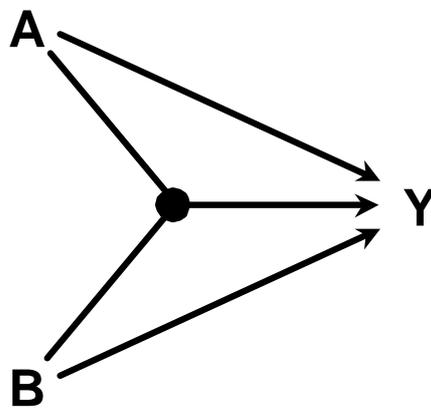


## Varianzanalyse

### Einleitung

Die Varianzanalyse ist ein sehr allgemein einsetzbares multivariates Analyseverfahren, mit dessen Hilfe Meßwerte einer abhängigen Variablen  $Y$ , die in der Regel von mehreren simultan wirksamen Faktoren (Variablen)  $A$ ,  $B$ , ... abhängen, analysiert werden können. Wie auch in der Regressionsanalyse wird für die abhängige Variable  $Y$  Intervallskalenniveau vorausgesetzt. Abweichend von der Regressionsanalyse wird aber für die unabhängigen Faktoren  $A$ ,  $B$ , ... lediglich Nominalskalenniveau verlangt.<sup>1</sup>

Die einer Varianzanalyse zugrunde liegende Kausalstruktur kann in Analogie zur Regressionsanalyse wie folgt dargestellt werden:



Wie der Name schon sagt, geht es in der Varianzanalyse ebenso wie in der Regressionsanalyse um die Erklärung der Varianz der abhängigen Variablen  $Y$  durch die unabhängigen Faktoren  $A$ ,  $B$ , ... Das Kausaldiagramm macht aber auch deutlich, daß im Unterschied zur Regressionsanalyse, in der üblicherweise nur die additive Kombination der linearen Effekte der unabhängigen Variablen berücksichtigt wird, auch die Interaktionseffekte zwischen den Faktoren zur Varianzerklärung herangezogen werden. Darüber hinaus wird die Form der Beziehung nicht von vornherein als linearer Zusammenhang unterstellt, was für nominalskalierte Faktoren natürlich auch keinen Sinn machen würde. Allerdings sind unter bestimmten Voraussetzungen an das Meßniveau der Faktoren mit Hilfe der Varianzanalyse Tests auf Linearität möglich (vgl. 1.5).

---

<sup>1</sup> Im folgenden beschränken wir die Diskussion der mehrfaktoriellen Varianzanalyse auf den zweifaktoriellen Fall, da bereits hier alle Besonderheiten der mehrfaktoriellen Konstellation erkennbar sind. Eine zusätzliche Berücksichtigung weiterer unabhängiger Faktoren würde nur zu einer hierarchischen Staffelung von Interaktionseffekten führen. So würde etwa die Einbeziehung eines weiteren Faktors  $C$  zu der Supervariablen  $A*B*C$  führen, deren erklärte Varianz in die Haupteffekte  $A$ ,  $B$  und  $C$ , die zweiseitigen Interaktionseffekte  $AB$ ,  $AC$  und  $BC$  und den dreiseitigen Interaktionseffekt  $ABC$  zu zerlegen wäre. Dabei ergeben sich die gleichen Probleme der Varianzzerlegung, wie sie im Kapitel 2 für den zweifaktoriellen Fall anhand der Supervariable  $A*B$  dargestellt werden - allerdings auf einer komplexeren Ebene.

Es ist nicht falsch, in der Varianzanalyse eine Verallgemeinerung der Regressionsanalyse zu sehen, in der die Beschränkung auf linear-additive Wirkungen der einzelnen unabhängigen Variablen aufgehoben ist. Dafür muß in Kauf genommen werden, daß die unabhängigen Faktoren lediglich nominalskaliert sind, jede Ausprägung eines Faktors also eine eigenständige kategoriale Bedeutung hat. Aber gerade dadurch ist die Varianzanalyse in idealer Weise geeignet, auch experimentelle Daten, die durch eine Vielzahl von experimentellen Faktoren erzeugt wurden, zu analysieren. Auch komplexe experimentelle Designs können daher in zugeordnete varianzanalytische Modelle übertragen und entsprechend analysiert werden.

Als Beispiel verwenden wir das im Skript eingehend diskutierte Experiment, in dem die langfristige Wirkung der Lektüre einer Tageszeitung auf die gesellschaftspolitischen Einstellungen der Versuchspersonen untersucht wurde. Wir konzipieren den eigentlichen experimentellen Stimulus Tageszeitung als Faktor A mit den drei Ausprägungen:

K: keine Tageszeitung

W: Die Welt

FR: Frankfurter Rundschau

und berücksichtigen als weiteres Merkmal, d.h. Faktor B, die Fachrichtung der als Versuchspersonen fungierenden Studenten mit den Ausprägungen:

wi: Wirtschaftswissenschaft

ing: Ingenieurwissenschaften

Die abhängige Variable spezifizieren wir als Einstellung zur Marktwirtschaft nach einer einjährigen, nicht selbst veranlaßten Lieferung der jeweiligen Tageszeitung. Das entsprechende experimentelle Design hat folgende Form:

Faktor A	keine Tageszeitung K	Die Welt W	Frankfurter Rundschau FR	
Faktor B				
Wirtschaftswissenschaft wi	$\bar{y}_{K,wi}$	$\bar{y}_{W,wi}$	$\bar{y}_{FR,wi}$	$\bar{y}_{wi}$
Ingenieurwissenschaften ing	$\bar{y}_{K,ing}$	$\bar{y}_{W,ing}$	$\bar{y}_{FR,ing}$	$\bar{y}_{ing}$
	$\bar{y}_K$	$\bar{y}_W$	$\bar{y}_{FR}$	$\bar{y}$

Durch das experimentelle Design werden also insgesamt sechs (3\*2) Versuchsgruppen erzeugt, für die jeweils Messungen der abhängigen Variablen Y (Einstellung zur Marktwirtschaft) vorliegen. Für jede der Versuchsgruppen kann dann der subgruppenspezifische Mittelwert  $\bar{y}_{i,j}$  in Abhängigkeit von den Ausprägungen der experimentellen Faktoren A= i und B= j berechnet werden.

Es können folgende Fragestellungen unterschieden werden:

- (1) Hat der Faktor A (Tageszeitung) einen Einfluß auf die Einstellung zur Marktwirtschaft (Y)?
- (2) Hat der Faktor B (Studienrichtung) einen Einfluß auf die Einstellung zur Marktwirtschaft (Y)?
- (3) Haben die Faktoren A und B einen gemeinsamen Einfluß auf die Einstellung zur Marktwirtschaft (Y)? Diese Frage nach dem Zusammenwirken der Faktoren A und B im Hinblick auf Veränderungen in der abhängigen Variablen Y beschreibt das eigentliche Ziel der Varianzanalyse. Insbesondere kann überprüft werden, ob zwischen den Faktoren A und B ein Interaktionseffekt existiert.

Diese Fragen können einmal als Fragen nach Mittelwertunterschieden in den einzelnen Versuchsgruppen verstanden werden:

- (1a) Gibt es signifikante Unterschiede in der Einstellung zur Marktwirtschaft zwischen Lesern der verschiedenen Tageszeitungen, d.h. unterscheiden sich die im Hinblick auf den Faktor A gebildeten Versuchsgruppen in ihren Mittelwerten  $\bar{y}_K$ ,  $\bar{y}_W$  und  $\bar{y}_{FR}$  signifikant voneinander?
- (2a) Gibt es signifikante Unterschiede in der Einstellung zur Marktwirtschaft zwischen Studenten der Wirtschaftswissenschaft und der Ingenieurwissenschaften, d.h. unterscheiden sich die im Hinblick auf Faktor B gebildeten Versuchsgruppen in ihren Mittelwerten  $\bar{y}_{wi}$  und  $\bar{y}_{ing}$  signifikant voneinander?<sup>2</sup>
- (3a) Gibt es signifikante Unterschiede in der Einstellung zur Marktwirtschaft zwischen den nach Faktor A und B gemeinsam gebildeten Versuchsgruppen, d.h. unterscheiden sich die Mittelwerte  $\bar{y}_{K,wi}$ ,  $\bar{y}_{K,ing}$ ,  $\bar{y}_{W,wi}$ ,  $\bar{y}_{W,ing}$ ,  $\bar{y}_{FR,wi}$  und  $\bar{y}_{FR,ing}$  signifikant voneinander? Findet man solche Unterschiede, so wird man weiter fragen, wie sich diese Unterschiede auf die Einflüsse des Faktors A, des Faktors B und einen eventuellen Interaktionseffekt zwischen den Faktoren A und B zurückführen lassen.

Andererseits können diese Fragen auch als Fragen nach dem Anteil der durch den jeweiligen Faktor erklärten Varianz gestellt werden:

- (1b) Wieviel Varianz in Y wird durch Faktor A erklärt?
- (2b) Wieviel Varianz in Y wird durch Faktor B erklärt?
- (3b) Wieviel Varianz in Y wird durch die Kombination beider Faktoren, d.h. durch die Supervariable A\*B erklärt?

---

<sup>2</sup> Dies ist exakt die Fragestellung des t-Tests. Der F-Test der Varianzanalyse kann als Verallgemeinerung des t-Tests auf mehrere Versuchsgruppen aufgefaßt werden (vgl. 1.2.1).

Hier wird die Grundidee der Varianzanalyse deutlich. In direkter Analogie zur Regressionsrechnung wird die gesamte Varianz der abhängigen Variablen in zwei Varianzanteile zerlegt, nämlich die durch die Faktoren erklärte Varianz und die nicht erklärte (Fehler-)Varianz. In der Varianzanalyse geschieht die Varianzzerlegung in Bezug auf die Summe der Abweichungsquadrate  $SSQ$  [sum of squares im SPSS-Output]. Je stärker die Mittelwertsunterschiede zwischen den Versuchsgruppen sind, um so größer wird  $SSQ_{\text{between}}$ , die Summe der Abweichungsquadrate zwischen den Gruppen und damit die durch den Faktor erklärte Varianz. Die Summe der Abweichungsquadrate innerhalb der Gruppen  $SSQ_{\text{within}}$  ist ein Maß für die Fehlervarianz, also die nicht durch den Faktor erklärbare Restvarianz. Die gesamte Summe der Abweichungsquadrate  $SSQ_{\text{total}}$  setzt sich aus diesen beiden Komponenten zusammen:

$$SSQ_{\text{total}} = SSQ_{\text{between}} + SSQ_{\text{within}}$$

Entsprechend können unsere obigen Fragestellungen auch formal varianzanalytisch formuliert werden und ergeben dann folgende Varianzzerlegungen:

$$(1c) \quad SSQ_{\text{total}} = SSQ_{\text{between}}(A) + SSQ_{\text{within}}(A)$$

$$(2c) \quad SSQ_{\text{total}} = SSQ_{\text{between}}(B) + SSQ_{\text{within}}(B)$$

$$(3c) \quad SSQ_{\text{total}} = SSQ_{\text{between}}(A*B) + SSQ_{\text{within}}(A*B)$$

(1c) und (2c) sind jeweils einfaktorielle Varianzanalysen - sie entsprechen der bivariaten Regression (mit jeweils einem Faktor als unabhängiger Variablen). Von besonderem Interesse ist die Konstellation (3c). In einer ersten Interpretation kann die Variable  $A*B$  als eine neu gebildete Supervariable verstanden werden, die alle durch Multiplikation der Ausprägungen der Faktoren A und B möglichen Kombinationen enthält (in unserem Beispiel also  $3*2 = 6$  Kombinationen). Es kann dann nach dem Zusammenwirken zwischen  $SSQ_{\text{between}}(A)$ ,  $SSQ_{\text{between}}(B)$  und  $SSQ_{\text{between}}(A*B)$  gefragt werden, d.h. man kann fragen, welche Anteile an der insgesamt von  $A*B$  erklärten Varianz (eindeutig?) auf den Faktor A bzw. den Faktor B zurückgeführt werden können. Methodisch entspricht dies der Frage nach der Möglichkeit der (eindeutig?) Zurechnung von kausalen Wirkungen einzelner Faktoren, die möglicherweise auch untereinander korreliert sind, auf eine abhängige Variable (vgl. 2.3).

