





Zweifaktorielle Varianzanalyse (ohne Messwiederholung)

- 1 Beispiel und Grundlagen
- 2 Effekte in Mittelwertsdiagrammen
- 3 Quadratsummenzerlegung
- 4 Modell
- 5 Testprozedur
- 6 Durchführung mittels SPSS

Einführende Literatur

-  Bortz, J. & Schuster, Ch. (2010). *Statistik für Human- und Sozialwissenschaftler* (7. Auflage). Berlin: Springer. [Kap. 14 & 15]
-  Diehl, J. M. & Arbinger, R. (2001). *Einführung in die Inferenzstatistik* (3. Auflage). Eschborn bei Frankfurt: Klotz Verlag. [Kap. 12 & 13]
-  Eid, M., Gollwitzer, M. & Schmitt, M. (2010). *Statistik und Forschungsmethoden*. Weinheim: Beltz. [Kap. 13.2]

Weiterführende Literatur

-  Keppel, G. & Wickens, T. D. (2004). *Design and analysis. A researchers handbook* (4th ed.). Upper Saddle River, NJ: Pearson. [Kap. 10-14]

Zweifaktorielle Varianzanalyse: Beispiel

➤ **Beispiel 1:** In einer Studie werden erneut die drei Lehrmethoden „konventionelles Lehrbuch“, „verzweigtes Lehrbuch“ und „verzweigte Computeranwendung im Web“ auf ihre Wirksamkeit hin untersucht, diesmal zur Vermittlung von Physikwissen. 60 Schüler werden den drei Lehrmethoden per Zufall zugewiesen. Zusätzlich werden die Schüler aufgrund ihrer früheren Physiknoten zusätzlich in solche mit hohem und niedrigem Leistungsniveau unterteilt. Am Ende des Unterrichts wird ein Wissenstest durchgeführt, in dem die Schüler max. 90 Punkte erzielen können.

➤ Es interessiert wieder die Frage, ob sich die Lehrmethoden in ihrer Wirksamkeit unterscheiden. Zudem interessiert, ob Studierende mit vorherigem hohem und niedrigem Leistungsniveau auch in diesem Wissenstest unterschiedlich abschneiden. Ferner lässt sich betrachten, ob die Lehrmethoden bei den Schülern mit hohem und niedrigem Leistungsniveau unterschiedlich wirksam sind.

| UV ₂ : Leistungs niveau | UV ₁ : Lehrmethode | | | | | |
|--|-------------------------------|----|-----------------------|----|---------------------------|----|
| | Lehrbuch konventionell | | Lehrbuch verzweigt | | Web-gestützt verzweigt | |
| niedrig | 34 | 35 | 50 | 62 | 61 | 69 |
| | 27 | 22 | 64 | 47 | 55 | 64 |
| | 25 | 27 | 49 | 53 | 52 | 47 |
| | 44 | 33 | 51 | 45 | 49 | 51 |
| | 16 | 56 | 36 | 68 | 44 | 72 |
| hoch | 85 | 60 | 60 | 73 | 45 | 55 |
| | 72 | 42 | 51 | 67 | 54 | 61 |
| | 75 | 41 | 52 | 88 | 72 | 59 |
| | 78 | 48 | 50 | 77 | 52 | 48 |
| | 65 | 38 | 40 | 68 | 79 | 49 |

Zweifaktorielle Varianzanalyse

- Wir untersuchen nun den Einfluss von **zwei unabhängigen Variablen (Faktoren)** simultan auf **eine abhängige Variable**. Im Beispiel weist der erste Faktor A drei Faktorstufen ($m = 3$) auf und der zweite Faktor B zwei Faktorstufen ($q = 2$).
- Dabei betrachten wir zunächst den Fall (wie im Beispiel), bei dem ...
 - die Faktorstufen vollständig „gekreuzt“ sind, d.h. alle möglichen $m \cdot q$ Kombinationen von Faktorstufen untersucht werden. Im Beispiel gibt es also 3×2 verschiedene Untersuchungsbedingungen (=Zellen des Versuchsplans), unter denen jeweils V_{pn} untersucht werden. Man spricht dann von einem **faktoriellen (Versuchs-) Plan (Design)**. Die Auswertung erfolgt später über eine Varianzanalyse, die auch als **3×2 ANOVA** bezeichnet wird (bzw. 2×3 ANOVA, da die Reihenfolge der Faktoren ohne Bedeutung ist).
 - beide Faktoren feste Faktoren darstellen.
 - die V_{pn} zufällig gezogen und den Zellen des Versuchsplans per Zufall zugewiesen werden. Die Stichproben (sechs Gruppen) in allen Zellen sind also unabhängig.
 - die Zahl der Personen in allen Zellen gleich ist.

Zweifaktorielle Varianzanalyse

- **Vor-** und **Nachteile** des zweifaktoriellen gegenüber dem einfaktoriellen Design:
 - (+) Es ermöglicht in der gleichen Studie die Untersuchung einer weiteren Einflussgröße (ökonomisch),
 - (-) was allerdings einen größeren versuchsplanerischen Aufwand und häufig auch eine größere Zahl an Versuchspersonen erfordert.
 - (+) Ein Teil der Varianz, die im einfaktoriellen Design Fehlervarianz (QS_{in}) darstellt, kann hier (evtl.) durch den zweiten Faktor aufgeklärt werden. Die Power ist entsprechend für den ersten Effekt größer.
 - (+) Es kann geprüft werden, ob der Faktor seinen Effekt unabhängig von der Ausprägung des anderen Faktors ausübt. Das sagt also etwas über die Generalisierbarkeit des Effektes. Hängt die Wirkung des einen Faktors davon ab, welche Ausprägung der andere Faktor aufweist, so spricht man von einer **Wechselwirkung (Interaktion)** der beiden Faktoren. Neben dem Effekt des Faktors A und dem Effekt des Faktors B (**Haupteffekte**) können wir also auch die Wechselwirkung der beiden Faktoren (wir schreiben dafür AxB) untersuchen. Letzteres wäre nicht möglich, wenn wir den Einfluss des Faktors A und des Faktors B in zwei verschiedenen Studien untersuchen würden.

Zweifaktorielle Varianzanalyse

- Im Folgenden wollen wir uns anhand von **Mittelwertsdiagrammen (Interaktionsdiagramme)** die möglichen Muster beim Zusammenspiel der drei Effekte betrachten:
 - **Haupteffekt** von **Faktor A**: Hier unterscheiden sich die Mittelwerte unter den Faktorstufen von *A* (Spaltenmittelwerte), wobei man den Faktor *B* ignoriert. Eine Lehrmethode ist also z.B. allgemein besser als eine andere.
 - **Haupteffekt** von **Faktor B**: Hier unterscheiden sich die Mittelwerte unter den Faktorstufen von *B* (Zeilenmittelwerte), wobei man den Faktor *A* ignoriert. Schüler mit einem höheren Leistungsniveau schneiden also z.B. generell besser ab als solche mit einem niedrigeren Leistungsniveau.
 - **Wechselwirkung (Interaktion)** von **Faktor A** und **Faktor B**: Der Effekt des Faktors *A* ist auf verschiedenen Stufen des Faktors *B* unterschiedlich stark (bzw. umgekehrt). D.h. die Differenzen der Zellenmittelwerte des Faktors *A* sind unter den verschiedenen Stufen von *B* unterschiedlich (und umgekehrt).
- Alle drei Effekte können unabhängig voneinander auftreten: Es kann also z.B. nur einen Haupteffekt für Faktor *A* oder z.B. nur einen Haupteffekt für Faktor *B* plus die Interaktion geben usw.

Zweifaktorielle Varianzanalyse

- Man kann das wichtige Konzept der **Interaktion** auf verschiedene Art charakterisieren:
Eine Interaktion besteht dann, wenn ...
 - der Effekt einer UV auf den Stufen einer anderen UV unterschiedlich ausfällt.
 - die Unterschiede der Mittelwerte eines Faktors A unter den Stufen eines anderen Faktors B nicht gleich groß sind.
 - die Effekte einer UV konditional sind zur Ausprägung einer anderen UV.
 - sich die Ausprägungen in der AV nicht durch eine additive Wirkung der Faktoren (plus Stichprobenfehler) erklären lassen.

- Eine Interaktion bedeutet nicht, dass sich die UVs untereinander beeinflussen!!!

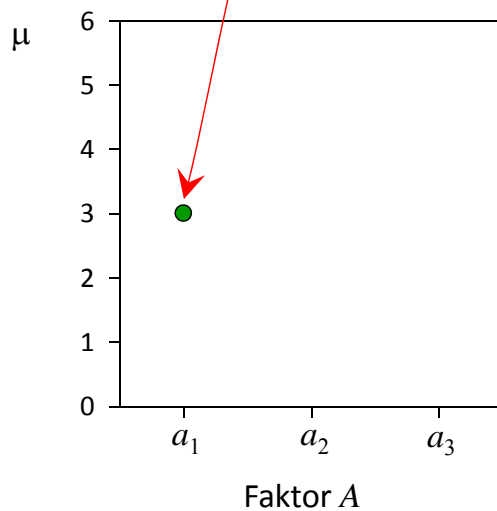
- 1 Beispiel und Grundlagen
- 2 Effekte in Mittelwertsdiagrammen
- 3 Quadratsummenzerlegung
- 4 Modell
- 5 Testprozedur
- 6 Durchführung mittels SPSS

Zweifaktorielle Varianzanalyse

- In den Mittelwertdiagrammen tragen wir alle $m \cdot q$ Zellenmittelwerte in der Weise ab, dass wir einen der beiden Faktoren (UVs) auf der Abszisse, den Mittelwert der AV auf der Ordinate und die Faktorstufen des zweiten Faktors als getrennte Profillinien abtragen.
- Wir gehen zunächst davon aus, dass es sich bei den Zellenmittelwerten um Mittelwerte in der Population μ handelt.

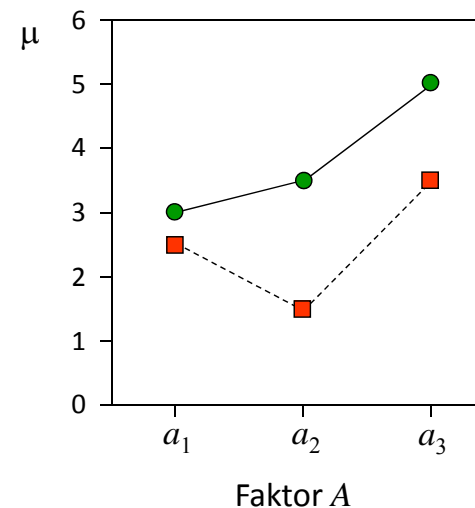
Mittelwert aller Vpn unter der Faktorkombination (a_1, b_1)

| | | Faktor A | | |
|----------|-------|----------|-------|-------|
| | | a_1 | a_2 | a_3 |
| Faktor B | b_1 | 3.0 | 3.5 | 5.0 |
| | b_2 | 2.5 | 1.5 | 3.5 |



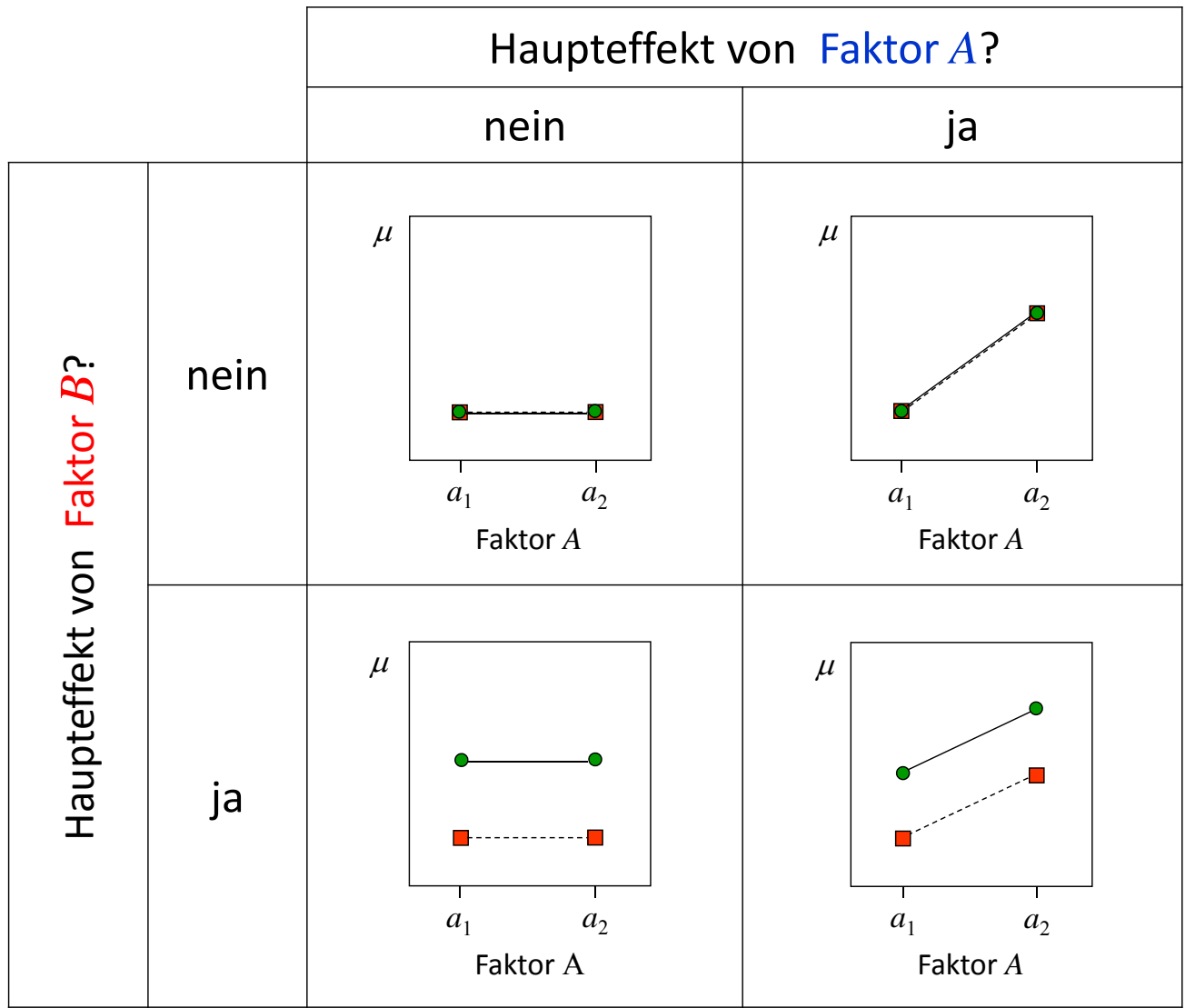
Faktor B

- — ● b_1
- - - - ■ b_2



Zweifaktorielle Varianzanalyse

Mögliche Haupteffekte einer 2 x 2 ANOVA ohne Wechselwirkungen



Abstand der beiden Kurven

Faktor B

- — ● b_1
- - - ■ b_2

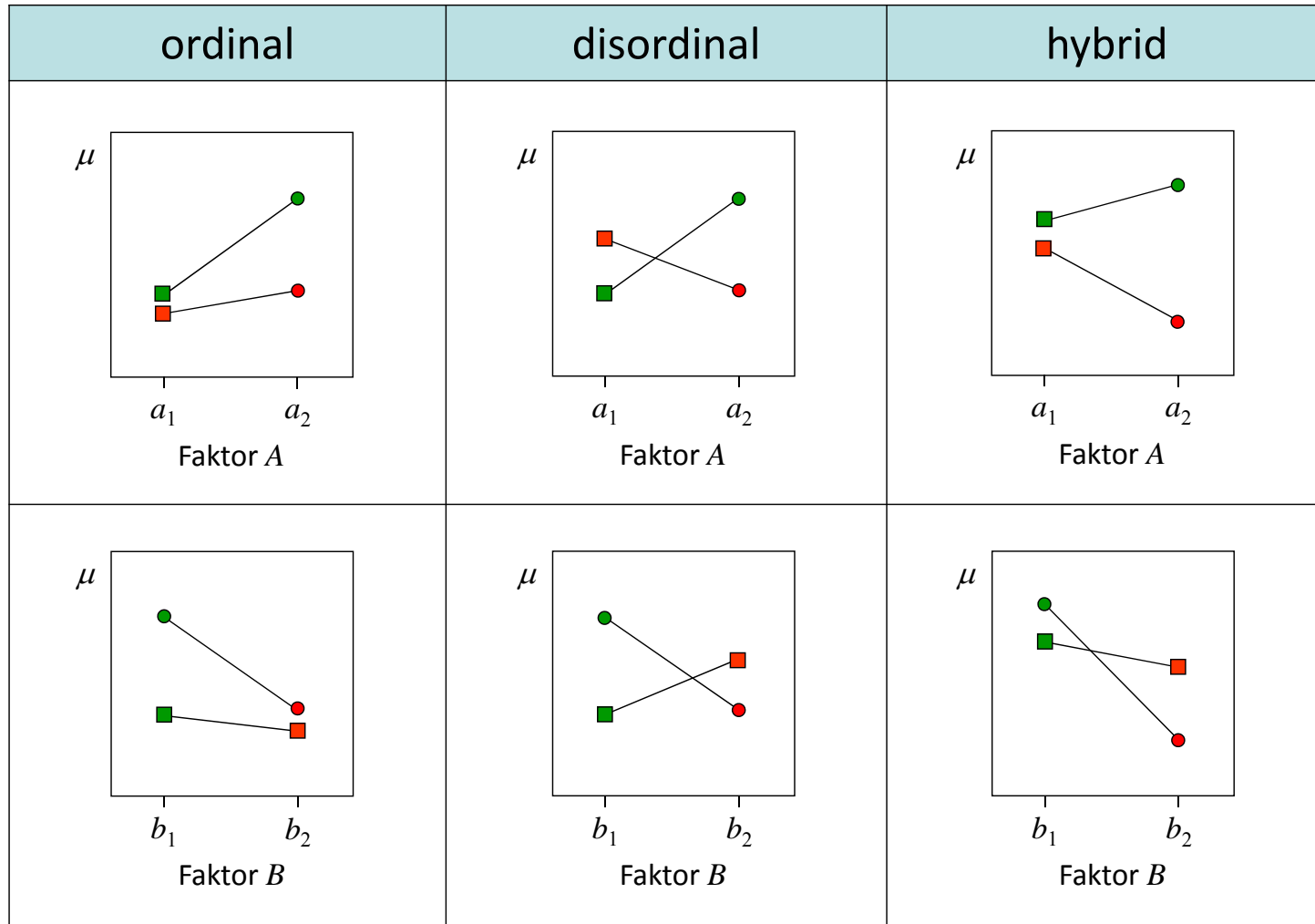
Ansteigen/Abfallen der Kurven

Zweifaktorielle Varianzanalyse

- Das Vorliegen einer Interaktion erkennt man (in der Population) daran, dass die Kurven nicht mehr überall parallel sind (sich „irgendwo“ schneiden).
- Man kann verschiedene Interaktionsformen unterscheiden, die man am besten identifiziert, indem man sich beide Mittelwertsdiagramme anschaut (d.h. das mit Faktor A und das mit Faktor B auf der Abszisse):
 - **Ordinale Interaktion:** Hier schneiden sich die Profilkurven in beiden Darstellungen nicht. Oder anders ausgedrückt: Die Mittelwerte des Faktors A haben unter den Faktorstufen von B immer die gleichen Rangordnung.
 - **Disordinale Interaktion:** Hier schneiden sich die Profilkurven in beiden Darstellungen. Die Rangordnung der Mittelwerte des Faktors A ist unter den Faktorstufen von B unterschiedlich und umgekehrt.
 - **Semidisordinale (=hybride) Interaktion:** Hier schneiden sich die Profilkurven in einer Darstellung und in der anderen nicht. Die Rangordnung der Mittelwerte des Faktors A ist unter den Faktorstufen von B unterschiedlich, nicht aber die Rangordnung des Faktors B unter den Stufen von A (oder umgekehrt).

Zweifaktorielle Varianzanalyse

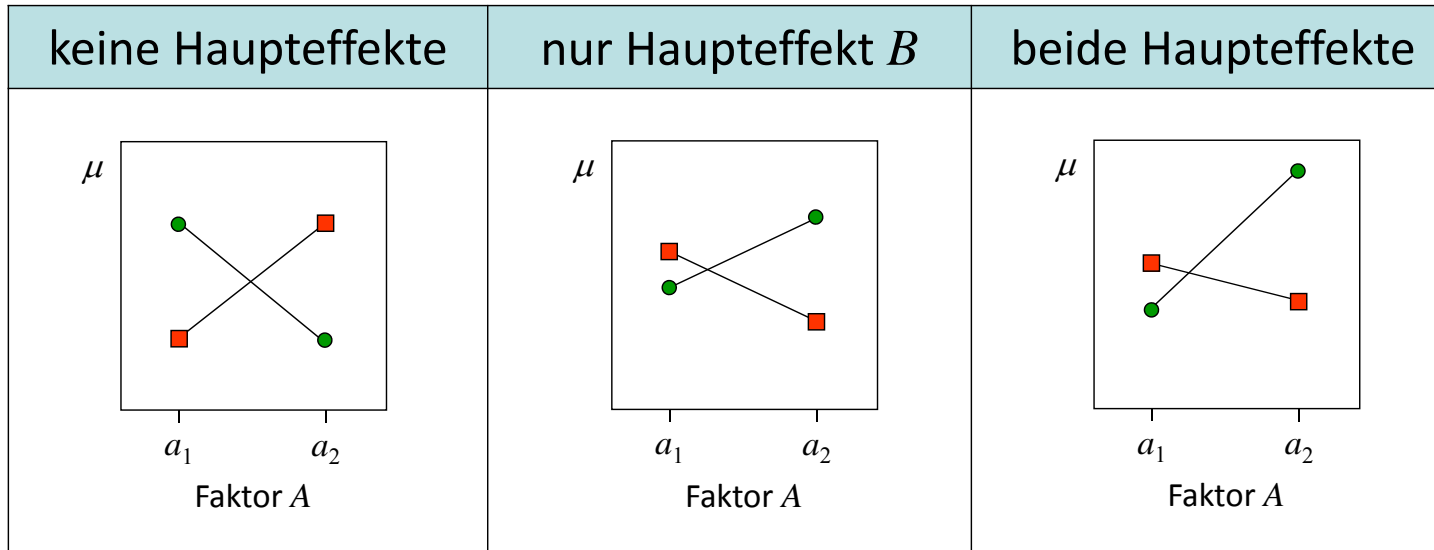
Faktor B
 ■—● b_1
 ■- - -● b_2



Formen von Interaktionen einer 2×2 ANOVA

- Bleiben die Rangordnungen der Zellenmittelwerte wie bei der ordinalen Interaktion gleich, so kann man die Interaktion durch eine monotone Transformation der AV zum Verschwinden bringen (removeable interaction), andernfalls nicht (crossover interaction).

Zweifaktorielle Varianzanalyse



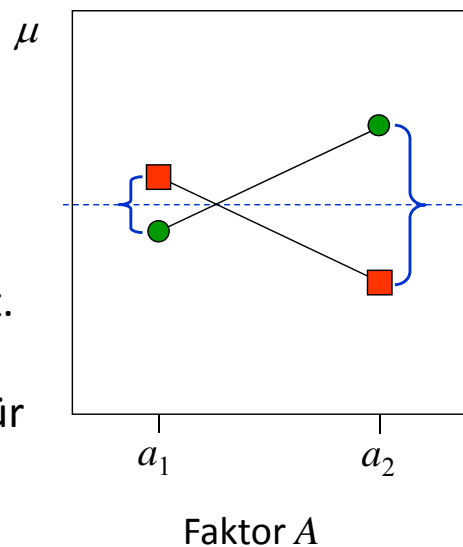
Disordinale Interaktionen in einer 2×2 ANOVA, gepaart mit unterschiedlichen Haupteffekten

Faktor B

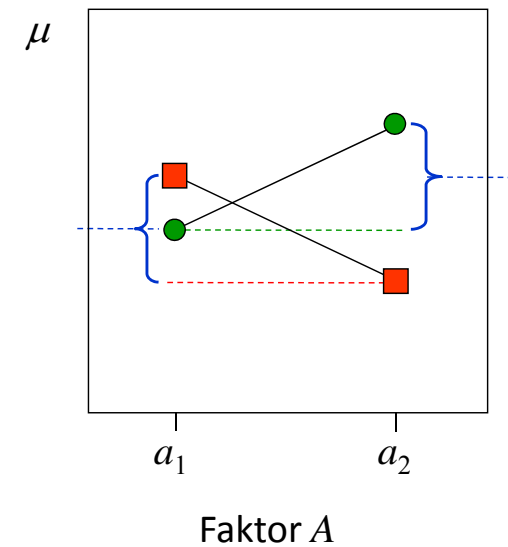
●—● b_1

■- - ■ b_2

Ein Haupteffekt für Faktor A liegt vor, wenn der Mittelwert beider Mittelwerte unter a_1 nicht mit dem unter a_2 übereinstimmt. Im Beispiel rechts liegt also kein Haupteffekt für Faktor A vor.



Ein Haupteffekt für Faktor B liegt vor, wenn der Mittelwert beider Mittelwerte unter b_1 nicht mit dem unter b_2 übereinstimmt. Im Beispiel rechts liegt also ein Haupteffekt für Faktor B vor.



Zweifaktorielle Varianzanalyse

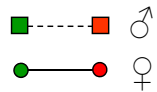
- Für die Deutung der Haupteffekte ist von Bedeutung, ob eine Wechselwirkung existiert und falls ja, welche Form diese aufweist. Im Allgemeinen bedeutet ja ein Haupteffekt, dass es Mittelwertsunterschiede in einem Faktor gibt, wenn man den anderen Faktor ignoriert (d.h. die Randmittel betrachtet). Er besagt also z.B., dass eine Lehrmethode E generell besser ist als eine andere D (d.h. z.B. sowohl bei Jungen als auch bei Mädchen) oder eine bestimmte psychologische Therapie F bei Phobien allgemein wirksamer ist als die Gabe eines Medikaments C oder gar keine Therapie (gleichgültig welche Phobie, wie stark die Phobie usw.)
- Liegt eine **ordinale** Interaktion vor, so können die Haupteffekte weiterhin in diesem Sinne interpretiert werden. Wenn man die Mittelwerte bezüglich des anderen Faktors getrennt betrachtet, so bleibt z.B. in allen Gruppen Lehrmethode E besser als D . Lediglich die Stärke der Überlegenheit von E gegenüber D ist dann z.B. bei Jungen und Mädchen unterschiedlich. Man würde also weiterhin immer Lehrmethode E empfehlen.
- Sind aber wie bei der **disordinalen** Interaktion die Rangfolgen der Zellenmittelwerte nicht in allen Stufen eines Faktors gleich, so gilt dies nicht mehr: Hier hängt nun z.B. die Überlegenheit einer Lehrmethode von den Ausprägungen des anderen Faktors ab: So könnte die Lehrmethode D bei Mädchen und die Lehrmethode E bei Jungen überlegen sein. Dies muss bei der Interpretation der Haupteffekte berücksichtigt werden, falls man diese überhaupt noch interpretieren will!

