Allgemeines Lineares Modell

BSc Psychologie, SoSe 2024

Joram Soch

(11) Zweifaktorielle Varianzanalyse

Anwendungsszenario

Modellformulierung

Modellschätzung

Modellevaluation

Selbstkontrollfragen

Anwendungsszenario

Modellformulierung

Modellschätzung

Modellevaluation

Selbstkontrollfragen

Randomisiertes zweifaktorielles Studiendesigns mit gekreuztem Design

Zweifaktorielle Varianzanalyse (ZVA) = two-way analysis of variance (ANOVA):

- Eine univariate abhängige Variable, bestimmt an randomisierten experimentellen Einheiten.
- Zwei diskrete unabhängige Variablen, die mindestens zweistufig sind.
- Die unabhängigen Variablen werden **Faktoren** genannt.
- Die Stufen der Faktoren werden auch **Level** der Faktoren genannt.
- Jedes Level eines Faktors wird mit allen Leveln des anderen Faktors kombiniert.
- Die Kombination zweier spezifischer Faktorlevel wird **Zelle** des Designs genannt.

Zweifaktorielle Studiendesigns werden üblicherweise anhand ihrer Faktorlevel bezeichnet:

2×2 ANOVA:	Faktor A mit Level 1,2	Faktor B mit Level 1,2
2×3 ANOVA:	Faktor A mit Level 1,2	Faktor B mit Level 1,2,3
4×2 ANOVA:	Faktor A mit Level 1,2,3,4	Faktor B mit Level 1,2
3×1 ANOVA:	Faktor A mit Level 1,2,3	Faktor B mit Level 1

- Generell sind **2 x 2 Designs** sehr populär, wir fokussieren auf diesen Fall.
- Die Zellen eines 2 x 2 Designs werden im Folgenden auch als **Gruppen** bezeichnet.

Konzeptuelles Design

		Faktor B	
		Level 1	Level 2
Faktor A	Level 1	A1B1	A1B2
	Level 2	A2B1	A2B2

Die Zellen/Gruppen des Designs sind hier mit A1B1, A1B2, A2B1 und A2B2 bezeichnet.

Anwendungsszenario

Datennotation

		Faktor B: Level 1	Faktor B: Level 2
Faktor A	Level 1	$y_{111} = 1.2$ $y_{112} = 0.4$ $y_{113} = 1.7$ $\vdots$ $y_{11n_{11}} = 2.1$	$y_{121} = 4.1$ $y_{122} = 1.8$ $y_{123} = 3.3$ $\vdots$ $y_{12n_{12}} = 5.9$
	Level 2	$y_{211} = 0.1$ $y_{212} = 0.8$ $y_{213} = 2.7$ $\vdots$ $y_{21n_{21}} = 1.4$	$y_{221} = 7.4$ $y_{222} = 6.2$ $y_{223} = 9.5$ $\vdots$ $y_{22n_{22}} = 6.1$

y_{ijk} bezeichnet die Datenvariable der $k$ten experimentellen Einheit ($k = 1, ..., n_{ij}$)

im $i$ten Level von Faktor A und $j$ten Level von Faktor B ($i = 1, 2, j = 1, 2$).

Haupteffekte (*main effects*) und Interaktionen (*interactions*)

Man unterscheidet intuitiv hinsichtlich der Gruppenmittelwerte *Haupteffekte* und *Interaktionen*:

- Intuitiv spricht man vom Vorliegen eines *Haupteffekts von Faktor A*, wenn sich die Gruppenmittelwerte zwischen Level 1 und Level 2 von Faktor A, jeweils gemittelt über die zwei Level von Faktor B, unterscheiden.
- Intuitiv spricht man vom Vorliegen eines *Haupteffekts von Faktor B*, wenn sich die Gruppenmittelwerte zwischen Level 1 und Level 2 von Faktor B, jeweils gemittelt über die zwei Level von Faktor A, unterscheiden.
- Intuitiv spricht man vom Vorliegen einer *Interaktion der Faktoren A und B*, wenn der Unterschied der Gruppenmittelwerte von Faktor A zwischen Level 1 und 2 unterschiedlich für Level 1 und Level 2 von Faktor B ausgeprägt ist bzw. wenn der Unterschied der Gruppenmittelwerte von Faktor B zwischen Level 1 und 2 unterschiedlich für Level 1 und Level 2 von Faktor A ausgeprägt ist.

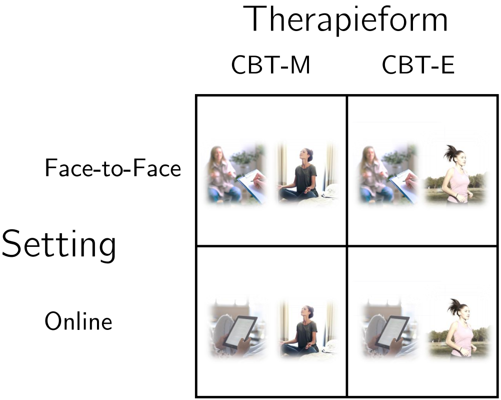
Intuitiv beziehen sich Haupteffekte also auf (marginale) Unterschiede (Differenzen), während sich Interaktionen auf Unterschiede von Unterschieden (Differenzen von Differenzen) beziehen.

Das Vorhandensein einer Interaktion besagt lediglich, dass sich die Unterschiede der Gruppenmittelwerte zwischen den Leveln eines experimentellen Faktors in Abhängigkeit von den Leveln des anderen experimentellen Faktors ändern, es macht aber keine Aussage darüber, warum dies so ist.

$\Rightarrow$ Haupteffekte und Interaktionen sind Datenmuster, keine wissenschaftlichen Theorien.

Anwendungsbeispiel

BDI Pre-Post-Differenzwertanalyse für zwei CBT-Settings und CBT-Varianten



Anwendungsbeispiel

BDI Pre-Post-Differenzwertanalyse für zwei CBT-Settings und CBT-Varianten

- Faktor (A) CBT-Setting mit Level (1) face-to-face (F2F) und Level (2) online (ONL).
- Faktor (B) CBT-Variante mit Level (1) mindfulness (MDN) und Level (2) exercise (EXC).
- $n_{ij} := 12,\ 1 \leq i, j \leq 2$ Patient:innen in jeder Bedingung, $n = 48$ Patient:innen insgesamt.

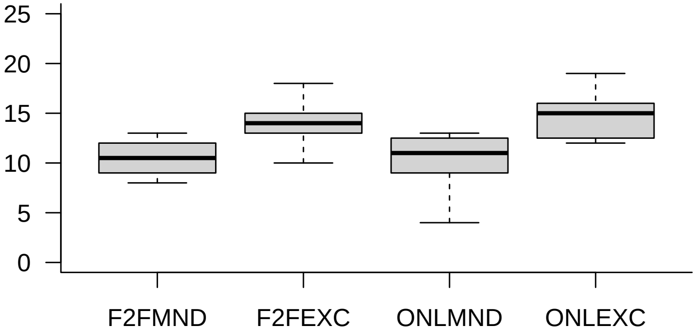
2 x 2 randomisiertes zweifaktorielles Studiendesign $\Rightarrow$ 2 x 2 ANOVA

- inferentielle Evidenz für einen Haupteffekt des CBT Settings?
- inferentielle Evidenz für einen Haupteffekt der CBT Variante?
- inferentielle Evidenz für Interaktion von CBT Setting und CBT Variante?

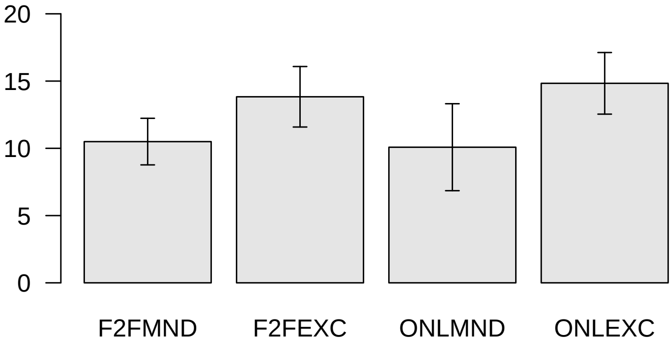
Beispieldatensatz: 4 Datenpunkte pro Gruppe

	Setting	Variant	dBDI
1	F2F	MND	12
2	F2F	MND	11
3	F2F	MND	8
4	F2F	MND	8
13	F2F	EXC	13
14	F2F	EXC	10
15	F2F	EXC	14
16	F2F	EXC	15
25	ONL	MND	4
26	ONL	MND	10
27	ONL	MND	12
28	ONL	MND	13
37	ONL	EXC	15
38	ONL	EXC	18
39	ONL	EXC	13
40	ONL	EXC	16

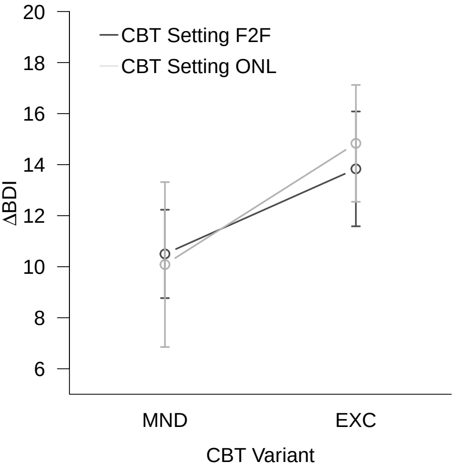
Visualisierung der Daten: Box-Plot



Visualisierung der Daten: Balkendiagramm



Visualisierung der Daten: Liniendiagramm



Anwendungsszenario

Modellformulierung

Modellschätzung

Modellevaluation

Selbstkontrollfragen

Modellformulierung

Modell der additiven ZVA

In Analogie zur einfaktoriellen Varianzanalyse (EVA) möchte man in der additiven ZVA die Gruppenerwartungswerte μ_{ij} mit $i = 1, ..., I$ für die Level von Faktor A und $j = 1, ..., J$ für die Level von Faktor B als Summe eines gruppenunspezifischen Erwartungswertes und den Effekten der Level von Faktor A und der Level von Faktor B modellieren.

