Programmierung und Deskriptive Statistik

BSc Psychologie WiSe 2022/23

Belinda Fleischmann

(11) Anwendungsbeispiel

Beispieldatensatz

Datenvorverarbeitung

Deskriptive Statistiken

Visualisierung

Parameterschätzung

Konfidenzintervalle

Hypothesentests

Beispieldatensatz

Datenvorverarbeitung

Deskriptive Statistiken

Visualisierung

Parameterschätzung

Konfidenzintervalle

Hypothesentests

Evidenzbasierte Evaluation von Psychotherapieformen bei Depression

Welche Therapieform ist bei Depression wirksamer?

Online Psychotherapie

Klassische Psychotherapie

Evidenzbasierte Evaluation von Psychotherapieformen bei Depression

Becks Depressions-Inventar (BDI) zur Depressionsdiagnostik

BDI-II Fragebogen

Name | Alter | Geschlecht m / w | Datum

Anleitung: Dieser Fragebogen enthält 21 Gruppen von Aussagen. Bitte lesen Sie jede dieser Gruppen vor Aussagen sorgfältig durch und suchen Sie sich dann in jeder Gruppe **eine Aussage** heraus, die am bester beschreibt, wie Sie sich **in den letzten zwei Wochen, einschließlich heute, gefühlt haben**. Kreuzen Sie die Zah neben der Aussage an, die Sie sich herausgesucht haben (0, 1, 2 oder 3). Falls in einer Gruppe mehrere Aussager gleichermaßen auf Sie zutreffen, kreuzen Sie die Aussage mit der höheren Zahl an. Achten Sie bitte darauf, das Sie in jeder Gruppe nicht mehr als eine Aussage ankreuzen, das gilt auch für Gruppe 16 (Veränderungen de Schlafgewohnheiten) oder Gruppe 18 (Veränderungen des Appetits).

1.) Traurigkeit
- **0** Ich bin nicht traurig.
- **1** Ich bin oft traurig.
- **2** Ich bin ständig traurig.
- **3** Ich bin so traurig oder unglücklich, dass ich es nicht aushalte.

2.) Pessimismus
- **0** Ich sehe nicht mutlos in die Zukunft.
- **1** Ich sehe mutloser in die Zukunft als sonst.
- **2** Ich bin mutlos und erwarte nicht, dass meine Situation besser wird.
- **3** Ich glaube, dass meine Zukunft hoffnungslos ist und nur noch schlechter wird.

3.) Versagensgefühle
- **0** Ich fühle mich nicht als Versager.
- **1** Ich habe häufiger Versagensgefühle.
- **2** Wenn ich zurückblicke, sehe ich eine Menge Fehlschläge.
- **3** Ich habe das Gefühl, als Mensch ein völliger Versager zu sein.

4.) Verlust von Freude
- **0** Ich kann die Dinge genauso gut genießen wie früher.
- **1** Ich kann die Dinge nicht mehr so genießen wie früher.
- **2** Dinge, die mir früher Freude gemacht haben, kann ich kaum mehr genießen.
- **3** Dinge, die mir früher Freude gemacht haben, kann ich überhaupt nicht mehr genießen.

5.) Schuldgefühle
- **0** Ich habe keine besonderen Schuldgefühle.
- **1** Ich habe oft Schuldgefühle wegen Dingen, die ich getan habe oder hätte tun sollen.
- **2** Ich habe die meiste Zeit Schuldgefühle.
- **3** Ich habe ständig Schuldgefühle.

6.) Bestrafungsgefühle
- **0** Ich habe nicht das Gefühl, für etwas bestraft zu sein.
- **1** Ich habe das Gefühl, vielleicht bestraft zu werden.
- **2** Ich erwarte, bestraft zu werden.
- **3** Ich habe das Gefühl, bestraft zu sein.

7.) Selbstablehnung
- **0** Ich halte von mir genauso viel wie immer.
- **1** Ich habe Vertrauen in mich verloren.
- **2** Ich bin von mir enttäuscht.
- **3** Ich lehne mich völlig ab.

8.) Selbstvorwürfe
- **0** Ich kritisiere oder tadle mich nicht meh als sonst.
- **1** Ich bin mir gegenüber kritischer als sonst.
- **2** Ich kritisiere mich für all meine Mängel.
- **3** Ich gebe mir die Schuld für alles Schlimme, was passiert.

9.) Selbstmordgedanken
- **0** Ich denke nicht daran, mir etwas anzutun.
- **1** Ich denke manchmal an Selbstmord, aber ich würde es nicht tun.
- **2** Ich möchte mich am liebsten umbringe
- **3** Ich würde mich umbringen, wenn ich die Gelegenheit dazu hätte.

10.) Weinen
- **0** Ich weine nicht öfter als früher.
- **1** Ich weine jetzt mehr als früher.
- **2** Ich weine beim geringsten Anlass.
- **3** Ich möchte gern weinen, aber ich kann nicht.

PEARSON PsychCorp

Summe Seite 1:

Bitte wende

11.) Unruhe
- **0** Ich bin nicht unruhiger als sonst.
- **1** Ich bin unruhiger als sonst.
- **2** Ich bin so unruhig, dass es mir schwerfällt, still zu sitzen.
- **3** Ich bin so unruhig, dass ich mich ständig bewegen oder etwas tun muss.