Zweifaktorielle Varianzanalyse

Faktor B



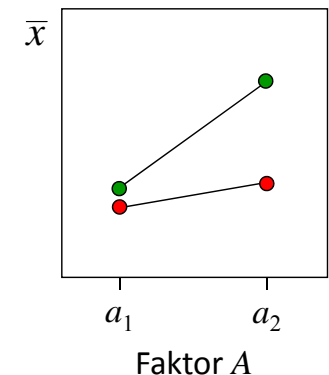
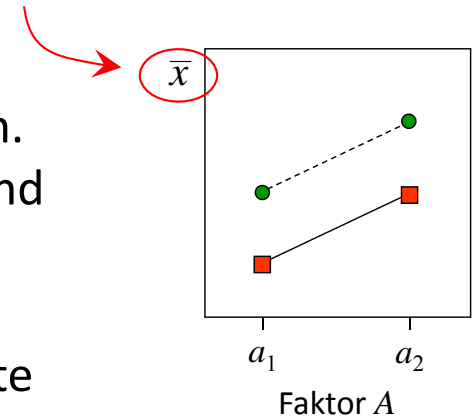
Faktor A



| | ordinal | disordinal | hybrid |
|--|---|---|---|
| | | | |
| | | | |
| Interpretation des Haupteffektes A Geschlecht ... | uneingeschränkt: Mädchen sind besser als Jungen. | eingeschränkt, da Jungen beim darbietenden Lernen und Mädchen beim entdeckenden Lernen besser sind. | eingeschränkt, da Jungen beim darbietenden Lernen und Mädchen beim entdeckenden Lernen besser sind. |
| Interpretation des Haupteffektes B Lehrmethode ... | uneingeschränkt: Entdecken des Lernen ist besser als darbietendes Lernen. | eingeschränkt, da beim darbietenden Lernen Jungen und bei entdeckenden Lernen Mädchen besser sind. | uneingeschränkt: Entdecken des Lernen ist besser als darbietendes Lernen. |

Zweifaktorielle Varianzanalyse

- Das Erkennen von Haupteffekten und Formen der Wechselwirkungen „mit bloßem Auge“ anhand der Mittelwertsdiagramme ist nur möglich, wenn wir die Populationsmittelwerte betrachten. Tatsächlich haben wir nur die entsprechenden Stichprobenstatistiken vorliegen.
- Betrachten wir also z.B. das Diagramm rechts, in dem nun die Stichprobenmittel abgetragen sind, so deutet dies auf zwei Haupteffekte hin. Tatsächlich kann es aber sein, dass ein oder beide Haupteffekte aufgrund von Stichprobenfehler zustande gekommen sind (bzw. nicht statistisch signifikant werden, weil die Effekte bei den vorliegenden Stichprobengrößen zu klein sind, als dass wir die Nullhypothese gleicher Mittelwerte auf dem gegebenen Signifikanzniveau zurückweisen können).
- Das gleiche gilt für das Mittelwertsdiagramm rechts, in dem z.B. in der Population auch keine Interaktion bestehen könnte oder die Interaktion nicht ordinal ausfällt, sondern sich die Kurven in der Population doch schneiden könnten (hybride oder disordinale Interaktion).
- Zu beachten ist auch, dass die Frage, ob sich zwei Kurven schneiden, auch von der Wahl des Wertebereichs der Faktoren abhängen kann (z.B. könnten sich die Kurven schneiden, wenn der Faktor A eine Medikamentendosierung darstellt und wir eine kleinere Dosierung als a_1 untersuchen würden).



- 1 Beispiel und Grundlagen
- 2 Effekte in Mittelwertsdiagrammen
- 3 **Quadratsummenzerlegung**
- 4 Modell
- 5 Testprozedur
- 6 Durchführung mittels SPSS

Zweifaktorielle Varianzanalyse

➤ Wir führen folgende Notation ein:

Faktor A

$j = 1 \dots m$ Faktorstufen (hier $m = 3$)

| | | Faktor A | | | | | |
|----------|---------|------------------------|-----|--------------------|-----|---------------|-----|
| | | Lehrbuch konventionell | | Lehrbuch verzweigt | | Web verzweigt | |
| Faktor B | niedrig | 34 | 35 | 50 | 62 | 61 | 69 |
| | | 27 | 22 | 64 | 47 | 55 | 64 |
| | | ... | ... | ... | ... | ... | ... |
| | hoch | 16 | 56 | x_{ijk} | 8 | 44 | 72 |
| | | 85 | 60 | 3 | 45 | 55 | ... |
| | | 72 | 42 | 51 | 67 | 54 | 61 |
| ... | ... | ... | ... | ... | ... | | |
| 65 | 38 | 40 | 68 | 79 | 49 | | |

Faktor B

$k = 1 \dots q$

Faktorstufen (hier $q = 2$)

Messwert der Person $i = 1$ in der Bedingung $j = 3$ und $k = 1$:
 $x_{131} = 61$.

$i = 1 \dots n_{jk}$ Personen in Zelle (j, k)
(hier $n_{\bullet\bullet} = n_{jk} = 10$ in allen Zellen)

$$n = \sum_{j=1}^m \sum_{k=1}^q n_{jk}$$

$$n = 60$$

$$n = m \cdot q \cdot n_{\bullet\bullet}$$

Zweifaktorielle Varianzanalyse

➤ Folgende Mittelwerte lassen sich bestimmen:

| | | Faktor A ($j = 1 \dots m$) | | | |
|-----------------------------|---------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|
| | | Lehrbuch konventionell | Lehrbuch verzweigt | Web verzweigt | |
| Faktor B $k = 1 \dots q$ | niedrig | $\bar{x}_{\cdot 11} = 31.9$ | $\bar{x}_{\cdot 21} = 52.5$ | $\bar{x}_{\cdot 31} = 56.4$ | $\bar{x}_{\cdot \cdot 1} = 46.933$ |
| | hoch | $\bar{x}_{\cdot 12} = 60.4$ | $\bar{x}_{\cdot 22} = 62.6$ | $\bar{x}_{\cdot 32} = 57.4$ | $\bar{x}_{\cdot \cdot 2} = 60.133$ |
| | | $\bar{x}_{\cdot \cdot 1} = 46.150$ | $\bar{x}_{\cdot \cdot 2} = 57.550$ | $\bar{x}_{\cdot \cdot 3} = 56.900$ | $\bar{x}_{\dots} = 53.533$ |

2 Zeilenmittel

$$\bar{x}_{\cdot \cdot k} = \frac{\sum_{i=1}^{n_{\cdot \cdot}} \sum_{j=1}^m x_{ijk}}{m \cdot n_{\cdot \cdot}}$$

6 Zellenmittel

$$\bar{x}_{\cdot \cdot jk} = \frac{\sum_{i=1}^{n_{\cdot \cdot}} x_{ijk}}{n_{\cdot \cdot}}$$

$i = 1 \dots n_{\cdot \cdot}$, d.h. gleicher Stichprobenumfang in allen Zellen

3 Spaltenmittel

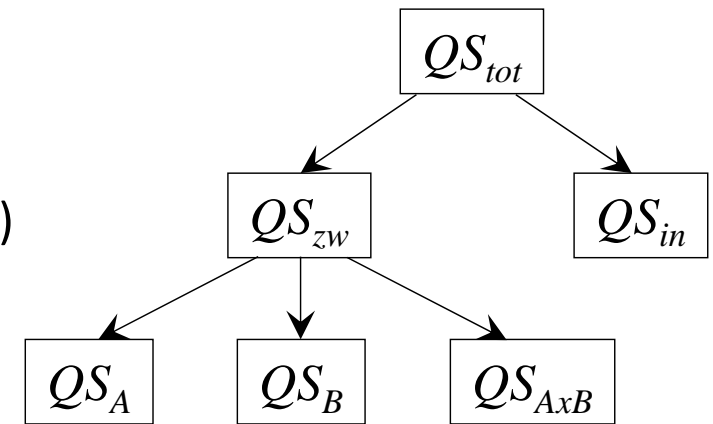
$$\bar{x}_{\cdot \cdot j} = \frac{\sum_{i=1}^{n_{\cdot \cdot}} \sum_{k=1}^q x_{ijk}}{q \cdot n_{\cdot \cdot}}$$

Gesamtmittel

$$\bar{x}_{\dots} = \frac{\sum_{i=1}^{n_{\cdot \cdot}} \sum_{j=1}^m \sum_{k=1}^q x_{ijk}}{n}$$

Zweifaktorielle Varianzanalyse

- Wir können nun wie bei der einfaktoriellen Varianzanalyse wieder die Gesamtvariabilität in den Daten, QS_{tot} , zerlegen in einen Teil, den wir durch die Faktoren erklären können (QS_{zw} , d.h. Unterschiede zwischen Zellen- bzw. Randmitteln) und einen Teil, den wir nicht erklären können (QS_{in} , Variabilität innerhalb der Zellen). Aufgrund des Vorliegens von zwei Faktoren A und B können wir die Zwischenvariabilität aufschlüsseln in die Effekte zulasten von A , B und der Interaktion AxB .



- Wir verwenden dabei die oben dargestellte Notation und vereinfachen dann immer durch folgende vereinfachte Schreibweise:

$\bar{x} = \bar{x}_{\dots}$ Gesamtmittelwert

$\bar{x}_{jk} = \bar{x}_{\cdot jk}$ Zellenmittelwerte

$\bar{x}_{j\cdot} = \bar{x}_{\cdot j\cdot}$ Spaltenmittelwerte (Faktor A)

$\bar{x}_{\cdot k} = \bar{x}_{\cdot\cdot k}$ Zeilenmittelwerte (Faktor B)

Zweifaktorielle Varianzanalyse

- Die **Gesamt-Quadratsumme** QS_{tot} ergibt sich als Summe der quadrierten Abweichungen aller Messwerte vom Gesamtmittelwert:

$$QS_{tot} = \sum_{i=1}^{n_{..}} \sum_{j=1}^m \sum_{k=1}^q (x_{ijk} - \bar{x}_{...})^2 = \sum_{i=1}^{n_{..}} \sum_{j=1}^m \sum_{k=1}^q (x_{ijk} - \bar{x})^2 \quad \text{mit} \quad df_{tot} = n - 1$$

- Die **Quadratsumme zwischen** QS_{zw} (hier: zwischen den Zellen) ergibt sich Summe der quadrierten Abweichungen der $m \cdot q$ Zellenmittelwerte vom Gesamtmittelwert:

$$QS_{zw} = \sum_{i=1}^{n_{..}} \sum_{j=1}^m \sum_{k=1}^q (\bar{x}_{\cdot jk} - \bar{x}_{...})^2 = n_{..} \cdot \sum_{j=1}^m \sum_{k=1}^q (\bar{x}_{jk} - \bar{x})^2 \quad \text{mit} \quad df_{zw} = m \cdot q - 1$$

- Die **Quadratsumme innerhalb** (der Residuen, Fehler) QS_{in} resultiert als Summe der quadrierten Abweichungen der Messwerte vom Zellenmittel:

$$QS_{in} = \sum_{i=1}^{n_{..}} \sum_{j=1}^m \sum_{k=1}^q (x_{ijk} - \bar{x}_{\cdot jk})^2 = \sum_{i=1}^{n_{..}} \sum_{j=1}^m \sum_{k=1}^q (x_{ijk} - \bar{x}_{jk})^2 \quad \text{mit} \quad df_{in} = m \cdot q \cdot (n_{..} - 1) = n - m \cdot q$$

Zweifaktorielle Varianzanalyse

- Die Quadratsumme QS_{zw} (die uns selbst nicht direkt interessiert) zerlegt sich nun in drei weitere Quadratsummen. Die **Quadratsumme für den Faktor A** (Spaltenfaktor) QS_A ergibt sich als Summe der quadrierten Abweichungen der m Spaltenmittelwerte vom Gesamtmittelwert:

$$QS_A = \sum_{i=1}^{n_{..}} \sum_{j=1}^m \sum_{k=1}^q (\bar{x}_{\cdot j \cdot} - \bar{x}_{\dots})^2 = n_{..} \cdot q \cdot \sum_{j=1}^m (\bar{x}_{j \cdot} - \bar{x})^2 \quad \text{mit} \quad df_A = m - 1$$

- Analog ergibt sich die **Quadratsumme für den Faktor B** (Zeilenfaktor) QS_B als Summe der quadrierten Abweichungen der q Zeilenmittelwerte vom Gesamtmittelwert:

$$QS_B = \sum_{i=1}^{n_{..}} \sum_{j=1}^m \sum_{k=1}^q (\bar{x}_{\cdot \cdot k} - \bar{x}_{\dots})^2 = n_{..} \cdot m \cdot \sum_{k=1}^q (\bar{x}_{\cdot k} - \bar{x})^2 \quad \text{mit} \quad df_B = q - 1$$

- Die **Quadratsumme der Interaktion** QS_{AxB} beinhaltet nun die Variabilität der Zellenmittel QS_{zw} , die nicht durch den Faktor A und nicht durch den Faktor B erklärt wird, also

$$QS_{AxB} = QS_{zw} - QS_A - QS_B$$

Zweifaktorielle Varianzanalyse

- Unter Rückgriff auf die Daten lässt sich die Interaktion QS_{AxB} alternativ auch bestimmen als Summe der quadrierten Abweichungen der $m \cdot q$ Zellenmittel vom Gesamtmittel, die verbleiben, wenn man vorher die Spalten- und Zeileneffekte subtrahiert hat:

$$\begin{aligned}
 QS_{AxB} &= \sum_{i=1}^{n_{..}} \sum_{j=1}^m \sum_{k=1}^q \left((\bar{x}_{\cdot jk} - \bar{x}_{\dots}) - \overbrace{(\bar{x}_{\cdot j\cdot} - \bar{x}_{\dots})} \quad - \quad \overbrace{(\bar{x}_{\cdot\cdot k} - \bar{x}_{\dots})} \right)^2 \\
 &= n_{..} \cdot \sum_{j=1}^m \sum_{k=1}^q \left((\bar{x}_{jk} - \bar{x}) - (\bar{x}_{j\cdot} - \bar{x}) - (\bar{x}_{\cdot k} - \bar{x}) \right)^2 \\
 &= n_{..} \cdot \sum_{j=1}^m \sum_{k=1}^q \left(\bar{x}_{jk} - \bar{x}_{j\cdot} - \bar{x}_{\cdot k} + \bar{x} \right)^2 \quad \text{mit} \quad df_{AxB} = (m-1) \cdot (q-1)
 \end{aligned}$$

- Wie oben dargestellt, gilt wieder die **Additivität** der Quadratsummen und der Freiheitsgrade:

$$QS_{tot} = QS_A + QS_B + QS_{AxB} + QS_{in} \qquad df_{tot} = df_A + df_B + df_{AxB} + df_{in}$$

- Die **mittleren Quadrate** werden wieder als Quotienten der QS und df gebildet:

$$MQ_{tot} = \frac{QS_{tot}}{df_{tot}}, \quad MQ_A = \frac{QS_A}{df_A}, \quad \dots$$

- 1 Beispiel und Grundlagen
- 2 Effekte in Mittelwertsdiagrammen
- 3 Quadratsummenzerlegung
- 4 Modell
- 5 Testprozedur
- 6 Durchführung mittels SPSS

Zweifaktorielle Varianzanalyse

- Die Modellgleichung für die einfaktorielle Varianzanalyse (ohne Messwiederholung) lautete:

$$x_{ij} = \mu + \alpha_j + e_{ij} \quad i = 1 \dots n_j \text{ (Vp-Index)}, j = 1 \dots m \text{ (Faktor)}$$

- Die Modellgleichung für die zweifaktorielle Varianzanalyse (ohne Messwiederholung) lautet entsprechend erweitert:

$$x_{ijk} = \mu + \alpha_j + \beta_k + \gamma_{jk} + e_{ijk} \quad \begin{array}{l} i = 1 \dots n_{jk} \text{ (Vp-Index)}, j = 1 \dots m \text{ (Faktor A)}, \\ k = 1 \dots q \text{ (Faktor B)} \end{array}$$

Manchmal findet sich statt γ_{jk} auch die Bezeichnung $\alpha\beta_{jk}$.

wobei bezeichnet ...

- μ den Gesamtmittelwert über alle $m \times q$ Bedingungen in der Population
- $\alpha_j = \mu_{j\bullet} - \mu$ den Effekt der Bedingung (Stufe) j des Faktors A in der Population
- $\beta_k = \mu_{\bullet k} - \mu$ den Effekt der Bedingung (Stufe) k des Faktors B in der Population
- $\gamma_{jk} = \mu_{jk} - \mu_{j\bullet} - \mu_{\bullet k} + \mu$ den Effekt der Bedingungskombination (j, k) der Interaktion $A \times B$ in der Population
- e_{ijk} den Residualwert der Person i unter der Bedingungskombination (j, k) .

Zweifaktorielle Varianzanalyse

- Bezeichnung: Zweifaktorielle Varianzanalyse (ohne Messwiederholung).
- Einsatzbereich: Untersuchung des Einflusses von zwei unabhängigen Variablen (Faktoren mit festen Effekten) auf die Mittelwerte in einer metrischen Variablen X beim Vorliegen unabhängiger Gruppen.
- Hypothesen: Die ungerichteten Hypothesen für die drei möglichen Effekte können wieder auf verschiedene Weisen dargestellt werden (vgl. einfaktorielle ANOVA) , z.B.
 - $H_0 : \sum_{j=1}^m \alpha_j^2 = 0$ und $H_1 : \sum_{j=1}^m \alpha_j^2 > 0$ Haupteffekt Faktor A
 - $H_0 : \sum_{k=1}^q \beta_k^2 = 0$ und $H_1 : \sum_{k=1}^q \beta_k^2 > 0$ Haupteffekt Faktor B
 - $H_0 : \sum_{j=1}^m \sum_{k=1}^q \gamma_{jk}^2 = 0$ und $H_1 : \sum_{j=1}^m \sum_{k=1}^q \gamma_{jk}^2 > 0$ Interaktion
- Voraussetzungen: In allen Teilpopulationen (Gruppen) ist das Merkmal X (die AV) normalverteilt und weist gleiche Varianzen auf (Homoskedastizität). Die folgenden Darstellungen gelten zudem nur für den Fall, dass die Stichprobengrößen in allen Zellen gleich sind. (Außerdem wird wieder davon ausgegangen, dass die Beobachtungen unabhängig sind.)

Zweifaktorielle Varianzanalyse

- Vorgehen: Bei der einfaktoriellen Varianzanalyse haben wir gezeigt, dass bei Gültigkeit der Annahmen MQ_{in} eine erwartungstreue Schätzung der Populationsresidualvarianz darstellt und dies für MQ_{zw} nur dann gilt, wenn die H_0 gilt.
- Analog ist auch im zweifaktoriellen Fall MQ_{in} eine erwartungstreue Schätzung der Populationsresidualvarianz σ_e^2 und die drei mittleren Quadrate der Effekte überschätzen die Populationsresidualvarianz umso mehr, je stärker die jeweiligen Effekte in der Population sind:

$$E(MQ_A) = \sigma_e^2 + q \cdot n_{..} \cdot \sum_{j=1}^m \frac{\alpha_j^2}{m-1}$$

$$E(MQ_B) = \sigma_e^2 + m \cdot n_{..} \cdot \sum_{k=1}^q \frac{\beta_k^2}{q-1}$$

$$E(MQ_{AxB}) = \sigma_e^2 + n_{..} \cdot \sum_{j=1}^m \sum_{k=1}^q \frac{\gamma_{jk}^2}{(m-1) \cdot (q-1)}$$

- Bei allen drei Effekten spricht also wieder ein Quotient MQ_{Effekt} / MQ_{in} nahe 1 jeweils für die betreffende H_0 . Je größer ein Quotient ist, desto stärker sprechen die Stichprobendaten gegen die betreffende H_0 .

Zweifaktorielle Varianzanalyse

- Erneut sind die Quadratsummen mit ihren entsprechenden Freiheitsgraden unter der H_0 χ^2 -verteilt und die Quotienten der mittleren Quadrate F -verteilt:

$$F_A = \frac{MQ_A}{MQ_{in}} \quad \text{mit } df_1 = m - 1 \quad \text{und } df_2 = n - m \cdot q$$

$$F_B = \frac{MQ_B}{MQ_{in}} \quad \text{mit } df_1 = q - 1 \quad \text{und } df_2 = n - m \cdot q$$

$$F_{AxB} = \frac{MQ_{AxB}}{MQ_{in}} \quad \text{mit } df_1 = (m - 1) \cdot (q - 1) \quad \text{und } df_2 = n - m \cdot q$$

- Entscheidungen: Zurückweisung der H_0 jeweils, wenn $F_{\text{Effekt}} > F_{\text{crit}} = F_{df_1; df_2; 1-\alpha}$ (bei zweiseitiger Hypothese), also z.B. $F_A > F_{\text{crit}} = F_{m-1; n-m \cdot q; 0.95}$.

Zweifaktorielle Varianzanalyse

➤ **Beispiel:** Zu bestimmen sind die QS , df und MQ sowie schließlich die F -Werte wie folgt:

$$QS_A = n_{..} \cdot q \cdot \sum_{j=1}^m (\bar{x}_{j\cdot} - \bar{x})^2 = 10 \cdot 2 \cdot ((46.15 - 53.533)^2 + (57.55 - 53.533)^2 + \dots) = 1639.63$$

$$QS_B = n_{..} \cdot m \cdot \sum_{k=1}^q (\bar{x}_{\cdot k} - \bar{x})^2 = 10 \cdot 3 \cdot ((46.933 - 53.533)^2 + (60.133 - 53.533)^2) = 2613.60$$

$$QS_{AxB} = n_{..} \cdot \sum_{j=1}^m \sum_{k=1}^q (\bar{x}_{jk} - \bar{x}_{j\cdot} - \bar{x}_{\cdot k} + \bar{x})^2 = 10 \cdot ((31.9 - 46.15 - 46.933 + 53.533)^2 + (60.4 - 46.15 - 60.133 + 53.533)^2 + \dots) = 1962.70$$

$$QS_{tot} = \sum_{i=1}^{n_{..}} \sum_{j=1}^m \sum_{k=1}^q (x_{ijk} - \bar{x})^2 = (34 - 53.533)^2 + (27 - 53.533)^2 + \dots = 14716.93$$

$$QS_{in} = QS_{tot} - QS_A - QS_B - QS_{AxB} = 14716.93 - 1639.63 - 2613.60 - 1962.70 = 8501.00$$

| | |
|---------------|---------------------|
| $n_{..} = 10$ | Zahl Vpn pro Zelle |
| $n = 60$ | Gesamtzahl Vpn |
| $m = 3$ | Zahl Faktorstufen A |
| $q = 2$ | Zahl Faktorstufen B |

| \bar{x} : | Lb kon | Lb ver | Web | |
|-------------|--------|--------|------|--------|
| niedrig | 31.9 | 52.5 | 56.4 | 46.933 |
| hoch | 60.4 | 62.6 | 57.4 | 60.133 |
| | 46.15 | 57.55 | 65.9 | 53.533 |

| | | Faktor A | | |
|----------|---------|----------|------|-----|
| | | Lb k | Lb v | Web |
| Faktor B | niedrig | 34 | 50 | 61 |
| | | 27 | 64 | 55 |
| | | ... | ... | ... |
| | hoch | 56 | 68 | 72 |
| | | 85 | 60 | 45 |
| | | 72 | 51 | 54 |
| | ... | ... | ... | |
| | 38 | 68 | 49 | |

➤ **Beispiel:** Als nächstes sind die df und MQ zu bestimmen:

$$df_A = m - 1 = 3 - 1 = 2$$

$$df_B = q - 1 = 2 - 1 = 1$$

$$df_{AxB} = (m - 1) \cdot (q - 1) = 2 \cdot 1 = 2$$

$$df_{tot} = n - 1 = 60 - 1 = 59$$

$$df_{in} = n - m \cdot q = 60 - 3 \cdot 2 = 54$$

$$MQ_A = \frac{QS_A}{df_A} = \frac{1639.633}{2} = 819.817$$

$$MQ_B = \frac{QS_B}{df_B} = \frac{2613.000}{1} = 2613.000$$

$$MQ_{AxB} = \frac{QS_{AxB}}{df_{AxB}} = \frac{1962.700}{2} = 981.350$$

$$MQ_{in} = \frac{QS_{in}}{df_{in}} = \frac{8501.000}{54} = 157.426$$

Zweifaktorielle Varianzanalyse

- **Beispiel:** Abschließend werden die F -Werte für die beiden Haupteffekte und die Wechselwirkung bestimmt:

$$\left. \begin{aligned} F_A &= \frac{MQ_A}{MQ_{in}} = \frac{819.817}{157.426} = 5.21 && \text{mit } df_1 = 2 \\ F_B &= \frac{MQ_B}{MQ_{in}} = \frac{2613.600}{157.426} = 16.60 && \text{mit } df_1 = 1 \\ F_{AxB} &= \frac{MQ_{AxB}}{MQ_{in}} = \frac{981.350}{157.426} = 6.23 && \text{mit } df_1 = 2 \end{aligned} \right\} df_2 = 54$$

- Die kritischen F -Werte bei $\alpha = 0.05$ für den Haupteffekt A und die Wechselwirkung AxB sind gleich: $F_{crit} = F_{2;54;0.95} = 3.17$. Für den Haupteffekt B lautet der kritische Wert $F_{crit} = F_{1;54;0.95} = 4.02$. Damit sind alle drei Effekte auf dem 5%-Niveau statistisch signifikant.
- Um sich die Art der Haupteffekte und die Wechselwirkung genauer zu betrachten werden Mittelwertdiagramme erstellt.

