Check for updates

Zweifaktorielle Varianzanalyse

Inhaltsverzeichnis

B. Rasch et al., *Quantitative Methoden 2*, https://doi.org/10.1007/978-3-662-63284-0_2

Lernziele

- Was versteht man unter Faktoren, Faktorstufen und Stufenkombinationen?
- Welche Effektarten kommen bei der zweifaktoriellen ANOVA vor und wie werden sie statistisch geprüft?
- Wie berechne ich die Effektgröße, die Teststärke und den optimalen Stichprobenumfang in einer zweifaktoriellen ANOVA?
- Was ist eine Wechselwirkung?
- Welche Formen der Wechselwirkung gibt es, und wie interpretiere ich sie?
- Welche Arten des Post-hoc-Vergleichs gibt es für die zweifaktorielle ANOVA, und wie wende ich sie an?
- Welchen Einfluss hat eine Erhöhung der Faktorenzahl auf die Ergebnisse einer mehrfaktoriellen ANOVA?

Im vorherigen ▶ Kap. 1 haben wir den Einfluss eines Faktors auf eine abhängige Variable untersucht. Menschliches Verhalten wird aber fast immer von mehr als einem Faktor bestimmt. Deshalb interessiert Sozialwissenschaftler oft die Wirkung von nicht nur einem, sondern mehreren Faktoren auf eine abhängige Variable. Von ganz besonderem Interesse ist ein mögliches Zusammenwirken der betrachteten Faktoren, das als Wechselwirkung zwischen zwei Faktoren bezeichnet wird. Weiterhin kann das Hinzuziehen eines zusätzlichen Faktors versuchsplanerische Gründe haben: Bei gegebenem systematischen Einfluss erklärt ein zusätzlicher Faktor einen Teil der Residualvarianz. Durch die Verkleinerung des unaufgeklärten Varianzanteils hebt sich ein möglicher Effekt des eigentlich interessierenden Faktors stärker heraus. Außerdem ist die gleichzeitige Betrachtung mehrerer Faktoren ökonomisch, da die Erhebung von Versuchspersonen oft sehr aufwendig ist.

Für das in den vorangehenden Kapiteln betrachtete Gedächtnisexperiment (vgl. Einleitung, Band 1) wäre ein denkbarer zweiter Faktor z. B. das Geschlecht: Gibt es einen generellen Unterschied in der Erinnerungsleistung zwischen Frauen und Männern? Oder reagieren Frauen nur in einer bestimmten Stufe der Verarbeitungstiefe, etwa in der strukturellen Verarbeitung, anders als Männer? Derartige Problemstellungen lassen sich mit der zweifaktoriellen Varianzanalyse bearbeiten und auswerten.

Die meisten Erkenntnisse aus ▶ Kap. 1 können wir in dieses Kapitel übertragen. Neu ist allerdings das Konzept der Wechselwirkung zwischen zwei Faktoren. Höhere mehrfaktorielle Varianzanalysen sind nicht mehr Teil dieses Buches. Sie lassen sich aber aufgrund der Ähnlichkeit mit den hier vorgestellten Konzepten verstehen.

Zu Beginn dieses Kapitels erfolgt eine kurze Darstellung der Bezeichnungen in einer zweifaktoriellen ANOVA. Auch sie ist in Teilen schon bekannt aus dem vorangehenden ▶ Kap. 1.

2.1 Nomenklatur

In einer zweifaktoriellen Varianzanalyse wird der erste betrachtete Faktor mit dem Buchstaben A, der zweite Faktor mit B gekennzeichnet. Welche inhaltliche Variable diesen Buchstaben zugeordnet wird, ist gleichgültig. Die Angabe der Stufenanzahl der Faktoren erfolgt durch zwei unterschiedliche Indizes:

- Faktor A hat p Stufen (Laufindex i)
- Faktor B hat q Stufen (Laufindex j)

Eine Beschreibung der verwendeten Auswertungsmethode in der Literatur beinhaltet immer auch die Angabe der Stufenanzahl der betrachteten Faktoren. Bei einer zweifaktoriellen ANOVA lautet die Bezeichnung:

- p×q-Varianzanalyse

Jeder dieser p×q Stufenkombinationen oder Zellen werden n Versuchspersonen zugeordnet. Das kleine n gibt die Anzahl der Versuchspersonen pro Zelle an. Ist n in jeder Zelle gleich groß, so ergibt sich die Gesamtanzahl der Versuchspersonen zu $N = p \cdot q \cdot n$.

