Allgemeines Lineares Modell

BSc Psychologie, SoSe 2024

Joram Soch

(14) Kovarianzanalyse

Vorbemerkungen

Additive Kovarianzanalyse

Kovarianzanalyse mit Interaktion

Selbstkontrollfragen

Vorbemerkungen

Additive Kovarianzanalyse

Kovarianzanalyse mit Interaktion

Selbstkontrollfragen

Vorbemerkungen

Variabilität der Werte einer abhängigen Variable

- Die Werte einer AV unterscheiden sich im Allgemeinen zwischen experimentellen Einheiten.
- Die Schwankungen der Werte einer AV bezeichnet man als *Datenvariabilität*.
- Ziel jeder Datenanalyse ist die Erklärung von Datenvariabilität durch Zerlegung.

Datenvariabilität und Allgemeines Lineares Modell

- Eine spezielle Art der Quantifizierung von Datenvariabilität ist die Stichprobenvarianz.
- Einen additiven Zugang zur Dekomposition von Datenvariabilität bietet das ALM.
- Den folgenden Überlegungen liegt datenanalytisch die Kovarianzanalyse zugrunde.

Gesamtvariabilität = Primärvariabilität + Residualvariabilität

Primärvariabilität

Systematische Veränderung der AV, die allein auf Variation der UV zurückzuführen ist.

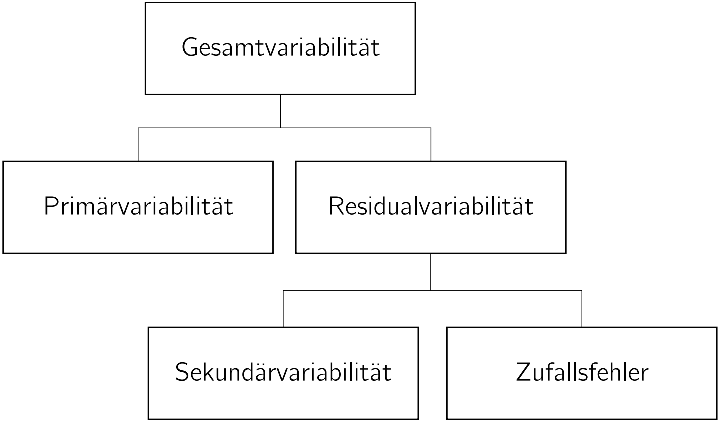
Residualvariabilität = Sekundärvariabilität + Zufallsfehler

Sekundärvariabilität

Systematische Veränderung der AV, die auf die Wirkung von unkontrollierten Störvariablen, nicht aber auf die Variation der UV zurückzuführen ist.

Zufallsfehler

Unsystematische Veränderung der AV, die weder auf die Variation der UV, noch auf den Einfluss von Störvariablen zurückzuführen ist.


Gesamtvariabilität
Primärvariabilität
Residualvariabilität
Sekundärvariabilität
Zufallsfehler

Vorbemerkungen

Parametrische und faktorielle ALM-Designs

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.
- $\Rightarrow$ einfache lineare Regression, multiple lineare Regression

Faktorielle ALM-Designs

- Designmatrizen besitzen Spalten mit 1en und 0en.
- Betaparameter repräsentieren Gruppenerwartungswerte.
- Betaparameterschätzer ergeben sich aus Gruppenstichprobenmitteln.
- $\Rightarrow$ T-Tests, einfaktorielle Varianzanalyse, mehrfaktorielle Varianzanalyse

Parametrisch-faktorielle ALM-Designs

- Designmatrizen mit mehreren parametrischen und faktoriellen Regressoren.
- $\Rightarrow$ Kovarianzanalyse $\Leftrightarrow$ Allgemeines Lineares Modell

Wir betrachten hier exemplarisch den Fall 1 Faktor + 1 parametrischer Regressor.

Vorbemerkungen

Additive Kovarianzanalyse

Kovarianzanalyse mit Interaktion

Selbstkontrollfragen

Additive Kovarianzanalyse

Modell der additiven einfaktoriellen Kovarianzanalyse

Im Modell der additiven einfaktoriellen Kovarianzanalyse möchte man die Erwartungswertparameter μ_{ij} der insgesamt $n = \sum_{i=1}^{I} n_i$ Datenvariablen

$$y_{ij} \sim N(\mu_{ij}, \sigma^2 I_n) \tag{1}$$

basierend auf den Leveln $i = 1, ..., I$ des experimentellen Faktors und den entsprechenden Werten x_{ij} für $j = 1, ..., n_i$ der Kovariate modellieren.