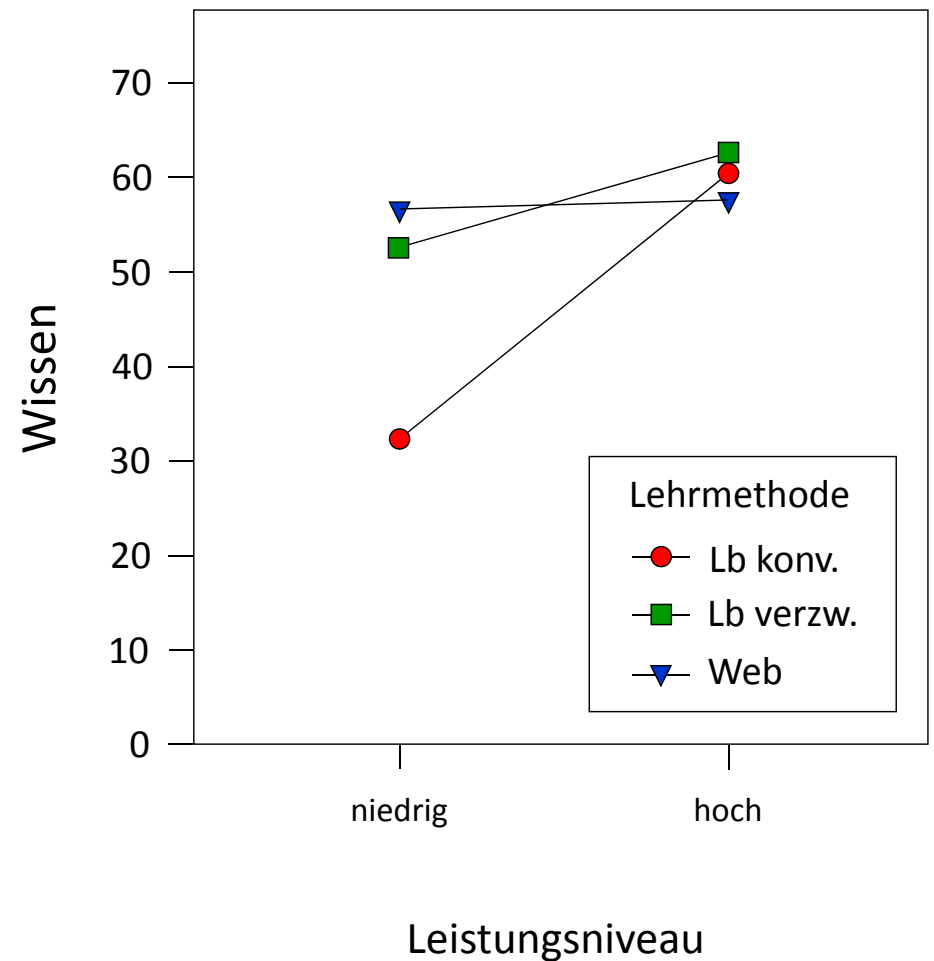
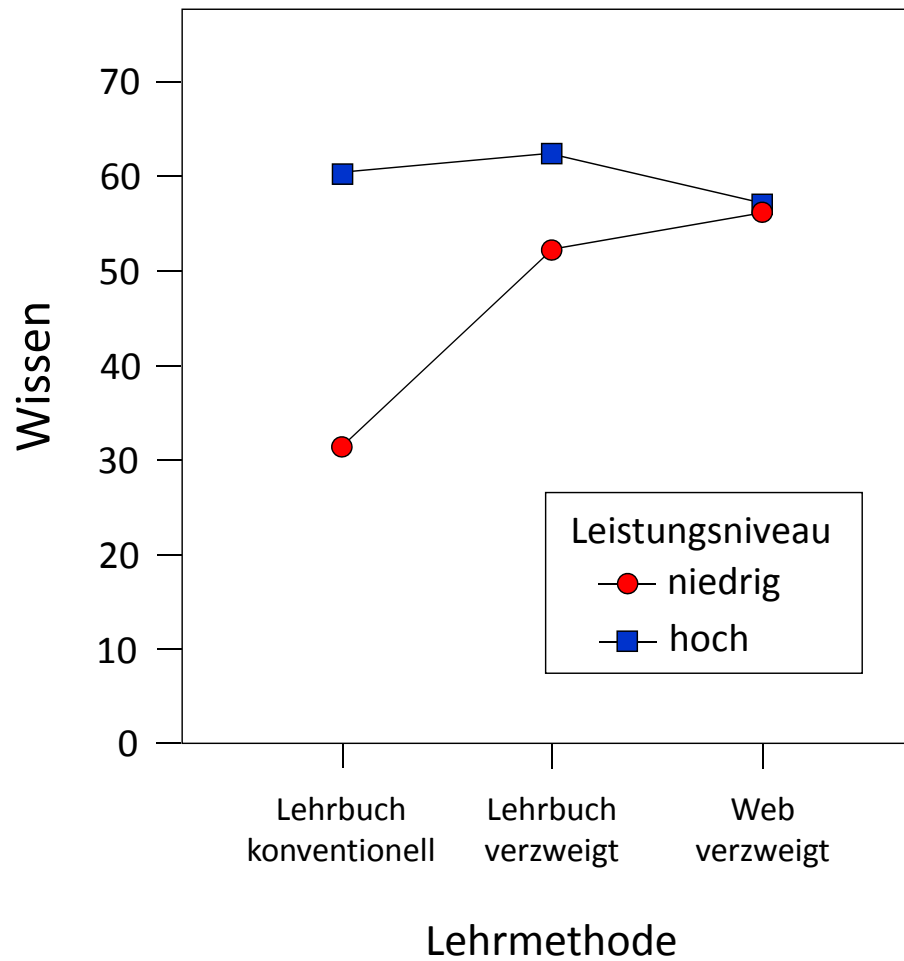
Zweifaktorielle Varianzanalyse

- **Beispiel:** Vorher sind in der varianzanalytischen Tabelle nochmals alle Ergebnisse zusammengetragen:

| Variation | Quadratsummen | Freiheitsgrade | Mittlere Quadrate | F | p |
|-----------------|---------------|----------------|-------------------|--------|-------|
| Lehrmethode | 1639.633 | 2 | 819.817 | 5.208 | < .05 |
| Leistungsniveau | 2613.600 | 1 | 2613.600 | 16.602 | < .05 |
| Interaktion | 1962.700 | 2 | 981.350 | 6.234 | < .05 |
| Innerhalb | 8501.000 | 54 | 157.426 | | |
| Total | 14716.933 | 59 | | | |

Zweifaktorielle Varianzanalyse

➤ **Beispiel:** Die Mittelwertdiagramme sehen wie folgt aus:



Zweifaktorielle Varianzanalyse

- Um nach dem Vorliegen des Haupteffektes für die Lehrmethode festzustellen, zwischen welchen Lehrmethoden signifikante Mittelwertsunterschiede bestehen, führen wir post hoc den Tukey-Test durch (s.u.). Dieser ergibt, dass die konventionelle Lehrbuchmethode signifikant schlechter abschneidet als die beiden verzweigten, die sich untereinander nicht unterscheiden.
- Die (in der Stichprobe) bestehende hybride Wechselwirkung schränkt aber die Deutung des Haupteffektes der Lehrmethode ein: Wir können dadurch nicht sagen, dass die konventionelle Lehrmethode (immer) am schlechtesten abschneidet.
- Die Unterlegenheit der konventionellen Lehrmethode resultiert vor allem aus dem schlechten Abschneiden der Schüler mit geringem Leistungsniveau. Bei Schülern mit hohen Physikfähigkeiten hingegen spielt die Lehrmethode offensichtlich keine große Rolle; sie lernen unter allen Lehrmethoden annähernd gleich gut.
- Bezüglich des zweiten Faktors Leistungsniveau sagt der Haupteffekt aus, dass Schüler mit höherem Leistungsniveau besser abschneiden. Diese Aussage muss durch die Wechselwirkung nicht eingeschränkt werden.

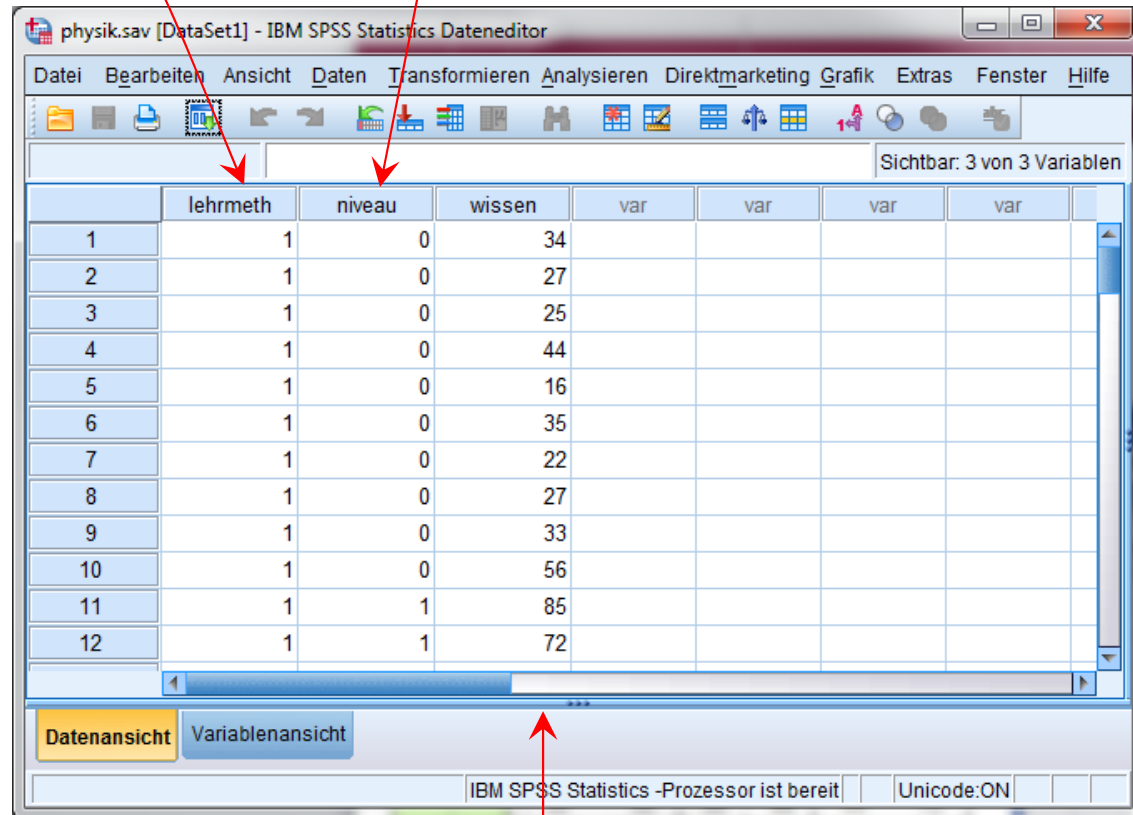
- 1 Beispiel und Grundlagen
- 2 Effekte in Mittelwertsdiagrammen
- 3 Quadratsummenzerlegung
- 4 Modell
- 5 Testprozedur
- 6 Durchführung mittels SPSS

Zweifaktorielle Varianzanalyse ohne Messwiederholung in SPSS

Faktor A: Lehrmethode, codiert
 1 = Lehrbuch konventionell
 2 = Lehrbuch verzweigt
 3 = Web verzweigt

Faktor B: Leistungsniveau, codiert:
 0 = niedrig
 1 = hoch

| Leistungsniveau | Lehrmethode | | | | | |
|-----------------|--------------------------|----|----------------------|----|------------------------|----|
| | konventionelles Lehrbuch | | verzweigtes Lehrbuch | | verzweigt Web-gestützt | |
| niedrig | 34 | 35 | 50 | 62 | 61 | 69 |
| | 27 | 22 | 64 | 47 | 55 | 64 |
| | 25 | 27 | 49 | 53 | 52 | 47 |
| | 44 | 33 | 51 | 45 | 49 | 51 |
| | 16 | 56 | 36 | 68 | 44 | 72 |
| | 85 | 60 | 60 | 73 | 45 | 55 |
| hoch | 72 | 42 | 51 | 67 | 54 | 61 |
| | 75 | 41 | 52 | 88 | 72 | 59 |
| | 78 | 48 | 50 | 77 | 52 | 48 |
| | 65 | 38 | 40 | 68 | 79 | 49 |



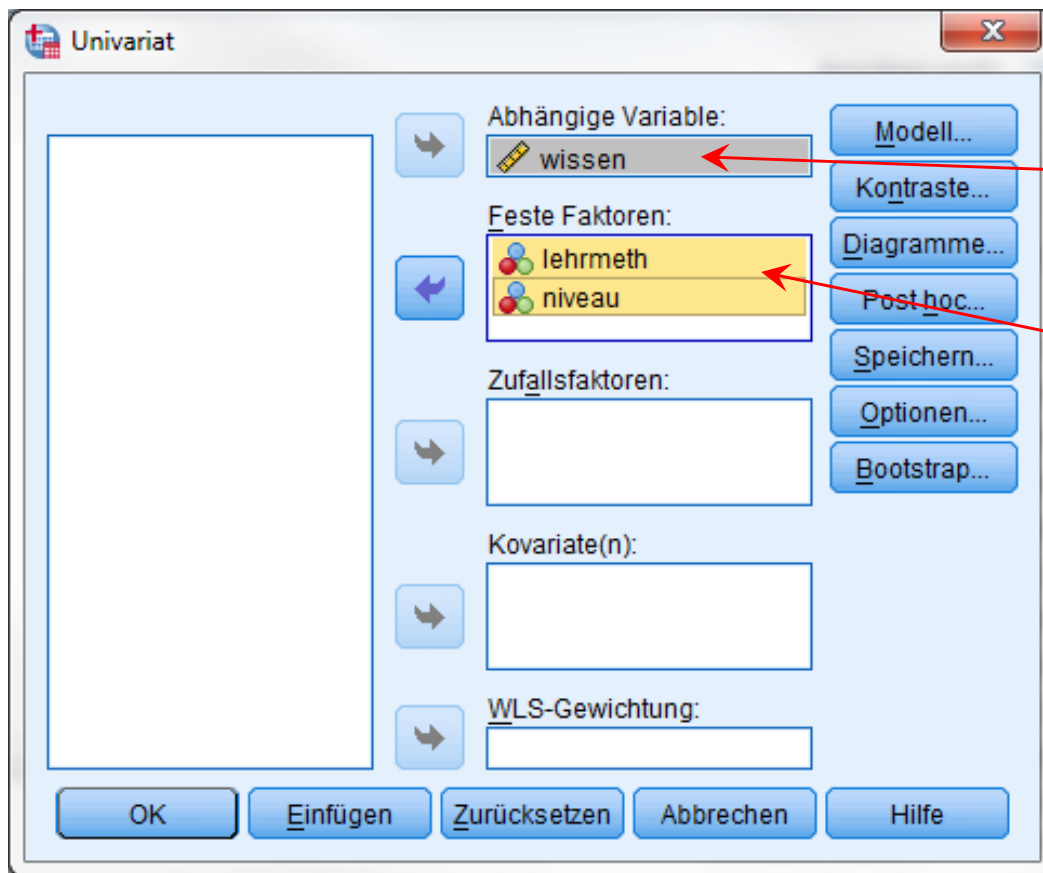
6 UV-Blöcke

| UV ₁ | UV ₂ |
|-----------------|-----------------|
| 1 | 0 |
| 1 | 1 |
| 2 | 0 |
| 2 | 1 |
| 3 | 0 |
| 3 | 1 |

Abhängige Variable: Wissen

Zweifaktorielle Varianzanalyse ohne Messwiederholung in SPSS

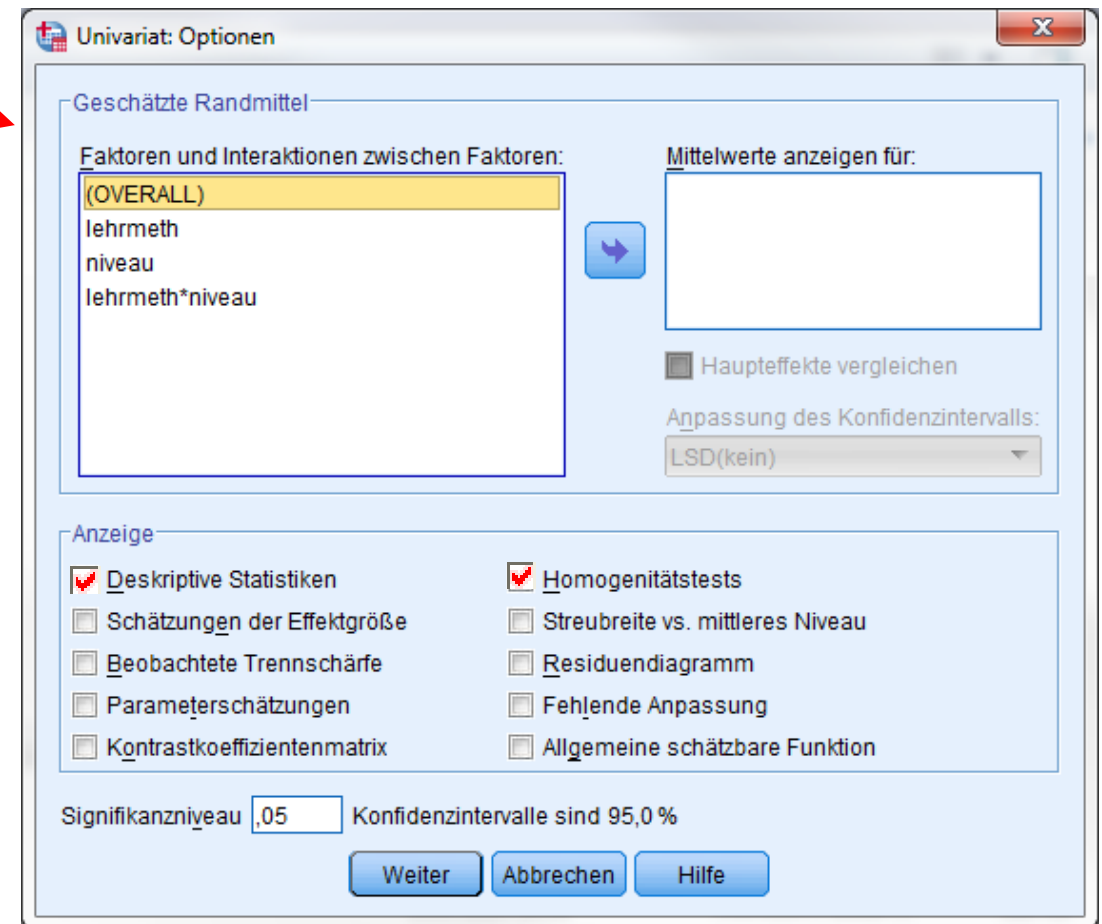
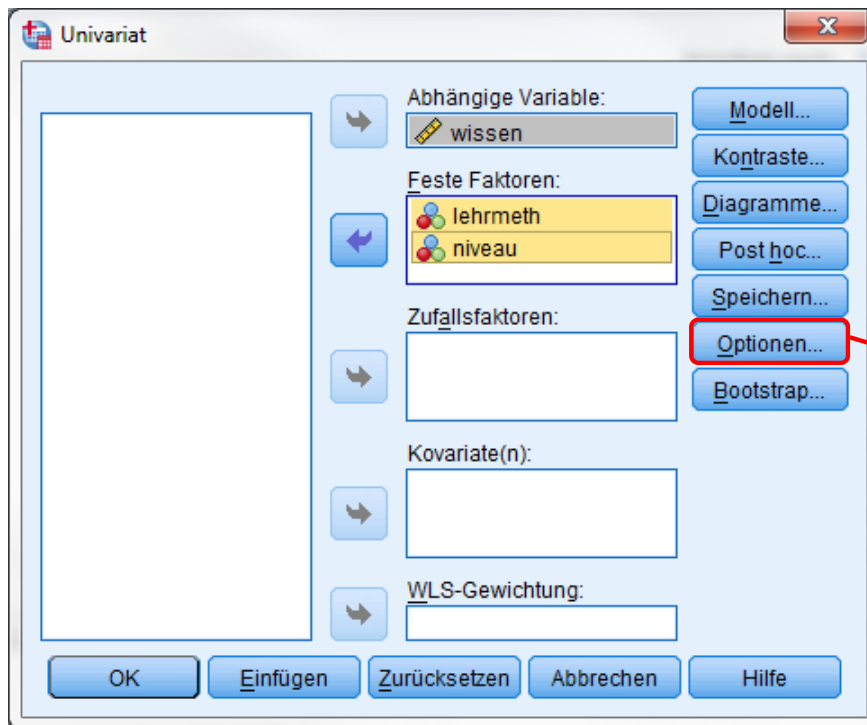
Die zwei- (und auch höher) faktorielle Varianzanalyse, bei der auf keinem Faktor Messwiederholung vorliegt, erhält man über die Prozedur Unianova unter `Analyse > Erweiterte Analyse > Allgemeines Lineares Modell > Univariat...`



Angabe der abhängigen Variablen

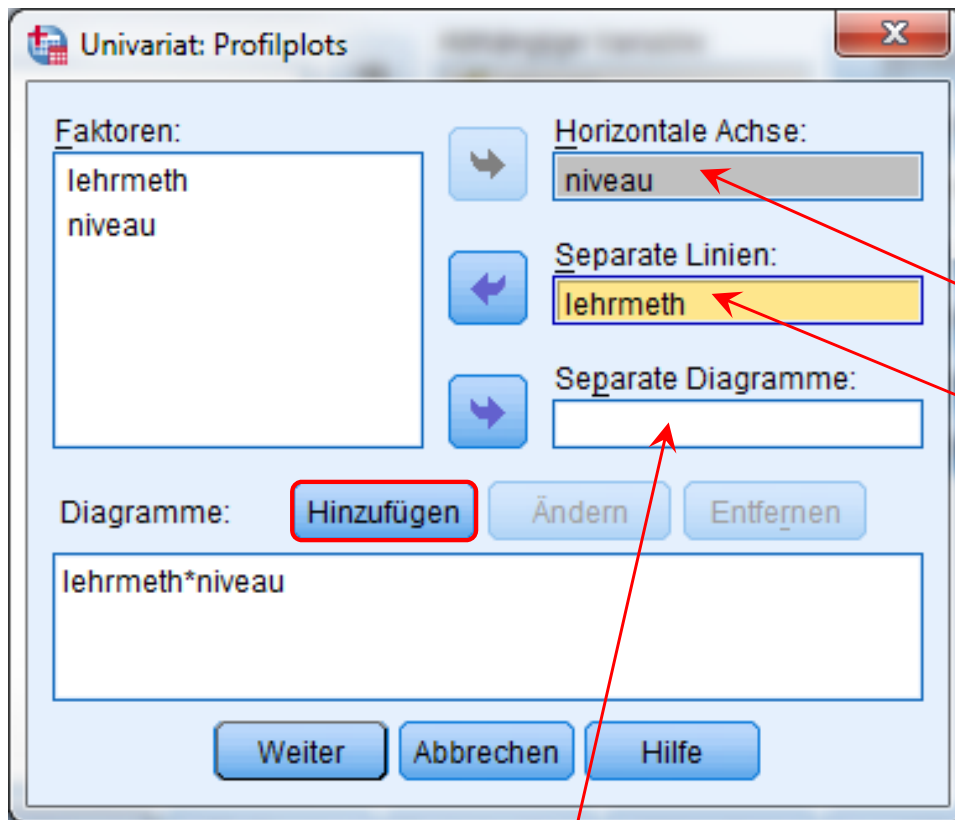
Angabe der unabhängigen Variablen (bei festen Effekten)

Zweifaktorielle Varianzanalyse ohne Messwiederholung in SPSS



Nützliche (Optionen) sind unter anderem „Deskriptive Statistiken“ und „Homogenitätstest“

Zweifaktorielle Varianzanalyse ohne Messwiederholung in SPSS



Unter (Di agramme) können grafische Mittelwertsprofile angefordert werden.

Dafür wird die UV, die auf der Abszisse abgetragen werden soll unter "Horizontale Achse" und die UV, die in die Linien abgebildet werden soll, unter "Separate Linien" spezifiziert.

Da in einem Aufruf mehrere Grafiken angefordert werden können, muss nach der Angabe jedes Diagramms der Button (Hi n z u f ü g e n) gedrückt werden!

Bei mehr als 2 UVs kann unter "Separate Diagramme" ein dritter Faktor abgegeben werden. Es werden dann für jede Faktorstufe dieser UV getrennte Diagramme erstellt.

Univariate Varianzanalyse

Deskriptive Statistiken

Abhängige Variable:wissen

| lehrmeth | niveau | Mittelwert | Standardabweichung | N |
|--------------------|-----------|------------|--------------------|----|
| 1 Lb konventionell | 0 niedrig | 31,90 | 11,493 | 10 |
| | 1 hoch | 60,40 | 17,174 | 10 |
| | Gesamt | 46,15 | 20,397 | 20 |
| 2 Lb verzweigt | 0 niedrig | 52,50 | 9,675 | 10 |
| | 1 hoch | 62,60 | 14,653 | 10 |
| | Gesamt | 57,55 | 13,149 | 20 |
| 3 Web verzweigt | 0 niedrig | 56,40 | 9,594 | 10 |
| | 1 hoch | 57,40 | 10,824 | 10 |
| | Gesamt | 56,90 | 9,968 | 20 |
| Gesamt | 0 niedrig | 46,93 | 14,769 | 30 |
| | 1 hoch | 60,13 | 14,115 | 30 |
| | Gesamt | 53,53 | 15,794 | 60 |

Hier werden zunächst alle Zellen und Randmittelwerte (mit Standardabweichung und Stichprobenumfängen) ausgegeben.

Der Levene-Test zeigt an, dass wir bei $\alpha = 0.10$ von Varianzhomogenität ausgehen können, $F(5, 54) = 1.89, p > .10$.