Im Prinzip stellen sich also bei der Varianzanalyse die gleichen Probleme der Varianzzurechnung wie bei der Regressionsrechnung. Allerdings wird bei der Varianzanalyse neben der additiven Wirkung der beiden Faktoren von vornherein ein Wechselwirkungseffekt (der sog. Interaktionseffekt) der beiden Faktoren mit berücksichtigt. Die Varianzanalyse kann daher als ein allgemeiner Fall der kausalen Abhängigkeitsanalyse gemäß dem oben skizzierten Kausaldiagramm angesehen werden, wobei die unabhängigen Variablen, d.h. die Faktoren A und B, lediglich über Nominalskalenniveau verfügen.

# 1. Einfaktorielle Varianzanalyse

## 1.1 Vorstellung der Beispieldaten

Um die grundlegenden Überlegungen zur Varianzanalyse möglichst einfach darzustellen, verwenden wir in den folgenden Beispielen eine Datenmatrix mit hypothetischen Zahlen. In der Matrix sind die abhängige intervallskalierte Variable Y sowie die Faktoren A und B mit drei bzw. zwei Kategorien enthalten. Die Datenmatrix enthält 18 Fälle:

A	B	Y
1	1	1
1	1	2
1	1	3
2	1	4
2	1	5
2	1	6
3	1	7
3	1	8
3	1	9
1	2	10
1	2	11
1	2	12
2	2	13
2	2	14
2	2	15
3	2	16
3	2	17
3	2	18

Es ergeben sich 6 Kombinationen mit jeweils unterschiedlichen Ausprägungen von A und B. Die Mittelwerte von Y für jede Kombination sind in der folgenden Tabelle aufgeführt (bei Verständnisschwierigkeiten bietet es sich an, die jeweiligen Mittelwerte  $\bar{y}_{AB}$  anhand der angegebenen Datenmatrix nachzurechnen):

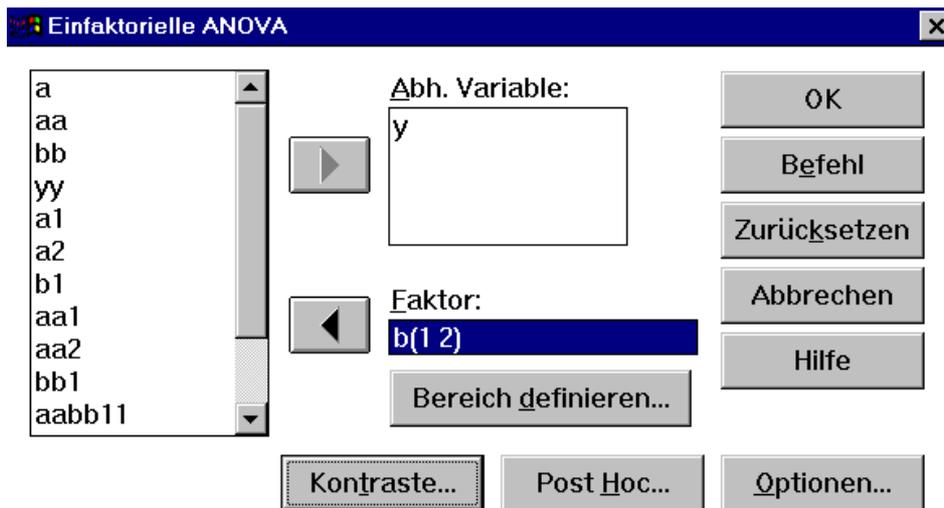
$\bar{Y}_{AB}$	A=1	A=2	A=3	
B=1	2	5	8	5
B=2	11	14	17	14
	6,5	9,5	12,5	9,5

## 1.2 Die Varianzanalyse als Verallgemeinerung des t-Tests

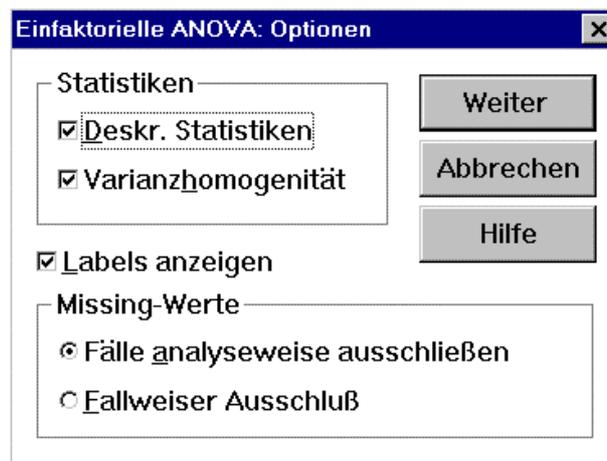
### 1.2.1 Ein Faktor mit zwei Ausprägungen

Wir betrachten zunächst den einfachsten Fall einer einfaktoriellen Varianzanalyse. Die unabhängige Variable B (der sog. Faktor) hat hier nur zwei Kategorien 1 und 2. Untersucht wird der Mittelwertunterschied der abhängigen Variable Y zwischen diesen beiden Kategorien. Diese Anordnung entspricht exakt dem t-Test auf Mittelwertunterschiede.

Die einfaktorielle Varianzanalyse wird in der SPSS-Menüleiste über die Menüfelder Statistik → Mittelwertvergleiche → Einfaktorielle ANOVA aufgerufen.



Die abhängige Variable (Y) und der Faktor (B) werden mit den Pfeiltasten in die entsprechenden Felder in der Bildmitte übertragen. Für den Faktor B muß noch genauer spezifiziert werden, welche Zahlen die zu untersuchenden Kategorien einschließen. Dies erfolgt über den Button Bereich definieren. In unserem Fall werden die Kategorien von 1 bis 2 berücksichtigt (da B nur 2 Kategorien hat) - wir werden weiter unten sehen, daß hier auch eine größere Spannweite möglich ist. Über den Button Optionen gelangen wir zu folgendem Dialogfeld:



Hier fordern wir zusätzlich die Deskriptiven Statistiken und den Levene Test auf Varianzhomogenität an (durch ankreuzen). Der Button Weiter führt uns zurück zum ersten Dialogfeld. Der Button OK führt die einfaktorielle Varianzanalyse mit den von uns getroffenen Einstellungen durch. Der von SPSS intern verarbeitete Befehl lautet oneway. Er wird in ein Syntaxfenster ausgegeben, wenn der Button Befehl anstelle von OK gedrückt wird:

```
oneway variables y by b(1,2) / statistics descriptives homogeneity.
```

Wir erhalten den folgenden Output:

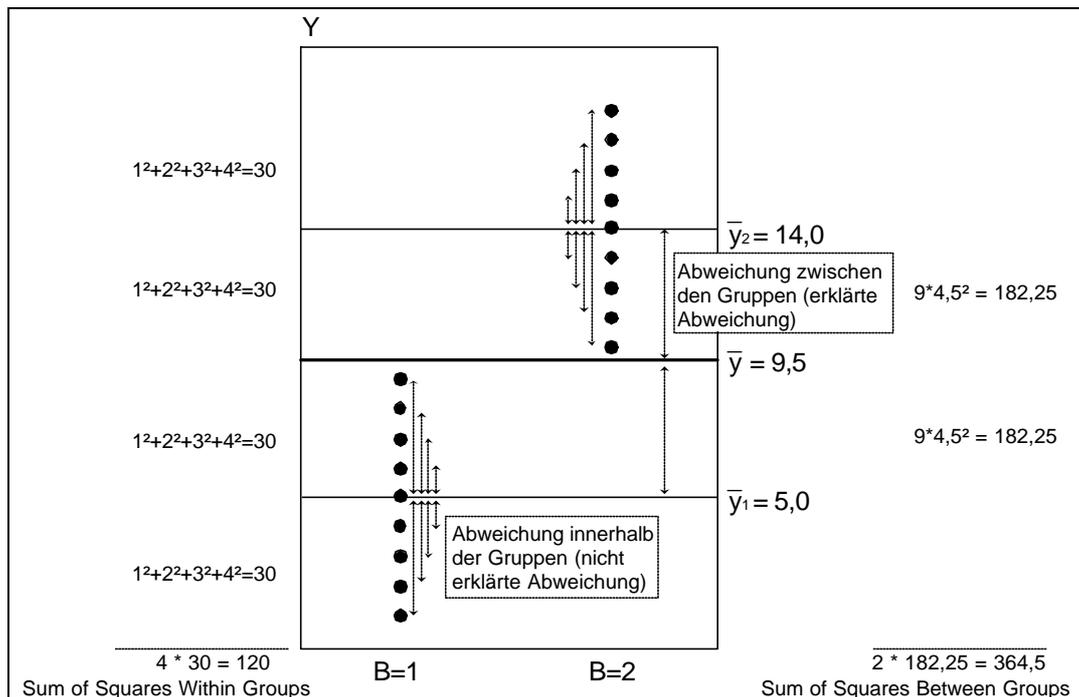
- - - - - O N E W A Y - - - - -

Variable Y  
By Variable B

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	1	364,5000	364,5000	48,6000	,0000
Within Groups	16	120,0000	7,5000		
Total	17	484,5000			

Die wichtigste Information erhalten wir bereits zu Beginn des Outputs. Unter der Überschrift Analysis of Variance wird das Ergebnis der Varianzzerlegung ausgegeben. Die Summe der Abweichungsquadrate (Sum of Squares) (vgl. 1.3) wird zerlegt in einen Anteil zwischen den Gruppen  $SSQ_{\text{Between}}$  (Between Groups) und einen Anteil innerhalb der Gruppen  $SSQ_{\text{Within}}$  (Within Groups). Die Gesamtstreuung  $SSQ_{\text{Total}}$  (die Summe aus Between Groups und Within Groups) wird als Total ausgegeben. Das Verhältnis von  $SSQ_{\text{Between}}$  zu  $SSQ_{\text{Total}}$  entspricht dem Anteil der erklärten Varianz, das Verhältnis von  $SSQ_{\text{Within}}$  zu  $SSQ_{\text{Total}}$  ergibt den Anteil der Fehlervarianz. In der folgenden Abbildung ist die Zerlegung der Varianz für diesen Fall grafisch dargestellt.



Aus der Summe der Abweichungsquadrate werden die mittleren Abweichungsquadrate MSQ (Mean Squares) berechnet, indem die Sum of Squares durch die Anzahl der Freiheitsgrade (D.F.) geteilt werden.<sup>3</sup> Die mittleren Abweichungsquadrate entsprechen den Varianzen.

<sup>3</sup> Für die Freiheitsgrade zwischen den Gruppen gilt:  $df_{\text{Between}} = (\text{Anzahl der Gruppen} - 1)$ . Für die Freiheitsgrade innerhalb der Gruppen gilt:  $df_{\text{Within}} = \text{Anzahl der Gruppen} \cdot (\text{Zahl der Beobachtungen pro Gruppe} - 1)$ . In unserem Beispiel ergibt sich somit  $df_{\text{Between}} = 2 - 1 = 1$  und  $df_{\text{Within}} = 2 \cdot (9 - 1) = 16$  (vgl. auch 1.3).

Wir erhalten  $MSQ_{\text{Between}} = 364,5000$  für die Varianz zwischen den Gruppen und  $MSQ_{\text{Within}} = 7,5000$  für die Varianz innerhalb der Gruppen. Entscheidend ist  $MSQ_{\text{Between}}$ , die Varianz zwischen den Gruppen. Je größer dieser Wert im Verhältnis zur  $MSQ_{\text{Within}}$ , der Varianz innerhalb der Gruppen ist, desto eher kann man auf signifikante Mittelwertunterschiede zwischen den Gruppen in der Grundgesamtheit schließen. Als statistischer Test wird ein F-Test durchgeführt. Die Prüfgröße F wird berechnet als Quotient aus  $MSQ_{\text{Between}}$  und  $MSQ_{\text{Within}}$  (vgl. auch 1.3):

$$F_{(1,16)} = \frac{\frac{SSQ_{\text{Between}}}{1}}{\frac{SSQ_{\text{Within}}}{16}} = \frac{MSQ_{\text{Between}}}{MSQ_{\text{Within}}}$$

Getestet wird die Nullhypothese ( $H_0$ ), die Gruppenmittelwerte seien gleich. In diesem Fall müßten  $MSQ_{\text{Between}}$  und  $MSQ_{\text{Within}}$  in der Stichprobe ungefähr gleich groß sein. Genaugenommen wird also Gleichheit zweier Varianzen getestet. Wenn die Prüfgröße F den kritischen Wert für ein gegebenes Signifikanzniveau  $\alpha$  überschreitet, wird die Nullhypothese verworfen.<sup>4</sup> In logischer Konsequenz muß dann die Alternativhypothese ( $H_1$ ) angenommen werden, die lautet, daß es signifikante Abweichungen zwischen den Gruppenmittelwerten gibt.