In einer Tabelle einer zweifaktoriellen Varianzanalyse stehen üblicherweise nicht alle Werte der einzelnen Versuchspersonen, sondern es werden nur die Mittelwerte der Zellen angegeben. Beim Umgang mit dieser Art der Darstellung ist immer zu bedenken, dass die Zahlen in den Zellen Mittelwerte sind und sich hinter diesen Mittelwerten immer die Werte mehrerer Versuchspersonen verbergen. Die ◘ Tab. 2.1 zeigt die allgemeine Darstellung einer zweifaktoriellen Varianzanalyse. Zum Aufbau: Faktor A wird im Gegensatz zum einfaktoriellen Fall (siehe z. B. ◘ Tab. 1.3) nun als Zeilenfaktor eingeführt, Faktor B kommt als Spaltenfaktor hinzu. Für die Nummerierung der Indizes gilt das Prinzip „Zeile vor Spalte".

Bezeichnungen der Mittelwerte:

- Mittelwerte der einzelnen Stufen des Faktors A: $\bar{A}_i$
- Mittelwerte der Stufen des Faktors B: $\bar{B}_j$
- Zellmittelwerte der Kombination der Faktorstufen: $\overline{AB}_{ij}$

Zur Veranschaulichung betrachten wir wieder den Beispieldatensatz des Gedächtnisexperiments aus ► Abschn. 1.2. Zusätzlich zu dem Faktor „Verarbeitungstiefe" fügen wir den Faktor „Geschlecht" hinzu. Der Faktor „Verarbeitungstiefe" ist als Faktor A gekennzeichnet und hat $p = 3$ Stufen. Der Faktor „Geschlecht" (B) hat $q = 2$ Stufen (männlich/weiblich). Die verwendete Auswertungsmethode ist demzufolge eine 3×2-Varianzanalyse. Um die Darstellung einfach zu halten, ordnen wir jeder Zelle nur 2 Versuchspersonen zu. Die Größe der Gesamtstichprobe ist also $N = p \cdot q \cdot n = 3 \cdot 2 \cdot 2 = 12$. Die Versuchspersonen erzielten die folgenden Werte (◘ Tab. 2.2):

Die allgemeine Schreibweise einer 3×2-Varianzanalyse angewandt auf das Zahlenbeispiel der vorangehenden Seite zeigt ◘ Tab. 2.3.

Hinter jedem Zellmittelwert verbergen sich in diesem Beispiel jeweils die Werte von zwei Versuchspersonen.

◘ **Tab. 2.1** Allgemeine Darstellung von Zellmittelwerten in einer zweifaktoriellen Varianzanalyse

Bedingung	**B_1**	**B_2**	**…**	**B_q**	
A_1	$\overline{AB}_{11}$	$\overline{AB}_{12}$	…	$\overline{AB}_{1q}$	$\bar{A}_1$
A_2	$\overline{AB}_{21}$	$\overline{AB}_{22}$	…	$\overline{AB}_{2q}$	$\bar{A}_2$
…	…	…	…	…	…
A_p	$\overline{AB}_{p1}$	$\overline{AB}_{p2}$	…	$\overline{AB}_{pq}$	$\bar{A}_p$
	$\bar{B}_1$	$\bar{B}_2$	…	$\bar{B}_q$	$\bar{G}$

2

Tab. 2.2 Messwerte des Beispieldatensatzes des Erinnerungsexperiments mit $N = 12$

Bedingung	Männlich	Weiblich	Mittelwerte
Strukturell	7 8	6 7	**7**
Bildhaft	10 11	11 12	**11**
Emotional	11 12	12 13	**12**
Mittelwerte	**9,83**	**10,17**	**10**

Tab. 2.3 Darstellung der Zellmittelwerte einer 3×2-Varianzanalyse mit den Beispieldaten

Bedingung	B_1: männlich	B_2: weiblich	Mittelwerte
A_1: strukturell	$\overline{AB}_{11} = 7{,}5$	$\overline{AB}_{12} = 6{,}5$	$\bar{A}_1 = 7$
A_2: bildhaft	$\overline{AB}_{21} = 10{,}5$	$\overline{AB}_{22} = 11{,}5$	$\bar{A}_2 = 11$
A_3: emotional	$\overline{AB}_{31} = 11{,}5$	$\overline{AB}_{32} = 12{,}5$	$\bar{A}_3 = 12$
Mittelwerte	$\bar{B}_1 = 9{,}83$	$\bar{B}_2 = 10{,}17$	$\bar{G} = 10$

Die Tab. 2.3 macht deutlich, dass es in einer zweifaktoriellen Varianzanalyse viele Mittelwerte gibt, deren Unterschiede auf Signifikanz überprüft werden müssen. Zum einen können sich die Gruppenmittelwerte des Faktors A unterscheiden. Ebenso ist ein Unterschied der Gruppenmittelwerte des Faktors B denkbar. Und drittens ist eine Prüfung des zusätzlichen Zusammenwirkens spezieller Stufen der beiden Faktoren nötig. Eine solche Wechselwirkung zwischen A und B zeigt sich in dem Vergleich der Unterschiede zwischen den Zellmittelwerten über die verschiedenen Stufen des Faktors A oder B (► Abschn. 2.2).