Dazu wählt man für jeden Erwartungswertparameter μ_{ij} zunächst einen Parameter μ_0, der einen Datenvariablen-unspezifischen Offset modelliert. Weiterhin wählt man einen Parameter α_i, der den Beitrag des $i$ten Levels des experimentellen Faktors modelliert und schließlich einen Gewichtungsparameter β_0, der den Beitrag der Kovariatenwerten x_{ij} zu μ_{ij} quantifiziert.

Das Modell für den Erwartungswertparameter von y_{ij} für $i = 1, ..., I, j = 1, ..., n_i$ nimmt im Fall der additiven einfaktoriellen Kovarianzanalyse also folgende Form an:

$$\mu_{ij} := \mu_0 + \alpha_i + \beta_0 x_{ij} \text{ mit } \alpha_1 := 0, \tag{2}$$

wobei $\alpha_1 := 0$ wie im Falle der einfaktoriellen Varianzanalyse eine Überparameterisierung des Modells verhindert.

Man beachte, dass die Form des ijten Erwartungswertparameters damit eine einfache lineare Regression für jedes Level $i = 1, .., I$ des experimentellen Faktors definiert, wobei $\mu_0 + \alpha_i$ als Faktorlevel-spezifische Offsetparameter interpretiert werden können und der Steigungsparameter β_0 für alle Faktorlevel identisch ist.

Definition (Modell der additiven einfaktoriellen Kovarianzanalyse)

y_{ij} mit $i = 1, ..., I$, $j = 1, ..., n_i$ sei die Zufallsvariable, die den zum $i$ten Level des Faktors gehörenden $j$ten Datenpunkt modelliert und x_{ij} sei der entsprechende Wert der Kovariate. Dann hat das *Modell der additiven einfaktoriellen Kovarianzanalyse* die strukturelle Form

$$y_{ij} = \mu_{ij} + \varepsilon_{ij} \text{ mit } \varepsilon_{ij} \sim N(0, \sigma^2) \text{ u.i.v. für } i = 1, ..., I, j = 1, ..., n_i, \sigma^2 > 0 \tag{3}$$

und die Datenverteilungsform

$$y_{ij} \sim N(\mu_{ij}, \sigma^2) \text{ u.v. für } i = 1, ..., I, j = 1, ..., n_i \text{ ,} \tag{4}$$

wobei

$$\mu_{ij} := \mu_0 + \alpha_i + \beta_0 x_{ij} \text{ mit } \alpha_1 := 0 \text{ .} \tag{5}$$

Bemerkungen

- Die Form des ijten Erwartungswertparameters

$$\mu_{ij} := \mu_0 + \alpha_i + \beta_0 x_{ij} \text{ mit } \alpha_1 := 0 \tag{6}$$

  definiert eine einfache lineare Regression für jedes $i = 1, .., I$, wobei

  - $\mu_0 + \alpha_i$ ein Faktorlevel-spezifischer Offsetparameter ist und
  - der Steigungsparameter β_0 für alle Faktorlevel identisch ist.

Theorem (Designmatrixform der additiven einfaktoriellen Kovarianzanalyse)

Gegeben sei die strukturelle Form eines additiven einfaktoriellen Kovarianzanalysemodells mit Anzahl der Faktorlevel I und es sei $n := \sum_{i=1}^{I} n_i$ die Gesamtanzahl an Datenvariablen. Dann hat dieses Modell die Designmatrixform

$$y \sim N(X\beta, \sigma^2 I_n) \tag{7}$$

mit

$$y := \begin{pmatrix} y_{11} \\ \vdots \\ y_{1n_1} \\ y_{21} \\ \vdots \\ y_{2n_2} \\ \\ \vdots \\ \\ y_{I1} \\ \vdots \\ y_{In_I} \end{pmatrix}, \quad X := \begin{pmatrix} 1 & 0 & & 0 & x_{11} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 1 & 0 & & 0 & x_{1n_1} \\ 1 & 1 & & 0 & x_{21} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 1 & 1 & & 0 & x_{2n_2} \\ \\ \vdots & \vdots & & \cdots & \vdots \\ \\ 1 & 0 & & 1 & x_{I1} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 1 & 0 & & 1 & x_{In_I} \end{pmatrix} \in \mathbb{R}^{n \times (I+1)}, \quad \beta := \begin{pmatrix} \mu_0 \\ \alpha_2 \\ \vdots \\ \alpha_I \\ \beta_0 \end{pmatrix} \in \mathbb{R}^{I+1} \text{ und } \sigma^2 > 0. \tag{8}$$

Bemerkungen

- Das Theorem ergibt sich direkt mit den Regeln der Matrixmultiplikation.