Wir bezeichnen den gruppenunspezifischen Erwartungswertparameter mit μ_0, den Effekt von Level i von Faktor A mit α_i und den Effekt von Level j von Faktor B mit β_j (β_j bezeichnet hier also nicht den $j$ten Eintrag des Betaparametervektors). Dann ergibt sich zum Beispiel für $I := J := 2$:

$\mu_{11} := \mu_0 + \alpha_1 + \beta_1$	$\mu_{12} := \mu_0 + \alpha_1 + \beta_2$
$\mu_{21} := \mu_0 + \alpha_2 + \beta_1$	$\mu_{22} := \mu_0 + \alpha_2 + \beta_2$

Wie im Falle der EVA ist diese Darstellung der Gruppenerwartungswerte μ_{ij} allerdings überparameterisiert. Um eine eindeutige Darstellung der μ_{ij} zu gewährleisten, bietet sich auch hier die Restriktion an, den Effekt des ersten Levels jedes Faktors als Null zu definieren:

$$\alpha_1 := \beta_1 := 0 \ . \tag{1}$$

Dies etabliert die Faktorlevelkombination A1B1 als Referenzgruppe. Es ergibt sich somit zum Beispiel für für $I := J := 2$:

$\mu_{11} := \mu_0$	$\mu_{12} := \mu_0 + \beta_2$
$\mu_{21} := \mu_0 + \alpha_2$	$\mu_{22} := \mu_0 + \alpha_2 + \beta_2$

Auch bei dieser Effektdarstellung des Modells der additiven 2 x 2 ZVA mit Referenzgruppe ändern sich die Interpretation der Parameter μ_0, α_2, β_2 im Vergleich zum überparameterisierten Fall ohne Referenzgruppe: μ_0 entspricht dem Erwartungswert der Faktorlevelkombination A1B1, α_2 der Differenz beim Übergang von Level 1 zu Level 2 von Faktor A und β_2 der Differenz beim Übergang von Level 1 zu Level 2 von Faktor B.

Definition (Modell der additiven ZVA mit Referenzgruppe)

y_{ijk} mit $i = 1, ..., I, j = 1, ..., J, k = 1, ..., n_{ij}$ sei die Zufallsvariable, die den $k$ten Datenpunkt zum $i$ten Level von Faktor A und dem $j$ten Level von Faktor B in einem ZVA-Anwendungsszenario modelliert. Dann hat das *Modell der additiven ZVA mit Referenzgruppe* die strukturelle Form

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

und die Datenverteilungsform

$$y_{ijk} \sim N(\mu_{ij}, \sigma^2) \text{ u.v. für } i = 1, ..., I, j = 1, ..., J, k = 1, ..., n_{ij} \tag{3}$$

mit

$$\mu_{ij} := \mu_0 + \alpha_i + \beta_j \text{ für } i = 1, ..., I, j = 1, ..., J \text{ mit } \alpha_1 := \beta_1 := 0 \tag{4}$$

und $\sigma^2 > 0$.

Bemerkungen

- Das Modell der additiven ZVA modelliert ausschließlich Haupteffekte, keine Interaktionen.

Parameterbeispiele

(1) Es sei $\mu_0 := 1, \alpha_2 := 1, \beta_2 := 0$. Dann gilt:

$$\mu_{11} = \mu_0 + \alpha_1 + \beta_1 = 1 + 0 + 0 = 1 \qquad \mu_{12} = \mu_0 + \alpha_1 + \beta_2 = 1 + 0 + 0 = 1$$
$$\mu_{21} = \mu_0 + \alpha_2 + \beta_1 = 1 + 1 + 0 = 2 \qquad \mu_{22} = \mu_0 + \alpha_2 + \beta_2 = 1 + 1 + 0 = 2$$

$\Rightarrow$ Haupteffekt von Faktor A, kein Haupteffekt von Faktor B

(2) Es sei $\mu_0 := 1, \alpha_2 := 0, \beta_2 := 1$. Dann gilt:

$$\mu_{11} = \mu_0 + \alpha_1 + \beta_1 = 1 + 0 + 0 = 1 \qquad \mu_{12} = \mu_0 + \alpha_1 + \beta_2 = 1 + 0 + 1 = 2$$
$$\mu_{21} = \mu_0 + \alpha_2 + \beta_1 = 1 + 0 + 0 = 1 \qquad \mu_{22} = \mu_0 + \alpha_2 + \beta_2 = 1 + 0 + 1 = 2$$

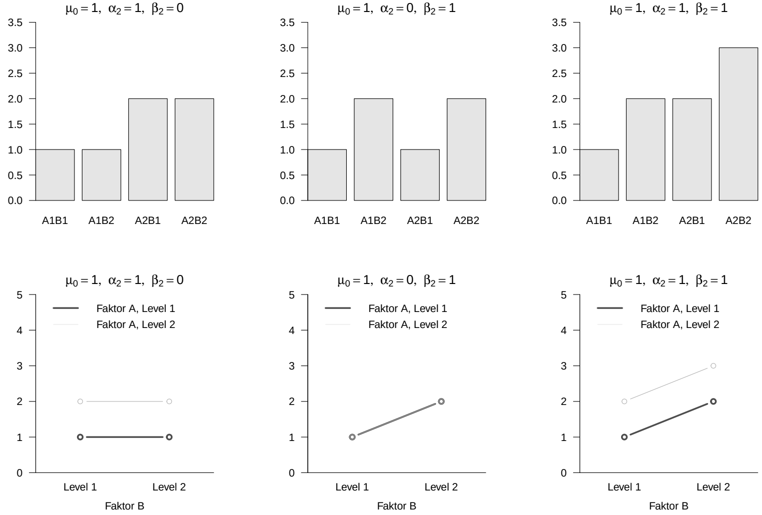
$\Rightarrow$ kein Haupteffekt von Faktor A, Haupteffekt von Faktor B

(3) Es sei $\mu_0 := 1, \alpha_2 := 1, \beta_2 := 1$. Dann gilt:

$$\mu_{11} = \mu_0 + \alpha_1 + \beta_1 = 1 + 0 + 0 = 1 \qquad \mu_{12} = \mu_0 + \alpha_1 + \beta_2 = 1 + 0 + 1 = 2$$
$$\mu_{21} = \mu_0 + \alpha_2 + \beta_1 = 1 + 1 + 0 = 2 \qquad \mu_{22} = \mu_0 + \alpha_2 + \beta_2 = 1 + 1 + 1 = 3$$

$\Rightarrow$ Haupteffekt von Faktor A, Haupteffekt von Faktor B

Parameterbeispiele



Modellformulierung

Theorem (Designmatrixform des Modells der additiven 2 x 2 ZVA mit Referenzgruppe)

Gegeben sei die strukturelle Form des Modells der additiven 2 x 2 ZVA mit Referenzgruppe. Dann hat dieses Modell die Designmatrixform

$$y \sim N(X\beta, \sigma^2 I_n), \tag{5}$$

wobei

$$y := \begin{pmatrix} y_{111} \\ \vdots \\ y_{11n_{11}} \\ y_{121} \\ \vdots \\ y_{12n_{12}} \\ y_{211} \\ \vdots \\ y_{21n_{21}} \\ y_{221} \\ \vdots \\ y_{22n_{22}} \end{pmatrix}, \ X = \begin{pmatrix} 1 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 1 \end{pmatrix} \in \mathbb{R}^{n \times 3}, \ \beta := \begin{pmatrix} \mu_0 \\ \alpha_2 \\ \beta_2 \end{pmatrix} \in \mathbb{R}^3 \text{ und } \sigma^2 > 0. \tag{6}$$

Bemerkung

- Das Theorem folgt direkt aus den Regel der Matrixmultiplikation.

Modell der ZVA mit Interaktion

In der ZVA mit Interaktion möchte man die Gruppenerwartungswerte μ_{ij} mit $i = 1, ..., I$ für die Level von Faktor A und $j = 1, ..., J$ für die Level von Faktor B als Summe eines gruppenunspezifischen Erwartungswertes, der Effekte der Level von Faktor A und Faktor B und der Interaktion der Level der Faktoren modellieren.

