

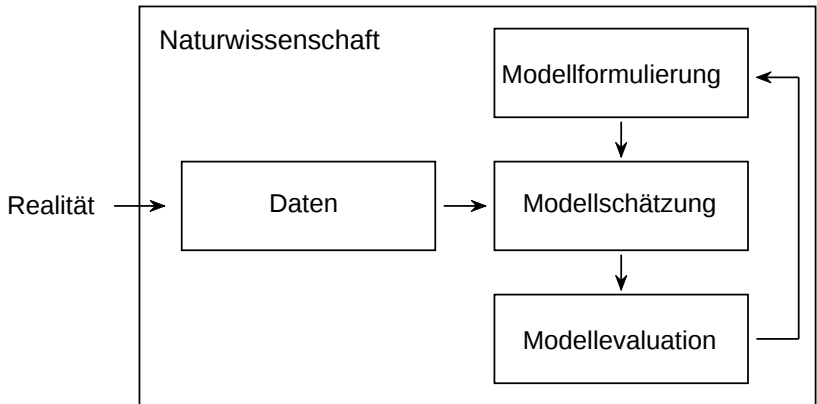


Inferenzstatistik

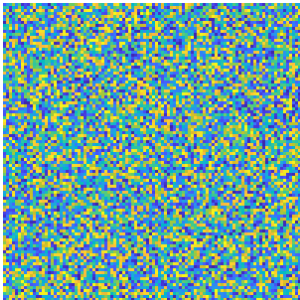
BSc Psychologie SoSe 2021

Prof. Dr. Dirk Ostwald

(12) Varianzanalysen

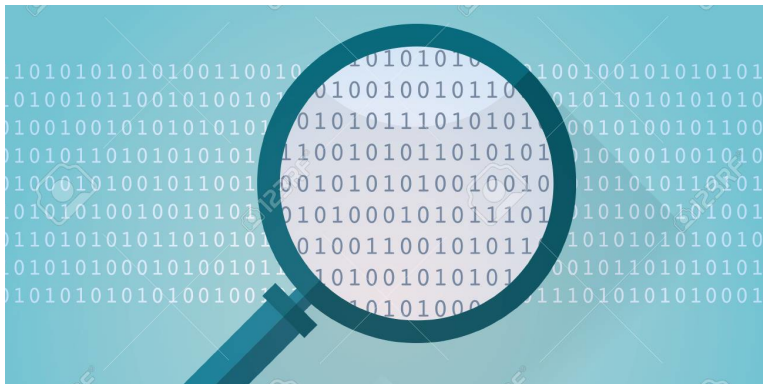


Rohdaten



Reduzierte Daten





Einfaktorielle Varianzanalyse

- Anwendungsszenario und Datendeskription
- Modellformulierung
- Modellschätzung
- Modellevaluation

Zweifaktorielle Varianzanalyse

- Anwendungsszenario und Datendeskription
- Modellformulierung
- Modellschätzung
- Modellevaluation

Selbstkontrollfragen

Die Darstellung basiert im Wesentlichen auf Kapitel 11 in DeGroot and Schervish (2012) und Kapitel 14 in Fahrmeir et al. (2016).

Einfaktorielle Varianzanalyse

- **Anwendungsszenario und Datendeskription**
- Modellformulierung
- Modellschätzung
- Modellevaluation

Zweifaktorielle Varianzanalyse

- Anwendungsszenario und Datendeskription
- Modellformulierung
- Modellschätzung
- Modellevaluation

Selbstkontrollfragen

Einfaktorielle Varianzanalyse = One-way Analysis of Variance (ANOVA)

Einfaktorielle Varianzanalyse = One-way Analysis of Variance (ANOVA)

Anwendungsszenario

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

Einfaktorielle Varianzanalyse = One-way Analysis of Variance (ANOVA)

Anwendungsszenario

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

Generalisierung des Zweistichproben-T-Tests bei unabhängigen Stichproben mit einfacher Nullhypothese für mehr als Zweistichproben

Einfaktorielle Varianzanalyse = One-way Analysis of Variance (ANOVA)

Anwendungsszenario

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

Generalisierung des Zweistichproben-T-Tests bei unabhängigen Stichproben mit einfacher Nullhypothese für mehr als Zweistichproben

Anwendungsbeispiele

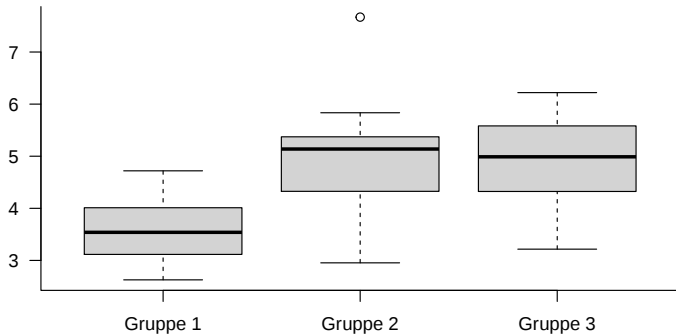
- BDI Score Analyse von drei Gruppen psychiatrischer Patient:innen nach PA, CBT, ST
 - Inferentielle Evidenz für Gruppenerwartungswertunterschiede?
 - Evidenz für unterschiedliche Therapiewirksamkeit?
- Tense Arousal Differenzwertanalyse der vier Filmclipbedingungen im Affect Datensatz
 - Inferentielle Evidenz für Gruppenerwartungswertunterschiede?
 - Evidenz für unterschiedliche Tense Arousal Induktion?

x_{ij} : Datenpunkt der j ten experimentellen in Einheit in der i ten Gruppe

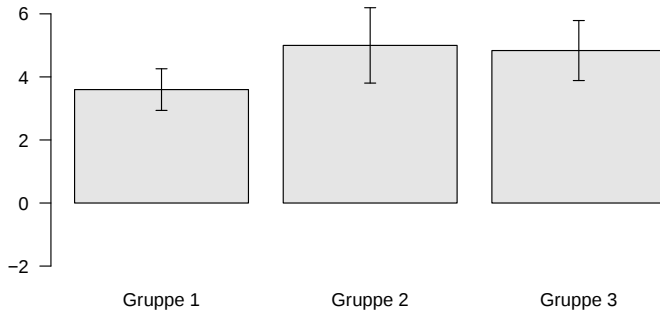
	x_{1j}	x_{2j}	x_{3j}
1	3.55	5.12	4.35
2	2.79	4.88	5.60
3	4.04	3.66	3.22
4	4.64	5.17	5.33
5	3.21	5.16	5.56
6	3.62	5.03	6.22
7	3.52	2.95	4.79
8	4.72	5.21	5.70
9	3.98	7.67	4.29
10	2.63	3.77	4.54
11	3.02	5.83	3.23
12	3.45	5.53	5.19

$i = 1, 2, 3$ mit $j = 1, \dots, 12$

Boxplot



Gruppenmittelwerte ($\pm$ Gruppenstandardabweichungen)



Einfaktorielle Varianzanalyse

- Anwendungsszenario und Datendeskription
- **Modellformulierung**
- Modellschätzung
- Modellevaluation

Zweifaktorielle Varianzanalyse

- Anwendungsszenario und Datendeskription
- Modellformulierung
- Modellschätzung
- Modellevaluation

Selbstkontrollfragen

Definition (Modell der einfaktoriellen Varianzanalyse)

Für $i = 1, \dots, p$ sei

$$X_{i1}, \dots, X_{im} \sim N(\mu_i, \sigma^2) \quad (1)$$

eine Stichprobe eines Normalverteilungsmodells von Größe m mit unbekanntem Erwartungswertparameter μ_i und unbekanntem Varianzparameter $\sigma^2 > 0$. Dann heißt (1) *Klassisches Modell der einfaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung*. Die Gesamtstichprobengröße bezeichnen wir mit $n := pm$.

Definition (Modell der einfaktoriellen Varianzanalyse)

Für $i = 1, \dots, p$ sei

$$X_{i1}, \dots, X_{im} \sim N(\mu_i, \sigma^2) \quad (1)$$

eine Stichprobe eines Normalverteilungsmodells von Größe m mit unbekanntem Erwartungswertparameter μ_i und unbekanntem Varianzparameter $\sigma^2 > 0$. Dann heißt (1) *Klassisches Modell der einfaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung*. Die Gesamtstichprobengröße bezeichnen wir mit $n := pm$. Äquivalent sei

$$X_{ij} = \mu_i + \varepsilon_{ij} \text{ mit } \varepsilon_{ij} \sim N(0, \sigma^2) \text{ für } i = 1, \dots, p \text{ und } j = 1, \dots, m \quad (2)$$

wobei

- i die Stichproben und j die experimentellen Einheiten indizieren,
- m die Stichprobengrößen und $n = pm$ die Gesamtstichprobengrößen sind,
- X_{ij} beobachtbare Zufallsvariablen sind,
- μ_i feste Erwartungswertparameter der Stichprobenvariablen sind, und
- ε_{ij} unabhängige normalverteilte nicht-beobachtbare Zufallsvariablen sind.

Dann heißt (2) *Generatives Modell der einfaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung*.

Definition (Modell der einfaktoriellen Varianzanalyse)

Für $i = 1, \dots, p$ sei

$$X_{i1}, \dots, X_{im} \sim N(\mu_i, \sigma^2) \quad (1)$$

eine Stichprobe eines Normalverteilungsmodells von Größe m mit unbekanntem Erwartungswertparameter μ_i und unbekanntem Varianzparameter $\sigma^2 > 0$. Dann heißt (1) *Klassisches Modell der einfaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung*. Die Gesamtstichprobengröße bezeichnen wir mit $n := pm$. Äquivalent sei

$$X_{ij} = \mu_i + \varepsilon_{ij} \text{ mit } \varepsilon_{ij} \sim N(0, \sigma^2) \text{ für } i = 1, \dots, p \text{ und } j = 1, \dots, m \quad (2)$$

wobei

- i die Stichproben und j die experimentellen Einheiten indizieren,
- m die Stichprobengrößen und $n = pm$ die Gesamtstichprobengrößen sind,
- X_{ij} beobachtbare Zufallsvariablen sind,
- μ_i feste Erwartungswertparameter der Stichprobenvariablen sind, und
- ε_{ij} unabhängige normalverteilte nicht-beobachtbare Zufallsvariablen sind.

Dann heißt (2) *Generatives Modell der einfaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung*.

Bemerkungen

- Der Einfachheit halber setzen wir hier die identische Stichprobengröße m voraus.

Definition (Modell der einfaktoriellen Varianzanalyse)

Für $i = 1, \dots, p$ sei

$$X_{i1}, \dots, X_{im} \sim N(\mu_i, \sigma^2) \quad (1)$$

eine Stichprobe eines Normalverteilungsmodells von Größe m mit unbekanntem Erwartungswertparameter μ_i und unbekanntem Varianzparameter $\sigma^2 > 0$. Dann heißt (1) *Klassisches Modell der einfaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung*. Die Gesamtstichprobengröße bezeichnen wir mit $n := pm$. Äquivalent sei

$$X_{ij} = \mu_i + \varepsilon_{ij} \text{ mit } \varepsilon_{ij} \sim N(0, \sigma^2) \text{ für } i = 1, \dots, p \text{ und } j = 1, \dots, m \quad (2)$$

wobei

- i die Stichproben und j die experimentellen Einheiten indizieren,
- m die Stichprobengrößen und $n = pm$ die Gesamtstichprobengrößen sind,
- X_{ij} beobachtbare Zufallsvariablen sind,
- μ_i feste Erwartungswertparameter der Stichprobenvariablen sind, und
- ε_{ij} unabhängige normalverteilte nicht-beobachtbare Zufallsvariablen sind.

Dann heißt (2) *Generatives Modell der einfaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung*.

Bemerkungen

- Der Einfachheit halber setzten wir hier die identische Stichprobengröße m voraus.
- Die Varianzparameter aller Stichproben werden als identisch vorausgesetzt.

Definition (Modell der einfaktoriellen Varianzanalyse)

Für $i = 1, \dots, p$ sei

$$X_{i1}, \dots, X_{im} \sim N(\mu_i, \sigma^2) \quad (1)$$

eine Stichprobe eines Normalverteilungsmodells von Größe m mit unbekanntem Erwartungswertparameter μ_i und unbekanntem Varianzparameter $\sigma^2 > 0$. Dann heißt (1) *Klassisches Modell der einfaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung*. Die Gesamtstichprobengröße bezeichnen wir mit $n := pm$. Äquivalent sei

$$X_{ij} = \mu_i + \varepsilon_{ij} \text{ mit } \varepsilon_{ij} \sim N(0, \sigma^2) \text{ für } i = 1, \dots, p \text{ und } j = 1, \dots, m \quad (2)$$

wobei

- i die Stichproben und j die experimentellen Einheiten indizieren,
- m die Stichprobengrößen und $n = pm$ die Gesamtstichprobengrößen sind,
- X_{ij} beobachtbare Zufallsvariablen sind,
- μ_i feste Erwartungswertparameter der Stichprobenvariablen sind, und
- ε_{ij} unabhängige normalverteilte nicht-beobachtbare Zufallsvariablen sind.

Dann heißt (2) *Generatives Modell der einfaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung*.

Bemerkungen

- Der Einfachheit halber setzten wir hier die identische Stichprobengröße m voraus.
- Die Varianzparameter aller Stichproben werden als identisch vorausgesetzt.
- Wir verzichten auf einen Beweise der Äquivalenz von (1) und (2).

Definition (Modell der einfaktoriellen Varianzanalyse)

Für $i = 1, \dots, p$ sei

$$X_{i1}, \dots, X_{im} \sim N(\mu_i, \sigma^2) \quad (1)$$

eine Stichprobe eines Normalverteilungsmodells von Größe m mit unbekanntem Erwartungswertparameter μ_i und unbekanntem Varianzparameter $\sigma^2 > 0$. Dann heißt (1) *Klassisches Modell der einfaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung*. Die Gesamtstichprobengröße bezeichnen wir mit $n := pm$. Äquivalent sei

$$X_{ij} = \mu_i + \varepsilon_{ij} \text{ mit } \varepsilon_{ij} \sim N(0, \sigma^2) \text{ für } i = 1, \dots, p \text{ und } j = 1, \dots, m \quad (2)$$

wobei

- i die Stichproben und j die experimentellen Einheiten indizieren,
- m die Stichprobengrößen und $n = pm$ die Gesamtstichprobengrößen sind,
- X_{ij} beobachtbare Zufallsvariablen sind,
- μ_i feste Erwartungswertparameter der Stichprobenvariablen sind, und
- ε_{ij} unabhängige normalverteilte nicht-beobachtbare Zufallsvariablen sind.

Dann heißt (2) *Generatives Modell der einfaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung*.

Bemerkungen

- Der Einfachheit halber setzten wir hier die identische Stichprobengröße m voraus.
- Die Varianzparameter aller Stichproben werden als identisch vorausgesetzt.
- Wir verzichten auf einen Beweise der Äquivalenz von (1) und (2).
- Es handelt sich um die Generalisierung des Zweistichproben-T-Test Modells für unabhängige Stichproben unter Annahme identischer Varianzparameter von $p = 2$ auf ein beliebiges $p \in \mathbb{N}$.

Definition (Modell der einfaktoriellen Varianzanalyse)

Für $i = 1, \dots, p$ sei

$$X_{i1}, \dots, X_{im} \sim N(\mu_i, \sigma^2) \quad (1)$$

eine Stichprobe eines Normalverteilungsmodells von Größe m mit unbekanntem Erwartungswertparameter μ_i und unbekanntem Varianzparameter $\sigma^2 > 0$. Dann heißt (1) *Klassisches Modell der einfaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung*. Die Gesamtstichprobengröße bezeichnen wir mit $n := pm$. Äquivalent sei

$$X_{ij} = \mu_i + \varepsilon_{ij} \text{ mit } \varepsilon_{ij} \sim N(0, \sigma^2) \text{ für } i = 1, \dots, p \text{ und } j = 1, \dots, m \quad (2)$$

wobei

- i die Stichproben und j die experimentellen Einheiten indizieren,
- m die Stichprobengrößen und $n = pm$ die Gesamtstichprobengrößen sind,
- X_{ij} beobachtbare Zufallsvariablen sind,
- μ_i feste Erwartungswertparameter der Stichprobenvariablen sind, und
- ε_{ij} unabhängige normalverteilte nicht-beobachtbare Zufallsvariablen sind.

Dann heißt (2) *Generatives Modell der einfaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung*.

Bemerkungen

- Der Einfachheit halber setzten wir hier die identische Stichprobengröße m voraus.
- Die Varianzparameter aller Stichproben werden als identisch vorausgesetzt.
- Wir verzichten auf einen Beweise der Äquivalenz von (1) und (2).
- Es handelt sich um die Generalisierung des Zweistichproben-T-Test Modells für unabhängige Stichproben unter Annahme identischer Varianzparameter von $p = 2$ auf ein beliebiges $p \in \mathbb{N}$.
- Die Gesamtheit aller Stichprobenvariablen bezeichnen wir mit $X := \{X_{ij}\}_{i=1, \dots, p, j=1, \dots, m}$.

Theorem (Effektdarstellung des einfaktoriellen Varianzanalysemodells)

Es sei das generative Modell der einfaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung wie oben definiert gegeben. Dann können die beobachtbaren Zufallsvariablen äquivalent in der Form

$$X_{ij} = \mu + \alpha_i + \varepsilon_{ij} \text{ mit } \varepsilon_{ij} \sim N(0, \sigma^2) \text{ für } i = 1, \dots, p, j = 1, \dots, m, n = pm \quad (3)$$

mit

$$\alpha_i := \mu_i - \mu, \quad \mu := \frac{1}{p} \sum_{i=1}^p \mu_i \text{ und } \sum_{i=1}^p \alpha_i = 0 \quad (4)$$

geschrieben werden. (3) und (4) heißen *Generatives Modell der einfaktoriellen Varianzanalyse in Effektdarstellung*.

Theorem (Effektdarstellung des einfaktoriellen Varianzanalysemodells)

Es sei das generative Modell der einfaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung wie oben definiert gegeben. Dann können die beobachtbaren Zufallsvariablen äquivalent in der Form

$$X_{ij} = \mu + \alpha_i + \varepsilon_{ij} \text{ mit } \varepsilon_{ij} \sim N(0, \sigma^2) \text{ für } i = 1, \dots, p, j = 1, \dots, m, n = pm \quad (3)$$

mit

$$\alpha_i := \mu_i - \mu, \quad \mu := \frac{1}{p} \sum_{i=1}^p \mu_i \text{ und } \sum_{i=1}^p \alpha_i = 0 \quad (4)$$

geschrieben werden. (3) und (4) heißen *Generatives Modell der einfaktoriellen Varianzanalyse in Effektdarstellung*.

Beweis

Für ein beliebig gewähltes μ gilt

$$\mu_i = \mu + (\mu_i - \mu) =: \mu + \alpha_i \quad (5)$$

Aus der Definition von μ in obigem Theorem und mit $\mu_i = \alpha_i + \mu$ folgt dann außerdem

$$\mu = \frac{1}{p} \sum_{i=1}^p \mu_i = \frac{1}{p} \sum_{i=1}^p (\mu + \alpha_i) = \mu + \frac{1}{p} \sum_{i=1}^p \alpha_i \Leftrightarrow \sum_{i=1}^p \alpha_i = 0. \quad (6)$$

□

Theorem (Effektdarstellung des einfaktoriellen Varianzanalysemodells)

Es sei das generative Modell der einfaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung wie oben definiert gegeben. Dann können die beobachtbaren Zufallsvariablen äquivalent in der Form

$$X_{ij} = \mu + \alpha_i + \varepsilon_{ij} \text{ mit } \varepsilon_{ij} \sim N(0, \sigma^2) \text{ für } i = 1, \dots, p, j = 1, \dots, m, n = pm \quad (3)$$

mit

$$\alpha_i := \mu_i - \mu, \quad \mu := \frac{1}{p} \sum_{i=1}^p \mu_i \text{ und } \sum_{i=1}^p \alpha_i = 0 \quad (4)$$

geschrieben werden. (3) und (4) heißen *Generatives Modell der einfaktoriellen Varianzanalyse in Effektdarstellung*.

Beweis

Für ein beliebig gewähltes μ gilt

$$\mu_i = \mu + (\mu_i - \mu) =: \mu + \alpha_i \quad (5)$$

Aus der Definition von μ in obigem Theorem und mit $\mu_i = \alpha_i + \mu$ folgt dann außerdem

$$\mu = \frac{1}{p} \sum_{i=1}^p \mu_i = \frac{1}{p} \sum_{i=1}^p (\mu + \alpha_i) = \mu + \frac{1}{p} \sum_{i=1}^p \alpha_i \Leftrightarrow \sum_{i=1}^p \alpha_i = 0. \quad (6)$$

□

Bemerkungen

- α_i heißt *i*ter Gruppeneffektparameter oder Effektparameter des *i*ten Faktorlevels.
- Die Effektdarstellung betont die Analogie zum zweifaktoriellen Varianzanalysemodell.