Additive Kovarianzanalyse

Anwendungsbeispiel

experimenteller Faktor/Studiengruppe ($I := 2,\ i = 1, 2$)

- Face-to-face- vs. Online-Psychotherapie

parametrischer Regressor/Kovariate ($x_{ij},\ j = 1, ..., n_i$)

- Dauer der Depressionssymptomatik zu Beginn der Intervention

abhängige Variable/primäres Ergebnismaß ($y_{ij},\ j = 1, ..., n_i$)

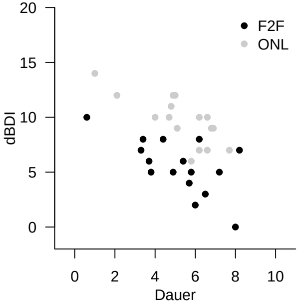
- BDI Pre-Post-Differenzwerte

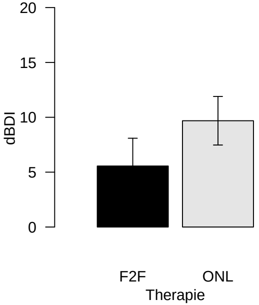
Additive Kovarianzanalyse

Beispieldatensatz ($n = 32$, $n_i = 16$ für $i = 1, 2$; hier: $i = 1, 2$, $j = 1, ..., 10$)

	Therapie	Dauer	dBDI
1	F2F	3.7	6
2	F2F	5.4	6
3	F2F	3.3	7
4	F2F	8.2	7
5	F2F	5.7	4
6	F2F	3.4	8
7	F2F	6.0	2
8	F2F	6.5	3
9	F2F	6.2	8
10	F2F	4.4	8
17	ONL	5.0	12
18	ONL	6.9	9
19	ONL	6.6	7
20	ONL	6.2	7
21	ONL	6.8	9
22	ONL	6.6	10
23	ONL	5.1	9
24	ONL	1.0	14
25	ONL	6.2	10
26	ONL	4.9	12

Visualisierung





Modellformulierung

Modell 1 | keine Berücksichtigung der Kovariate

$$y_{ij} \sim N(\mu_{ij}, \sigma^2) \text{ mit } \mu_{ij} := \mu_0 + \alpha_i \text{ und } \alpha_1 := 0 \text{ und } \sigma^2 > 0 \tag{9}$$

$$\Leftrightarrow \quad y \sim N(X\beta, \sigma^2 I_n) \text{ mit } X := \begin{pmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 1 & 1 \\ \vdots & \vdots \\ 1 & 1 \end{pmatrix}, \ \beta := \begin{pmatrix} \mu_0 \\ \alpha_2 \end{pmatrix} \text{ und } \sigma^2 > 0 \tag{10}$$

Modell 2 | Additive Berücksichtigung der Kovariate

$$y_{ij} \sim N(\mu_{ij}, \sigma^2) \text{ mit } \mu_{ij} := \mu_0 + \alpha_i + \beta_0 x_{ij} \text{ und } \alpha_1 := 0 \text{ und } \sigma^2 > 0 \tag{11}$$

$$\Leftrightarrow \quad y \sim N(X\beta, \sigma^2 I_n) \text{ mit } X := \begin{pmatrix} 1 & 0 & x_{11} \\ \vdots & \vdots & \vdots \\ 1 & 0 & x_{1n_1} \\ 1 & 1 & x_{21} \\ \vdots & \vdots & \vdots \\ 1 & 1 & x_{2n_2} \end{pmatrix}, \ \beta := \begin{pmatrix} \mu_0 \\ \alpha_2 \\ \beta_0 \end{pmatrix} \text{ und } \sigma^2 > 0 \tag{12}$$