12.) Interessenverlust
- **0** Ich habe das Interesse an anderen Menschen oder an Tätigkeiten nicht verloren.
- **1** Ich habe weniger Interesse an anderen Menschen oder an Dingen als sonst.
- **2** Ich habe das Interesse an anderen Menschen oder Dingen zum größten Teil verloren.
- **3** Es fällt mir schwer, mich überhaupt für irgend etwas zu interessieren.

13.) Entschlussunfähigkeit
- **0** Ich bin so entschlussfreudig wie immer.
- **1** Es fällt mir schwerer als sonst, Entscheidungen zu treffen.
- **2** Es fällt mir sehr viel schwerer als sonst, Entscheidungen zu treffen.
- **3** Ich habe Mühe, überhaupt Entscheidungen zu treffen.

14.) Wertlosigkeit
- **0** Ich fühle mich nicht wertlos.
- **1** Ich halte mich für weniger wertvoll und nützlich als sonst.
- **2** Verglichen mit anderen Menschen fühle ich mich viel weniger wert.
- **3** Ich fühle mich völlig wertlos.

15.) Energieverlust
- **0** Ich habe so viel Energie wie immer.
- **1** Ich habe weniger Energie als sonst.
- **2** Ich habe so wenig Energie, dass ich kaum noch etwas schaffe.
- **3** Ich habe keine Energie mehr, um überhaupt noch etwas zu tun.

16.) Veränderungen der Schlafgewohnheiten
- **0** Meine Schlafgewohnheiten haben sich nicht verändert.
- **1a** Ich schlafe etwas mehr als sonst
- **1b** Ich schlafe etwas weniger als sonst.
- **2a** Ich schlafe viel mehr als sonst.
- **2b** Ich schlafe viel weniger als sonst.
- **3a** Ich schlafe fast den ganzen Tag.
- **3b** Ich wache 1-2 Stunden früher auf als gewöhnlich und kann dann nicht mehr einschlafen.

17.) Reizbarkeit
- **0** Ich bin nicht reizbarer als sonst.
- **1** Ich bin reizbarer als sonst.
- **2** Ich bin viel reizbarer als sonst.
- **3** Ich fühle mich dauernd gereizt.

18.) Veränderungen des Appetits
- **0** Mein Appetit hat sich nicht verändert.
- **1a** Mein Appetit ist etwas schlechter als sonst.
- **1b** Mein Appetit ist etwas größer als sonst.
- **2a** Mein Appetit ist viel schlechter als sonst.
- **2b** Mein Appetit ist viel größer als sonst.
- **3a** Ich habe überhaupt keinen Appetit.
- **3b** Ich habe ständig Heißhunger.

19.) Konzentrationsschwierigkeiten
- **0** Ich kann mich so gut konzentrieren wie immer.
- **1** Ich kann mich nicht mehr so gut konzentrieren wie sonst.
- **2** Es fällt mir schwer, mich längere Zeit auf irgend etwas zu konzentrieren.
- **3** Ich kann mich überhaupt nicht mehr konzentrieren.

20.) Ermüdung oder Erschöpfung
- **0** Ich fühle mich nicht müder oder erschöpfter als sonst.
- **1** Ich werde schneller müde oder erschöpft als sonst.
- **2** Für viele Dinge, die ich üblicherweise tue, bin ich zu müde oder erschöpft.
- **3** Ich bin so müde oder erschöpft, dass ich fast nichts mehr tun kann.

21.) Verlust an sexuellem Interesse
- **0** Mein Interesse an Sexualität hat sich in letzter Zeit nicht verändert.
- **1** Ich interessiere mich weniger für Sexualität als früher.
- **2** Ich interessiere mich jetzt viel weniger für Sexualität.
- **3** Ich habe das Interesse an Sexualität völlig verloren.

Summe Seite 2:

Übertrag Seite 1: | Gesamt Seite 1+2:

0 - 8 keine Depression

9 - 13 minimale Depression

14 - 19 leichte Depression

20 - 28 mittelschwere Depression

29 - 63 schwere Depression

Beispiel: Evaluation von Psychotherapieformen bei Depression

Experimentelle Bedingung (Gruppen von n = 50)		Psychotherapie	
Klassisch	Pre-BDI		Post-BDI
Online	Pre-BDI		Post-BDI

Beispieldatensatz

Einlesen des Datensatzes mit `read.table()`

```
fname          = file.path(data_dir, "psychotherapie_datensatz.csv")
D              = read.table(fname, sep = ",", header = TRUE)
```

Daten der ersten acht Proband:innen jeder Gruppe

	Bedingung	Pre.BDI	Post.BDI
1	Klassisch	17	9
2	Klassisch	20	14
3	Klassisch	16	13
4	Klassisch	18	12
5	Klassisch	21	12
6	Klassisch	17	14
7	Klassisch	17	12
8	Klassisch	17	9
51	Online	22	16
52	Online	19	15
53	Online	21	13
54	Online	18	15
55	Online	19	13
56	Online	17	16
57	Online	20	13
58	Online	19	16