Einfaktorielle Varianzanalyse

- Anwendungsszenario und Datendeskription
- Modellformulierung
- **Modellschätzung**
- Modellevaluation

Zweifaktorielle Varianzanalyse

- Anwendungsszenario und Datendeskription
- Modellformulierung
- Modellschätzung
- Modellevaluation

Selbstkontrollfragen

Theorem (Parameterschätzer des einfaktoriellen Varianzanalysemodells)

$\mu_1, \dots, \mu_p$ und $\sigma^2 > 0$ seien die Parameter des einfaktoriellen Varianzanalysemodells wie oben definiert. Dann sind die Maximum Likelihood Schätzer von $\mu_1, \dots, \mu_p$ und $\sigma^2 > 0$ gegeben durch

$$\hat{\mu}_i := \frac{1}{m} \sum_{j=1}^m X_{ij} \text{ und } \hat{\sigma}_{\text{ML}}^2 := \frac{1}{pm} \sum_{i=1}^p \sum_{j=1}^m (X_{ij} - \bar{X}_i)^2, \quad (7)$$

respektive, wobei die Erwartungswertparameterschätzer unverzerzte Schätzer sind und der Varianzschätzer ein verzerrter Schätzer ist.

Theorem (Parameterschätzer des einfaktoriellen Varianzanalysemodells)

$\mu_1, \dots, \mu_p$ und $\sigma^2 > 0$ seien die Parameter des einfaktoriellen Varianzanalysemodells wie oben definiert. Dann sind die Maximum Likelihood Schätzer von $\mu_1, \dots, \mu_p$ und $\sigma^2 > 0$ gegeben durch

$$\hat{\mu}_i := \frac{1}{m} \sum_{j=1}^m X_{ij} \text{ und } \hat{\sigma}_{\text{ML}}^2 := \frac{1}{pm} \sum_{i=1}^p \sum_{j=1}^m (X_{ij} - \bar{X}_i)^2, \quad (7)$$

respektive, wobei die Erwartungswertparameterschätzer unverzerrte Schätzer sind und der Varianzschätzer ein verzerrter Schätzer ist. Ein unverzerrter Schätzer von $\sigma^2 > 0$ ist gegeben durch

$$\hat{\sigma}^2 := \frac{1}{n-p} \sum_{i=1}^p \sum_{j=1}^m (X_{ij} - \bar{X}_i)^2. \quad (8)$$

Theorem (Parameterschätzer des einfaktoriellen Varianzanalysemodells)

$\mu_1, \dots, \mu_p$ und $\sigma^2 > 0$ seien die Parameter des einfaktoriellen Varianzanalysemodells wie oben definiert. Dann sind die Maximum Likelihood Schätzer von $\mu_1, \dots, \mu_p$ und $\sigma^2 > 0$ gegeben durch

$$\hat{\mu}_i := \frac{1}{m} \sum_{j=1}^m X_{ij} \text{ und } \hat{\sigma}_{\text{ML}}^2 := \frac{1}{pm} \sum_{i=1}^p \sum_{j=1}^m (X_{ij} - \bar{X}_i)^2, \quad (7)$$

respektive, wobei die Erwartungswertparameterschätzer unverzerrte Schätzer sind und der Varianzschätzer ein verzerrter Schätzer ist. Ein unverzerrter Schätzer von $\sigma^2 > 0$ ist gegeben durch

$$\hat{\sigma}^2 := \frac{1}{n - p} \sum_{i=1}^p \sum_{j=1}^m (X_{ij} - \bar{X}_i)^2. \quad (8)$$

Bemerkungen

- Wir verzichten auf einen Beweis.
- Die Stichprobenmittelwerte sind unverzerrte Schätzer der wahren, aber unbekannten, Stichprobenerwartungswertparameter, die Stichprobenvarianz der Gesamtstichprobe ist ein unverzerrter Schätzer des Varianzparameters.

Einfaktorielle Varianzanalyse

- Anwendungsszenario und Datendeskription
- Modellformulierung
- Modellschätzung
- **Modellevaluation**

Zweifaktorielle Varianzanalyse

- Anwendungsszenario und Datendeskription
- Modellformulierung
- Modellschätzung
- Modellevaluation

Selbstkontrollfragen

Hypothesen und Überblick

- Primäres Ziel der einfaktoriellen Varianzanalyse ist das Testen der Nullhypothese

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_p \quad (9)$$

bzw.

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_p = 0. \quad (10)$$

Hypothesen und Überblick

- Primäres Ziel der einfaktoriellen Varianzanalyse ist das Testen der Nullhypothese

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_p \quad (9)$$

bzw.

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_p = 0. \quad (10)$$

- Die Nullhypothese besagt also, dass zwischen den Gruppen keine Erwartungswertparameterunterschiede vorliegen bzw. dass die Gruppeneffektparameter identisch Null sind.

Hypothesen und Überblick

- Primäres Ziel der einfaktoriellen Varianzanalyse ist das Testen der Nullhypothese

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_p \quad (9)$$

bzw.

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_p = 0. \quad (10)$$

- Die Nullhypothese besagt also, dass zwischen den Gruppen keine Erwartungswertparameterunterschiede vorliegen bzw. dass die Gruppeneffektparameter identisch Null sind.
- Die Alternativhypothese lautet somit

$$H_1 : \mu_i \neq \mu_j \text{ für mindestens ein Paar } (i, j) \text{ mit } i \neq j \quad (11)$$

bzw.

$$H_1 : \alpha_i \neq 0 \text{ für mindestens ein } i \in \mathbb{N}_p. \quad (12)$$

Hypothesen und Überblick

- Primäres Ziel der einfaktoriellen Varianzanalyse ist das Testen der Nullhypothese

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_p \quad (9)$$

bzw.

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_p = 0. \quad (10)$$

- Die Nullhypothese besagt also, dass zwischen den Gruppen keine Erwartungswertparameterunterschiede vorliegen bzw. dass die Gruppeneffektparameter identisch Null sind.
- Die Alternativhypothese lautet somit

$$H_1 : \mu_i \neq \mu_j \text{ für mindestens ein Paar } (i, j) \text{ mit } i \neq j \quad (11)$$

bzw.

$$H_1 : \alpha_i \neq 0 \text{ für mindestens ein } i \in \mathbb{N}_p. \quad (12)$$

- Die Alternativhypothese besagt also, dass sich mindestens zwei Erwartungswertparameter unterscheiden bzw. mindestens ein Gruppeneffektparameter ungleich Null ist.

Hypothesen und Überblick

- Primäres Ziel der einfaktoriellen Varianzanalyse ist das Testen der Nullhypothese

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_p \quad (9)$$

bzw.

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_p = 0. \quad (10)$$

- Die Nullhypothese besagt also, dass zwischen den Gruppen keine Erwartungswertparameterunterschiede vorliegen bzw. dass die Gruppeneffektparameter identisch Null sind.
- Die Alternativhypothese lautet somit

$$H_1 : \mu_i \neq \mu_j \text{ für mindestens ein Paar } (i, j) \text{ mit } i \neq j \quad (11)$$

bzw.

$$H_1 : \alpha_i \neq 0 \text{ für mindestens ein } i \in \mathbb{N}_p. \quad (12)$$

- Die Alternativhypothese besagt also, dass sich mindestens zwei Erwartungswertparameter unterscheiden bzw. mindestens ein Gruppeneffektparameter ungleich Null ist.
- Ein kritischer Wert-basierter Test dieser Hypothesen kann mithilfe der *F-Teststatistik* und ihrer Verteilung konstruiert werden. Die F-Teststatistik wiederum beruht auf einer *Quadratsummenzerlegung*. Wir führen beide Konzepte im Folgenden ein.

Theorem (Quadratsummenzerlegung bei einfaktorieller Varianzanalyse)

Für $i = 1, \dots, p$ und $j = 1, \dots, m$ bezeichne X_{ij} die j te Stichprobenvariable der i ten Stichproben-
gruppe eines einfaktoriellen Varianzanalysemodells. Weiterhin seien

$$\bar{X} := \frac{1}{n} \sum_{i=1}^p \sum_{j=1}^m X_{ij} \text{ und } \bar{X}_i := \frac{1}{m} \sum_{j=1}^m X_{ij} \quad (13)$$

den *Gesamtmittelwert* (*grand mean*) und das i te Stichprobenmittel, respektive.

Theorem (Quadratsummenzerlegung bei einfaktorieller Varianzanalyse)

Für $i = 1, \dots, p$ und $j = 1, \dots, m$ bezeichne X_{ij} die j te Stichprobenvariable der i ten Stichproben-
gruppe eines einfaktoriellen Varianzanalysemodells. Weiterhin seien

$$\bar{X} := \frac{1}{n} \sum_{i=1}^p \sum_{j=1}^m X_{ij} \text{ und } \bar{X}_i := \frac{1}{m} \sum_{j=1}^m X_{ij} \quad (13)$$

den *Gesamtmittelwert* (*grand mean*) und das i te Stichprobenmittel, respektive. Schließlich seien

$$\text{SQT} := \sum_{i=1}^p \sum_{j=1}^m (X_{ij} - \bar{X})^2 \quad \text{die Total Sum of Squares}$$

$$\text{SQB} := \sum_{i=1}^p m(\bar{X}_i - \bar{X})^2 \quad \text{die Between-Group Sum of Squares}$$

$$\text{SQW} := \sum_{i=1}^p \sum_{j=1}^m (X_{ij} - \bar{X}_i)^2 \quad \text{die Within-Group Sum of Squares}$$

Theorem (Quadratsummenzerlegung bei einfaktorieller Varianzanalyse)

Für $i = 1, \dots, p$ und $j = 1, \dots, m$ bezeichne X_{ij} die j te Stichprobenvariable der i ten Stichprobengruppe eines einfaktoriellen Varianzanalysemodells. Weiterhin seien

$$\bar{X} := \frac{1}{n} \sum_{i=1}^p \sum_{j=1}^m X_{ij} \text{ und } \bar{X}_i := \frac{1}{m} \sum_{j=1}^m X_{ij} \quad (13)$$

den *Gesamtmittelwert* (*grand mean*) und das i te Stichprobenmittel, respektive. Schließlich seien

$$\text{SQT} := \sum_{i=1}^p \sum_{j=1}^m (X_{ij} - \bar{X})^2 \quad \text{die Total Sum of Squares}$$

$$\text{SQB} := \sum_{i=1}^p m(\bar{X}_i - \bar{X})^2 \quad \text{die Between-Group Sum of Squares}$$

$$\text{SQW} := \sum_{i=1}^p \sum_{j=1}^m (X_{ij} - \bar{X}_i)^2 \quad \text{die Within-Group Sum of Squares}$$

Dann gilt

$$\text{SQT} = \text{SQW} + \text{SQB}. \quad (14)$$

Bemerkungen

- SQT repräsentiert die totale Variabilität der Daten um den Gesamtmittelwert.

Bemerkungen

- SQT repräsentiert die totale Variabilität der Daten um den Gesamtmittelwert.
- SQB repräsentiert die Variabilität der Gruppenmittelwerte um den Gesamtmittelwert.

Bemerkungen

- SQT repräsentiert die totale Variabilität der Daten um den Gesamtmittelwert.
- SQB repräsentiert die Variabilität der Gruppenmittelwerte um den Gesamtmittelwert.
- SQW repräsentiert die Variabilität der Daten um ihre jeweiligen Gruppenmittelwerte.

Bemerkungen

- SQT repräsentiert die totale Variabilität der Daten um den Gesamtmittelwert.
- SQB repräsentiert die Variabilität der Gruppenmittelwerte um den Gesamtmittelwert.
- SQW repräsentiert die Variabilität der Daten um ihre jeweiligen Gruppenmittelwerte.
- Die totale Variabilität wird hier also in zwei unabhängige Beiträge von Variabilität zerlegt.
- SQW wird auch *Residualvariabilität* oder *SQR* genannt, weil sie die verbleibende Variabilität nach Schätzung der Gruppenerwartungswertparameter quantifiziert und gilt, dass

$$(n - p)\hat{\sigma}^2 = \text{SQW} = \text{SQR}. \quad (15)$$

Bemerkungen

- SQT repräsentiert die totale Variabilität der Daten um den Gesamtmittelwert.
- SQB repräsentiert die Variabilität der Gruppenmittelwerte um den Gesamtmittelwert.
- SQW repräsentiert die Variabilität der Daten um ihre jeweiligen Gruppenmittelwerte.
- Die totale Variabilität wird hier also in zwei unabhängige Beiträge von Variabilität zerlegt.
- SQW wird auch *Residualvariabilität* oder *SQR* genannt, weil sie die verbleibende Variabilität nach Schätzung der Gruppenerwartungswertparameter quantifiziert und gilt, dass

$$(n - p)\hat{\sigma}^2 = \text{SQW} = \text{SQR}. \quad (15)$$

- Die Quadratsummenzerlegung ergibt sich aus algebraischen Identitäten wie unten gezeigt.

Beweis

$$\begin{aligned}\text{SQT} &= \sum_{i=1}^p \sum_{j=1}^m (X_{ij} - \bar{X})^2 \\&= \sum_{i=1}^p \sum_{j=1}^m (X_{ij} - \bar{X}_i + \bar{X}_i - \bar{X})^2 \\&= \sum_{i=1}^p \sum_{j=1}^m \left((X_{ij} - \bar{X}_i) + (\bar{X}_i - \bar{X}) \right)^2 \\&= \sum_{i=1}^p \sum_{j=1}^m \left((X_{ij} - \bar{X}_i)^2 + 2(X_{ij} - \bar{X}_i)(\bar{X}_i - \bar{X}) + (\bar{X}_i - \bar{X})^2 \right) \\&= \sum_{i=1}^p \left(\sum_{j=1}^m (X_{ij} - \bar{X}_i)^2 + \sum_{j=1}^m 2(X_{ij} - \bar{X}_i)(\bar{X}_i - \bar{X}) + \sum_{j=1}^m (\bar{X}_i - \bar{X})^2 \right) \\&= \sum_{i=1}^p \left(\sum_{j=1}^m (X_{ij} - \bar{X}_i)^2 + 2(\bar{X}_i - \bar{X}) \sum_{j=1}^m (X_{ij} - \bar{X}_i) + m(\bar{X}_i - \bar{X})^2 \right)\end{aligned}$$

Beweis (fortgeführt)

$$\begin{aligned} &= \sum_{i=1}^p \left(\sum_{j=1}^m (X_{ij} - \bar{X}_i)^2 + 2(\bar{X}_i - \bar{X}) \sum_{j=1}^m \left(X_{ij} - \frac{1}{m} \sum_{j=1}^m X_{ij} \right) + m(\bar{X}_i - \bar{X})^2 \right) \\ &= \sum_{i=1}^p \left(\sum_{j=1}^m (X_{ij} - \bar{X}_i)^2 + 2(\bar{X}_i - \bar{X}) \left(\sum_{j=1}^m X_{ij} - \sum_{j=1}^m \left(\frac{1}{m} \sum_{j=1}^m X_{ij} \right) \right) + m(\bar{X}_i - \bar{X})^2 \right) \\ &= \sum_{i=1}^p \left(\sum_{j=1}^m (X_{ij} - \bar{X}_i)^2 + 2(\bar{X}_i - \bar{X}) \left(\sum_{j=1}^m X_{ij} - \frac{m}{m} \sum_{j=1}^m X_{ij} \right) + m(\bar{X}_i - \bar{X})^2 \right) \\ &= \sum_{i=1}^p \left(\sum_{j=1}^m (X_{ij} - \bar{X}_i)^2 + 2(\bar{X}_i - \bar{X}) \left(\sum_{j=1}^m X_{ij} - \sum_{j=1}^m X_{ij} \right) + m(\bar{X}_i - \bar{X})^2 \right) \\ &= \sum_{i=1}^p \left(\sum_{j=1}^m (X_{ij} - \bar{X}_i)^2 + m(\bar{X}_i - \bar{X})^2 \right) \\ &= \sum_{i=1}^p \sum_{j=1}^m (X_{ij} - \bar{X}_i)^2 + \sum_{i=1}^p m(\bar{X}_i - \bar{X})^2 \\ &= \text{SQW} + \text{SQB}. \end{aligned}$$

Definition (F-Teststatistik)

Es seien das Modell der einfaktoriellen Varianzanalyse sowie die Between-Group Sum of Squares SQB und die Within-Group Sum of Squares SQW definiert wie oben. Dann ist die *F-Teststatistik* definiert als

$$F := \frac{SQB/(p-1)}{SQW/(n-p)} \quad (16)$$

Der Zähler und der Nenner dieser Definition werden auch als *Mean SQB* und *Mean SQW*, respektive, bezeichnet.

Definition (F-Teststatistik)

Es seien das Modell der einfaktoriellen Varianzanalyse sowie die Between-Group Sum of Squares SQB und die Within-Group Sum of Squares SQW definiert wie oben. Dann ist die *F-Teststatistik* definiert als

$$F := \frac{SQB/(p-1)}{SQW/(n-p)} \quad (16)$$

Der Zähler und der Nenner dieser Definition werden auch als *Mean SQB* und *Mean SQW*, respektive, bezeichnet.

Bemerkungen

- Die F-Teststatistik vergleicht die Variabilität zwischen Gruppen mit der Variabilität innerhalb von Gruppen.

Definition (F-Teststatistik)

Es seien das Modell der einfaktoriellen Varianzanalyse sowie die Between-Group Sum of Squares SQB und die Within-Group Sum of Squares SQW definiert wie oben. Dann ist die *F-Teststatistik* definiert als

$$F := \frac{SQB/(p-1)}{SQW/(n-p)} \quad (16)$$

Der Zähler und der Nenner dieser Definition werden auch als *Mean SQB* und *Mean SQW*, respektive, bezeichnet.

Bemerkungen

- Die F-Teststatistik vergleicht die Variabilität zwischen Gruppen mit der Variabilität innerhalb von Gruppen.
- Große Variabilität der Gruppenmittelwerte um den Gesamtmittelwert bei gleichzeitiger kleiner Residualvariabilität resultiert in hohen *F*-Werten $\Rightarrow$ Bei Nichtzutreffen der Nullhypothese erwartet man also hohe *F*-Werte; bei Zutreffen der Nullhypothese eher geringe *F*-Werte.

Definition (F-Teststatistik)

Es seien das Modell der einfaktoriellen Varianzanalyse sowie die Between-Group Sum of Squares SQB und die Within-Group Sum of Squares SQW definiert wie oben. Dann ist die *F-Teststatistik* definiert als

$$F := \frac{SQB/(p-1)}{SQW/(n-p)} \quad (16)$$

Der Zähler und der Nenner dieser Definition werden auch als *Mean SQB* und *Mean SQW*, respektive, bezeichnet.

Bemerkungen

- Die F-Teststatistik vergleicht die Variabilität zwischen Gruppen mit der Variabilität innerhalb von Gruppen.
- Große Variabilität der Gruppenmittelwerte um den Gesamtmittelwert bei gleichzeitiger kleiner Residualvariabilität resultiert in hohen F -Werten $\Rightarrow$ Bei Nichtzutreffen der Nullhypothese erwartet man also hohe F -Werte; bei Zutreffen der Nullhypothese eher geringe F -Werte.
- Die F-Teststatistik misst die Größe der SQB pro zusätzlicher Gruppe in Einheiten der geschätzten Varianz $\hat{\sigma}^2$.

Definition (F-Teststatistik)

Es seien das Modell der einfaktoriellen Varianzanalyse sowie die Between-Group Sum of Squares SQB und die Within-Group Sum of Squares SQW definiert wie oben. Dann ist die *F-Teststatistik* definiert als

$$F := \frac{SQB/(p-1)}{SQW/(n-p)} \quad (16)$$

Der Zähler und der Nenner dieser Definition werden auch als *Mean SQB* und *Mean SQW*, respektive, bezeichnet.

Bemerkungen

- Die F-Teststatistik vergleicht die Variabilität zwischen Gruppen mit der Variabilität innerhalb von Gruppen.
- Große Variabilität der Gruppenmittelwerte um den Gesamtmittelwert bei gleichzeitiger kleiner Residualvariabilität resultiert in hohen F -Werten $\Rightarrow$ Bei Nichtzutreffen der Nullhypothese erwartet man also hohe F -Werte; bei Zutreffen der Nullhypothese eher geringe F -Werte.
- Die F-Teststatistik misst die Größe der SQB pro zusätzlicher Gruppe in Einheiten der geschätzten Varianz $\hat{\sigma}^2$.
- Ein F -Werte ≤ 1 besagt, dass die SQB pro zusätzlicher Gruppe kleiner bzw. gleich der Residualvarianz ist.

Definition (F-Teststatistik)

Es seien das Modell der einfaktoriellen Varianzanalyse sowie die Between-Group Sum of Squares SQB und die Within-Group Sum of Squares SQW definiert wie oben. Dann ist die *F-Teststatistik* definiert als

$$F := \frac{SQB/(p-1)}{SQW/(n-p)} \quad (16)$$

Der Zähler und der Nenner dieser Definition werden auch als *Mean SQB* und *Mean SQW*, respektive, bezeichnet.

Bemerkungen

- Die F-Teststatistik vergleicht die Variabilität zwischen Gruppen mit der Variabilität innerhalb von Gruppen.
- Große Variabilität der Gruppenmittelwerte um den Gesamtmittelwert bei gleichzeitiger kleiner Residualvariabilität resultiert in hohen F -Werten $\Rightarrow$ Bei Nichtzutreffen der Nullhypothese erwartet man also hohe F -Werte; bei Zutreffen der Nullhypothese eher geringe F -Werte.
- Die F-Teststatistik misst die Größe der SQB pro zusätzlicher Gruppe in Einheiten der geschätzten Varianz $\hat{\sigma}^2$.
- Ein F -Werte ≤ 1 besagt, dass die SQB pro zusätzlicher Gruppe kleiner bzw. gleich der Residualvarianz ist.
- Ohne Beweis halten wir fest, dass für $p = 2$ gilt, dass $F = T^2$, wobei T die T-Teststatistik bezeichnet.

