch wichtiges Szenario (Neymann-Pearson Lemma)
- Praktische Relevanz eher gering

Einfache Nullhypothese, zusammengesetzte Alternativhypothese $H_0 : \mu = \mu_0, H_1 : \mu \neq \mu_0$

- Zweiseitiger Einstichproben-T-Test mit ungerichteter Hypothese
- Ungerichtete Fragestellung nach einem Unterschied

Zusammengesetzte Nullhypothese/Alternativhypothese $H_0 : \mu \leq \mu_0, H_1 : \mu > \mu_0$

- Einseitiger Einstichproben-T-Test mit gerichteter Hypothese
- Gerichtete Fragestellung nach einem positiven Unterschied

Zusammengesetzte Nullhypothese/Alternativhypothese $H_0 : \mu \geq \mu_0, H_1 : \mu < \mu_0$

- Gerichtete Fragestellung nach einem negativen Unterschied
- Qualitativ äquivalente Theorie zum umgekehrten Fall

6. Geben Sie die Definition des Einstichproben-T-Test Modells wieder.

Definition (Einstichproben-T-Test Modell)

$y_i, i = 1, \ldots, n$ seien Zufallsvariablen, die die n Datenpunkte eines Einstichproben-T-Test Anwendungsszenarios modellieren. Dann hat das *Einstichproben-T-Test Modell* die strukturelle Form

$$y_i = \mu + \varepsilon_i \text{ mit } \varepsilon_i \sim N(0, \sigma^2) \text{ u.i.v. für } i = 1, \ldots, n \text{ mit } \mu \in \mathbb{R} \text{ und } \sigma^2 > 0, \tag{3}$$

die Datenverteilungsform

$$y_i \sim N(\mu, \sigma^2) \text{ u.i.v. für } i = 1, \ldots, n \text{ mit } \mu \in \mathbb{R} \text{ und } \sigma^2 > 0, \tag{4}$$

und für den Datenvektor $y = (y_1, \ldots, y_n)^T$ die Designmatrixform

$$y = X\beta + \varepsilon \text{ mit } X := 1_n \in \mathbb{R}^{n \times 1}, \beta := \mu \in \mathbb{R}, \varepsilon \sim N(0_n, \sigma^2 I_n), \text{ und } \sigma^2 > 0. \tag{5}$$

Anmerkung

- Das Modell ist identisch mit dem Modell unabhängiger und identisch normalverteilter Zufallsvariablen.

6. Geben Sie die Definition des Einstichproben-T-Test Modells wieder.

Modellformulierung eines beispielhaften wahren Modells + Datensimulation in R:

```
# Libraries
library(MASS)                                 # Multivariate Normalverteilung

# Modellformulierung
n      = 40                                   # Anzahl von Datenpunkten
p      = 1                                    # Anzahl von Betaparameter
X      = matrix(rep(1,n), nrow = n)           # Designmatrix
I_n    = diag(n)                              # n x n Einheitsmatrix
beta   = 5                                    # wahrer, aber unbekannter, Betaparameter
sigsqr = 14                                   # wahrer, aber unbekannter, Varianzparameter

# Datenrealisierung / Datensimulation
y      = mvrnorm(1, X %*% beta, sigsqr*I_n)  # eine Realisierung eines n-dimensionalen ZVs
```

7. Geben Sie das Theorem zur Parameterschätzung im Einstichproben-T-Test Modell wieder.

Theorem (Parameterschätzung im Einstichproben-T-Test Modell)

Gegeben sei die Designmatrixform des Einstichproben-T-Test Modells. Dann ergeben sich für den Betaparameterschätzer

$$\hat{\beta} = \frac{1}{n}\sum_{i=1}^{n} y_i =: \bar{y}, \tag{6}$$

und für den Varianzparameterschätzer

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n} (y_i - \bar{y})^2 =: s_y^2 \tag{7}$$

Anmerkung

- Die Formen von $\hat{\beta}$ und $\hat{\sigma}^2$ wurden in Einheit (6) Modellschätzung hergeleitet.
- $\bar{y}$ und s_y^2 bezeichnen das Stichprobenmittel und die Stichprobenvarianz der $y_1, ..., y_n$.

7. Geben Sie das Theorem zur Parameterschätzung im Einstichproben-T-Test Modell wieder.

Modellschätzung in R am Beispiel der ONLINE-Gruppe im Beispieldatensatz:

```r
# Dateneinlesen
fname       = file.path(getwd(), "9_Daten", "10_Einfaktorielle_Varianzanalyse_Daten.csv")   # Dateiname
D           = read.table(fname, sep = ",", header = TRUE)          # Dataframe
y           = D$BDI[D$Condition == "ONL"]                          # BDI Differenzwerte in der ONL Gruppe

# Modellformulierung
n          = length(y)                                             # Anzahl Datenpunkte
p          = 1                                                     # Anzahl Betaparameter
X          = matrix(rep(1,n), nrow = n)                            # Designmatrix

# Modellschätzung
beta_hat   = solve(t(X) %*% X) %*% t(X) %*% y                      # Betaparameterschätzer
eps_hat    = y - X %*% beta_hat                                    # Residuenvektor
sigsqr_hat = (t(eps_hat) %*% eps_hat) /(n-p)                       # Varianzparameterschätzer

# Ausgabe
cat("hat{beta}   : "  , beta_hat,                                  # Betaparameterschätzer
    "bar{y}      : ", mean(y),                                     # Stichprobenmittel
    "\nhat{sigsqr} : ", sigsqr_hat,                                # Varianzparameterschätzer
    "s_y^2       : ", var(y))                                      # Stichprobenvarianz
```

```
> hat{beta}   :  5.93 bar{y}      :  5.92
> hat{sigsqr} :  25.7 s_y^2       :  25.7
```

8. Geben Sie das Theorem zur T-Teststatistik des Einstichproben-T-Tests wieder.

Theorem (T-Teststatistik des Einstichproben-T-Tests)

Gegeben sei die Designmatrixform des Einstichproben-T-Test Modells. Dann ergibt sich für die T-Teststatistik mit

$$c := 1 \text{ und } c^T \beta_0 =: \mu_0, \tag{8}$$

dass

$$T = \sqrt{n}\left(\frac{\bar{y} - \mu_0}{s_y}\right) \tag{9}$$

und es gilt

$$T \sim t(\delta, n-1) \text{ mit } \delta = \sqrt{n}\left(\frac{\mu - \mu_0}{\sigma}\right) \tag{10}$$

Wiederholung: T-Teststatistik aus Einheit (7)

Theorem (T-Teststatistik)

Es sei

$$y = X\beta + \varepsilon \text{ mit } \varepsilon \sim N(0_n, \sigma^2 I_n) \tag{11}$$

das ALM in generativer Form. Weiterhin seien

$$\hat{\beta} := (X^T X)^{-1} X^T y \text{ und } \hat{\sigma}^2 := \frac{(y - X\hat{\beta})^T (y - X\hat{\beta})}{n - p} \tag{12}$$

die Betaparameter- und Varianzparameterschätzer, respektive. Schließlich sei für einen *Kontrastgewichtsvektor* $c \in \mathbb{R}^p$ und einen *Nullhypothesenbetaparameter* $\beta_0 \in \mathbb{R}^p$ die *T-Teststatistik* definiert als

$$T := \frac{c^T \hat{\beta} - c^T \beta_0}{\sqrt{\hat{\sigma}^2 c^T (X^T X)^{-1} c}}. \tag{13}$$

Dann gilt

$$T \sim t(\delta, n - p) \text{ mit } \delta := \frac{c^T \beta - c^T \beta_0}{\sqrt{\sigma^2 c^T (X^T X)^{-1} c}} \tag{14}$$

Herleitung (Beweis) der T-Teststatistik des Einstichproben-T-Tests

Mit dem T-Teststatistik Theorem in Einheit (7) Modellevaluation gilt

$$T = \frac{c^T\hat{\beta} - c^T\beta_0}{\sqrt{\hat{\sigma}^2 c^T(X^TX)^{-1}c}} = \frac{1^T\bar{y} - 1^T\mu_0}{\sqrt{s_y^2 1^T(1_n^T 1_n)^{-1}1}} = \sqrt{n}\left(\frac{\bar{y}-\mu_0}{s_y}\right).$$