Beispieldatensatz

Datensatzübersicht mit `View()`

D

Filter

	Bedingung	Pre BDI	Post BDI
1	Klassisch	17	9
2	Klassisch	20	14
3	Klassisch	16	13
4	Klassisch	18	12
5	Klassisch	21	12
6	Klassisch	17	14
7	Klassisch	17	12
8	Klassisch	17	9
9	Klassisch	18	11
10	Klassisch	18	14
11	Klassisch	20	10
12	Klassisch	17	15
13	Klassisch	16	17
14	Klassisch	18	12
15	Klassisch	16	10
16	Klassisch	18	13

Beispieldatensatz

Datenvorverarbeitung

Deskriptive Statistiken

Visualisierung

Parameterschätzung

Konfidenzintervalle

Hypothesentests

Datenvorverarbeitung

- Studienfokus ist die **Veränderung** der Depressionsymptomatik durch Therapieformen.
- Für jede Proband:in der ergibt sich diese Veränderung als **Differenz** von Post.BDI - Pre.BDI.
- Eine Reduktion der Depressionssymptomatik ergibt dabei eine **negative Zahl**.
- Es ist sinnvoller, Verbesserungen mit **positiven Zahlen** zu repräsentieren.
- Als Maß des Therapieeffekts bei Proband:in i bietet sich also an

$$\Delta\text{BDI[i]} := -(\text{Post.BDI[i]} - \text{Pre.BDI[i]}) \tag{1}$$

- Wir betrachten in der Folge also das ΔBDI Maß mit folgender Interpretation

$\Delta\text{BDI} > 0$	Verminderung der Depressionsymptomatik	Wirksame Therapie
$\Delta\text{BDI} = 0$	Keine Veränderung der Depressionsymptomatik	Wirkungslose Therapie
$\Delta\text{BDI} < 0$	Verstärkung der Depressionsymptomatik	Schädigende Therapie

Beispieldatensatz

Datenvorverarbeitung

Hinzufügen einer ΔBDI Spalte zum Dataframe

```
fname        = file.path(data_dir, "psychotherapie_datensatz.csv") # Einlesen
D            = read.table(fname, sep = ",", header = TRUE)                  # Rohdaten
D$Delta.BDI = -(D$Post.BDI - D$Pre.BDI)                                     # \Delta BDI Maß
```

Daten der ersten acht Proband:innen jeder Gruppe

	Bedingung	Pre.BDI	Post.BDI	Delta.BDI
1	Klassisch	17	9	8
2	Klassisch	20	14	6
3	Klassisch	16	13	3
4	Klassisch	18	12	6
5	Klassisch	21	12	9
6	Klassisch	17	14	3
7	Klassisch	17	12	5
8	Klassisch	17	9	8
51	Online	22	16	6
52	Online	19	15	4
53	Online	21	13	8
54	Online	18	15	3
55	Online	19	13	6
56	Online	17	16	1
57	Online	20	13	7
58	Online	19	16	3

Beispieldatensatz

Datenvorverarbeitung

Deskriptive Statistiken

Visualisierung

Parameterschätzung

Konfidenzintervalle

Hypothesentests

Bedingungsabhängige Auswertung deskriptiver Statistiken

```
# Initialisierung eines Dataframes
tp            = c("Klassisch", "Online")            # 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$Delta.BDI[D$Bedingung == 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
}
```

Deskriptive Statistiken

Bedingungsabhängige Auswertung deskriptiver Statistiken

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

```
>            n Max Min Median Mean  Var  Std
> Klassisch 50  12  -1      6 6.16 7.08 2.66
> Online    50   9   1      5 4.92 3.91 1.98
```

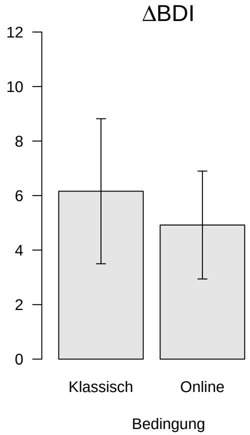
- Die Anzahl der Proband:innen in beiden Therapiegruppen ist gleich.
- Die Spannbreite der ΔBDI Daten ist in der klassischen Therapieform leicht erhöht.
- Median und Mittelwert nehmen für die klassische Therapieform leicht höhere Werte an.
- Ein ΔBDI Mittelwertsunterschied von 1 ist klinisch wohl eher vernachlässigbar.
- Median und Mittelwert sind in beiden Therapieformen ähnlich (unimodale Verteilung).
- Die Variabilitätsmaße zeigen eine etwas erhöhte Varabilität in der klassischen Therapieform.