Theorem (Verteilung der F-Teststatistik bei H_0)

Es seien das Modell der einfaktoriellen Varianzanalyse sowie die F-Teststatistik F wie oben definiert und es gelte zudem

$$H_0 : \mu_1 = \dots = \mu_p \text{ bzw. } H_0 : \alpha_1 = \dots = \alpha_p = 0. \quad (17)$$

Dann ist die F-Teststatistik ist nach einer f -Verteilung mit Freiheitsgradparametern $p - 1$ und $n - p$ verteilt,

$$F \sim f(p - 1, n - p). \quad (18)$$

Theorem (Verteilung der F-Teststatistik bei H_0)

Es seien das Modell der einfaktoriellen Varianzanalyse sowie die F-Teststatistik F wie oben definiert und es gelte zudem

$$H_0 : \mu_1 = \dots = \mu_p \text{ bzw. } H_0 : \alpha_1 = \dots = \alpha_p = 0. \quad (17)$$

Dann ist die F-Teststatistik ist nach einer f -Verteilung mit Freiheitsgradparametern $p - 1$ und $n - p$ verteilt,

$$F \sim f(p - 1, n - p). \quad (18)$$

Bemerkungen

- Wir verzichten auf einen Beweis.
- Trifft H_0 nicht zu, so ist die F-Teststatistik analog zur T-Teststatistik nach einer *nichtzentralen f-Verteilung* verteilt. Wir gehen auf diesen Fall nicht genauer ein, d.h. wir verzichten hier auf eine Analyse der Testgütefunktion.

Zur Erinnerung an (6) Transformationen der Normalverteilung

Definition (f -Zufallsvariable)

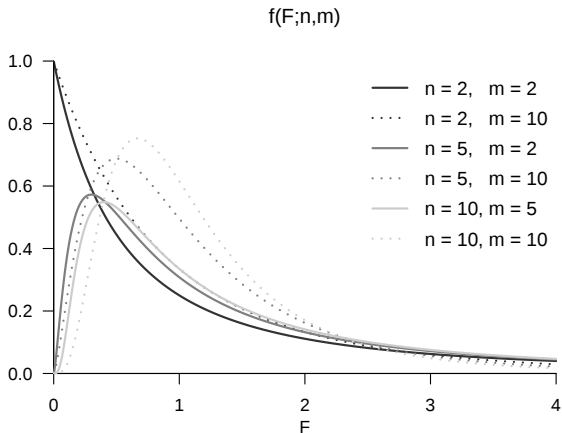
F sei eine Zufallsvariable mit Ergebnisraum $\mathbb{R}_{>0}$ und WDF

$$p_F : \mathbb{R} \rightarrow \mathbb{R}_{>0}, f \mapsto p_F(f) := m^{\frac{m}{2}} n^{\frac{n}{2}} \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right) \Gamma\left(\frac{n}{2}\right)} \frac{f^{\frac{m}{2}-1}}{\left(1 + \frac{m}{n} f\right)^{\frac{m+n}{2}}}, \quad (19)$$

wobei Γ die Gammafunktion bezeichne. Dann sagen wir, dass F einer f -Verteilung mit Freiheitsgradparametern n und m unterliegt und nennen F eine f -Zufallsvariable mit Freiheitsgradparametern n, m . Wir kürzen dies mit $F \sim f(n, m)$ ab. Die WDF einer f -Zufallsvariable mit Freiheitsgradparametern n, m bezeichnen wir mit $f(F; n, m)$, die KVF einer f -Zufallsvariable mit Freiheitsgradparametern n, m bezeichnen wir mit $\varphi(F; n, m)$, und die inverse KVF einer f -Zufallsvariable mit Freiheitsgradparametern n, m bezeichnen wir mit $\varphi^{-1}(F; n, m)$.

Zur Erinnerung an (6) Transformationen der Normalverteilung

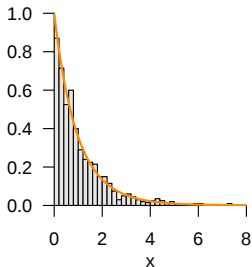
Wahrscheinlichkeitsdichtefunktionen der f -Verteilung



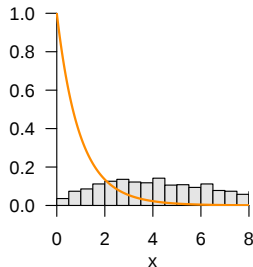
Simulation der Verteilung der F-Teststatistik

$H_0 : \mu_i := 3 \text{ für } i = 1, 2, 3 \text{ und } H_1 := \mu_1 = 4, \mu_i = 3 \text{ für } i = 2, 3 \text{ und } \sigma^2 := 1$

$\mu_1 = \mu_2 = \mu_3$



$\mu_1 \neq \mu_2 = \mu_3$



Theorem (F-Test und F-Testumfangkontrolle)

Es sei das Modell der einfaktoriellen Varianzanalyse gegeben und die F-Teststatistik wie oben definiert. Weiterhin sei der kritische Wert-basierte Test

$$\phi(X) := 1_{\{F \geq k\}} := \begin{cases} 1 & F \geq k \\ 0 & F < k \end{cases} \quad (20)$$

definiert. Dann ist ϕ genau dann ein Level- α_0 -Test mit Testumfang α , wenn

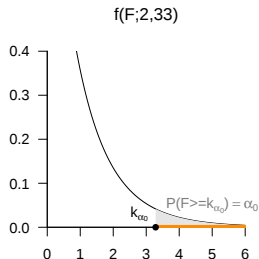
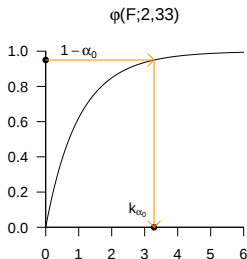
$$k = k_{\alpha_0} := \varphi^{-1}(1 - \alpha_0; p - 1, n - p) \quad (21)$$

gilt.

Bemerkungen

- Ein Beweis kann in Analogie zur Testumfangkontrolle bei T-Tests formuliert werden.

Wahl von $k_{\alpha_0} := \varphi^{-1}(1 - \alpha_0; p - 1, n - p)$ mit $p = 3, n = 12$ $\alpha_0 := 0.05$ und Ablehnungsbereich



Praktisches Vorgehen

Praktisches Vorgehen

- Man nimmt an, dass ein vorliegender Datensatz von p Gruppendatensätzen $x_{11}, \dots, x_{1n_1}, x_{21}, \dots, x_{2n_2}, \dots, x_{p1}, \dots, x_{pn_p}$ Realisationen von $X_{ij} \sim N(\mu_i, \sigma^2)$ mit unbekannten Parametern $\mu_i, i = 1, \dots, p$ und $\sigma^2 > 0$ sind.

Praktisches Vorgehen

- Man nimmt an, dass ein vorliegender Datensatz von p Gruppendatensätzen $x_{11}, \dots, x_{1n_1}, x_{21}, \dots, x_{2n_2}, \dots, x_{p1}, \dots, x_{pn_p}$ Realisationen von $X_{ij} \sim N(\mu_i, \sigma^2)$ mit unbekannten Parametern $\mu_i, i = 1, \dots, p$ und $\sigma^2 > 0$ sind.
- Man möchte entscheiden ob $H_0 : \mu_1 = \dots = \mu_p$ eher zutrifft oder eher nicht.

Praktisches Vorgehen

- Man nimmt an, dass ein vorliegender Datensatz von p Gruppendatensätzen $x_{11}, \dots, x_{1n_1}, x_{21}, \dots, x_{2n_2}, \dots, x_{p1}, \dots, x_{pn_p}$ Realisationen von $X_{ij} \sim N(\mu_i, \sigma^2)$ mit unbekannten Parametern $\mu_i, i = 1, \dots, p$ und $\sigma^2 > 0$ sind.
- Man möchte entscheiden ob $H_0 : \mu_1 = \dots = \mu_p$ eher zutrifft oder eher nicht.
- Man wählt ein Signifikanzniveau α_0 und bestimmt den zugehörigen Freiheitsgradparameter-abhängigen kritischen Wert k_{α_0} . Zum Beispiel gilt bei Wahl von $\alpha_0 := 0.05, p = 3, m = 12, i = 1, 2, 3$ und somit $n = 36$, dass $k_{\alpha_0} = \varphi^{-1}(1 - 0.05; 2, 33) \approx 3.28$ ist.

Praktisches Vorgehen

- Man nimmt an, dass ein vorliegender Datensatz von p Gruppendatensätzen $x_{11}, \dots, x_{1n_1}, x_{21}, \dots, x_{2n_2}, \dots, x_{p1}, \dots, x_{pn_p}$ Realisationen von $X_{ij} \sim N(\mu_i, \sigma^2)$ mit unbekannten Parametern $\mu_i, i = 1, \dots, p$ und $\sigma^2 > 0$ sind.
- Man möchte entscheiden ob $H_0 : \mu_1 = \dots = \mu_p$ eher zutrifft oder eher nicht.
- Man wählt ein Signifikanzniveau α_0 und bestimmt den zugehörigen Freiheitsgradparameter-abhängigen kritischen Wert k_{α_0} . Zum Beispiel gilt bei Wahl von $\alpha_0 := 0.05, p = 3, m = 12, i = 1, 2, 3$ und somit $n = 36$, dass $k_{\alpha_0} = \varphi^{-1}(1 - 0.05; 2, 33) \approx 3.28$ ist.
- Anhand der SQB und SQW sowie von n und p berechnet man den realisierten Wert der F-Teststatistik, den wir hier mit $\tilde{f}$ bezeichnen.

Praktisches Vorgehen

- Man nimmt an, dass ein vorliegender Datensatz von p Gruppendatensätzen $x_{11}, \dots, x_{1n_1}, x_{21}, \dots, x_{2n_2}, \dots, x_{p1}, \dots, x_{pn_p}$ Realisationen von $X_{ij} \sim N(\mu_i, \sigma^2)$ mit unbekannten Parametern $\mu_i, i = 1, \dots, p$ und $\sigma^2 > 0$ sind.
- Man möchte entscheiden ob $H_0 : \mu_1 = \dots = \mu_p$ eher zutrifft oder eher nicht.
- Man wählt ein Signifikanzniveau α_0 und bestimmt den zugehörigen Freiheitsgradparameter-abhängigen kritischen Wert k_{α_0} . Zum Beispiel gilt bei Wahl von $\alpha_0 := 0.05, p = 3, m = 12, i = 1, 2, 3$ und somit $n = 36$, dass $k_{\alpha_0} = \varphi^{-1}(1 - 0.05; 2, 33) \approx 3.28$ ist.
- Anhand der SQB und SQW sowie von n und p berechnet man den realisierten Wert der F-Teststatistik, den wir hier mit $\tilde{f}$ bezeichnen.
- Wenn $\tilde{f}$ größer gleich k_{α_0} ist, lehnt man die Nullhypothese ab, andernfalls nicht.

Praktisches Vorgehen

- Man nimmt an, dass ein vorliegender Datensatz von p Gruppendatensätzen $x_{11}, \dots, x_{1n_1}, x_{21}, \dots, x_{2n_2}, \dots, x_{p1}, \dots, x_{pn_p}$ Realisationen von $X_{ij} \sim N(\mu_i, \sigma^2)$ mit unbekannten Parametern $\mu_i, i = 1, \dots, p$ und $\sigma^2 > 0$ sind.
- Man möchte entscheiden ob $H_0 : \mu_1 = \dots = \mu_p$ eher zutrifft oder eher nicht.
- Man wählt ein Signifikanzniveau α_0 und bestimmt den zugehörigen Freiheitsgradparameter-abhängigen kritischen Wert k_{α_0} . Zum Beispiel gilt bei Wahl von $\alpha_0 := 0.05, p = 3, m = 12, i = 1, 2, 3$ und somit $n = 36$, dass $k_{\alpha_0} = \varphi^{-1}(1 - 0.05; 2, 33) \approx 3.28$ ist.
- Anhand der SQB und SQW sowie von n und p berechnet man den realisierten Wert der F-Teststatistik, den wir hier mit $\tilde{f}$ bezeichnen.
- Wenn $\tilde{f}$ größer gleich k_{α_0} ist, lehnt man die Nullhypothese ab, andernfalls nicht.
- Die oben entwickelte Theorie garantiert dann, dass man im Mittel in höchstens $\alpha_0 \cdot 100$ von 100 Fällen die Nullhypothese fälschlicherweise ablehnt.

Praktisches Vorgehen

- Man nimmt an, dass ein vorliegender Datensatz von p Gruppendatensätzen $x_{11}, \dots, x_{1n_1}, x_{21}, \dots, x_{2n_2}, \dots, x_{p1}, \dots, x_{pn_p}$ Realisationen von $X_{ij} \sim N(\mu_i, \sigma^2)$ mit unbekannten Parametern $\mu_i, i = 1, \dots, p$ und $\sigma^2 > 0$ sind.
- Man möchte entscheiden ob $H_0 : \mu_1 = \dots = \mu_p$ eher zutrifft oder eher nicht.
- Man wählt ein Signifikanzniveau α_0 und bestimmt den zugehörigen Freiheitsgradparameter-abhängigen kritischen Wert k_{α_0} . Zum Beispiel gilt bei Wahl von $\alpha_0 := 0.05, p = 3, m = 12, i = 1, 2, 3$ und somit $n = 36$, dass $k_{\alpha_0} = \varphi^{-1}(1 - 0.05; 2, 33) \approx 3.28$ ist.
- Anhand der SQB und SQW sowie von n und p berechnet man den realisierten Wert der F-Teststatistik, den wir hier mit $\tilde{f}$ bezeichnen.
- Wenn $\tilde{f}$ größer gleich k_{α_0} ist, lehnt man die Nullhypothese ab, andernfalls nicht.
- Die oben entwickelte Theorie garantiert dann, dass man im Mittel in höchstens $\alpha_0 \cdot 100$ von 100 Fällen die Nullhypothese fälschlicherweise ablehnt.
- Schließlich ergibt sich der assoziierte p-Wert der realisierter F-Teststatistik $\tilde{f}$ in Analogie zum Einstichproben-T-Test mit zusammengesetzter Nullhypothese zu

$$\text{p-Wert} = \mathbb{P}(F \geq \tilde{f}) = 1 - \varphi(\tilde{f}; p - 1, n - p) \quad (22)$$

Anwendungsbeispiel

x_{ij} : Datenpunkt der j ten experimentellen in Einheit in der i ten Gruppe

	x_{1j}	x_{2j}	x_{3j}
1	3.55	5.12	4.35
2	2.79	4.88	5.60
3	4.04	3.66	3.22
4	4.64	5.17	5.33
5	3.21	5.16	5.56
6	3.62	5.03	6.22
7	3.52	2.95	4.79
8	4.72	5.21	5.70
9	3.98	7.67	4.29
10	2.63	3.77	4.54
11	3.02	5.83	3.23
12	3.45	5.53	5.19

Anwendungsbeispiel - Manuelle Berechnung

```
p      = ncol(X)           # Gruppenanzahl
m      = nrow(X)           # Gruppengröße bei gleichen Gruppengrößen
n      = p*m               # Gesamtstichprobengröße
x_bar  = mean(as.matrix(X)) # Gesamtmittelwert
x_bar_i = colMeans(as.matrix(X)) # Gruppenmittelwerte
alpha_0 = 0.05             # Signifikanzlevel
k_alpha_0 = qf(1 - alpha_0, p-1, n-p) # kritischer Wert
SQT     = sum((as.vector(as.matrix(X)) - x_bar)^2) # Total Sum of Squares
SQB     = sum(m*(x_bar_i - x_bar)^2) # Between-Group Sum of Squares
SQW     = sum((t(as.matrix(X)) - as.vector(x_bar_i))^2) # Within-Group Sum of Squares
MSB     = SQB/(p-1)         # Between-Group Mean Squares
MSW     = SQW/(n-p)         # Within-Group Mean Squares
f_tilde = MSB/MSW           # Realisierung der F-Teststatistik
if(f_tilde >= k_alpha_0){   # \tilde{f} >= k_alpha_0
  phi = 1                   # Testwert
} else {
  phi = 0                   # Testwert
}
p_value = 1 - pf(f_tilde, p-1, n-p) # p-Wert
```

Anwendungsbeispiel - Manuelle Berechnung

```
p      = ncol(X)           # Gruppenanzahl
m      = nrow(X)           # Gruppengröße bei gleichen Gruppengrößen
n      = p*m               # Gesamtstichprobengröße
x_bar  = mean(as.matrix(X)) # Gesamtmittelwert
x_bar_i = colMeans(as.matrix(X)) # Gruppenmittelwerte
alpha_0 = 0.05             # Signifikanzlevel
k_alpha_0 = qf(1 - alpha_0, p-1, n-p) # kritischer Wert
SQT    = sum((as.vector(as.matrix(X)) - x_bar)^2) # Total Sum of Squares
SQB    = sum(m*(x_bar_i - x_bar)^2) # Between-Group Sum of Squares
SQW    = sum((t(as.matrix(X)) - as.vector(x_bar_i))^2) # Within-Group Sum of Squares
MSB    = SQB/(p-1)         # Between-Group Mean Squares
MSW    = SQW/(n-p)         # Within-Group Mean Squares
f_tilde = MSB/MSW          # Realisierung der F-Teststatistik
if(f_tilde >= k_alpha_0){  # \tilde{f} >= k_alpha_0
  phi = 1                  # Testwert
} else {
  phi = 0                  # Testwert
}
p_value = 1 - pf(f_tilde, p-1, n-p) # p-Wert
```

```
Gruppenanzahl      = 3
Gruppengrößen      = 12
Gesamtstichprobengröße = 36
Gesamtmittelwert   = 4.48
Gruppenmittelwerte  = 3.6 5 4.84
alpha_0            = 0.05
k_alpha_0          = 3.28
SQT                = 44.6
SQB                = 14.1
SQW                = 30.5
MSB                = 7.05
MSW                = 0.923
f_tilde            = 7.63
phi                = 1
p-Wert             = 0.00189
```

Anwendungsbeispiel - Automatische Berechnung

```
XAOV = data.frame(measurements = c(X$X_1j, X$X_2j, X$X_3j),  
                  group        = as.factor(c(rep(1,m),  
                                              rep(2,m),  
                                              rep(3,m))))  
res.aov = aov(measurements ~ group, data = XAOV)  
summary(res.aov)
```

*# Reformatierung des Datensatzes:
Datenpunktgruppenlabel als Faktorlevelspalte
Benutzung von R's aov Funktion
Ausgabe*

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
group	2	14.1	7.05	7.63	0.0019 **
Residuals	33	30.5	0.92		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Einfaktorielle Varianzanalyse

- Anwendungsszenario und Datendeskription
- Modellformulierung
- Modellschätzung
- Modellevaluation

Zweifaktorielle Varianzanalyse

- **Anwendungsszenario und Datendeskription**
- Modellformulierung
- Modellschätzung
- Modellevaluation

Selbstkontrollfragen

Zweifaktorielle Varianzanalyse = Two-way Analysis of Variance (ANOVA)

Zweifaktorielle Varianzanalyse = Two-way Analysis of Variance (ANOVA)

Anwendungsszenario

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

Zweifaktorielle Varianzanalyse = Two-way Analysis of Variance (ANOVA)

Anwendungsszenario

- **Zweifaktorielle Versuchspläne mit crossed design (cf. Modul A.2)**
 - Eine univariate Messvariable bestimmt an individuellen experimentellen Einheiten.

Zweifaktorielle Varianzanalyse = Two-way Analysis of Variance (ANOVA)

Anwendungsszenario

- **Zweifaktorielle Versuchspläne mit crossed design (cf. Modul A.2)**
 - Eine univariate Messvariable bestimmt an individuellen experimentellen Einheiten.
 - Zwei diskrete experimentelle Variablen, die mindestens zweistufig sind.

Zweifaktorielle Varianzanalyse = Two-way Analysis of Variance (ANOVA)

Anwendungsszenario

- **Zweifaktorielle Versuchspläne mit crossed design (cf. Modul A.2)**
 - Eine univariate Messvariable bestimmt an individuellen experimentellen Einheiten.
 - Zwei diskrete experimentelle Variablen, die mindestens zweistufig sind.
 - Die experimentellen Variablen werden *Faktoren* genannt.

Zweifaktorielle Varianzanalyse = Two-way Analysis of Variance (ANOVA)

Anwendungsszenario

- **Zweifaktorielle Versuchspläne mit crossed design (cf. Modul A.2)**
 - Eine univariate Messvariable bestimmt an individuellen experimentellen Einheiten.
 - Zwei diskrete experimentelle Variablen, die mindestens zweistufig sind.
 - Die experimentellen Variablen werden *Faktoren* genannt.
 - Die Stufen der Faktoren werden auch *Faktorlevel* genannt.