Weiterhin gilt mit demselben Theorem

$$\delta = \frac{c^T\beta - c^T\beta_0}{\sqrt{\sigma^2 c^T(X^TX)^{-1}c}} = \frac{1^T\mu - 1^T\mu_0}{\sqrt{\sigma^2 1^T(1_n^T 1_n)^{-1}1}} = \sqrt{n}\left(\frac{\mu-\mu_0}{\sigma}\right)$$

Veranschaulichung Einstichproben T-Test mit $n = 5$

Anmerkung: Theoretische Aspekte sind türkis und mit Daten geschätzte Größen sind orange

Modellformulierung (Wie bei ALM für u.i.v. ZVen)

$$y \sim N(X\beta, \sigma^2 I_n) \text{ mit } X := 1_5 \in \mathbb{R}^{5\times 1}, \beta := \mu \in \mathbb{R}^1, \varepsilon \sim N(0_5, \sigma^2 I_5).$$

$$y = X\beta + \varepsilon \Leftrightarrow \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \beta + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \mu + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \end{pmatrix} = \begin{pmatrix} \mu + \varepsilon_1 \\ \mu + \varepsilon_2 \\ \mu + \varepsilon_3 \\ \mu + \varepsilon_4 \\ \mu + \varepsilon_5 \end{pmatrix}$$

Modellschätzung

$$\hat{\beta} = \bar{y}, \text{ und } \hat{\sigma}^2 = s_y^2$$

Veranschaulichung Einstichproben T-Test mit $n = 5$ (fortgeführt)

Modellevaluation

Für die Modellevaluation wollen wir die T-Teststatistik berechnen. Dafür wählen wir im Fall eines Einstichproben-T-Tests einen Kontrastgewichtsvektor von $c := 1$, wobei $c^T\beta_0 = \mu_0$).

Wir berechnen die T-Teststatistik T mit den Stichproben-Daten mit der Formel

$$T = \sqrt{n}\left(\frac{\bar{y} - \mu_0}{s_y}\right),$$

wobei wir eine Theorie darüber haben, wie T verteilt ist. Nämlich

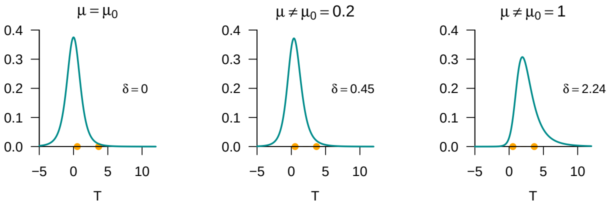
$$T \sim t(\delta, 5-1) \text{ mit } \delta = \sqrt{5}\left(\frac{\mu - \mu_0}{\sigma}\right)$$

Veranschaulichung Einstichproben T-Test mit $n = 5$ (fortgeführt)

Wir ziehen zwei mögliche Szenarien davon, was der wahre, aber unbekannte Parameter und die jeweils damit verbundenen (wahren, aber unbekannten) Verteilungen der T-Teststatistik sein könnten, in Betracht.

Nullhypothesen-Szenario: $c^T\beta = c^T\beta_0 \Leftrightarrow \mu = \mu_0 \Leftrightarrow \delta = 0 \Leftrightarrow T$ ist in Wahrheit zentral-t-verteilt.

Alternativhytpothesen-Szenario: $c^T\beta \neq c^T\beta_0 \Leftrightarrow \mu \neq \mu_0 \Leftrightarrow \delta \neq 0 \Leftrightarrow T$ ist in Wahrheit nicht-zentral-t-verteilt.



Wenn wir basierend auf Daten Parameterwerte geschätzt haben (in diesem Fall also Stichprobenmittel berechnet) und mit den Schätzern die T-Teststatistik berechnen, könnten wir zum Beispiel einen Wert von $T = 0.55$ oder einen Wert von $T = 3.67$ erhalten.

Anmerkung: Theoretische Aspekte sind türkis und mit Daten geschätzte Größen sind orange

9. Geben Sie die Definition des zweiseitigen Einstichproben-T-Tests (mit ungerichteter Hypothese) wieder.

Definition (Zweiseitiger Einstichproben-T-Tests)

Gegeben sei das Einstichproben-T-Test Modell. Für ein $\mu_0 \in \mathbb{R}$ seien die einfache Nullhypothese und die zusammengesetzte Alternativhypothese als

$$H_0 : \mu = \mu_0 \Leftrightarrow \Theta_0 := \{\mu_0\} \text{ und } H_1 : \mu \neq \mu_0 \Leftrightarrow \Theta_1 := \mathbb{R} \setminus \{\mu_0\}, \tag{15}$$

definiert. Weiterhin sei die T-Teststatistik definiert als

$$T := \sqrt{n}\left(\frac{\bar{y} - \mu_0}{s_y}\right) \tag{16}$$

Dann ist der *zweiseitige Einstichproben-T-Tests* definiert als der kritische Wert-basierten Test

$$\phi(y) := 1_{\{|T| \geq k\}} = \begin{cases} 1 & |T| \geq k \\ 0 & |T| < k \end{cases}. \tag{17}$$

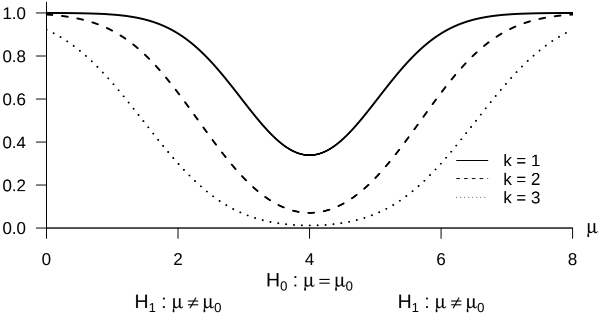
Bemerkungen

- Genauer gesagt handelt es sich um den *zweiseitigen Einstichproben-T-Tests mit ungerichteter Hypothese*.

10. Skizzieren Sie die Testgütefunktion des zweiseitigen Einstichproben-T-Tests.

Testgütefunktion q_ϕ für $\sigma^2 = 9, \mu_0 = 4, n = 12$ und $k = 1, 2, 3$.

$$q_\phi(\mu) = \mathbb{P}_\mu(\phi = 1)$$



11. Geben Sie das Theorem zur Testumfangkontrolle im zweiseitigen Level-α_0-Einstichproben-T-Test wieder.

Theorem (Testumfangkontrolle)

ϕ sei ein zweiseitiger Einstichproben-T-Test. Dann ist ϕ ein Level-α_0-Test mit Testumfang α_0, wenn der kritische Wert definiert ist durch

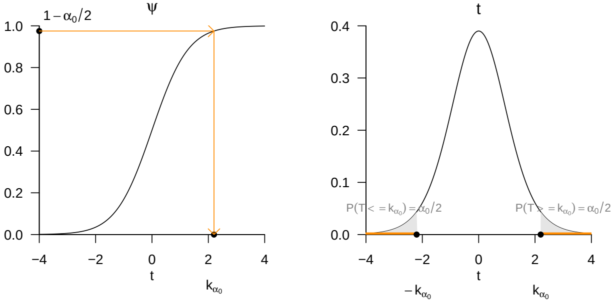
$$k_{\alpha_0} := \psi^{-1}\left(1 - \frac{\alpha_0}{2}; n-1\right), \tag{18}$$

wobei $\psi^{-1}(\cdot; n-1)$ die inverse KVF der t-Verteilung mit $n-1$ Freiheitsgraden ist.