Levene-Test auf Gleichheit der Fehlervarianzen^a

Abhängige Variable:wissen

| F | df1 | df2 | Sig. |
|-------|-----|-----|------|
| 1,889 | 5 | 54 | ,111 |

Prüft die Nullhypothese, daß die Fehlervarianz der abhängigen Variablen über Gruppen hinweg gleich ist.

a. Design: Konstanter Term + lehrmeth + niveau + lehrmeth * niveau

Zweifaktorielle Varianzanalyse ohne Messwiederholung in SPSS

Tests der Zwischensubjekteffekte

Abhängige Variable:wissen

| Quelle | Quadratsumme vom Typ III | df | Mittel der Quadrate | F | Sig. |
|-----------------------------|--------------------------|----|---------------------|---------|------|
| Korrigiertes Modell | 6215,933 ^a | 5 | 1243,187 | 7,897 | ,000 |
| Konstanter Term | 171949,067 | 1 | 171949,067 | 1,092E3 | ,000 |
| lehrmeth | 1639,633 | 2 | 819,817 | 5,208 | ,009 |
| niveau | 2613,600 | 1 | 2613,600 | 16,602 | ,000 |
| lehrmeth * niveau | 1962,700 | 2 | 981,350 | 6,234 | ,004 |
| Fehler | 8501,000 | 54 | 157,426 | | |
| Gesamt | 186666,000 | 60 | | | |
| Korrigierte Gesamtvariation | 14716,933 | 59 | | | |

Faktor A

Faktor B

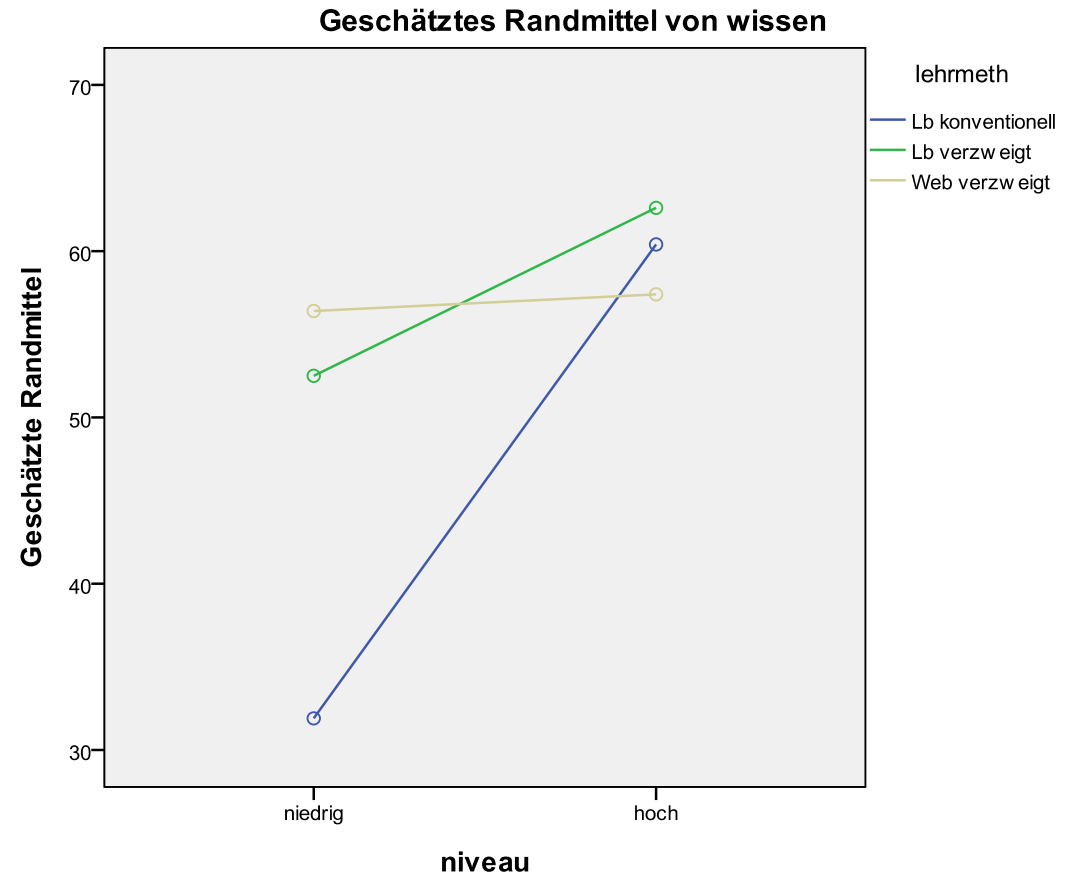
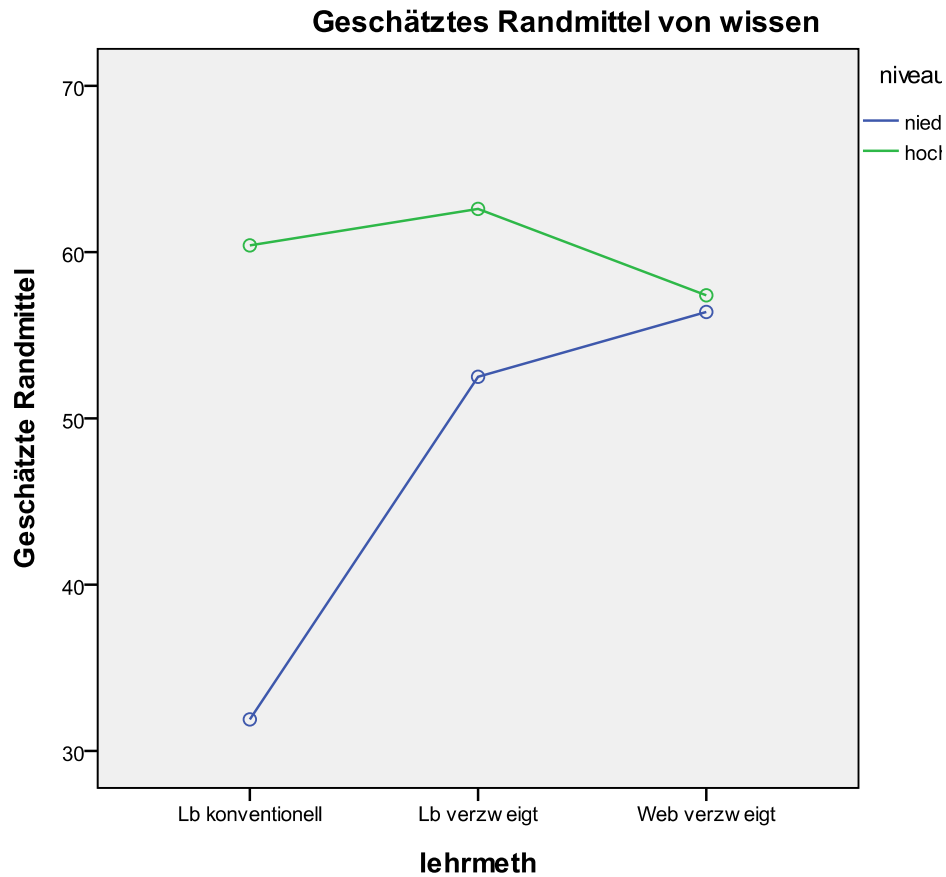
Interaktion $A \times B$

Residual/Fehler

a. R-Quadrat = ,422 (korrigiertes R-Quadrat = ,369)

Die Tabelle erhält in der vertrauten Anordnung die Ergebnisse der ANOVA.

Zweifaktorielle Varianzanalyse ohne Messwiederholung in SPSS



Die Standardausgabe der Mittelwertdiagramme der Prozedur Unianova ist noch optimierbar ...

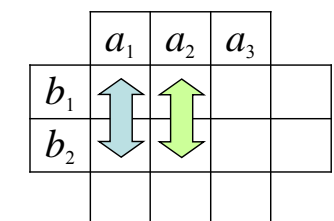
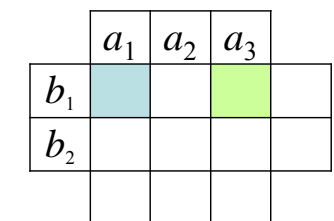
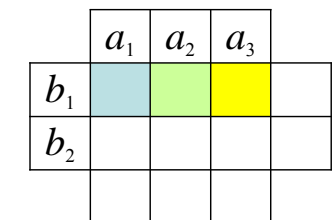
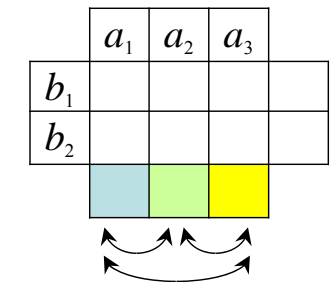
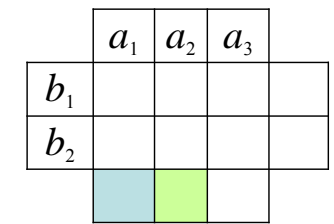
- 7 Kontraste
- 8 Effektstärken
- 9 Ungleiche Stichprobenumfänge
- 10 Designs mit zufälligen Faktoren

Zweifaktorielle Varianzanalyse: Kontraste

- Wie schon im einfaktoriellen Fall sind auch hier nicht alle Fragen mit der Durchführung der Varianzanalyse beantwortet. Prinzipiell können alle Formen von multiplen Mittelwertsvergleichen (a priori und a posteriori), die bereits dargestellt wurden, auch hier eingesetzt werden. Durch die Tatsache, dass wir es mit zwei Faktoren zu tun haben, werden prinzipiell entsprechend mehr solcher Vergleiche möglich.
- Die Frage, welche solcher Mittelwertsvergleiche sinnvoll sind, hängt von den erwarteten bzw. empirisch sich ergebenden Effekten ab.
- Liegt ein **Haupteffekt** vor (und keine Interaktion, die der Deutung entgegensteht) und weist der Faktor mehr als zwei Stufen auf, so wird man (analog zur einfaktoriellen Varianzanalyse) wissen wollen, zwischen welchen Gruppen (Randmittelwerten) Unterschiede bestehen. Dies ist z.B. mittels a priori geplanten Kontrasten oder a posteriori über den Vergleich aller paarweisen Mittelwerte prüfbar.
- Liegt eine **Wechselwirkung** vor, so sind die Effekte eines Faktors auf den Stufen des anderen Faktors unterschiedlich. Hier kann man prüfen, ob sich Zellenmittelwerte innerhalb einer festen Faktorstufe unterscheiden (**einfache Haupteffekte** oder **einfache Kontraste**) und ob Unterschiede zwischen Zellenmittelwerten auf verschiedenen Faktorstufen gleich sind (**Interaktionskontraste**).

Zweifaktorielle Varianzanalyse: Kontraste

| Nr. | Fragestellung | Bezeichnung |
|-----|---|---|
| (1) | Unterscheiden sich bestimmte a priori festgelegte Randmittel des Faktors A (mit $m > 2$, sonst gleich Haupteffekt)? | Treatmentkontrast (comparison of marginal means) |
| (2) | Welche der Randmittel unterscheiden sich bei Faktor A (mit $m > 2$, sonst eindeutig) a posteriori? | w.o. |
| (3) | Übt ein Faktor A auf einer festen Abstufung des Faktors B einen Effekt aus? | einfacher Haupteffekt (simple [main] effect) |
| (4) | Falls (3) gegeben ist: Welche der Zellenmittel von Faktor A unterscheiden sich innerhalb einer festen Stufe von B ? | einfacher Kontrast (simple contrast) |
| (5) | Unterscheiden sich die Unterschiede in den Zellenmitteln von B auf unterschiedlichen Abstufungen von Faktor A ? | Interaktionskontrast (interaction contrast) |



Zweifaktorielle Varianzanalyse: Kontraste

- Nehmen wir an, in unserer Beispiel-Studie hätte sich folgendes abweichende Ergebnis ergeben:

Deskriptive Statistiken

Abhängige Variable:wissen

| lehrmeth | niveau | Mittelwert | Standardabweichung | N |
|--------------------|-----------|------------|--------------------|----|
| 1 Lb konventionell | 0 niedrig | 34,90 | 12,242 | 10 |
| | 1 hoch | 51,10 | 16,278 | 10 |
| | Gesamt | 43,00 | 16,297 | 20 |
| 2 Lb verzweigt | 0 niedrig | 52,50 | 9,675 | 10 |
| | 1 hoch | 62,60 | 14,653 | 10 |
| | Gesamt | 57,55 | 13,149 | 20 |
| 3 Web verzweigt | 0 niedrig | 56,40 | 9,594 | 10 |
| | 1 hoch | 63,20 | 12,435 | 10 |
| | Gesamt | 59,80 | 11,358 | 20 |
| Gesamt | 0 niedrig | 47,93 | 13,950 | 30 |
| | 1 hoch | 58,97 | 15,130 | 30 |
| | Gesamt | 53,45 | 15,464 | 60 |

Levene-Test auf Gleichheit der Fehlervarianzen^a

Abhängige Variable:wissen

| F | df1 | df2 | Sig. |
|-------|-----|-----|------|
| 1,399 | 5 | 54 | ,239 |

Prüft die Nullhypothese, daß die Fehlervarianz der abhängigen Variablen über Gruppen hinweg gleich ist.

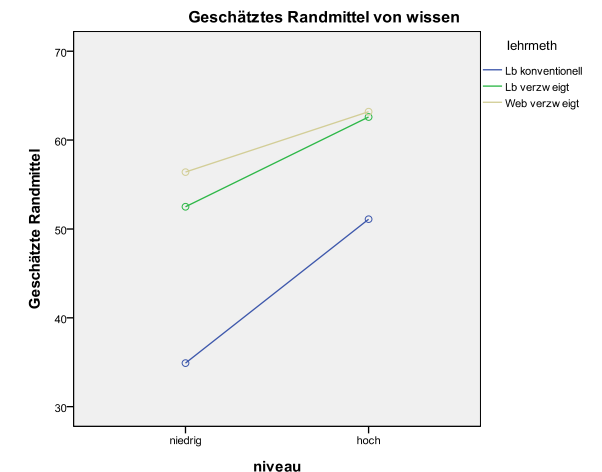
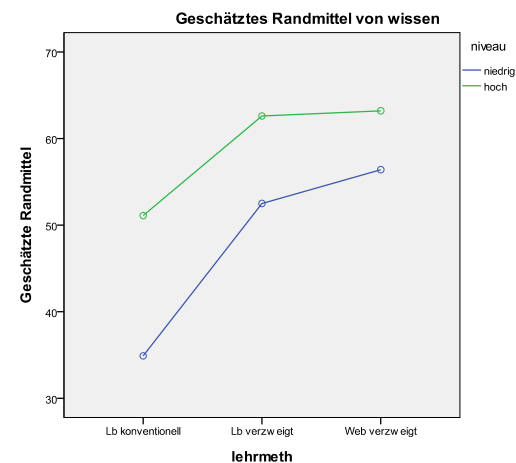
a. Design: Konstanter Term + lehrmeth + niveau + lehrmeth * niveau

Tests der Zwischensubjekteffekte

Abhängige Variable:wissen

| Quelle | Quadratsumme vom Typ III | df | Mittel der Quadrate | F | Sig. |
|-----------------------------|--------------------------|----|---------------------|---------|------|
| Korrigiertes Modell | 5380,150 ^a | 5 | 1076,030 | 6,657 | ,000 |
| Konstanter Term | 171414,150 | 1 | 1,714E5 | 1,060E3 | ,000 |
| lehrmeth | 3326,700 | 2 | 1663,350 | 10,290 | ,000 |
| niveau | 1826,017 | 1 | 1826,017 | 11,297 | ,001 |
| lehrmeth * niveau | 227,433 | 2 | 113,717 | ,704 | ,499 |
| Fehler | 8728,700 | 54 | 161,643 | | |
| Gesamt | 185523,000 | 60 | | | |
| Korrigierte Gesamtvariation | 14108,850 | 59 | | | |

a. R-Quadrat = ,381 (korrigiertes R-Quadrat = ,324)



Zweifaktorielle Varianzanalyse: Treatmentkontraste

- Es liegt ein Haupteffekt der Lehrmethode, $F(2, 54) = 10.29, p < .05$, sowie ein Haupteffekt des Leistungsniveaus vor, $F(1, 54) = 11.30, p < .05$, aber keine Interaktion, $F(2, 54) < 1, ns$.
- Da der Faktor Leistungsniveau nur 2 Stufen aufweist, kann man direkt nach Inspektion der Mittelwerte bzw. des Diagramms folgern, dass Schüler mit einem höheren Leistungsniveau signifikant besser in dem Wissenstest abgeschnitten haben.
- Der Faktor Lehrmethode weist 3 Stufen auf. Der signifikante Haupteffekt besagt hier, dass sich zumindest die Mittelwerte zweier Lehrmethoden unterscheiden, aber nicht welche.
- Es könnte z.B. a priori die Hypothese bestehen, dass die verzweigte Lehrbuchmethode besser als die konventionelle Lehrbuchmethode ist. Diese Hypothese lässt sich in den als **Treatmentkontrast** bezeichneten Vergleich $L = -1 \cdot \bar{x}_{1\bullet} + 1 \cdot \bar{x}_{2\bullet} + 0 \cdot \bar{x}_{3\bullet}$ überführen und daraus wie bekannt die Quadratsumme

$$QS_L = \frac{q \cdot n_{\bullet\bullet} \cdot L^2}{\sum_{j=1}^m c_j^2} = MQ_L \text{ mit } df = 1$$

bestimmen. (Statt n_{\bullet} wie in der einfaktoriellen Varianzanalyse gehen hier $q \cdot n_{\bullet\bullet}$ Messwerte in jeden Spalten-Randmittelwert ein.)

Bei einem Vergleich von Stufen des Zeilenfaktors B resultiert entsprechend:

$$QS_L = \frac{m \cdot n_{\bullet\bullet} \cdot L^2}{\sum_{k=1}^q c_k^2}$$

Zweifaktorielle Varianzanalyse: Treatmentkontraste

| \bar{x}_{jk} | Lb kon | Lb ver | Web | |
|----------------|--------|--------|--------|--------|
| niedrig | 34.9 | 52.5 | 56.4 | 47.933 |
| hoch | 51.1 | 62.6 | 63.2 | 58.967 |
| | 43.000 | 57.550 | 59.800 | 53.450 |

- In der Prüfgröße

$$F_L = \frac{MQ_L}{MQ_{in}} \text{ mit } df_1 = 1, df_2 = n - q \cdot m$$

wird im Nenner MQ_{in} aus der zweifaktoriellen Varianzanalyse eingesetzt.

- Im modifizierten **Beispiel** resultiert für den Kontrast $L = -1 \cdot \bar{x}_{1\cdot} + 1 \cdot \bar{x}_{2\cdot} + 0 \cdot \bar{x}_{3\cdot}$:

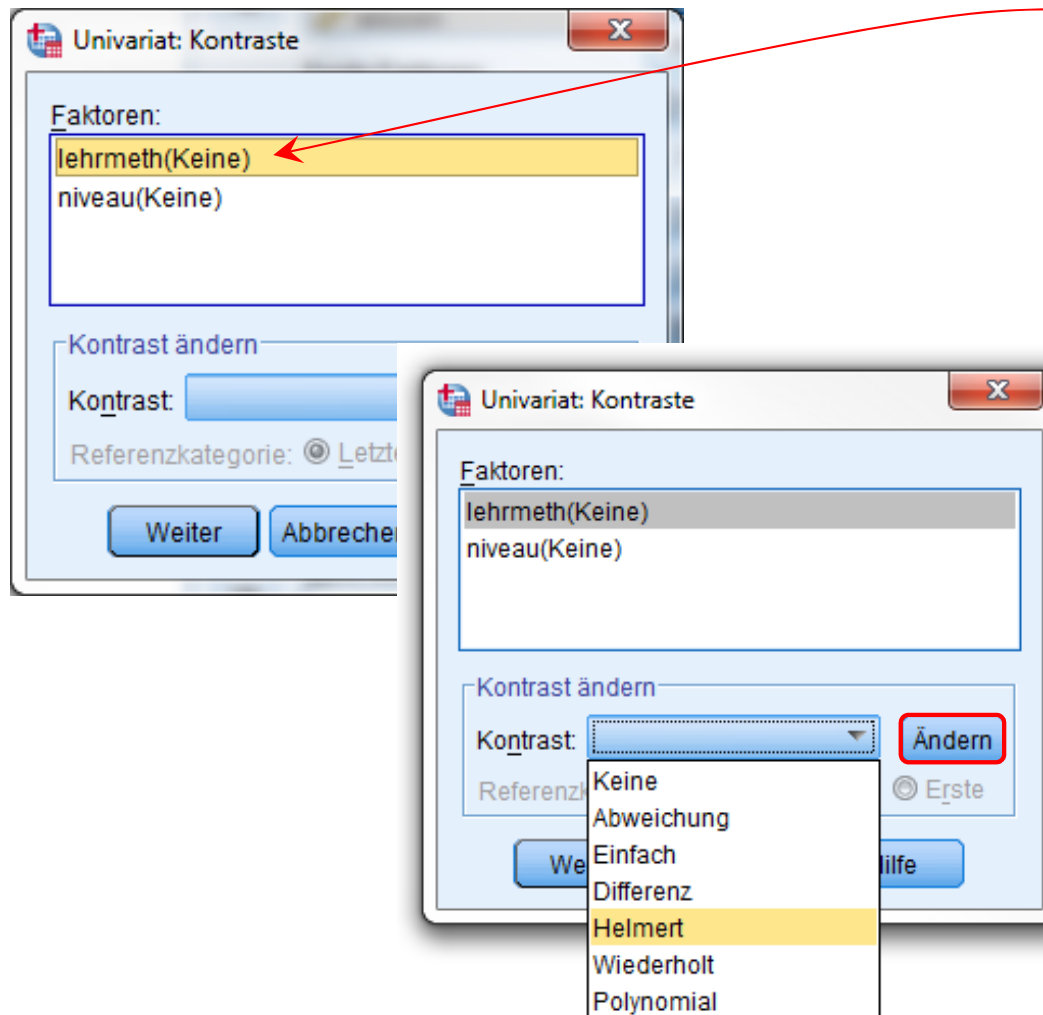
$$QS_L = \frac{q \cdot n_{\cdot\cdot} \cdot L^2}{\sum_{j=1}^m c_j^2} = \frac{2 \cdot 10 \cdot (-1 \cdot 43 + 1 \cdot 57.55 + 0 \cdot 59.8)^2}{(-1)^2 + 1^2 + 0^2} = \frac{20 \cdot (-14.55)^2}{2} = 2117.025 = MQ_L$$

$$F_L = \frac{MQ_L}{MQ_{in}} = \frac{2117.025}{161.643} = 13.097 \quad F_L > F_{crit} = F_{1, n-q \cdot m, 1-\alpha} = F_{1, 54, 0.95} = 4.02$$

- Die konventionelle Lehrbuchmethode ist also der verzweigten Lehrbuchmethode statistisch signifikant unterlegen.
- Analog sind auch komplexere Kontraste in den Randmitteln prüfbar, z.B. beide Lehrbuchmethoden vs. Web ($c_1 = -1, c_2 = -1, c_3 = 2$) oder Trends (in diesem Beispiel nicht möglich).

Zweifaktorielle Varianzanalyse: Treatmentkontraste

- SPSS stellt in der Prozedur Unianova unter (Kontraste) bestimmte vordefinierte a priori Treatmentkontraste bereit, die auf der kommenden Folie näher beschrieben sind.



Dazu ist hier der Faktor anzugeben, innerhalb dessen der Kontrast analysiert werden soll, dann der „Kontrast“ ausgewählt werden (z.B. „Helmert“) und dann mittels (Ändern) bestätigt werden. Ggf. ist noch die Referenzgruppe („Referenzkategorie“) abweichend auf „Erste“ einzustellen.

Falls gewünscht kann ebenso unabhängig davon ein Treatmentkontrast (z.B. „Differenz“) für den zweiten Faktor angefordert werden.

Zweifaktorielle Varianzanalyse: Treatmentkontraste

- SPSS stellt in der Prozedur Unianova folgende Kontraste bereit (am Beispiel wird verdeutlicht, welche Randmittel bei einem Faktor verglichen werden, der vier Gruppen aufweist):

| Bezeichnung | Vergleiche | Beispiel |
|-------------|---|----------------------------|
| Abweichung | Vergleich aller Gruppenmittelwerte – bis auf die der Gruppe mit der höchsten (oder wahlweise auch niedrigsten) Codierung – mit dem Gesamtmittelwert. | 1–G, 2–G, 3–G |
| Einfach | Vergleich aller Gruppenmittelwerte mit einer Referenzgruppe (z.B. der Kontrollgruppe), die entweder die höchste oder die niedrigste Codierung aufweisen muss. | 1–4, 2–4, 3–4 |
| Differenz | Sukzessiver Vergleich aller Gruppenmittelwerte mit den Mitteln aus den vorangegangenen Gruppen/Stufen. | 2–1, 3–{1,2}, 4–{1,2,3} |
| Helmert | Sukzessiver Vergleich aller Gruppenmittelwerte mit dem Mitteln aus den nachfolgenden Gruppen/Stufen. | 1–{2,3,4}, 2–{3,4}, 3–4 |
| Wiederholt | Vergleich jedes Gruppenmittelwertes mit dem Mittelwert der nachfolgenden Gruppe/Stufe. | 1–2, 2–3, 3–4 |
| Polynomial | Prüfung auf linearen, quadratischen und kubischen Trend. | |

G = Gesamtmittelwert. Bei „Abweichung“ und „Einfach“ ist im Beispiel angenommen, dass die Referenzgruppe die mit der höchsten Codierung ist. Die Kontraste Differenz, Helmert und Polynomial sind orthogonal.

Zweifaktorielle Varianzanalyse: Treatmentkontraste

- **Beispiel:** Die Hypothese, dass die verzweigte Lehrbuchmethode besser als die konventionelle Lehrbuchmethode ist, läuft auf einen Vergleich der Gruppen/Stufen 1 und 2 des Faktors *A* (Lehrmethode) hinaus. Dieser wird z.B. beim Kontrast „Differenz“ oder „Wiederholt“ ausgegeben:



Es ergibt sich ein statistisch signifikanter Kontrast; die verzweigte Lehrbuchmethode ist besser als die konventionelle.

| | | Abhängig... |
|------------------------------------|--|--|
| | | wissen |
| lehrmeth Differenzkontrast | | |
| Stufe 2 gegen Stufe 1 | Kontrastschätzer | 14,550 |
| | Hypothesenwert | 0 |
| | Differenz (Schätzung - Hypothesen) | 14,550 |
| | Standardfehler | 4,020 |
| | Sig. | ,001 |
| | 95% Konfidenzintervall für die Differenz | Untergrenze: 6,489 Obergrenze: 22,611 |
| Niveau 3 vs. vorhergehendes | | |
| | Kontrastschätzer | 9,525 |
| | Hypothesenwert | 0 |
| | Differenz (Schätzung - Hypothesen) | 9,525 |
| | Standardfehler | 3,482 |
| | Sig. | ,008 |
| | 95% Konfidenzintervall für die Differenz | Untergrenze: 2,544 Obergrenze: 16,506 |

L

- Findet sich der benötigte Paarvergleich nicht unter den vordefinierten (z.B. konventionelles Lehrbuch vs. Web) oder soll ein komplexerer Kontrast bestimmt werden, so ist dies in SPSS so ohne Weiteres nicht möglich. Im Prinzip sind solche Vergleiche mit dem `/LMATRIX-` Subkommando möglich, auf das wir später zurückkommen.

Zweifaktorielle Varianzanalyse: Treatmentkontraste

- Möglich ist auch, a posteriori alle möglichen Randmittel paarweise auf Unterschiedlichkeit zu prüfen.
- Dazu kann wieder der Tukey-Test herangezogen werden, bei der die Prüfgröße für den Vergleich der Spalten a und b

$$q_{T_{ab}} = \frac{\bar{x}_{a\bullet} - \bar{x}_{b\bullet}}{\sqrt{\frac{MQ_{in}}{q \cdot n_{\bullet\bullet}}}}$$

| \bar{x} | Lb kon | Lb ver | Web | |
|-----------|--------|--------|--------|--------|
| niedrig | 34.9 | 52.5 | 56.4 | 47.933 |
| hoch | 51.1 | 62.6 | 63.2 | 58.967 |
| | 43.000 | 57.550 | 59.800 | 53.450 |

Beim Vergleich aller Mittel des Zeilenfaktors B resultiert entsprechend:

$$q_{T_{ab}} = \frac{\bar{x}_{\bullet a} - \bar{x}_{\bullet b}}{\sqrt{\frac{MQ_{in}}{m \cdot n_{\bullet\bullet}}}}$$

sich von der im einfaktoriellen Fall nur dadurch unterscheidet, dass im Nenner statt n_{\bullet} jetzt $q \cdot n_{\bullet\bullet}$ Werte eingehen und MQ_{in} der zweifaktoriellen ANOVA eingesetzt wird.

Die Prüfgröße folgt unter den Voraussetzungen der Varianzanalyse einer studentisierten Range-Verteilung mit m Gruppen und $n - q \cdot m$ Freiheitsgraden. Der kritische Wert $q_{crit} = q_{m; n - q \cdot m; 1 - \alpha}$ kann entsprechenden Tabellen entnommen werden (z.B. Tabelle H in Diehl & Arbinger, 2001).