Zweifaktorielle Varianzanalyse = Two-way Analysis of Variance (ANOVA)

Anwendungsszenario

- **Zweifaktorielle Versuchspläne mit crossed design (cf. Modul A.2)**
 - Eine univariate Messvariable bestimmt an individuellen experimentellen Einheiten.
 - Zwei diskrete experimentelle Variablen, die mindestens zweistufig sind.
 - Die experimentellen Variablen werden *Faktoren* genannt.
 - Die Stufen der Faktoren werden auch *Faktorlevel* genannt.
 - Jedes Level eines Faktors wird mit allen Level des anderen Faktors kombiniert

Zweifaktorielle Varianzanalyse = Two-way Analysis of Variance (ANOVA)

Anwendungsszenario

- **Zweifaktorielle Versuchspläne mit crossed design (cf. Modul A.2)**
 - Eine univariate Messvariable bestimmt an individuellen experimentellen Einheiten.
 - Zwei diskrete experimentelle Variablen, die mindestens zweistufig sind.
 - Die experimentellen Variablen werden *Faktoren* genannt.
 - Die Stufen der Faktoren werden auch *Faktorlevel* genannt.
 - Jedes Level eines Faktors wird mit allen Level des anderen Faktors kombiniert
- Zweifaktorielle Versuchspläne 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

Zweifaktorielle Varianzanalyse = Two-way Analysis of Variance (ANOVA)

Anwendungsszenario

- **Zweifaktorielle Versuchspläne mit crossed design (cf. Modul A.2)**
 - Eine univariate Messvariable bestimmt an individuellen experimentellen Einheiten.
 - Zwei diskrete experimentelle Variablen, die mindestens zweistufig sind.
 - Die experimentellen Variablen werden *Faktoren* genannt.
 - Die Stufen der Faktoren werden auch *Faktorlevel* genannt.
 - Jedes Level eines Faktors wird mit allen Level des anderen Faktors kombiniert
- Zweifaktorielle Versuchspläne 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
- In der Psychologie sind 2×2 Designs sehr populär, wir fokussieren auf diesen Fall.

Zweifaktorielle Varianzanalyse = Two-way Analysis of Variance (ANOVA)

Anwendungsszenario

- **Zweifaktorielle Versuchspläne mit crossed design (cf. Modul A.2)**
 - Eine univariate Messvariable bestimmt an individuellen experimentellen Einheiten.
 - Zwei diskrete experimentelle Variablen, die mindestens zweistufig sind.
 - Die experimentellen Variablen werden *Faktoren* genannt.
 - Die Stufen der Faktoren werden auch *Faktorlevel* genannt.
 - Jedes Level eines Faktors wird mit allen Level des anderen Faktors kombiniert
- Zweifaktorielle Versuchspläne 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
- In der Psychologie sind 2×2 Designs sehr populär, wir fokussieren auf diesen Fall.
- Die Zellen eines 2×2 Designs werden im folgenden auch als Gruppen bezeichnet.

Anwendungsszenario 2 × 2 ANOVA

Konzeptuelles Design

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

Anwendungsszenario 2 × 2 ANOVA

Datennotation

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

x_{ijk} bezeichnet den Datenpunkt der k ten experimentellen Einheit ($k = 1, \dots, m$)
im i ten Level von Faktor A und j ten Level von Faktor B ($i = 1, 2, j = 1, 2$)

Anwendungsszenario 2 × 2 ANOVA

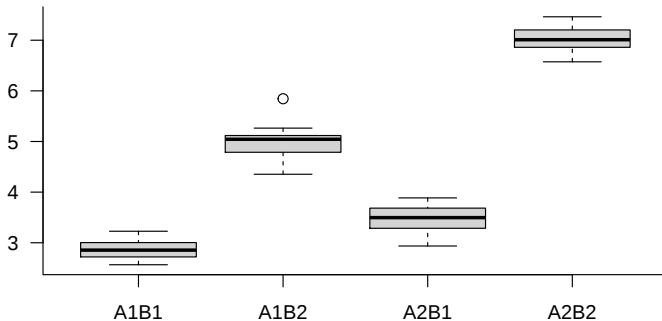
Datentabellenform mit $m := 12$

	x_{11k}	x_{12k}	x_{21k}	x_{22k}
1	2.86	5.04	3.29	6.88
2	2.62	4.96	3.69	7.33
3	3.01	4.58	2.94	6.76
4	3.20	5.05	3.61	6.57
5	2.75	5.05	3.68	6.84
6	2.88	5.01	3.89	7.45
7	2.85	4.35	3.43	7.06
8	3.23	5.07	3.72	6.99
9	2.99	5.84	3.28	6.93
10	2.57	4.61	3.35	7.46
11	2.69	5.26	2.94	7.07
12	2.83	5.17	3.56	7.03

x_{ijk} bezeichnet den Datenpunkt der k ten experimentellen Einheit ($k = 1, \dots, m$)
im i ten Level von Faktor A und j ten Level von Faktor B ($i = 1, 2, j = 1, 2$)

Anwendungsszenario 2 x 2 ANOVA

Gruppendatendeskription durch Boxplot



Anwendungsszenario 2×2 ANOVA

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

Anwendungsszenario 2 × 2 ANOVA

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.

Anwendungsszenario 2 × 2 ANOVA

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.

Anwendungsszenario 2 × 2 ANOVA

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.

Anwendungsszenario 2×2 ANOVA

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.

Anwendungsszenario 2 × 2 ANOVA

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

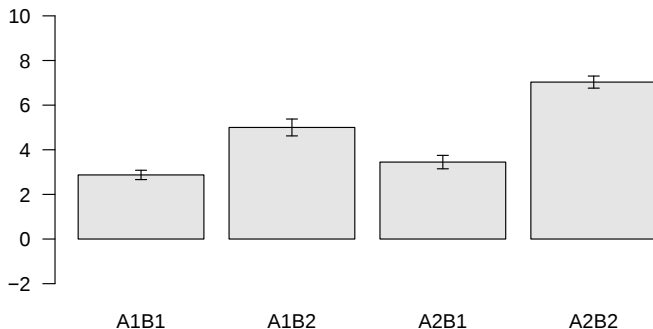
- Intuitiv spricht man vom Vorliegen eines *Haupteffekts von Faktor A*, wenn sich die Gruppenmittelwerte zwischen Level 1 und Level 2 von Faktor A, jeweils gemittelt über die zwei Level von Faktor B, unterscheiden.
- Intuitiv spricht man vom Vorliegen eines *Haupteffekts von Faktor B*, wenn sich die 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. Haupteffekte und Interaktionen sind lediglich Datenmuster, keine mechanistischen wissenschaftlichen Theorien.

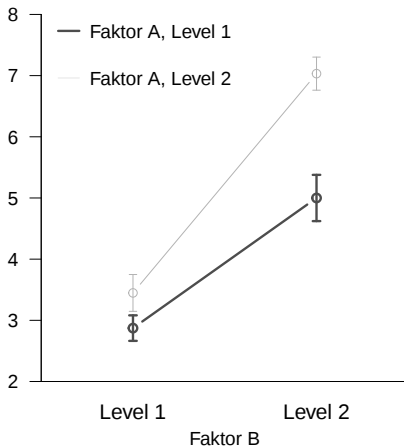
Anwendungsszenario 2 × 2 ANOVA

Darstellung der Gruppenmittelwerte und Gruppenstandardabweichungen durch Balkendiagramm



Anwendungsszenario 2 × 2 ANOVA

Darstellung der Gruppenmittelwerte und Gruppenstandardabweichungen durch Liniendiagramm



Anwendungsbeispiele

BDI Score Pre-Post-Differenzwertanalyse bei Klassischer und Online PT für zwei Altersklassen

- Faktor (A) Therapie mit Level (1) Klassische PT und Level (2) Online PT
- Faktor (B) Altersklasse mit Level (1) Junge Erwachsene und (2) Ältere Erwachsene
- 2×2 faktorieller Versuchsplan mit crossed design $\Rightarrow 2 \times 2$ ANOVA

Anwendungsbeispiele

BDI Score Pre-Post-Differenzwertanalyse bei Klassischer und Online PT für zwei Altersklassen

- Faktor (A) Therapie mit Level (1) Klassische PT und Level (2) Online PT
- Faktor (B) Altersklasse mit Level (1) Junge Erwachsene und (2) Ältere Erwachsene
- 2×2 faktorieller Versuchsplan mit crossed design $\Rightarrow 2 \times 2$ ANOVA
 - Inferentielle Evidenz für einen Haupteffekt der Therapieart?
 - Inferentielle Evidenz für einen Haupteffekt der Altersklasse?
 - Inferentielle Evidenz für Interaktion (Verschiedene Therapieeffekte zwischen Altersklassen)?

Anwendungsbeispiele

BDI Score Pre-Post-Differenzwertanalyse bei Klassischer und Online PT für zwei Altersklassen

- Faktor (A) Therapie mit Level (1) Klassische PT und Level (2) Online PT
- Faktor (B) Altersklasse mit Level (1) Junge Erwachsene und (2) Ältere Erwachsene
- 2×2 faktorieller Versuchsplan mit crossed design $\Rightarrow 2 \times 2$ ANOVA
 - Inferentielle Evidenz für einen Haupteffekt der Therapieart?
 - Inferentielle Evidenz für einen Haupteffekt der Altersklasse?
 - Inferentielle Evidenz für Interaktion (Verschiedene Therapieeffekte zwischen Altersklassen)?

TA Score Pre-Post-Differenzwertanalyse der vier Filmclips und zwei Studien im Affect Datensatz

- Faktor (A) Filmclips mit Level (1) Frontline, (2) Halloween, (3) Geographic, (4) Parenthood
- Faktor (B) Studien mit Level (1) Flat und (2) Maps
- 4×2 faktorieller Versuchsplan mit crossed design $\Rightarrow 4 \times 2$ ANOVA

Anwendungsbeispiele

BDI Score Pre-Post-Differenzwertanalyse bei Klassischer und Online PT für zwei Altersklassen

- Faktor (A) Therapie mit Level (1) Klassische PT und Level (2) Online PT
- Faktor (B) Altersklasse mit Level (1) Junge Erwachsene und (2) Ältere Erwachsene
- 2×2 faktorieller Versuchsplan mit crossed design $\Rightarrow 2 \times 2$ ANOVA
 - Inferentielle Evidenz für einen Haupteffekt der Therapieart?
 - Inferentielle Evidenz für einen Haupteffekt der Altersklasse?
 - Inferentielle Evidenz für Interaktion (Verschiedene Therapieeffekte zwischen Altersklassen)?

TA Score Pre-Post-Differenzwertanalyse der vier Filmclips und zwei Studien im Affect Datensatz

- Faktor (A) Filmclips mit Level (1) Frontline, (2) Halloween, (3) Geographic, (4) Parenthood
- Faktor (B) Studien mit Level (1) Flat und (2) Maps
- 4×2 faktorieller Versuchsplan mit crossed design $\Rightarrow 4 \times 2$ ANOVA
 - Inferentielle Evidenz für einen Haupteffekt der Filmclipbedingung?
 - Inferentielle Evidenz für einen Haupteffekt der Studie?
 - Inferentielle Evidenz für eine Interaktion (Änderung der Filmclipeffekte zwischen den Studien)?

Einfaktorielle Varianzanalyse

- Anwendungsszenario und Datendeskription
- Modellformulierung
- Modellschätzung
- Modellevaluation

Zweifaktorielle Varianzanalyse

- Anwendungsszenario und Datendeskription
- **Modellformulierung**
- Modellschätzung
- Modellevaluation

Definition (Modell der additiven zweifaktoriellen Varianzanalyse)

Es sei

$$X_{ij1}, \dots, X_{ijm} \sim N(\mu_{ij}, \sigma^2) \text{ für } i = 1, \dots, p \text{ und } j = 1, \dots, q \quad (23)$$

eine Stichprobe eines Normalverteilungsmodells der Größe m mit unbekanntem Erwartungswertparameter μ_{ij} und unbekanntem Varianzparameter $\sigma^2 > 0$. Dann heißt (23) das *Klassische Modell der additiven zweifaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung*.

Definition (Modell der additiven zweifaktoriellen Varianzanalyse)

Es sei

$$X_{ij1}, \dots, X_{ijm} \sim N(\mu_{ij}, \sigma^2) \text{ für } i = 1, \dots, p \text{ und } j = 1, \dots, q \quad (23)$$

eine Stichprobe eines Normalverteilungsmodells der Größe m mit unbekanntem Erwartungswertparameter μ_{ij} und unbekanntem Varianzparameter $\sigma^2 > 0$. Dann heißt (23) das *Klassische Modell der additiven zweifaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung*. Gilt für die Erwartungswertparameter in der Erwartungswertparameterdarstellung der zweifaktoriellen Varianzanalyse, dass

$$\mu_{ij} = \mu + \alpha_i + \beta_j \text{ für } i = 1, \dots, p, j = 1, \dots, q \text{ mit } \sum_{i=1}^p \alpha_i = 0 \text{ und } \sum_{j=1}^q \beta_j = 0, \quad (24)$$

dann heißen (23) und (24) *Klassisches Modell der additiven zweifaktoriellen Varianzanalyse in Effektdarstellung*. Die Parameter $\alpha_1, \dots, \alpha_p$ heißen die *Haupteffektparameter von Faktor A*, die Parameter $\beta_1, \dots, \beta_q$ heißen die *Haupteffektparameter von Faktor B*, und wir bezeichnen die Gesamtstichprobengröße der zweifaktoriellen Varianzanalyse mit $n := pqm$.

Definition (Modell der additiven zweifaktoriellen Varianzanalyse)

Es sei

$$X_{ij1}, \dots, X_{ijm} \sim N(\mu_{ij}, \sigma^2) \text{ für } i = 1, \dots, p \text{ und } j = 1, \dots, q \quad (23)$$

eine Stichprobe eines Normalverteilungsmodells der Größe m mit unbekanntem Erwartungswertparameter μ_{ij} und unbekanntem Varianzparameter $\sigma^2 > 0$. Dann heißt (23) das *Klassische Modell der additiven zweifaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung*. Gilt für die Erwartungswertparameter in der Erwartungswertparameterdarstellung der zweifaktoriellen Varianzanalyse, dass

$$\mu_{ij} = \mu + \alpha_i + \beta_j \text{ für } i = 1, \dots, p, j = 1, \dots, q \text{ mit } \sum_{i=1}^p \alpha_i = 0 \text{ und } \sum_{j=1}^q \beta_j = 0, \quad (24)$$

dann heißen (23) und (24) *Klassisches Modell der additiven zweifaktoriellen Varianzanalyse in Effektdarstellung*. Die Parameter $\alpha_1, \dots, \alpha_p$ heißen die *Haupteffektparameter von Faktor A*, die Parameter $\beta_1, \dots, \beta_q$ heißen die *Haupteffektparameter von Faktor B*, und wir bezeichnen die Gesamtstichprobengröße der zweifaktoriellen Varianzanalyse mit $n := pqm$. Schließlich heißt

$$X_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ij} \text{ mit } \sum_{i=1}^p \alpha_i = 0, \sum_{j=1}^q \beta_j = 0 \text{ und } \varepsilon_{ij} \sim N(0, \sigma^2) \quad (25)$$

Generatives Modell der additiven zweifaktoriellen Varianzanalyse in Effektdarstellung.

Bemerkungen

- Der Einfachheit halber setzen wir hier die identische Stichprobengröße m voraus.

Bemerkungen

- Der Einfachheit halber setzen wir hier die identische Stichprobengröße m voraus.
- Es seien

$$\bar{\mu} := \frac{1}{pq} \sum_{i=1}^p \sum_{j=1}^q \mu_{ij}, \quad \bar{\mu}_i := \frac{1}{q} \sum_{j=1}^q \mu_{ij}, \quad \text{und} \quad \bar{\mu}_j := \frac{1}{p} \sum_{i=1}^p \mu_{ij}. \quad (26)$$

Dann gelten folgende Beziehungen zwischen den Parametern der Erwartungswertparameterdarstellung und der Effektdarstellung für $i = 1, \dots, p$ und $j = 1, \dots, q$:

$$\mu = \bar{\mu}, \quad \alpha_i = \bar{\mu}_i - \mu, \quad \beta_j = \bar{\mu}_j - \mu \quad (27)$$

Bemerkungen

- Der Einfachheit halber setzen wir hier die identische Stichprobengröße m voraus.
- Es seien

$$\bar{\mu} := \frac{1}{pq} \sum_{i=1}^p \sum_{j=1}^q \mu_{ij}, \quad \bar{\mu}_i := \frac{1}{q} \sum_{j=1}^q \mu_{ij}, \quad \text{und} \quad \bar{\mu}_j := \frac{1}{p} \sum_{i=1}^p \mu_{ij}. \quad (26)$$

Dann gelten folgende Beziehungen zwischen den Parametern der Erwartungswertparameterdarstellung und der Effektdarstellung für $i = 1, \dots, p$ und $j = 1, \dots, q$:

$$\mu = \bar{\mu}, \quad \alpha_i = \bar{\mu}_i - \mu, \quad \beta_j = \bar{\mu}_j - \mu \quad (27)$$

In der Effektdarstellung der additiven zweifaktoriellen Varianzanalyse gilt also

- Der Parameter μ encodiert den Mittelwert der Gruppenerwartungswertparameter μ_{ij} .

Bemerkungen

- Der Einfachheit halber setzen wir hier die identische Stichprobengröße m voraus.
- Es seien

$$\bar{\mu} := \frac{1}{pq} \sum_{i=1}^p \sum_{j=1}^q \mu_{ij}, \quad \bar{\mu}_i := \frac{1}{q} \sum_{j=1}^q \mu_{ij}, \quad \text{und} \quad \bar{\mu}_j := \frac{1}{p} \sum_{i=1}^p \mu_{ij}. \quad (26)$$

Dann gelten folgende Beziehungen zwischen den Parametern der Erwartungswertparameterdarstellung und der Effektdarstellung für $i = 1, \dots, p$ und $j = 1, \dots, q$:

$$\mu = \bar{\mu}, \quad \alpha_i = \bar{\mu}_i - \mu, \quad \beta_j = \bar{\mu}_j - \mu \quad (27)$$

In der Effektdarstellung der additiven zweifaktoriellen Varianzanalyse gilt also

- Der Parameter μ encodiert den Mittelwert der Gruppenerwartungswertparameter μ_{ij} .
- Die Parameter α_i encodieren die Differenz der über j gemittelten Gruppenerwartungswerte μ_{ij} vom Mittelwert der Gruppenerwartungswerte. Die Haupteffektparameter α_i erklären also den Unterschied zwischen den Mittelwerten in der jeweiligen Faktorstufe von Faktor A zum Gesamtmittel ohne Berücksichtigung der Faktorstufe des anderen Faktors.

Bemerkungen

- Der Einfachheit halber setzen wir hier die identische Stichprobengröße m voraus.
- Es seien

$$\bar{\mu} := \frac{1}{pq} \sum_{i=1}^p \sum_{j=1}^q \mu_{ij}, \quad \bar{\mu}_i := \frac{1}{q} \sum_{j=1}^q \mu_{ij}, \quad \text{und} \quad \bar{\mu}_j := \frac{1}{p} \sum_{i=1}^p \mu_{ij}. \quad (26)$$

Dann gelten folgende Beziehungen zwischen den Parametern der Erwartungswertparameterdarstellung und der Effektdarstellung für $i = 1, \dots, p$ und $j = 1, \dots, q$:

$$\mu = \bar{\mu}, \quad \alpha_i = \bar{\mu}_i - \mu, \quad \beta_j = \bar{\mu}_j - \mu \quad (27)$$

In der Effektdarstellung der additiven zweifaktoriellen Varianzanalyse gilt also

- Der Parameter μ encodiert den Mittelwert der Gruppenerwartungswertparameter μ_{ij} .
- Die Parameter α_i encodieren die Differenz der über j gemittelten Gruppenerwartungswerte μ_{ij} vom Mittelwert der Gruppenerwartungswerte. Die Haupteffektparameter α_i erklären also den Unterschied zwischen den Mittelwerten in der jeweiligen Faktorstufe von Faktor A zum Gesamtmittel ohne Berücksichtigung der Faktorstufe des anderen Faktors.
- Die Parameter β_j encodieren die Differenz der über i gemittelten Gruppenerwartungswerte μ_{ij} vom Mittelwert der Gruppenerwartungswerte. Die Haupteffektparameter β_j erklären also den Unterschied zwischen den Mittelwerten in der jeweiligen Faktorstufe von Faktor B zum Gesamtmittel ohne Berücksichtigung der Faktorstufe des anderen Faktors.

Bemerkungen

- Der Einfachheit halber setzen wir hier die identische Stichprobengröße m voraus.
- Es seien

$$\bar{\mu} := \frac{1}{pq} \sum_{i=1}^p \sum_{j=1}^q \mu_{ij}, \quad \bar{\mu}_i := \frac{1}{q} \sum_{j=1}^q \mu_{ij}, \quad \text{und} \quad \bar{\mu}_j := \frac{1}{p} \sum_{i=1}^p \mu_{ij}. \quad (26)$$

Dann gelten folgende Beziehungen zwischen den Parametern der Erwartungswertparameterdarstellung und der Effektdarstellung für $i = 1, \dots, p$ und $j = 1, \dots, q$:

$$\mu = \bar{\mu}, \quad \alpha_i = \bar{\mu}_i - \mu, \quad \beta_j = \bar{\mu}_j - \mu \quad (27)$$

In der Effektdarstellung der additiven zweifaktoriellen Varianzanalyse gilt also

- Der Parameter μ encodiert den Mittelwert der Gruppenerwartungswertparameter μ_{ij} .
- Die Parameter α_i encodieren die Differenz der über j gemittelten Gruppenerwartungswerte μ_{ij} vom Mittelwert der Gruppenerwartungswerte. Die Haupteffektparameter α_i erklären also den Unterschied zwischen den Mittelwerten in der jeweiligen Faktorstufe von Faktor A zum Gesamtmittel ohne Berücksichtigung der Faktorstufe des anderen Faktors.
- Die Parameter β_j encodieren die Differenz der über i gemittelten Gruppenerwartungswerte μ_{ij} vom Mittelwert der Gruppenerwartungswerte. Die Haupteffektparameter β_j erklären also den Unterschied zwischen den Mittelwerten in der jeweiligen Faktorstufe von Faktor B zum Gesamtmittel ohne Berücksichtigung der Faktorstufe des anderen Faktors.