2.2 Effektarten

In der zweifaktoriellen Varianzanalyse gibt es drei Arten von Effekten:
- Haupteffekt A
- Haupteffekt B
- Wechselwirkung A×B

Die zweifaktorielle Varianzanalyse entspricht in ihrer Vorgehensweise der bereits bekannten einfaktoriellen ANOVA: Sie überprüft Mittelwertsunterschiede mit dem *F*-Test auf Signifikanz. Allerdings gibt es nun durch das Hinzufügen eines weiteren Faktors insgesamt drei Arten von Mittelwertsunterschieden (vgl. Tab. 2.1):

1. Mittelwertsunterschiede in den Stufen des Faktors A ($\bar{A}_i$)
2. Mittelwertsunterschiede in den Stufen des Faktors B ($\bar{B}_j$)
3. Mittelwertsunterschiede in den Bedingungskombinationen ($\overline{AB}_{ij}$), die nicht durch 1. und 2. erklärbar sind.

Bei einer zweifaktoriellen Varianzanalyse lassen sich deshalb drei Effekte unterscheiden: der Haupteffekt A, der Haupteffekt B und die Wechselwirkung der beiden Faktoren A×B.

Haupteffekt A

Der Haupteffekt A kennzeichnet den Einfluss des Faktors A auf die abhängige Variable, unabhängig von Faktor B. Er beschreibt die Unterschiede zwischen den Stufenmittelwerten des Faktors A, gemittelt über die Stufen des Faktors B. Der Haupteffekt A entspricht also in der Interpretation und den Mittelwerten dem bekannten Haupteffekt A der einfaktoriellen Varianzanalyse aus ► Kap. 1. In unserem Beispiel ist es unabhängig vom Geschlecht weiterhin möglich, den Einfluss der Verarbeitungstiefe auf die Erinnerungsleistung zu betrachten. In ◘ Tab. 2.2 ist zu erkennen, dass sich an den Mittelwerten des Faktors A (Verarbeitungstiefe) im Vergleich zur einfaktoriellen ANOVA nichts geändert hat (► Abschn. 1.2).

Haupteffekt B

Der Haupteffekt B kennzeichnet den Einfluss des Faktors B auf die abhängige Variable, unabhängig von Faktor A. Er beschreibt die Unterschiede zwischen den Stufenmittelwerten des Faktors B, gemittelt über die Stufen des Faktors A. Dies entspricht einer einfaktoriellen Betrachtungsweise des Faktors B. Unabhängig vom Faktor „Verarbeitungstiefe" ist es nun zusätzlich möglich, die Erinnerungsleistung von Frauen und Männern zu vergleichen.

Wechselwirkung A×B

Die Wechselwirkung A×B oder Interaktion beschreibt den gemeinsamen Einfluss von bestimmten Stufen der zwei Faktoren auf die abhängige Variable. Sie erfasst das Zusammenwirken von Faktorstufen. Mit anderen Worten: Die Interaktion verzeichnet die Einflüsse auf die abhängige Variable, die nur durch die gemeinsame und gleichzeitige Wirkung zweier Faktorstufen entstehen und die nicht durch den generellen Einfluss der zwei Faktoren erklärt werden können. Das heißt, es wird überprüft, ob die Wirkung des Faktors A auf allen Stufen des Faktors B identisch ist oder nicht bzw. ob die Wirkung des Faktors B auf allen Stufen des Faktors A identisch ist oder nicht. Die Wechselwirkung ist deshalb unabhängig von den zwei Haupteffekten. Sie berechnet die Unterschiede zwischen den beobachteten Zellmittelwerten und den aufgrund der Haupteffekte erwarteten Werten.

Die Wechselwirkung beschreibt den Einfluss der Kombination bestimmter Stufen der Faktoren A und B, der nicht durch die Wirkung der Faktoren allein zu erklären ist.

Haupteffekt A, Haupteffekt B und Wechselwirkung sind voneinander unabhängig und können deshalb auch getrennt voneinander untersucht werden. Jeder Effekt kann allein oder in Kombination mit einem oder beiden anderen Effekten in einer Untersuchung auftreten.