Zweifaktorielle Varianzanalyse: Treatmentkontraste

➤ Im modifizierten **Beispiel** resultiert:

$$q_{T_{12}} = \frac{\bar{x}_{1\cdot} - \bar{x}_{2\cdot}}{\sqrt{\frac{MQ_{in}}{q \cdot n_{\cdot\cdot}}}} = \frac{43 - 57.55}{\sqrt{\frac{161.643}{2 \cdot 10}}} = \frac{-14.55}{2.843} = -5.118$$

$$q_{T_{13}} = \frac{\bar{x}_{1\cdot} - \bar{x}_{3\cdot}}{\sqrt{\frac{MQ_{in}}{q \cdot n_{\cdot\cdot}}}} = \frac{43 - 59.8}{\sqrt{\frac{161.643}{2 \cdot 10}}} = \frac{-2.25}{2.843} = -5.909$$

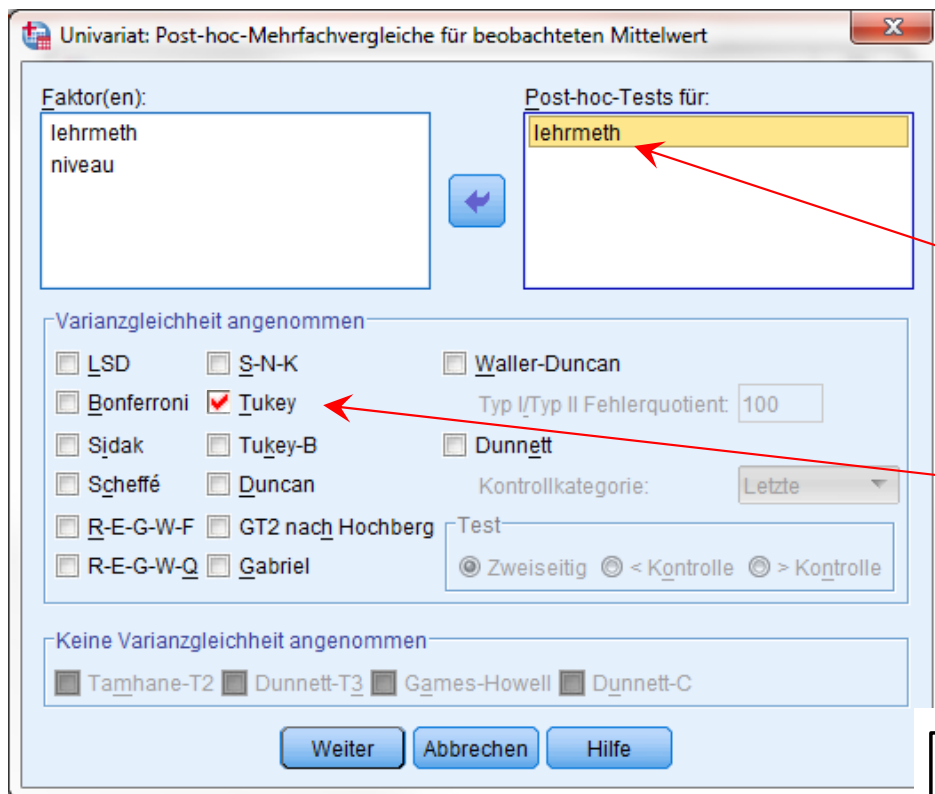
$$q_{T_{23}} = \frac{\bar{x}_{2\cdot} - \bar{x}_{3\cdot}}{\sqrt{\frac{MQ_{in}}{q \cdot n_{\cdot\cdot}}}} = \frac{57.55 - 59.8}{\sqrt{\frac{161.643}{2 \cdot 10}}} = \frac{-16.8}{2.843} = -0.791$$

| \bar{x} | Lb kon | Lb ver | Web | |
|-----------|--------|--------|--------|--------|
| niedrig | 34.9 | 52.5 | 56.4 | 47.933 |
| hoch | 51.1 | 62.6 | 63.2 | 58.967 |
| | 43.000 | 57.550 | 59.800 | 53.450 |

| Quelle | Quadratsumme vom Typ III | df | Mittel der Quadrate |
|---------------------|--------------------------|----|---------------------|
| Korrigiertes Modell | 5380,150 ^a | 5 | 1076,030 |
| Konstanter Term | 171414,150 | 1 | 1,714E5 |
| lehmeth | 3326,700 | 2 | 1663,350 |
| niveau | 1826,017 | 1 | 1826,017 |
| lehmeth * niveau | 227,433 | 2 | 113,717 |
| Fehler | 8728,700 | 54 | 161,643 |

➤ Tabellen entnimmt man einen kritischen q -Wert der studentisierten Range-Verteilung bei 3 Gruppen und $df = n - q \cdot m = 60 - 2 \cdot 3 = 54$ sowie $\alpha = 0.05$ von $q_{crit} = 3.408$. Nur für die beiden ersten Vergleiche gilt $|q_{T_{ab}}| > q_{crit}$. Entsprechend unterscheiden sich alle Lehrmethoden statistisch signifikant bis auf die beiden verzweigten.

Zweifaktorielle Varianzanalyse: Treatmentkontraste



Der Tukey-Test lässt sich in der Prozedur Unianova unter (Post-Hoc...) anfordern.

Hier sind die Faktoren anzugeben, deren Randmittel geprüft werden sollen.

Hier ist der Tukey-Test zu aktivieren (bzw. ein alternatives Verfahren).

Der Output rechts ist völlig analog zum einfaktoriellen Fall aufgebaut.

Multiple Comparisons

wissen
Tukey-HSD

| (I)lehrmeth | (J)lehrmeth | Mittlere Differenz (I-J) | Standard fehler | Sig. | 95%-Konfidenzintervall | |
|--------------------|--------------------|--------------------------|-----------------|------|------------------------|------------|
| | | | | | Untergrenze | Obergrenze |
| 1 Lb konventionell | 2 Lb verzweigt | -14,55* | 4,020 | ,002 | -24,24 | -4,86 |
| | 3 Web verzweigt | -16,80* | 4,020 | ,000 | -26,49 | -7,11 |
| 2 Lb verzweigt | 1 Lb konventionell | 14,55* | 4,020 | ,002 | 4,86 | 24,24 |
| | 3 Web verzweigt | -2,25 | 4,020 | ,842 | -11,94 | 7,44 |
| 3 Web verzweigt | 1 Lb konventionell | 16,80* | 4,020 | ,000 | 7,11 | 26,49 |
| | 2 Lb verzweigt | 2,25 | 4,020 | ,842 | -7,44 | 11,94 |

Grundlage: beobachtete Mittelwerte.
Der Fehlerterm ist Mittel der Quadrate(Fehler) = 161,643

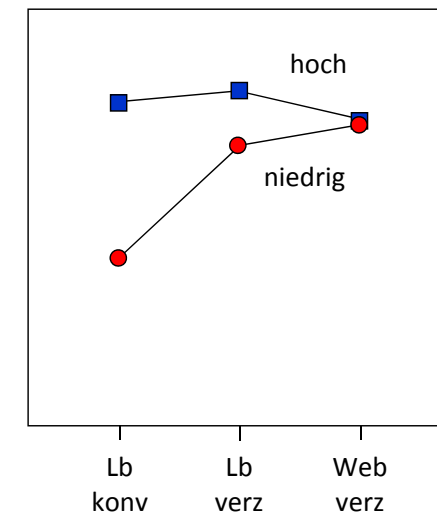
*. Die mittlere Differenz ist auf dem ,05-Niveau signifikant.

Zweifaktorielle Varianzanalyse: Einfache Haupteffekte

- Kommen wir auf die Original-Ergebnisse unseres **Beispiels** zurück. Dort lag eine Interaktion vor. In diesem Fall wollen wir die Interaktion näher charakterisieren.
- Dazu können wir uns dafür interessieren, ob ein Faktor auf einer festen Abstufung eines anderen Faktors einen Effekt ausübt (z.B. ob sich die Wirksamkeit der Lehrmethoden bei Schülern auf einem niedrigerem Leistungsniveau unterscheidet oder ob das Leistungsniveau bei der konventionellen Lehrmethode einen Effekt hat).
- Diese Effekte bezeichnet man auch als **einfache Haupteffekte** [simple (main) effects].

| | | Faktor A | | | |
|-----------|---------|----------|--------|------|--------|
| \bar{x} | | Lb kon | Lb ver | Web | |
| Faktor B | niedrig | 31.9 | 52.5 | 56.4 | 46.933 |
| | hoch | 60.4 | 62.6 | 57.4 | 60.133 |
| | | 46.15 | 57.55 | 56.9 | 53.533 |

| | | Faktor A | | | |
|-----------|---------|----------|--------|------|--------|
| \bar{x} | | Lb kon | Lb ver | Web | |
| Faktor B | niedrig | 31.9 | 52.5 | 56.4 | 46.933 |
| | hoch | 60.4 | 62.6 | 57.4 | 60.133 |
| | | 46.15 | 57.55 | 56.9 | 53.533 |



Zweifaktorielle Varianzanalyse: Einfache Haupteffekte

- Einfache Haupteffekte werden wie der F -Test in der einfaktoriellen Varianzanalyse bestimmt, mit dem einzigen Unterschied, dass im F -Test im Nenner die MQ_{in} aus der zweifaktoriellen Varianzanalyse eingesetzt wird.
- Die Berechnungen der erforderlichen Quadratsummen ergeben sich entsprechend als Summe der quadrierten Abweichungen aller Zellenmittelwerte in der betreffenden Zeile bzw. Spalte vom Zeilen- bzw. Spaltenmittelwert:

$$QS_{A|b_k} = n_{..} \cdot \sum_{j=1}^m (\bar{x}_{jk} - \bar{x}_{.k})^2 \text{ mit } df = m - 1$$

$$QS_{B|a_j} = n_{..} \cdot \sum_{k=1}^q (\bar{x}_{jk} - \bar{x}_{j.})^2 \text{ mit } df = q - 1$$

| | | Faktor A | | | |
|-------------|---------|----------|--------|------|--------|
| $\bar{x} :$ | | Lb kon | Lb ver | Web | |
| Faktor B | niedrig | 31.9 | 52.5 | 56.4 | 46.933 |
| | hoch | 60.4 | 62.6 | 57.4 | 60.133 |
| | | 46.15 | 57.55 | 56.9 | 53.533 |

| | | Faktor A | | | |
|-------------|---------|----------|--------|------|--------|
| $\bar{x} :$ | | Lb kon | Lb ver | Web | |
| Faktor B | niedrig | 31.9 | 52.5 | 56.4 | 46.933 |
| | hoch | 60.4 | 62.6 | 57.4 | 60.133 |
| | | 46.15 | 57.55 | 56.9 | 53.533 |

$A|b_k$ bezeichnet den Effekt des Faktors A auf der Stufe k des Faktors B.

Zweifaktorielle Varianzanalyse: Einfache Haupteffekte

- Im **Beispiel** können also prinzipiell zwei einfache Haupteffekte des Faktors *A* und drei einfache Haupteffekte des Faktors *B* untersucht werden.

| | | Faktor A | | | |
|----------|---------|-------------|--------|--------|--------|
| | | \bar{x} : | Lb kon | Lb ver | |
| Faktor B | niedrig | 31.9 | 52.5 | 56.4 | 46.933 |
| | hoch | 60.4 | 62.6 | 57.4 | 60.133 |
| | | 46.15 | 57.55 | 56.9 | 53.533 |

- Beim einfachen Haupteffekt A/b_1 wird beispielsweise untersucht, ob sich die Wirksamkeit der Lehrmethoden bei Schülern mit niedrigem Leistungsniveau unterscheidet (Es handelt sich also um den Effekt des Faktors *A* auf der Faktorstufe 1 des Faktors *B*):

$$\begin{aligned}
 QS_{A|b_1} &= n_{..} \cdot \sum_{j=1}^m (\bar{x}_{j1} - \bar{x}_{.1})^2 \\
 &= 10 \cdot \left((31.9 - 46.933)^2 + (52.5 - 46.933)^2 + (56.4 - 46.933)^2 \right) = 3466.067 \text{ mit } df = m - 1 = 2
 \end{aligned}$$

$$F_{A|b_1} = \frac{QS_{A|b_1} / df_{A|b_1}}{MQ_{in}} = \frac{3466.067 / 2}{157.426} = 11.01$$

$$F_{crit} = F_{m-1; n-m \cdot q; 1-\alpha} = F_{2; 54; 0.95} = 3.17$$

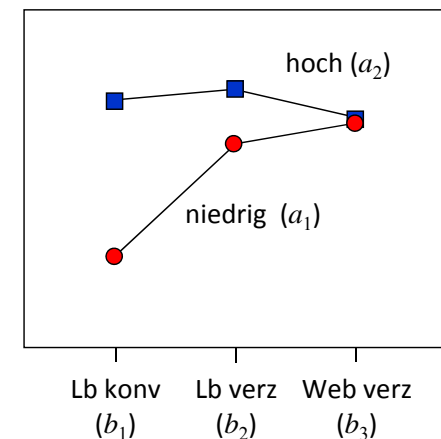
- Es resultiert also ein einfacher Haupteffekt: Die Lehrmethoden unterscheiden sich statistisch signifikant bei Schülern mit niedrigem Leistungsniveau.

Zweifaktorielle Varianzanalyse: Einfache Haupteffekte

- Führt man die Berechnungen analog für die anderen vier einfachen Haupteffekte aus, so kann man die Ergebnisse wieder in der üblichen varianzanalytischen Tabelle darstellen:

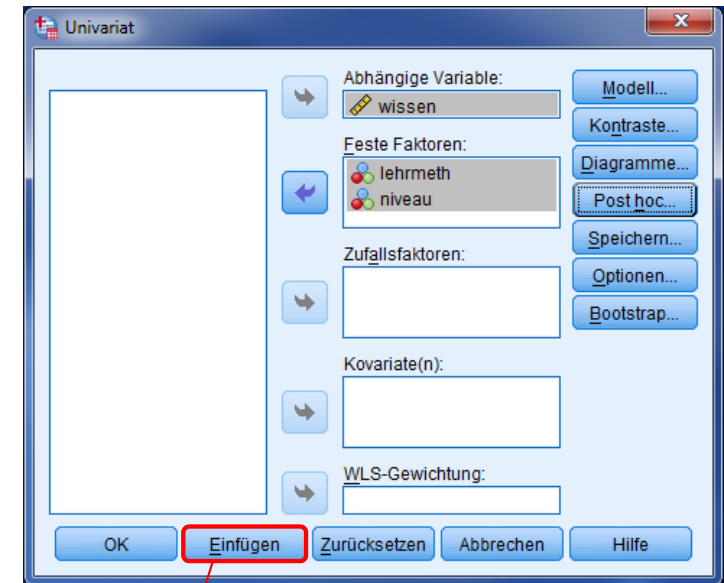
| | Variation | Quadratsummen | Freiheitsgrade | Mittlere Quadrate | F | p |
|---------|---------------------------------------|---------------|----------------|-------------------|--------|-------|
| $A b_1$ | Lehrmethode (niedriges Leistungsniv.) | 3466.067 | 2 | 1733.033 | 11.009 | 0.000 |
| $A b_2$ | Lehrmethode (hohes Leistungsniv.) | 136.267 | 2 | 68.133 | 0.433 | 0.651 |
| $B a_1$ | Leistungsniveau (Lb konventionell) | 4061.250 | 1 | 4061.250 | 25.798 | 0.000 |
| $B a_2$ | Leistungsniveau (Lb verzweigt) | 510.050 | 1 | 510.050 | 3.240 | 0.077 |
| $B a_3$ | Leistungsniveau (Web verzweigt) | 5.000 | 1 | 5.000 | 0.032 | 0.859 |
| | Innerhalb | 8501.000 | 54 | 157.426 | | |

- Sinnvoll ist hier wieder, gezielt nur die inhaltlich interessanten Effekte zu prüfen. Die im Kontext der geplanten Kontraste dargestellten Überlegungen zur α -Adjustierung gelten hier und im Folgenden analog.



Zweifaktorielle Varianzanalyse: Einfache Haupteffekte mittels SPSS

- Ein einfacher Haupteffekt kann in SPSS in der Prozedur Unianova nur über die Syntax angefordert werden.
- Hierzu fügt man nach Spezifizierung der Variablen zunächst wie gehabt die Syntax über die Prozedur ein. Dieser ist dann das Subkommando `/EMMEANS` (estimated marginal means) wie folgt hinzuzufügen:



```
UNIANOVA wissen BY lehrmeth niveau  
  /METHOD=SSTYPE(3)  
  /INTERCEPT=INCLUDE  
  /CRITERIA=ALPHA(0.05)  
  /DESIGN=lehrmeth niveau lehrmeth*niveau  
  /EMMEANS=TABLES(lehrmeth*niveau) COMPARE (lehrmeth).
```

UV₁ UV₂

UV, deren einfachen Haupteffekte bestimmt werden sollen. Hier im Bsp. `lehrmeth` oder `niveau`.

Zweifaktorielle Varianzanalyse: Einfache Haupteffekte mittels SPSS

- Als Ergebnis werden die einfachen Haupteffekte für die Lehrmethoden ausgegeben:

Tests auf Univariate

Abhängige Variable:wissen

| niveau | | Quadratsumme | df | Mittel der Quadrate | F | Sig. |
|-----------|----------|--------------|----|---------------------|--------|------|
| 0 niedrig | Kontrast | 3466,067 | 2 | 1733,033 | 11,009 | ,000 |
| | Fehler | 8501,000 | 54 | 157,426 | | |
| 1 hoch | Kontrast | 136,267 | 2 | 68,133 | ,433 | ,651 |
| | Fehler | 8501,000 | 54 | 157,426 | | |

Kontrast $A|b_1$

Kontrast $A|b_2$

F prüft den Effekt von lehrmeth. Diese Test basiert auf den linear unabhängigen, paarweisen Vergleichen bei den geschätzten Randmitteln. Diese Tests beruhen auf den linear unabhängigen paarweisen Vergleichen zwischen den geschätzten Randmitteln.

- Bei den Schüler mit niedrigen Leistungsniveaus unterscheiden sich also die drei Lehrmethoden statistisch signifikant. Bei den Schülern mit hohem Leistungsniveau hingegen findet sich kein statistisch signifikanter Unterschied der Lehrmethoden.

Zweifaktorielle Varianzanalyse: Einfache Kontraste

- Weist ein Faktor mehr als zwei Faktorstufen auf und ergibt sich für diesen Faktor ein signifikanter einfacher Haupteffekt auf einer Stufe des anderen Faktors, so interessieren wir uns häufig dafür, welche Gruppen sich unterscheiden.

| | | Faktor A | | | |
|-------------|--------|----------|------|--------|--|
| \bar{x} : | Lb kon | Lb ver | Web | | |
| niedrig | 31.9 | 52.5 | 56.4 | 46.933 | |
| hoch | 60.4 | 62.6 | 57.4 | 60.133 | |
| | 46.15 | 57.55 | 56.9 | 53.533 | |

- Im Beispiel wissen wir, dass sich die drei Lehrmethoden in ihrer Effektivität bei Schülern mit schwächerem Leistungsniveau unterscheiden. Wir wissen noch nicht, zwischen welchen Lehrmethoden bei den schwächeren Schülern ein Unterschied besteht.
- Hier können wir einen **einfachen Kontrast (simple contrast)** bestimmen, z.B. den Kontrast A_1 bezüglich des Vergleichs der beiden Lehrbuchmethoden bei schwächeren Schülern.

$$QS_{A_1|b_1} = \frac{n_{..} \cdot L^2}{\sum_{j=1}^m c_j^2} = \frac{n_{..} \cdot (-1 \cdot \bar{x}_{11} + 1 \cdot \bar{x}_{21} + 0 \cdot \bar{x}_{31})^2}{(-1)^2 + 1^2 + 0^2} = \frac{10 \cdot (-31.9 + 52.5)^2}{2} = \frac{10 \cdot 20.6^2}{2} = 2121.8 \text{ mit } df = 1$$

$$F_L = \frac{MQ_L}{MQ_{in}} = \frac{2121.8}{157.426} = 13.48 \quad F_{crit} = F_{1;n-m;q;1-\alpha} = F_{1;54;0.95} = 4.02$$

Die verzweigte Lehrbuchmethode ist also bei schwächeren Schülern statistisch signifikant besser als die unverzweigte Lehrbuchmethode.

Zweifaktorielle Varianzanalyse: Einfache Kontraste mittels SPSS

- Wenn wie oben durch das Subkommando `\EMMEANS` die einfachen Haupteffekte angefordert werden, gibt SPSS automatisch auch alle paarweisen einfachen Kontraste (ohne α -Adjustierung) aus:

| | | Faktor A | | |
|-------------|--|----------|--------|------|
| \bar{x} : | | Lb kon | Lb ver | Web |
| niedrig | | 31.9 | 52.5 | 56.4 |
| hoch | | 60.4 | 62.6 | 57.4 |

Paarweise Vergleiche

Abhängige Variable:wissen

| niveau | (I)lehrmeth | (J)lehrmeth | Mittlere Differenz (I-J) | Standardfehler | Sig. ^a | 95% Konfidenzintervall für die Differenz ^a | |
|-----------|--------------------|--------------------|--------------------------|----------------|-------------------|---|------------|
| | | | | | | Untergrenze | Obergrenze |
| 0 niedrig | 1 Lb konventionell | 2 Lb verzweigt | -20,600* | 5,611 | ,001 | -31,850 | -9,350 |
| | | 3 Web verzweigt | -24,500* | 5,611 | ,000 | -35,750 | -13,250 |
| | 2 Lb verzweigt | 1 Lb konventionell | 20,600* | 5,611 | ,001 | 9,350 | 31,850 |
| | | 3 Web verzweigt | -3,900 | 5,611 | ,490 | -15,150 | 7,350 |
| | 3 Web verzweigt | 1 Lb konventionell | 24,500* | 5,611 | ,000 | 13,250 | 35,750 |
| | | 2 Lb verzweigt | 3,900 | 5,611 | ,490 | -7,350 | 15,150 |
| 1 hoch | 1 Lb konventionell | 2 Lb verzweigt | -2,200 | 5,611 | ,697 | -13,450 | 9,050 |
| | | 3 Web verzweigt | 3,000 | 5,611 | ,595 | -8,250 | 14,250 |
| | 2 Lb verzweigt | 1 Lb konventionell | 2,200 | 5,611 | ,697 | -9,050 | 13,450 |
| | | 3 Web verzweigt | 5,200 | 5,611 | ,358 | -6,050 | 16,450 |
| | 3 Web verzweigt | 1 Lb konventionell | -3,000 | 5,611 | ,595 | -14,250 | 8,250 |
| | | 2 Lb verzweigt | -5,200 | 5,611 | ,358 | -16,450 | 6,050 |

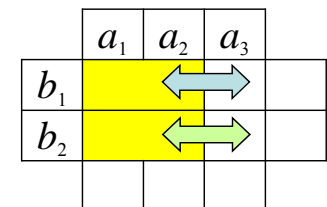
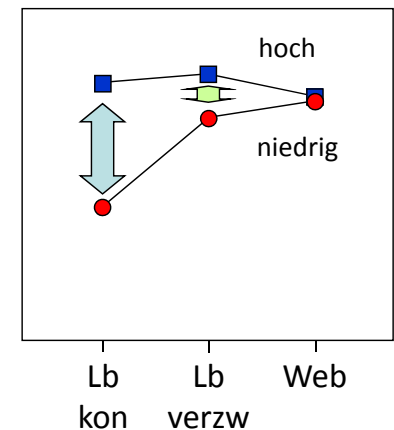
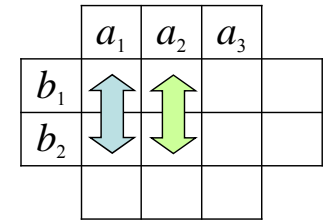
Basiert auf den geschätzten Randmitteln

*. Die mittlere Differenz ist auf dem ,05-Niveau signifikant.

a. Anpassung für Mehrfachvergleiche: Geringste signifikante Differenz (entspricht keinen Anpassungen).

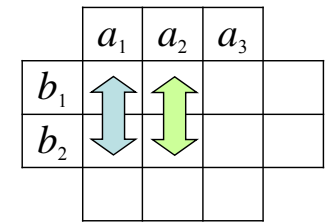
Zweifaktorielle Varianzanalyse: Interaktionskontraste

- Ferner können wir uns bei einer vorliegenden Interaktion dafür interessieren, ob Unterschiede in den Zellenmitteln eines Faktors sich auf bestimmten Abstufungen des anderen Faktors unterscheiden. Die entsprechenden Kontraste bezeichnet man als **Interaktionskontraste**.
- Wir wollen also z.B. wissen, ob der Unterschied zwischen besseren und schlechteren Schülern bei beiden Lehrbuchmethoden gleich ist. Sind also die besseren Schüler bei beiden Methoden gleichermaßen besser?
- Es sei darauf hingewiesen, dass man diese Frage nicht anhand der Signifikanzprüfungen der einfachen Haupteffekte entscheiden kann. Es kann also z.B. sein, dass ein einfacher Haupteffekt signifikant ist und der gleiche einfache Haupteffekt unter einer anderen Faktorstufe nicht, sich die beiden aber trotzdem nicht statistisch signifikant unterscheiden.
- Wir könnten auch komplexer wissen wollen, ob die Überlegenheit der computergestützten Methode gegenüber den beiden anderen Lehrmethoden bei besseren und schwächeren Schülern gleich ist.



Zweifaktorielle Varianzanalyse: Interaktionskontraste

- Allgemein kann man Interaktionskontraste aufstellen und prüfen, indem man einen Kontrast in Faktor A mit einem Kontrast in Faktor B kombiniert.



- Bezeichnen wir beide Kontraste mit

$$L_A = \sum_{j=1}^m c_j \cdot \bar{x}_{j\bullet} \quad \text{und} \quad L_B = \sum_{k=1}^q d_k \cdot \bar{x}_{\bullet k}$$

dann ergibt sich als Interaktionskontrast

$$L_{AxB} = \sum_{j=1}^m \sum_{k=1}^q c_j \cdot d_k \cdot \bar{x}_{jk}$$

| | | | | |
|-----------------|----|-------|----|---|
| | | c_j | | |
| $c_j \cdot d_k$ | | -1 | 1 | 0 |
| d_k | -1 | 1 | -1 | 0 |
| | 1 | -1 | 1 | 0 |

- **Beispiel-Fragestellung:** Ist der Unterschied zwischen besseren und schlechteren Schülern bei beiden Lehrbuchmethoden gleich?

| \bar{x}_{jk} | Lb kon | Lb ver | Web |
|----------------|--------|--------|------|
| niedrig | 31.9 | 52.5 | 56.4 |
| hoch | 60.4 | 62.6 | 57.4 |

$$L_{AxB} = \sum_{j=1}^m \sum_{k=1}^q c_j \cdot d_k \cdot \bar{x}_{jk} = (-1) \cdot (-1) \cdot 31.9 + (-1) \cdot 1 \cdot 60.4 + \dots + 0 \cdot 1 \cdot 57.4 = -18.4$$

$$= 1 \cdot 31.9 - 1 \cdot 60.4 - 1 \cdot 52.5 + 1 \cdot 62.6 + 0 \cdot 56.4 + 0 \cdot 57.4 = -18.4$$

Zweifaktorielle Varianzanalyse: Interaktionskontraste

- Entsprechend erhalten wir die Quadratsumme

$$QS_{L_{AxB}} = \frac{n_{..} \cdot L_{AxB}^2}{\sum_{j=1}^m \sum_{k=1}^q (c_j \cdot d_k)^2} = \frac{10 \cdot (-18.4)^2}{1^2 + (-1)^2 + (-1)^2 + 1^2 + 0^2 + 0^2} = \frac{3385.6}{4} = 846.4 \text{ mit } df = 1$$

und den F -Wert

$$F_{L_{AxB}} = \frac{MQ_{L_{AxB}}}{MQ_{in}} = \frac{846.4}{157.426} = 5.38 \quad F_{crit} = F_{1;n-m \cdot q;1-\alpha} = F_{1;54;0.95} = 4.02$$

| | c_j | | |
|-----------------|-------|----|---|
| $c_j \cdot d_k$ | -1 | 1 | 0 |
| d_k | -1 | 1 | 0 |
| | 1 | -1 | 0 |

Der Unterschied zwischen den besseren und schlechteren Schülern ist also bei der unverzweigten Lehrbuchmethode statistisch signifikant größer als bei der verzweigten.

- Interaktionskontraste sind Komponenten der Interaktion AxB und unabhängig zu beiden Haupteffekten. Bildet man eine Menge von $m - 1$ orthogonalen Kontrasten für den Faktor A und eine Menge von $q - 1$ orthogonalen Kontrasten für den Faktor B und dann für alle möglichen $(m - 1) \cdot (q - 1)$ Kombinationen Interaktionskontraste, so addieren sich deren Quadratsummen zu QS_{AxB} auf. Die so gebildeten Interaktionskontraste sind dann also ebenfalls alle orthogonal (nicht-überlappend) und enthalten keine redundanten Informationen.