Wir verzichten auf Beweise dieser Tatsachen.

Bemerkungen (fortgeführt)

- Die Bedingungen $\sum_{i=1}^p \alpha_i = 0$ und $\sum_{j=1}^q \beta_j = 0$ stellen die Eindeutigkeit der Darstellung der pq Parameter μ_{ij} durch die $1 + p + q$ Parameter μ, α_i, β_j sicher. Dass diese ohne weitere Bedingungen an die α_i, β_j nicht gegeben ist, macht man sich an folgendem Beispiel klar

Bemerkungen (fortgeführt)

- Die Bedingungen $\sum_{i=1}^p \alpha_i = 0$ und $\sum_{j=1}^q \beta_j = 0$ stellen die Eindeutigkeit der Darstellung der pq Parameter μ_{ij} durch die $1 + p + q$ Parameter μ, α_i, β_j sicher. Dass diese ohne weitere Bedingungen an die α_i, β_j nicht gegeben ist, macht man sich an folgendem Beispiel klar

- Für $p = 2$ und $q = 2$ gelte

$$\mu_{11} = 1, \mu_{12} = 2, \mu_{21} = 3, \mu_{22} = 4. \quad (28)$$

Dann können die μ_{ij} sowohl durch

$$\mu = 0, \alpha_1 = 1, \alpha_2 = 3, \beta_1 = 0, \beta_2 = 1, \quad (29)$$

als auch durch

$$\mu = 1, \alpha_1 = 0, \alpha_2 = 2, \beta_1 = 0, \beta_2 = 1 \quad (30)$$

dargestellt werden

Bemerkungen (fortgeführt)

- Die Bedingungen $\sum_{i=1}^p \alpha_i = 0$ und $\sum_{j=1}^q \beta_j = 0$ stellen die Eindeutigkeit der Darstellung der pq Parameter μ_{ij} durch die $1 + p + q$ Parameter μ, α_i, β_j sicher. Dass diese ohne weitere Bedingungen an die α_i, β_j nicht gegeben ist, macht man sich an folgendem Beispiel klar

- Für $p = 2$ und $q = 2$ gelte

$$\mu_{11} = 1, \mu_{12} = 2, \mu_{21} = 3, \mu_{22} = 4. \quad (28)$$

Dann können die μ_{ij} sowohl durch

$$\mu = 0, \alpha_1 = 1, \alpha_2 = 3, \beta_1 = 0, \beta_2 = 1, \quad (29)$$

als auch durch

$$\mu = 1, \alpha_1 = 0, \alpha_2 = 2, \beta_1 = 0, \beta_2 = 1 \quad (30)$$

dargestellt werden

- Wir verzichten hier auf einen Beweis, dass die Bedingungen $\sum_{i=1}^p \alpha_i = 0$ und $\sum_{j=1}^q \beta_j = 0$ eine eindeutige Darstellung der μ_{ij} erlauben.

Parameterbeispiele bei additiven 2×2 Varianzanalysen

(1) Es sei $\mu := 1, \alpha_1 := 1, \beta_1 := 0$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = -1$; aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 0$.
- Es ergibt sich also die Erwartungswertparameterdarstellung

$$\begin{aligned}\mu_{11} &= \mu + \alpha_1 + \beta_1 = 1 + 1 + 0 = \mathbf{2} & \mu_{12} &= \mu + \alpha_1 + \beta_2 = 1 + 1 - 0 = \mathbf{2} \\ \mu_{21} &= \mu + \alpha_2 + \beta_1 = 1 - 1 + 0 = \mathbf{0} & \mu_{22} &= \mu + \alpha_2 + \beta_2 = 1 - 1 + 0 = \mathbf{0}\end{aligned}$$

Parameterbeispiele bei additiven 2×2 Varianzanalysen

(1) Es sei $\mu := 1, \alpha_1 := 1, \beta_1 := 0$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = -1$; aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 0$.
- Es ergibt sich also die Erwartungswertparameterdarstellung

$$\begin{aligned}\mu_{11} &= \mu + \alpha_1 + \beta_1 = 1 + 1 + 0 = \mathbf{2} & \mu_{12} &= \mu + \alpha_1 + \beta_2 = 1 + 1 - 0 = \mathbf{2} \\ \mu_{21} &= \mu + \alpha_2 + \beta_1 = 1 - 1 + 0 = \mathbf{0} & \mu_{22} &= \mu + \alpha_2 + \beta_2 = 1 - 1 + 0 = \mathbf{0}\end{aligned}$$

(2) Es sei $\mu := 1, \alpha_1 := 0, \beta_1 := 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = 0$; aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = -1$.
- Es ergibt sich also die Erwartungswertparameterdarstellung

$$\begin{aligned}\mu_{11} &= \mu + \alpha_1 + \beta_1 = 1 + 0 + 1 = \mathbf{2} & \mu_{12} &= \mu + \alpha_1 + \beta_2 = 1 + 0 - 1 = \mathbf{0} \\ \mu_{21} &= \mu + \alpha_2 + \beta_1 = 1 + 0 + 1 = \mathbf{2} & \mu_{22} &= \mu + \alpha_2 + \beta_2 = 1 + 0 - 1 = \mathbf{0}\end{aligned}$$

Parameterbeispiele bei additiven 2×2 Varianzanalysen

(1) Es sei $\mu := 1, \alpha_1 := 1, \beta_1 := 0$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = -1$; aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 0$.
- Es ergibt sich also die Erwartungswertparameterdarstellung

$$\begin{aligned}\mu_{11} &= \mu + \alpha_1 + \beta_1 = 1 + 1 + 0 = \mathbf{2} & \mu_{12} &= \mu + \alpha_1 + \beta_2 = 1 + 1 - 0 = \mathbf{2} \\ \mu_{21} &= \mu + \alpha_2 + \beta_1 = 1 - 1 + 0 = \mathbf{0} & \mu_{22} &= \mu + \alpha_2 + \beta_2 = 1 - 1 + 0 = \mathbf{0}\end{aligned}$$

(2) Es sei $\mu := 1, \alpha_1 := 0, \beta_1 := 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = 0$; aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = -1$.
- Es ergibt sich also die Erwartungswertparameterdarstellung

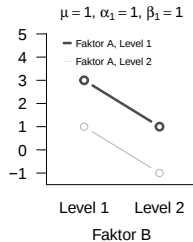
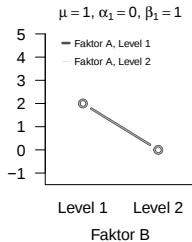
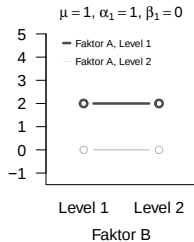
$$\begin{aligned}\mu_{11} &= \mu + \alpha_1 + \beta_1 = 1 + 0 + 1 = \mathbf{2} & \mu_{12} &= \mu + \alpha_1 + \beta_2 = 1 + 0 - 1 = \mathbf{0} \\ \mu_{21} &= \mu + \alpha_2 + \beta_1 = 1 + 0 + 1 = \mathbf{2} & \mu_{22} &= \mu + \alpha_2 + \beta_2 = 1 + 0 - 1 = \mathbf{0}\end{aligned}$$

(3) Es sei $\mu := 1, \alpha_1 := 1, \beta_1 := 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = -1$; aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = -1$.
- Es ergibt sich also die Erwartungswertparameterdarstellung

$$\begin{aligned}\mu_{11} &= \mu + \alpha_1 + \beta_1 = 1 + 1 + 1 = \mathbf{3} & \mu_{12} &= \mu + \alpha_1 + \beta_2 = 1 + 1 - 1 = \mathbf{1} \\ \mu_{21} &= \mu + \alpha_2 + \beta_1 = 1 - 1 + 1 = \mathbf{1} & \mu_{22} &= \mu + \alpha_2 + \beta_2 = 1 - 1 - 1 = \mathbf{-1}\end{aligned}$$

Parameterbeispiele bei additiven 2×2 Varianzanalysen



Definition (Modell der zweifaktoriellen Varianzanalyse)

Es sei

$$X_{ij1}, \dots, X_{ijm} \sim N(\mu_{ij}, \sigma^2) \text{ für } i = 1, \dots, p \text{ und } j = 1, \dots, q \quad (31)$$

eine Stichprobe eines Normalverteilungsmodells der Größe m mit unbekanntem Erwartungswertparameter μ_{ij} und unbekanntem Varianzparameter $\sigma^2 > 0$. Dann heißt (31) das *Klassische Modell der zweifaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung*.

Definition (Modell der zweifaktoriellen Varianzanalyse)

Es sei

$$X_{ij1}, \dots, X_{ijm} \sim N(\mu_{ij}, \sigma^2) \text{ für } i = 1, \dots, p \text{ und } j = 1, \dots, q \quad (31)$$

eine Stichprobe eines Normalverteilungsmodells der Größe m mit unbekanntem Erwartungswertparameter μ_{ij} und unbekanntem Varianzparameter $\sigma^2 > 0$. Dann heißt (31) das *Klassische Modell der zweifaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung*. Gilt für die Erwartungswertparameter in der Erwartungswertparameterdarstellung der zweifaktoriellen Varianzanalyse, dass

$$\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} \text{ für } i = 1, \dots, p, j = 1, \dots, q, \quad (32)$$

mit

$$\sum_{i=1}^p \alpha_i = 0, \sum_{j=1}^q \beta_j = 0, \sum_{i=1}^p \gamma_{ij} = 0, \sum_{j=1}^q \gamma_{ij} = 0 \text{ und } n := pqm \quad (33)$$

dann heißen (31), (32) und (33) *Klassisches Modell der zweifaktoriellen Varianzanalyse in Effektdarstellung*. $\alpha_1, \dots, \alpha_p$ heißen weiterhin die *Haupteffektparameter von Faktor A*, $\beta_1, \dots, \beta_q$ heißen weiterhin die *Haupteffektparameter von Faktor B*, und $\gamma_{11}, \dots, \gamma_{pq}$ heißen die *Interaktionseffektparameter der Faktoren A und B*.

Definition (Modell der zweifaktoriellen Varianzanalyse)

Es sei

$$X_{ij1}, \dots, X_{ijm} \sim N(\mu_{ij}, \sigma^2) \text{ für } i = 1, \dots, p \text{ und } j = 1, \dots, q \quad (31)$$

eine Stichprobe eines Normalverteilungsmodells der Größe m mit unbekanntem Erwartungswertparameter μ_{ij} und unbekanntem Varianzparameter $\sigma^2 > 0$. Dann heißt (31) das *Klassische Modell der zweifaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung*. Gilt für die Erwartungswertparameter in der Erwartungswertparameterdarstellung der zweifaktoriellen Varianzanalyse, dass

$$\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} \text{ für } i = 1, \dots, p, j = 1, \dots, q, \quad (32)$$

mit

$$\sum_{i=1}^p \alpha_i = 0, \sum_{j=1}^q \beta_j = 0, \sum_{i=1}^p \gamma_{ij} = 0, \sum_{j=1}^q \gamma_{ij} = 0 \text{ und } n := pqm \quad (33)$$

dann heißen (31), (32) und (33) *Klassisches Modell der zweifaktoriellen Varianzanalyse in Effektdarstellung*. $\alpha_1, \dots, \alpha_p$ heißen weiterhin die *Haupteffektparameter von Faktor A*, $\beta_1, \dots, \beta_q$ heißen weiterhin die *Haupteffektparameter von Faktor B*, und $\gamma_{11}, \dots, \gamma_{pq}$ heißen die *Interaktionseffektparameter der Faktoren A und B*. Schließlich heißt

$$X_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk} \text{ mit } \varepsilon_{ijk} \sim N(0, \sigma^2) \quad (34)$$

mit den Parameterbedingungen (33) *Generatives Modell der zweifaktoriellen Varianzanalyse in Effektdarstellung*.

Bemerkungen

- Es seien

$$\bar{\mu} := \frac{1}{pq} \sum_{i=1}^p \sum_{j=1}^q \mu_{ij}, \quad \bar{\mu}_i := \frac{1}{q} \sum_{j=1}^q \mu_{ij}, \quad \text{und} \quad \bar{\mu}_j := \frac{1}{p} \sum_{i=1}^p \mu_{ij}. \quad (35)$$

Dann gelten folgende Beziehungen zwischen den Parametern der Erwartungswertparameterdarstellung und der Effektdarstellung für $i = 1, \dots, p$ und $j = 1, \dots, q$:

Bemerkungen

- Es seien

$$\bar{\mu} := \frac{1}{pq} \sum_{i=1}^p \sum_{j=1}^q \mu_{ij}, \quad \bar{\mu}_i := \frac{1}{q} \sum_{j=1}^q \mu_{ij}, \quad \text{und} \quad \bar{\mu}_j := \frac{1}{p} \sum_{i=1}^p \mu_{ij}. \quad (35)$$

Dann gelten folgende Beziehungen zwischen den Parametern der Erwartungswertparameterdarstellung und der Effektdarstellung für $i = 1, \dots, p$ und $j = 1, \dots, q$:

$$\mu = \bar{\mu}, \quad \alpha_i = \bar{\mu}_i - \mu, \quad \beta_j = \bar{\mu}_j - \mu, \quad \gamma_{ij} = \mu_{ij} - (\mu + \alpha_i + \beta_j) \quad (36)$$

Bemerkungen

- Es seien

$$\bar{\mu} := \frac{1}{pq} \sum_{i=1}^p \sum_{j=1}^q \mu_{ij}, \quad \bar{\mu}_i := \frac{1}{q} \sum_{j=1}^q \mu_{ij}, \quad \text{und} \quad \bar{\mu}_j := \frac{1}{p} \sum_{i=1}^p \mu_{ij}. \quad (35)$$

Dann gelten folgende Beziehungen zwischen den Parametern der Erwartungswertparameterdarstellung und der Effektdarstellung für $i = 1, \dots, p$ und $j = 1, \dots, q$:

$$\mu = \bar{\mu}, \quad \alpha_i = \bar{\mu}_i - \mu, \quad \beta_j = \bar{\mu}_j - \mu, \quad \gamma_{ij} = \mu_{ij} - (\mu + \alpha_i + \beta_j) \quad (36)$$

- In der Effektdarstellung der zweifaktoriellen Varianzanalyse gilt also
 - Die Parameter μ und die Haupteffektparameter α_i und β_j enkodieren die gleichen Aspekte wie in der additiven zweifaktoriellen Varianzanalyse.

Bemerkungen

- Es seien

$$\bar{\mu} := \frac{1}{pq} \sum_{i=1}^p \sum_{j=1}^q \mu_{ij}, \quad \bar{\mu}_i := \frac{1}{q} \sum_{j=1}^q \mu_{ij}, \quad \text{und} \quad \bar{\mu}_j := \frac{1}{p} \sum_{i=1}^p \mu_{ij}. \quad (35)$$

Dann gelten folgende Beziehungen zwischen den Parametern der Erwartungswertparameterdarstellung und der Effektdarstellung für $i = 1, \dots, p$ und $j = 1, \dots, q$:

$$\mu = \bar{\mu}, \quad \alpha_i = \bar{\mu}_i - \mu, \quad \beta_j = \bar{\mu}_j - \mu, \quad \gamma_{ij} = \mu_{ij} - (\mu + \alpha_i + \beta_j) \quad (36)$$

- In der Effektdarstellung der zweifaktoriellen Varianzanalyse gilt also
 - Die Parameter μ und die Haupteffektparameter α_i und β_j kodieren die gleichen Aspekte wie in der additiven zweifaktoriellen Varianzanalyse.
 - Die Interaktionsparameter γ_{ij} kodieren den Unterschied zwischen dem Erwartungswertparameter in der jeweiligen Faktorlevelkombination zu dem Teil des Gesamtmittelwertes, der nicht durch Addition der Haupteffektparameter beschrieben werden kann.

Zweifaktorielle Varianzanalyse | Modellformulierung

Bemerkungen

- Es seien

$$\bar{\mu} := \frac{1}{pq} \sum_{i=1}^p \sum_{j=1}^q \mu_{ij}, \quad \bar{\mu}_i := \frac{1}{q} \sum_{j=1}^q \mu_{ij}, \quad \text{und} \quad \bar{\mu}_j := \frac{1}{p} \sum_{i=1}^p \mu_{ij}. \quad (35)$$

Dann gelten folgende Beziehungen zwischen den Parametern der Erwartungswertparameterdarstellung und der Effektdarstellung für $i = 1, \dots, p$ und $j = 1, \dots, q$:

$$\mu = \bar{\mu}, \quad \alpha_i = \bar{\mu}_i - \mu, \quad \beta_j = \bar{\mu}_j - \mu, \quad \gamma_{ij} = \mu_{ij} - (\mu + \alpha_i + \beta_j) \quad (36)$$

- In der Effektdarstellung der zweifaktoriellen Varianzanalyse gilt also
 - Die Parameter μ und die Haupteffektparameter α_i und β_j enkodieren die gleichen Aspekte wie in der additiven zweifaktoriellen Varianzanalyse.
 - Die Interaktionsparameter γ_{ij} enkodieren den Unterschied zwischen dem Erwartungswertparameter in der jeweiligen Faktorlevelkombination zu dem Teil des Gesamtmittelwertes, der nicht durch Addition der Haupteffektparameter beschrieben werden kann.
- Die Bedingungen

$$\sum_{i=1}^p \alpha_i = 0, \quad \sum_{j=1}^q \beta_j = 0, \quad \sum_{i=1}^p \gamma_{ij} = 0 \quad \text{und} \quad \sum_{j=1}^q \gamma_{ij} = 0 \quad (37)$$

stellen die Eindeutigkeit der Darstellung der pq Parameter μ_{ij} durch die $1 + p + q + qp$ Parameter $\mu, \alpha_i, \beta_j, \gamma_{ij}$ sicher. Wir verzichten wiederum auf einen Beweis dieser Tatsache.

Zweifaktorielle Varianzanalyse | Modellformulierung

Bemerkungen

- Es seien

$$\bar{\mu} := \frac{1}{pq} \sum_{i=1}^p \sum_{j=1}^q \mu_{ij}, \quad \bar{\mu}_i := \frac{1}{q} \sum_{j=1}^q \mu_{ij}, \quad \text{und} \quad \bar{\mu}_j := \frac{1}{p} \sum_{i=1}^p \mu_{ij}. \quad (35)$$

Dann gelten folgende Beziehungen zwischen den Parametern der Erwartungswertparameterdarstellung und der Effektdarstellung für $i = 1, \dots, p$ und $j = 1, \dots, q$:

$$\mu = \bar{\mu}, \quad \alpha_i = \bar{\mu}_i - \mu, \quad \beta_j = \bar{\mu}_j - \mu, \quad \gamma_{ij} = \mu_{ij} - (\mu + \alpha_i + \beta_j) \quad (36)$$

- In der Effektdarstellung der zweifaktoriellen Varianzanalyse gilt also
 - Die Parameter μ und die Haupteffektparameter α_i und β_j enkodieren die gleichen Aspekte wie in der additiven zweifaktoriellen Varianzanalyse.
 - Die Interaktionsparameter γ_{ij} enkodieren den Unterschied zwischen dem Erwartungswertparameter in der jeweiligen Faktorlevelkombination zu dem Teil des Gesamtmittelwertes, der nicht durch Addition der Haupteffektparameter beschrieben werden kann.
- Die Bedingungen

$$\sum_{i=1}^p \alpha_i = 0, \quad \sum_{j=1}^q \beta_j = 0, \quad \sum_{i=1}^p \gamma_{ij} = 0 \quad \text{und} \quad \sum_{j=1}^q \gamma_{ij} = 0 \quad (37)$$

stellen die Eindeutigkeit der Darstellung der pq Parameter μ_{ij} durch die $1 + p + q + qp$ Parameter $\mu, \alpha_i, \beta_j, \gamma_{ij}$ sicher. Wir verzichten wiederum auf einen Beweis dieser Tatsache.

- Die Gesamtheit aller Stichprobenvariablen bezeichnen wir mit $X := \{X_{ijk}\}_{i=1, \dots, p, j=1, \dots, q, k=1, \dots, m}$.

Parameterbeispiele bei 2×2 Varianzanalysen

(1) Es sei $\mu := 0, \alpha_1 := 0, \beta_1 := 0, \gamma_{11} = 1$

Parameterbeispiele bei 2×2 Varianzanalysen

(1) Es sei $\mu := 0, \alpha_1 := 0, \beta_1 := 0, \gamma_{11} = 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = 0$

Parameterbeispiele bei 2×2 Varianzanalysen

(1) Es sei $\mu := 0, \alpha_1 := 0, \beta_1 := 0, \gamma_{11} = 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = 0$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 0$.