In unserem Beispiel ist dies der Fall. Die Wahrscheinlichkeit, daß die Prüfgröße  $F_{(1,16)}$  den Wert  $48,6000$  rein zufällig überschreitet, ist  $0,0000$ , d.h. in den ersten 4 Stellen nach dem Komma nicht von Null verschieden. Wir verwerfen also die Nullhypothese. Die Irrtumswahrscheinlichkeit dabei beträgt annähernd Null. Das bedeutet, daß sich die Mittelwerte der Gruppen hochsignifikant unterscheiden.

Wie groß die Mittelwertunterschiede in der Stichprobe tatsächlich sind, sehen wir im nächsten Teil des Outputs:

Group	Count	Mean	Standard Deviation	Standard Error	95 Pct	Conf Int	for Mean
Grp 1	9	5,0000	2,7386	,9129	2,8949	TO	7,1051
Grp 2	9	14,0000	2,7386	,9129	11,8949	TO	16,1051
Total	18	9,5000	5,3385	1,2583	6,8452	TO	12,1548

Der Mittelwert (Mean) der Gruppe 1 beträgt  $5,0000$ . Der Mittelwert der Gruppe 2 liegt bei  $14,0000$ . Die Standardabweichung (Standard Deviation) ist in beiden Gruppen gleich  $2,7386$ . Zusätzlich wird für jeden Mittelwert ein Konfidenzintervall angegeben, in dem der Mittelwert in der Grundgesamtheit mit 95%iger Sicherheit liegen wird. In der Gruppe 1 liegt der Mittelwert in der Grundgesamtheit wahrscheinlich zwischen  $2,8949$  und  $7,1051$ . Der Mittelwert der Gruppe 2 beträgt wahrscheinlich mindestens  $11,8949$  und höchstens  $16,1051$ .

---

<sup>4</sup> Die kritischen Werte der F-Verteilung sind in einschlägigen Statistikbüchern in Abhängigkeit der Freiheitsgrade tabelliert. Der kritische Wert ist allgemein  $F_{dfB,dfW,\alpha}$ , in unserem Fall gilt bei einer Irrtumswahrscheinlichkeit von  $\alpha = 0,05$ :  $F_{1,16,0,05} = 4,49$ .

Abschließend wird ein Levene Test auf Varianzhomogenität durchgeführt. Die Gleichheit der Varianzen der einzelnen Gruppen ist eine notwendige Voraussetzung für die Anwendung des F-Testes in der einfaktoriellen Varianzanalyse (vgl. 1.3). Insofern ist die Platzierung dieses Tests am Ende des Outputs unglücklich. Die Nullhypothese ( $H_0$ ) des Levene Testes lautet, die Varianzen der Gruppen sind gleich (also homogen). Auch hier wird ein F-Wert berechnet (Statistic) und die zugehörige Wahrscheinlichkeit angegeben.

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
,0000	1	16	1,000

In unserem Fall ist die Prüfgröße gleich Null und die zugehörige Wahrscheinlichkeit ist gleich eins -  $H_0$  sollte also auf keinen Fall verworfen werden. Wie oben im Output zu sehen, ist die Standardabweichung (Wurzel der Varianz) in beiden Gruppen tatsächlich identisch und deshalb wird die Prüfgröße zu Null.

Als abschließendes Ergebnis läßt sich festhalten, daß signifikante Mittelwertunterschiede zwischen den Gruppen bestehen und daß die Voraussetzungen zur Anwendung der einfaktoriellen Varianzanalyse erfüllt sind.

Dieses Ergebnis läßt sich in diesem Fall auch über einen t-Test auf Mittelwertunterschiede reproduzieren. Wir betrachten direkt den Output des t-Tests (vgl. auch die Beispieldatei zum t-Test):

t-tests for Independent Samples of B

Variable	Number of Cases	Mean	SD	SE of Mean
Y				
B 1	9	5,0000	2,739	,913
B 2	9	14,0000	2,739	,913

Mean Difference = -9,0000

Levene's Test for Equality of Variances: F= ,000 P= 1,000

t-test for Equality of Means					95%
Variances	t-value	df	2-Tail Sig	SE of Diff	CI for Diff
Equal	-6,97	16	,000	1,291	(-11,737; -6,263)
Unequal	-6,97	16,00	,000	1,291	(-11,737; -6,263)

Der t-Wert beträgt  $-6,97$  und ist bei einer Irrtumswahrscheinlichkeit von angenähert  $0,000$  signifikant. Das bedeutet, daß die Differenz der Mittelwerte (Mean Difference) signifikant von Null verschieden ist. In diesem Fall gilt im übrigen, daß der F-Wert der Varianzanalyse ( $48,5$ ; s.o.) gleich dem quadrierten t-Wert ( $6,97^2$ ) ist (von Rundungsfehlern abgesehen).

### 1.2.2 Ein Faktor mit mehr als zwei Ausprägungen

Analog zu 1.2.1 verläuft die Varianzanalyse auch mit einem Faktor, der mehr als 2 Kategorien enthält. Wir untersuchen nun die Mittelwertunterschiede der Variable Y für den Faktor A mit drei Ausprägungen. Die Vorgehensweise gleicht der unter 1.2.1, mit dem Unterschied, daß als Bereich die Zahlen 1 bis 3 angegeben werden. Der SPSS-Befehl lautet dafür:

```
oneway variables y by a(1,3) / statistics descriptives homogeneity.
```

Wir erhalten den folgenden Output:

```

- - - - - O N E W A Y - - - - -

Variable Y
By Variable A

Analysis of Variance

Source          D.F.      Sum of Squares      Mean Squares      F Ratio      F Prob.
Between Groups      2         108,0000           54,0000           2,1514      ,1508
Within Groups      15         376,5000           25,1000
Total              17         484,5000

```

Die Interpretation der Varianzzerlegung gleicht ebenfalls der unter 1.2.1. Es fällt auf, daß sich die Freiheitsgrade geändert haben. Die  $df_{\text{Between}}$  betragen nun 2 (= 3 Kategorien - 1), die  $df_{\text{Within}}$  betragen hier 15 (= 3 · (6 Fälle - 1)). Die Prüfgröße F, der Quotient aus  $MSQ_{\text{Between}}$  und  $MSQ_{\text{Within}}$  ( $54/25,1$ ), beträgt hier nur noch 2,1514 und ist somit kleiner als der kritische Wert  $F_{2,15,0.05} = 3,68$ . Die Irrtumswahrscheinlichkeit bei Zurückweisen von  $H_0$  beträgt 0,1508. Es liegen also keine signifikanten Mittelwertunterschiede vor.

Group	Count	Mean	Standard Deviation	Standard Error	95 Pct Conf Int for Mean
Grp 1	6	6,5000	5,0100	2,0453	1,2424 TO 11,7576
Grp 2	6	9,5000	5,0100	2,0453	4,2424 TO 14,7576
Grp 3	6	12,5000	5,0100	2,0453	7,2424 TO 17,7576
Total	18	9,5000	5,3385	1,2583	6,8452 TO 12,1548

Wir sehen, daß die drei Gruppen durchaus unterschiedliche Mittelwerte aufweisen. Aufgrund der geringen Fallzahlen und der hohen Standardabweichung in den Gruppen sind jedoch die Konfidenzintervalle so breit, daß sie sich überschneiden. Es wäre z.B. möglich, daß alle drei Mittelwerte in der Grundgesamtheit 10,0 betragen, denn dieser Wert ist von allen drei Konfidenzintervallen eingeschlossen.

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
,0000	2	15	1,000

Der Levene Test auf Varianzhomogenität fällt genauso aus wie unter 1.2.1. Wir können also davon ausgehen, daß die Varianzen in beiden Gruppen gleich sind..

Wir halten fest, daß keine signifikanten Mittelwertunterschiede von Y in den Gruppen von A festzustellen sind. Allerdings ändert sich dieses Ergebnis, wenn zusätzlich Faktor B in einer mehrfaktoriellen Varianzanalyse kontrolliert wird (vgl. 2.1 und die Bemerkung in 1.3).

Um die Parallele zur Regressionsanalyse zu verdeutlichen, untersuchen wir den Einfluß von A auf Y erneut, diesmal jedoch im Rahmen eines linearen Regressionsmodells (vgl. auch die Beispieldatei zur linearen Regression). Da A nominalskaliert mit mehr als zwei Ausprägungen ist, müssen wir zunächst zwei geeignete Dummy-Variablen bilden. Zur Konstruktion der Dummy-Variablen verwenden wir die Effekt-Codierung und wählen die Ausprägung A= 3 als Referenzkategorie. Wir erhalten zwei Dummy-Variablen nach folgendem Schema:

Ursprungs- variable	Dummy-Variablen	
	A1	A2
A		
1	1	0
2	0	1
3	-1	-1

Wenn A den Wert 3 annimmt, nehmen A1 und A2 den Wert -1 an. Für A = 1 wird A1 = 1, für A = 2 gilt A2 = 1. Die lineare Regression mit Y als abhängiger und A2 und A3 als unabhängigen Variablen liefert folgendes Ergebnis (gekürzt):

```
* * * * M U L T I P L E   R E G R E S S I O N   * * * *
Equation Number 1      Dependent Variable..  Y
Block Number  1.  Method:  Enter      A2      A3

Multiple R              ,47213
R Square                ,22291
Adjusted R Square       ,11930
Standard Error          5,00999

Analysis of Variance
      DF      Sum of Squares      Mean Square
Regression      2          108,00000          54,00000
Residual        15          376,50000          25,10000

F =          2,15139      Signif F =  ,1508
```

$SSQ_{\text{Between}}$  erscheint hier als durch die multiple Regression erklärte Varianz (Regression 108,00) und  $SSQ_{\text{Within}}$  als Fehlervarianz (Residual 376,00). Die Prüfgröße F beträgt genau wie bei der einfaktoriellen Varianzanalyse 2,1514 und ist somit nicht signifikant. Die erklärte Varianz  $R^2$  (R Square) beträgt wie oben 22,3%.

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
A1	-3,000000	1,669997	-,472134	-1,796	,0926
A2	1,86237E-16	1,669997	2,931E-17	,000	1,0000
(Constant)	9,500000	1,180866		8,045	,0000

Die Mittelwertdifferenzen zwischen den einzelnen Gruppen sind leicht mit den B-Werten zu reproduzieren. Dazu verwenden wir die Regressionsgleichung für den geschätzten Wert  $\hat{Y}$  und setzen für A1 und A2 die entsprechenden Werte der Effekt-Codierung ein:

Regressionsgleichung  $\hat{Y} = c + b_1 a_1 + b_2 a_2 = 9,5 - 3,0 \cdot a_1 + 0,0 \cdot a_2$   
für A = 1  $\hat{Y} = 9,5 - 3,0 \cdot 1 + 0,0 \cdot 0 = 6,5$   
für A = 2  $\hat{Y} = 9,5 - 3,0 \cdot 0 + 0,0 \cdot 1 = 9,5$   
für A = 3  $\hat{Y} = 9,5 - 3,0 \cdot -1 + 0,0 \cdot -1 = 12,5$

Die Schätzer  $\hat{Y}$  entsprechen genau den Gruppenmittelwerten von Y für den Faktor A. Die Konstante entspricht dem Gesamtmittelwert.

Häufig interessiert man sich neben der Frage, ob es überhaupt Mittelwertunterschiede in der abhängigen Variablen für die einzelnen Kategorien des Faktors gibt, auch dafür, welche Gruppenmittelwerte sich signifikant voneinander unterscheiden. Hier wäre es zunächst naheliegend, eine Reihe von einzelnen t-Tests für unterschiedliche Paarungen von Gruppen durchzuführen (in diesem Fall A(1,2), A(1,3) und A(2,3)). Diese Vorgehensweise ist jedoch aus statistischen Gründen nicht zulässig, da bei wiederholter Anwendung von t-Tests auf die gleiche Stichprobe bereits rein zufällig ein Teil der Tests signifikant ausfallen könnte. Statt dessen verwendet man für die Untersuchung der spezifischen Mittelwertunterschiede zwischen einzelnen Kategorienpaarungen spezielle Tests für multiple Mittelwertvergleiche. Da es für Faktor A an dieser Stelle keinen Sinn macht, die einzelnen Mittelwertunterschiede genauer zu untersuchen (wir wissen bereits, daß sich kein einziger Mittelwert signifikant von den anderen unterscheidet), behandeln wir diese multiplen Vergleichstests im Rahmen des folgenden Beispiels aus der Forschungspraxis (siehe 1.4). Vorher wollen wir aber die allgemeinen Grundlagen des varianzanalytischen Modells darstellen.