Zweifaktorielle Varianzanalyse: Interaktionskontraste

- Im obigen Beispiel für einen Interaktionskontrast haben wir geprüft, ob sich ein Kontrast in bestimmten Gruppen des anderen Faktors gleich zeigt (Interaktion zweier Kontraste).
- Im Beispiel: Unterscheiden sich schlechtere und bessere Schüler bei beiden Lehrbuchmethoden in gleicher Weise? Die Web-Gruppe a_3 blieb also unberücksichtigt.

| | a_1 | a_2 | a_3 | |
|-------|-------|-------|-------|--|
| b_1 | ↕ | ↕ | | |
| b_2 | ↕ | ↕ | | |
| | | | | |

- Bei einer größeren Zahl an Gruppen in den Faktoren, z.B. einem 3×3 Design, kann man auch testen, ob sich ein Kontrast in allen Gruppen des anderen Faktors gleich zeigt (Interaktion Kontrast \times Faktor). Wie man diese Kontraste bestimmt, kann man in Keppel & Wickens (2004, Kap. 13.4) nachlesen.

| | a_1 | a_2 | a_3 | |
|-------|-------|-------|-------|--|
| b_1 | ↕ | ↕ | ↕ | |
| b_2 | ↕ | ↕ | ↕ | |
| b_3 | | | | |
| | | | | |

Zweifaktorielle Varianzanalyse: Interaktionskontraste in SPSS

- Interaktionskontraste kann man in SPSS in der Prozedur Unianova nur über die Syntax anfordern. Hierzu fügt man wieder (wie bei /EMMEANS) nach Spezifizierung der Variablen zunächst wie gehabt die Syntax über die Prozedur ein. Der Syntax ist dann das Subkommando /LMATRIX wie folgt hinzuzufügen:

```
UNIANOVA wissen BY niveau lehrmeth  
  /METHOD=SSTYPE(3)  
  /INTERCEPT=INCLUDE  
  /CRITERIA=ALPHA(0.05)  
  /DESIGN=niveau lehrmeth niveau*lehrmeth  
  /LMATRIX niveau*lehrmeth 1 -1 0  
                                     -1 1 0.
```

$UV_1 * UV_2$

| | | | | | | |
|-----------------|-------|----------|-------|---|----------|--|
| | | Faktor B | | | | |
| | | | c_j | | | |
| $c_j \cdot d_k$ | | -1 | 1 | 0 | | |
| | d_k | -1 | -1 | 0 | Faktor A | |
| | | 1 | 1 | 0 | | |

Hier ist die Matrix der $c_j \cdot d_k$ -Koeffizienten einzugeben.

Die Reihenfolge der Faktorstufen richtet sich nach der Reihenfolge der Variablen hinter dem BY Schlüsselwort in der ersten Zeile des Befehls. (Die Reihenfolge im LMATRIX-Befehl ist ohne Bedeutung.) Im Falle von BY A B müssen also die Koeffizienten in der Reihenfolge der Zellen (a_1, b_1) , (a_1, b_2) , (a_1, b_3) , (a_2, b_1) ... folgen.

Benutzerdefinierter Hypothesentest

Kontrastergebnisse (K-Matrix)^a

| Kontrast | | Abhängig... |
|----------|--|-------------|
| | | wissen |
| L1 | Kontrastschätzer | -18,400 |
| | Hypothesenwert | 0 |
| | Differenz (Schätzung - Hypothesen) | -18,400 |
| | Standardfehler | 7,935 |
| | Sig. | ,024 |
| | 95% Konfidenzintervall für die Differenz | |
| | Untergrenze | -34,309 |
| | Obergrenze | -2,491 |

L

Die Ergebnistabelle ist aufgebaut wie beim Treatmentkontrast.

a. Basiert auf der (L')-Matrix der benutzerdefinierten Kontrastkoeffizienten Nummer 1

7 Kontraste

8 Effektstärken

9 Ungleiche Stichprobenumfänge

10 Designs mit zufälligen Faktoren

Zweifaktorielle Varianzanalyse: Effektstärken

- Als **Effektstärkemaß** kann als Verallgemeinerung des einfaktoriellen Falles wieder das Verhältnis der durch den Effekt erklärten Varianz an der Gesamtvarianz herangezogen werden:

$$\hat{\eta}_{\text{Effekt}}^2 = \frac{QS_{\text{Effekt}}}{QS_{\text{tot}}} = \frac{QS_{\text{Effekt}}}{QS_A + QS_B + QS_{AxB} + QS_{in}}$$

- Im zweifaktoriellen Fall bestehen drei Effekte, deren Stärke entsprechend quantifiziert werden können, im Beispiel:

$$\hat{\eta}_A^2 = \frac{QS_A}{QS_{\text{tot}}} = \frac{1639.633}{14716.933} = 0.11$$

$$\hat{\eta}_B^2 = \frac{QS_B}{QS_{\text{tot}}} = \frac{2613.600}{14716.933} = 0.18$$

$$\hat{\eta}_{AxB}^2 = \frac{QS_{AxB}}{QS_{\text{tot}}} = \frac{1962.700}{14716.933} = 0.13$$

| Quelle | Quadratsumme vom Typ III |
|-----------------------------|--------------------------|
| Korrigiertes Modell | 6215,933 ^a |
| Konstanter Term | 171949,067 |
| lehrmeth | 1639,633 |
| niveau | 2613,600 |
| lehrmeth * niveau | 1962,700 |
| Fehler | 8501,000 |
| Gesamt | 186666,000 |
| Korrigierte Gesamtvariation | 14716,933 |

= QS_{tot}

- Dabei summieren sich die Effektstärken zur insgesamt aufgeklärten Effektstärke $\hat{\eta}_{\text{zw}}^2$ auf:

$$\hat{\eta}_{\text{zw}}^2 = \frac{QS_{\text{zw}}}{QS_{\text{tot}}} = \frac{6215.933}{14716.933} = 0.42 = \hat{\eta}_A^2 + \hat{\eta}_B^2 + \hat{\eta}_{AxB}^2$$

Zweifaktorielle Varianzanalyse: Effektstärken

- Bei der Wahl des Effektstärkemaßes ist die Vergleichbarkeit über Studien hinweg relevant. Die Frage stellt sich besonders bei Studien, bei der nur eine Teilmenge der UVs überlappen.
- Um die Vergleichbarkeit zu erhöhen, empfehlen einige Autoren statt der Verwendung von $\hat{\eta}^2$ den Einsatz des als **partielles $\hat{\eta}_p^2$** bezeichneten Effektstärkemaßes, bei dem im Nenner die Quadratsummen aller anderen Effekte herausgezogen (herauspartialisiert) werden:

$$\hat{\eta}_{p_{\text{Effekt}}}^2 = \frac{QS_{\text{Effekt}}}{QS_{\text{Effekt}} + QS_{\text{in}}} = \frac{F_{\text{Effekt}} \cdot df_{\text{Effekt}}}{F_{\text{Effekt}} \cdot df_{\text{Effekt}} + df_{\text{in}}}$$

$$\text{z.B. } \hat{\eta}_{p_A}^2 = \frac{QS_A}{QS_A + QS_{\text{in}}} = \frac{QS_A}{QS_{\text{tot}} - QS_B - QS_{A \times B}}$$

- Das partielle Maß gibt also an, wie viel Variabilität ein Effekt an der Variabilität erklärt, die nicht durch die anderen Faktoren erklärt wird. Offensichtlich gilt $\hat{\eta}_{p_{\text{Effekt}}}^2 \geq \hat{\eta}_{\text{Effekt}}^2$. Die Additivität gilt nicht mehr.

- Im **Beispiel** ergibt sich für Faktor A:

$$\hat{\eta}_{p_A}^2 = \frac{QS_A}{QS_A + QS_{\text{in}}} = \frac{1639.633}{1639.633 + 8501} = 0.16$$

sowie $\hat{\eta}_{p_B}^2 = 0.24$ und $\hat{\eta}_{p_{A \times B}}^2 = 0.19$

| Quelle | Quadratsumme vom Typ III |
|---------------------|--------------------------|
| Korrigiertes Modell | 6215,933 ^a |
| Konstanter Term | 171949,067 |
| lehmeth | 1639,633 |
| niveau | 2613,600 |
| lehmeth * niveau | 1962,700 |
| Fehler | 8501,000 |

Zweifaktorielle Varianzanalyse: Effektstärken

- Wie schon bei der einfaktoriellen Varianzanalyse gilt, dass die η -Koeffizienten deskriptiv eine anschauliche Deutung ermöglichen, aber verzerrte Schätzungen der Populationseffektstärken

$$\eta_{\text{Effekt}}^2 = \omega_{\text{Effekt}}^2 = \frac{\sigma_{\text{Effekt}}^2}{\sigma_{\text{tot}}^2} \quad \text{bzw.} \quad \eta_{p_{\text{Effekt}}}^2 = \omega_{p_{\text{Effekt}}}^2 = \frac{\sigma_{\text{Effekt}}^2}{\sigma_{\text{Effekt}}^2 + \sigma_{\text{in}}^2}$$

darstellen. Bessere Schätzungen liefern wieder ω -Koeffizienten. Diese benötigen zur Berechnung die folgenden Schätzungen der Populationsvarianzen, für die im Falle fixer Effekte für A und B gilt (Howell, 2010, S. 438ff):

$$\hat{\sigma}_A^2 = \frac{m-1}{n} \cdot (MQ_A - MQ_{in})$$

$$\hat{\sigma}_B^2 = \frac{q-1}{n} \cdot (MQ_B - MQ_{in})$$

$$\hat{\sigma}_{AxB}^2 = \frac{(m-1) \cdot (q-1)}{n} \cdot (MQ_{AxB} - MQ_{in})$$

$$\hat{\sigma}_{in}^2 = MQ_{in}$$

$$\hat{\sigma}_{\text{tot}}^2 = \hat{\sigma}_A^2 + \hat{\sigma}_B^2 + \hat{\sigma}_{AxB}^2 + \hat{\sigma}_{in}^2$$

Zweifaktorielle Varianzanalyse: Effektstärken

- Für die ω -Koeffizienten gilt (Howell, 2010, S. 438ff; Keppel & Wickens, 2004, S. 232ff):

$$\hat{\omega}_{\text{Effekt}}^2 = \frac{\hat{\sigma}_{\text{Effekt}}^2}{\hat{\sigma}_{\text{tot}}^2} = \frac{(F_{\text{Effekt}} - 1) \cdot df_{\text{Effekt}}}{(F_A - 1) \cdot df_A + (F_B - 1) \cdot df_B + (F_{AxB} - 1) \cdot df_{AxB} + n}$$

$$\hat{\omega}_{P_{\text{Effekt}}}^2 = \frac{\hat{\sigma}_{\text{Effekt}}^2}{\hat{\sigma}_{\text{Effekt}}^2 + \hat{\sigma}_{\text{in}}^2} = \frac{(F_{\text{Effekt}} - 1) \cdot df_{\text{Effekt}}}{(F_{\text{Effekt}} - 1) \cdot df_{\text{Effekt}} + n}$$

| Quelle | Quadratsumme vom Typ III | df | Mittel der Quadrate |
|-------------------|--------------------------|----|---------------------|
| lehrmeth | 1639,633 | 2 | 819,817 |
| niveau | 2613,600 | 1 | 2613,600 |
| lehrmeth * niveau | 1962,700 | 2 | 981,350 |
| Fehler | 8501,000 | 54 | 157,426 |

- Im **Beispiel** resultieren als Varianzschätzungen:

$$\hat{\sigma}_A^2 = \frac{m-1}{n} \cdot (MQ_A - MQ_{in}) = \frac{3-1}{60} \cdot (819.817 - 157.426) = 22.080$$

$$\hat{\sigma}_B^2 = \frac{q-1}{n} \cdot (MQ_B - MQ_{in}) = \frac{2-1}{60} \cdot (2613.6 - 157.426) = 40.936$$

$$\hat{\sigma}_{AxB}^2 = \frac{(m-1) \cdot (q-1)}{n} \cdot (MQ_{AxB} - MQ_{in}) = \frac{(3-1) \cdot (2-1)}{60} \cdot (981.350 - 157.426) = 27.464$$

$$\hat{\sigma}_{in}^2 = MQ_{in} = 157.426$$

$$\hat{\sigma}_{\text{tot}}^2 = \hat{\sigma}_A^2 + \hat{\sigma}_B^2 + \hat{\sigma}_{AxB}^2 + \hat{\sigma}_{in}^2 = 22.080 + 40.936 + 27.464 + 157.426 = 247.906$$

Zweifaktorielle Varianzanalyse: Effektstärken

- Entsprechend ergibt sich z.B. für den Faktor A:

$$\hat{\omega}_A^2 = \frac{\hat{\sigma}_A^2}{\hat{\sigma}_{total}^2} = \frac{22.080}{247.906} = 0.09$$

$$\hat{\omega}_{p_A}^2 = \frac{\hat{\sigma}_A^2}{\hat{\sigma}_A^2 + \hat{\sigma}_{in}^2} = \frac{22.080}{22.080 + 157.426} = 0.12$$

und für die anderen Effekte:

$$\hat{\omega}_B^2 = 0.17 \quad \hat{\omega}_{p_B}^2 = 0.21$$

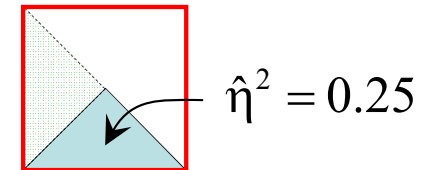
$$\hat{\omega}_{AxB}^2 = 0.11 \quad \hat{\omega}_{p_{AxB}}^2 = 0.15$$

- Die Konventionen lauten wieder in Anlehnung an Cohen (1988): ≈ 0.01 „kleiner Effekt“, ≈ 0.06 „mittlerer Effekt“ und ≈ 0.14 „großer Effekt“.

Zweifaktorielle Varianzanalyse: Effektstärken

- Wann sollte man partielle Maße verwenden? Dadurch, dass die Varianzen der Effekte immer an anderen „Gesamt“-Varianzen normiert werden, resultiert einerseits der Nachteil der Nicht-Additivität und andererseits die schlechte Vergleichbarkeit der Effekte innerhalb einer Studie (weswegen man sie zu diesem Zweck nicht heranziehen sollte).

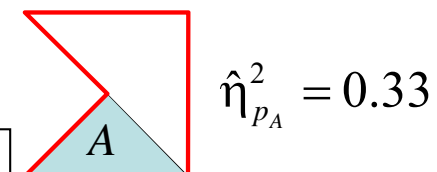
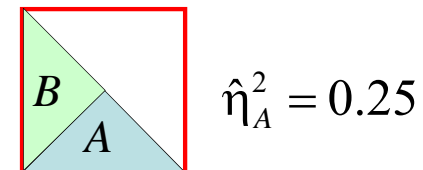
- Nehmen wir an, in einer Studie S hätte sich bei Schülern einfaktoriell ein Effekt dreier Lehrmethoden auf die Leistung der Stärke $\hat{\eta}^2 = 0.25$ ergeben.



- Wir führen nun eine Studie N durch, um zu untersuchen, ob die Leistung nicht nur von der Wirksamkeit der Lehrmethoden (Faktor A) sondern auch vom Geschlecht (Faktor B; oder der Vorerfahrung der Schüler) abhängt. Wenn wir die Stärke des Effekts der Lehrmethode im Vergleich zur Studie S betrachten wollten, sollten wir dazu ein partielles Maß verwenden?

- In den Beispielen ist die Variabilität, die auf die zweiten Faktoren zurückgeht, auch bereits in der Studie S enthalten (in der Fehlervarianz QS_{in}). Daher ist es sinnvoller, kein partielles Maß zu verwenden, denn wir können davon ausgehen, dass dieser Faktor keine zusätzliche Varianz erzeugt. Würden wir ein partielles Maß verwenden, so würde die Varianz des Effektes B im Nenner herauspartialisiert, die aber in der ersten Studie enthalten ist. (Relativ würde unser Maß dadurch zu groß werden.)

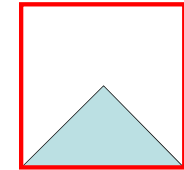
Studie N



(In den Abbildungen wird eine ggf. bestehende Interaktion nicht berücksichtigt.)

Zweifaktorielle Varianzanalyse: Effektstärken

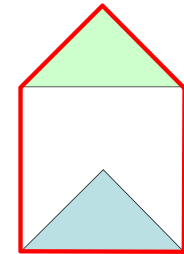
- Nehmen wir an, in einer Studie S hätte sich bei Schülern einfaktoriellement ein Effekt dreier Lehrmethoden auf die Leistung der Stärke $\hat{\eta}^2 = 0.25$ ergeben.



$$\hat{\eta}^2 = 0.25$$

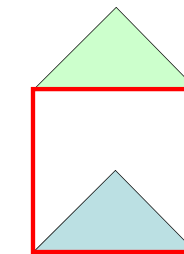
- Was wäre der Fall, wenn wir in unserer Studie N als zweiten Faktor B untersuchen wollten, wie es sich auswirkt, wenn man zusätzlich experimentelle Übungen anbietet (oder die Unterrichtsdauer verlängert oder vorab klare Lehrziele formuliert)?

- In diesen Beispielen führen unsere zweiten Faktoren B (z.B. „Übungen vs. nicht“) zu zusätzlicher Varianz, denn in der Studie S hatte keiner der Schüler an solchen Übungen teilgenommen, es wurden vorab keine Lehrziele formuliert etc. In diesem Fall führen partielle Maße zu einer besseren Vergleichbarkeit. (Im Falle nicht-partieller Maße würde der Effekt in Studie N relativ unterschätzt.)



$$\hat{\eta}_A^2 = 0.20$$

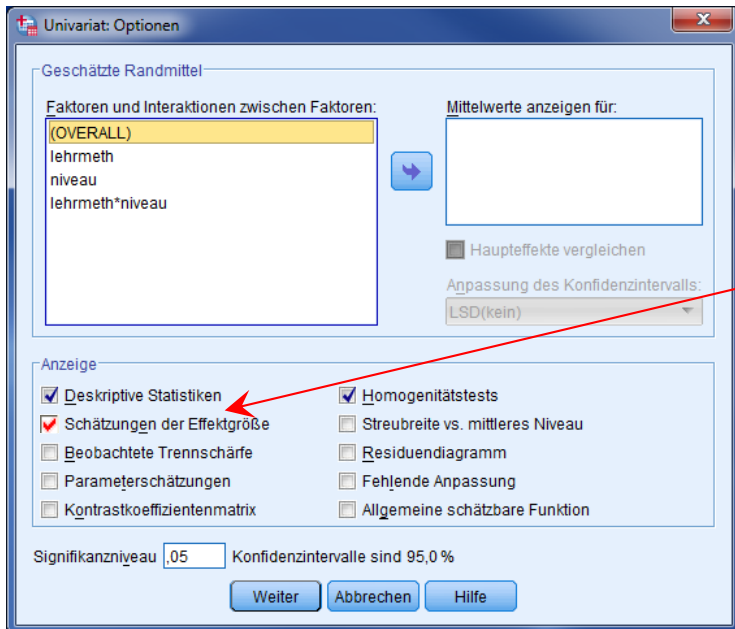
- Häufiger ist aber nicht klar, in welchem Ausmaß ein zusätzlicher Faktor die Varianz erhöht, z.B. bei Faktoren wie z.B. sozioökonomischer Status der Schüler, Erfahrung der Lehrer usw.



$$\hat{\eta}_{PA}^2 = 0.25$$

- Außerdem ist natürlich nicht immer klar, womit (wenn überhaupt) verglichen werden soll. Insofern ist die Entscheidung schwieriger. In jedem Fall muss berichtet werden, welches Maß berechnet wurde. (Ggf. können die anderen meist auch rekonstruiert werden.)

Zweifaktorielle Varianzanalyse: Effektstärken



In der Prozedur Unianova kann man unter (Opti onen) „Schätzer der Effektgrößen“ anfordern. In diesem Fall gibt SPSS (nur) die partiellen Eta-Effektstärken $\hat{\eta}_{P_{Effekt}}^2$ aus.

Tests der Zwischensubjekteffekte

Abhängige Variable:wissen

| Quelle | Quadratsumme vom Typ III | df | Mittel der Quadrate | F | Sig. | Partielles Eta-Quadrat |
|-----------------------------|--------------------------|----|---------------------|----------|------|------------------------|
| Korrigiertes Modell | 6215,933 ^a | 5 | 1243,187 | 7,897 | ,000 | ,422 |
| Konstanter Term | 171949,067 | 1 | 171949,067 | 1092,254 | ,000 | ,953 |
| lehrmeth | 1639,633 | 2 | 819,817 | 5,208 | ,009 | ,162 |
| niveau | 2613,600 | 1 | 2613,600 | 16,602 | ,000 | ,235 |
| lehrmeth * niveau | 1962,700 | 2 | 981,350 | 6,234 | ,004 | ,188 |
| Fehler | 8501,000 | 54 | 157,426 | | | |
| Gesamt | 186666,000 | 60 | | | | |
| Korrigierte Gesamtvariation | 14716,933 | 59 | | | | |

a. R-Quadrat = ,422 (korrigiertes R-Quadrat = ,369)

Die übrigen Effektstärken müssen also auf der Basis der nebenstehenden Tabelle per Hand bestimmt werden.

Zweifaktorielle Varianzanalyse

- Wir können Effektstärkemaße nicht nur für die beiden Haupteffekte und die Interaktion angeben, sondern auch für die anderen Effekte, die wir betrachten können. Das partielle Omega für den **Haupteffekt A** sah wie folgt aus (illustriert am **Beispiel**):

$$\hat{\omega}_{p_A}^2 = \frac{(F_A - 1) \cdot df_A}{(F_A - 1) \cdot df_A + n} = \frac{(5.208 - 1) \cdot 2}{(5.208 - 1) \cdot 2 + 60} = 0.12$$

- Analog lassen sich jetzt z.B. auch die Effektstärken für **einfache Haupteffekte** angeben, z.B. die Stärke des Effektes der Lehrmethoden bei schwächeren Schülern (vgl. Keppel & Wickens, 2004, S. 258ff):

$$\hat{\omega}_{p_{A|b_1}}^2 = \frac{(F_{A|b_1} - 1) \cdot df_{A|b_1}}{(F_{A|b_1} - 1) \cdot df_{A|b_1} + m \cdot n_{..}} = \frac{(11.01 - 1) \cdot 2}{(11.01 - 1) \cdot 2 + 3 \cdot 10} = 0.40$$

oder für einen **einfachen Kontrast**, z.B. den Vergleich der beiden Lehrbuchmethoden innerhalb der schwächeren Schüler:

$$\hat{\omega}_{p_L}^2 = \frac{(F_L - 1) \cdot df_L}{(F_L - 1) \cdot df_L + 2 \cdot n_{..}} = \frac{(13.48 - 1) \cdot 1}{(13.48 - 1) \cdot 1 + 2 \cdot 10} = 0.38$$

Der einzige Unterschied besteht im letzten Term im Nenner, der davon abhängt, wie viele Messwerte eingehen.

- 7 Kontraste
- 8 Effektstärken
- 9 **Ungleiche Stichprobenumfänge**
- 10 Designs mit zufälligen Faktoren

Zweifaktorielle Varianzanalyse mit ungleichen Zellenbesetzungen

- Alle obigen Berechnungen gelten nur für den erstrebenswerten Fall gleicher Stichprobengrößen in allen Zellen. In diesem Fall spricht man auch von einem **orthogonalen** oder **balancierten** faktoriellen **Design**.
- Ungleiche Zellbesetzungen können geplant auftreten (z.B. weil eine Gruppe leichter zu rekrutieren ist) oder ungeplant resultieren (z.B. durch Vpn-Ausfälle, Datenprobleme oder durch die nachträgliche Untersuchung von Merkmalen als Faktor, z.B. Geschlecht).
- Unabhängig von der Frage nach der statistischen Auswertung, können sich durch Vpn-Ausfälle erhebliche Probleme für die Validität der Studie ergeben, wenn diese mit dem Treatment (-kombinationen) im Zusammenhang stehen sind (z.B. wenn Patienten mit einer stärker ausgeprägten Störung eine Therapie häufig abbrechen). Im Bezug auf die Validität generell unproblematisch sind Vpn-Ausfälle nur, wenn sie zufällig erfolgt sind.
- Liegen unbalancierte Designs vor, so ergibt sich bei konventioneller Berechnung der Rand- und Quadratsummen das Problem, dass die Haupteffekte und die Wechselwirkung (teilweise) nicht mehr unabhängig voneinander sind (konfundiert sind) und sich wechselseitig beeinflussen.
- Dies wird im Folgenden an konstruierten Datensätzen demonstriert.