Additive Kovarianzanalyse

Implementation

```
fname      = file.path(getwd(), "Daten", "Kovarianzanalyse-1.csv")       # Datensatzdateiname
D          = read.table(fname, sep = ",", header = TRUE)                 # Datensatz
n_i        = c(sum(D$Therapie == "F2F"), sum(D$Therapie == "ONL"))       # Anzahl Datenpunkte pro Gruppe
y          = D$dBDI                                                      # Daten
n          = length(y)                                                   # Gesamtanzahldatenpunkte
XS         = list()                                                      # Modell 1 und 2 Liste
CS         = list()                                                      # Kontrastgewichtsvektorenliste
XS[[1]]    = matrix(c(rep(1,n_i[1]),rep(1,n_i[2]),                       # \mu_0    Regressor
                      rep(0,n_i[1]),rep(1,n_i[2])), nrow = n)            # \alpha_2 Regressor
XS[[2]]    = matrix(c(rep(1,n_i[1]),rep(1,n_i[2]),                       # \mu_0    Regressor
                      rep(0,n_i[1]),rep(1,n_i[2]),                       # \alpha_2 Regressor
                      D$Dauer), nrow = n)                                # x        Regressor
CS[[1]]    = matrix(c(0,1)  , nrow = 2)                                  # Kontrastgewichtsvektor \alpha_2
CS[[2]]    = matrix(c(0,1,0), nrow = 3)                                  # Kontrastgewichtsvektor \alpha_2
B          = list()                                                      # Betaparameterschätzerliste
S          = rep(NaN,2)                                                  # Varianzparameterschätzervektor
T          = rep(NaN,2)                                                  # T-Statistik-Vektor
for(i in 1:2){                                                           # Iteration über Modelle
    X           = XS[[i]]                                                # Designmatrix
    p           = ncol(X)                                                # Anzahl Betaparameter
    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
    c           = CS[[i]]                                                # Kontrastgewichtsvektor
    t_num       = t(c) %*% beta_hat                                      # Zähler der Zweistichproben-T-Teststatistik
    t_den       = sqrt(sigsqr_hat %*% t(c) %*% solve(t(X) %*% X) %*% c) # Nenner der Zweistichproben-T-Teststatistik
    t           = t_num/t_den                                            # Wert   der Zweistichproben-T-Teststatistik
    B[[i]]      = beta_hat                                               # Betaparameterschätzer
    S[i]        = sigsqr_hat                                             # Varianzparameterschätzer
    T[i]        = t                                                      # T-Statistik
}
```

Additive Kovarianzanalyse

Ergebnisse

```
> Modell 1 | Betaparameterschätzer    :  5.56 4.12
> Modell 2 | Betaparameterschätzer    :  10.1 4.2 -0.86
> Modell 1 | Varianzparameterschätzer :  5.65
> Modell 2 | Varianzparameterschätzer :  3.14
> Modell 1 | T-Statistik ONL-Therapie :  4.91
> Modell 2 | T-Statistik ONL-Therapie :  6.7
```

In Modell 2 ist der Varianzparameterschätzer kleiner als in Modell 1.

$\Rightarrow$ Der parametrische Regressor erklärt Datenvarianz.

Die Betaparameterschätzer für den ONL-Therapieeffekt sind in Modell 1 und Modell 2 ähnlich.

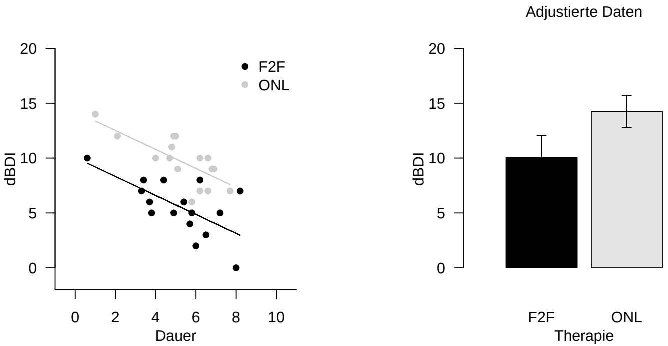
$\Rightarrow$ Die T-Statistik (Signal-zu-Rauschen-Verhältnis) für den ONL-Therapieeffekt ist Modell 2 höher.