Einstichproben-T-Tests - Modellevaluation - Testumfangkontrolle

Veranschaulichung

Wahl von $k_{\alpha_0} := \psi^{-1}(1-\alpha_0/2; n-1)$ mit $n = 12$, $\alpha_0 := 0.05$ und Ablehnungsbereich



12. Erläutern Sie das praktische Vorgehen bei Durchführung eines zweiseitigen Level-α_0-Einstichproben-T-Tests.

- Man nimmt an, dass ein Datensatz $\upsilon_1, ..., \upsilon_n$ eine Realisation von $y_i \sim N(\mu, \sigma^2)$ u.i.v. für $i = 1, ..., n$ mit unbekannten Parametern μ und $\sigma^2 > 0$ ist.
- Man möchte entscheiden ob für ein $\mu_0 \in \mathbb{R}$ eher $H_0 : \mu = \mu_0$ oder $H_1 : \mu \neq \mu_0$ zutrifft.
- Man wählt ein Signifikanzniveau α_0 und bestimmt den zugehörigen Freiheitsgradparameter-abhängigen kritischen Wert k_{α_0}.
- Anhand von $n, \mu_0, \bar{\upsilon}$ und s_υ berechnet man die Realisierung der T-Teststatistik

$$t := \sqrt{n}\left(\frac{\bar{\upsilon} - \mu_0}{s_\upsilon}\right) \tag{19}$$

- Wenn t größer-gleich k_{α_0} ist oder wenn t kleiner- gleich $-k_{\alpha_0}$ ist, lehnt man die Nullhypothese ab, andernfalls lehnt man sie nicht ab.
- Die oben entwickelte Theorie garantiert dann, dass man in höchstens $\alpha_0 \cdot 100$ von 100 Fällen die Nullhypothese fälschlicherweise ablehnt.

13. Geben Sie die Definition des p-Wertes Werts für einen zweiseitigen Einstichproben-T-Test wieder.

- Der p-Wert ist das kleinste Signifikanzlevel α_0, bei welchem man die Nullhypothese basierend auf einem vorliegendem Wert der Teststatistik ablehnen würde.
- Bei $T = t$ würde H_0 für jedes α_0 mit $|t| \geq \psi^{-1}(1 - \alpha_0/2; n - 1)$ abgelehnt werden. Für diese α_0 gilt

$$\alpha_0 \geq 2\mathbb{P}(T \geq |t|).$$

- Das kleinste $\alpha_0 \in [0, 1]$ mit $\alpha_0 \geq 2\mathbb{P}(T \geq |t|)$ ist dann $\alpha_0 = 2\mathbb{P}(T \geq |t|)$, also folgt

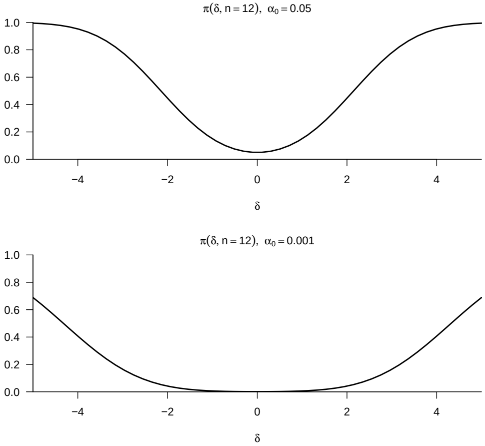
$$\text{p-Wert} = 2\mathbb{P}(T \geq |t|) = 2(1 - \psi(|t|; n - 1)).$$

- Im Gegensatz zum Z-Test hängt bei T-Tests der p-Wert auch von der Stichprobengröße ab.
- Zum Beispiel ist für $T = 2.00$ und $n = 10$ der p-Wert 0.076, für $T = 2.00$ und $n = 100$ ist der p-Wert dagegen 0.048.

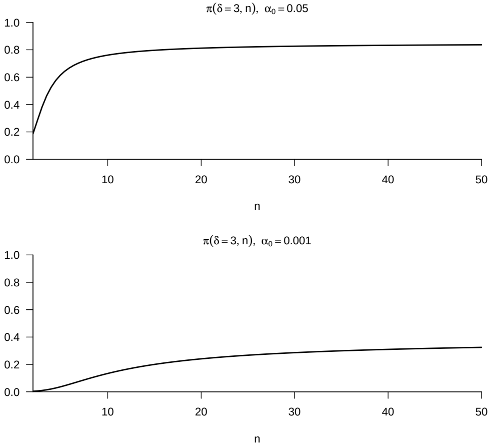
14. Von welchen Werten hängt die Powerfunktion eines zweiseitigen Einstichproben-T-Tests ab?

Die Powerfunktion des zweiseitigen T-Tests mit einfacher Nullhypothese hängt bei fesgelegtem α_0 vom wahren, aber unbekannten, Parameterwert $\delta = \sqrt{n}\frac{\mu - \mu_0}{\sigma}$ und von der Stichprobengröße n ab

15. Skizzieren Sie die Powerfunktion des zweiseitigen Einstichproben-T-Tests bei fester Stichprobengröße.



16. Skizzieren Sie die Powerfunktion des zweiseitigen Einstichproben-T-Tests bei festem Nichtzentralitätsparameter.



17. Betrachten Sie die PostBDI-PreBDI Differenzwertdaten der Waitlist Control Bedingung im Beispieldatensatz. Erstellen Sie ein Histogramm dieser Daten und evaluieren Sie Ihnen bekannte deskriptive Statistiken zu diesen Daten. Führen Sie einen zweiseitigen Einstichproben-T-Test mit Nullhypothesenparameter $\mu_0 = 0$ durch. Dokumentieren Sie Ihre Ergebnisse. Was folgern Sie aus den sich ergebenen Resultaten?

Dateneinlesen

```
fname       = file.path(getwd(), "9_Daten", "10_Einfaktorielle_Varianzanalyse_Daten.csv")
D           = read.table(fname, sep = ",", header = TRUE)
```

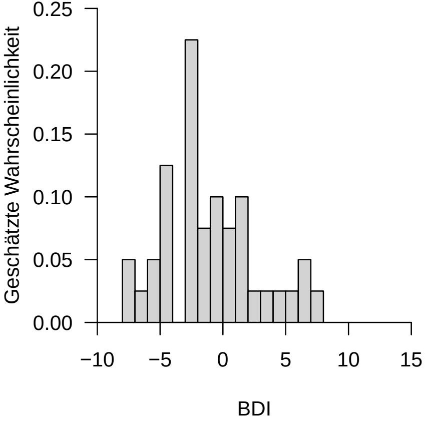
	X	ID	Condition	PreBDI	PostBDI	BDI	Age	Duration
81	81	81	WLC	27	31	-4	59	21
82	82	82	WLC	29	35	-6	28	21
83	83	83	WLC	33	35	-2	41	14
84	84	84	WLC	24	29	-5	64	18
85	85	85	WLC	31	23	8	74	19
86	86	86	WLC	30	38	-8	62	18
87	87	87	WLC	32	32	0	34	17
88	88	88	WLC	28	32	-4	59	17
89	89	89	WLC	30	30	0	37	12
90	90	90	WLC	30	32	-2	77	23
91	91	91	WLC	27	29	-2	30	13
92	92	92	WLC	33	31	2	45	18
93	93	93	WLC	33	29	4	35	22
94	94	94	WLC	31	26	5	26	19
95	95	95	WLC	34	33	1	70	17
96	96	96	WLC	31	35	-4	52	19
97	97	97	WLC	25	29	-4	60	21
98	98	98	WLC	27	26	1	44	14
99	99	99	WLC	25	32	-7	71	24
100	100	100	WLC	28	30	-2	64	18
101	101	101	WLC	27	25	2	41	23
102	102	102	WLC	29	30	-1	77	23
103	103	103	WLC	26	28	-2	59	21
104	104	104	WLC	29	29	0	22	20
105	105	105	WLC	27	26	1	56	20
106	106	106	WLC	35	36	-1	45	13
107	107	107	WLC	31	29	2	25	17
108	108	108	WLC	32	25	7	52	18
109	109	109	WLC	30	31	-1	78	23
110	110	110	WLC	34	31	3	63	21
111	111	111	WLC	27	27	0	53	21
112	112	112	WLC	28	21	7	35	18
113	113	113	WLC	34	28	6	67	13
114	114	114	WLC	27	32	-5	59	21
115	115	115	WLC	28	30	-2	70	17
116	116	116	WLC	28	30	-2	59	17
117	117	117	WLC	28	32	-4	49	12
118	118	118	WLC	28	26	2	50	14
119	119	119	WLC	31	33	-2	43	17
120	120	120	WLC	28	30	-2	47	21