Wir bezeichnen den gruppenunspezifischen Erwartungswertparameter mit μ_0, den Effekt von Level i von Faktor A mit α_i, den Effekt von Level j von Faktor B mit β_j, und die Interaktion von Level i von Faktor A mit Level j von Faktor B mit γ_{ij}. Dann ergibt sich zum Beispiel für $I := J := 2$:

$\mu_{11} := \mu_0 + \alpha_1 + \beta_1 + \gamma_{11}$	$\mu_{12} := \mu_0 + \alpha_1 + \beta_2 + \gamma_{12}$
$\mu_{21} := \mu_0 + \alpha_2 + \beta_1 + \gamma_{21}$	$\mu_{22} := \mu_0 + \alpha_2 + \beta_2 + \gamma_{22}$

Wie in der die additiven ZVA ist diese Darstellung der Gruppenerwartungswerte μ_{ij} multipel überparameterisiert. Um eine eindeutige Darstellung der μ_{ij} zu gewährleisten, bietet sich auch hier die Restriktion an, den Effekt des ersten Levels jedes Faktors und jeder Interaktion als Null zu definieren:

$$\alpha_1 := \beta_1 := \gamma_{i1} := \gamma_{1j} := 0 \text{ für } i = 1, ..., I, j = 1, .., J \tag{7}$$

Dies etabliert die Faktorlevelkombination A1B1 als Referenzgruppe. Es ergibt sich somit zum Beispiel für $I := J := 2$:

$\mu_{11} := \mu_0$	$\mu_{12} := \mu_0 + \beta_2$
$\mu_{21} := \mu_0 + \alpha_2$	$\mu_{22} := \mu_0 + \alpha_2 + \beta_2 + \gamma_{22}$

Auch bei dieser Effektdarstellung des Modells der 2 x 2 ZVA mit Interaktion und Referenzgruppe ändert sich die Interpretation der Parameter $\mu_0, \alpha_2, \beta_2, \gamma_{22}$ im Vergleich zum überparameterisierten Fall ohne Referenzgruppe: μ_0 entspricht dem Erwartungswert der Faktorlevelkombination A1B1, α_2 der Differenz beim Übergang von Level 1 zu Level 2 von Faktor A, β_2 der Differenz beim Übergang von Level 1 zu Level 2 von Faktor B und γ_{22} der Differenz beim Übergang von Level 1 zu Level 2 von Faktor B im Unterschiede zum Übergang von Level 1 zu Level 2 von Faktor A.

Modellformulierung

Definition (Modell der ZVA mit Interaktion und Referenzgruppe)

y_{ijk} mit $i = 1, ..., I, j = 1, ..., J, k = 1, ..., n_{ij}$ sei die Zufallsvariable, die den $k$ten Datenpunkt zum $i$ten Level von Faktor A und dem $j$ten Level von Faktor B in einem ZVA-Anwendungsszenario modelliert. Dann hat das *Modell der ZVA mit Interaktion und Referenzgruppe* die strukturelle Form

$$y_{ijk} = \mu_{ij} + \varepsilon_{ijk} \text{ mit } \varepsilon_{ijk} \sim N(0, \sigma^2) \text{ u.i.v. für } i = 1, ..., I, j = 1, ..., J, k = 1, ..., n_{ij} \tag{8}$$

und die Datenverteilungsform

$$y_{ijk} \sim N(\mu_{ij}, \sigma^2) \text{ u.v. für } i = 1, ..., I, j = 1, ..., J, k = 1, ..., n_{ij} \tag{9}$$

mit

$$\mu_{ij} := \mu_0 + \alpha_i + \beta_j + \gamma_{ij} \tag{10}$$

sowie

$$\alpha_1 := \beta_1 := \gamma_{i1} := \gamma_{1j} := 0 \text{ für } i = 1, ..., I, j = 1, ..., J \tag{11}$$

und $\sigma^2 > 0$.

Bemerkungen

- Das Modell der ZVA mit Interaktion beinhaltet durch die Terme γ_{ij} für $i = 1, ..., I, j = 1, ..., J$ Interaktionseffekte, die Abweichungen vom Modell der additiven ZVA mit ausschließlich Haupteffekten repräsentieren.

Modellformulierung

Parameterbeispiele

(1) Es sei $\mu_0 := 1, \alpha_2 := 0, \beta_2 := 0, \gamma_{22} = 2$. Dann gilt:

$\mu_{11} = \mu_0 + \alpha_1 + \beta_1 + \gamma_{11} = 1 + 0 + 0 + 0 = 1$ $\quad$ $\mu_{12} = \mu_0 + \alpha_1 + \beta_2 + \gamma_{12} = 1 + 0 + 0 + 0 = 1$

$\mu_{21} = \mu_0 + \alpha_2 + \beta_1 + \gamma_{21} = 1 + 0 + 0 + 0 = 1$ $\quad$ $\mu_{22} = \mu_0 + \alpha_2 + \beta_2 + \gamma_{22} = 1 + 0 + 0 + 2 = 3$

$\Rightarrow$ kein Haupteffekt von Faktor A, kein Haupteffekt von Faktor B, Interaktion von A und B

(2) Es sei $\mu_0 := 1, \alpha_2 := 1, \beta_2 := 1, \gamma_{22} = -2$. Dann gilt:

$\mu_{11} = \mu_0 + \alpha_1 + \beta_1 + \gamma_{11} = 1 + 0 + 0 + 0 = 1$ $\quad$ $\mu_{12} = \mu_0 + \alpha_1 + \beta_2 + \gamma_{12} = 1 + 0 + 1 + 0 = 2$

$\mu_{21} = \mu_0 + \alpha_2 + \beta_1 + \gamma_{21} = 1 + 0 + 1 + 0 = 2$ $\quad$ $\mu_{22} = \mu_0 + \alpha_2 + \beta_2 + \gamma_{22} = 1 + 1 + 1 - 2 = 1$

$\Rightarrow$ Haupteffekt von Faktor A, Haupteffekt von Faktor B, negative Interaktion von A und B

(3) Es sei $\mu_0 := 1, \alpha_2 := 1, \beta_2 := 0, \gamma_{22} = 1$. Dann gilt:

$\mu_{11} = \mu_0 + \alpha_1 + \beta_1 + \gamma_{11} = 1 + 0 + 0 + 0 = 1$ $\quad$ $\mu_{12} = \mu_0 + \alpha_1 + \beta_2 + \gamma_{12} = 1 + 0 + 0 + 0 = 1$

$\mu_{21} = \mu_0 + \alpha_2 + \beta_1 + \gamma_{21} = 1 + 1 + 0 + 0 = 2$ $\quad$ $\mu_{22} = \mu_0 + \alpha_2 + \beta_2 + \gamma_{22} = 1 + 1 + 0 + 1 = 3$

$\Rightarrow$ Haupteffekt von Faktor A, kein Haupteffekt von Faktor B, Interaktion von A und B

(4) Es sei $\mu_0 := 1, \alpha_2 := 0, \beta_2 := 1, \gamma_{22} = 1$. Dann gilt:

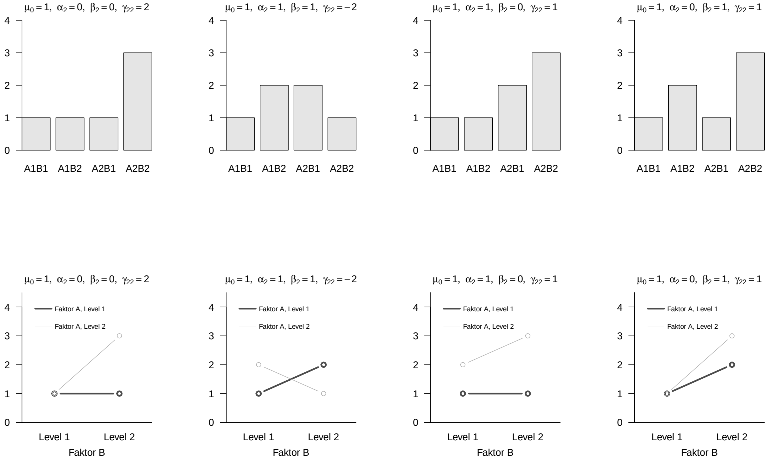
$\mu_{11} = \mu_0 + \alpha_1 + \beta_1 + \gamma_{11} = 1 + 0 + 0 + 0 = 1$ $\quad$ $\mu_{12} = \mu_0 + \alpha_1 + \beta_2 + \gamma_{12} = 1 + 0 + 1 + 0 = 2$

$\mu_{21} = \mu_0 + \alpha_2 + \beta_1 + \gamma_{21} = 1 + 0 + 0 + 0 = 1$ $\quad$ $\mu_{22} = \mu_0 + \alpha_2 + \beta_2 + \gamma_{22} = 1 + 0 + 1 + 1 = 3$

$\Rightarrow$ kein Haupteffekt von Faktor A, Haupteffekt von Faktor B, Interaktion von A und B

Modellformulierung

Parameterbeispiele



Theorem (Designmatrixform des Modells der 2 x 2 ZVA mit Interaktion und Referenzgruppe)

Gegeben sei die strukturelle Form eines 2 x 2 ZVA-Modells mit Interaktion und Referenzgruppe und es sei

$$n := \sum_{i=1}^{2}\sum_{i=1}^{2} n_{ij} \tag{12}$$

die Gesamtanzahl an Datenvariablen. Dann hat dieses Modell die Designmatrixform

$$y \sim N(X\beta, \sigma^2 I_n), \text{ mit} \tag{13}$$

$$y := \begin{pmatrix} y_{111} \\ \vdots \\ y_{11n_{11}} \\ y_{121} \\ \vdots \\ y_{12n_{12}} \\ y_{211} \\ \vdots \\ y_{21n_{21}} \\ y_{221} \\ \vdots \\ y_{22n_{22}} \end{pmatrix},\ X = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & 1 \end{pmatrix} \in \mathbb{R}^{n\times 4},\ \beta := \begin{pmatrix} \mu_0 \\ \alpha_2 \\ \beta_2 \\ \gamma_{22} \end{pmatrix} \in \mathbb{R}^4 \text{ und } \sigma^2 > 0. \tag{14}$$

Beispiel

Es seien

$$I := 2, J := 2 \text{ und } n_{ij} := 4 \text{ für } i = 1,2, j = 1,2, \text{ also } n = 16. \tag{15}$$

Dann gilt

$$y = X\beta + \varepsilon \text{ mit } \varepsilon \sim N(0_{16}, \sigma^2 I_{16}) \tag{16}$$

mit

$$y := \begin{pmatrix} y_{111} \\ y_{112} \\ y_{113} \\ y_{114} \\ y_{121} \\ y_{122} \\ y_{123} \\ y_{124} \\ y_{211} \\ y_{212} \\ y_{213} \\ y_{214} \\ y_{221} \\ y_{222} \\ y_{223} \\ y_{224} \end{pmatrix}, \ X := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \in \mathbb{R}^{16 \times 4}, \ \beta := \begin{pmatrix} \mu_0 \\ \alpha_2 \\ \beta_2 \\ \gamma_{22} \end{pmatrix} \in \mathbb{R}^4 \text{ und } \sigma^2 > 0. \tag{17}$$