Zweifaktorielle Varianzanalyse mit ungleichen Zellenbesetzungen

gleiche Stichprobenumfänge

| Rohdaten | | | | |
|----------|-------|-------|-------|-------|
| | a_1 | a_2 | a_3 | |
| b_1 | 1 2 3 | 1 2 3 | 1 2 3 | 1 2 3 |
| | 1 2 3 | 1 2 3 | 1 2 3 | 1 2 3 |
| | 1 2 3 | 1 2 3 | 1 2 3 | 1 2 3 |
| b_2 | 4 5 6 | 4 5 6 | 4 5 6 | 4 5 6 |
| | 4 5 6 | 4 5 6 | 4 5 6 | 4 5 6 |
| | 4 5 6 | 4 5 6 | 4 5 6 | 4 5 6 |

Mittelwerte

| \bar{x}_{jk} | $\bar{x}_{\bullet k}$ | | | |
|----------------------|-----------------------|-----|-----|-----|
| | 2 | 2 | 2 | 2 |
| | 5 | 5 | 5 | 5 |
| $\bar{x}_{j\bullet}$ | 3.5 | 3.5 | 3.5 | 3.5 |

Stichprobenumfänge

| n_{jk} | $n_{\bullet k}$ | | | |
|----------------|-----------------|----|----|-----|
| | 18 | 18 | 18 | 54 |
| | 18 | 18 | 18 | 54 |
| $n_{j\bullet}$ | 36 | 36 | 36 | 108 |

ungleiche, disproportionale Stichprobenumfänge

| | a_1 | a_2 | a_3 |
|-------|-------|-------|-------|
| b_1 | 1 2 3 | 1 2 3 | 1 2 3 |
| | 1 2 3 | | 1 2 3 |
| b_2 | 4 5 6 | 4 5 6 | 4 5 6 |
| | 4 5 6 | | |
| | 4 5 6 | | |

| \bar{x}_{jk} | $\bar{x}_{\bullet k}$ | | | |
|----------------------|-----------------------|-----|---|------|
| | 2 | 2 | 2 | 2 |
| | 5 | 5 | 5 | 5 |
| $\bar{x}_{j\bullet}$ | 4 | 3.5 | 3 | 3.58 |

| n_{jk} | $n_{\bullet k}$ | | | |
|----------------|-----------------|----|----|----|
| | 9 | 6 | 12 | 27 |
| | 18 | 6 | 6 | 30 |
| $n_{j\bullet}$ | 27 | 12 | 18 | 57 |

Beispiel für eine Beeinflussung des Haupteffektes A durch den Haupteffekt B bei ungleichem n .

$$\bar{x}_{j\bullet} = \frac{\sum_{k=1}^q \sum_{i=1}^{n_{jk}} x_{ijk}}{\sum_{k=1}^q n_{jk}} = \frac{\sum_{k=1}^q \sum_{i=1}^{n_{jk}} x_{ijk}}{n_{j\bullet}}, \quad \bar{x}_{\bullet k} = \frac{\sum_{j=1}^m \sum_{i=1}^{n_{jk}} x_{ijk}}{\sum_{j=1}^m n_{jk}} = \frac{\sum_{j=1}^m \sum_{i=1}^{n_{jk}} x_{ijk}}{n_{\bullet k}}$$

Zweifaktorielle Varianzanalyse mit ungleichen Zellenbesetzungen

Rohdaten

gleiche
Stichprobenumfänge

| | a_1 | a_2 | a_3 |
|-------|-------------|-------------|-------------|
| b_1 | 1 2 3 1 2 3 | 3 4 5 3 4 5 | 2 3 4 2 3 4 |
| b_2 | 3 4 5 3 4 5 | 1 2 3 1 2 3 | 2 3 4 2 3 4 |

Mittelwerte

| \bar{x}_{jk} | | | | $\bar{x}_{\bullet k}$ |
|----------------------|---|---|---|-----------------------|
| | 2 | 4 | 3 | 3 |
| $\bar{x}_{j\bullet}$ | 4 | 2 | 3 | 3 |
| | 3 | 3 | 3 | 3 |

Stichprobenumfänge

| n_{jk} | | | | $n_{\bullet k}$ |
|----------------|----|----|----|-----------------|
| | 18 | 18 | 18 | 54 |
| $n_{j\bullet}$ | 18 | 18 | 18 | 54 |
| | 36 | 36 | 36 | 108 |

ungleiche,
disproportionale
Stichprobenumfänge

| | a_1 | a_2 | a_3 |
|-------|-------------|-------------|-------------|
| b_1 | 1 2 3 1 2 3 | 3 4 5 3 4 5 | 2 3 4 2 3 4 |
| b_2 | 3 4 5 3 4 5 | 1 2 3 1 2 3 | 2 3 4 |

| \bar{x}_{jk} | | | | $\bar{x}_{\bullet k}$ |
|----------------------|-----|-----|---|-----------------------|
| | 2 | 4 | 3 | 3.2 |
| $\bar{x}_{j\bullet}$ | 4 | 2 | 3 | 3 |
| | 3.5 | 2.8 | 3 | 3.09 |

| n_{jk} | | | | $n_{\bullet k}$ |
|----------------|----|----|----|-----------------|
| | 6 | 12 | 12 | 30 |
| $n_{j\bullet}$ | 18 | 18 | 3 | 39 |
| | 24 | 30 | 15 | 69 |

Beispiel für eine Beeinflussung beider Haupteffekte durch eine Interaktion bei ungleichem n .

Zweifaktorielle Varianzanalyse mit ungleichen Zellenbesetzungen

- Bei ungleichen, aber **proportionalen Stichprobenumfängen** müssen in allen Zeilen und allen Spalten die Stichprobenumfänge im gleichen Verhältnis stehen (bzw. formal gelten: $n_{jk} = n_{j\bullet} \cdot n_{\bullet k} / n$, d.h. es besteht kein Zusammenhang zwischen den Faktoren in Bezug auf die Zellbesetzungen). Im untenstehenden Beispiel ist der Stichprobenumfang unter der Faktorstufe b_2 immer doppelt so groß wie unter b_1 .
- In diesem Fall kann ein Haupteffekt nicht mehr durch den anderen Haupteffekt, sondern nur noch durch die Interaktion beeinflusst werden.

| | | Rohdaten | | | | | | Mittelwerte | | | | Stichprobenumfänge | | | | | | | |
|---|-------|----------|-------|---|-------|---|-------|-------------|----------------|---|---|--------------------|---|-----------------------|---|----|----|-----------------|---|
| ungleiche, proportionale Stichprobenumfänge | | | a_1 | | a_2 | | a_3 | | | | | | | | | | | | |
| | b_1 | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 |
| | | 2 | 1 | 2 | 3 | | | | | | | | | | | | | | |
| | b_2 | 4 | 5 | 6 | 4 | 5 | 6 | 4 | 5 | 6 | 4 | 5 | 6 | 4 | 5 | 6 | 4 | 5 | 6 |
| 4 | | 5 | 6 | 4 | 5 | 6 | 4 | 5 | 6 | 4 | 5 | 6 | 4 | 5 | 6 | 4 | 5 | 6 | |
| 4 | | 5 | 6 | 4 | 5 | 6 | 4 | 5 | 6 | 4 | 5 | 6 | 4 | 5 | 6 | 4 | 5 | 6 | |
| | | | | | | | | | \bar{x}_{jk} | | | | | $\bar{x}_{\bullet k}$ | | | | | |
| | | | | | | | | | 2 | 2 | 2 | 2 | 2 | | | | | | |
| | | | | | | | | | 5 | 5 | 5 | 5 | 5 | | | | | | |
| | | | | | | | | | 4 | 4 | 4 | 4 | 4 | | | | | | |
| | | | | | | | | | | | | | | n_{jk} | | | | | |
| | | | | | | | | | | | | | | 9 | 3 | 6 | 18 | | |
| | | | | | | | | | | | | | | 18 | 6 | 12 | 36 | | |
| | | | | | | | | | | | | | | 27 | 9 | 18 | 54 | | |
| | | | | | | | | | | | | | | | | | | $n_{\bullet k}$ | |

Zweifaktorielle Varianzanalyse mit ungleichen Zellenbesetzungen

- Wenn wir die Randmittel (und in der Folge dann die Quadratsummen) wie bisher über alle Messwerte der Personen in der betreffenden Zeile bzw. Spalte bestimmen, so kann ein Haupteffekt A oder eine Interaktion AxB auf den (anderen) Haupteffekt durchschlagen.

- Eine Alternative besteht darin, die Randmittel über die durch die Faktorstufen definierten **Gruppen (-mittelwerte)** zu bestimmen. Letztere Mittelwerte bezeichnet man auch als **adjustierte** Mittelwerte (estimated means, adjusted means) oder **ungewichtete** (seltener „gleich gewichtete“) Mittelwerte, z.B. für die Spalten:

$$\tilde{x}_{j\bullet} = \frac{\sum_{k=1}^q \bar{x}_{jk}}{q} \quad \text{z.B. } \tilde{x}_{1\bullet} = \frac{\bar{x}_{11} + \bar{x}_{12}}{2} = \frac{2+4}{2} = 3$$

| | | | | | |
|------------------------|-----|-----|---|-----------------------|-------------------------|
| \bar{x}_{jk} | | | | $\bar{x}_{\bullet k}$ | $\tilde{x}_{\bullet k}$ |
| | 2 | 4 | 3 | 3 | 3 |
| | 4 | 2 | 3 | 3 | 3 |
| $\bar{x}_{j\bullet}$ | 3.5 | 2.8 | 3 | | |
| $\tilde{x}_{j\bullet}$ | 3 | 3 | 3 | | |

| | | | | |
|----------------|----|----|----|-----------------|
| n_{jk} | | | | $n_{\bullet k}$ |
| | 6 | 12 | 12 | 30 |
| | 18 | 18 | 3 | 39 |
| $n_{j\bullet}$ | 24 | 30 | 15 | 69 |

- Durch die Bestimmung der ungewichteten Mittel gibt es keinen überproportionalen Einfluss der größeren Teilgruppen auf die Randmittel mehr, hier z.B. der Gruppe (1, 2) relativ zur Gruppe (1, 1) bei der Bestimmung des Mittelwertes der Faktorstufe a_1 . Wir haben also im Beispiel den Effekt von Faktor B auf die Randmittel von A kontrolliert.

Zweifaktorielle Varianzanalyse mit ungleichen Zellenbesetzungen

- Entsprechend werden die Mittelwerte über die Messwerte als (um die Zellen- n 's) **gewichtete** Mittelwerte bezeichnet, z.B. für die Spalten:

$$\bar{x}_{j\bullet} = \frac{\sum_{k=1}^q \sum_{i=1}^{n_{jk}} x_{ijk}}{n_{j\bullet}} = \frac{\sum_{k=1}^q n_{jk} \cdot \bar{x}_{jk}}{n_{j\bullet}}$$

Da:

$$\bar{x}_{jk} = \frac{\sum_{i=1}^{n_{jk}} x_{ijk}}{n_{jk}}$$

z.B. $\bar{x}_{1\bullet} = \frac{n_{11} \cdot \bar{x}_{11} + n_{12} \cdot \bar{x}_{12}}{n_{1\bullet}} = \frac{6 \cdot 2 + 18 \cdot 4}{24} = 3.5$

| | | | | | |
|------------------------|-----|-----|---|-----------------------|-------------------------|
| \bar{x}_{jk} | | | | $\bar{x}_{\bullet k}$ | $\tilde{x}_{\bullet k}$ |
| | 2 | 4 | 3 | 3 | 3 |
| | 4 | 2 | 3 | 3 | 3 |
| $\bar{x}_{j\bullet}$ | 3.5 | 2.8 | 3 | | |
| $\tilde{x}_{j\bullet}$ | 3 | 3 | 3 | | |

| | | | | |
|----------------|----|----|----|-----------------|
| n_{jk} | | | | $n_{\bullet k}$ |
| | 6 | 12 | 12 | 30 |
| | 18 | 18 | 3 | 39 |
| $n_{j\bullet}$ | 24 | 30 | 15 | 69 |

Zweifaktorielle Varianzanalyse mit ungleichen Zellenbesetzungen

- Insgesamt gibt es verschiedene Methoden, nicht-orthogonale Designs per Varianzanalyse auszuwerten. Nach einer Unterscheidung von Overall und Spiegel (1969) werden die drei prominentesten Methoden mit 1, 2 und 3 bezeichnet (wobei die Nummerierungen in der Literatur nicht immer einheitlich verwendet werden, siehe auch unten bei SPSS).
- Sie führen nur im Falle gleicher Zellbesetzungen zum gleichen – dem in der Vorlesung ausführlich dargestellten – Ergebnis.
- Bei Methode 1 beziehen sich die Tests auf die ungewichteten Zeilen- und Spaltenmittel. Dies führt dazu, dass jeder Effekt in Bezug auf alle anderen Effekte adjustiert wird (d.h. die anderen Effekte statistisch kontrolliert werden bzw. der Einfluss der anderen Effekte herauspartialisiert wird). Diese Methode ...
 - ist angemessen, wenn die unterschiedlichen Zellbesetzungen irrelevant für die Schlüsse der Studie sind, was ja fast immer der Fall ist.
 - prüft die gleichen Nullhypothesen, wie im orthogonalen Fall (d.h. die Hypothesen sind unabhängig von den Zellbesetzungen).
 - führt dazu, dass sich die Quadratsummen nicht mehr additiv verhalten.
 - ist die Standardeinstellung in SPSS (dort als `TYPE III sum of squares` bezeichnet).

Zweifaktorielle Varianzanalyse mit ungleichen Zellenbesetzungen

- In den 70er Jahren gab es kontroverse Diskussionen darüber, welche der drei (oder noch anderen) Methoden vorzuziehen ist. Heute wird von fast allen Autoren Methode 1 empfohlen.
- **Methode 3** (hierarchical oder sequential approach, SPSS: TYPE I) bezieht sich auf die gewichteten Zeilen- und Spaltenmittelwerte. Dabei muss eine Reihenfolge der Faktoren festgelegt werden und jeder Effekt wird dann (nur) in Bezug auf die vorangegangenen adjustiert. Je nach Reihenfolge unterscheiden sich die Ergebnisse. Diese Methode ...
 - kommt nur in Frage, wenn die Zellbesetzungen inhaltlich bedeutsam sind (d.h. proportional zu denen in der Population sind).
 - testet stichprobengewichtete Nullhypothesen.
 - führt zu additiven Quadratsummen.
- (Bei **Methode 2** (SPSS: TYPE II) sind die Haupteffekte wechselseitig adjustiert und der Interaktionseffekt in Bezug auf die Haupteffekte (nicht aber umgekehrt). Sie testet eine sehr komplizierte stichprobengewichtete Nullhypothese (Howell & McCaughy, 1982). Wenn in der Population keine Interaktion besteht, weist diese Methode eine höhere Power für die Haupteffekte auf; andernfalls überschätzt sie die Haupteffekte. Spielt heute praktisch kaum eine Rolle.)

Zweifaktorielle Varianzanalyse mit ungleichen Zellenbesetzungen

- Die Durchführung der drei Methoden kann allgemein nicht mehr in einfacher Weise dargestellt werden. Die Berechnungen basieren heute alle auf einem verallgemeinerten Regressionsansatz mit mehr als einem Prädiktor, den man auch als **Allgemeines Lineares Modell** (ALM) bezeichnet. Das Vorgehen wird hier nicht näher dargestellt (vgl. Keppel & Wickens, 2004 und Zahn, 2010, in Bezug auf unser folgendes Beispiel) und erst wieder im Rahmen der Vorlesung **Multivariate Verfahren** aufgegriffen.
- Die Ergebnisse von **Robustheitsuntersuchungen** bei Verstößen gegen die Annahmen der Varianzanalyse wurden im einfaktoriellen Fall ausführlicher dargestellt. Sie gelten prinzipiell auch für den zweifaktoriellen Fall. Bei ungleichen Zellbesetzungen reagiert die Varianzanalyse noch empfindlicher auf Varianzheterogenität.
- Auch im zweifaktoriellen Fall sind bei Varianzheterogenität Korrekturen nach Welch oder Brown-Forsyth möglich, aber in SPSS nicht implementiert.
- Im Falle ungleicher Zellbesetzungen müssen auch entsprechend die Kontraste adjustiert werden (vgl. Diehl & Arbinger, 2001, Kap. 13.7).

Zweifaktorielle Varianzanalyse mit ungleichen Zellenbesetzungen

- **Beispiel** (fiktiv aus Maxwell & Delaney, 2004; vgl. auch Zahn, 2010): In einem Unternehmen soll untersucht werden, wie sich die Ausbildung (Faktor A ; mit und ohne College-Abschluss) und das Geschlecht der Beschäftigten (Faktor B) auf das Einstiegsgehalt (gerundet in Tausend \$ pro Jahr) auswirkt. In dem Unternehmen wird eine Zufallsstichprobe der im letzten Jahr eingestellten MitarbeiterInnen gezogen. Es resultieren die nebenstehenden Daten.

- Folgende Mittelwerte resultieren:

| | ohne Ab. | mit Ab. | $\bar{x}_{\bullet k}$ | $\tilde{x}_{\bullet k}$ |
|--------------------------|----------|---------|-----------------------|-------------------------|
| Frauen | 17.00 | 25.00 | 22.33 | 21.00 |
| Männer | 20.00 | 27.00 | 22.10 | 23.50 |
| $\bar{x}_{j\bullet} :$ | 18.91 | 25.55 | | |
| $\tilde{x}_{j\bullet} :$ | 18.50 | 26.00 | | |

| | College abgeschlossen? | |
|--------|------------------------|----|
| | nein | ja |
| Frauen | 15 | 24 |
| | 17 | 26 |
| | 20 | 25 |
| | 16 | 24 |
| | | 27 |
| | | 24 |
| | | 27 |
| | | 23 |
| Männer | 19 | 25 |
| | 18 | 29 |
| | 21 | 27 |
| | 20 | |
| | 21 | |
| | 22 | |
| | 19 | |

Zweifaktorielle Varianzanalyse mit ungleichen Zellenbesetzungen

- In diesem Beispiel kommt man bzgl. des Effektes des Geschlechts deskriptiv zu unterschiedlichen Schlüssen, wenn man die gewichteten und die ungewichteten Randmittel heranzieht.
- Auf der Basis der gewichteten Mittel kommt man zu dem Schluss, dass die Frauen mehr verdienen ($22.33 > 22.10$); bei den ungewichteten Mitteln ist es anders herum ($21.00 < 23.50$).
- Der höhere Verdienst in den gewichteten Randmitteln kommt offensichtlich dadurch zustande, dass bei den Frauen ein größerer Anteil einen College-Abschluss hat als bei den Männern und bei einem College-Abschluss (wie man an den Spaltenrandmitteln sieht) besser bezahlt wird.

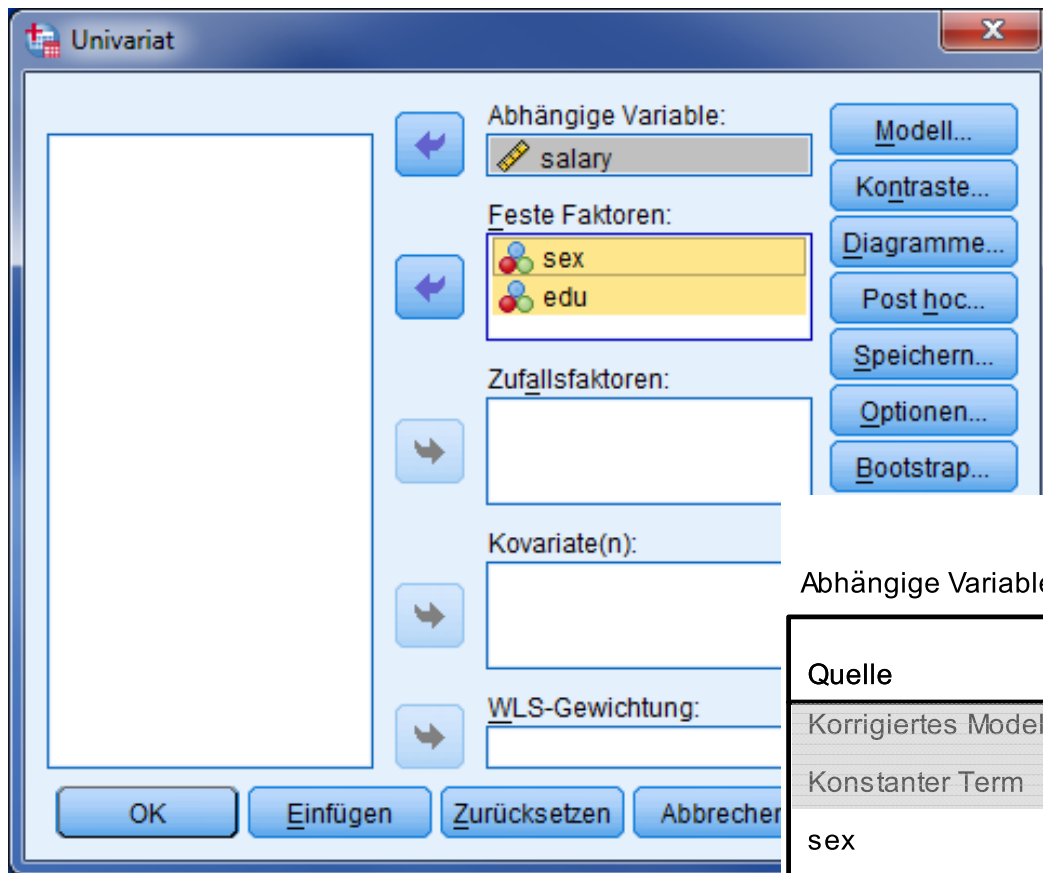
| | ohne Ab. | mit Ab. | $\bar{x}_{\bullet k}$ | $\tilde{x}_{\bullet k}$ |
|--------------------------|----------|---------|-----------------------|-------------------------|
| Frauen | 17.00 | 25.00 | 22.33 | 21.00 |
| Männer | 20.00 | 27.00 | 22.10 | 23.50 |
| $\bar{x}_{j\bullet} :$ | 18.91 | 25.55 | | |
| $\tilde{x}_{j\bullet} :$ | 18.50 | 26.00 | | |

| | ohne Ab. | mit Ab. | $n_{\bullet k}$ |
|------------------|----------|---------|-----------------|
| Frauen | 4 | 8 | 12 |
| Männer | 7 | 3 | 10 |
| $n_{j\bullet} :$ | 11 | 11 | 22 |

Zweifaktorielle Varianzanalyse mit ungleichen Zellenbesetzungen

- Vermutlich werden wir uns in dem Beispiel dafür interessieren, ob das Geschlecht einen Effekt hat, wenn für den Einfluss der Ausbildung kontrolliert worden ist (bzw. umgekehrt). D.h. wir wollen nicht, dass der Geschlechtseffekt durch die Varianz kontaminiert wird, die durch den Zusammenhang zwischen dem Geschlecht und der Ausbildung zustande kommt. Oder nochmal anders ausgedrückt: Wir wollen wissen, ob sich Lohnunterschiede zwischen Männern und Frauen bei gleicher Ausbildung ergeben.
- Dabei interessiert dann nicht, dass Frauen im Durchschnitt mehr verdienen, denn dieser Effekt verschwindet (kehrt sich um), wenn wir für die Ausbildung kontrollieren.
- In diesem Fall würden wir entsprechend die Varianzanalyse-Methode 1 (SPSS: Type III Quadratsummen) durchführen.

Zweifaktorielle Varianzanalyse mit ungleichen Zellenbesetzungen in SPSS



Die Quadratsummen des Typs III sind bei SPSS voreingestellt, daher sind keine weiteren Veränderungen erforderlich.

Tests der Zwischensubjekteffekte

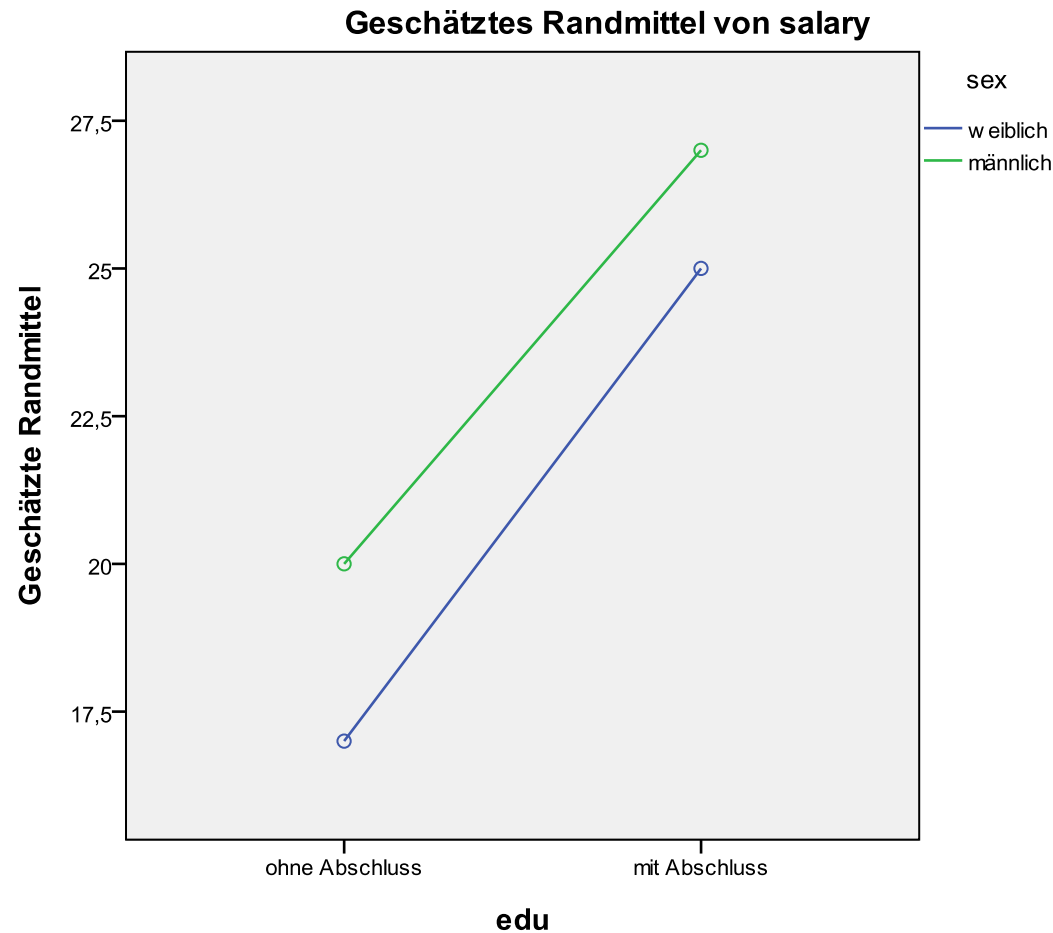
Abhängige Variable:salary

| Quelle | Quadratsumme vom Typ III | df | Mittel der Quadrate | F | Sig. |
|-----------------------------|--------------------------|----|---------------------|----------|------|
| Korrigiertes Modell | 273,864 ^a | 3 | 91,288 | 32,864 | ,000 |
| Konstanter Term | 9305,790 | 1 | 9305,790 | 3350,084 | ,000 |
| sex | 29,371 | 1 | 29,371 | 10,573 | ,004 |
| edu | 264,336 | 1 | 264,336 | 95,161 | ,000 |
| sex * edu | 1,175 | 1 | 1,175 | ,423 | ,524 |
| Fehler | 50,000 | 18 | 2,778 | | |
| Gesamt | 11193,000 | 22 | | | |
| Korrigierte Gesamtvariation | 323,864 | 21 | | | |

Es zeigt sich, dass beide Haupteffekte, nicht aber die Wechselwirkung statistisch signifikant ausfallen.

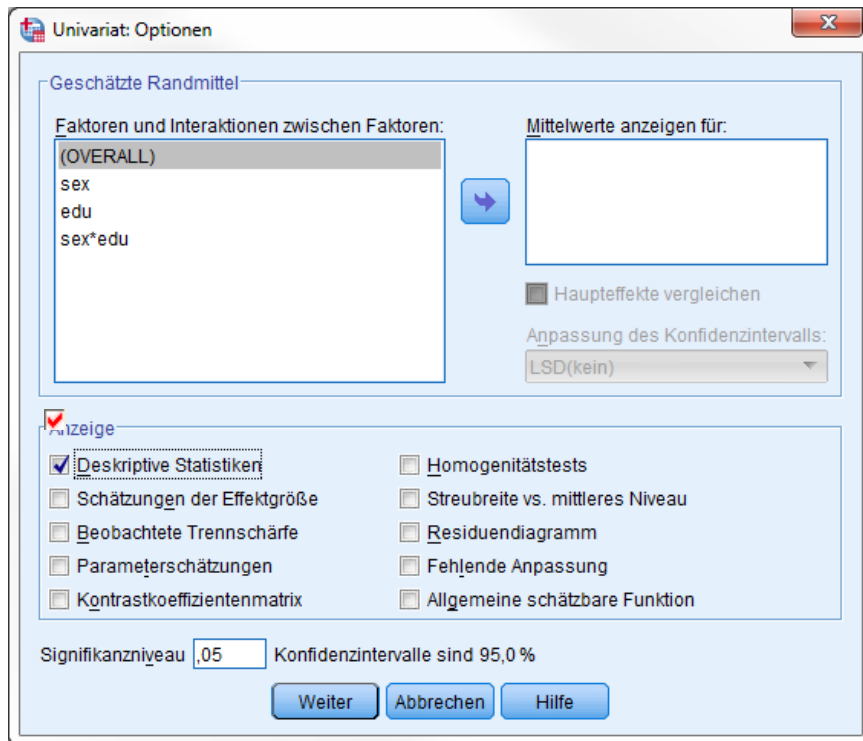
a. R-Quadrat = ,846 (korrigiertes R-Quadrat = ,820)

Zweifaktorielle Varianzanalyse mit ungleichen Zellenbesetzungen in SPSS



Die unter (Diagramme. . .) erzeugten Mittelwertsdiagramme stellen nur die Zellenmittelwerte dar und sind insofern von der Methode der Bestimmung der Randmittel unbeeinflusst und korrekt. (Die Beschriftung mit „Geschätztes Randmittel“, die SPSS immer wählt, ist im zweifaktoriellen Falle unsinnig.)