Parameterbeispiele bei 2 x 2 Varianzanalysen

(1) Es sei $\mu := 0$, $\alpha_1 := 0$, $\beta_1 := 0$, $\gamma_{11} = 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = 0$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 0$.
- Aus $\sum_{i=1}^p \gamma_{ij} = 0$ für $j = 1, 2$, also $\gamma_{11} + \gamma_{21} = 0$ und $\gamma_{12} + \gamma_{22} = 0$, folgt $\gamma_{21} = -1$

Parameterbeispiele bei 2 x 2 Varianzanalysen

(1) Es sei $\mu := 0, \alpha_1 := 0, \beta_1 := 0, \gamma_{11} = 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = 0$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 0$.
- Aus $\sum_{i=1}^p \gamma_{ij} = 0$ für $j = 1, 2$, also $\gamma_{11} + \gamma_{21} = 0$ und $\gamma_{12} + \gamma_{22} = 0$, folgt $\gamma_{21} = -1$
- Aus $\sum_{j=1}^q \gamma_{ij} = 0$ für $i = 1, 2$, also $\gamma_{11} + \gamma_{12} = 0$ und $\gamma_{21} + \gamma_{22} = 0$, folgt $\gamma_{12} = -1$

Parameterbeispiele bei 2 x 2 Varianzanalysen

(1) Es sei $\mu := 0, \alpha_1 := 0, \beta_1 := 0, \gamma_{11} = 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = 0$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 0$.
- Aus $\sum_{i=1}^p \gamma_{ij} = 0$ für $j = 1, 2$, also $\gamma_{11} + \gamma_{21} = 0$ und $\gamma_{12} + \gamma_{22} = 0$, folgt $\gamma_{21} = -1$
- Aus $\sum_{j=1}^q \gamma_{ij} = 0$ für $i = 1, 2$, also $\gamma_{11} + \gamma_{12} = 0$ und $\gamma_{21} + \gamma_{22} = 0$, folgt $\gamma_{12} = -1$
- Mit $\gamma_{12} + \gamma_{22} = 0$ folgt dann auch $\gamma_{22} = 1$

Parameterbeispiele bei 2 x 2 Varianzanalysen

(1) Es sei $\mu := 0, \alpha_1 := 0, \beta_1 := 0, \gamma_{11} = 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = 0$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 0$.
- Aus $\sum_{i=1}^p \gamma_{ij} = 0$ für $j = 1, 2$, also $\gamma_{11} + \gamma_{21} = 0$ und $\gamma_{12} + \gamma_{22} = 0$, folgt $\gamma_{21} = -1$
- Aus $\sum_{j=1}^q \gamma_{ij} = 0$ für $i = 1, 2$, also $\gamma_{11} + \gamma_{12} = 0$ und $\gamma_{21} + \gamma_{22} = 0$, folgt $\gamma_{12} = -1$
- Mit $\gamma_{12} + \gamma_{22} = 0$ folgt dann auch $\gamma_{22} = 1$

Parameterbeispiele bei 2 x 2 Varianzanalysen

(1) Es sei $\mu := 0, \alpha_1 := 0, \beta_1 := 0, \gamma_{11} = 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = 0$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 0$.
- Aus $\sum_{i=1}^p \gamma_{ij} = 0$ für $j = 1, 2$, also $\gamma_{11} + \gamma_{21} = 0$ und $\gamma_{12} + \gamma_{22} = 0$, folgt $\gamma_{21} = -1$
- Aus $\sum_{j=1}^q \gamma_{ij} = 0$ für $i = 1, 2$, also $\gamma_{11} + \gamma_{12} = 0$ und $\gamma_{21} + \gamma_{22} = 0$, folgt $\gamma_{12} = -1$
- Mit $\gamma_{12} + \gamma_{22} = 0$ folgt dann auch $\gamma_{22} = 1$

Die Wahl von $\mu := 1, \alpha_1 := 1, \beta_1 := 0, \gamma_{11} := 1$ legt also $\alpha_2, \beta_2, \gamma_{21}, \gamma_{12}, \gamma_{22}$ fest.

Parameterbeispiele bei 2 x 2 Varianzanalysen

(1) Es sei $\mu := 0, \alpha_1 := 0, \beta_1 := 0, \gamma_{11} = 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = 0$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 0$.
- Aus $\sum_{i=1}^p \gamma_{ij} = 0$ für $j = 1, 2$, also $\gamma_{11} + \gamma_{21} = 0$ und $\gamma_{12} + \gamma_{22} = 0$, folgt $\gamma_{21} = -1$
- Aus $\sum_{j=1}^q \gamma_{ij} = 0$ für $i = 1, 2$, also $\gamma_{11} + \gamma_{12} = 0$ und $\gamma_{21} + \gamma_{22} = 0$, folgt $\gamma_{12} = -1$
- Mit $\gamma_{12} + \gamma_{22} = 0$ folgt dann auch $\gamma_{22} = 1$

Die Wahl von $\mu := 1, \alpha_1 := 1, \beta_1 := 0, \gamma_{11} := 1$ legt also $\alpha_2, \beta_2, \gamma_{21}, \gamma_{12}, \gamma_{22}$ fest.

Es ergibt sich also die Erwartungswertparameterdarstellung

$$\mu_{11} = \mu + \alpha_1 + \beta_1 + \gamma_{11} = 0 + 0 + 0 + 1 = 1 \qquad \mu_{12} = \mu + \alpha_1 + \beta_2 + \gamma_{12} = 0 + 0 + 0 - 1 = -1$$

$$\mu_{21} = \mu + \alpha_2 + \beta_1 + \gamma_{21} = 0 + 0 + 0 - 1 = -1 \qquad \mu_{22} = \mu + \alpha_2 + \beta_2 + \gamma_{22} = 0 + 0 + 0 + 1 = 1$$

Parameterbeispiele bei 2×2 Varianzanalysen

(2) Es sei $\mu := 0, \alpha_1 := 0, \beta_1 := 0, \gamma_{11} = -1$

Parameterbeispiele bei 2×2 Varianzanalysen

(2) Es sei $\mu := 0, \alpha_1 := 0, \beta_1 := 0, \gamma_{11} = -1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = 0$

Parameterbeispiele bei 2×2 Varianzanalysen

(2) Es sei $\mu := 0, \alpha_1 := 0, \beta_1 := 0, \gamma_{11} = -1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = 0$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 0$.

Parameterbeispiele bei 2 x 2 Varianzanalysen

(2) Es sei $\mu := 0, \alpha_1 := 0, \beta_1 := 0, \gamma_{11} = -1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = 0$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 0$.
- Aus $\sum_{i=1}^p \gamma_{ij} = 0$ für $j = 1, 2$, also $\gamma_{11} + \gamma_{21} = 0$ und $\gamma_{12} + \gamma_{22} = 0$, folgt $\gamma_{21} = 1$

Parameterbeispiele bei 2 x 2 Varianzanalysen

(2) Es sei $\mu := 0, \alpha_1 := 0, \beta_1 := 0, \gamma_{11} = -1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = 0$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 0$.
- Aus $\sum_{i=1}^p \gamma_{ij} = 0$ für $j = 1, 2$, also $\gamma_{11} + \gamma_{21} = 0$ und $\gamma_{12} + \gamma_{22} = 0$, folgt $\gamma_{21} = 1$
- Aus $\sum_{j=1}^q \gamma_{ij} = 0$ für $i = 1, 2$, also $\gamma_{11} + \gamma_{12} = 0$ und $\gamma_{21} + \gamma_{22} = 0$, folgt $\gamma_{12} = 1$

Parameterbeispiele bei 2 x 2 Varianzanalysen

(2) Es sei $\mu := 0, \alpha_1 := 0, \beta_1 := 0, \gamma_{11} = -1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = 0$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 0$.
- Aus $\sum_{i=1}^p \gamma_{ij} = 0$ für $j = 1, 2$, also $\gamma_{11} + \gamma_{21} = 0$ und $\gamma_{12} + \gamma_{22} = 0$, folgt $\gamma_{21} = 1$
- Aus $\sum_{j=1}^q \gamma_{ij} = 0$ für $i = 1, 2$, also $\gamma_{11} + \gamma_{12} = 0$ und $\gamma_{21} + \gamma_{22} = 0$, folgt $\gamma_{12} = 1$
- Mit $\gamma_{12} + \gamma_{22} = 0$ folgt dann auch $\gamma_{22} = -1$

Parameterbeispiele bei 2 x 2 Varianzanalysen

(2) Es sei $\mu := 0, \alpha_1 := 0, \beta_1 := 0, \gamma_{11} = -1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = 0$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 0$.
- Aus $\sum_{i=1}^p \gamma_{ij} = 0$ für $j = 1, 2$, also $\gamma_{11} + \gamma_{21} = 0$ und $\gamma_{12} + \gamma_{22} = 0$, folgt $\gamma_{21} = 1$
- Aus $\sum_{j=1}^q \gamma_{ij} = 0$ für $i = 1, 2$, also $\gamma_{11} + \gamma_{12} = 0$ und $\gamma_{21} + \gamma_{22} = 0$, folgt $\gamma_{12} = 1$
- Mit $\gamma_{12} + \gamma_{22} = 0$ folgt dann auch $\gamma_{22} = -1$

Parameterbeispiele bei 2 x 2 Varianzanalysen

(2) Es sei $\mu := 0, \alpha_1 := 0, \beta_1 := 0, \gamma_{11} = -1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = 0$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 0$.
- Aus $\sum_{i=1}^p \gamma_{ij} = 0$ für $j = 1, 2$, also $\gamma_{11} + \gamma_{21} = 0$ und $\gamma_{12} + \gamma_{22} = 0$, folgt $\gamma_{21} = 1$
- Aus $\sum_{j=1}^q \gamma_{ij} = 0$ für $i = 1, 2$, also $\gamma_{11} + \gamma_{12} = 0$ und $\gamma_{21} + \gamma_{22} = 0$, folgt $\gamma_{12} = 1$
- Mit $\gamma_{12} + \gamma_{22} = 0$ folgt dann auch $\gamma_{22} = -1$

Die Wahl von $\mu := 1, \alpha_1 := 1, \beta_1 := 0, \gamma_{11} := 1$ legt also $\alpha_2, \beta_2, \gamma_{21}, \gamma_{12}, \gamma_{22}$ fest.

Parameterbeispiele bei 2 x 2 Varianzanalysen

(2) Es sei $\mu := 0, \alpha_1 := 0, \beta_1 := 0, \gamma_{11} = -1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = 0$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 0$.
- Aus $\sum_{i=1}^p \gamma_{ij} = 0$ für $j = 1, 2$, also $\gamma_{11} + \gamma_{21} = 0$ und $\gamma_{12} + \gamma_{22} = 0$, folgt $\gamma_{21} = 1$
- Aus $\sum_{j=1}^q \gamma_{ij} = 0$ für $i = 1, 2$, also $\gamma_{11} + \gamma_{12} = 0$ und $\gamma_{21} + \gamma_{22} = 0$, folgt $\gamma_{12} = 1$
- Mit $\gamma_{12} + \gamma_{22} = 0$ folgt dann auch $\gamma_{22} = -1$

Die Wahl von $\mu := 1, \alpha_1 := 1, \beta_1 := 0, \gamma_{11} := 1$ legt also $\alpha_2, \beta_2, \gamma_{21}, \gamma_{12}, \gamma_{22}$ fest.

Es ergibt sich also die Erwartungswertparameterdarstellung

$$\mu_{11} = \mu + \alpha_1 + \beta_1 + \gamma_{11} = 0 + 0 + 0 - 1 = -1 \qquad \mu_{12} = \mu + \alpha_1 + \beta_2 + \gamma_{12} = 0 + 0 + 0 + 1 = 1$$

$$\mu_{21} = \mu + \alpha_2 + \beta_1 + \gamma_{21} = 0 + 0 + 0 + 1 = 1 \qquad \mu_{22} = \mu + \alpha_2 + \beta_2 + \gamma_{22} = 0 + 0 + 0 - 1 = -1$$

Parameterbeispiele bei 2×2 Varianzanalysen

(3) Es sei $\mu := 0, \alpha_1 := 1, \beta_1 := 0, \gamma_{11} = 1$

Parameterbeispiele bei 2 x 2 Varianzanalysen

(3) Es sei $\mu := 0, \alpha_1 := 1, \beta_1 := 0, \gamma_{11} = 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = -1$

Die Wahl von $\mu := 1, \alpha_1 := 1, \beta_1 := 0, \gamma_{11} := 1$ legt also $\alpha_2, \beta_2, \gamma_{21}, \gamma_{12}, \gamma_{22}$ fest.

Parameterbeispiele bei 2 x 2 Varianzanalysen

(3) Es sei $\mu := 0, \alpha_1 := 1, \beta_1 := 0, \gamma_{11} = 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = -1$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 0$.

Die Wahl von $\mu := 1, \alpha_1 := 1, \beta_1 := 0, \gamma_{11} := 1$ legt also $\alpha_2, \beta_2, \gamma_{21}, \gamma_{12}, \gamma_{22}$ fest.

Parameterbeispiele bei 2 x 2 Varianzanalysen

(3) Es sei $\mu := 0, \alpha_1 := 1, \beta_1 := 0, \gamma_{11} = 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = -1$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 0$.
- Aus $\sum_{i=1}^p \gamma_{ij} = 0$ für $j = 1, 2$, also $\gamma_{11} + \gamma_{21} = 0$ und $\gamma_{12} + \gamma_{22} = 0$, folgt $\gamma_{21} = -1$

Die Wahl von $\mu := 1, \alpha_1 := 1, \beta_1 := 0, \gamma_{11} := 1$ legt also $\alpha_2, \beta_2, \gamma_{21}, \gamma_{12}, \gamma_{22}$ fest.

Parameterbeispiele bei 2 x 2 Varianzanalysen

(3) Es sei $\mu := 0, \alpha_1 := 1, \beta_1 := 0, \gamma_{11} = 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = -1$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 0$.
- Aus $\sum_{i=1}^p \gamma_{ij} = 0$ für $j = 1, 2$, also $\gamma_{11} + \gamma_{21} = 0$ und $\gamma_{12} + \gamma_{22} = 0$, folgt $\gamma_{21} = -1$
- Aus $\sum_{j=1}^q \gamma_{ij} = 0$ für $i = 1, 2$, also $\gamma_{11} + \gamma_{12} = 0$ und $\gamma_{21} + \gamma_{22} = 0$, folgt $\gamma_{12} = -1$

Die Wahl von $\mu := 1, \alpha_1 := 1, \beta_1 := 0, \gamma_{11} := 1$ legt also $\alpha_2, \beta_2, \gamma_{21}, \gamma_{12}, \gamma_{22}$ fest.

Parameterbeispiele bei 2 x 2 Varianzanalysen

(3) Es sei $\mu := 0, \alpha_1 := 1, \beta_1 := 0, \gamma_{11} = 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = -1$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 0$.
- Aus $\sum_{i=1}^p \gamma_{ij} = 0$ für $j = 1, 2$, also $\gamma_{11} + \gamma_{21} = 0$ und $\gamma_{12} + \gamma_{22} = 0$, folgt $\gamma_{21} = -1$
- Aus $\sum_{j=1}^q \gamma_{ij} = 0$ für $i = 1, 2$, also $\gamma_{11} + \gamma_{12} = 0$ und $\gamma_{21} + \gamma_{22} = 0$, folgt $\gamma_{12} = -1$
- Mit $\gamma_{12} + \gamma_{22} = 0$ folgt dann auch $\gamma_{22} = 1$

Die Wahl von $\mu := 1, \alpha_1 := 1, \beta_1 := 0, \gamma_{11} := 1$ legt also $\alpha_2, \beta_2, \gamma_{21}, \gamma_{12}, \gamma_{22}$ fest.

Parameterbeispiele bei 2 x 2 Varianzanalysen

(3) Es sei $\mu := 0, \alpha_1 := 1, \beta_1 := 0, \gamma_{11} = 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = -1$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 0$.
- Aus $\sum_{i=1}^p \gamma_{ij} = 0$ für $j = 1, 2$, also $\gamma_{11} + \gamma_{21} = 0$ und $\gamma_{12} + \gamma_{22} = 0$, folgt $\gamma_{21} = -1$
- Aus $\sum_{j=1}^q \gamma_{ij} = 0$ für $i = 1, 2$, also $\gamma_{11} + \gamma_{12} = 0$ und $\gamma_{21} + \gamma_{22} = 0$, folgt $\gamma_{12} = -1$
- Mit $\gamma_{12} + \gamma_{22} = 0$ folgt dann auch $\gamma_{22} = 1$

Die Wahl von $\mu := 1, \alpha_1 := 1, \beta_1 := 0, \gamma_{11} := 1$ legt also $\alpha_2, \beta_2, \gamma_{21}, \gamma_{12}, \gamma_{22}$ fest.

Parameterbeispiele bei 2 x 2 Varianzanalysen

(3) Es sei $\mu := 0, \alpha_1 := 1, \beta_1 := 0, \gamma_{11} = 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = -1$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 0$.
- Aus $\sum_{i=1}^p \gamma_{ij} = 0$ für $j = 1, 2$, also $\gamma_{11} + \gamma_{21} = 0$ und $\gamma_{12} + \gamma_{22} = 0$, folgt $\gamma_{21} = -1$
- Aus $\sum_{j=1}^q \gamma_{ij} = 0$ für $i = 1, 2$, also $\gamma_{11} + \gamma_{12} = 0$ und $\gamma_{21} + \gamma_{22} = 0$, folgt $\gamma_{12} = -1$
- Mit $\gamma_{12} + \gamma_{22} = 0$ folgt dann auch $\gamma_{22} = 1$

Die Wahl von $\mu := 1, \alpha_1 := 1, \beta_1 := 0, \gamma_{11} := 1$ legt also $\alpha_2, \beta_2, \gamma_{21}, \gamma_{12}, \gamma_{22}$ fest.

Es ergibt sich also die Erwartungswertparameterdarstellung

$$\mu_{11} = \mu + \alpha_1 + \beta_1 + \gamma_{11} = 0 + 1 + 0 + 1 = \mathbf{2} \qquad \mu_{12} = \mu + \alpha_1 + \beta_2 + \gamma_{12} = 0 + 1 + 0 - 1 = \mathbf{0}$$

$$\mu_{21} = \mu + \alpha_2 + \beta_1 + \gamma_{21} = 0 - 1 + 0 - 1 = \mathbf{-2} \qquad \mu_{22} = \mu + \alpha_2 + \beta_2 + \gamma_{22} = 0 - 1 + 0 + 1 = \mathbf{0}$$

Parameterbeispiele bei 2×2 Varianzanalysen

(4) Es sei $\mu := 1, \alpha_1 := 1, \beta_1 := -1, \gamma_{11} = 1$

Parameterbeispiele bei 2×2 Varianzanalysen

(4) Es sei $\mu := 1, \alpha_1 := 1, \beta_1 := -1, \gamma_{11} = 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = -1$

Parameterbeispiele bei 2×2 Varianzanalysen

(4) Es sei $\mu := 1, \alpha_1 := 1, \beta_1 := -1, \gamma_{11} = 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = -1$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 1$.

Parameterbeispiele bei 2 x 2 Varianzanalysen

(4) Es sei $\mu := 1, \alpha_1 := 1, \beta_1 := -1, \gamma_{11} = 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = -1$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 1$.
- Aus $\sum_{i=1}^p \gamma_{ij} = 0$ für $j = 1, 2$, also $\gamma_{11} + \gamma_{21} = 0$ und $\gamma_{12} + \gamma_{22} = 0$, folgt $\gamma_{21} = -1$

Parameterbeispiele bei 2 x 2 Varianzanalysen

(4) Es sei $\mu := 1, \alpha_1 := 1, \beta_1 := -1, \gamma_{11} = 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = -1$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 1$.
- Aus $\sum_{i=1}^p \gamma_{ij} = 0$ für $j = 1, 2$, also $\gamma_{11} + \gamma_{21} = 0$ und $\gamma_{12} + \gamma_{22} = 0$, folgt $\gamma_{21} = -1$
- Aus $\sum_{j=1}^q \gamma_{ij} = 0$ für $i = 1, 2$, also $\gamma_{11} + \gamma_{12} = 0$ und $\gamma_{21} + \gamma_{22} = 0$, folgt $\gamma_{12} = -1$

Parameterbeispiele bei 2 x 2 Varianzanalysen

(4) Es sei $\mu := 1, \alpha_1 := 1, \beta_1 := -1, \gamma_{11} = 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = -1$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 1$.
- Aus $\sum_{i=1}^p \gamma_{ij} = 0$ für $j = 1, 2$, also $\gamma_{11} + \gamma_{21} = 0$ und $\gamma_{12} + \gamma_{22} = 0$, folgt $\gamma_{21} = -1$
- Aus $\sum_{j=1}^q \gamma_{ij} = 0$ für $i = 1, 2$, also $\gamma_{11} + \gamma_{12} = 0$ und $\gamma_{21} + \gamma_{22} = 0$, folgt $\gamma_{12} = -1$
- Mit $\gamma_{12} + \gamma_{22} = 0$ folgt dann auch $\gamma_{22} = 1$

Parameterbeispiele bei 2 x 2 Varianzanalysen

(4) Es sei $\mu := 1, \alpha_1 := 1, \beta_1 := -1, \gamma_{11} = 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = -1$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 1$.
- Aus $\sum_{i=1}^p \gamma_{ij} = 0$ für $j = 1, 2$, also $\gamma_{11} + \gamma_{21} = 0$ und $\gamma_{12} + \gamma_{22} = 0$, folgt $\gamma_{21} = -1$
- Aus $\sum_{j=1}^q \gamma_{ij} = 0$ für $i = 1, 2$, also $\gamma_{11} + \gamma_{12} = 0$ und $\gamma_{21} + \gamma_{22} = 0$, folgt $\gamma_{12} = -1$
- Mit $\gamma_{12} + \gamma_{22} = 0$ folgt dann auch $\gamma_{22} = 1$

Parameterbeispiele bei 2 x 2 Varianzanalysen

(4) Es sei $\mu := 1, \alpha_1 := 1, \beta_1 := -1, \gamma_{11} = 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = -1$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 1$.
- Aus $\sum_{i=1}^p \gamma_{ij} = 0$ für $j = 1, 2$, also $\gamma_{11} + \gamma_{21} = 0$ und $\gamma_{12} + \gamma_{22} = 0$, folgt $\gamma_{21} = -1$
- Aus $\sum_{j=1}^q \gamma_{ij} = 0$ für $i = 1, 2$, also $\gamma_{11} + \gamma_{12} = 0$ und $\gamma_{21} + \gamma_{22} = 0$, folgt $\gamma_{12} = -1$
- Mit $\gamma_{12} + \gamma_{22} = 0$ folgt dann auch $\gamma_{22} = 1$

Die Wahl von $\mu := 1, \alpha_1 := 1, \beta_1 := 0, \gamma_{11} := 1$ legt also $\alpha_2, \beta_2, \gamma_{21}, \gamma_{12}, \gamma_{22}$ fest.