Modellformulierung

Beispiel

```
# Modellformulierung
library(MASS)                                   # multivariate Normalverteilung
I      = 2                                      # Anzahl Level Faktor A
J      = 2                                      # Anzahl Level Faktor B
n_ij   = 4                                      # Anzahl von Datenpunkten der i,jten Gruppe
n      = I*J*n_ij                               # Anzahl Datenpunkte insgesamt
p      = 1 + (I-1)+(J-1)+(I*J-3)                # Anzahl Parameter insgesamt
D      = matrix(c(1,0,0,0,                      # prototypische Designmatrix für balancierte Designs
                  1,0,1,0,
                  1,1,0,0,
                  1,1,1,1),
                nrow  = p,
                byrow = TRUE)
C      = matrix(rep(1,n_ij), nrow = n_ij)       # prototypischer Zellenvektor für balancierte Designs
X      = kronecker(D,C)                         # Kroneckerprodukt-Designmatrix-Erzeugung für balancierte Designs
I_n    = diag(n)                                # n x n Einheitsmatrix
beta   = matrix(c(1,1,1,1), nrow = p)           # \beta = (\mu_0,\alpha_2,\alpha_3,\alpha_4)
sigsqr = 10                                     # \sigma^2

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

```
>       [,1] [,2] [,3] [,4]
>  [1,]    1    0    0    0
>  [2,]    1    0    0    0
>  [3,]    1    0    0    0
>  [4,]    1    0    0    0
>  [5,]    1    0    1    0
>  [6,]    1    0    1    0
>  [7,]    1    0    1    0
>  [8,]    1    0    1    0
>  [9,]    1    1    0    0
> [10,]    1    1    0    0
> [11,]    1    1    0    0
> [12,]    1    1    0    0
> [13,]    1    1    1    1
> [14,]    1    1    1    1
> [15,]    1    1    1    1
> [16,]    1    1    1    1
```

Anwendungsszenario

Modellformulierung

Modellschätzung

Modellevaluation

Selbstkontrollfragen

Theorem (Betaparameterschätzung im additiven 2 x 2 ZVA-Modell mit Referenzgruppe)

Gegeben sei die Designmatrixform eines balancierten additiven 2 x 2 ZVA-Modells mit Referenzgruppe. Dann ergibt sich für den Betaparameterschätzer

$$\hat{\beta} := \begin{pmatrix} \hat{\mu}_0 \\ \hat{\alpha}_2 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} \frac{3}{4}\bar{y}_{11} + \frac{1}{4}(\bar{y}_{12} + \bar{y}_{21}) - \frac{1}{4}\bar{y}_{22} \\ \frac{1}{2}(\bar{y}_{21} + \bar{y}_{22}) - \frac{1}{2}(\bar{y}_{11} + \bar{y}_{12}) \\ \frac{1}{2}(\bar{y}_{12} + \bar{y}_{22}) - \frac{1}{2}(\bar{y}_{11} + \bar{y}_{21}) \end{pmatrix}, \tag{18}$$

wobei

$$\bar{y}_{ij} := \frac{1}{n_{ij}} \sum_{k=1}^{n_{ij}} y_{ijk} \text{ für } 1 \le i, j \le 2 \tag{19}$$

das Stichprobenmittel der i, jten Gruppe des 2 x 2 ZVA-Designs bezeichnet.

Beweis

Wir bestimmen zunächst $X^{\mathrm{T}}y,\ X^{\mathrm{T}}X$ und $(X^{\mathrm{T}}X)^{-1}$ bei konstantem n_{ij} für $1 \le i, j \le 2.$

Beweis (fortgeführt)

$$
X^{\mathrm{T}}y = \begin{pmatrix} 1 & \cdots & 1 & 1 & \cdots & 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & \cdots & 1 & 1 & \cdots & 1 \\ 0 & \cdots & 0 & 1 & \cdots & 1 & 0 & \cdots & 0 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} y_{111} \\ \vdots \\ y_{11n_{11}} \\ y_{121} \\ \vdots \\ y_{12n_{12}} \\ y_{211} \\ \vdots \\ y_{21n_{21}} \\ y_{221} \\ \vdots \\ y_{22n_{22}} \end{pmatrix}
$$

$$
= \begin{pmatrix} \sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{k=1}^{n_{ij}} y_{ijk} \\ \sum_{j=1}^{2} \sum_{k=1}^{n_{2j}} y_{2jk} \\ \sum_{i=1}^{2} \sum_{k=1}^{n_{i2}} y_{i2k} \end{pmatrix} \tag{20}
$$

Modellschätzung

Beweis (fortgeführt)

$$
X^{\mathrm{T}}X = \begin{pmatrix} 1 & \cdots & 1 & 1 & \cdots & 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & \cdots & 1 & 1 & \cdots & 1 \\ 0 & \cdots & 0 & 1 & \cdots & 1 & 0 & \cdots & 0 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 1 \end{pmatrix} \tag{21}
$$

$$
= n_{ij} \begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix}
$$

Beweis (fortgeführt)

Ohne Beweis halten wir weiterhin fest, dass

$$(X^{\mathrm{T}}X)^{-1} = n_{ij}\begin{pmatrix} 4 & 2 & 2 \\ 2 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix}^{-1} = \frac{1}{n_{ij}}\begin{pmatrix} \frac{3}{4} & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{pmatrix}. \tag{22}$$

Es ergibt sich also

$$\hat{\beta} = \begin{pmatrix} \hat{\mu}_0 \\ \hat{\alpha}_2 \\ \hat{\beta}_2 \end{pmatrix} = \frac{1}{n_{ij}}\begin{pmatrix} \frac{3}{4} & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{pmatrix}\begin{pmatrix} \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{n_{ij}} y_{ijk} \\ \sum_{j=1}^{2}\sum_{k=1}^{n_{2j}} y_{2jk} \\ \sum_{i=1}^{2}\sum_{k=1}^{n_{i2}} y_{i2k} \end{pmatrix}. \tag{23}$$

Beweis (fortgeführt)

Damit ergibt sich dann

$$
\begin{aligned}
\hat{\mu}_0 &= \frac{1}{n_{ij}}\left(\frac{3}{4}\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{n_{ij}} y_{ijk} - \frac{1}{2}\sum_{j=1}^{2}\sum_{k=1}^{n_{2j}} y_{2jk} - \frac{1}{2}\sum_{i=1}^{2}\sum_{k=1}^{n_{i2}} y_{i2k}\right) \\
&= \frac{1}{n_{ij}}\left(\frac{3}{4}\sum_{k=1}^{n_{11}} y_{11k} + \frac{3}{4}\sum_{k=1}^{n_{12}} y_{12k} + \frac{3}{4}\sum_{k=1}^{n_{21}} y_{21k} + \frac{3}{4}\sum_{k=1}^{n_{22}} y_{22k}\right) \\
&+ \frac{1}{n_{ij}}\left(-\frac{1}{2}\sum_{k=1}^{n_{21}} y_{21k} - \frac{1}{2}\sum_{k=1}^{n_{22}} y_{22k} - \frac{1}{2}\sum_{k=1}^{n_{12}} y_{12k} - \frac{1}{2}\sum_{k=1}^{n_{22}} y_{22k}\right) \\
&= \frac{1}{n_{ij}}\left(\frac{3}{4}\sum_{k=1}^{n_{11}} y_{11k} + \frac{1}{4}\sum_{k=1}^{n_{12}} y_{12k} + \frac{1}{4}\sum_{k=1}^{n_{21}} y_{21k} - \frac{1}{4}\sum_{k=1}^{n_{22}} y_{22k}\right) \\
&= \frac{3}{4}\bar{y}_{11} + \frac{1}{4}(\bar{y}_{12} + \bar{y}_{21}) - \frac{1}{4}\bar{y}_{22}
\end{aligned}
\tag{24}
$$

Beweis (fortgeführt)

sowie

$$
\begin{aligned}
\hat{\alpha}_2 &= \frac{1}{n_{ij}} \left(-\frac{1}{2} \sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{k=1}^{n_{ij}} y_{ijk} + \sum_{j=1}^{2} \sum_{k=1}^{n_{2j}} y_{2jk} \right) \\
&= \frac{1}{n_{ij}} \left(-\frac{1}{2} \sum_{k=1}^{n_{11}} y_{11k} - \frac{1}{2} \sum_{k=1}^{n_{12}} y_{12k} - \frac{1}{2} \sum_{k=1}^{n_{21}} y_{21k} - \frac{1}{2} \sum_{k=1}^{n_{22}} y_{22k} + \sum_{k=1}^{n_{21}} y_{21k} + \sum_{k=1}^{n_{22}} y_{22k} \right) \\
&= \frac{1}{n_{ij}} \left(-\frac{1}{2} \sum_{k=1}^{n_{11}} y_{11k} - \frac{1}{2} \sum_{k=1}^{n_{12}} y_{12k} + \frac{1}{2} \sum_{k=1}^{n_{21}} y_{21k} + \frac{1}{2} \sum_{k=1}^{n_{22}} y_{22k} \right) \\
&= \frac{1}{2}(\bar{y}_{21} + \bar{y}_{22}) - \frac{1}{2}(\bar{y}_{11} + \bar{y}_{12})
\end{aligned}
\tag{25}
$$

und analog für $\hat{\beta}_2$. $\square$

Beispiel