Zweifaktorielle Varianzanalyse mit ungleichen Zellenbesetzungen in SPSS



Fordern wir wie üblich unter (Optionen) deskriptive Statistiken an, so erhalten wir immer die gewichteten Randmittel, die in diesem Fall (Type III der Quadratsummen) nicht zur Interpretation herangezogen werden dürfen.

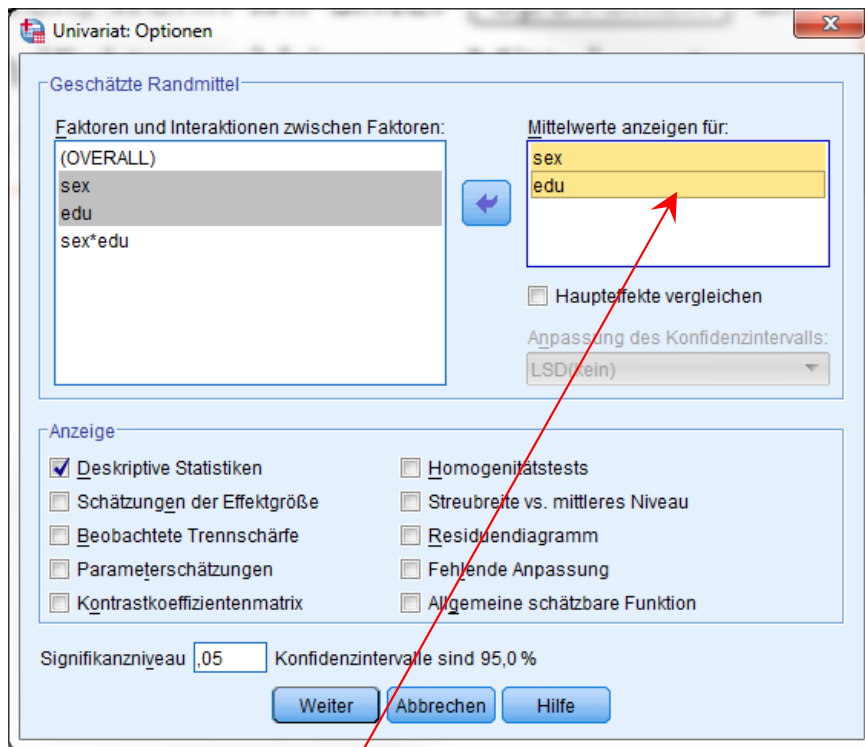
| | ohne Ab. | mit Ab. | $\bar{x}_{\cdot k}$ | $\tilde{x}_{\cdot k}$ |
|------------------------|----------|---------|---------------------|-----------------------|
| Frauen | 17.00 | 25.00 | 22.33 | 21.00 |
| Männer | 20.00 | 27.00 | 22.10 | 23.50 |
| $\bar{x}_{j\cdot}$: | 18.91 | 25.55 | 22.23 | |
| $\tilde{x}_{j\cdot}$: | 18.50 | 26.00 | | |

Deskriptive Statistiken

Abhängige Variable:salary

| sex | edu | Mittelwert | Standardabweichung | N |
|------------|------------------|------------|--------------------|----|
| 1 weiblich | 0 ohne Abschluss | 17,00 | 2,160 | 4 |
| | 1 mit Abschluss | 25,00 | 1,512 | 8 |
| | Gesamt | 22,33 | 4,271 | 12 |
| 2 männlich | 0 ohne Abschluss | 20,00 | 1,414 | 7 |
| | 1 mit Abschluss | 27,00 | 2,000 | 3 |
| | Gesamt | 22,10 | 3,695 | 10 |
| Gesamt | 0 ohne Abschluss | 18,91 | 2,212 | 11 |
| | 1 mit Abschluss | 25,55 | 1,809 | 11 |
| | Gesamt | 22,23 | 3,927 | 22 |

Zweifaktorielle Varianzanalyse mit ungleichen Zellenbesetzungen in SPSS



Die ungewichteten Randmittel werden ausgegeben, wenn wir unter (Optionen) die Variablen (Faktoren) hier zu „Mittelwerte anzeigen für“ hinüberziehen.

| | ohne Ab. | mit Ab. | $\bar{x}_{\cdot k}$ | $\tilde{x}_{\cdot k}$ |
|------------------------|----------|---------|---------------------|-----------------------|
| Frauen | 17.00 | 25.00 | 22.33 | 21.00 |
| Männer | 20.00 | 27.00 | 22.10 | 23.50 |
| $\bar{x}_{j\cdot} :$ | 18.91 | 25.55 | 22.23 | |
| $\tilde{x}_{j\cdot} :$ | 18.50 | 26.00 | | |

Geschätzte Randmittel

1. sex

Abhängige Variable:salary

| sex | Mittelwert | Standardfehler | 95%-Konfidenzintervall | |
|------------|------------|----------------|------------------------|------------|
| | | | Untergrenze | Obergrenze |
| 1 weiblich | 21,000 | ,510 | 19,928 | 22,072 |
| 2 männlich | 23,500 | ,575 | 22,292 | 24,708 |

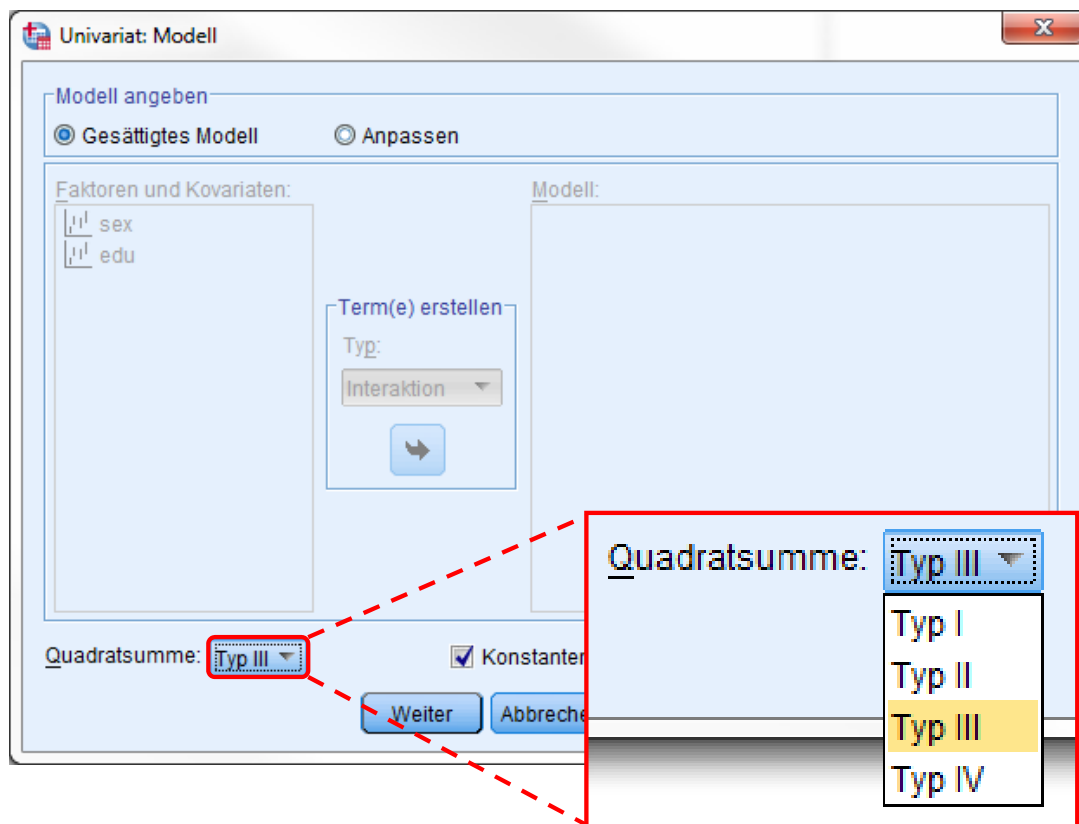
2. edu

Abhängige Variable:salary

| edu | Mittelwert | Standardfehler | 95%-Konfidenzintervall | |
|------------------|------------|----------------|------------------------|------------|
| | | | Untergrenze | Obergrenze |
| 0 ohne Abschluss | 18,500 | ,522 | 17,403 | 19,597 |
| 1 mit Abschluss | 26,000 | ,564 | 24,815 | 27,185 |

Zweifaktorielle Varianzanalyse mit ungleichen Zellbesetzungen in SPSS

- In dieser Studie haben die unterschiedlichen Fallzahlen in den Zellen aufgrund der Zufallsziehung eine inhaltliche Bedeutung. Sie spiegeln den Zusammenhang zwischen dem Geschlecht und der Ausbildung in dem Unternehmen wider. Es zeigt sich, dass mehr Frauen mit Abschluss eingestellt werden, als das bei den Männern der Fall ist.



- Man könnte sich also auch fragen, wer in dem Unternehmen mehr verdient, Männer oder Frauen? D.h. man würde bei dieser Frage den Zusammenhang zwischen dem Geschlecht und der Ausbildung ignorieren (d.h. die Tatsache, dass der höhere durchschnittliche Verdienst der Frauen an ihrer besseren Ausbildung liegt).

Für die Analyse der gewichteten Mittelwerte muss unter (Model I) abweichend die Quadratsumme auf Typ I geändert werden.

Zweifaktorielle Varianzanalyse mit ungleichen Zellbesetzungen in SPSS

Tests der Zwischensubjekteffekte

Abhängige Variable: salary

| Quelle | Quadratsumme vom Typ I | df | Mittel der Quadrate | F | Sig. |
|-----------------------------|------------------------|----|---------------------|----------|------|
| Korrigiertes Modell | 273,864 ^a | 3 | 91,288 | 32,864 | ,000 |
| Konstanter Term | 10869,136 | 1 | 10869,136 | 3912,889 | ,000 |
| sex | ,297 | 1 | ,297 | ,107 | ,747 |
| edu | 272,392 | 1 | 272,392 | 98,061 | ,000 |
| sex * edu | 1,175 | 1 | 1,175 | ,423 | ,524 |
| Fehler | 50,000 | 18 | 2,778 | | |
| Gesamt | 11193,000 | 22 | | | |
| Korrigierte Gesamtvariation | 323,864 | 21 | | | |

a. R-Quadrat = ,846 (korrigiertes R-Quadrat = ,820)

Da nur der erste Faktor beim Quadratsummen-Typ I nicht adjustiert ist, müssen wir dafür sorgen, dass er als erste Variable im Dialog unter „Fester Faktor“ angewählt wird (oder in der Syntax die Reihenfolge unter /DESIGN= entsprechend geändert wird).

Hier zeigt sich, dass das geringfügig höhere Gehalt der Frauen nicht statistisch signifikant ausfällt. Zur Deutung sind jetzt die gewichteten Mittelwerte heranzuziehen.

```
UNIANOVA salary BY sex edu
  /METHOD=SSTYPE(1)
  /INTERCEPT=INCLUDE
  /PLOT=PROFILE(edu*sex)
  /PRINT=DESCRIPTIVE
  /CRITERIA=ALPHA(.05)
  /DESIGN=sex edu sex*edu.
```

- 7 Kontraste
- 8 Effektstärken
- 9 Ungleiche Stichprobenumfänge
- 10 Designs mit zufälligen Faktoren

Zweifaktorielle Varianzanalyse: Modelle mit Zufallseffekten

- **Beispiel** (fiktiv aus Howell, 2010): Es soll untersucht werden, ob es Personen schneller gelingt, große Buchstaben zu identifizieren als kleine Buchstaben (Faktor *A*: Großschreibung). Wir untersuchen dies anhand von fünf zufällig ausgewählten Buchstaben (Faktor *B*: Buchstabe). Je 10 zufällig zugewiesene Vpn wurden in jeder Zelle des faktoriellen Plans untersucht. Abhängige Variable ist die Reaktionszeit, die die Vp benötigt, um einen am Computerbildschirm dargebotenen Buchstaben korrekt zu identifizieren (in Zehntelsekunden).

- In diesem Beispiel ist der Faktor *A* (Großschreibung) wie in den vorangegangenen Beispielen ein fester Faktor; der Faktor *B* (Buchstabe) ist hingegen ein **zufälliger** Faktor (vgl. ausführlicher Gliederungspunkt 6 bei der einfaktoriellen Varianzanalyse).

| | | Faktor A: Großschreibung | |
|---------------------|---|---------------------------------|----------------------------------|
| | | Großbuchstabe | Kleinbuchstabe |
| Faktor B: Buchstabe | A | 9 6 10 6 7 8 8 4 5 7 | 8 4 7 5 9 6 6 6 7 7 |
| | G | 7 6 6 6 8 9 6 11 3 7 | 10 8 4 10 7 7 10 7 6 7 |
| | D | 11 8 14 13 10 13 6 11 13 11 | 14 18 13 17 12 11 14 22 16 11 |
| | K | 12 16 9 12 19 11 11 23 10 11 | 20 16 18 20 14 16 15 16 22 19 |
| | W | 10 14 10 14 11 19 5 11 15 11 | 21 17 22 22 18 19 15 16 22 21 |

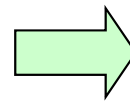
Zweifaktorielle Varianzanalyse: Modelle mit Zufallseffekten

- Prinzipiell können in einem zweifaktoriellen Design einer oder beide Faktoren zufällig sein. Wir betrachten hier nur den häufigeren Fall, dass ein Faktor fest und einer zufällig ist, was auch als **gemischtes Modell** (=mixed model) bezeichnet wird (vgl. allgemeiner Keppel & Wickens, 2004, Kap. 24).
- Während im einfaktoriellen Fall die Auswertung eines zufälligen Faktors keine (gegenüber dem festen Faktor) modifizierte Auswertung erforderte, ist dies im zwei- (oder höher-) faktoriellen Fall anders.
- Dies hängt mit folgendem Sachverhalt zusammen: Nehmen wir an, (wie im Beispiel) dass B der zufällige Faktor ist. Dann hängen die Randmittel von A (und damit der Haupteffekt von A) auch von der spezifischen Zufallsauswahl von B ab. Würde man z.B. eine Stufe in B eliminieren (oder durch eine andere austauschen), so ändern sich im Allgemeinen auch die Randmittel von A (Maxwell & Delaney, 2004, Kap. 10).

Zweifaktorielle Varianzanalyse: Modelle mit Zufallseffekten

- Um dies zu demonstrieren, nehmen wir an, in der Population bestünden die unten links dargestellten Mittelwerte μ für alle möglichen Kombinationen der Faktorstufen des festen und zufälligen Faktors. Es läge eine Interaktion vor, aber keine Haupteffekte für A und B .
- Wählen wir nun eine Zufallsauswahl aus den Faktorstufen von B , so könnte das Ergebnis wie rechts resultieren. Man sieht, dass die Zufallsauswahl des Faktors B einen Einfluss auf die Randmittel (den Haupteffekt) des Faktors A aufweist. Der Haupteffekt für A ist mit der Interaktion konfundiert.

| | Groß | Klein | $\mu_{\bullet k}$ |
|------------------|------|-------|-------------------|
| A | 7.7 | 12.3 | 10 |
| B | 10.0 | 10.0 | 10 |
| C | 11.1 | 8.9 | 10 |
| D | 7.0 | 13.0 | 10 |
| ... | | | |
| Y | 12.3 | 7.7 | 10 |
| Z | 11.3 | 8.7 | 10 |
| $\mu_{j\bullet}$ | 10 | 10 | |



Zufallsauswahl
der Faktorstufen
von Faktor B

| | Groß | Klein | $\mu_{\bullet k}$ |
|------------------|------|-------|-------------------|
| A | 7.7 | 12.3 | 10 |
| D | 7.0 | 13.0 | 10 |
| Y | 12.3 | 7.7 | 10 |
| $\mu_{j\bullet}$ | 9 | 11 | |

Zweifaktorielle Varianzanalyse: Modelle mit Zufallseffekten

- Das **Modell** lautet wie im zweifaktoriellen Fall

$$x_{ijk} = \mu + \alpha_j + \beta_k + \gamma_{jk} + e_{ijk}$$

wobei davon ausgegangen wird, dass nicht nur die x_{ijk} bzw. e_{ijk} normalverteilt sind, formal geschrieben $e_{ijk} \sim N(0, \sigma_e^2)$, sondern auch die $\beta_k \sim N(0, \sigma_\beta^2)$ und alle Zufallseffekte unabhängig voneinander sind.

- Beim zufälligen Faktor B sind wir nicht mehr an den Unterschieden zwischen bestimmten Faktorstufen k interessiert, sondern an der Varianz σ_β^2 der Effekte β_k in der Population. Entsprechend lautet die Nullhypothese nicht wie beim Faktor A $\mu_1 = \mu_2 = \dots = \mu_m$, sondern für den Haupteffekt B : $H_0 : \sigma_\beta^2 = 0$ $H_1 : \sigma_\beta^2 \neq 0$.
- Dies gilt auch für die Interaktion: $H_0 : \sigma_\gamma^2 = 0$ $H_1 : \sigma_\gamma^2 \neq 0$.

Zweifaktorielle Varianzanalyse: Modelle mit Zufallseffekten

- Vergleicht man die bekannten Erwartungswerte in der Tabelle links mit denen im gemischten Modell rechts, so sieht man auch hier, dass in den mittleren Quadraten des festen Faktors dessen Effekt und der der Interaktion konfundiert sind. Eine Lösung besteht darin, bei der Testung des Faktors A in der F -Prüfgröße MQ_A an $MQ_{A \times B}$ zu relativieren.

| | Beide Faktoren fest | Mixed Model: Faktor A fest; Faktor B zufällig |
|------------------------|---|--|
| $E(MQ_A) =$ | $\sigma_e^2 + q \cdot n_{..} \cdot \sum_{j=1}^m \frac{\alpha_j^2}{m-1} = \sigma_e^2 + q \cdot n_{..} \cdot \theta_\alpha^2$ | $\sigma_e^2 + q \cdot n_{..} \cdot \theta_\alpha^2 + n_{..} \cdot \sigma_\gamma^2$ |
| $E(MQ_B) =$ | $\sigma_e^2 + m \cdot n_{..} \cdot \sum_{k=1}^q \frac{\beta_k^2}{q-1} = \sigma_e^2 + m \cdot n_{..} \cdot \theta_\beta^2$ | $\sigma_e^2 + m \cdot n_{..} \cdot \sigma_\beta^2$ |
| $E(MQ_{A \times B}) =$ | $\sigma_e^2 + n_{..} \cdot \sum_{j=1}^m \sum_{k=1}^q \frac{\gamma_{jk}^2}{(m-1) \cdot (q-1)} = \sigma_e^2 + n_{..} \cdot \theta_\gamma^2$ | $\sigma_e^2 + n_{..} \cdot \sigma_\gamma^2$ |
| $E(MQ_{in}) =$ | σ_e^2 | σ_e^2 |

Zweifaktorielle Varianzanalyse: Modelle mit Zufallseffekten

- Die Effekte werden also im gemischten Modell wie folgt getestet:

$$F_A = \frac{MQ_A}{MQ_{AxB}} \quad (\text{fester Faktor}) \quad \text{mit } df_1 = m - 1; df_2 = (m - 1) \cdot (q - 1)$$

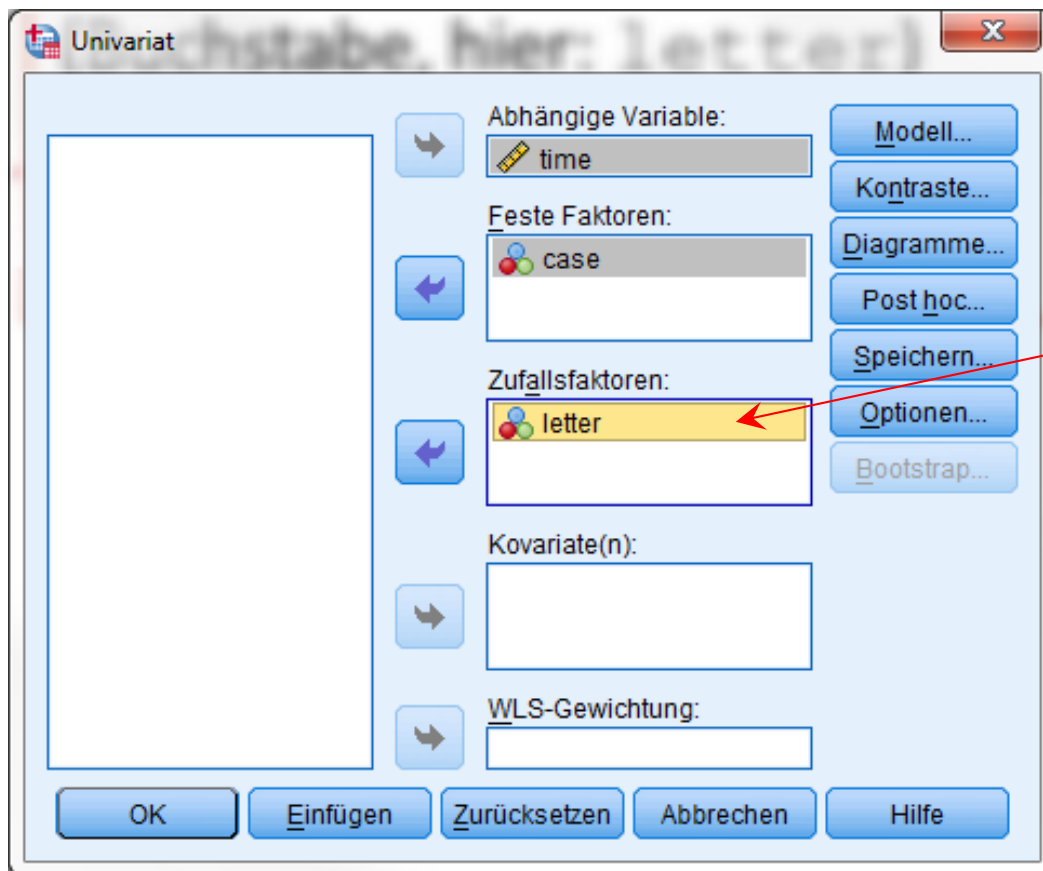
$$F_B = \frac{MQ_B}{MQ_{in}} \quad (\text{zufälliger Faktor}) \quad \text{mit } df_1 = q - 1; df_2 = n - m \cdot q$$

$$F_{AxB} = \frac{MQ_{AxB}}{MQ_{in}} \quad \text{mit } df_1 = (m - 1) \cdot (q - 1); df_2 = n - m \cdot q$$

- Im **Beispiel** resultiert:

| Variation | Quadratsummen | Freiheitsgrade | Mittlere Quadrate | F | p |
|-------------------|---------------|----------------|-------------------|--------|-------|
| A: Großschreibung | 240.25 | 1 | 240.250 | 5.050 | < .05 |
| B: Buchstabe | 1514.94 | 4 | 378.735 | 47.191 | < .05 |
| AxB | 190.30 | 4 | 47.575 | 5.928 | < .05 |
| Innerhalb | 722.30 | 90 | 8.026 | | |
| Total | 2667.79 | | | | |

Zweifaktorielle Varianzanalyse: Modelle mit Zufallseffekten in SPSS



In diesem Fall ist der Faktor *B* (Buchstabe, hier: letter) unter „Zufallsfaktoren“ einzutragen.

Zweifaktorielle Varianzanalyse: Modelle mit Zufallseffekten in SPSS

Deskriptive Statistiken

Abhängige Variable:time

| case | letter | Mittelwert | Standardabweichung | N |
|--------------|--------|------------|--------------------|-------|
| 1 Großbuch. | 1 A | 7,00 | 1,826 | 10 |
| | 2 G | 6,90 | 2,132 | 10 |
| | 3 D | 11,00 | 2,494 | 10 |
| | 4 K | 13,40 | 4,502 | 10 |
| | 5 W | 12,00 | 3,742 | 10 |
| | Gesamt | | 10,06 | 4,007 |
| 2 Kleinbuch. | 1 A | 6,50 | 1,434 | 10 |
| | 2 G | 7,60 | 1,955 | 10 |
| | 3 D | 14,80 | 3,490 | 10 |
| | 4 K | 17,60 | 2,591 | 10 |
| | 5 W | 19,30 | 2,669 | 10 |
| | Gesamt | | 13,16 | 5,787 |
| Gesamt | 1 A | 6,75 | 1,618 | 20 |
| | 2 G | 7,25 | 2,023 | 20 |
| | 3 D | 12,90 | 3,538 | 20 |
| | 4 K | 15,50 | 4,174 | 20 |
| | 5 W | 15,65 | 4,902 | 20 |
| | Gesamt | | 11,61 | 5,191 |

Nebenstehend die üblichen deskriptiven Statistiken.

Zweifaktorielle Varianzanalyse: Modelle mit Zufallseffekten in SPSS

Tests der Zwischensubjekteffekte

Abhängige Variable:time

| Quelle | | Quadratsumme vom Typ III | df | Mittel der Quadrate | F | Sig. |
|-----------------|-----------|--------------------------|----|----------------------|--------|------|
| Konstanter Term | Hypothese | 13479,210 | 1 | 13479,210 | 35,590 | ,004 |
| | Fehler | 1514,940 | 4 | 378,735 ^a | | |
| case | Hypothese | 240,250 | 1 | 240,250 | 5,050 | ,088 |
| | Fehler | 190,300 | 4 | 47,575 ^b | | |
| letter | Hypothese | 1514,940 | 4 | 378,735 | 7,961 | ,035 |
| | Fehler | 190,300 | 4 | 47,575 ^b | | |
| case * letter | Hypothese | 190,300 | 4 | 47,575 | 5,928 | ,000 |
| | Fehler | 722,300 | 90 | 8,026 ^c | | |

a. MS(letter)




b. MS(case * letter)

c. MS(Fehler)

Die Ergebnisse sind bis auf eine Ausnahme so wie oben berechnet. SPSS bestimmt den Test für den zufälligen Faktor abweichend.

Will man den Effekt des Faktors B „letter“ wie oben dargestellt testen, so muss statt $F = MQ_B / MQ_{A \times B} = MQ_{\text{letter}} / MQ_{\text{case} * \text{letter}} = 378.735 / 47.575 = 7.961$ (bei SPSS) entsprechend $F = MQ_B / MQ_{in} = MQ_{\text{letter}} / MQ_{\text{Fehler}} = 378.735 / 8.026 = 47.191$ bestimmt werden.

Zitierte Quellen:

-  Maxwell, S. E. & Delaney, H. D. (2004). *Designing experiments and analyzing data. A model comparison approach* (2nd ed.). New York: Psychology Press.
-  Overall, J. E. & Spiegel, D. K. (1969). Concerning least squares analysis of experimental data. *Psychological Bulletin*, 72, 311-322.
-  Zahn, I. (2010). Working with unbalanced cell sizes in multiple regression with categorical predictors. Internet am 30.6.2014 [<http://sites.psu.edu/stat461psbsp2013/wp-content/uploads/sites/1906/2013/03/InteractionsAndTypesOfSS.pdf>]