Einstichproben-T-Tests - SKF 17 - Anwendungsbeispiel - Histogramm

Histogramm

```
# Datensatz von Interesse
BDI_WLC      = D$BDI[D$Condition == "WLC"]   # BDI Differenzwerte in der WLC Gruppe

# Histogrammparameter
h            = 1                             # gewünschte Klassenbreite
b_0          = min(BDI_WLC)                  # b_0
b_k          = max(BDI_WLC)                  # b_0
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(0,0.25)                     # y-Achsenlimits
xlimits      = c(-10,15)                      # x-Achsenlimits

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

# Histogramm
hist(
BDI_WLC,                                     # Delta.BDI Werte von Therapiebedingung i
breaks     = b,                              # Histogrammklassen
freq       = F,                              # normierte relative Häufigkeit
xlim       = xlimits,                        # x-Achsenlimits
ylim       = ylimits,                        # y-Achsenlimits
xlab       = "BDI",                          # x-Achsenbeschriftung
ylab       = "Geschätzte Wahrscheinlichkeit",# y-Achsenbeschriftung
main       = "")                             # Titelbeschriftung

# PDF Speicherung
dev.copy2pdf(
file         = file.path(getwd(), "9_Abbildungen", "alm_9_WLC_histogramm.pdf"),
width        = 4,
height       = 4)
```

Einstichproben-T-Tests - SKF 17 - Anwendungsbeispiel - Histogramm



Einstichproben-T-Tests - SKF 17 - Anwendungsbeispiel - Deskriptive

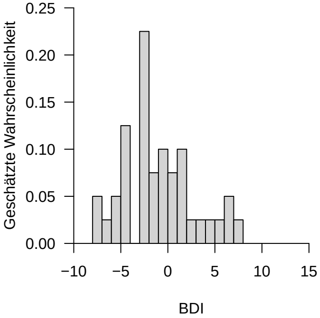
Deskriptive Statistiken

```
# Initialisierung eines Dataframes
tp              = c("WLC")                              # Therapiebedingungen
ntp             = length(tp)                            # Anzahl Therapiebedingungen
S               = data.frame(                           # Dataframeerzeugung
                    n         = rep(NaN,ntp),           # Stichprobengrößen
                    Max       = rep(NaN,ntp),           # Maxima
                    Min       = rep(NaN,ntp),           # Minima
                    Median    = rep(NaN,ntp),           # Mediane
                    Mean      = rep(NaN,ntp),           # Mittelwerte
                    Var       = rep(NaN,ntp),           # Varianzen
                    Std       = rep(NaN,ntp),           # Standardabweichungen
                    row.names = tp)                     # Therapiebedingungen

# Iterationen über Therapiebedingungen
for(i in 1:ntp){
  data          = D$BDI[D$Condition == tp[i]]           # Daten
  S$n[i]        = length(data)                          # Stichprobengröße
  S$Max[i]      = max(data)                             # Maxima
  S$Min[i]      = min(data)                             # Minima
  S$Median[i]   = median(data)                          # Mediane
  S$Mean[i]     = mean(data)                            # Mittelwerte
  S$Var[i]      = var(data)                             # Varianzen
  S$Std[i]      = sd(data)                              # Standardabweichungen
}
# Ausgabe
print.AsIs(S)
```

```
>      n Max Min Median   Mean  Var  Std
> WLC 40   8  -8     -1 -0.525 14.7 3.84
```

Einstichproben-T-Tests - SKF 17 - Anwendungsbeispiel - Deskriptive

Deskriptive Statistiken




```
# Ausgabe
print.AsIs(S)
```

```
>      n Max Min Median   Mean  Var  Std
> WLC 40   8  -8     -1 -0.525 14.7 3.84
```

Einstichproben-T-Tests - SKF 17 - Anwendungsbeispiel - T-Test (manuell)

T-Test

```r
# Dateneinlesen
fname       = file.path(getwd(), "9_Daten", "10_Einfaktorielle_Varianzanalyse_Daten.csv")  # Dateiname
D           = read.table(fname, sep = ",", header = TRUE)         # Dataframe
y           = D$BDI[D$Condition == "WLC"]                         # BDI Differenzwerte in der WLC Gruppe

# Modellformulierung
n          = length(y)                                            # Anzahl Datenpunkte
p          = 1                                                    # Anzahl Betaparameter
X          = matrix(rep(1,n), nrow = n)                           # Designmatrix

# Modellschätzung
beta_hat   = solve(t(X) %*% X) %*% t(X) %*% y                     # Betaparameterschätzer
eps_hat    = y - X %*% beta_hat                                   # Residuenvektor
sigsqr_hat = (t(eps_hat) %*% eps_hat) /(n-p)                      # Varianzparameterschätzer

# Modellevaluation
c          = matrix(c(1),nrow = p)                                # Kontrastgewichtsvektor
mu_0       = 0                                                    # Nullhypothese H_0
alpha_0    = 0.05                                                 # Signifikanzniveau
k_alpha_0  = qt(1 - (alpha_0/2), n-1)                             # kritischer Wert
t_num      = t(c) %*% beta_hat - mu_0                             # T-Teststatistik Zähler
t_den      = sqrt(sigsqr_hat %*% t(c)*solve(t(X) %*% X)%*%c)      # T-Teststatistik Nenner
t          = t_num/t_den                                          # 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))                               # p-Wert
```

```
> fg        =  39
> t         =  -0.866
> alpha_0   =  0.05
> k_alpha_0 =  2.02
> phi       =  0
> p-Wert    =  0.392
```

T-Test mit R-Funktion “‘t.test()“‘

```
# Automatischer Einstichproben-T-Test
varphi    = t.test(                             # ?t.test für Details
            y,                                  # Datensatz
            alternative = c("two.sided"),       # H_1: \mu \neq \mu_0
            mu          = 0,                    # \mu_0 (sic!)
            conf.level  = 1-alpha_0)            # \delta = 1 - \alpha_0 (sic!)

# Ausgabe
print(varphi)
```

```
>
>   One Sample t-test
>
> data:  y
> t = -0.9, df = 39, p-value = 0.4
> alternative hypothesis: true mean is not equal to 0
> 95 percent confidence interval:
>  -1.752  0.702
> sample estimates:
> mean of x
>    -0.525
```

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

```
> [1] "c(t = -0.865516026971161)"
```

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

```
> [1] "0.392048815383875"
```

17. Was folgern Sie aus den sich ergebenen Resultaten?

- Die Wahrscheinlichkeit, dass wir die Nullhypothese ablehnen würden, obwohl sie zutrifft liegt über dem von uns definierten Signifikanzlevel $\alpha_0 = 0.05$
- Im Sinne von Hypothesentests lehnen wir die Nullhypothese nicht ab.
- Die BDI Differenswerte in der WLC Gruppe wird als "nicht signifikant" nicht null bezeichnet.
- Im Rahmen der frequentistischen Inferenz wird die Schlussfolgerung gezogen, dass die BDI-Werte in der WLC Gruppe nach (post) Therapie "nicht signifikant" geringer sind als vor (prä) Therapie.
- Bei einem Signifikanzniveau von 0.05 nehmen wir an, dass wir in höchstens 0.05×100, also 5 von 100 Fällen die Nullhypothese fälschlich ablehnen würde.

18. Erläutern Sie das Anwendungsszenario eines Zweistichproben-T-Tests.

Im Anwendungsszenario eines Zweistichproben-T-Tests betrachten wir **zwei Gruppen** (Stichproben) randomisierter experimenteller Einheiten. Wir nehmen an, dass die Datenpunkte in beiden Gruppen unabhängig und identisch normalverteilte Realisierungen von ZVen sind.