```
# Datenreformatierung
fname      = file.path(getwd(), "Daten", "Zweifaktorielle_Varianzanalyse_Daten.csv")
D          = read.table(fname, sep = ",", header = TRUE)     # Datensatz
A1B1       = D$dBDI[D$Setting == "F2F" & D$Variant == "MND"] # face-to-face, mindfulness
A1B2       = D$dBDI[D$Setting == "F2F" & D$Variant == "EXC"] # face-to-face, exercise
A2B1       = D$dBDI[D$Setting == "ONL" & D$Variant == "MND"] # online,       mindfulness
A2B2       = D$dBDI[D$Setting == "ONL" & D$Variant == "EXC"] # online,       exercise

# Datenmatrix für Gruppenmittelwerte
n_ij       = length(A1B1)                                    # Anzahl der Datenpunkte pro Gruppe
Y          = matrix(c(A1B1,A1B2,A2B1,A2B2), nrow = n_ij)     # n_ij x 4 Datenmatrix
bar_y      = colMeans(Y)                                     #    1 x 4 Zellenmittelwerte

# Modellschätzung
I          = 2                                               # Anzahl Level Faktor A (Therapiesetting)
J          = 2                                               # Anzahl Level Faktor B (Therapievariante)
n          = I*J*n_ij                                        # Anzahl Datenpunkte insgesamt
p          = 1 + (I-1)+(J-1)+(I*J-3)                         # Anzahl Parameter insgesamt
D          = matrix(c(1,0,0,                                 # prototypische Designmatrix für balancierte Designs
                      1,0,1,
                      1,1,0,
                      1,1,1),
                    nrow  = p,
                    byrow = TRUE)
C          = matrix(rep(1,n_ij),nrow = n_ij)                 # prototypischer Zellenvektor für balancierte Designs
X          = kronecker(D,C)                                  # Kroneckerprodukt-Designmatrix
y          = matrix(c(A1B1,A1B2,A2B1,A2B2), nrow = n)        # Datenvektor
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
```

Modellschätzung

Betaparameterschätzer und Stichprobenmittelwerte

```
> hat{beta}                                                 :  10.1 0.292 4.04
> hat{sigsqr}                                               :  6.07
> bar{y}_11, bar{y}_12, bar{y}_21, bar{y}_22                :  10.5 13.8 10.1 14.8
> 3/4bar{y}_11 + 1/4(bar{y}_12 + bar{y}_21) - 1/4bar{y}_22 :  10.1
> 1/2(bar{y}_21 + bar{y}_22) - 1/2(bar{y}_11 + bar{y}_12)  :  0.292
> 1/2(bar{y}_12 + bar{y}_22) - 1/2(bar{y}_11 + bar{y}_21)  :  4.04
```

Theorem (Betaparameterschätzung im 2 x 2 ZVA-Modell mit Interaktion und Referenzgruppe)

Gegeben sei die Designmatrixform eines balancierten 2 x 2 ZVA-Modells mit Interaktion und Referenzgruppe. Dann ergibt sich für den Betaparameterschätzer

$$\hat{\beta} := \begin{pmatrix} \hat{\mu}_0 \\ \hat{\alpha}_2 \\ \hat{\beta}_2 \\ \hat{\gamma}_{22} \end{pmatrix} = \begin{pmatrix} \bar{y}_{11} \\ \bar{y}_{21} - \bar{y}_{11} \\ \bar{y}_{12} - \bar{y}_{11} \\ \bar{y}_{11} + \bar{y}_{22} - \bar{y}_{12} - \bar{y}_{21} \end{pmatrix}, \tag{26}$$

wobei

$$\bar{y}_{ij} := \frac{1}{n_{ij}} \sum_{k=1}^{n_{ij}} y_{ijk} \text{ für } 1 \le i, j \le 2 \tag{27}$$

das Stichprobenmittel der i, jten Gruppe des 2 x 2 ZVA-Designs bezeichnet.

Beweis

Wir bestimmen zunächst $X^{\mathrm{T}}y,\ X^{\mathrm{T}}X$ und $(X^{\mathrm{T}}X)^{-1}$ bei konstantem n_{ij} für $1 \le i, j \le 2$.

Modellschätzung

Beweis (fortgeführt)

$$
X^{\mathrm{T}}y = \begin{pmatrix}
1 & \cdots & 1 & 1 & \cdots & 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \\
0 & \cdots & 0 & 0 & \cdots & 0 & 1 & \cdots & 1 & 1 & \cdots & 1 \\
0 & \cdots & 0 & 1 & \cdots & 1 & 0 & \cdots & 0 & 1 & \cdots & 1 \\
0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & \cdots & 1
\end{pmatrix}
\begin{pmatrix}
y_{111} \\ \vdots \\ y_{11n_{11}} \\ y_{121} \\ \vdots \\ y_{12n_{12}} \\ y_{211} \\ \vdots \\ y_{21n_{21}} \\ y_{221} \\ \vdots \\ y_{22n_{22}}
\end{pmatrix}
\tag{28}
$$

$$
= \begin{pmatrix}
\sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{k=1}^{n_{ij}} y_{ijk} \\
\sum_{j=1}^{2} \sum_{k=1}^{n_{2j}} y_{2jk} \\
\sum_{i=1}^{2} \sum_{k=1}^{n_{i2}} y_{i2k} \\
\sum_{k=1}^{n_{22}} y_{22k}
\end{pmatrix}
$$

Beweis (fortgeführt)

$$
X^{\mathrm{T}}X =
\begin{pmatrix}
1 & \cdots & 1 & 1 & \cdots & 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \\
0 & \cdots & 0 & 0 & \cdots & 0 & 1 & \cdots & 1 & 1 & \cdots & 1 \\
0 & \cdots & 0 & 1 & \cdots & 1 & 0 & \cdots & 0 & 1 & \cdots & 1 \\
0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & \cdots & 1
\end{pmatrix}
\begin{pmatrix}
1 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots \\
1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 \\
\vdots & \vdots & \vdots & \vdots \\
1 & 0 & 1 & 0 \\
1 & 1 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots \\
1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 \\
\vdots & \vdots & \vdots & \vdots \\
1 & 1 & 1 & 1
\end{pmatrix}
\tag{29}
$$

$$
= n_{ij}
\begin{pmatrix}
4 & 2 & 2 & 1 \\
2 & 2 & 1 & 1 \\
2 & 1 & 2 & 1 \\
1 & 1 & 1 & 1
\end{pmatrix}
$$

Beweis (fortgeführt)

Ohne Beweis halten wir weiterhin fest, dass

$$(X^{\mathrm{T}}X)^{-1} = n_{ij}\begin{pmatrix} 4 & 2 & 2 & 1 \\ 2 & 2 & 1 & 1 \\ 2 & 1 & 2 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}^{-1} = \frac{1}{n_{ij}}\begin{pmatrix} 1 & -1 & -1 & 1 \\ -1 & 2 & 1 & -2 \\ -1 & 1 & 2 & -2 \\ 1 & -2 & -2 & 4 \end{pmatrix}. \tag{30}$$

Es ergibt sich also

$$\hat{\beta} = \begin{pmatrix} \hat{\mu}_0 \\ \hat{\alpha}_2 \\ \hat{\beta}_2 \\ \hat{\gamma}_{22} \end{pmatrix} = \frac{1}{n_{ij}}\begin{pmatrix} 1 & -1 & -1 & 1 \\ -1 & 2 & 1 & -2 \\ -1 & 1 & 2 & -2 \\ 1 & -2 & -2 & 4 \end{pmatrix}\begin{pmatrix} \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{n_{ij}} y_{ijk} \\ \sum_{j=1}^{2}\sum_{k=1}^{n_{2j}} y_{2jk} \\ \sum_{i=1}^{2}\sum_{k=1}^{n_{i2}} y_{i2k} \\ \sum_{k=1}^{n_{22}} y_{22k} \end{pmatrix}. \tag{31}$$

Beweis (fortgeführt)

Damit ergibt sich dann

$$
\begin{aligned}
\hat{\mu}_0 &= \frac{1}{n_{ij}} \left(\sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{k=1}^{n_{ij}} y_{ijk} - \sum_{j=1}^{2} \sum_{k=1}^{n_{2j}} y_{2jk} - \sum_{i=1}^{2} \sum_{k=1}^{n_{i2}} y_{i2k} + \sum_{k=1}^{n_{22}} y_{22k} \right) \\
&= \frac{1}{n_{ij}} \left(\sum_{k=1}^{n_{11}} y_{11k} + \sum_{k=1}^{n_{12}} y_{12k} + \sum_{k=1}^{n_{21}} y_{21k} + \sum_{k=1}^{n_{22}} y_{22k} \right) \\
&+ \frac{1}{n_{ij}} \left(- \sum_{k=1}^{n_{21}} y_{21k} - \sum_{k=1}^{n_{22}} y_{22k} - \sum_{k=1}^{n_{12}} y_{12k} - \sum_{k=1}^{n_{22}} y_{22k} + \sum_{k=1}^{n_{22}} y_{22k} \right) \\
&= \frac{1}{n_{11}} \sum_{k=1}^{n_{11}} y_{11k} \\
&= \bar{y}_{11}
\end{aligned}
\tag{32}
$$

Beweis (fortgeführt)

sowie

$$
\begin{aligned}
\hat{\alpha}_2 &= \frac{1}{n_{ij}} \left(-\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{n_{ij}} y_{ijk} + 2\sum_{j=1}^{2}\sum_{k=1}^{n_{2j}} y_{2jk} + 1\sum_{i=1}^{2}\sum_{k=1}^{n_{i2}} y_{i2k} - 2\sum_{k=1}^{n_{22}} y_{22k} \right) \\
&= \frac{1}{n_{ij}} \left(-\sum_{k=1}^{n_{11}} y_{11k} - \sum_{k=1}^{n_{12}} y_{12k} - \sum_{k=1}^{n_{21}} y_{21k} - \sum_{k=1}^{n_{22}} y_{22k} \right) \\
&+ \frac{1}{n_{ij}} \left(2\sum_{k=1}^{n_{21}} y_{21k} + 2\sum_{k=1}^{n_{22}} y_{22k} + \sum_{k=1}^{n_{12}} y_{12k} + \sum_{k=1}^{n_{22}} y_{22k} - 2\sum_{k=1}^{n_{22}} y_{22k} \right) \\
&= \frac{1}{n_{ij}} \left(\sum_{k=1}^{n_{21}} y_{21k} - \sum_{k=1}^{n_{11}} y_{11k} \right) \\
&= \bar{y}_{21} - \bar{y}_{11}
\end{aligned}
\tag{33}
$$

und analog für $\hat{\beta}_2$.

Beweis (fortgeführt)

Schließlich ergibt sich

$$
\begin{aligned}
\hat{\gamma}_{22} &= \frac{1}{n_{ij}} \left(\sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{k=1}^{n_{ij}} y_{ijk} - 2 \sum_{j=1}^{2} \sum_{k=1}^{n_{2j}} y_{2jk} - 2 \sum_{i=1}^{2} \sum_{k=1}^{n_{i2}} y_{i2k} + 4 \sum_{k=1}^{n_{22}} y_{22k} \right) \\
&= \frac{1}{n_{ij}} \left(\sum_{k=1}^{n_{11}} y_{11k} + \sum_{k=1}^{n_{12}} y_{12k} + \sum_{k=1}^{n_{21}} y_{21k} + \sum_{k=1}^{n_{22}} y_{22k} \right) \\
&+ \frac{1}{n_{ij}} \left(-2 \sum_{k=1}^{n_{21}} y_{21k} - 2 \sum_{k=1}^{n_{22}} y_{22k} - 2 \sum_{k=1}^{n_{12}} y_{12k} - 2 \sum_{k=1}^{n_{22}} y_{22k} + 4 \sum_{k=1}^{n_{22}} y_{22k} \right) \\
&= \frac{1}{n_{ij}} \left(\sum_{k=1}^{n_{11}} y_{11k} + \sum_{k=1}^{n_{22}} y_{22k} - \sum_{k=1}^{n_{12}} y_{12k} - \sum_{k=1}^{n_{21}} y_{21k} \right) \\
&= \bar{y}_{11} + \bar{y}_{22} - \bar{y}_{12} - \bar{y}_{21}.
\end{aligned}
\tag{34}
$$

$\square$

Modellschätzung

Beispiel