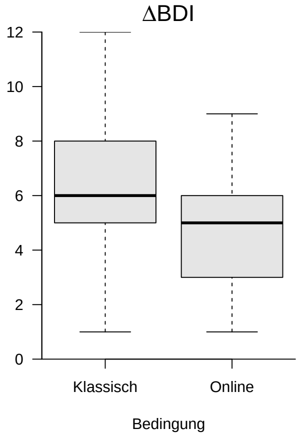
Deskriptive Statistiken

Bedingungsabhängige Visualisierung deskriptiver Statistiken

```r
# 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    = 1.5)                             # Titeltextvergrößerungsfaktor

# Linkes Panel: Balkendiagramm mit Fehlerbalken
mw          = S$Mean                             # Gruppenmittelwert
sd          = S$Std                              # Gruppenstandardabweichung
names(mw)   = tp                                 # barplot braucht x-Werte als names
x = barplot(                                     # Ausgabe der x-Ordinaten (?barplot für Details)
  mw,                                            # Mittelwerte = Balkenhöhe
  col         = "gray90",                        # Balkenfarbe
  ylim        = c(0,12),                         # y-Achsenbegrenzung
  xlim        = c(0,3),                          # x-Achsenbegrenzung
  xlab        = "Bedingung",                     # x-Achsenbeschriftung
  main        = TeX("$\\Delta BDI$"))            # Titel

arrows(                                          # arrows() für Fehlerbalken (siehe ?arrows)
  x0          = x,                               # arrow start x-ordinate
  y0          = mw - sd,                         # arrow start y-ordinate
  x1          = x,                               # arrow end   x-ordinate
  y1          = mw + sd,                         # arrow end   y-ordinate
  code        = 3,                               # Pfeilspitzen beiderseits
  angle       = 90,                              # Pfeilspitzenwinkel -> Linie
  length      = 0.05)                            # Linielänge

# Rechtes Panel: Boxplot
boxplot(
  D$Delta.BDI ~ D$Bedingung,                     # Gruppierung der Delta.BDI Daten nach D$Bedingung
  ylim        = c(0,12),                         # y-Achsenbegrenzung
  col         = "gray90",                        # Boxfarbe
  ylab        = "",                              # y-Achsenbeschriftung
  xlab        = "Bedingung",                     # x-Achsenbeschriftung
  main        = TeX("$\\Delta BDI$"))            # Titel
```

Bedingungsabhängige Visualisierung deskriptiver Statistiken





Beispieldatensatz

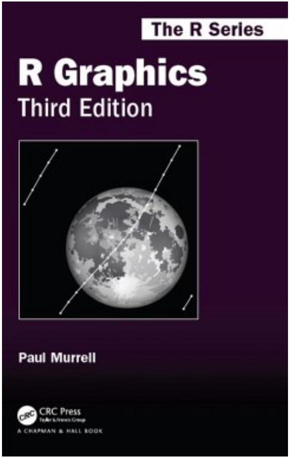
Deskriptive Statistiken

Visualisierung

Parameterschätzung

Konfidenzintervalle

Hypothesentests


The R Series
R Graphics
Third Edition
Paul Murrell
CRC Press
A CHAPMAN & HALL BOOK

R Funktionalitäten für Abbildungen

Base Graphics

- Erstellung und bedarfsgerechte Anpassung von Abbildungen
- Eher low-level, fine tuning orientiert

Lattice und ggplot2

- Erstellung und bedarfsgerechte Anpassung von Abbildungen
- Eher high level, an der eigenen Philosophie orientiert

Base Graphics, lattice und ggplot2 können ähnliche Abbildungen generieren

LaTeX Typesetting ist in allen Paketen unterentwickelt

R Funktionalitäten für Abbildungen

Base Graphics

- **Erstellung und bedarfsgerechte Anpassung von Abbildungen**
- **Eher low-level, fine tuning orientiert**

Lattice und ggplot2

- Erstellung und bedarfsgerechte Anpassung von Abbildungen
- Eher high level, an der eigenen Philosophie orientiert

Base Graphics, lattice und ggplot2 können ähnliche Abbildungen generieren

LaTeX Typesetting ist in allen Paketen unterentwickelt

Visualisierung

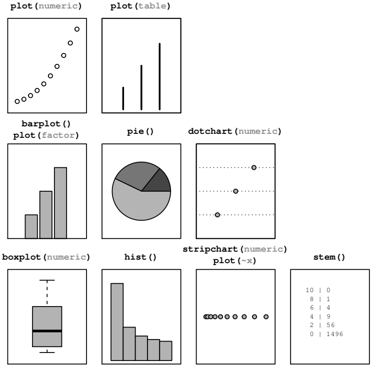


Figure 2.5
High-level base graphics plotting functions for producing plots of a single variable. Where the function can be used to produce more than one type of plot, the relevant data type is shown (in gray). For example, `plot(numeric)` means that this is what the `plot()` produces when it is given a single numeric argument.

Murrell (2019)

Visualisierung

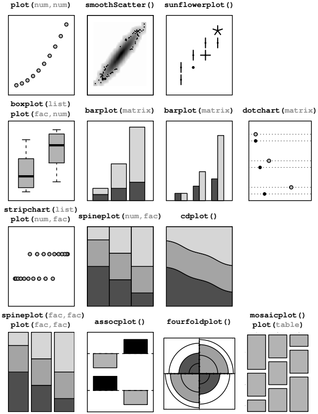


Figure 2.6
High-level base graphics plotting functions for producing plots of two variables. Where the function can be used to produce more than one type of plot, the relevant data type is shown (in gray). For example `plot(num,fac)` represents calling the `plot()` function with a numeric vector as the first argument and a factor as the second argument.