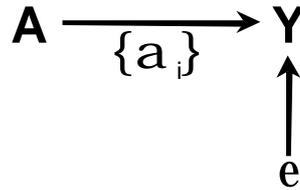
### 1.3 Das Modell der einfaktorielle Varianzanalyse im Überblick

Wir gehen von einem Faktor A mit I Ausprägungen aus.  $n_i$  bezeichnet die Gruppengröße der i-ten Versuchsgruppe und  $y_{iz}$  den Meßwert für die z-te Versuchsperson in der i-ten Gruppe. Die Grundgleichung der Varianzanalyse lautet:

$$y_{iz} = \mu + \alpha_i + \varepsilon_{iz} \quad (1.3.1)$$

Der Meßwert der z-ten Versuchsperson in der i-ten Gruppe setzt sich also aus dem Gesamtmittelwert  $\mu$ , dem gruppenspezifischen Effekt  $\alpha_i$  und einer Zufallskomponente  $\varepsilon_{iz}$  zusammen,

wobei wir für  $\varepsilon$  die üblichen Eigenschaften eines Zufallsfehlers (Unabhängigkeit, Mittelwert = 0) unterstellen. Kausalanalytisch kann das Modell also folgendermaßen skizziert werden:



Das Diagramm macht die Parallele zur bivariaten Regression und der Gleichung  $Y = a + bX + \varepsilon$  deutlich.

$$\begin{aligned} \text{Gruppenmittelwert der } i\text{-ten Gruppe } y_{i\cdot} &:= \sum_{z=1}^{n_i} y_{iz} / n_i \\ \text{Gesamtmittelwert } y_{\cdot\cdot} &:= \sum_{i=1}^I \sum_{z=1}^{n_i} y_{iz} / N \\ &= \sum_{i=1}^I y_{i\cdot} / I \end{aligned}$$

wobei  $N = \sum_{i=1}^I n_i$  die Gesamtfallzahl darstellt. Für den Fall gleich großer Gruppen (balanciertes Design; vgl. 2.1) gilt offensichtlich:  $n \cdot I = N$ .

Folgende Umformung stellt einen Bezug zur Grundgleichung (1.3.1.) her:

$$y_{iz} = y_{\cdot\cdot} + (y_{i\cdot} - y_{\cdot\cdot}) + (y_{iz} - y_{i\cdot}) \quad (1.3.2)$$

$$\begin{matrix} (m) & (a_i) & (e_{iz}) \end{matrix}$$

Die Zerlegung in die Summe der Abweichungsquadrate SSQ ergibt sich wie folgt:

$$\begin{aligned} \text{SSQ}_{\text{total}} &:= \sum_{i=1}^I \sum_{z=1}^{n_i} (y_{iz} - y_{\cdot\cdot})^2 \\ \text{SSQ}_{\text{between}} &:= \sum_{i=1}^I n_i (y_{i\cdot} - y_{\cdot\cdot})^2 \end{aligned}$$

Die erklärte Varianz (Varianz zwischen den Gruppen) ergibt sich also, indem für jede Gruppe der jeweilige Gruppenmittelwert als bester Schätzer verwendet wird.

$$\text{SSQ}_{\text{within}} := \sum_{i=1}^I \sum_{z=1}^{n_i} (y_{iz} - y_{i\cdot})^2$$

Die Fehlervarianz (Varianz innerhalb der Gruppen) ergibt sich danach als individuelle Abweichung vom jeweiligen Gruppenmittelwert.

Gemäß dem Fundamentaltheorem der Varianzanalyse gilt die folgende Varianzerlegung:

$$\text{SSQ}_{\text{total}} = \text{SSQ}_{\text{between}} + \text{SSQ}_{\text{within}}$$

Dies entspricht der Grundgleichung der Regressionsanalyse, die die Gesamtvarianz als Summe von erklärter Varianz und Fehlervarianz darstellt (vgl. Skript zur Vorlesung, Kap. 3.1.1).

Um Signifikanztests durchführen zu können, wird in der Varianzanalyse die folgende Verteilungsvoraussetzung gemacht:

**Die Fehlerterme  $e_{iz}$  sind unabhängig, normalverteilt und haben die gleiche Varianz.**

Auch diese Annahmen sind exakt parallel zu den entsprechenden Voraussetzungen für die Fehlerterme im Regressionsmodell. Die Annahme der Varianzhomogenität wird auch als Homoskedastizität bezeichnet.

Vor Durchführung einer Varianzanalyse müssen die Verteilungsvoraussetzungen überprüft werden, wenn auch gesagt werden kann, daß das Verfahren relativ robust gegenüber Verletzungen dieser Voraussetzungen ist. Liegt eine Zufallsstichprobe vor, ist die Unabhängigkeitsannahme immer erfüllt. Die Normalverteilungsannahme in Bezug auf die abhängige Variable kann in SPSS durch graphische Tests oder den Kolmogoroff-Smirnov-Test geprüft werden. Zur Überprüfung der Varianzhomogenität steht im Rahmen der ONEWAY- und der ANOVA-Prozedur der Levene-Test zur Verfügung.

Signifikanztests im Rahmen der Varianzanalyse werden mit Hilfe der F-Verteilung durchgeführt. Dazu muß zunächst auf die Zahl der Freiheitsgrade eingegangen werden, die den jeweiligen Komponenten der Varianzzerlegung zugeordnet sind.

SSQ	df
$\text{SSQ}_{\text{between}}$	I - 1
$\text{SSQ}_{\text{within}}$	N - I
$\text{SSQ}_{\text{total}}$	N - 1

Mit Hilfe der Freiheitsgrade können die mittleren Abweichungsquadrate MSQ [mean squares] aus den summierten Abweichungsquadraten berechnet werden, indem man durch die jeweiligen Freiheitsgrade dividiert:

$$\text{MSQ} = \frac{\text{SSQ}}{\text{df}}$$

Um zu testen, ob die vorliegenden Mittelwertsunterschiede zwischen den Versuchsgruppen signifikant sind, d.h. ob die durch Faktor A erklärte Varianz signifikant ist, wird der folgende F-Wert gebildet:

$$F_{(I-1, N-I)} = \frac{\frac{SSQ_{\text{between}}}{I-1}}{\frac{SSQ_{\text{within}}}{N-I}} = \frac{MSQ_{\text{between}}}{MSQ_{\text{within}}} \quad (1.3.3)$$

Sind die Verteilungsvoraussetzungen für die Varianzanalyse erfüllt, ist die obige Statistik F-verteilt. Der SPSS-Output der Prozeduren ONEWAY und ANOVA gibt an der entsprechenden Stelle gleich den zugehörigen p-Wert (F Prob.) an, der angibt, wie wahrscheinlich die Daten unter Annahme der Nullhypothese gleiche Gruppenmittelwerte sind. Für das Standard-Signifikanzniveau von 5% muß der angegebene p-Wert also kleiner als 0,05 sein.

Bemerkung: Der Nenner des F-Wertes enthält die Fehlervarianz, wie sie in dem jeweils spezifizierten varianzanalytischen Modell definiert ist. Da aber in dem einfaktoriellen Modell für Faktor A in 1.2.1 der zweite Faktor B nicht kontrolliert ist, erscheint dessen Varianz als Fehlervarianz. Entsprechend hoch ist der mittlere quadratische Fehler ( $MSQ_{\text{within}}$ ) mit 25.1, der verhindert, daß die Varianzerklärung durch A signifikant ist ( $p = .1508$ ). Kontrolliert man dagegen in der zweifaktoriellen Varianzanalyse gleichzeitig den Faktor B, bleibt nur ein mittlerer quadratischer Fehler ( $MSQ_{\text{within}}$ ) von 1.0, so daß der von A erklärte Varianzanteil nun hochsignifikant ist ( $p = 0.000$ ) [vgl. 2.1], obwohl der Anteil der durch A erklärten Varianz unverändert geblieben ist.

#### **1.4 Beispiel aus der Forschungspraxis: Wahlabsicht und Links-Rechts-Selbsteinstufung**

Im ALLBUS 1996 wird (wie auch in allen zurückliegenden ALLBUS-Studien) die Frage nach der Wahlabsicht gestellt, falls am nächsten Sonntag Bundestagswahl wäre ("Sonntagsfrage"). Zu einem früheren Zeitpunkt in der Befragung wurden die Befragten gebeten, sich auf einer Links-Rechts-Skala mit 10 Ausprägungen (1 = links ... 10 = rechts) selbst einzustufen. Es liegt nahe, zu vermuten, daß je nach Parteipräferenz eine Einstufung eher links oder eher rechts erfolgt. Die Wählergruppen sollten sich also in ihren Mittelwerten auf der Links-Rechts-Skala unterscheiden. Die Wähler rechter Parteien sollten höhere Mittelwerte aufweisen als die Wähler linker Parteien.

Diese Hypothese wurde mit einer einfaktoriellen ANOVA untersucht. Als abhängige Variable wurde die Links-Rechts-Selbsteinstufung (V112) eingetragen, als Faktor wurde die Wahlabsicht (V325) mit den Ausprägungen 1 = CDU-CSU, 2 = SPD, 3 = F.D.P., 6 = Bündnis 90/Die Grünen, 7 = Republikaner und 8 = PDS eingetragen.

Wir erhalten folgenden Output:

- - - - - O N E W A Y - - - - -

Variable V112 LINKS-RECHTS-SELBSTEINSTUFUNG, BEFR.  
By Variable V325 WAHLABSICHT, BUNDESTAGSWAHL; BEFR.

Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	5	2294,0057	458,8011	179,9129	,0000
Within Groups	2235	5699,5374	2,5501		
Total	2240	7993,5431			

Die Prüfgröße F beträgt 179,9 und liegt damit weit über dem kritischen Wert von ca. 2,21. Es gibt also signifikante Unterschiede zwischen den Mittelwerten der einzelnen Gruppen. Die Streuung zwischen den Gruppen hat mit 2294 einen Anteil an der Gesamtstreuung (7994) von 28,7%. Das bedeutet, daß die Wahlabsicht 28,7% der Varianz der Links-Rechts-Selbsteinstufung erklärt.

Die nächste Tabelle vermittelt einen ersten Eindruck, welche (potentiellen) Wähler sich eher links bzw. eher rechts einstufen.

Group	Count	Mean	Standard Deviation	Standard Error	95 Pct Conf Int	for Mean
CDU-CSU	781	6,2292	1,5940	,0570	6,1172 TO	6,3412
SPD	721	4,5687	1,6503	,0615	4,4480 TO	4,6893
F.D.P.	206	5,8155	1,4996	,1045	5,6095 TO	6,0215
BUENDNIS	358	4,0810	1,4976	,0791	3,9253 TO	4,2367
REPUBLIK	49	7,2245	1,9818	,2831	6,6552 TO	7,7937
PDS	126	3,3016	1,5660	,1395	3,0255 TO	3,5777
Total	2241	5,1709	1,8891	,0399	5,0927 TO	5,2492

Wir erinnern uns, daß kleinere Mittelwerte eine eher linke Einstufung bedeuten. Potentielle PDS-Wähler stufen sich also im Schnitt mit 3,3 am weitesten links auf der Links-Rechts-Skala ein, Wähler der Republikaner mit 7,2 erwartungsgemäß am weitesten rechts. Es fällt weiterhin auf, daß sich keines der ausgegebenen Konfidenzintervalle überschneidet. Dies ist ein erster Hinweis darauf, daß nicht nur die Mittelwerte der Wähler der Extremparteien, sondern evt. alle Mittelwerte signifikant unterschiedlich sind.

Der Test auf Varianzhomogenität zeigt an, daß die Voraussetzung der Varianzhomogenität in diesem Fall nicht gegeben ist. Mindestens die Varianz einer der Gruppen unterscheidet sich signifikant von den anderen. Hier fällt auf, daß vor allem die Republikaner über eine sehr inhomogene Anhängerschaft verfügen.

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
2,5698	5	2235	,025

Generell ist anzumerken, daß bei sehr großen Fallzahlen (hier 2241) bereits sehr feine Unterschiede in den Varianzen als signifikant ausgewiesen werden. Dennoch sollten die folgenden Signifikanztests wegen der fehlenden Varianzhomogenität mit Vorsicht interpretiert werden.