```
# Datenreformatierung
fname      = file.path(getwd(), "Daten", "Zweifaktorielle_Varianzanalyse_Daten.csv")
D          = read.table(fname, sep = ",", header = TRUE)     # Datensatz
A1B1       = D$dBDI[D$Setting == "F2F" & D$Variant == "MND"] # face-to-face, mindfulness
A1B2       = D$dBDI[D$Setting == "F2F" & D$Variant == "EXC"] # face-to-face, exercise
A2B1       = D$dBDI[D$Setting == "ONL" & D$Variant == "MND"] # online,       mindfulness
A2B2       = D$dBDI[D$Setting == "ONL" & D$Variant == "EXC"] # online,       exercise

# Datenmatrix für Gruppenmittelwerte
n_ij       = length(A1B1)                                    # Anzahl der Datenpunkte pro Gruppe
Y          = matrix(c(A1B1,A1B2,A2B1,A2B2), nrow = n_ij)     # n_ij x 4 Datenmatrix
bar_y      = colMeans(Y)                                     #    1 x 4 Zellenmittelwerte

# Modellschätzung
I          = 2                                               # Anzahl Level Faktor A (Therapiesetting)
J          = 2                                               # Anzahl Level Faktor B (Therapievariante)
n          = I*J*n_ij                                        # Anzahl Datenpunkte insgesamt
p          = 1 + (I-1)+(J-1)+(I*J-3)                         # Anzahl Parameter insgesamt
D          = matrix(c(1,0,0,0,                               # prototypische Designmatrix für balancierte Designs
                      1,0,1,0,
                      1,1,0,0,
                      1,1,1,1),
                    nrow  = p,
                    byrow = TRUE)
C          = matrix(rep(1,n_ij), nrow = n_ij)                # prototypischer Zellenvektor für balancierte Designs
X          = kronecker(D,C)                                  # Kroneckerprodukt-Designmatrix
y          = matrix(c(A1B1,A1B2,A2B1,A2B2), nrow = n)        # Datenvektor
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
```

Modellschätzung

Betaparameterschätzer und Stichprobenmittelwerte

```
> hat{beta}                                      :  10.5 -0.417 3.33 1.42
> hat{sigsqr}                                    :  5.94
> bar{y}_11, bar{y}_12, bar{y}_21, bar{y}_22     :  10.5 13.8 10.1 14.8
> bar{y}_11                                      :  10.5
> bar{y}_21 - bar{y}_11                          :  -0.417
> bar{y}_12 - bar{y}_11                          :  3.33
> bar{y}_11 + bar{y}_22 - bar{y}_12 + bar{y}_21 :  1.42
```

Anwendungsszenario

Modellformulierung

Modellschätzung

Modellevaluation

Selbstkontrollfragen

Überblick

Wie bei der EVA kann auch bei der ZVA eine Modellevaluationstheorie mithilfe einer Quadratsummenzerlegung entwickelt werden (siehe Einheit (10) in *Allgemeines Lineares Modell*). Mit zunehmender Designkomplexität wird eine solche Theorie allerdings zunehmend unübersichtlich. Prinzipiell können alle Quadratsummenzerlegung-basierten F-Statistiken auf Likelihood-basierte Modellvergleiche zurückgeführt werden. Allerdings fehlt dazu in diesem Kurs bisher eine allgemeine Kontrasttheorie. Wir beschränken und im Folgenden deshalb auf die

(1) Evaluation der Haupteffekte im additiven Modell der 2×2 ZVA mit Referenzgruppe und die
(2) Evaluation der Interaktion im Modell der 2×2 ZVA mit Interaktion und Referenzgruppe.

Evaluation der Haupteffekte im additiven Modell der 2 x 2 ZVA mit Referenzgruppe

(1) Statistische Modelle und Teststatistiken

(2) Testhypothesen und Tests

(3) Testumfangkontrollen und p-Werte

Evaluation der Interaktion im Modell der 2 x 2 ZVA mit Interaktion und Referenzgruppe

(1) Statistisches Modell und Teststatistik

(2) Testhypothese und Test

(3) Testumfangkontrolle und p-Wert

Modellevaluation: Haupteffekte (1) Teststatistiken

Theorem (Teststatistiken für Haupteffekte)

Gegeben sei die Designmatrixform des additiven Modells der 2 x 2 ZVA mit Referenzgruppe

$$y = X\beta + \varepsilon \text{ mit } \varepsilon \sim N(0_n, \sigma^2 I_n)\,, \tag{35}$$

wobei die Spalten von X bezeichnet seien durch

$$X := \begin{pmatrix} X_{\mu_0} & X_{\alpha_2} & X_{\beta_2} \end{pmatrix} \in \mathbb{R}^{n\times 3}\,. \tag{36}$$

Dann gelten:

(A) Eine *F-Teststatistik für den Haupteffekt von Faktor A*, F_A, ist die F-Statistik unter der Partitionierung

$$X := \begin{pmatrix} X_{\mu_0} & X_{\beta_2} & X_{\alpha_2} \end{pmatrix},\ \beta := \begin{pmatrix} \mu_0 \\ \beta_2 \\ \alpha_2 \end{pmatrix} \quad \text{und} \quad X_0 := \begin{pmatrix} X_{\mu_0} & X_{\beta_2} \end{pmatrix},\ \beta_0 := \begin{pmatrix} \mu_0 \\ \beta_2 \end{pmatrix}\,. \tag{37}$$

(B) Eine *F-Teststatistik für den Haupteffekt von Faktor B*, F_B, ist die F-Statistik unter der Partitionierung

$$X := \begin{pmatrix} X_{\mu_0} & X_{\alpha_2} & X_{\beta_2} \end{pmatrix},\ \beta := \begin{pmatrix} \mu_0 \\ \alpha_2 \\ \beta_2 \end{pmatrix} \quad \text{und} \quad X_0 := \begin{pmatrix} X_{\mu_0} & X_{\alpha_2} \end{pmatrix},\ \beta_0 := \begin{pmatrix} \mu_0 \\ \alpha_2 \end{pmatrix}\,. \tag{38}$$

Bemerkung

- Der Beweis ergibt sich aus dem Theorem zur Verteilung der F-Statistik bei Partitionierung eines ALM (siehe Einheit (8) in *Allgemeines Lineares Modell*).

Modellevaluation: Haupteffekte (2) Testhypothesen

Definition (Testhypothesen und Tests)

Gegeben sei das Modell der additiven 2 x 2 ZVA mit Referenzgruppe. Die F-Teststatistiken für die Haupteffekte von Faktor A und B seien mit F_A und F_B bezeichnet und wie oben definiert. Dann gilt:

(A) Der kritische-Wert-basierte Test

$$\phi_A(y) := 1_{\{F_A \geq k\}} \text{ mit Nullhypothese } H_0^A : \alpha_2 = 0 \tag{39}$$

definiert den *F-Test des Haupteffekts von Faktor A*.

(B) Der kritische-Wert-basierte Test

$$\phi_B(y) := 1_{\{F_B \geq k\}} \text{ mit Nullhypothese } H_0^B : \beta_2 = 0 \tag{40}$$

definiert den *F-Test des Haupteffekts von Faktor B*.

Bemerkung

- Die F-Tests beziehen sich auf die im vorangegangenen Theorem definierten F-Statistiken.

Modellevaluation: Haupteffekte (3) Testumfangkontrolle

Theorem (Testumfangkontrolle und p-Werte)

Mit obigen Definitionen und der KVF $\varphi(\cdot; n_1, n_2)$ der f-Verteilung gelten:

(A) ϕ_A ist ein Level-α_0-Test mit Testumfang α_0, wenn der kritische Wert definiert ist durch

$$k_{\alpha_0}^A := \varphi^{-1}(1 - \alpha_0; 1, n - 3). \tag{41}$$

Der zu einem beobachteten Wert f_A von F_A assoziierte p-Wert ist gegeben durch

$$\text{p-Wert} := 1 - \varphi(f_A; 1, n - 3). \tag{42}$$

(B) ϕ_B ist ein Level-α_0-Test mit Testumfang α_0, wenn der kritische Wert definiert ist durch

$$k_{\alpha_0}^B := \varphi^{-1}(1 - \alpha_0; 1, n - 3). \tag{43}$$

Der zu einem beobachteten Wert f_B von F_B assoziierte p-Wert ist gegeben durch

$$\text{p-Wert} := 1 - \varphi(f_B; 1, n - 3). \tag{44}$$

Bemerkung

- Wir verzichten auf einen Beweis.

Modellevaluation: Haupteffekte