Murrell (2019)

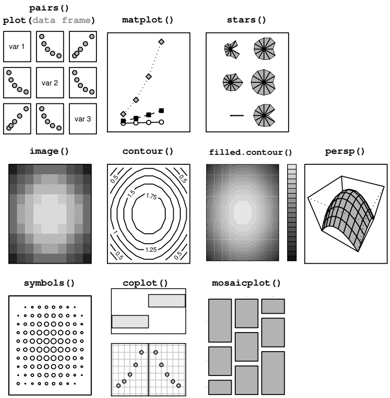


Figure 2.7
High-level base graphics plotting functions for producing plots of many variables. Where the function can be used to produce more than one type of plot, the relevant data type is shown (in gray).

Murrell (2019)

Visualisierung

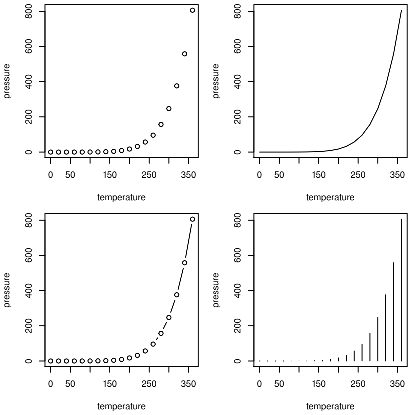


Figure 2.2
Four variations on a scatterplot. In each case, the plot is produced by a call to the `plot()` function with the same data; all that changes is the value of the `type` argument. At top-left, `type="p"` to give **p**oints (data symbols), at top-right, `type="l"` to give **l**ines, at bottom-left, `type="b"` to give **b**oth, and at bottom-right, `type="h"` to give **h**istogram-like vertical lines.

Murrell (2019)

Visualisierung

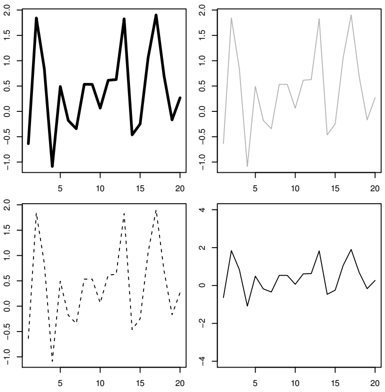


Figure 2.9
Standard arguments for high-level functions. All four plots are produced by calls to the `plot()` function with the same data, but with different standard plot function arguments specified: the top-left plot makes use of the `lwd` argument to control line thickness; the top-right plot uses the `col` argument to control line color; the bottom-left plot makes use of the `lty` argument to control line type; and the bottom-right plot uses the `ylim` argument to control the scale on the y-axis.

Murrell (2019)

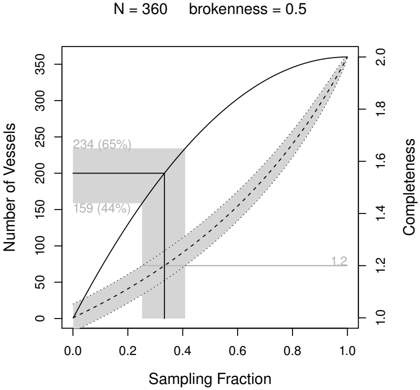


Figure 1.3
A customized scatterplot produced using R. This is created by starting with a simple scatterplot and augmenting it by adding an additional y-axis and several additional sets of lines, polygons, and text labels.

Murrell (2019)

Code Outline

```
# Initialisierung einer neuen Abbildung
dev.new()

# Abbildungsparameter
par(
z.B. Arrangement von Panels, Begrenzungsstile, Schriftfonts, etc
)

# Higher-level Abbildungsfunktion wie plot(), hist(), barplot(), ...
plot(
z.B. x- und y-Daten, Achsenlimits, Achsenbeschriftungen, Titel, Farben, etc.
Jeder Aufruf einer higher-level Graphikfunktion belegt ein neues Subpanel!
)

# Hinzufügen weiterer Daten mit lower-level Abbildungsfunktionen zum aktuellen Panel
z.B. points(), lines(), abline()

# Weitere Graphikannotation zu aktuellem Panel
z.B. legend(), text()

# Speichern der Abbildung (Größenverhältnisse erst hier final festgelegt)
z.B. dev.copy2pdf()
```