2.2.1 Prüfung der Effekte

Die Prüfung der Effekte auf Signifikanz erfordert eine Zerlegung der Gesamtvarianz (◘ Abb. 2.1). In der zweifaktoriellen Varianzanalyse teilt sich die Gesamtvarianz ebenfalls in systematische Varianz und Residualvarianz auf. Allerdings wird die systematische Varianz nicht mehr allein durch einen Haupteffekt A erzeugt, sondern der Haupteffekt B und die Wechselwirkung A×B treten hinzu. Die systematische Varianz enthält also auf Populationsebene drei Komponenten bzw. Effektvarianzen:

$$\sigma^2_{\text{sys}} = \sigma^2_{\alpha} + \sigma^2_{\beta} + \sigma^2_{\alpha\times\beta}$$

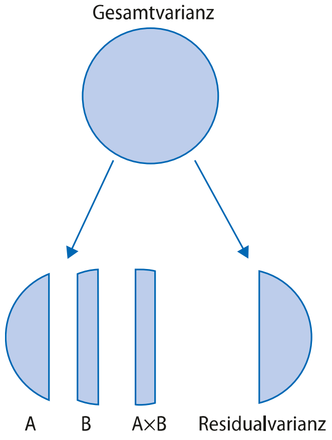

Abb. 2.1 Zerlegung der Gesamtvarianz in einer zweifaktoriellen Varianzanalyse ohne Messwiederholung

Um die drei in der systematischen Varianz enthaltenen Effekte getrennt voneinander untersuchen zu können, benötigen wir drei Schätzer. Wie bereits bei der einfaktoriellen Varianzanalyse ist die Residualvarianz in den einzelnen Schätzern der Effektvarianz mit enthalten. Es gibt neben der globalen Zwischenvarianz nun Zwischenvarianzen zu jedem der drei möglichen Effekte.

Zur Unterscheidung bezeichnen wir diese drei Zwischenvarianzen mit den jeweiligen Buchstaben der zugehörigen Effekte A, B oder A×B. In den Erwartungswerten ist jeweils die Effektvarianz zur Residualvarianz addiert. Die Erwartungswerte der geschätzten Varianzen lauten:

$$\text{Haupteffekt A}: E\left(\hat{\sigma}_{\mathrm{A}}^2\right) = n \cdot q \cdot \sigma_{\alpha}^2 + \sigma_{\varepsilon}^2$$

$$\text{Haupteffekt B}: E\left(\hat{\sigma}_{\mathrm{B}}^2\right) = n \cdot p \cdot \sigma_{\beta}^2 + \sigma_{\varepsilon}^2$$

$$\text{Wechselwirkung A} \times \text{B}: E\left(\hat{\sigma}_{\mathrm{A\times B}}^2\right) = n \cdot \sigma_{\alpha \times \beta}^2 + \sigma_{\varepsilon}^2$$

Die drei geschätzten Zwischenvarianzen enthalten neben der Effektvarianz immer auch Residualvarianz.

Die Prüfvarianz wird wie bei der einfaktoriellen Varianzanalyse aus der „Varianz innerhalb“ bzw. Residualvarianz geschätzt.

$$\text{Prüfvarianz}: E\left(\hat{\sigma}_{Res}^2\right) = \sigma_{\varepsilon}^2$$

Die Zwischenvarianzen werden einzeln an der geschätzten Residualvarianz auf Signifikanz geprüft.

Die Signifikanzprüfung der einzelnen Effekte verläuft nach dem bereits bekannten Grundprinzip der *F*-Bruch-Bildung (▶ Abschn. 1.2.7): Der *F*-Bruch muss so gebildet werden, dass sich der Erwartungswert der Varianz im Zähler nur in dem gesuchten Effekt von dem Erwartungswert der Nennervarianz unterscheidet. So lässt sich mithilfe der oben vorgestellten Erwartungswerte der Varianzen für jeden der drei Effekte ein *F*-Bruch konstruieren, der dieser Bedingung genügt (für die Herleitung vgl. ▶ Abschn. 1.2.7).

$$\text{Haupteffekt A}: F_{\mathrm{A}(df_{\mathrm{A}};df_{\mathrm{Res}})} = \frac{\hat{\sigma}_A^2}{\hat{\sigma}_{\mathrm{Res}}^2} \text{ mit } df_{\mathrm{A}} = p - 1$$

$$\text{Haupteffekt B}: F_{\mathrm{B}(df_{\mathrm{B}};df_{\mathrm{Res}})} = \frac{\hat{\sigma}_{\mathrm{B}}^2}{\hat{\sigma}_{\mathrm{Res}}^2} \text{ mit } df_{\mathrm{B}} = q - 1$$

$$\text{Wechselwirkung A} \times \text{B}: F_{\mathrm{A\times B}(df_{\mathrm{A\times B}};df_{\mathrm{Res}})} = \frac{\hat{\sigma}_{\mathrm{A\times B}}^2}{\hat{\sigma}_{\mathrm{Res}}^2} \text{ mit } df_{\mathrm{A\times B}} = (p-1)\cdot(q-1)$$