Prinzipiell kann man die beobachteten Daten für den geschätzten Dauer Effekt durch

$$\hat{y}_{ij}^{\text{adj}} = y_{ij} - \hat{\beta}_0 x_{ij} \tag{13}$$

"korrigieren" und erhält so einen "adjustierten Datensatz".

Visualisierung



Vorbemerkungen

Additive Kovarianzanalyse

Kovarianzanalyse mit Interaktion

Selbstkontrollfragen

Kovarianzanalyse mit Interaktion

Modell der einfaktoriellen Kovarianzanalyse mit Interaktion

Wie im Modell der additiven einfaktoriellen Kovarianzanalyse möchte man im Modell der einfaktoriellen Kovarianzanalyse mit Interaktion die Erwartungswertparameter μ_{ij} der insgesamt $n = \sum_{i=1}^{I} n_i$ Datenvariablen

$$y_{ij} \sim N(\mu_{ij}, \sigma^2 I_n) \tag{14}$$

basierend auf den Leveln $i = 1, ..., I$ des experimentellen Faktors und den entsprechenden Werten x_{ij} für $j = 1, ..., n_i$ der Kovariate modellieren. Über das Szenario der additiven einfaktoriellen Kovarianzanalyse hinaus möchte man dabei explizit einen Faktorlevel-spezifischen Einfluss der Kovariate auf die Erwartungswertparameter zulassen.

Dazu wählt man für jeden Erwartungswertparameter μ_{ij} zunächst wieder einen Parameter μ_0, der einen Datenvariablen-unspezifischen Offset modelliert. Weiterhin wählt man einen Parameter α_i, der den Beitrag des $i$ten Levels des experimentellen Faktors modelliert. Den Wichtungsparameter des Beitrags der Kovariaten x_{ij} zu μ_{ij} modelliert man nun hier mithilfe eines Faktorlevel-unspezifischen Parameter β_0 <u>und</u> eines Faktorlevel-spezifischen Parameters γ_i für $i = 1, ..., I$. Das Modell für den Erwartungswertparameter von y_{ij} für $i = 1, ..., I, j = 1, ..., n_i$ nimmt im Fall der einfaktoriellen Kovarianzanalyse mit Interaktion damit folgende Form an:

$$\mu_{ij} := \mu_0 + \alpha_i + (\beta_0 + \gamma_i)x_{ij} \text{ mit } \alpha_1 := \gamma_1 := 0, \tag{15}$$

wobei $\alpha_1 := \gamma_1 := 0$ wie im Fall der zweiaktoriellen Varianzanalyse mit Interaktion eine Überparmeterisierung des Modells verhindert.

Man beachte, dass die Form des ijten Erwartungswertparameters damit eine einfache lineare Regression für jedes Level $i = 1, .., I$ des experimentellen Faktors definiert, bei $\mu_0 + \alpha_i$ als Faktorlevel-spezifische Offsetparameter interpretiert werden können und sich die Steigungsparameter $\beta_0 + \gamma_i$ über die Faktorlevel unterscheiden können. Es werden also Unterschiede (zwischen den Faktorleveln) von Unterschieden (zwischen den Erwartungswertparametern eines Faktorlevels in Abhängigkeit vom Wert der Kovariate), d.h. Interaktionen, mit modelliert.

Definition (Modell der einfaktoriellen Kovarianzanalyse mit Interaktion)

y_{ij} mit $i = 1, ..., I$, $j = 1, ..., n_i$ sei die Zufallsvariable, die den zum $i$ten Level des Faktors gehörenden $j$ten Datenpunkt modelliert und x_{ij} sei der entsprechende Wert der Kovariate. Dann hat das *Modell einfaktoriellen Kovarianzanalyse mit Interaktion* die strukturelle Form

$$y_{ij} = \mu_{ij} + \varepsilon_{ij} \text{ mit } \varepsilon_{ij} \sim N(0, \sigma^2) \text{ u.i.v. für } i = 1, ..., I, j = 1, ..., n_i, \sigma^2 > 0 \tag{16}$$

und die Datenverteilungsform

$$y_{ij} \sim N(\mu_{ij}, \sigma^2) \text{ u.v. für } i = 1, ..., I, j = 1, ..., n_i, \tag{17}$$

wobei

$$\mu_{ij} := \mu_0 + \alpha_i + (\beta_0 + \gamma_i)x_{ij} \text{ mit } \alpha_1 := \gamma_1 := 0. \tag{18}$$