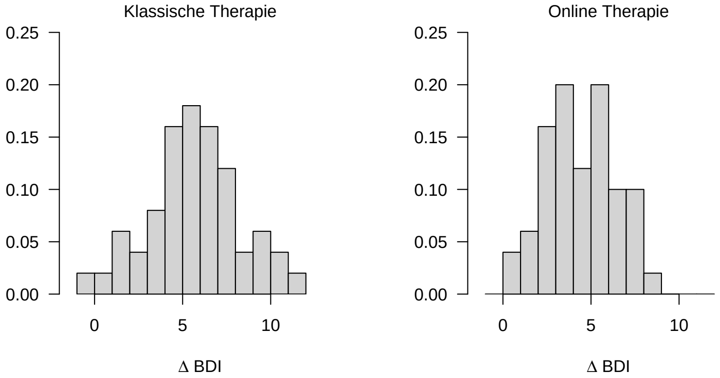
Visualisierung

Histogramme

```
# Histogrammparameter
h            = 1                            # gewünschte Klassenbreite
b_0          = min(D$Delta.BDI)             # b_0
b_k          = max(D$Delta.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("Klassisch" , "Online")    # Therapiebedingungen
labs         = c("Klassische Therapie",     # Abbildungslabel
                "Online Therapie")
# 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$Delta.BDI[D$Bedingung == therapie[i]],  # Delta.BDI Werte von Therapiebedingung i
  breaks     = b,                           # Histogrammklassen
  freq       = F,                           # normierte relative Häufigkeit
  xlim       = xlimits,                     # x-Achsenlimits
  ylim       = ylimits,                     # y-Achsenlimits
  xlab       = TeX("$\\Delta$ BDI"),        # x-Achsenbeschriftung
  ylab       = "",                          # y-Achsenbeschriftung
  main       = labs[i])                     # Titelbeschriftung
}

# PDF Speicherung
dev.copy2pdf(
file         = file.path(abb_dir, "pds_11_histogramm.pdf"),
width        = 8,
height       = 4)
```

Visualisierung

Histogramme



Beispieldatensatz

Deskriptive Statistiken

Visualisierung

Parameterschätzung

Konfidenzintervalle

Hypothesentests

Modellannahmen für Parameterschätzung und Konfidenzintervalle

Motiviert durch die therapieabhängige Visualisierung der ΔBDI Daten und unseren wissenschaftssoziologischen Kontext legen wir nun das Normalverteilungsmodell zugrunde.

Wir nehmen also an, dass die Δ BDI Werte Realisierungen von unabhängig verteilten Zufallsvariablen

$$\upsilon_{ij} \sim N\left(\mu_i, \sigma_i^2\right), i = 1, 2, j = 1, ..., 50 \tag{2}$$

sind, wobei i die Therapiebedingung (1 = Klassisch, 2 = Online) und j den Proband:innen Index in der $i$ten experimentellen Bedingung bezeichnen. Innerhalb einer Bedingung sind diese Zufallsvariablen also unabhängig und identisch verteilt.

Dies entspricht der Annahme, dass sich der ΔBDI Wert einer Proband:in durch Addition einer normalverteilten Fehlervariable mit Erwartungswertparameter 0 und Varianzparameter σ_i^2 zu den innerhalb einer Therapiebedingung identischen Wert μ_i ergibt.

Parameterschätzung

Zur Parameterschätzung im vorliegenden Modell nutzen wir

- den Maximum Likelihood Schätzer für μ_i
- den Varianzschätzer für σ_i^2

```
# Initialisierung eines Dataframes
tp                 = c("Klassisch", "Online")               # Therapiebedingungen
ntp                = length(tp)                             # Anzahl Therapiebedingungen
theta_hats         = data.frame(                            # Dataframeerzeugung
                       mu_ML      = rep(NaN,ntp),           # ML Schätzer für \mu_i
                       sigsqr_VAR = rep(NaN,ntp))           # Varianzschätzer für \sigma^2_i

# Iterationen über Therapiebedingungen
for(i in 1:ntp){
  data                     = D$Delta.BDI[D$Bedingung == tp[i]]    # Daten
  theta_hats$mu_ML[i]      = mean(data)                           # ML Schätzer für \mu_i
  theta_hats$sigsqr_VAR[i] = var(data)                            # Varianzschätzer für \sigma_2^_i
}

# Ausgabe
print(theta_hats)
```

```
>   mu_ML sigsqr_VAR
> 1  6.16       7.08
> 2  4.92       3.91
```

Tipps für μ_i und σ_i^2 auf Grundlage dieser unverzerrten Schätzer sind also

$$\hat{\mu}_1 = 6.16, \quad \hat{\mu}_2 = 4.92, \quad \hat{\sigma}_1^2 = 7.08, \quad \hat{\sigma}_2^2 = 3.91. \tag{3}$$

Die mit diesen Tipps assoziierte Unsicherheit ist hier nicht angegeben.

Konfidenzintervalle

Konfidenzintervalle für die Erwartungswertparameterschätzer

```
# Analyseparameter
t       = c("Klassisch", "Online")                    # Therapiebedingungen
ntp     = length(tp)                                  # Anzahl an Therapiebedingungen
n       = 50                                          # Anzahl von Beobachtungen pro Therapiebedingung
C       = data.frame(                                 # Dataframeerzeugung
            G_u       = rep(NaN,ntp),                 # untere KI Grenze
            mu_hat    = rep(NaN,ntp),                 # Erwartungswertparameterschätzer
            G_o       = rep(NaN,ntp),                 # obere KI Grenze
            row.names = tp)                           # Therapiebedingungen


# Konfidenzintervallparameter
delta   = 0.95                                        # Konfidenzlevel
psi_inv = qt((1+delta)/2,n-1)                         # \psi^-1((\delta + 1)/2, n-1)

# Konfidenzintervallevaluation
for(i in 1:ntp){
   data       = D$Delta.BDI[D$Bedingung == t[i]]   # Stichprobenrealisierung
   X_bar      = mean(data)                          # Stichprobenmittel
   S          = sd(data)                            # Stichprobenstandardabweichung
   C$G_u[i]    = X_bar - (S/sqrt(n))*psi_inv        # untere KI Grenze
   C$mu_hat[i] = X_bar                              # Erwartungswertparameterschätzer
   C$G_o[i]    = X_bar + (S/sqrt(n))*psi_inv        # obere KI Grenze
}

# Ausgabe
print.AsIs(C)
```

```
>             G_u mu_hat  G_o
> Klassisch 5.40   6.16 6.92
> Online    4.36   4.92 5.48
```