Gruppe 1: $y_{1j} \sim N(\mu_1, \sigma^2)$, und Gruppe 2: $y_{1j} \sim N(\mu_2, \sigma^2)$ für $j = 1, ..., n$

Wir nehmen weiterhin an, dass wir daran interessiert sind, Unsicherheit, die mit dem inferentiellen Vergleich von μ_1 und μ_2 verbunden ist, im Sinne eines Hypothesentests zu quantifizieren.

19. Geben Sie die Definition des Zweistichproben-T-Test Modells wieder.

Definition (Zweistichproben-T-Test Modell)

y_{ij} mit $i = 1,2$ und $j = 1,...,n_i$ seien Zufallsvariablen, die die Datenpunkte eines Zweistichproben-T-Test Anwendungsszenarios modellieren. Dann hat das *Zweistichproben-T-Test Modell* die strukturelle Form

$$y_{ij} = \mu_i + \varepsilon_{ij} \text{ mit } \varepsilon_{ij} \sim N(0, \sigma^2) \text{ u.i.v. für } i = 1,2, j = 1,...,n_i \text{ mit } \mu_i \in \mathbb{R} \text{ und } \sigma^2 > 0, \tag{20}$$

die Datenverteilungsform

$$y_{ij} \sim N(\mu_i, \sigma^2) \text{ u.i.v. für } i = 1,2, j = 1,...,n_i \text{ mit } \mu_i \in \mathbb{R} \text{ und } \sigma^2 > 0, \tag{21}$$

und für den Datenvektor $y = (y_{11},...,y_{1n_1}, y_{21},...,y_{2n_2})^T$ und $n := n_1 + n_2$ die Designmatrixform

$$y = X\beta + \varepsilon \text{ mit } X := \begin{pmatrix} 1_{n_1} & 0_{n_1} \\ 0_{n_2} & 1_{n_2} \end{pmatrix} \in \mathbb{R}^{n \times 2}, \beta := \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \in \mathbb{R}^2, \varepsilon \sim N(0_n, \sigma^2 I_n), \sigma^2 > 0. \tag{22}$$

Bemerkungen

- i indiziert die Gruppen, j indiziert die Daten in jeder Gruppe.
- n_1 und n_2 repräsentieren die Gruppengrößen, n repräsentiert die Gesamtanzahl an Datenpunkten.
- Es ist $p = 2$.

20. Geben Sie das Theorem zur Parameterschätzung im Zweistichproben-T-Test Modell wieder.

Theorem (Parameterschätzung im Zweistichproben-T-Test Modell)

Gegeben sei die Designmatrixform des Zweistichproben-T-Test Modells. Dann ergeben sich für den Betaparameterschätzer

$$\hat{\beta} = \begin{pmatrix} \frac{1}{n_1}\sum_{j=1}^{n_1} y_{1j} \\ \frac{1}{n_2}\sum_{j=1}^{n_2} y_{2j} \end{pmatrix} =: \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} \tag{23}$$

und für den Varianzparameterschätzer

$$\hat{\sigma}^2_{12} = \frac{\sum_{j=1}^{n_1}(y_{1j}-\bar{y}_1)^2 + \sum_{j=1}^{n_2}(y_{2j}-\bar{y}_2)^2}{n_1+n_2-2} =: s^2_{12} \tag{24}$$

Bemerkungen

- $\bar{y}_1$ und $\bar{y}_2$ bezeichnen die gruppenspezifischen Stichprobenmittel.
- s^2_{12} wird als *gepoolte Stichprobenvarianz* bezeichnet.
- Für einen Datensatz $y = (y_1, y_2)^T \in \mathbb{R}^{n_1+n_2}$ gilt allgemeinen, dass $s^2_y \neq s^2_{12}$; die gepoolte Stichprobenvarianz und die Stichprobenvarianz eines konkatenierten Datensatzes sind im Allgemeinen also nicht identisch. Wir wollen das Konzept der gepoolten Stichprobenvarianz hier aber nicht weiter vertiefen.

21. Geben Sie das Theorem zur T-Teststatistik des Zweistichproben-T-Tests wieder.

Theorem (T-Teststatistik des Zweistichproben-T-Tests)

Gegeben sei die Designmatrixform des Zweistichproben-T-Tests. Dann ergibt sich für die T-Teststatistik mit

$$c := (1, -1)^T \text{ und } c^T\beta_0 =: \mu_0, \tag{25}$$

dass

$$T = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \left(\frac{\bar{y}_1 - \bar{y}_2 - \mu_0}{s_{12}} \right) \tag{26}$$

und es gilt

$$T \sim t(\delta, n_1 + n_2 - 2) \text{ mit } \delta = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \left(\frac{\mu_1 - \mu_2 - \mu_0}{\sigma} \right). \tag{27}$$

22. Erläutern Sie mögliche Hypothesenszenarien eines Zweistichproben-T-Tests.

$H_0 : \mu_1 - \mu_2 = \mu_0$ und $H_1 : \mu_1 - \mu_2 \neq \mu_0$

- Zweiseitiger Zweistichproben T-Test mit ungerichteter Hypothese
- Ungerichtete Fragestellungen nach einem Unterschied zwischen μ_1 und μ_2
- Für $\mu_0 := 0$ gelten dabei insbesondere
  - $H_0 : \mu_1 - \mu_2 = 0 \Leftrightarrow H_0 : \mu_1 = \mu_2$
  - $H_1 : \mu_1 - \mu_2 \neq 0 \Leftrightarrow H_1 : \mu_1 \neq \mu_2$

$H_0 : \mu_1 - \mu_2 \leq \mu_0$ und $H_1 : \mu_1 - \mu_2 > \mu_0$

- Einseitiger Zweistichproben T-Test mit gerichteter Hypothese
- Gerichtete Fragestellungen nach einem positiven Unterschied zwischen μ_1 und μ_2

$H_0 : \mu_1 - \mu_2 \geq \mu_0$ und $H_1 : \mu_1 - \mu_2 < \mu_0$

- Einseitiger Zweistichproben T-Test mit gerichteter Hypothese
- Gerichtete Fragestellungen nach einem negativen Unterschied zwischen μ_1 und μ_2

Veranschaulichung Zweistichproben T-Test mit $n = n_1 + n_2 = 2 + 3 = 5$

Anmerkung: Theoretische Aspekte sind türkis und mit Daten geschätzte Größen sind orange

Modellformulierung (Ähnlich wie bei ALM für u.i.v. ZVen, aber mit Aufteilung in 2 Gruppen, indiziert mit $i = 1, 2$ und Gruppengrößen $n_1 = 2$ und $n_2 = 3$)

$$y_{ij} = \mu_i + \varepsilon_{ij} \text{ mit } \varepsilon_{ij} \sim N(0, \sigma^2) u.i.v. \text{ für } i = 1, 2 \text{ und } j = 1, ..., n_i$$

$$y = X\beta + \varepsilon \text{ mit } X := \begin{pmatrix} 1_2 & 0_2 \\ 0_3 & 1_3 \end{pmatrix} \in \mathbb{R}^{5\times 2}, \beta := \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \in \mathbb{R}^2, \varepsilon \sim N(0_5, \sigma^2 I_5), \sigma^2 > 0.$$

$$y = X\beta + \varepsilon \Leftrightarrow \begin{pmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{22} \\ y_{23} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \varepsilon_{23} \end{pmatrix} = \begin{pmatrix} 1\mu_1 + 0\mu_2 + \varepsilon_{11} \\ 1\mu_1 + 0\mu_2 + \varepsilon_{12} \\ 0\mu_1 + 1\mu_2 + \varepsilon_{21} \\ 0\mu_1 + 1\mu_2 + \varepsilon_{22} \\ 0\mu_1 + 1\mu_2 + \varepsilon_{23} \end{pmatrix} = \begin{pmatrix} 1\mu_1 + \varepsilon_{11} \\ 1\mu_1 + \varepsilon_{12} \\ 1\mu_2 + \varepsilon_{21} \\ 1\mu_2 + \varepsilon_{22} \\ 1\mu_2 + \varepsilon_{23} \end{pmatrix}$$

Modellschätzung

$$\hat{\beta} = \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix}, \text{ und } \hat{\sigma}^2_{12} = s^2_{12}$$

Anmerkung:

- s^2_{12} ist die gepoolte Stichprobenvarianz

Veranschaulichung Zweistichproben T-Test mit $n = n_1 + n_2 = 2 + 3 = 5$

Modellevaluation

Für die Modellevaluation wollen wir die T-Teststatistik berechnen. Dafür wählen wir diesem Fall eines Zweistichproben-T-Tests einen Kontrastgewichtsvektor von $c := (1, -1)^T$, wobei $c^T \beta_0 = \mu_0$.