Bemerkungen

- Die Form des ijten Erwartungswertparameters

$$\mu_{ij} := \mu_0 + \alpha_i + (\beta_0 + \gamma_i)x_{ij} \text{ mit } \alpha_1 := \gamma_1 := 0 \tag{19}$$

  definiert eine einfache lineare Regression für jedes $i = 1, .., I$, wobei

  - μ_0 und α_i Faktorlevel-unspezifische und -spezifische Offsetparameter sind und
  - β_0 und γ_i Faktorlevel-unspezifische und -spezifische Steigungsparameter sind.

Theorem (Designmatrixform der einfaktoriellen Kovarianzanalyse mit Interaktion)

Gegeben sei die strukturelle Form eines einfaktoriellen Kovarianzanalysemodells mit Interaktion mit Anzahl der Faktorlevel I und es sei $n := \sum_{i=1}^N n_i$ die Gesamtanzahl an Datenvariablen. Dann hat dieses Modell die Designmatrixform

$$y \sim N(X\beta, \sigma^2 I_n) \tag{20}$$

mit

$$y := \begin{pmatrix} y_{11} \\ \vdots \\ y_{1n_1} \\ y_{21} \\ \vdots \\ y_{2n_2} \\ \\ \vdots \\ \\ y_{I1} \\ \vdots \\ y_{In_I} \end{pmatrix}, \quad X := \begin{pmatrix} 1 & 0 & & 0 & x_{11} & 0 & & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & 0 & & 0 & x_{1n_1} & 0 & & 0 \\ 1 & 1 & & 0 & x_{21} & x_{21} & & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & 1 & & 0 & x_{2n_2} & x_{2n_2} & & 0 \\ \\ \vdots & \vdots & & \cdots & \vdots & \vdots & & \cdots \\ \\ 1 & 0 & & 1 & x_{I1} & 0 & & x_{I1} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & 0 & & 1 & x_{In_I} & 0 & & x_{In_I} \end{pmatrix} \in \mathbb{R}^{n \times (2I)}, \quad \beta := \begin{pmatrix} \mu_0 \\ \alpha_2 \\ \vdots \\ \alpha_I \\ \beta_0 \\ \gamma_2 \\ \vdots \\ \gamma_I \end{pmatrix} \in \mathbb{R}^{2I} \text{ und } \sigma^2 > 0. \tag{21}$$

Bemerkungen

- Das Theorem ergibt sich direkt mit den Regeln der Matrixmultiplikation.

Anwendungsbeispiel

experimenteller Faktor/Studiengruppe ($I := 2,\ i = 1,2$)

- Face-to-face- vs. Online-Psychotherapie

parametrischer Regressor/Kovariate ($x_{ij},\ j = 1,...,n_i$)

- Digitalaffinität der Studienteilnehmer

abhängige Variable/primäres Ergebnismaß ($y_{ij},\ j = 1,...,n_i$)

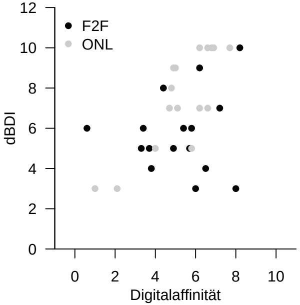
- BDI Pre-Post-Differenzwerte

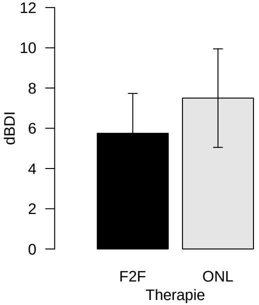
Kovarianzanalyse mit Interaktion

Beispieldatensatz ($n = 32$, $n_i = 16$ für $i = 1, 2$; hier: $i = 1, 2$, $j = 1, ..., 10$)

	Therapie	Dig.Affin	dBDI
1	F2F	3.7	5
2	F2F	5.4	6
3	F2F	3.3	5
4	F2F	8.2	10
5	F2F	5.7	5
6	F2F	3.4	6
7	F2F	6.0	3
8	F2F	6.5	4
9	F2F	6.2	9
10	F2F	4.4	8
17	ONL	5.0	9
18	ONL	6.9	10
19	ONL	6.6	7
20	ONL	6.2	7
21	ONL	6.8	10
22	ONL	6.6	10
23	ONL	5.1	7
24	ONL	1.0	3
25	ONL	6.2	10
26	ONL	4.9	9