Konfidenzintervalle für die Varianzparameterschätzer

```r
# Analyseparameter
t       = c("Klassisch", "Online")                    # Therapiebedingungen
ntp     = length(tp)                                  # Anzahl an Therapiebedingungen
n       = 50                                          # Anzahl von Beobachtungen pro Therapiebedingung
C       = data.frame(                                 # Dataframeerzeugung
          G_u        = rep(NaN,ntp),                  # untere KI Grenze
          sigsqr_hat = rep(NaN,ntp),                  # Varianzparameterschätzer
          G_o        = rep(NaN,ntp),                  # obere KI Grenze
          row.names  = tp)                            # Therapiebedingungen


# Konfidenzintervallparameter
delta   = 0.95                                        # Konfidenzlevel
xi_1    = qchisq((1-delta)/2, n - 1)                  # \Xi^2((1-\delta)/2; n - 1)
xi_2    = qchisq((1+delta)/2, n - 1)                  # \Xi^2((1+\delta)/2; n - 1)

# Konfidenzintervallevaluation
for(i in 1:ntp){
   data             = D$Delta.BDI[D$Bedingung == t[i]] # Stichprobenrealisierung
   S2               = var(data)                       # Stichprobenvarianz
   C$G_u[i]         = (n-1)*S2/xi_2                   # untere KI Grenze
   C$sigsqr_hat[i]  = S2                              # Varianzparameterschätzer
   C$G_o[i]         = (n-1)*S2/xi_1                   # obere KI Grenze
}

# Ausgabe
print.AsIs(C)
```

```
>            G_u sigsqr_hat   G_o
> Klassisch 4.94       7.08 10.99
> Online    2.73       3.91  6.07
```

Hypothesentest

Anwendungsszenario und statistisches Modell

Wir nehmen an, dass die Δ BDI Werte, also der uns vorliegende Datensatz Realisierungen von unabhängig verteilten Zufallsvariablen

$$\upsilon_{ij} \sim N\left(\mu_i, \sigma^2\right), i = 1, 2, j = 1, ..., 50 \tag{4}$$

sind, wobei i die Therapiebedingung (1 = Klassisch, 2 = Online) und j den Proband:innen Index in der $i$ten experimentellen Bedingung bezeichnen. Innerhalb einer Bedingung sind diese Zufallsvariablen also unabhängig und identisch verteilt.

Die Parameter μ_1, μ_2, σ^2 sind unbekannt sind. Wir beabsichtigen das Quantifizieren der Unsicherheit beim inferentiellen Vergleich von μ_1 mit μ_2.

Dafür können wir einen Zweistichproben-T-Test bei unabhängigen Stichproben unter Annahme identischer Varianz durchführen.

Wir wollen die Hypothesen $H_0 : \mu_1 = \mu_2$ und $H_1 : \mu_1 \neq \mu_2$ mit einem Signifikanzniveau von $\alpha_0 = 0.05$ testen und verwenden dafür einen **zweiseitigen Zweistichproben-Test**.

(1) Modell und Testhypothesen

Wir betrachten die einfache Nullhpothese und die zusammengesetzte Alternativhypothese

$$H_0 : \mu_1 - \mu_2 = 0 \Leftrightarrow \mu_1 = \mu_2 \text{ und} \tag{5}$$

$$H_1 : \mu_1 - \mu_2 \neq 0 \Leftrightarrow \mu_1 \neq \mu_2 \tag{6}$$

(2) Definition und Analyse der Teststatistik

Die T-Teststatistik für den Zweistichproben-T-Test ist gegeben durch

$$T = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \left(\frac{\bar{y}_1 - \bar{y}_2 - \mu_0}{s_{12}} \right) \tag{7}$$

und dessen Verteilung durch

$$T \sim t(, n_1 + n_2 - 2) \text{ mit } = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \left(\frac{\mu_1 - \mu_2 - \mu_0}{\sigma} \right). \tag{8}$$

wobei $\bar{y}_1$ und $\bar{y}_2$ die Stichprobenmittel der Gruppen Klassische und Online-Therapie, respektive sind, n_1 und n_2 die jeweiligen Stichprobengrößen und $s_{12} = \sqrt{s_{12}^2}$ die *gepoolte Stichprobenstandardabweichung*. Die *gepoolte Stichprobenvarianz* ist gegeben durch

$$s_{12}^2 := \frac{\sum_{j=1}^{n_1} (y_{1j} - \bar{y}_1)^2 + \sum_{j=1}^{n_2} (y_{2j} - \bar{y}_2)^2}{n_1 + n_2 - 2} \tag{9}$$