Parameterbeispiele bei 2 x 2 Varianzanalysen

(4) Es sei $\mu := 1, \alpha_1 := 1, \beta_1 := -1, \gamma_{11} = 1$

- Aus der Bedingung $\alpha_1 + \alpha_2 = 0$ folgt $\alpha_2 = -1$
- Aus der Bedingung $\beta_1 + \beta_2 = 0$ folgt $\beta_2 = 1$.
- Aus $\sum_{i=1}^p \gamma_{ij} = 0$ für $j = 1, 2$, also $\gamma_{11} + \gamma_{21} = 0$ und $\gamma_{12} + \gamma_{22} = 0$, folgt $\gamma_{21} = -1$
- Aus $\sum_{j=1}^q \gamma_{ij} = 0$ für $i = 1, 2$, also $\gamma_{11} + \gamma_{12} = 0$ und $\gamma_{21} + \gamma_{22} = 0$, folgt $\gamma_{12} = -1$
- Mit $\gamma_{12} + \gamma_{22} = 0$ folgt dann auch $\gamma_{22} = 1$

Die Wahl von $\mu := 1, \alpha_1 := 1, \beta_1 := 0, \gamma_{11} := 1$ legt also $\alpha_2, \beta_2, \gamma_{21}, \gamma_{12}, \gamma_{22}$ fest.

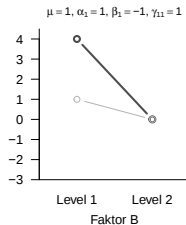
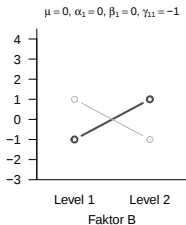
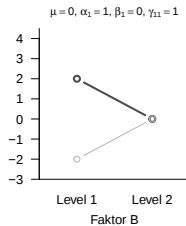
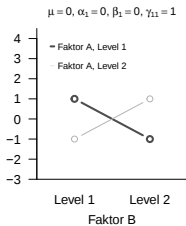
Es ergibt sich also die Erwartungswertparameterdarstellung

$$\mu_{11} = \mu + \alpha_1 + \beta_1 + \gamma_{11} = 1 + 1 + 1 + 1 = 4 \quad \mu_{12} = \mu + \alpha_1 + \beta_2 + \gamma_{12} = 1 + 1 - 1 - 1 = 0$$

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

Zweifaktorielle Varianzanalyse | Modellformulierung

Parameterbeispiele bei 2 x 2 Varianzanalysen



Einfaktorielle Varianzanalyse

- Anwendungsszenario und Datendeskription
- Modellformulierung
- Modellschätzung
- Modellevaluation

Zweifaktorielle Varianzanalyse

- Anwendungsszenario und Datendeskription
- Modellformulierung
- **Modellschätzung**
- Modellevaluation

Selbstkontrollfragen

Theorem (Parameterschätzer der zweifaktoriellen Varianzanalyse)

Für $i = 1, \dots, p$ und $j = 1, \dots, q$ seien μ , α_i , β_j , γ_{ij} und $\sigma^2 > 0$ die Parameter des Modells der zweifaktoriellen Varianzanalyse in (klassischer oder generativer) Effektdarstellung wie oben definiert.

Theorem (Parameterschätzer der zweifaktoriellen Varianzanalyse)

Für $i = 1, \dots, p$ und $j = 1, \dots, q$ seien $\mu, \alpha_i, \beta_j, \gamma_{ij}$ und $\sigma^2 > 0$ die Parameter des Modells der zweifaktoriellen Varianzanalyse in (klassischer oder generativer) Effektdarstellung wie oben definiert. Weiterhin seien

$$\bar{X} := \frac{1}{pqn} \sum_{i=1}^p \sum_{j=1}^q \sum_{k=1}^m X_{ijk} \quad (38)$$

und für $i = 1, \dots, p, j = 1, \dots, q$

$$\bar{X}_i := \frac{1}{qm} \sum_{j=1}^q \sum_{k=1}^m X_{ijk}, \bar{X}_j := \frac{1}{pm} \sum_{i=1}^p \sum_{k=1}^m X_{ijk} \text{ und } \bar{X}_{ij} := \frac{1}{m} \sum_{k=1}^m X_{ijk} \quad (39)$$

definiert.

Theorem (Parameterschätzer der zweifaktoriellen Varianzanalyse)

Für $i = 1, \dots, p$ und $j = 1, \dots, q$ seien $\mu, \alpha_i, \beta_j, \gamma_{ij}$ und $\sigma^2 > 0$ die Parameter des Modells der zweifaktoriellen Varianzanalyse in (klassischer oder generativer) Effektdarstellung wie oben definiert. Weiterhin seien

$$\bar{X} := \frac{1}{pqn} \sum_{i=1}^p \sum_{j=1}^q \sum_{k=1}^m X_{ijk} \quad (38)$$

und für $i = 1, \dots, p, j = 1, \dots, q$

$$\bar{X}_i := \frac{1}{qm} \sum_{j=1}^q \sum_{k=1}^m X_{ijk}, \bar{X}_j := \frac{1}{pm} \sum_{i=1}^p \sum_{k=1}^m X_{ijk} \text{ und } \bar{X}_{ij} := \frac{1}{m} \sum_{k=1}^m X_{ijk} \quad (39)$$

definiert. Dann sind die Maximum Likelihood Schätzer von μ, α_i, β_j und γ_{ij} gegeben durch

$$\hat{\mu} = \bar{X}, \quad \hat{\alpha}_i = \bar{X}_i - \bar{X}, \quad \hat{\beta}_j = \bar{X}_j - \bar{X} \text{ und } \hat{\gamma}_{ij} = \bar{X}_{ij} - (\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j) \quad (40)$$

Theorem (Parameterschätzer der zweifaktoriellen Varianzanalyse)

Für $i = 1, \dots, p$ und $j = 1, \dots, q$ seien $\mu, \alpha_i, \beta_j, \gamma_{ij}$ und $\sigma^2 > 0$ die Parameter des Modells der zweifaktoriellen Varianzanalyse in (klassischer oder generativer) Effektdarstellung wie oben definiert. Weiterhin seien

$$\bar{X} := \frac{1}{pqn} \sum_{i=1}^p \sum_{j=1}^q \sum_{k=1}^m X_{ijk} \quad (38)$$

und für $i = 1, \dots, p, j = 1, \dots, q$

$$\bar{X}_i := \frac{1}{qm} \sum_{j=1}^q \sum_{k=1}^m X_{ijk}, \bar{X}_j := \frac{1}{pm} \sum_{i=1}^p \sum_{k=1}^m X_{ijk} \text{ und } \bar{X}_{ij} := \frac{1}{m} \sum_{k=1}^m X_{ijk} \quad (39)$$

definiert. Dann sind die Maximum Likelihood Schätzer von μ, α_i, β_j und γ_{ij} gegeben durch

$$\hat{\mu} = \bar{X}, \quad \hat{\alpha}_i = \bar{X}_i - \bar{X}, \quad \hat{\beta}_j = \bar{X}_j - \bar{X} \text{ und } \hat{\gamma}_{ij} = \bar{X}_{ij} - (\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j) \quad (40)$$

Ein unverzerrter Schätzer von $\sigma^2 > 0$ ist gegeben durch

$$\hat{\sigma}^2 = \frac{1}{pq(m-1)} \sum_{i=1}^p \sum_{j=1}^q \sum_{k=1}^m (X_{ijk} - \bar{X}_{ij})^2. \quad (41)$$

Theorem (Parameterschätzer der zweifaktoriellen Varianzanalyse)

Für $i = 1, \dots, p$ und $j = 1, \dots, q$ seien $\mu, \alpha_i, \beta_j, \gamma_{ij}$ und $\sigma^2 > 0$ die Parameter des Modells der zweifaktoriellen Varianzanalyse in (klassischer oder generativer) Effektdarstellung wie oben definiert. Weiterhin seien

$$\bar{X} := \frac{1}{pqn} \sum_{i=1}^p \sum_{j=1}^q \sum_{k=1}^m X_{ijk} \quad (38)$$

und für $i = 1, \dots, p, j = 1, \dots, q$

$$\bar{X}_i := \frac{1}{qm} \sum_{j=1}^q \sum_{k=1}^m X_{ijk}, \bar{X}_j := \frac{1}{pm} \sum_{i=1}^p \sum_{k=1}^m X_{ijk} \text{ und } \bar{X}_{ij} := \frac{1}{m} \sum_{k=1}^m X_{ijk} \quad (39)$$

definiert. Dann sind die Maximum Likelihood Schätzer von μ, α_i, β_j und γ_{ij} gegeben durch

$$\hat{\mu} = \bar{X}, \quad \hat{\alpha}_i = \bar{X}_i - \bar{X}, \quad \hat{\beta}_j = \bar{X}_j - \bar{X} \text{ und } \hat{\gamma}_{ij} = \bar{X}_{ij} - (\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j) \quad (40)$$

Ein unverzerrter Schätzer von $\sigma^2 > 0$ ist gegeben durch

$$\hat{\sigma}^2 = \frac{1}{pq(m-1)} \sum_{i=1}^p \sum_{j=1}^q \sum_{k=1}^m (X_{ijk} - \bar{X}_{ij})^2. \quad (41)$$

- Wir verzichten auf einen Beweis.

Einfaktorielle Varianzanalyse

- Anwendungsszenario und Datendeskription
- Modellformulierung
- Modellschätzung
- Modellevaluation

Zweifaktorielle Varianzanalyse

- Anwendungsszenario und Datendeskription
- Modellformulierung
- Modellschätzung
- **Modellevaluation**

Selbstkontrollfragen

Hypothesen und Überblick

- Primäre Ziele der zweifaktoriellen Varianzanalyse sind das Testen von Haupteffekt- und Interaktionshypothesen

Hypothesen und Überblick

- Primäre Ziele der zweifaktoriellen Varianzanalyse sind das Testen von Haupteffekt- und Interaktionshypothesen
- Die Null- und Alternativhypothesen bezüglich der Haupteffektparameter von Faktor A lauten

$$H_0^A : \alpha_i = 0 \text{ für alle } i = 1, \dots, p \text{ und } H_1^A : \alpha_i \neq 0 \text{ für mindestens zwei } i \in \mathbb{N}_p.$$

Hypothesen und Überblick

- Primäre Ziele der zweifaktoriellen Varianzanalyse sind das Testen von Haupteffekt- und Interaktionshypothesen
- Die Null- und Alternativhypothesen bezüglich der Haupteffektparameter von Faktor A lauten

$$H_0^A : \alpha_i = 0 \text{ für alle } i = 1, \dots, p \text{ und } H_1^A : \alpha_i \neq 0 \text{ für mindestens zwei } i \in \mathbb{N}_p.$$

- Die Null- und Alternativhypothesen bezüglich der Haupteffektparameter von Faktor B lauten

$$H_0^B : \beta_j = 0 \text{ für alle } j = 1, \dots, q \text{ und } H_1^B : \beta_j \neq 0 \text{ für mindestens zwei } j \in \mathbb{N}_q.$$

Hypothesen und Überblick

- Primäre Ziele der zweifaktoriellen Varianzanalyse sind das Testen von Haupteffekt- und Interaktionshypothesen
- Die Null- und Alternativhypothesen bezüglich der Haupteffektparameter von Faktor A lauten

$$H_0^A : \alpha_i = 0 \text{ für alle } i = 1, \dots, p \text{ und } H_1^A : \alpha_i \neq 0 \text{ für mindestens zwei } i \in \mathbb{N}_p.$$

- Die Null- und Alternativhypothesen bezüglich der Haupteffektparameter von Faktor B lauten

$$H_0^B : \beta_j = 0 \text{ für alle } j = 1, \dots, q \text{ und } H_1^B : \beta_j \neq 0 \text{ für mindestens zwei } j \in \mathbb{N}_q.$$

- Die Null- und Alternativhypothesen bezüglich der Interaktionsparameter lauten

$$H_0^{A \times B} : \gamma_{ij} = 0 \text{ für alle } i = 1, \dots, p, j = 1, \dots, q$$

$$H_1^{A \times B} : \gamma_{ij} \neq 0 \text{ für mindestens zwei } (i, j) \in \mathbb{N}_p \times \mathbb{N}_q.$$

Die Interaktionsalternativhypothese macht keine Aussage über die Form der Interaktion.

Hypothesen und Überblick

- Primäre Ziele der zweifaktoriellen Varianzanalyse sind das Testen von Haupteffekt- und Interaktionshypothesen
- Die Null- und Alternativhypothesen bezüglich der Haupteffektparameter von Faktor A lauten

$$H_0^A : \alpha_i = 0 \text{ für alle } i = 1, \dots, p \text{ und } H_1^A : \alpha_i \neq 0 \text{ für mindestens zwei } i \in \mathbb{N}_p.$$

- Die Null- und Alternativhypothesen bezüglich der Haupteffektparameter von Faktor B lauten

$$H_0^B : \beta_j = 0 \text{ für alle } j = 1, \dots, q \text{ und } H_1^B : \beta_j \neq 0 \text{ für mindestens zwei } j \in \mathbb{N}_q.$$

- Die Null- und Alternativhypothesen bezüglich der Interaktionsparameter lauten

$$H_0^{A \times B} : \gamma_{ij} = 0 \text{ für alle } i = 1, \dots, p, j = 1, \dots, q$$

$$H_1^{A \times B} : \gamma_{ij} \neq 0 \text{ für mindestens zwei } (i, j) \in \mathbb{N}_p \times \mathbb{N}_q.$$

Die Interaktionsalternativhypothese macht keine Aussage über die Form der Interaktion.

- Die Alternativhypothesen sind jeweils zweimal ausgeprägt, da die Effektsummen nach Voraussetzung 0 sind.

Hypothesen und Überblick

- Primäre Ziele der zweifaktoriellen Varianzanalyse sind das Testen von Haupteffekt- und Interaktionshypothesen
- Die Null- und Alternativhypothesen bezüglich der Haupteffektparameter von Faktor A lauten

$$H_0^A : \alpha_i = 0 \text{ für alle } i = 1, \dots, p \text{ und } H_1^A : \alpha_i \neq 0 \text{ für mindestens zwei } i \in \mathbb{N}_p.$$

- Die Null- und Alternativhypothesen bezüglich der Haupteffektparameter von Faktor B lauten

$$H_0^B : \beta_j = 0 \text{ für alle } j = 1, \dots, q \text{ und } H_1^B : \beta_j \neq 0 \text{ für mindestens zwei } j \in \mathbb{N}_q.$$

- Die Null- und Alternativhypothesen bezüglich der Interaktionsparameter lauten

$$H_0^{A \times B} : \gamma_{ij} = 0 \text{ für alle } i = 1, \dots, p, j = 1, \dots, q$$

$$H_1^{A \times B} : \gamma_{ij} \neq 0 \text{ für mindestens zwei } (i, j) \in \mathbb{N}_p \times \mathbb{N}_q.$$

Die Interaktionsalternativhypothese macht keine Aussage über die Form der Interaktion.

- Die Alternativhypothesen sind jeweils zweimal ausgeprägt, da die Effektsummen nach Voraussetzung 0 sind.
- In Analogie zur einfaktoriellen Varianzanalyse können mithilfe entsprechend definierter F-Teststatistiken und ihrer Verteilungen kritische Wert-basierte Tests konstruiert werden. Die F-Teststatistiken beruhen auch im zweifaktoriellen Fall auf einer Quadratsummenzerlegung.

Theorem (Quadratsummenzerlegung bei zweifaktorieller Varianzanalyse)

Für $i = 1, \dots, p$, $j = 1, \dots, q$ und $k = 1, \dots, m$ bezeichne X_{ijk} die k te Stichprobenvariable bei Faktorlevelkombination ij eines zweifaktoriellen Varianzanalysemodells und es seien die Parameterschätzer $\hat{\mu}$, $\hat{\alpha}_i$, $\hat{\beta}_j$, $\hat{\gamma}_{ij}$, $\hat{\sigma}^2$ wie in der Parameterschätzung des Modells definiert gegeben.

Theorem (Quadratsummenzerlegung bei zweifaktorieller Varianzanalyse)

Für $i = 1, \dots, p, j = 1, \dots, q$ und $k = 1, \dots, m$ bezeichne X_{ijk} die k te Stichprobenvariable bei Faktorlevelkombination ij eines zweifaktoriellen Varianzanalysemodells und es seien die Parameterschätzer $\hat{\mu}, \hat{\alpha}_i, \hat{\beta}_j, \hat{\gamma}_{ij}, \hat{\sigma}^2$ wie in der Parameterschätzung des Modells definiert gegeben. Schließlich seien

$$\text{SQT} := \sum_{i=1}^p \sum_{j=1}^q \sum_{k=1}^m (X_{ijk} - \hat{\mu})^2 \quad \begin{array}{ll} \text{die Total} & \text{Sum of Squares} \end{array}$$

$$\text{SQA} := qm \sum_{i=1}^p \hat{\alpha}_i^2 \quad \begin{array}{ll} \text{die Faktor A} & \text{Sum of Squares} \end{array}$$

$$\text{SQB} := pm \sum_{j=1}^q \hat{\beta}_j^2 \quad \begin{array}{ll} \text{die Faktor B} & \text{Sum of Squares} \end{array}$$

$$\text{SQG} := m \sum_{i=1}^p \sum_{j=1}^q \hat{\gamma}_{ij}^2 \quad \begin{array}{ll} \text{die Interaktion} & \text{Sum of Squares} \end{array}$$

$$\text{SQR} := pq(m-1)\hat{\sigma}^2 \quad \begin{array}{ll} \text{die Residual} & \text{Sum of Squares} \end{array}$$

Theorem (Quadratsummenzerlegung bei zweifaktorieller Varianzanalyse)

Für $i = 1, \dots, p, j = 1, \dots, q$ und $k = 1, \dots, m$ bezeichne X_{ijk} die k te Stichprobenvariable bei Faktorlevelkombination ij eines zweifaktoriellen Varianzanalysemodells und es seien die Parameterschätzer $\hat{\mu}, \hat{\alpha}_i, \hat{\beta}_j, \hat{\gamma}_{ij}, \hat{\sigma}^2$ wie in der Parameterschätzung des Modells definiert gegeben. Schließlich seien

$$\text{SQT} := \sum_{i=1}^p \sum_{j=1}^q \sum_{k=1}^m (X_{ijk} - \hat{\mu})^2 \quad \begin{array}{ll} \text{die Total} & \text{Sum of Squares} \end{array}$$

$$\text{SQA} := qm \sum_{i=1}^p \hat{\alpha}_i^2 \quad \begin{array}{ll} \text{die Faktor A} & \text{Sum of Squares} \end{array}$$

$$\text{SQB} := pm \sum_{j=1}^q \hat{\beta}_j^2 \quad \begin{array}{ll} \text{die Faktor B} & \text{Sum of Squares} \end{array}$$

$$\text{SQG} := m \sum_{i=1}^p \sum_{j=1}^q \hat{\gamma}_{ij}^2 \quad \begin{array}{ll} \text{die Interaktion} & \text{Sum of Squares} \end{array}$$

$$\text{SQR} := pq(m-1)\hat{\sigma}^2 \quad \begin{array}{ll} \text{die Residual} & \text{Sum of Squares} \end{array}$$

Dann gilt

$$\text{SQT} = \text{SQA} + \text{SQB} + \text{SQG} + \text{SQR}. \quad (42)$$

- Wir verzichten auf einen Beweis.

Definition (F-Teststatistiken für Haupteffekte)

Es seien das Modell der zweifaktoriellen Varianzanalyse sowie die Faktor A Sum of Squares SQA, die Faktor B Sum of Squares SQB und die Residual Sum of Squares SQR wie oben definiert. Dann sind die F-Teststatistiken für die Haupteffekte von Faktor A und Faktor B definiert als

$$F_A := \frac{SQA/(p-1)}{SQR/(pq(m-1))} \text{ und } F_B := \frac{SQB/(q-1)}{SQR/(pq(m-1))} \quad (43)$$

respektive. Die Zähler und Nenner dieser Definitionen werden auch als *Mean SQA*, *Mean SQB*, *Mean SQR* bezeichnet.