Um der oben aufgeworfenen Frage nachzugehen, welche der 6 potentiellen Wählergruppen sich von den anderen signifikant unterscheiden, könnte man auf die Idee kommen, paarweise t-Tests (z.B. CDU-CSU vs. B90/Grüne ...) durchzuführen. Dies ist jedoch zum einen unpraktikabel, denn es lassen sich 15 verschiedene Paarungen bilden, was einen unverhältnismäßig hohen Arbeitsaufwand bedeuten würde. Zum anderen ist eine Aneinanderreihung von t-Tests für Subgruppen einer kategorialen Variable auch aus statistischer Sicht problematisch. Bei einer gegebenen Irrtumswahrscheinlichkeit von z.B. 5% ist es zu erwarten, dass wenigstens einer der 15 Tests rein zufällig signifikant ausfällt. Um diese Probleme zu umgehen, stellt die SPSS-Prozedur der einfaktoriellen ANOVA eine Reihe von multiplen Vergleichstests zur Verfügung, die gleichzeitig alle Mittelwertdifferenzen untersuchen und die alle (bis auf einen Test!) auch das Problem der Irrtumswahrscheinlichkeit bei wiederholten statistischen Tests berücksichtigen.

Die verschiedenen Tests können über den Button Post Hoc Mittelwertvergleiche angefordert werden. Der Test auf geringste signifikante Differenz (LSD-Test) ist nichts anderes als eine Aneinanderreihung von t-Tests ohne Korrektur der Gesamtfehlerrate - daher sollte dieser Test nicht verwendet werden. Die übrigen Tests berücksichtigen alle die Gesamtfehlerrate, allerdings jeweils auf unterschiedliche Weise.<sup>5</sup> Wir weisen nachfolgend das Ergebnis des Tukey-Tests auf ehrlich signifikante Differenz (Tukey-HSD) aus:

```
Multiple Range Tests: Tukey-HSD test with significance level ,050
```

```
The difference between two means is significant if  
MEAN(J)-MEAN(I) >= 1,1292 * RANGE * SQRT(1/N(I) + 1/N(J))  
with the following value(s) for RANGE: 4,04
```

Im Output jedes Tests ist die zugrunde liegende Formel angegeben. In der ersten Zeile steht das zugrunde liegende Signifikanzniveau  $\alpha$  (0,05). Bei Übertragung des Befehls in ein Syntaxfenster kann diese Voreinstellung für  $\alpha$  geändert werden.

---

<sup>5</sup> Auf die speziellen Unterschiede soll hier nicht eingegangen werden. Einen kurzen Überblick geben hier z.B. Brosius & Brosius (1995, S. 430).

(\*) Indicates significant differences which are shown in the lower triangle

Mean	V325	
3,3016	PDS	
4,0810	BUENDNIS	*
4,5687	SPD	* *
5,8155	F.D.P.	* * *
6,2292	CDU-CSU	* * * *
7,2245	REPUBLIK	* * * * *

Das Testergebnis ist in dieser Kreuztabelle festgehalten. Jede signifikante Differenz zweier Mittelwerte auf dem Signifikanzniveau  $\alpha$  (hier 0,05) wird durch einen Stern gekennzeichnet. Wir sehen, daß sich die potentiellen Wählergruppen aller Parteien signifikant in ihren Mittelwerten unterscheiden. Interessanterweise entspricht die Reihenfolge der Mittelwerte von links nach rechts durchaus der Einschätzung, die man ad hoc äußern würde: Am linken Rand die PDS, dann der Reihenfolge nach B90/Grüne, SPD, F.D.P, CDU-CSU und ganz rechts die Republikaner.

### 1.5 Beispiel aus der Forschungspraxis: Bildung und normative Geschlechterrollenorientierung (Test auf Linearität)

Im ALLBUS 1996 wurde unter anderem auch die Schulbildung der Befragten sowie deren normative Geschlechterrollenorientierung (NGRO) mit einer Likertskala erhoben. Wir unterscheiden vier Bildungsgruppen:

1 = ohne Abschluß

2 = Volks-/Hauptschulabschluß

3 = Mittlere Reife/Realschulabschluß

4 = Abitur/Fachhochschulreife/Hochschulabschluß

Wir untersuchen im folgenden die Hypothese, daß sich die NGRO je nach Schulbildung des Befragten unterscheidet. Auch die vermutete Richtung der Beziehung läßt sich spezifizieren: Befragte mit einer höheren Schulbildung sollten eine liberalere NGRO (einen höheren Wert auf der Likertskala) aufweisen.<sup>6</sup> Wir untersuchen diese Frage zunächst mit einer einfaktoriellen Va-

---

<sup>6</sup> Allerdings kann eine solche Richtungshypothese nicht unmittelbar mit dem Grundmodell der Varianzanalyse getestet werden, da der unabhängige Faktor (hier Bildung) als nominalskalierte Variable behandelt wird, selbst wenn tatsächlich ein höheres Meßniveau (hier Ordinalskalenniveau) realisiert ist. Für den später folgenden Linearitätstest muß hingegen sogar Intervallniveau für die Skalierung der Bildungsvariablen angenommen werden.

riananalyse und dem unter 1.4 bereits vorgestellten Tukey-HSD-Test auf multiple Mittelwertunterschiede.

Wir erhalten die folgenden Ergebnisse (der Output ist gekürzt wiedergegeben):

```

- - - - - O N E W A Y - - - - -

Variable  NGRO
By Variable  BILDUNG

Analysis of Variance

Source          D.F.      Sum of Squares      Mean Squares      F Ratio      F Prob.
Between Groups      3      5705,9566      1901,9855      122,8268      ,0000
Within Groups     2980     46145,6168      15,4851
Total              2983     51851,5734
```

Die einfaktorielle Varianzanalyse zeigt, daß die Mittelwerte der NGRO nicht in allen Bildungsgruppen gleich sind.  $H_0$  kann mit einer Irrtumswahrscheinlichkeit von angenähert 0,000 zurückgewiesen werden. Die durch die Gruppenzugehörigkeit (=Bildung) erklärte Varianz der NGRO beträgt 11,0% (5705,9566 / 51851,5734).

Der Levene-Test auf Homogenität der Varianzen fällt günstig aus. Die Hypothese, alle Varianzen innerhalb der Gruppen sind gleich ( $H_0$ ), kann nicht zurückgewiesen werden, da die Irrtumswahrscheinlichkeit in diesem Fall 62,1% beträgt.

```

Levene Test for Homogeneity of Variances

Statistic    df1    df2      2-tail Sig.
,5907         3     2980      ,621
```

Die Voraussetzung zur Anwendung der Varianzanalyse sind somit erfüllt.

Multiple Range Tests: Tukey-HSD test with significance level ,050

```

The difference between two means is significant if
MEAN(J)-MEAN(I) >= 2,7825 * RANGE * SQRT(1/N(I) + 1/N(J))
with the following value(s) for RANGE: 3,65
```

(\*) Indicates significant differences which are shown in the lower triangle

```

o V R A
h o e b
n l a i
e k l t
s s u
A - c r
b / h /
s H u F

Mean      BILDUNG
2,2500    ohne Abs
4,8217    Volks-/H *
7,0000    Realschu * *
7,8261    Abitur/F * * *
```

Das Ergebnis des Tukey-HSD-Tests bestätigt auch unsere Vermutung bezüglich der Richtung des Zusammenhanges. Tatsächlich sind die Mittelwertunterschiede zwischen allen benachbarten

Gruppen auf dem 5%-Niveau signifikant und die NGRO wird mit zunehmendem Bildungsniveau liberaler.

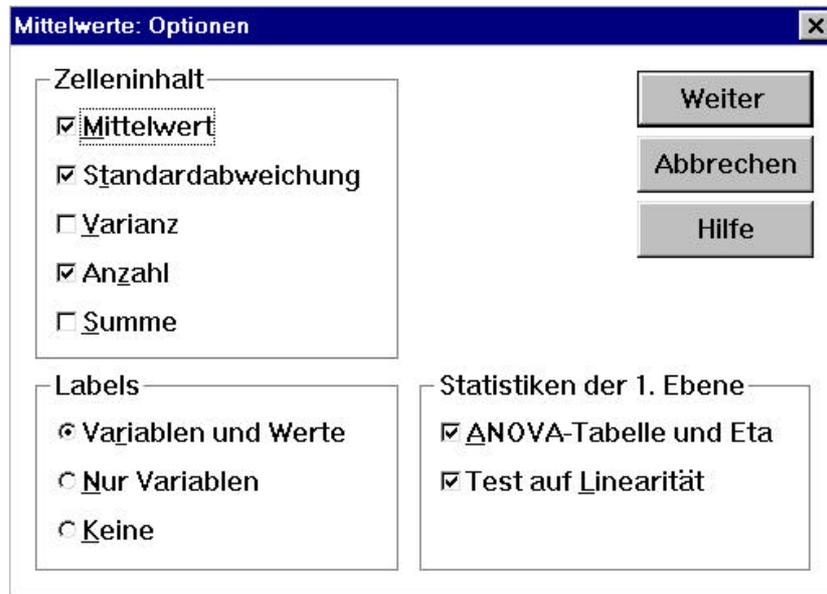
Es wäre möglich, daß die NGRO mit steigender Bildung nicht nur zunimmt, sondern daß dieser Zusammenhang auch linear ist. Dazu müssen wir allerdings Intervallskalenniveau bei der Messung der Bildungsvariablen unterstellen. Für den Linearitätstest kann auf die SPSS-Prozedur Mittelwerte (Means) zurückgegriffen werden (Statistik → Mittelwertvergleiche → Mittelwerte):<sup>7</sup>



Diese Prozedur berechnet für jede Ausprägung der unabhängigen Variablen die Mittelwerte der abhängigen Variablen (genauso wie die Varianzanalyse). Unter dem Button Optionen ... kann ein zusätzlicher Linearitätstest und eine ANOVA-Tabelle und Eta ausgewählt werden:

---

<sup>7</sup> Alternativ ist auch ein Linearitätstest mit der Prozedur ONEWAY (Mittelwerte → einfaktorische ANOVA) möglich. Bei dieser Prozedur muß unter dem Button Kontraste das Feld Polynomial angeklickt werden. Neben der Linearität können hier auch Zusammenhänge höherer Ordnung (quadratisch, kubisch etc.) getestet werden.



Wir erhalten den folgenden Output:

- - Description of Subpopulations - -

Summaries of        NGRO  
By levels of        BILDUNG

Variable	Value	Label	Mean	Std Dev	Cases
For Entire Population			6,1149	4,1692	2984
BILDUNG	1,00	ohne Abschluß	2,2500	3,6615	52
BILDUNG	2,00	Volks-/Hauptschule	4,8217	3,9805	1352
BILDUNG	3,00	Realschule	7,0000	3,9476	913
BILDUNG	4,00	Abitur/FHR	7,8261	3,8445	667

Total Cases = 3518  
Missing Cases = 534 or 15,2 Pct

Zunächst wird für jede Ausprägung der Variablen Bildung der Mittelwert der NGRO (Mean) sowie die gruppenspezifische Standardabweichung (Std Dev) und die Fallzahl (Cases) ausgegeben. Diesen Output erhält man von Means per Voreinstellung. Als nächstes wird die angeforderte ANOVA-Tabelle ausgegeben, die als zusätzliche Information die Summe der quadratischen Abweichungen innerhalb der Gruppen (Sum of Sq) ausgibt.

- - Analysis of Variance - -

Dependent Variable NGRO  
By levels of BILDUNG

Value	Label	Mean	Std Dev	Sum of Sq	Cases
1,00	ohne Abschluß	2,2500	3,6615	683,7500	52
2,00	Volks-/Hauptschule	4,8217	3,9805	21406,0407	1352
3,00	Realschule	7,0000	3,9476	14212,0000	913
4,00	Abitur/FHR	7,8261	3,8445	9843,8261	667
Within Groups Total		6,1149	3,9351	46145,6168	2984

Der uns eigentlich interessierende Test auf Linearität steht am Ende des Outputs. Das Grundprinzip des Tests beruht auf einer Gegenüberstellung der durch die unabhängige Variable insgesamt erklärten Varianz  $\eta^2$  (Eta Squared) und der durch den linearen Anteil der Beziehung erklärten Varianz  $R^2$  (R Squared).