(3) Definition des Tests

$$\phi(v) := 1_{\{|T| \geq k\}} = \begin{cases} 1 & |T| \geq k \\ 0 & |T| < k \end{cases}. \tag{10}$$

(4) Analyse der Testgütefunktion

Theorem (Testgütefunktion)

Es sei ϕ der im obigen Modell formulierte Zweistichproben-T-Test. Dann ist die Testgütefunktion von ϕ gegeben durch

$$q_\phi : \mathbb{R}^2 \to [0,1], (\mu_1, \mu_2) \mapsto q_\phi(\mu_1, \mu_2)$$
$$:= 1 - \psi(k; \delta, n_1 + n_2 - 2) + \psi(-k; \delta, n_1 + n_2 - 2) \quad (11)$$

wobei $\psi(\cdot; \delta, n_1 + n_2 - 2)$ die KVF der nichtzentralen t-Verteilung mit Nichtzentralitätsparameter

$$\delta := \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \frac{\mu_1 - \mu_2}{\sigma} \quad (12)$$

und Freiheitsgradparameter $n_1 + n_2 - 2$ bezeichnet.

(5) Testumfangkontrolle

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

$$k_{\alpha_0} := \psi^{-1}\left(1 - \frac{\alpha_0}{2}; n_1 + n_2 - 2\right), \tag{13}$$

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

Manueller Zweistichproben-T-Test

```
# Datenauswahl
x_1       = D$Delta.BDI[D$Bedingung == "Klassisch"]           # \Delta.BDI Daten Klassische Therapie
x_2       = D$Delta.BDI[D$Bedingung == "Online"]              # \Delta.BDI Daten Klassische Therapie
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
```

Manueller Zweistichproben-T-Test

```r
# Ausgabe
cat("\nx_bar_1            = ", x_bar_1,
    "\nx_bar_2            = ", x_bar_2,
    "\nFreiheitsgrade     = ", n_1 + n_2 - 2,
    "\nSignifikanzlevel   = ", alpha_0,
    "\nKritischer Wert    = ", k_alpha_0,
    "\nTeststatistik      = ", t,
    "\nTestwert           = ", phi,
    "\np-Wert             = ", pval)
```

```
>
> x_bar_1          =  6.16
> x_bar_2          =  4.92
> Freiheitsgrade   =  98
> Signifikanzlevel =  0.05
> Kritischer Wert  =  1.98
> Teststatistik    =  2.65
> Testwert         =  1
> p-Wert           =  0.00951
```

Folgendes können wir aus dieser Zusammenfassung ablesen:

- Die Stichprobenmittel der zwei Datensätze sind 6.16 und 4.92
- Die Anzahl der Freiheitsgrade ist 98
- Das Signifikanzniveau dieses Tests α_0 ist 0.05
- Der kritische Wert k_{α_0} ist 1.98
- Der Wert der T-Teststatistik T ist 2.65
- Das Ergebnis des Tests ϕ ist 1 (Wie zu erwarten, da hier $T > k_{\alpha_0}$)
- $\Rightarrow$ Wir lehnen die Nullhypothese $H_0 : \mu_1 = \mu_2$ ab.
- Der p-Wert beträgt 0.00951

R Implementation des Zweistichproben-T-Tests

```
# 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 = 3, df = 98, p-value = 0.01
> alternative hypothesis: true difference in means is not equal to 0
> 95 percent confidence interval:
>  0.31 2.17
> sample estimates:
> mean of x mean of y
>      6.16      4.92
```

Folgendes können wir aus dieser Zusammenfassung ablesen:

- Es wurden die Daten verwendet, die in den Variablen `x_1` und `x_2` gespeichert sind.
- Der Wert der T-Teststatistik T ist 2.65
- Die Anzahl der Freiheitsgrade, engl.: degrees of freedom (df) ist 98
- Die Alternativhypothese ist, dass die "wahre" Differenz zwischen μ_1 und μ_2 gleich 0 ist.
- Das 95%-Konfidenzintervall der geschätzten Erwartungswertdifferenz ist $[0.31, 2.17]$.
- Die Stichprobenmittel der zwei Datensätze sind 6.16 und 4.92
- Der p-Wert beträgt (gerundet) 0.01.
- $\Rightarrow$ Da dieser Wert kleiner als das festgelegte Signifikanzniveau 0.05 ist, kann die die Nullhypothese $H_0 : \mu_1 = \mu_2$ abgelehnt werden.

R Implementation des Zweistichproben-T-Tests

Die Werte, die in der automatischen Ausgabe angezeigt werden sind gerundet. Die exakten Werte sind jedoch im Objekt `varphi` gespeichert und wir können diese aufrufen.

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

```
> [1] "c(t = 2.64516155336263)"
```

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

```
> [1] "0.00951137026459394"
```

References

Murrell, Paul. 2019. *R Graphics*. Third edition. The R Series. Boca Raton: CRC Press, Taylor & Francis Group.