```
# Modellevaluation
I           = 2                                         # Anzahl Level Faktor A (Therapiesetting)
J           = 2                                         # Anzahl Level Faktor B (Therapievariante)
n_ij        = length(A1B1)                              # balanciertes ZVA-Design
n           = I*J*n_ij                                  # Anzahl Datenpunkte
p           = 3                                         # Anzahl Parameter vollständiges Modell
y           = matrix(c(A1B1,A1B2,A2B1,A2B2), nrow = n)  # Datenvektor
D           = matrix(c(1,0,0,  1,0,1,  1,1,0,  1,1,1),
                     nrow = I*J, byrow = TRUE)          # prototypische Designmatrix
C           = matrix(rep(1,n_ij), nrow = n_ij)          # prototypischer Zellenvektor für balancierte Designs
X           = kronecker(D,C)                            # Kroneckerprodukt-Designmatrix
XH          = list(X[,c(1,3,2)], X)                     # Modellvarianten
alpha_0     = 0.05                                      # Signifikanzlevel
Eff         = rep(NaN,2)                                # F-Teststatistik Arrayinitialisierung
k_alpha_0   = rep(NaN,2)                                # kritischer Wert Arrayinitialisierung
phi         = rep(NaN,2)                                # Testwert Arrayinitialisierung
p_vals      = rep(NaN,2)                                # p-Wert Arrayinitialisierung
for(i in 1:2){                                          # Iteration über Modellvarianten
    X           = XH[[i]]                               # Designmatrix vollständiges Modell
    X_0         = X[,-3]                                # Designmatrix reduziertes Modell
    p           = ncol(X)                               # Anzahl Parameter vollständiges Modell
    p_0         = ncol(X_0)                             # Anzahl Parameter reduziertes Modell
    p_1         = p - p_0                               # Anzahl zusätzlicher Parameter im vollständigen Modell
    beta_hat_0  = solve(t(X_0)%*%X_0)%*%t(X_0)%*%y      # Betaparameterschätzer reduziertes Modell
    beta_hat    = solve(t(X) %*%X )%*%t(X) %*%y         # Betaparameterschätzer vollständiges Modell
    eps_hat_0   = y-X_0%*%beta_hat_0                    # Residuenvektor reduziertes Modell
    eps_hat     = y - X%*%beta_hat                      # Residuenvektor vollständiges Modell
    eh0_eh0     = t(eps_hat_0) %*% eps_hat_0            # residuelle QS reduziertes Modell
    eh_eh       = t(eps_hat) %*% eps_hat                # esiduelle QS vollständiges Modell
    sigsqr_hat  = eh_eh/(n-p)                           # Varianzparameterschätzer vollständiges Modell
    Eff[i]      = ((eh0_eh0-eh_eh)/p_1)/sigsqr_hat      # F-Statistik
    k_alpha_0[i] = qf(1-alpha_0, p_1, n-p)              # kritischer Wert
    if(Eff[i] >= k_alpha_0[i]){ phi[i] = 1 }            # H_0 ablehnen
    else {                      phi[i] = 0 }            # H_0 nicht ablehnen
    p_vals[i]   = 1 - pf(Eff[i], p_1,n-p)               # p-Wert
}
data.frame("f"=Eff, "k"= k_alpha_0, "phi"=phi, "p-Wert"= p_vals, row.names = c("Setting", "Variant"))
```

```
>              f    k phi   p.Wert
> Setting  0.172 4.06   0 6.80e-01
> Variant 33.004 4.06   1 7.46e-07
```

Modellevaluation: Haupteffekte

Haupteffekte der additiven 2 x 2 ZVA mit R's aov() Funktion

```
# Daten einlesen
fname      = file.path(getwd(), "Daten", "Zweifaktorielle_Varianzanalyse_Daten.csv")
D          = read.table(fname, sep = ",", header = TRUE)

# R's aov-Funktion
res.aov    = aov(dBDI ~ Setting + Variant, data = D)          # Modellformulierung und Modellschätzung
summary(res.aov)                                               # Modellevaluation
```

```
>             Df Sum Sq Mean Sq F value  Pr(>F)
> Setting      1      1     1.0    0.17    0.68
> Variant      1    196   196.0   33.00 7.5e-07 ***
> Residuals   45    267     5.9
> ---
> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Haupteffekte der additiven 2 x 2 ZVA mit R's lm() und anova() Funktionen

```
# Dateneinlesen
fname      = file.path(getwd(), "Daten", "Zweifaktorielle_Varianzanalyse_Daten.csv")
D          = read.table(fname, sep = ",", header = TRUE)

# R's lm() und anova() Funktionen
glm        = lm(dBDI ~ Setting + Variant, data = D)           # Modellformulierung und Modellschätzung
anova(glm)                                                     # Modellevaluation
```

```
> Analysis of Variance Table
>
> Response: dBDI
>           Df Sum Sq Mean Sq F value  Pr(>F)
> Setting    1      1     1.0    0.17    0.68
> Variant    1    196   196.0   33.00 7.5e-07 ***
> Residuals 45    267     5.9
> ---
> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Evaluation der Haupteffekte im additiven Modell der 2 x 2 ZVA mit Referenzgruppe

(1) Statistische Modelle und Teststatistiken

(2) Testhypothesen und Tests

(3) Testumfangkontrollen und p-Werte

Evaluation der Interaktion im Modell der 2 x 2 ZVA mit Interaktion und Referenzgruppe

(1) Statistisches Modell und Teststatistik

(2) Testhypothese und Test

(3) Testumfangkontrolle und p-Wert

Theorem (Teststatistik für die Interaktion)

Gegeben sei die Designmatrixform des additiven Modells der 2 x 2 ZVA mit Interaktion Referenzgruppe

$$y = X\beta + \varepsilon \text{ mit } \varepsilon \sim N(0_n, \sigma^2 I_n)\,, \tag{45}$$

wobei die Spalten von X bezeichnet seien durch

$$X := \begin{pmatrix} X_{\mu_0} & X_{\alpha_2} & X_{\beta_2} & X_{\gamma_{22}} \end{pmatrix} \in \mathbb{R}^{n \times 4}\,. \tag{46}$$

Dann gilt: Eine *F-Teststatistik für die Interaktion von Faktor A und B* wird als $F_{A \times B}$ bezeichnet und ist die F-Statistik unter der Partitionierung

$$\begin{gathered} X := \begin{pmatrix} X_{\mu_0} & X_{\alpha_2} & X_{\beta_2} & X_{\gamma_{22}} \end{pmatrix},\ \beta := \begin{pmatrix} \mu_0 \\ \alpha_2 \\ \beta_2 \\ \gamma_{22} \end{pmatrix},\ p := 4 \quad \text{und} \\ X_0 := \begin{pmatrix} X_{\mu_0} & X_{\alpha_2} & X_{\beta_2} \end{pmatrix},\ \beta_0 := \begin{pmatrix} \mu_0 \\ \alpha_2 \\ \beta_2 \end{pmatrix},\ p_0 := 3\,. \end{gathered} \tag{47}$$

Bemerkung

- Der Beweis ergibt sich aus dem Theorem zur Verteilung der F-Statistik bei Partitionierung eines ALM (siehe Einheit (8) in *Allgemeines Lineares Modell*).

Definition (Testhypothese und Test)

Die F-Teststatistik für die Interaktion von Faktor A und Faktor B sei mit $F_{A\times B}$ bezeichnet und wie oben definiert. Dann definiert der kritische-Wert-basierte Test

$$\phi_{A\times B}(y) := 1_{\{F_{A\times B}\geq k\}} \text{ mit Nullhypothese } H_0^{A\times B} : \gamma_{22} = 0 \tag{48}$$

den *F-Test der Interaktion von Faktor A und Faktor B*.

Theorem (Testumfangkontrolle und p-Wert)

Mit der obigen Definition und der KVF $\varphi(\cdot; n_1, n_2)$ der f-Verteilung gilt, dass $\phi_{A\times B}$ ein Level-α_0-Test mit Testumfang α_0 ist, wenn der kritische Wert definiert ist durch

$$k_{\alpha_0}^{A\times B} := \varphi^{-1}(1-\alpha_0; 1, n-4). \tag{49}$$

Der zu einem beobachteten Wert $f_{A\times B}$ von $F_{A\times B}$ assoziierte p-Wert ist gegeben durch

$$\text{p-Wert} := 1 - \varphi(f_{A\times B}; 1, n-4). \tag{50}$$

Modellevaluation: Interaktion