Die Nennerfreiheitsgrade berechnen sich zu: $df_{Res} = p \cdot q \cdot (n-1)$

Tabellen E im Anhang A2 von Band 1

Die Prüfung auf Signifikanz erfolgt anhand der Tabellen E im Anhang A2 von Band 1, in denen die Wahrscheinlichkeiten der jeweiligen *F*-Werte unter Annahme der Nullhypothese in Abhängigkeit von Zähler- und Nennerfreiheitsgraden abgetragen sind. Der folgende Abschnitt erläutert ausführlich die genaue Durchführung sowie die Berechnung der einzelnen Varianzen. Er behandelt die einzelnen Tests der drei Effekte in einer zweifaktoriellen Varianzanalyse, ihre Nullhypothesen, die Berechnung der Varianzen mithilfe der Quadratsummen und ihren Freiheitsgraden sowie die Bestimmung der Wahr-

B., Friese, M., Hofmann, W. & Naumann, E. (2021). Quantitative Methoden 2. Einführung in die Statistik für Psychologie, Sozial- & Erziehungswissenschaften (5. Aufl.).

scheinlichkeit der F-Werte. Als illustrierendes Beispiel dient uns wieder das bekannte Gedächtnisexperiment.

2.2.2 Die Schätzung der Residualvarianz

Auch in der zweifaktoriellen Varianzanalyse schätzt die „Varianz innerhalb" die Residualvarianz. Sie ist wie bei der einfaktoriellen ANOVA ein Maß für die mittlere quadrierte Abweichung der einzelnen Personenwerte von ihrem Gruppenmittelwert. Sie berechnet sich aus der Abweichung jedes einzelnen Werts von seinem jeweiligen Zellmittelwert. Diese Varianzen werden addiert und durch die Anzahl der Zellen geteilt, also ihr Mittelwert gebildet. Das Ergebnis heißt „Varianz innerhalb" bzw. geschätzte Residualvarianz.

$$\hat{\sigma}^2_{\text{Res}} = \hat{\sigma}^2_{\text{innerhalb}} = \frac{\hat{\sigma}^2_{11} + \hat{\sigma}^2_{12} + \ldots + \hat{\sigma}^2_{1q} + \hat{\sigma}^2_{21} + \ldots + \hat{\sigma}^2_{pq}}{p \cdot q}$$

Die Residualvarianz in der Population wird durch die durchschnittliche Varianz innerhalb der Bedingungskombinationen geschätzt.

Nach der mathematischen Voraussetzung der Varianzhomogenität sollten die Residualvarianzen in jeder Zelle bzw. Bedingungskombination theoretisch gleich groß sein. Da sie das empirisch meistens nicht genau sind, wird hier zur Schätzung der Residualvarianz die durchschnittliche Varianz innerhalb der Zellen gebildet.

Wie ergeben sich die Freiheitsgrade der geschätzten Residualvarianz in einer zweifaktoriellen Varianzanalyse? Jede einzelne Residualvarianz innerhalb einer Zelle hat $n - 1$ Freiheitsgrade, d. h., $n - 1$ Summanden können pro Zellvarianz frei variieren. Insgesamt hat die geschätzte Residualvarianz deshalb folgende Freiheitsgrade:

$$df_{\text{Res}} = p \cdot q \cdot (n-1)$$

Auch in diesem Kapitel sollen analog zu ► Kap. 1 die Berechnungen der einzelnen Varianzen über die Quadratsummen vorgestellt werden. Zuerst erfolgt die Aufspaltung der einzelnen Residualvarianzen aus der obigen Formel in Quadratsummen und Freiheitsgrade. Unter der Annahme, dass die Anzahl der Versuchspersonen in jeder Gruppe gleich ist, lässt sich $n - 1$ mit in den Nenner aufnehmen:

$$\hat{\sigma}^2_{\text{Res}} = \frac{\frac{QS_{11}}{n-1} + \frac{QS_{12}}{n-1} + \ldots + \frac{QS_{pq}}{n-1}}{p \cdot q} = \frac{\sum_{i=1}^{p} \sum_{j=1}^{q} QS_{ij}}{p \cdot q \cdot (n-1)}$$