Visualisierung





Kovarianzanalyse mit Interaktion

Modellformulierung

Modell 1 | additive Berücksichtigung der Kovariate

$$y_{ij} \sim N(\mu_{ij}, \sigma^2) \text{ mit } \mu_{ij} := \mu_0 + \alpha_i + \beta_0 x_{ij} \text{ und } \alpha_1 := 0 \text{ und } \sigma^2 > 0 \tag{22}$$

$$\Leftrightarrow \quad y \sim N(X\beta, \sigma^2 I_n) \text{ mit } X := \begin{pmatrix} 1 & 0 & x_{11} \\ \vdots & \vdots & \vdots \\ 1 & 0 & x_{1n_1} \\ 1 & 1 & x_{21} \\ \vdots & \vdots & \vdots \\ 1 & 1 & x_{2n_2} \end{pmatrix}, \ \beta := \begin{pmatrix} \mu_0 \\ \alpha_2 \\ \beta_0 \end{pmatrix} \text{ und } \sigma^2 > 0 \tag{23}$$

Modell 2 | additive und interaktive Berücksichtigung der Kovariate

$$y_{ij} \sim N(\mu_{ij}, \sigma^2) \text{ mit } \mu_{ij} := \mu_0 + \alpha_i + (\beta_0 + \gamma_i) x_{ij} \text{ und } \alpha_1 := \gamma_1 := 0 \text{ und } \sigma^2 > 0 \tag{24}$$

$$\Leftrightarrow \quad y \sim N(X\beta, \sigma^2 I_n) \text{ mit } X := \begin{pmatrix} 1 & 0 & x_{11} & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & x_{1n_1} & 0 \\ 1 & 1 & x_{21} & x_{21} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & x_{2n_2} & x_{2n_2} \end{pmatrix}, \ \beta := \begin{pmatrix} \mu_0 \\ \alpha_2 \\ \beta_0 \\ \gamma_2 \end{pmatrix} \text{ und } \sigma^2 > 0 \tag{25}$$

Kovarianzanalyse mit Interaktion

Implementation

```r
fname     = file.path("./Daten/Kovarianzanalyse-2.csv")            # Datensatzdateiname
D         = read.table(fname, sep = ",", header = TRUE)            # Datensatz
n_i       = c(sum(D$Therapie == "F2F"), sum(D$Therapie == "ONL"))  # Anzahl Datenpunkte pro Gruppe
y         = D$dBDI                                                 # Daten
n         = length(y)                                              # Gesamtanzahl Datenpunkte
XS        = list()                                                 # Model 1 und 2 Liste
CS        = list()                                                 # Kontrastgewichtsvektorenliste
XS[[1]]   = matrix(c(rep(1,n_i[1]), rep(1,n_i[2]),                 # \mu_0    Regressor
                     rep(0,n_i[1]), rep(1,n_i[2]),                 # \alpha_2 Regressor
                     D$Dig.Affin), nrow = n)                       # \beta_2  Regressor
XS[[2]]   = matrix(c(rep(1,n_i[1]), rep(1,n_i[2]),                 # \mu_0    Regressor
                     rep(0,n_i[1]), rep(1,n_i[2]),                 # \alpha_2 Regressor
                     D$Dig.Affin,                                  # \beta_2  Regressor
                     D$Dig.Affin*(c(rep(0,n_i[1]),rep(1,n_i[2])))), # \gamma_2 Regressor
                   nrow = n)                                       # Anzahl Zeilen der Designmatrix
CS[[1]]   = matrix(c(0,1,0),   nrow = 3)                           # Kontrastgewichtsvektor \alpha_2
CS[[2]]   = matrix(c(0,1,0,0), nrow = 4)                           # Kontrastgewichtsvektor \alpha_2
B         = list()                                                 # Betaparameterschätzerliste
S         = rep(NaN,2)                                             # Varianzparameterschätzervektor
T         = rep(NaN,2)                                             # T-Statistik-Vektor
for(i in 1:2){                                                     # Iteration über Modelle
    X           = XS[[i]]                                          # Designmatrix
    p           = ncol(X)                                          # Anzahl Betaparameter
    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
    c           = CS[[i]]                                          # Kontrastgewichtsvektor
    t_num       = t(c) %*% beta_hat                                # Zähler der Zweistichproben-T-Teststatistik
    t_den       = sqrt(sigsqr_hat %*% t(c) %*% solve(t(X) %*% X) %*%c)# Nenner der Zweistichproben-T-Teststatistik
    t           = t_num/t_den                                      # Wert der   Zweistichproben-T-Teststatistik
    B[[i]]      = beta_hat                                         # Betaparameterschätzer
    S[i]        = sigsqr_hat                                       # Varianzparameterschätzer
    T[i]        = t                                                # T-Statistik
}
```