Wir berechnen die T-Teststatistik T mit den Stichproben-Daten mit der Formel

$$T = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \left(\frac{\bar{y}_1 - \bar{y}_2 - \mu_0}{s_{12}} \right),$$

wobei wir eine Theorie darüber haben, wie T verteilt ist. Nämlich

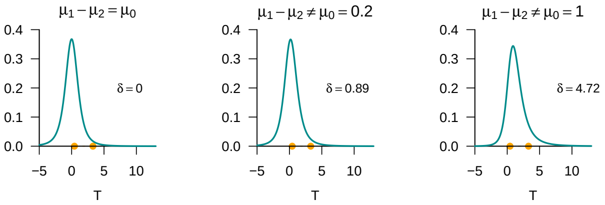
$$T \sim t(\delta, 2 + 3 - 1) \text{ mit } \delta = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \left(\frac{\mu_1 - \mu_2 - \mu_0}{\sigma} \right)$$

Veranschaulichung Zweistichproben T-Test mit $n = n_1 + n_2 = 2 + 3 = 5$

Wir ziehen zwei mögliche Szenarien davon, was der wahre, aber unbekannte Parameter sein könnte und die damit verbundenen jeweiligen (wahren, aber unbekannten) Verteilungen der T-Teststatistik in Betracht.

Nullhypothesen-Szenario: $c^T\beta = c^T\beta_0$ bzw. $\mu_1 - \mu_2 = \mu_0 \Leftrightarrow \delta = 0 \Leftrightarrow T$ ist in Wahrheit zentral-t-verteilt.

Alternativhytpothesen-Szenario: $c^T\beta = c^T\beta_0$ bzw. $\mu_1 - \mu_2 \neq 0 \Leftrightarrow \delta \neq 0 \Leftrightarrow T$ ist in Wahrheit nicht-zentral-t-verteilt.



Wenn wir basierend auf Daten Parameterwerte geschätzt haben (in diesem Fall also Stichprobenmittel berechnet) und mit den Schätzern die T-Teststatistik berechenen, könnten wir zum Beispiel einen Wert von $T = 0.43$ oder einen Wert von $T = 4.22$ erhalten.

Anmerkung: Theoretische Aspekte sind türkis und mit Daten geschätzte Größen sind orange

23. Geben Sie die Definition des zweiseitigen Zweistichproben-T-Tests (mit ungerichteter Hypothese) wieder.

Definition (Zweiseitiger Zweistichproben-T-Test)

Gegeben sei das Zweistichproben-T-Test Modell. Für ein $\mu_0 \in \mathbb{R}$ seien die einfache Nullhypothese und die zusammgensetzte Alternativhypothese gegeben durch

$$H_0 : \mu_1 - \mu_2 = \mu_0 \Leftrightarrow \Theta_0 := \{(\mu_1, \mu_2) \in \mathbb{R}^2 | \mu_1 - \mu_2 = \mu_0\} \tag{28}$$

und

$$H_1 : \mu_1 - \mu_2 \neq \mu_0 \Leftrightarrow \Theta_1 := \{(\mu_1, \mu_2) \in \mathbb{R}^2 | \mu_1 - \mu_2 \neq \mu_0\}, \tag{29}$$

respektive. Weiterhin sei die T-Teststatistik definiert durch

$$T := \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \left(\frac{\bar{y}_1 - \bar{y}_2}{s_{12}} \right) \tag{30}$$

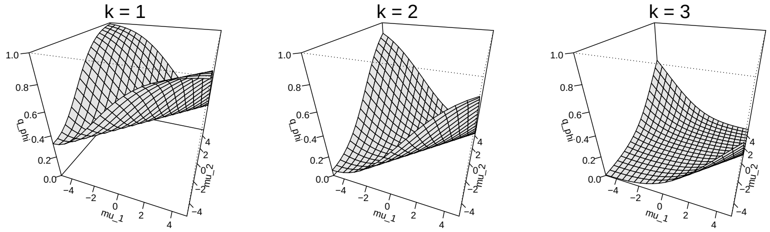
Dann ist der zweiseitige *Zweistichproben-T-Teststatistik* definiert als der kritischen Wert-basierte Test

$$\phi(y) := 1_{\{|T| \geq k\}}. \tag{31}$$

24. Skizzieren Sie die Testgütefunktion des zweiseitigen Zweistichproben-T-Tests.

Testgütefunktion q_ϕ für $\sigma^2 = 9, n_1 = 12, n_2 = 12.$

$$q_\phi(\mu) = \mathbb{P}_\mu(\phi = 1)$$



25. Geben Sie das Theorem zur Testumfangkontrolle im zweiseitigen Level-α_0-Zweistichproben-T-Test wieder.

Theorem (Testumfangkontrolle)

ϕ sei der im obigen Testszenario definierte Test. Dann ist ϕ ein Level-α_0-Test mit Testumfang α_0, wenn der kritische Wert definiert ist durch

$$k_{\alpha_0} := \psi^{-1}\left(1 - \frac{\alpha_0}{2}; n_1 + n_2 - 2\right), \tag{32}$$

wobei $\psi^{-1}(\cdot; n_1 + n_2 - 2)$ die inverse KVF der t-Verteilung mit $n_1 + n_2 - 2$ Freiheitsgraden ist.

Bemerkungen

- Das Resultat folgt in Analogie zum Einstichproben-T-Test.
- Im Vergleich zum Einstichproben-T-Testfall gilt lediglich

$$n - 1 \hookrightarrow n_1 + n_2 - 2. \tag{33}$$

26. Erläutern Sie das praktische Vorgehen bei Durchführung eines zweiseitigen Level-α_0-Zweistichproben-T-Tests.

- Man nimmt an, dass die Daten zweier Gruppen $\upsilon_{11}, \dots, \upsilon_{1n_1}$ und $\upsilon_{21}, \dots, \upsilon_{2n_2}$ Realisationen von $y_{1j} \sim N(\mu_1, \sigma^2)$ u.i.v. für $j = 1, \dots, n_1$ und $y_{2j} \sim N(\mu_2, \sigma^2)$ u.i.v. für $j = 1, \dots, n_2$ mit unbekannten Parametern μ_1, μ_2, σ^2 sind.
- Man möchte entscheiden, ob eher $H_0 : \mu_1 - \mu_2 = \mu_0$ oder $H_1 : \mu_1 - \mu_2 \neq \mu_0$ zutrifft.
- Man wählt ein Signifikanzniveau α_0 und bestimmt den zugehörigen Freiheitsgradparameter-abhängigen kritischen Wert k_{α_0}. Zum Beispiel gilt bei Wahl von $\alpha_0 := 0.05$ und $n_1 = 12, n_2 = 12$, also Freiheitsgradparameter 12+12-2 = 22, dass $k_{0.05} = \psi^{-1}(1 - 0.05/2; 22) \approx 2.07$ ist.
- Anhand von $n_1, n_2, \bar{\upsilon}_1, \bar{\upsilon}_2$ und der gepoolten Stichprobenstandardabweichung s_{12} berechnet man die Realisierung der Zweistichproben-T-Teststatistik

$$t := \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \left(\frac{\bar{\upsilon}_1 - \bar{\upsilon}_2}{s_{12}} \right) \tag{34}$$

- Wenn t größer-gleich k_{α_0} ist oder wenn t kleiner- gleich $-k_{\alpha_0}$ ist, lehnt man die Nullhypothese ab, andernfalls lehnt man sie nicht ab.
- Die oben entwickelte Theorie des Zweistichproben-T-Tests garantiert dann, dass man in höchstens $\alpha_0 \cdot 100$ von 100 Fällen die Nullhypothese fälschlicherweise ablehnt.