Durch Einsetzen der Formel der Quadratsumme ergibt sich die Berechnung der geschätzten Residualvarianz:

$$\hat{\sigma}^2_{\text{Res}} = \frac{\sum_{i=1}^{p} \sum_{j=1}^{q} \sum_{m=1}^{n} \left(x_{ijm} - \overline{AB}_{ij} \right)^2}{p \cdot q \cdot (n-1)} = \frac{QS_{Res}}{df_{Res}}$$

Schätzung der Residualvarianz über ihre Quadratsumme und Freiheitsgrade

Der Zähler dieses Bruchs, die QS_{Res}, drückt aus, dass von jedem Messwert sein jeweiliger Zellmittelwert abgezogen wird. Das geschieht innerhalb jeder Zelle, also jeder Kombination der Faktorstufen. Zur Veranschaulichung berechnen wir die geschätzte Residualvarianz aus obigem Beispiel. Die Werte entstammen den ◘ Tab. 2.2 und 2.3:

$$\hat{\sigma}^2_{\text{Res}} = \frac{(7-7{,}5)^2 + (8-7{,}5)^2 + (6-6{,}5)^2 + \ldots + (13-12{,}5)^2}{3 \cdot 2 \cdot (2-1)}$$

$$\hat{\sigma}^2_{\text{Res}} = \frac{0{,}25 + 0{,}25 + 0{,}25 + \ldots + 0{,}25}{6} = 0{,}5$$

Die Freiheitsgrade belaufen sich zu $df_{Res} = 3 \cdot 2 \cdot (2-1) = 6$

2.2.3 Prüfung des Haupteffekts A auf Signifikanz

Der vom Faktor B unabhängige systematische Einfluss des Faktors A auf die abhängige Variable heißt Haupteffekt A. Die Nullhypothese dieses Haupteffekts lautet bei p Stufen des Faktors A (analog zur einfaktoriellen ANOVA): Alle Populationsmittelwerte der Stufen des Faktors A (gemittelt über die Stufen des Faktors B) sind gleich:

Hypothesenpaar zum Haupteffekt A

$$\text{H}_0: \mu_1 = \mu_2 = \ldots = \mu_p$$

$$\text{H}_1: \neg H_0$$

Die Hypothesen lassen sich auch über Varianzen ausdrücken:

$$\text{H}_0: \sigma^2_\alpha = 0$$

$$\text{H}_1: \sigma^2_\alpha > 0$$

F-Bruch für den Haupteffekt A

Die Alternativhypothese ist wie bei allen Varianzanalysen zweiseitig formuliert, da Varianzen aufgrund ihrer Quadrierung nur positive Werte annehmen können und so keine Aussage über die Richtung der Abweichungen zulassen. Der *F*-Bruch für die Prüfung des Haupteffekts A ist (► Abschn. 2.2.1):

$$F_{\text{A}(df_{\text{A}};df_{\text{Res}})} = \frac{\hat{\sigma}^2_{\text{A}}}{\hat{\sigma}^2_{\text{Res}}}$$

Trifft die Nullhypothese zu, so sollte der *F*-Wert den Wert eins annehmen, da die geschätzte Varianz des Haupteffekts A unter dieser Annahme nur aus Residualvarianz besteht. Allerdings können die Mittelwerte zufällig variieren und so die Zwischenvarianz des Faktors A erhöhen. Die *F*-Verteilung der Nullhypothese gibt Auskunft, wie wahrscheinlich ein empirischer *F*-Wert unter Annahme der Nullhypothese ist. Ist er hinreichend unwahrscheinlich, wird die Nullhypothese verworfen und die Alternativhypothese angenommen. Der Haupteffekt A ist signifikant. Ein signifikanter Haupteffekt A bedeutet, dass der Faktor A (mit einer tolerierten Fehlerwahrscheinlichkeit) in irgendeiner Weise einen systematischen Einfluss auf die abhängige Variable hat.

Die Berechnung im Einzelnen:

$\hat{\sigma}^2_{\text{A}}$ resultiert aus der QS_{A} und ihren Freiheitsgraden:

Schätzung der Zwischenvarianz des Faktors A über die Quadratsummen und Freiheitsgrade

$$\hat{\sigma}_{\mathrm{A}}^{2} = \frac{QS_{\mathrm{A}}}{\mathrm{d}f_{\mathrm{A}}}$$

Die Quadratsumme des Faktors A entspricht der quadrierten Abweichung der Mittelwerte in den Stufen des Faktors A vom Gesamtmittelwert, multipliziert mit *n* und der Anzahl der Bedingungen des Faktors B:

$$QS_{\mathrm{A}} = \sum_{i=1}^{p} n \cdot q \cdot \left(\bar{A}_i - \bar{G}\right)^2$$

Die Freiheitsgrade ergeben sich aus der Anzahl der Summanden der QS_{A} minus eins. Die Anzahl der Summanden entspricht der Anzahl der Stufen des Faktors A:

$$df_{\mathrm{A}} = p - 1$$

In unserem Beispiel hat der Faktor A „Verarbeitungstiefe" $p = 3$ Stufen. Wir wollen die Frage prüfen, ob die Verarbeitungsbedingung einen systematischen Einfluss auf die Erinnerungsleistung hat. Die abhängige Variable ist die Anzahl erinnerter Wörter.