Kovarianzanalyse mit Interaktion

Ergebnisse

```
> Modell 1 | Betaparameterschätzer    :  2.73 1.7 0.58
> Modell 2 | Betaparameterschätzer    :  5.03 -3.47 0.14 0.99
> Modell 1 | Varianzparameterschätzer :  3.92
> Modell 2 | Varianzparameterschätzer :  3.15
> Modell 1 | T-Statistik ONL-Therapie :  2.43
> Modell 2 | T-Statistik ONL-Therapie :  -1.79
```

In Modell 1 wird der ONL-Therapieeffekt positiv, in Modell 2 negativ geschätzt.

Entsprechend ist die T-Statistik für ONL-Therapie in Modell 1 positiv, in Modell 2 negativ.

Die wahren, aber unbekannten Betaparameterwerte sind $\beta := (5, -3, 0, 1)^{\mathrm{T}}$.

$\Rightarrow$ Die Abhängigkeit von Digitalaffinität kommt bei ONL zum Tragen, bei F2F nicht.

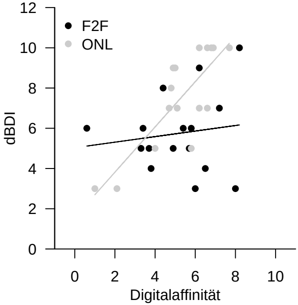
$\Rightarrow$ Der Haupteffekt der Therapieart ist schwierig zu beurteilen.

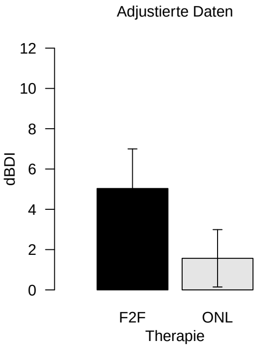
$\Rightarrow$ ONL-Therapie ist bei niedriger Digitalaffinität nicht so wirksam wie F2F-Therapie.

$\Rightarrow$ ONL-Therapie ist bei hoher Digitalaffinität wirksamer als F2F-Therapie.

Kovarianzanalyse mit Interaktion

Visualisierung





Vorbemerkungen

Additive Kovarianzanalyse

Kovarianzanalyse mit Interaktion

Selbstkontrollfragen

Selbstkontrollfragen

1. Erläutern Sie die Begriffe der Gesamt-/Primär-/Sekundär- und Residualvaribilität.
2. Geben Sie jeweils zwei Beispiele für parametrische bzw. faktorielle ALM-Designs an.
3. Geben Sie die Definition des Modells der additiven einfaktoriellen Kovarianzanalyse (EKA) wieder.
4. Erläutern Sie die Bedeutung der Parameter μ_0, α_i und β_0 im Modell der additiven EKA.
5. Geben Sie die Designmatrixform des Modells der additiven EKA für $I := 2$ und $n_1 = n_2 = 3$ an.
6. Erläutern Sie den Unterschied zwischen dem Modell der einfaktoriellen Varianzanalyse und dem Modell der additiven einfaktoriellen Kovarianzanalyse.
7. Geben Sie die Definition des Modells der einfaktoriellen Kovarianzanalyse (EKA) mit Interaktion wieder.
8. Erläutern Sie die Bedeutung der Parameter μ_0, α_i, β_0 und γ_i im Modell der EKA mit Interaktion.
9. Geben Sie die Designmatrixform des Modells der EKA mit Interaktion für $I := 2$ und $n_1 = n_2 = 3$ an.
10. Erläutern Sie den Unterschied zwischen dem Modell der additiven einfaktoriellen Kovarianzanalyse und dem Modell der einfaktoriellen Kovarianzanalyse mit Interaktion.