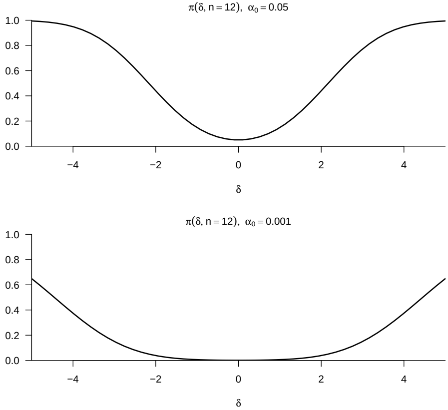
27. Geben Sie die Definition des p-Wertes Werts für einen zweiseitigen Zweistichproben-T-Test wieder.

- Per Definition ist der p-Wert das kleinste Signifikanzlevel α_0, bei welchem man die Nullhypothese basierend auf einem vorliegendem Wert der Teststatistik ablehnen würde.
- Bei $T = t$ würde H_0 für jedes α_0 mit $|t| \geq \psi^{-1}(1 - \alpha_0/2; n_1 + n_2 - 2)$ abgelehnt werden. Für diese α_0 gilt, wie bereits mehrfach gezeigt,
$$\alpha_0 \geq 2\mathbb{P}(T \geq |t|).$$
- Das kleinste $\alpha_0 \in [0, 1]$ mit $\alpha_0 \geq 2\mathbb{P}(T \geq |t|)$ ist dann $\alpha_0 = 2\mathbb{P}(T \geq |t|)$, also folgt
$$\text{p-Wert} = 2\mathbb{P}(T \geq |t|) = 2(1 - \psi(|t|; n_1 + n_2 - 2)).$$
- Im Vergleich zum Einstichprobenfall gilt lediglich $n \hookrightarrow n_1 + n_2 - 2$.

28. Von welchen Werten hängt die Powerfunktion eines zweiseitigen Zweistichproben-T-Tests ab?

Bei festgelegten α_0 hängt die Powerfunktion des zweiseitigen Zweistichproben-T-Tests mit einfacher Nullhypothese vom unbekannten Wert d und von der Summe der Stichprobengrößen n ab. De-facto handelt es sich also um die gleiche Powerfunktion wie beim zweiseitigen Einstichproben-T-Test mit dem einzigen Unterschied, dass für den Freiheitsgradparameter $n-2$ anstelle von $n-1$ gilt.

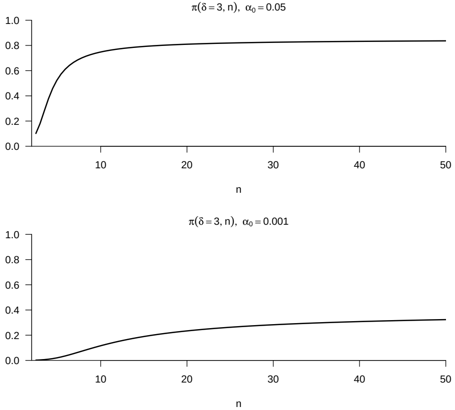
29. Skizzieren Sie die Powerfunktion des zweiseitigen Zweistichproben-T-Tests bei fester Stichprobengröße.



Anmerkung:

- Untescheid zu SKF 15: Freiheitsgradparameter $n-2$ anstelle von $n-1$

30. Skizzieren Sie die Powerfunktion des zweiseitigen Zweistichproben-T-Tests bei festem Nichtzentralitätsparameter.



Anmerkung:

- Untescheid zu SKF 15: Freiheitsgradparameter $n-2$ anstelle von $n-1$

31. Betrachten Sie die Daten zum Alter der Patient:innen in der Face-to-Face und Online Therapie Bedingung im Beispieldatensatz. Erstellen gruppenspezifische Histogramme dieser Daten und evaluieren Sie gruppenspezifische deskriptive Statistiken zu diesen Daten. Führen Sie einen zweiseitigen Zweistichproben-T-Test mit Nullhypothesenparameter $\mu_0 = 0$ durch. Dokumentieren Sie Ihre Ergebnisse. Was folgern Sie aus den sich ergebenen Resultaten?

Zweistichproben-T-Tests - SKF 31 - Anwendungsbeispiel - Dateneinlesen

Dateneinlesen | $j = 1, \ldots, 20$ für jede Gruppe

```
fname       = file.path(getwd(), "9_Daten", "10_Einfaktorielle_Varianzanalyse_Daten.csv")
D           = read.table(fname, sep = ",", header = TRUE)
```

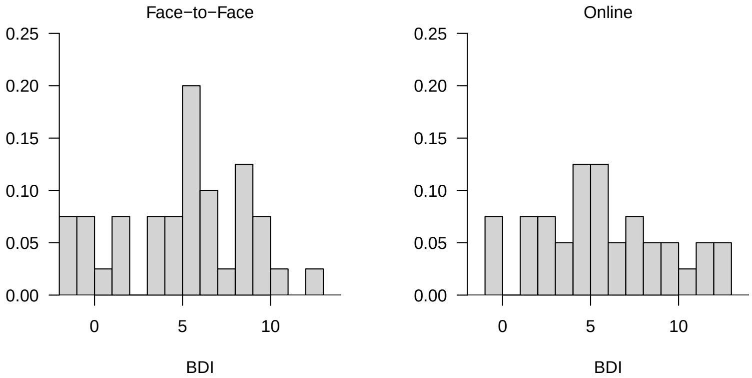
	X	ID	Condition	PreBDI	PostBDI	BDI	Age	Duration
1	1	1	F2F	29	25	4	66	22
2	2	2	F2F	32	31	1	29	23
3	3	3	F2F	28	26	2	71	14
4	4	4	F2F	36	26	10	77	21
5	5	5	F2F	32	27	5	55	24
6	6	6	F2F	28	29	-1	50	24
7	7	7	F2F	33	27	6	31	16
8	8	8	F2F	33	27	6	20	17
9	9	9	F2F	33	25	8	73	23
10	10	10	F2F	30	26	4	28	13
11	11	11	F2F	36	27	9	21	23
12	12	12	F2F	32	25	7	76	16
13	13	13	F2F	29	29	0	38	18
14	14	14	F2F	24	22	2	30	18
15	15	15	F2F	35	28	7	44	21
16	16	16	F2F	31	22	9	48	17
17	17	17	F2F	31	26	5	46	14
18	18	18	F2F	34	25	9	51	16
19	19	19	F2F	34	25	9	71	20
20	20	20	F2F	33	27	6	23	13
41	41	41	ONL	31	27	4	70	23
42	42	42	ONL	31	25	6	63	16
43	43	43	ONL	34	29	5	41	23
44	44	44	ONL	34	29	5	28	14
45	45	45	ONL	30	24	6	43	22
46	46	46	ONL	30	33	-3	76	15
47	47	47	ONL	33	25	8	68	12
48	48	48	ONL	34	21	13	66	22
49	49	49	ONL	32	26	6	77	20
50	50	50	ONL	35	27	8	80	23
51	51	51	ONL	33	33	0	56	20
52	52	52	ONL	30	26	4	22	22
53	53	53	ONL	33	27	6	40	16
54	54	54	ONL	28	26	2	37	17
55	55	55	ONL	37	25	12	27	19
56	56	56	ONL	38	26	12	23	13
57	57	57	ONL	31	28	3	42	14
58	58	58	ONL	29	33	-4	40	19
59	59	59	ONL	34	29	5	30	21
60	60	60	ONL	32	30	2	57	22