Definition (F-Teststatistiken für Haupteffekte)

Es seien das Modell der zweifaktoriellen Varianzanalyse sowie die Faktor A Sum of Squares SQA, die Faktor B Sum of Squares SQB und die Residual Sum of Squares SQR wie oben definiert. Dann sind die F-Teststatistiken für die Haupteffekte von Faktor A und Faktor B definiert als

$$F_A := \frac{SQA/(p-1)}{SQR/(pq(m-1))} \text{ und } F_B := \frac{SQB/(q-1)}{SQR/(pq(m-1))} \quad (43)$$

respektive. Die Zähler und Nenner dieser Definitionen werden auch als *Mean SQA*, *Mean SQB*, *Mean SQR* bezeichnet.

Theorem (Verteilungen der F-Teststatistiken für Haupteffekte)

Es seien das Modell der zweifaktoriellen Varianzanalyse sowie die F-Teststatistiken F_A und F_B für die Haupteffekte von Faktor A und Faktor B wie oben definiert. Es gelte zudem

$$H_0^A : \alpha_i = 0 \text{ für alle } i = 1, \dots, p \text{ bzw. } H_0^B : \beta_j = 0 \text{ für alle } j = 1, \dots, q. \quad (44)$$

Dann sind die F-Teststatistiken nach f -Verteilungen mit Freiheitsgradparametern $p-1, pq(m-1)$ bzw. $q-1, pq(m-1)$ verteilt,

$$F_A \sim f(p-1, pq(m-1)) \text{ bzw. } F_B \sim f(q-1, pq(m-1)) \quad (45)$$

- Wir verzichten auf einen Beweis.

Definition (F-Teststatistik für Interaktionseffekte)

Es seien das Modell der zweifaktoriellen Varianzanalyse sowie die Interaktions Sum of Squares SQG und die Residual Sum of Squares SQR wie oben definiert. Dann ist die F-Teststatistik für Interaktionseffekte definiert als

$$F_{A \times B} := \frac{SQG/(p-1)(q-1)}{SQR/(pq(m-1))}. \quad (46)$$

Die Zähler und Nenner dieser Definitionen werden auch als *Mean SQG* und *Mean SQR* bezeichnet.

Definition (F-Teststatistik für Interaktionseffekte)

Es seien das Modell der zweifaktoriellen Varianzanalyse sowie die Interaktions Sum of Squares SQG und die Residual Sum of Squares SQR wie oben definiert. Dann ist die F-Teststatistik für Interaktionseffekte definiert als

$$F_{A \times B} := \frac{SQG/(p-1)(q-1)}{SQR/(pq(m-1))}. \quad (46)$$

Die Zähler und Nenner dieser Definitionen werden auch als *Mean SQG* und *Mean SQR* bezeichnet.

Theorem (Verteilung der F-Teststatistik für Interaktionseffekte)

Es seien das Modell der zweifaktoriellen Varianzanalyse sowie die F-Teststatistik $F_{A \times B}$ für Interaktionseffekte wie oben definiert. Es gelte zudem

$$H_0^{A \times B} : \gamma_{ij} = 0 \text{ für alle } i = 1, \dots, p, j = 1, \dots, q \quad (47)$$

Dann ist die F-Teststatistik nach einer f -Verteilung mit Freiheitsgradparametern $(p-1)(q-1), pq(m-1)$ verteilt,

$$F_{A \times B} \sim f((p-1)(q-1), pq(m-1)). \quad (48)$$

- Wir verzichten auf einen Beweis.

Theorem (F-Tests und F-Testsumfangkontrolle)

Es sei das Modell der zweifaktoriellen Varianzanalyse gegeben und die F-Teststatistiken F_A , F_B und $F_{A \times B}$ wie oben definiert. Weiterhin seien die kritischen Wert-basierten Tests

$$\phi_A := 1_{\{F_A \geq k_A\}}, \phi_B := 1_{\{F_B \geq k_B\}} \text{ und } \phi_{A \times B} := 1_{\{F_{A \times B} \geq k_{A \times B}\}} \quad (49)$$

definiert.

Theorem (F-Tests und F-Testumfangkontrolle)

Es sei das Modell der zweifaktoriellen Varianzanalyse gegeben und die F-Teststatistiken F_A , F_B und $F_{A \times B}$ wie oben definiert. Weiterhin seien die kritischen Wert-basierten Tests

$$\phi_A := 1_{\{F_A \geq k_A\}}, \phi_B := 1_{\{F_B \geq k_B\}} \text{ und } \phi_{A \times B} := 1_{\{F_{A \times B} \geq k_{A \times B}\}} \quad (49)$$

definiert. Dann sind ϕ_A , ϕ_B und $\phi_{A \times B}$ genau dann Level- α_0 -Tests mit Testumfang α_0 , wenn

$$\begin{aligned} k_A &:= k_{\alpha_0}^A := \varphi^{-1}(1 - \alpha_0; p - 1, pq(m - 1)) \\ k_B &:= k_{\alpha_0}^B := \varphi^{-1}(1 - \alpha_0; q - 1, pq(m - 1)) \end{aligned} \quad (50)$$

und

$$k_{A \times B} := k_{\alpha_0}^{A \times B} := \varphi^{-1}(1 - \alpha_0; (p - 1)(q - 1), pq(m - 1)), \quad (51)$$

respektive.

- Wir verzichten auf einen Beweis.

Praktisches Vorgehen

- Man hat einen Datensatz $x_{ijk}, i = 1, \dots, p, j = 1, \dots, q, k = 1, \dots, m$ eines zweifaktoriellen Versuchsplans mit crossed design vorliegen und nimmt an, dass die x_{ijk} Realisationen von $X_{ijk} \sim N(\mu_{ij}, \sigma^2)$ mit unbekannten Parametern μ_{ij} und σ^2 und entsprechenden Haupteffekt und Interaktionsparametern sind.

Praktisches Vorgehen

- Man hat einen Datensatz x_{ijk} , $i = 1, \dots, p$, $j = 1, \dots, q$, $k = 1, \dots, m$ eines zweifaktoriellen Versuchsplans mit crossed design vorliegen und nimmt an, dass die x_{ijk} Realisationen von $X_{ijk} \sim N(\mu_{ij}, \sigma^2)$ mit unbekannten Parametern μ_{ij} und σ^2 und entsprechenden Haupteffekt und Interaktionsparametern sind.
- Man möchte entscheiden ob die Haupteffekts- und Interaktionsnullhypothesen H_0^A , H_0^B und $H_0^{A \times B}$ eher zutreffen oder eher nicht.

Praktisches Vorgehen

- Man hat einen Datensatz x_{ijk} , $i = 1, \dots, p$, $j = 1, \dots, q$, $k = 1, \dots, m$ eines zweifaktoriellen Versuchsplans mit crossed design vorliegen und nimmt an, dass die x_{ijk} Realisationen von $X_{ijk} \sim N(\mu_{ij}, \sigma^2)$ mit unbekannten Parametern μ_{ij} und σ^2 und entsprechenden Haupteffekt und Interaktionsparametern sind.
- Man möchte entscheiden ob die Haupteffekts- und Interaktionsnullhypothesen H_0^A , H_0^B und $H_0^{A \times B}$ eher zutreffen oder eher nicht.
- Man wählt ein Signifikanzniveau α_0 und bestimmt die zugehörigen Freiheitsgradparameter-abhängigen kritischen Werte k_A , k_B und $k_{A \times B}$.

Praktisches Vorgehen

- Man hat einen Datensatz x_{ijk} , $i = 1, \dots, p$, $j = 1, \dots, q$, $k = 1, \dots, m$ eines zweifaktoriellen Versuchsplans mit crossed design vorliegen und nimmt an, dass die x_{ijk} Realisationen von $X_{ijk} \sim N(\mu_{ij}, \sigma^2)$ mit unbekannten Parametern μ_{ij} und σ^2 und entsprechenden Haupteffekt und Interaktionsparametern sind.
- Man möchte entscheiden ob die Haupteffekts- und Interaktionsnullhypothesen H_0^A , H_0^B und $H_0^{A \times B}$ eher zutreffen oder eher nicht.
- Man wählt ein Signifikanzniveau α_0 und bestimmt die zugehörigen Freiheitsgradparameter-abhängigen kritischen Werte k_A , k_B und $k_{A \times B}$.
- Anhand der SQA, SQB, SQG und SQR sowie p , q und n berechnet man die realisierten Werte der F-Teststatistiken, die wir hier mit $\tilde{f}_A$, $\tilde{f}_B$ und $\tilde{f}_{A \times B}$ bezeichnen.

Praktisches Vorgehen

- Man hat einen Datensatz x_{ijk} , $i = 1, \dots, p$, $j = 1, \dots, q$, $k = 1, \dots, m$ eines zweifaktoriellen Versuchsplans mit crossed design vorliegen und nimmt an, dass die x_{ijk} Realisationen von $X_{ijk} \sim N(\mu_{ij}, \sigma^2)$ mit unbekannten Parametern μ_{ij} und σ^2 und entsprechenden Haupteffekt und Interaktionsparametern sind.
- Man möchte entscheiden ob die Haupteffekts- und Interaktionsnullhypothesen H_0^A , H_0^B und $H_0^{A \times B}$ eher zutreffen oder eher nicht.
- Man wählt ein Signifikanzniveau α_0 und bestimmt die zugehörigen Freiheitsgradparameter-abhängigen kritischen Werte k_A , k_B und $k_{A \times B}$.
- Anhand der SQA, SQB, SQG und SQR sowie p , q und n berechnet man die realisierten Werte der F-Teststatistiken, die wir hier mit $\tilde{f}_A$, $\tilde{f}_B$ und $\tilde{f}_{A \times B}$ bezeichnen.
- Wenn für $i \in \{A, B, A \times B\}$ $\tilde{f}_i$ größer gleich $k_{\alpha_0}^i$ ist, lehnt man die jeweilige Nullhypothese ab, sonst nicht.

Praktisches Vorgehen

- Man hat einen Datensatz x_{ijk} , $i = 1, \dots, p$, $j = 1, \dots, q$, $k = 1, \dots, m$ eines zweifaktoriellen Versuchsplans mit crossed design vorliegen und nimmt an, dass die x_{ijk} Realisationen von $X_{ijk} \sim N(\mu_{ij}, \sigma^2)$ mit unbekannten Parametern μ_{ij} und σ^2 und entsprechenden Haupteffekt und Interaktionsparametern sind.
- Man möchte entscheiden ob die Haupteffekts- und Interaktionsnullhypothesen H_0^A , H_0^B und $H_0^{A \times B}$ eher zutreffen oder eher nicht.
- Man wählt ein Signifikanzniveau α_0 und bestimmt die zugehörigen Freiheitsgradparameter-abhängigen kritischen Werte k_A , k_B und $k_{A \times B}$.
- Anhand der SQA, SQB, SQG und SQR sowie p , q und n berechnet man die realisierten Werte der F-Teststatistiken, die wir hier mit $\tilde{f}_A$, $\tilde{f}_B$ und $\tilde{f}_{A \times B}$ bezeichnen.
- Wenn für $i \in \{A, B, A \times B\}$ $\tilde{f}_i$ größer gleich $k_{\alpha_0}^i$ ist, lehnt man die jeweilige Nullhypothese ab, sonst nicht.
- Die oben entwickelte Theorie garantiert dann, dass man im Mittel in höchstens $\alpha_0 \cdot 100$ von 100 Fällen die jeweilige Nullhypothese fälschlicherweise ablehnt.

Praktisches Vorgehen

- Man hat einen Datensatz x_{ijk} , $i = 1, \dots, p$, $j = 1, \dots, q$, $k = 1, \dots, m$ eines zweifaktoriellen Versuchsplans mit crossed design vorliegen und nimmt an, dass die x_{ijk} Realisationen von $X_{ijk} \sim N(\mu_{ij}, \sigma^2)$ mit unbekannten Parametern μ_{ij} und σ^2 und entsprechenden Haupteffekt und Interaktionsparametern sind.
- Man möchte entscheiden ob die Haupteffekts- und Interaktionsnullhypothesen H_0^A , H_0^B und $H_0^{A \times B}$ eher zutreffen oder eher nicht.
- Man wählt ein Signifikanzniveau α_0 und bestimmt die zugehörigen Freiheitsgradparameter-abhängigen kritischen Werte k_A , k_B und $k_{A \times B}$.
- Anhand der SQA, SQB, SQG und SQR sowie p , q und n berechnet man die realisierten Werte der F-Teststatistiken, die wir hier mit $\tilde{f}_A$, $\tilde{f}_B$ und $\tilde{f}_{A \times B}$ bezeichnen.
- Wenn für $i \in \{A, B, A \times B\}$ $\tilde{f}_i$ größer gleich $k_{\alpha_0}^i$ ist, lehnt man die jeweilige Nullhypothese ab, sonst nicht.
- Die oben entwickelte Theorie garantiert dann, dass man im Mittel in höchstens $\alpha_0 \cdot 100$ von 100 Fällen die jeweilige Nullhypothese fälschlicherweise ablehnt.
- Schließlich ergeben sich die assoziierten p-Werte der realisierten F-Teststatistiken in Analogie zum einfaktoriellen Varianzanalysefall zu

$$\begin{array}{lll} p_A\text{-Wert} & = \mathbb{P}\left(F_A \geq \tilde{f}_A\right) & = 1 - \varphi\left(\tilde{f}_A; p - 1, pq(m - 1)\right) \\ p_B\text{-Wert} & = \mathbb{P}\left(F_B \geq \tilde{f}_B\right) & = 1 - \varphi\left(\tilde{f}_B; q - 1, pq(m - 1)\right) \\ p_{A \times B}\text{-Wert} & = \mathbb{P}\left(F_{A \times B} \geq \tilde{f}_{A \times B}\right) & = 1 - \varphi\left(\tilde{f}_{A \times B}; (p - 1)(q - 1), pq(m - 1)\right) \end{array}$$

Anwendungsbeispiel

	x_{11k}	x_{12k}	x_{21k}	x_{22k}
1	2.86	5.04	3.29	6.88
2	2.62	4.96	3.69	7.33
3	3.01	4.58	2.94	6.76
4	3.20	5.05	3.61	6.57
5	2.75	5.05	3.68	6.84
6	2.88	5.01	3.89	7.45
7	2.85	4.35	3.43	7.06
8	3.23	5.07	3.72	6.99
9	2.99	5.84	3.28	6.93
10	2.57	4.61	3.35	7.46
11	2.69	5.26	2.94	7.07
12	2.83	5.17	3.56	7.03

x_{ijk} bezeichnet den Datenpunkt der k ten experimentellen Einheit ($k = 1, \dots, m$)
im i ten Level von Faktor A und j ten Level von Faktor B ($i = 1, 2, j = 1, 2$)

Anwendungsbeispiel - Manuelle Berechnung

```
# Analyseparameter
p           = 2
q           = 2
m           = 12
n           = 2*2*12
alpha_0     = 0.05

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

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

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

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

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

# \mu Schätzer
# \alpha_i Schätzer
# \beta_j Schätzer
# \gamma_{ij} Schätzer

# \sigma^2 Schätzers

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

Anwendungsbeispiel - Manuelle Berechnung

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

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

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

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

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

Anwendungsbeispiel - Manuelle Berechnung

Zweifaktorielle Varianzanalyse

Anzahl Level Faktor A = 2
Anzahl Level Faktor B = 2
Gruppengrößen = 12
Gesamtstichprobengröße = 48
 α_0 = 0.05

Haupteffekt Faktor A
SQA = 20.4
MSA = 20.4
F_A = 233
 k_{A,α_0} = 4.06
 ϕ_A = 1
p_A-Wert = 0

Haupteffekt Faktor B
SQB = 97.9
MSB = 97.9
F_B = 1118
 k_{B,α_0} = 4.06
 ϕ_B = 1
p_B-Wert = 0

Interaktion A x B
SQG = 6.37
MSG = 6.37
F_AxB = 72.8
 k_{AxB,α_0} = 4.06
 ϕ_{AxB} = 1
p_AxB-Wert = 7.07e-11

Anwendungsbeispiel - Automatische Berechnung

```
XAOV = data.frame(X = c(X$A1B1, X$A1B2, X$A2B1, X$A2B2), # Reformatierung des Datensatzes
                  A = as.factor(c(rep(1,2*m), rep(2,2*m))), # Faktor A Level Labels
                  B = as.factor(c(rep(1,m), rep(2,m), rep(1,m), rep(2,m)))) # Faktor B Level Labels

res.aov = aov(X ~ A + B + A:B, data = XAOV) # Benutzung von R's aov Funktion
summary(res.aov)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
A	1	20.4	20.4	233.1	< 2e-16 ***
B	1	97.9	97.9	1117.8	< 2e-16 ***
A:B	1	6.4	6.4	72.8	7.1e-11 ***
Residuals	44	3.9	0.1		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Einfaktorielle Varianzanalyse

- Anwendungsszenario und Datendeskription
- Modellformulierung
- Modellschätzung
- Modellevaluation

Zweifaktorielle Varianzanalyse

- Anwendungsszenario und Datendeskription
- Modellformulierung
- Modellschätzung
- Modellevaluation

Selbstkontrollfragen

Einfaktorielle Varianzanalysen

1. Erläutern Sie das Anwendungsszenario für einfaktorielle Varianzanalysen.
2. Definieren Sie das klassische Modell der einfaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung.
3. Definieren Sie das generative Modell der einfaktoriellen Varianzanalyse in Effektdarstellung.
4. Erläutern Sie die Beziehung zwischen dem klassischen Modell der einfaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung und dem generativen Modell der einfaktoriellen Varianzanalyse in Effektdarstellung.
5. Definieren Sie Null- und Alternativhypothese der einfaktoriellen Varianzanalyse.
6. Definieren Sie die Total Sum of Squares, die Between-Group Sum of Squares, und die Within-Group Sum of Squares der einfaktoriellen Varianzanalyse.
7. Erläutern Sie die intuitive Bedeutung der Total Sum of Squares, der Between-Group Sum of Squares, und der Within-Group Sum of Squares in der einfaktoriellen Varianzanalyse.
8. Geben Sie die Quadratsummenzerlegung der einfaktoriellen Varianzanalyse wieder.
9. Definieren Sie die F-Teststatistik der einfaktoriellen Varianzanalyse.
10. Geben Sie das Theorem zur Verteilung der F-Teststatistik bei Zutreffen der Nullhypothese einer einfaktoriellen Varianzanalyse wieder.
11. Definieren Sie den Level- α_0 -Test der einfaktoriellen Varianzanalyse.
12. Erläutern Sie das praktische Vorgehen bei Durchführung einer einfaktoriellen Varianzanalyse.

Selbstkontrollfragen

Zweifaktorielle Varianzanalysen

1. Erläutern Sie das Anwendungsszenario für zweifaktorielle Varianzanalysen.
2. Erläutern Sie die Begriffe Faktor und Faktorlevel bei zweifaktoriellen Varianzanalysen.
3. Wieviele Level besitzt der zweite experimentelle Faktor (Faktor B) einer 2×4 ANOVA?
4. Erläutern Sie die intuitive Bedeutung von Haupt- und Interaktionseffekten einer 2×2 ANOVA.
5. Definieren Sie die klassische Modell der zweifaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung.
6. Definieren Sie das klassische Modell der zweifaktoriellen Varianzanalyse in Effektdarstellung.
7. Definieren Sie das generative Modell der zweifaktoriellen Varianzanalyse in Effektdarstellung.
8. Erläutern Sie die Beziehung zwischen dem klassischen Modell der zweifaktoriellen Varianzanalyse in Erwartungswertparameterdarstellung und dem klassischen Modell der zweifaktoriellen Varianzanalyse in Effektdarstellung.
9. Definieren Sie Null- und Alternativhypothesen einer zweifaktoriellen Varianzanalyse für Haupteffektparameter.
10. Definieren Sie Null- und Alternativhypothesen einer zweifaktoriellen Varianzanalyse für Interaktionsparameter.
11. Definieren Sie die F-Teststatistiken einer zweifaktoriellen Varianzanalyse für Haupteffekte.
12. Geben Sie das Theorem zur Verteilung der F-Teststatistiken für Haupteffekte bei Zutreffen der Nullhypothesen einer zweifaktoriellen Varianzanalyse wieder.
13. Definieren Sie die F-Teststatistiken einer zweifaktoriellen Varianzanalyse für Interaktionseffekte.
14. Geben Sie das Theorem zur Verteilung der F-Teststatistiken für Interaktionseffekte bei Zutreffen der Nullhypothese einer zweifaktoriellen Varianzanalyse wieder.
15. Definieren Sie die Level- α_0 -Tests einer zweifaktoriellen Varianzanalyse.
16. Erläutern Sie das praktische Vorgehen bei Durchführung einer zweifaktoriellen Varianzanalyse.

References

- DeGroot, Morris H., and Mark J. Schervish. 2012. *Probability and Statistics*. 4th ed. Boston: Addison-Wesley.
- Fahrmeir, Ludwig, Christian Heumann, Rita Künstler, Iris Pigeot, and Gerhard Tutz. 2016. *Statistik*. Springer-Lehrbuch. Berlin, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-50372-0>.