Source	Sum of Squares	d.f.	Mean Square	F	Sig.
Between Groups	5705,9566	3	1901,9855	122,8268	,0000
Linearity	5368,2849	1	5368,2849	346,6741	,0000
Dev. from Linearity	337,6717	2	168,8359	10,9031	,0000
R = ,3218		R Squared = ,1035			
Within Groups	46145,6168	2980	15,4851		
Eta = ,3317		Eta Squared = ,1100			

Die erklärte Varianz  $\eta^2$  ergibt sich (wie in den vorherigen Beispielen) aus dem Quotienten der quadratischen Abweichung zwischen den Gruppen (Sum of Squares Between Groups) und der gesamten Abweichung (Between Groups + Within Groups).  $R^2$  ist der quadrierte Pearsonsche Korrelationskoeffizient  $r$  (hier  $R!$ ). Der Korrelationskoeffizient  $r$  ist ein Maß für den linearen Zusammenhang zweier Variablen.  $R^2$  kann niemals größer werden als  $\eta^2$ . Je stärker  $R^2$  von  $\eta^2$  (nach unten) abweicht, desto mehr Varianz wird durch nicht-lineare Einflüsse erklärt. Die Prozedur Means gibt zwei Signifikanztests aus. In der Zeile Linearity wird der lineare Anteil der Beziehung ( $r$ ) auf Signifikanz getestet. In der Zeile Dev. From Linearity wird der restliche Anteil der Beziehung (der nicht-lineare Anteil) auf Signifikanz getestet. In unserem Fall ist sowohl der lineare Anteil als auch der nicht-lineare Anteil der Beziehung signifikant. Allerdings gehen 10,35% von insgesamt 11% erklärter Varianz auf den linearen Anteil der Beziehung zurück. Es wäre also möglich, daß bereits eine kleine Modifikation der Ausprägungen der Bildungsvariable zu einem nicht mehr signifikanten nicht-linearen Anteil der Beziehung führt.

In unserem Fall ist zu erwägen, ob der Unterschied im Bildungsniveau zwischen mittlerer Reife und Abitur nicht als geringer als der Unterschied zwischen Volks-/Hauptschulabschluß und mittlerer Reife angesehen werden kann. Wir reduzieren daher den Abstand auf der Bildungsskala (als Intervallskala interpretiert!) zwischen den beiden höchsten Bildungsabschlüssen um

die Hälfte (aus Bildung = 4 wird Bildung = 3,5) und führen einen erneuten Linearitätstest mit Means durch:

- - Analysis of Variance - -

Dependent Variable NGRO  
By levels of BILDUNG2

Value	Label	Mean	Std Dev	Sum of Sq	Cases
1,00	ohne Abschluß	2,2500	3,6615	683,7500	52
2,00	Volks-/Hauptschule	4,8217	3,9805	21406,0407	1352
3,00	Realschule	7,0000	3,9476	14212,0000	913
3,50	Abitur/FHR	7,8261	3,8445	9843,8261	667
-----					
Within Groups Total		6,1149	3,9351	46145,6168	2984

Source	Sum of Squares	d.f.	Mean Square	F	Sig.
Between Groups	5705,9566	3	1901,9855	122,8268	,0000
Linearity	5675,7477	1	5675,7477	366,5295	,0000
Dev. from Linearity	30,2089	2	15,1044	,9754	,3772
R = ,3308		R Squared = ,1095			
Within Groups	46145,6168	2980	15,4851		
Eta = ,3317		Eta Squared = ,1100			

Zu beachten ist, daß sich durch die (für eine ordinale Variable zulässige) Transformation der Bildungsvariablen die erklärte Varianz  $\eta^2$  nicht geändert hat - wohl aber die durch den linearen Anteil erklärte Varianz  $R^2$ . Die Abweichung von der Linearität ist nun nicht mehr signifikant. Natürlich handelt es sich bei unserer Vorgehensweise nicht um einen Test der Beziehung zwischen Bildung und NGRO auf Linearität im strengen Sinne, da wir die Bildungsvariable so reskaliert haben, d.h. die Abstände zwischen den Rangplätzen so verändert haben, daß die Beziehung angenähert linear wird. Eine solche Vorgehensweise ist aber oft sinnvoll, wenn man eine ordinalskalierte Variable als unabhängige Variable zusammen mit anderen intervallskalierten Variablen in ein multiples Regressionsmodell aufnehmen will.

## 2. Mehrfaktorielle Varianzanalyse

Im folgenden wenden wir uns dem Problem der Varianzzerlegung in der mehrfaktoriellen Varianzanalyse am Beispiel der zweifaktoriellen Varianzanalyse mit den Faktoren A und B zu. Im Mittelpunkt steht die Frage, unter welchen Umständen die Zuordnung auf die beiden Haupteffekte A und B und den Interaktionseffekt AB eindeutig ist. Wie wir sehen werden, ist dies nur für das sog. balancierte Design, bei dem die Gruppengröße für alle Versuchsgruppen gleich ist, möglich. Nachdem wir diesen Spezialfall, der nur im Experiment (angenähert) verwirklicht werden kann, in den Abschnitten 2.1 und 2.2 anhand unseres Beispiels erläutert haben, stellen wir in 2.3 das allgemeine Modell der mehrfaktoriellen Varianzanalyse dar, in dem eine eindeu-

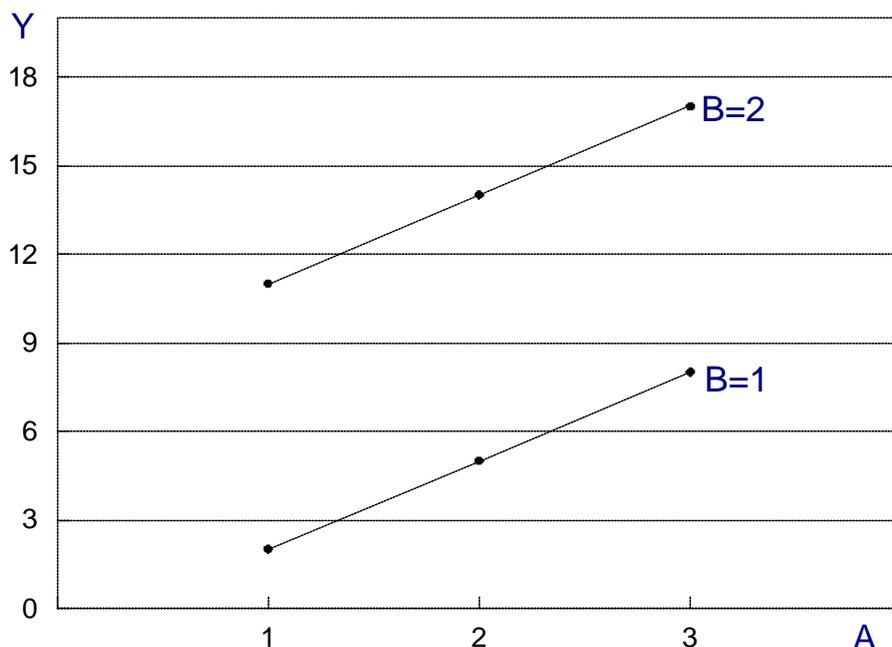
tige Zuordnung der erklärten Varianz auf die Komponenten A,B und AB nicht mehr möglich ist. Auch für solche nicht balancierten und allgemein nicht orthogonalen Designs sind aber Methoden entwickelt worden, die mit Hilfe zusätzlicher Annahmen eine eindeutige Varianzzerlegung ermöglichen. Wir stellen die experimentelle, die hierarchische und die regressionsanalytische Methode vor und illustrieren alle drei Vorgehensweisen jeweils an demselben Beispieldatensatz (2.4). Im Anschluß daran folgt ein Beispiel aus der Forschungspraxis (2.5). Abschließend wird demonstriert, wie die Varianzanalyse als Spezialfall der Regressionsanalyse mit entsprechender Codierung der Effekte als Dummy-Variablen verstanden werden kann (2.6). Damit erweist sich die Varianzanalyse als Spezialfall des allgemeinen linearen Modells.

### 2.1 *Balanciertes Design ohne Interaktionseffekte*

Wir demonstrieren zunächst den einfachsten Fall einer mehrfaktoriellen Varianzanalyse: zwei Faktoren A und B in einem balancierten Design. Ein Design ist balanciert, wenn die Zellbesetzungszahlen für alle Versuchsgruppen gleich sind, wie dies in unserem Beispiel aus 1. der Fall ist.

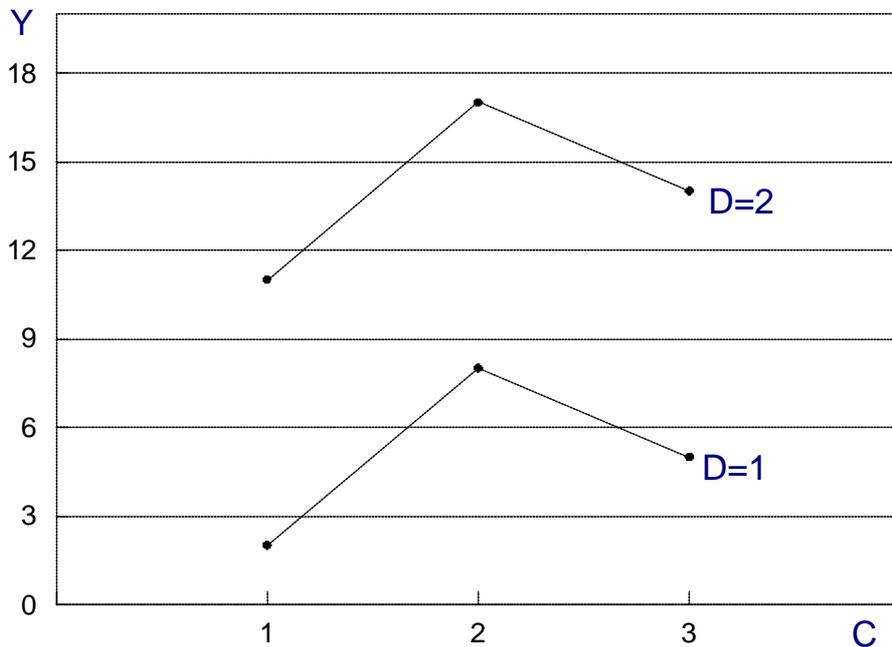
N	A=1	A=2	A=3	N <sub>B</sub>
B=1	3	3	3	9
B=2	3	3	3	9
N <sub>A</sub>	6	6	6	18

Darüber hinaus wurde das Beispiel so konstruiert, daß die abhängige Variable keinen Interaktionseffekt zwischen A und B enthält. Dies kann grafisch anhand des sog. Interaktionsdiagrammes für die Gruppenmittelwerte überprüft werden.



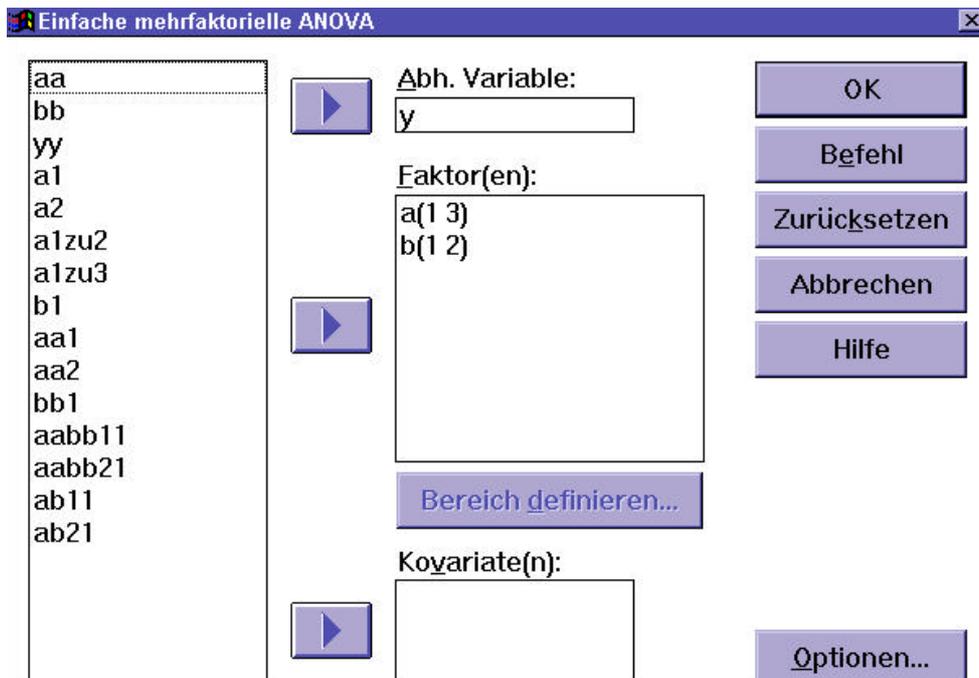
Wie man sieht, ist der Abstand zwischen den beiden Kurven für B=1 und B=2 konstant, d.h. der Effekt von B ist unabhängig von der Ausprägung von A. Zusätzlich - dies ist für das Ver-

schwinden des Interaktionseffektes nicht notwendig - erkennt man, daß die Beziehung zwischen A und Y linear ist, da beide Kurven Geraden darstellen. Ein (hypothetisches) Beispiel für eine nicht-lineare Beziehung ohne Interaktionseffekt wäre etwa:

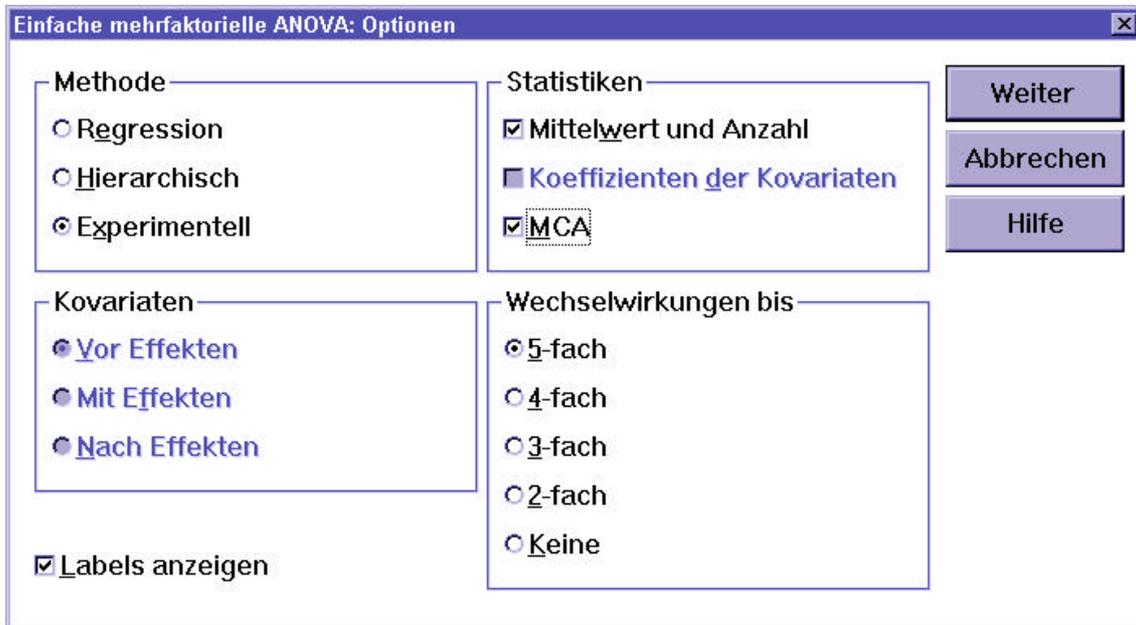


Auch hier verlaufen beide Kurven parallel, d.h. der Effekt von D ist unabhängig von C - ein Interaktionseffekt existiert nicht.

Die mehrfaktorielle Varianzanalyse (ANOVA) wird in SPSS über die Menüfelder ANOVA-Modelle → Einfach mehrfaktoriell ... aufgerufen. Es erscheint das folgende Dialogfenster:



Wie schon bei der einfaktoriellen Varianzanalyse, muß hier zunächst eine abhängige Variable (Y) spezifiziert werden. In den Kästen Faktoren können nun einer oder mehrere Faktoren (hier A und B) eingetragen werden. Wie bei ONEWAY muß auch hier für jeden Faktor der Bereich der Kategorien angegeben werden (über den Button Bereich definieren ...). Hinter dem Button Optionen verbirgt sich folgendes Fenster:



Wichtig ist hier vor allem die Beachtung des Kastens Methode. Die Voreinstellung ist hier Regression, wir wählen in unserem Beispiel Experimentell.<sup>8</sup> Als zusätzliche Statistiken wählen wir die MCA (Multiple Klassifikationsanalyse; s.u.) aus. Nach der Ausführung der ANOVA mit den getroffenen Einstellungen erhalten wir zunächst eine Tabelle mit den Gruppenmittelwerten (CELL MEANS):

```
-> anova variables y by a(1,3) b(1,2) / method experim / statistics mca mean.
```

```
* * * C E L L M E A N S * * *
```

```
      Y  
    by A  
      B
```

```
Total Population
```

```
      9,50  
    (   18)
```

Der Gesamtmittelwert von Y in der Stichprobe ist 9,50. Es folgen die Gruppenmittelwerte von Y für die Faktoren A und B getrennt:

---

<sup>8</sup> Von der gewählten Methode hängt es ab, wie die Varianz zerlegt wird, sofern es sich nicht um ein balanciertes Design handelt, für das die Varianzzerlegung eindeutig und damit unabhängig von der gewählten Methode ist (vgl. 2.3).

A		
1	2	3
6,50	9,50	12,50
( 6)	( 6)	( 6)
B		
1	2	
5,00	14,00	
( 9)	( 9)	

Die dritte Tabelle enthält die Gruppenmittelwerte von Y für jede Kombination von A und B:

		B	
		1	2
A	1	2,00	11,00
		( 3)	( 3)
	2	5,00	14,00
		( 3)	( 3)
	3	8,00	17,00
		( 3)	( 3)

Es folgt die eigentliche Varianzanalyse:

\*\*\* ANALYSIS OF VARIANCE \*\*\*

Y  
by A  
B

EXPERIMENTAL sums of squares  
Covariates entered FIRST

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig of F
Main Effects	472,500	3	157,500	157,500	,000
A	108,000	2	54,000	54,000	,000
B	364,500	1	364,500	364,500	,000
2-Way Interactions	,000	2	,000	,000	1,00
A B	,000	2	,000	,000	1,00
Explained	472,500	5	94,500	94,500	,000
Residual	12,000	12	1,000		
Total	484,500	17	28,500		

18 cases were processed.  
0 cases (,0 pct) were missing.

Zunächst wird die gesamte Erklärkraft der beiden Faktoren A und B gemeinsam erfaßt. Dies geschieht dadurch, daß die "Supervariable" A\*B mit sechs Ausprägungen (entsprechend den sechs Versuchsgruppen) gebildet wird. Gemäß der einfaktoriellen Varianzanalyse erhält man dann die eindeutige Zerlegung in durch A und B gemeinsam erklärte Varianz  $SSQ_{\text{Between}(A*B)}$  und Fehlervarianz  $SSQ_{\text{Within}(A*B)}$ . Im Output finden wird  $SSQ_{\text{Between}(A*B)}$  als  $SSQ_{\text{Explained}}$  und  $SSQ_{\text{Within}(A*B)}$  als  $SSQ_{\text{Residual}}$ . Die Gesamtsumme der Abweichungsquadrate wird weiterhin mit  $SSQ_{\text{Total}}$  bezeichnet und beträgt, da es sich um die gleiche abhängige Variable wie in 1.2.1

und 1.2.2 handelt, weiterhin 484.5. A und B erklären also insgesamt 472.5 von 484.5, d.h. 97,5% der Gesamtvarianz.

Aus 1.2.1 und 1.2.2 übernehmen wir folgende Werte für die Erklärungskraft von A und B als Einzelfaktoren:

$$\begin{array}{rcl} \text{SSQ}_{\text{Between}}(\text{A}) & = & 108.00 \quad 22,2\% \\ \text{SSQ}_{\text{Between}}(\text{B}) & = & \frac{364.50}{472.50} \quad \frac{75,3\%}{97,5\%} \end{array}$$

Die Werte entsprechen exakt den Haupteffekten (*main effects*) der Faktoren A und B in der zweifaktoriellen Varianzanalyse. Der zweiseitige Interaktionseffekt (*2-way Interactions*) von A und B ist gleich Null:  $\text{SSQ}_{\text{Between}}(\text{AB}) = 0$ . Es liegt also eine eindeutige Zerlegung der insgesamt erklärten Varianz  $\text{SSQ}_{\text{Between}}(\text{A*B})$  in die beiden Komponenten  $\text{SSQ}_{\text{Between}}(\text{A})$  und  $\text{SSQ}_{\text{Between}}(\text{B})$  vor. Außerdem wird bestätigt, was wir bereits aufgrund unserer inhaltlichen Überlegungen wissen: es gibt keinen Interaktionseffekt in der Wirkung von A und B auf Y.

Auch in der mehrfaktoriellen Varianzanalyse können die einzelnen Varianzbestandteile auf Signifikanz geprüft werden. Dazu wird wiederum der F-Test verwendet, wobei als Referenzgröße im Nenner der F-Statistik nun die Residualvarianz des mehrfaktoriellen Modells erscheint. Daher können sich Abweichungen von den entsprechenden Tests im Modell der einfaktoriellen Varianzanalyse ergeben (vgl. die abschließende Bemerkung in 1.3). Dies sei am Beispiel des Tests für den Haupteffekt von A, nämlich  $\text{SSQ}_{\text{Between}}(\text{A})$ , erläutert. Der Wert für  $\text{SSQ}_{\text{Between}}(\text{A})$  bleibt im Vergleich zur einfaktoriellen Varianzanalyse unverändert, allerdings ist  $\text{SSQ}_{\text{Residual}}$  deutlich kleiner als  $\text{SSQ}_{\text{Within}}(\text{A})$  in der einfaktoriellen Analyse, weil nun der Anteil von B, der in der einfaktoriellen Varianzanalyse zur Fehlervarianz geschlagen wurde, in der mehrfaktoriellen Varianzanalyse - also bei Kontrolle von B - als eigener Haupteffekt  $\text{SSQ}_{\text{Between}}(\text{B})$  erscheint. Damit ergibt sich als F-Test für den Haupteffekt von A:

$$F_{(2,12)} = \frac{\frac{\text{SSQ}_{\text{between}}(\text{A})}{\text{df}_A}}{\frac{\text{SSQ}_{\text{residual}}}{\text{df}_{\text{residual}}}} = \frac{\frac{108,0}{2}}{\frac{12,0}{12}} = \frac{54,0}{1,0} = 54,0$$

Damit ist der Effekt von A hochsignifikant (*Sig of F = 0,000*).

Im Zusammenhang mit experimentellen Designs wird diese Möglichkeit zur Verstärkung eines Signifikanztests durch Kontrolle von Fehlervarianz in sog. Block-Designs ausgenutzt. Dazu werden neben der eigentlich interessierenden Wirkung des experimentellen Stimulus (hier Faktor A) Blöcke gebildet (hier Faktor B), die einen Teil der Fehlervarianz binden und so die Chance erhöhen, den Einfluß des experimentellen Stimulus als signifikant nachzuweisen. Möchte etwa ein Versandhaus die Wirkung zweier unterschiedlicher Katalogformen (Faktor A) auf die Bestellungen (Y) testen, so wird sich das Experiment in der Regel auf eine große Zahl von Gebieten (Faktor B) mit unterschiedlicher Umsatzstärke (Z) als Blöcken erstrecken. Je

heterogener diese Blöcke sind, d.h. um so stärker sie sich in ihrer Kaufkraft unterscheiden, um so mehr erhöht sich die Stärke des Tests des Haupteffektes von A bei Kontrolle von B. Verfügt man allerdings, etwa auf Grundlage einer Kundendatei, über direkte Messungen des bisherigen Bestellvolumens der Kunden (Z), so wäre eine Kovarianzanalyse mit Z als Kontrollvariablen vorzuziehen.

Falls in einer mehrfaktoriellen Varianzanalyse keine signifikanten Interaktionseffekte auftreten, empfiehlt sich eine multiple Klassifikationsanalyse (MCA), um die Haupteffekte der einzelnen Faktoren bei Kontrolle aller anderen Faktoren darzustellen.

```
*** MULTIPLE CLASSIFICATION ANALYSIS ***
      Y
    by  A
      B

Grand Mean =      9,50

Variable + Category      N      Unadjusted      Adjusted for
                        Dev'n  Eta      Independents
                        Dev'n  Beta

A
  1          6      -3,00      -3,00
  2          6         ,00         ,00
  3          6       3,00       3,00
                        ,47          ,47

B
  1          9      -4,50      -4,50
  2          9       4,50       4,50
                        ,87          ,87

Multiple R Squared          ,975
Multiple R                   ,988
```

Der Output enthält zunächst Informationen über die Haupteffekte der Faktoren ohne Kontrolle (unadjusted) der übrigen Faktoren, also Angaben, die in anderer Form bereits in der einfaktoriellen Varianzanalyse vorliegen. So geben die nicht adjustierten Abweichungen gerade die Mittelwerte für die einzelnen Ausprägungen eines Faktors an und zwar als Abweichung vom Gesamtmittelwert. So entsprechen im Falle von Faktor A die Abweichungen -3, 0, 3 den Mittelwerten 6.5, 9.5 und 12.5.  $\eta$  (Eta) ist ein nominales Bestimmtheitsmaß mit einer varianzanalytischen Interpretation analog zu  $r$ , das aber im Gegensatz zu  $r$  auch den nicht linearen Anteil der Beziehung mißt (vgl. 1.5). Der Wert von .47 für  $\eta$  besagt, daß Faktor A bivariat genau  $.47^2$  der Varianz von Y erklärt (erklärter Varianzanteil 22,3%; siehe 1.2.2).