Zweistichproben-T-Tests - SKF 31 - Anwendungsbeispiel - Histogramme

Histogramme

```r
# Histogrammparameter
h           = 1                              # gewünschte Klassenbreite
b_0         = min(D$BDI)                     # b_0
b_k         = max(D$BDI)                     # b_0
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,.25)                       # y-Achsenlimits
xlimits     = c(-2,14)                       # x-Achsenlimits
therapie    = c("F2F" , "ONL")               # Therapiebedingungen
labs        = c("Face-to-Face",              # Abbildungslabel
                "Online")
# Abbildungsparameter
par(                                         # für Details siehe ?par
mfcol       = c(1,2),                        # 1 x 2 Panelstruktur
family      = "sans",                        # Serif-freier Fonttyp
pty         = "m",                           # Maximale Abbildungsregion
bty         = "l",                           # L förmige Box
las         = 1,                             # Horizontale Achsenbeschriftung
xaxs        = "i",                           # x-Achse bei y = 0
yaxs        = "i",                           # y-Achse bei x = 0
font.main   = 1,                             # Non-Bold Titel
cex         = 1,                             # Textvergrößerungsfaktor
cex.main    = 1)                             # Titeltextvergrößerungsfaktor
# Iteration über Therapiebedingungen
for(i in 1:2){
  hist(
  D$BDI[D$Condition == therapie[i]],         # Werte von Therapiebedingung i
  breaks     = b,                            # Histogrammklassen
  freq       = F,                            # normierte relative Häufigkeit
  xlim       = xlimits,                      # x-Achsenlimits
  ylim       = ylimits,                      # y-Achsenlimits
  xlab       = TeX("BDI"),                   # x-Achsenbeschriftung
  ylab       = "",                           # y-Achsenbeschriftung
  main       = labs[i])                      # Titelbeschriftung
}
# PDF Speicherung
dev.copy2pdf(
file        = file.path(getwd(), "9_Abbildungen", "alm_9_F2F_ONL_histogramme_skf31.pdf"),
width       = 8,
height      = 4)
```

Zweistichproben-T-Tests - SKF 31 - Anwendungsbeispiel - Histogramme



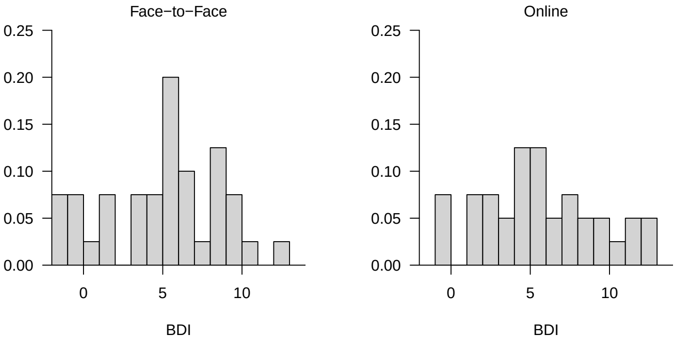
Zweistichproben-T-Tests - SKF 31 - Anwendungsbeispiel - Deskriptive

Deskriptive Statistiken

```
# Initialisierung eines Dataframes
tp            = c("F2F", "ONL")                          # Therapiebedingungen
ntp           = length(tp)                               # Anzahl Therapiebedingungen
S             = data.frame(                              # Dataframeerzeugung
                  n         = rep(NaN,ntp),              # Stichprobengrößen
                  Max       = rep(NaN,ntp),              # Maxima
                  Min       = rep(NaN,ntp),              # Minima
                  Median    = rep(NaN,ntp),              # Mediane
                  Mean      = rep(NaN,ntp),              # Mittelwerte
                  Var       = rep(NaN,ntp),              # Varianzen
                  Std       = rep(NaN,ntp),              # Standardabweichungen
                  row.names = tp)                        # Therapiebedingungen

# Iterationen über Therapiebedingungen
for(i in 1:ntp){
  data        = D$BDI[D$Condition == tp[i]]              # Daten
  S$n[i]      = length(data)                             # Stichprobengröße
  S$Max[i]    = max(data)                                # Maxima
  S$Min[i]    = min(data)                                # Minima
  S$Median[i] = median(data)                             # Mediane
  S$Mean[i]   = mean(data)                               # Mittelwerte
  S$Var[i]    = var(data)                                # Varianzen
  S$Std[i]    = sd(data)                                 # Standardabweichungen
}
```

Zweistichproben-T-Tests - SKF 31 - Anwendungsbeispiel - Deskriptive

Deskriptive Statistiken




```
# Ausgabe
print.AsIs(S)
```

```
>      n Max Min Median Mean  Var  Std
> F2F 40  13  -3      6 5.28 14.8 3.85
> ONL 40  18  -5      6 5.92 25.7 5.07
```

Zweistichproben-T-Tests - SKF 31 - Anwendungsbeispiel - T-Test (manuell)

T-Test

```r
# Dateneinlesen
fname      = file.path(getwd(), "9_Daten", "data_9_t_tests.csv")  # Dateiname
D          = read.table(fname, sep = ",", header = TRUE)           # Dataframe
y_1        = D$BDI[D$Condition == "F2F"]                           # BDI Differenzwerte in der F2F Gruppe
y_2        = D$BDI[D$Condition == "ONL"]                           # BDI Differenzwerte in der ONL Gruppe

# Modellformulierung
n_1        = length(y_1)                                           # Anzahl Datenpunkte Gruppe 1 (F2F)
n_2        = length(y_1)                                           # Anzahl Datenpunkte Gruppe 2 (ONL)
n          = n_1 + n_2                                             # Gesamtanzahl Datenpunkte
y          = matrix(c(y_1, y_2), nrow = n)                         # Datenvektor
p          = 2                                                     # Anzahl Betaparameter
X          = matrix(c(rep(1,n_1), rep(0,n_1),                      # Designmatrix
                 rep(0,n_2), rep(1,n_2)),
                 nrow  = n)

# Modellschätzung
beta_hat   = solve(t(X) %*% X) %*% t(X) %*% y                      # Betaparameterschätzer
eps_hat    = y - X %*% beta_hat                                    # Residuenvektor
sigsqr_hat = (t(eps_hat) %*% eps_hat) /(n-p)                       # Varianzparameterschätzer

# Modellevaluation
c          = matrix(c(1,-1), nrow = 2)                             # Kontrastgewichtsvektor
mu_0       = 0                                                     # Nullhypothese H_0
alpha_0    = 0.05                                                  # Signifikanzniveau
k_alpha_0  = qt(1 - (alpha_0/2), n-1)                              # kritischer Wert
t_num      = t(c) %*% beta_hat - mu_0                              # T-Teststatistik Zähler
t_den      = sqrt(sigsqr_hat*t(c) %*% solve(t(X) %*% X)%*%c)       # T-Teststatistik Nenner
t          = t_num/t_den                                           # 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
```

```
> fg        =  78
> t         =  -0.646
> alpha_0   =  0.05
> k_alpha_0 =  1.99
> phi       =  0
> p-Wert    =  0.52
```

T-Test mit R-Funktion “‘t.test()“‘

```
# Automatischer Zweistichproben-T-Test
varphi    = t.test(                                 # ?t.test für Details
            y_1,                                    # Datensatz y_1
            y_2,                                    # Datensatz y_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:  y_1 and y_2
> t = -0.6, df = 78, p-value = 0.5
> alternative hypothesis: true difference in means is not equal to 0
> 95 percent confidence interval:
>  -2.65  1.35
> sample estimates:
> mean of x mean of y
>      5.28      5.92
```

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

```
> [1] "c(t = -0.646128613569434)"
```

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

```
> [1] "0.520092604281206"
```

31. Was folgern Sie aus den sich ergebenen Resultaten?

- Die Wahrscheinlichkeit, dass wir die Nullhypothese ablehnen würden, obwohl sie zutrifft liegt über dem von uns definierten Signifikanzlevel $\alpha_0 = 0.05$
- Im Sinne von Hypothesentests lehnen wir die Nullhypothese nicht ab.
- Die Differenz zwischen den BDI Differenswerten in der F2F und der ONLINE Gruppe ist "nicht signifikant" nicht null.
- Im Rahmen der frequentistischen Inferenz wird die Schlussfolgerung gezogen, dass die post-BDI-Werte in in den beiden Gruppen sich nicht signifikant unterscheiden.
- Bei einem Signifikanzniveau von 0.05 nehmen wir an, dass wir in höchstens 0.05×100, also 5 von 100 Fällen die Nullhypothese fälschlich ablehnen würden.