```r
# Modellevaluation
I           = 2                                            # Anzahl Level Faktor A (Therapiesetting)
J           = 2                                            # Anzahl Level Faktor B (Therapievariante)
n_ij        = length(A1B1)                                 # balanciertes ZVA-Design
n           = I*J*n_ij                                     # Anzahl Datenpunkte
p           = 4                                            # Anzahl Parameter vollständiges Modell
y           = matrix(c(A1B1,A1B2,A2B1,A2B2), nrow = n)     # Datenvektor
D           = matrix(c(1,0,0,0,  1,0,1,0,  1,1,0,0,  1,1,1,1),
                     nrow = I*J, byrow=TRUE)               # prototypische Designmatrix
C           = matrix(rep(1,n_ij),nrow = n_ij)              # prototypischer Zellenvektor für balancierte Designs
X           = kronecker(D,C)                               # Kroneckerprodukt-Designmatrix
alpha_0     = 0.05                                         # Signifikanzlevel
# X         = X                                            # Designmatrix vollständiges Modell
X_0         = X[,-4]                                       # Designmatrix reduziertes Modell
p           = ncol(X)                                      # Anzahl Parameter vollständiges Modell
p_0         = ncol(X_0)                                    # Anzahl Parameter reduziertes Modell
p_1         = p - p_0                                      # Anzahl zusätzlicher Parameter im vollständigen Modell
beta_hat_0 = solve(t(X_0)%*%X_0)%*%t(X_0)%*%y              # Betaparameterschätzer reduziertes Modell
beta_hat    = solve(t(X) %*%X )%*%t(X) %*%y                # Betaparameterschätzer vollständiges Modell
eps_hat_0   = y-X_0%*%beta_hat_0                           # Residuenvektor reduziertes Modell
eps_hat     = y - X%*%beta_hat                             # Residuenvektor vollständiges Modell
eh0_eh0     = t(eps_hat_0) %*% eps_hat_0                   # residuelle QS reduziertes Modell
eh_eh       = t(eps_hat) %*% eps_hat                       # residuelle QS vollständiges Modell
sigsqr_hat = eh_eh/(n-p)                                   # Varianzparameterschätzer vollständiges Modell
f           = ((eh0_eh0-eh_eh)/p_1)/sigsqr_hat             # F-Statistik
k_alpha_0   = qf(1-alpha_0, p_1, n-p)                      # kritischer Wert
if(f >= k_alpha_0){phi = 1} else {phi = 0}                 # Test
p_val       = 1 - pf(f, p_1,n-p)                           # p-Wert
data.frame("f"=f, "k"=k_alpha_0, "phi"=phi, "p-Wert"=p_val, row.names=c("Setting x Variant"))
```

```
>                      f    k phi p.Wert
> Setting x Variant 1.01 4.06   0  0.319
```

Modellevaluation: Interaktion

Interaktion im Modell der 2 x 2 ZVA mit Interaktion mit R's aov() Funktion

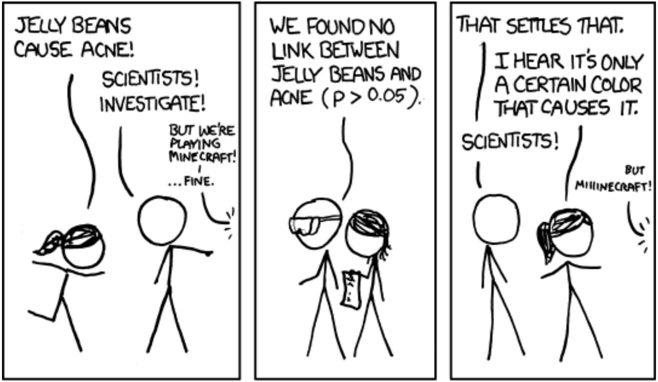
```
# Dateneinlesen
fname      = file.path(getwd(), "Daten", "Zweifaktorielle_Varianzanalyse_Daten.csv")
D          = read.table(fname, sep = ",", header = TRUE)

# R's aov Funktion
res.aov    = aov(dBDI ~ Setting + Variant + Setting:Variant, data = D)     # Modellformulierung und Modellschätzung
summary(res.aov)                                                          # Modellevaluation
```

```
>                 Df Sum Sq Mean Sq F value Pr(>F)
> Setting          1      1     1.0    0.17   0.68
> Variant          1    196   196.0   33.01  8e-07 ***
> Setting:Variant  1      6     6.0    1.01   0.32
> Residuals       44    261     5.9
> ---
> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Interaktion im Modell der 2 x 2 ZVA mit Interaktion mit R's lm() und anova() Funktionen

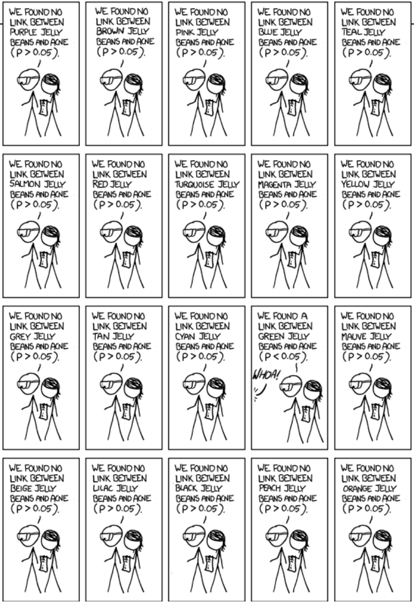
```
# Dateneinlesen
fname      = file.path(getwd(), "Daten", "Zweifaktorielle_Varianzanalyse_Daten.csv")
D          = read.table(fname, sep = ",", header = TRUE)

# R's lm() und anova() Funktionen
glm        = lm(dBDI ~ Setting + Variant + Setting:Variant, data = D)     # Modellformulierung und Modellschätzung
anova(glm)                                                               # Modellevaluation
```

```
> Analysis of Variance Table
>
> Response: dBDI
>                 Df Sum Sq Mean Sq F value Pr(>F)
> Setting          1      1     1.0    0.17   0.68
> Variant          1    196   196.0   33.01  8e-07 ***
> Setting:Variant  1      6     6.0    1.01   0.32
> Residuals       44    261     5.9
> ---
> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

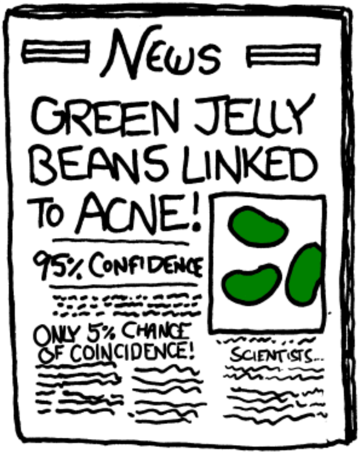


Webcomic zum p-Wert (Quelle: *xkcd*: "Significant" (#882); Lizenz: CC-BY-NC 2.5).

Modellevaluation



Webcomic zum p-Wert (Quelle: *xkcd*: "Significant" (#882); Lizenz: CC-BY-NC 2.5).



Webcomic zum p-Wert (Quelle: *xkcd*: “Significant” (#882); Lizenz: CC-BY-NC 2.5).

Anwendungsszenario

Modellformulierung

Modellschätzung

Modellevaluation

Selbstkontrollfragen

Selbstkontrollfragen

1. Erläutern Sie das Anwendungsszenario der zweifaktoriellen Varianzanalyse (ZVA).
2. Aus wie vielen Datenpunkten besteht ein Datensatz eines 3 x 4 ZVA-Designs mit 10 Datenpunkten pro Zelle?
3. Erläutern Sie die intuitive Bedeutung eines Haupteffektes in einem ZVA-Design.
4. Erläutern Sie die intuitive Bedeutung einer Interaktion ein einem ZVA-Design.
5. Geben Sie die Definition des additiven Modells der ZVA mit Referenzgruppe (RG) wieder.
6. Erläutern Sie die Bedeutung der Parameter μ_0, α_2 und β_2 im additiven Modell der ZVA mit RG.
7. Bestimmen Sie μ_{ij} für $\mu_0 := 2, \alpha_2 := -1, \beta_2 := 3$ im additiven Modell der ZVA mit RG.
8. Geben Sie die Designmatrixform des Modells einer additiven 2 x 2 ZVA mit RG für $n_{ij} := 1$ an.
9. Geben Sie die Designmatrixform des Modells einer additiven 2 x 2 ZVA mit RG für $n_{ij} := 3$ an.
10. Geben Sie die Definition des Modells der ZVA mit Interaktion und Referenzgruppe wieder.
11. Erläutern Sie die Bedeutung der Parameter μ_0, α_2, β_2 und γ_{22} im Modell der ZVA mit Interaktion und RG.
12. Bestimmen Sie μ_{ij} für $\mu_0 := 2, \alpha_2 := 1, \beta_2 := -1, \gamma_{22} := 3$ im Modell der ZVA mit Interaktion und RG.
13. Geben Sie die Designmatrixform des Modells einer 2 x 2 ZVA mit Interaktion und RG für $n_{ij} := 1$ an.
14. Geben Sie die Designmatrixform des Modells einer 2 x 2 ZVA mit Interaktion und RG für $n_{ij} := 3$ an.
15. Geben Sie das Theorem zur Betaparameterschätzung im additiven 2 x 2 ZVA-Modell mit RG wieder.
16. Geben Sie das Theorem zur Betaparameterschätzung im 2 x 2 ZVA-Modell mit Interaktion und RG wieder.