Die Nullhypothese behauptet, dass die Versuchspersonen in allen drei Bedingungen im Mittel die gleiche Anzahl Wörter erinnern. Die Alternativhypothese nimmt an, dass sie in den drei Bedingungen unterschiedlich viele Wörter erinnern.

Das Hypothesenpaar des Haupteffekts A, ausgedrückt in Mittelwerten, lautet deshalb:

$$\mathrm{H}_0 : \mu_{strukturell} = \mu_{bildhaft} = \mu_{emotional}$$

$$\mathrm{H}_1 : \neg H_0$$

Die Hypothesen, ausgedrückt in Varianzen:

$$\mathrm{H}_0 : \sigma_{\alpha}^2 = 0$$

$$\mathrm{H}_1 : \sigma_{\alpha}^2 > 0$$

Mit diesen Informationen lässt sich die Zwischenvarianz für den Haupteffekt A berechnen:

$$\hat{\sigma}_{\mathrm{A}}^{2} = \frac{QS_{\mathrm{A}}}{df_{\mathrm{A}}} = \frac{\sum_{i=1}^{p} n \cdot q \cdot \left(\bar{A}_i - \bar{G}\right)^2}{p-1}$$

$$\hat{\sigma}_{\mathrm{A}}^{2} = \frac{2 \cdot 2 \cdot \left[(7-10)^2 + (11-10)^2 + (12-10)^2\right]}{3-1} = 28$$

Die Freiheitsgrade ergeben sich zu $df_{\mathrm{A}} = p - 1 = 2$.

Die geschätzte Residualvarianz in unserem Beispiel beträgt:

2

$\hat{\sigma}^2_{\text{Res}} = 0{,}5$ (Rechnung in ▶ Abschn. 2.2.2)
Es resultiert folgender *F*-Bruch:

$$F_{\text{A}(2;6)} = \frac{\hat{\sigma}^2_{\text{A}}}{\hat{\sigma}^2_{\text{Res}}} = \frac{28}{0{,}5} = 56$$

Die Wahrscheinlichkeit, dass ein *F*-Wert von $F_{(2;6)} = 56$ unter Gültigkeit der Nullhypothese auftritt, ist $p < 0{,}001$. Diese Wahrscheinlichkeit ist kleiner als das von uns a priori festgesetzte α-Niveau von 5 %. Oder anders betrachtet: Der empirische *F*-Wert ist größer als der kritische *F*-Wert von $F_{(2;6)} = 5{,}14$. Das Ergebnis ist signifikant. Wir können die Nullhypothese verwerfen und die Alternativhypothese annehmen: Die Verarbeitungstiefe hat einen (irgendwie gearteten) Einfluss auf die Anzahl der Wörter, die die Versuchspersonen erinnern. Es bleibt weiteren Analysen überlassen, die genaue Struktur des systematischen Einflusses von Faktor A zu untersuchen, über die die bloße Feststellung eines signifikanten Haupteffekts keinerlei Aussage macht.

Das Ergebnis entspricht nicht genau der in ▶ Kap. 1 ermittelten Wahrscheinlichkeit des Haupteffekts A aus der einfaktoriellen Varianzanalyse. Das liegt an der Veränderung der Größe der Residualvarianz beim Hinzuziehen eines weiteren Faktors. Die Zusammenhänge zwischen dem Einbezug mehrerer Faktoren in eine Auswertung und der Veränderung der Residualvarianz erläutert ▶ Abschn. 2.6. Die Zwischenvarianz des Faktors A ist trotz Hinzufügens weiterer Faktoren immer mit der „Varianz zwischen" im einfaktoriellen Fall identisch.

2.2.4 Prüfung des Haupteffekts B auf Signifikanz

Der vom Faktor A unabhängige systematische Einfluss des Faktors B auf die abhängige Variable heißt Haupteffekt B. Die Prüfung des Haupteffekts B erfolgt nach demselben Muster wie die im vorangehenden Abschnitt erläuterte Prüfung des Haupteffekts A. Die Nullhypothese dieses Haupteffekts lautet bei *q* Stufen des Faktors B: Alle Populationsmittelwerte der Stufen des Faktors B (gemittelt über die Stufen des Faktors A) sind gleich:

Hypothesenpaar zum Haupteffekt B

$$\text{H}_0: \mu_1 = \mu_2 = \ldots = \mu_q \text{ oder } \sigma^2_\beta = 0$$

$$\text{H}_1: \neg H_0 \text{ oder } \sigma^2_\beta > 0$$

Bildung des *F*-Bruchs:

F-Bruch für den Haupteffekt B

$$F_{B(df_B;df_{\text{Res}})} = \frac{\hat{\sigma}^2_B}{\hat{\sigma}^2_{\text{Res}}}$$

Die Berechnung der Zwischenvarianz des Faktors B durch die Quadratsumme und Freiheitsgrade verläuft nach folgendem Muster:

Schätzung der Zwischenvarianz des Haupteffekts B durch Quadratsumme und Freiheitsgrade

$$\hat{\sigma}^2_{\text{B}} = \frac{QS_{\text{B}}}{df_{\text{B}}}$$

$$QS_{\mathrm{B}} = \sum_{j=1}^{q} n \cdot p \cdot \left(\bar{B}_j - \bar{G}\right)^2 ; df_{\mathrm{B}} = q - 1$$

In unserem Beispiel hat der Faktor B „Geschlecht“ $q = 2$ Stufen. Gibt es einen bedeutsamen Unterschied in der Erinnerungsleistung zwischen Männern und Frauen? Die abhängige Variable ist ebenfalls die Anzahl erinnerter Wörter. Die Nullhypothese geht davon aus, dass Männer und Frauen im Durchschnitt gleich viele Wörter erinnern, die Alternativhypothese behauptet das Nichtzutreffen dieser Aussage.

$$\mathrm{H}_0 : \mu_{\text{männlich}} = \mu_{\text{weiblich}} \text{ oder } \sigma_\beta^2 = 0$$

$$\mathrm{H}_1 : \mu_{\text{männlich}} \neq \mu_{\text{weiblich}} \text{ oder } \sigma_\beta^2 > 0$$

Auch die Berechnung der Zwischenvarianz des Faktors B verläuft analog zu der von Faktor A im letzten Abschnitt:

$$\hat{\sigma}_{\mathrm{B}}^2 = \frac{QS_{\mathrm{B}}}{\mathrm{d}f_{\mathrm{B}}}$$

$$\hat{\sigma}_{\mathrm{B}}^2 = \frac{\sum_{j=1}^{q} n \cdot p \cdot \left(\bar{B}_j - \bar{G}\right)^2}{q-1} = \frac{2 \cdot 3 \cdot \left[(9{,}83 - 10)^2 + (10{,}17 - 10)^2\right]}{2-1} = 0{,}35$$

$$df_{\mathrm{B}} = q - 1$$

$$\hat{\sigma}_{\mathrm{Res}}^2 = 0{,}5$$

$$F_{\mathrm{B}(1;6)} = \frac{\hat{\sigma}_{\mathrm{B}}^2}{\hat{\sigma}_{\mathrm{Res}}^2} = \frac{0{,}35}{0{,}5} = 0{,}7$$

Der empirische F-Wert $F_{(1;6)} = 0{,}7$ ist kleiner als eins. Ein F-Wert von eins bedeutet, dass kein systematischer Einfluss des betrachteten Faktors vorliegt, die Zwischenvarianz besteht ausschließlich aus Residualvarianz (► Abschn. 1.2.5). In der Theorie ist ein F-Wert kleiner als eins nicht möglich. Durch das Arbeiten mit Schätzern und deren Ungenauigkeiten kann ein solches Ergebnis aber in der Empirie durchaus vorkommen, wie an der Verteilung der F-Werte in ► Abschn. 1.2.8 deutlich zu sehen ist. Ein empirischer F-Wert von kleiner eins ist aber per definitionem niemals statistisch signifikant. Der kritische F-Wert liegt bei einem Signifikanzniveau von 5 % bei $F_{krit(1;6)} = 5{,}99$. Trotz dieser deutlichen Nichtsignifikanz muss vor der Annahme der Nullhypothese die Berechnung der Teststärke erfolgen. Die Vorgehensweise dafür erläutert ► Abschn. 2.3.2.

2.2.5 Prüfung der Wechselwirkung A×B auf Signifikanz

Die Wechselwirkung A×B entspricht dem Effekt, der durch das Zusammenwirken bestimmter Stufen der beiden Faktoren auf die abhängige Variable