Interessanter als diese nicht adjustierten, d.h. auf dem bivariaten Modell der einfaktoriellen Varianzanalyse basierenden Informationen, sind die zugehörigen adjustierten Werte, die sich durch Kontrolle der Haupteffekte aller anderen an der Varianzanalyse beteiligten Faktoren ergeben. Die entsprechenden Werte für A ändern sich in unserem Beispiel nicht, da A und B unkorreliert sind, die Kontrolle von B also weder die Mittelwerte in den Kategorien von A, noch die Erklärkraft von A verändert. Im allgemeinen ergeben sich aber veränderte Werte, die dann wie die partiellen Koeffizienten der Regressionsanalyse interpretiert werden können (vgl. die

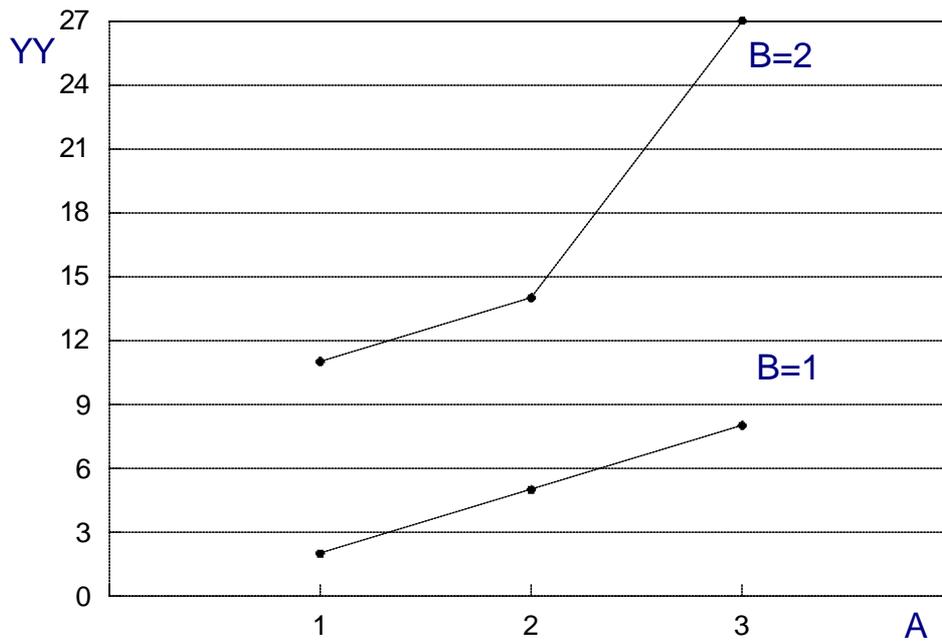
Beispielanalyse in 2.6). Die zugehörige (im allgemeinen numerisch veränderte) partielle Korrelation wird als  $\beta$  (Beta) ausgegeben. Das multiple Bestimmtheitsmaß  $R^2$  gibt den Anteil der insgesamt von A und B erklärten Varianz an, also 97,5% - auch dies in direkter Analogie zur multiplen Regressionsanalyse.

## 2.2 *Balanciertes Design mit Interaktionseffekt*

Die folgende zweifaktorielle Varianzanalyse mit den Faktoren A und B verwendet YY als abhängige Variable.

Beispieldaten		
A	B	YY
1	1	1
1	1	2
1	1	3
2	1	4
2	1	5
2	1	6
3	1	7
3	1	8
3	1	9
1	2	10
1	2	11
1	2	12
2	2	13
2	2	14
2	2	15
3	2	26
3	2	27
3	2	28

Da sich die Gruppengröße somit nicht verändert, liegt wieder ein balanciertes Design vor. Wir erwarten also wieder eine eindeutige Zerlegung der erklärten Varianz  $SSQ_B(A*B)$  in ihre Komponenten. Allerdings bedeutet die Verwendung von YY als abhängiger Variablen, daß nun ein Interaktionseffekt zwischen A und B existiert. Dies kann graphisch anhand des Interaktionsdiagrammes überprüft werden:



Im Unterschied zu Y wurden bei YY die Werte für die sechste Gruppe (A=3, B=2) jeweils um 10 erhöht, so daß der Mittelwert von 17 auf 27 ansteigt. Das bedeutet, daß die beiden Kurven für B=1 und B=2 nicht mehr parallel verlaufen. Die Wirkung von Faktor A auf YY hängt also von der Ausprägung des Faktors B ab.

Zunächst zeigen wir die Ergebnisse der beiden einfaktoriellen Varianzanalysen mit YY als abhängiger Variable im Überblick:

<b>Oneway: A ® YY</b>		erklärte Varianz
SSQ <sub>Between</sub> (A)	388,0	32,8%
SSQ <sub>Within</sub> (A)	796,0	
SSQ <sub>Total</sub> (A)	1184,0	

<b>Oneway: B ® YY</b>		erklärte Varianz
SSQ <sub>Between</sub> (B)	684,5	57,8%
SSQ <sub>Within</sub> (B)	499,5	
SSQ <sub>Total</sub> (B)	1184,0	

Die mehrfaktorielle Varianzanalyse (die Ausführung in SPSS verläuft analog zu 2.1) ergibt:

\* \* \* A N A L Y S I S O F V A R I A N C E \* \* \*

YY  
by A  
B

EXPERIMENTAL sums of squares  
Covariates entered FIRST

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig of F
Main Effects	1072,500	3	357,500	357,500	,000
A	388,000	2	194,000	194,000	,000
B	684,500	1	684,500	684,500	,000
2-Way Interactions	100,000	2	50,000	50,000	,000
A B	100,000	2	50,000	50,000	,000
Explained	1172,500	5	234,500	234,500	,000
Residual	12,000	12	1,000		
Total	1184,500	17	69,676		

18 cases were processed.  
0 cases (,0 pct) were missing.

Zunächst ist festzuhalten, daß im mehrfaktoriellen Fall die Haupteffekte von A und B genauso groß sind, wie in den oben dargestellten einfaktoriellen Varianzanalysen (dort als  $SSQ_{\text{Between}(A)}$  und  $SSQ_{\text{Between}(B)}$  ausgegeben und hier als Main Effects A bzw. B bezeichnet.). Zusätzlich erhalten wir nun jedoch mit  $SSQ_B(AB) = 100,0$  (2-Way Interactions) einen signifikanten Interaktionseffekt. Die von A und B gemeinsam erklärte Varianz  $SSQ_B(A*B)$  (Explained) ergibt sich aus der Summe der beiden Haupteffekte A und B und des Interaktionseffektes AB ( $1172,5 = 388,0 + 684,5 + 100,0$ ) und beträgt 99,0%.

Durch den Interaktionseffekt AB ist die Gesamterklärungskraft von A und B in der zweifaktoriellen Varianzanalyse also leicht höher als die Summe der Erklärungskraft der Einzeleffekte A und B aus den einfaktoriellen Varianzanalyse.

Wegen des Vorliegens eines signifikanten Interaktionseffekts ( $F = 50,0$ ; Sig of  $F = 0,00$ ) verzichten wir auf eine multiple Klassifikationsanalyse.

Entscheidend für die beiden hier besprochenen mehrfaktoriellen Varianzanalysen ist das Vorliegen eines balancierten Designs, das die eindeutige Zerlegbarkeit der erklärten Varianz in Haupt- und Interaktionseffekte sicherstellt. Um die Komplikationen, die sich bei nicht balancierten (und noch allgemeiner: nicht orthogonalen) Designs ergeben, besser verstehen zu können, beschäftigen wir uns zunächst mit einer formalen Darstellung der mehrfaktoriellen Varianzanalyse.

### 2.3 Die Zuordnung der Varianzkomponenten im nicht orthogonalen Design

Wir beschränken uns im folgenden auf die zweifaktorielle Varianzanalyse. In Erweiterung von 1.3 gilt die Grundgleichung:

$$y_{ijz} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijz} \quad (2.3.1)$$

Neben den Haupteffekten  $\alpha_i$  und  $\beta_j$  muß hier auch der Interaktionseffekt  $\alpha\beta_{ij}$  berücksichtigt werden.  $\varepsilon_{ijz}$  ist weiterhin ein Zufallsfehler. Bezogen auf die Gruppenmittelwerte

$y_{i\bullet\bullet}$  (Mittelwert für A = i; summiert über j und z)

$y_{\bullet j\bullet}$  (Mittelwert für B = j; summiert über i und z)

$y_{ij\bullet}$  (Mittelwert für AB = ij; summiert über z)

ergibt sich:

$$y_{ijz} = \underbrace{y_{\bullet\bullet\bullet}}_{(\mu)} + \underbrace{(y_{i\bullet\bullet} - y_{\bullet\bullet\bullet})}_{(\alpha_i)} + \underbrace{(y_{\bullet j\bullet} - y_{\bullet\bullet\bullet})}_{(\beta_j)} + \underbrace{(y_{ij\bullet} - y_{i\bullet\bullet} - y_{\bullet j\bullet} + y_{\bullet\bullet\bullet})}_{(\alpha\beta_{ij})} + \underbrace{(y_{ijz} - y_{ij\bullet})}_{(\varepsilon_{ijz})} \quad (2.3.2)$$

Nach dem Fundamentaltheorem der Varianzanalyse gilt nach wie vor

$$SSQ_{\text{total}}(A*B) = SSQ_{\text{Between}}(A*B) + SSQ_{\text{within}}(A*B) \quad (2.3.3)$$

mit den zugehörigen Freiheitsgraden

SSQ	df	df im Beispiel (I = 3, J = 1, N = 18)
between	I · J - 1	5
within	N - I · J	12
total	N - 1	17

Angestrebt wird eine eindeutige Zerlegung von  $SSQ_{\text{Between}}(A*B)$  in

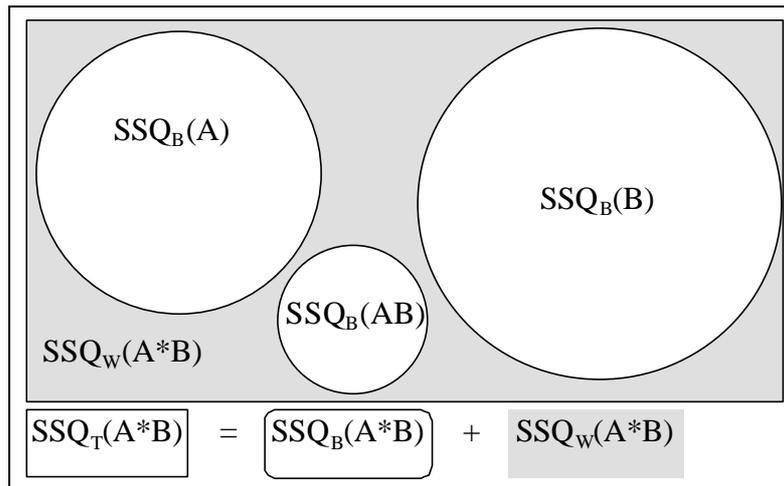
		df	df im Beispiel
Haupteffekt von A	$SSQ_{\text{Between}}(A)$	I - 1	2
Haupteffekt von B	$SSQ_{\text{Between}}(B)$	J - 1	1
Interaktionseffekt AB	$SSQ_{\text{Between}}(AB)$	(I - 1) · (J - 1)	2
insgesamt von A*B erklärte Varianz	$SSQ_{\text{Between}}(A*B)$	I · J - 1	5

Wie in 2.1 und 2.2 demonstriert, ist eine solche Zerlegung (mit oder ohne Interaktionseffekt) im Falle eines balancierten Designs möglich.

Ein balanciertes Design ist durch identische Größen der Versuchsgruppen charakterisiert, in unserem Beispiel:

N	A=1	A=2	A=3	
B=1	3	3	3	9
B=2	3	3	3	9
	6	6	6	18

Graphisch kann die Eindeutigkeit der Zerlegung von  $SSQ_{\text{Between}}(A*B)$  wie folgt veranschaulicht werden:

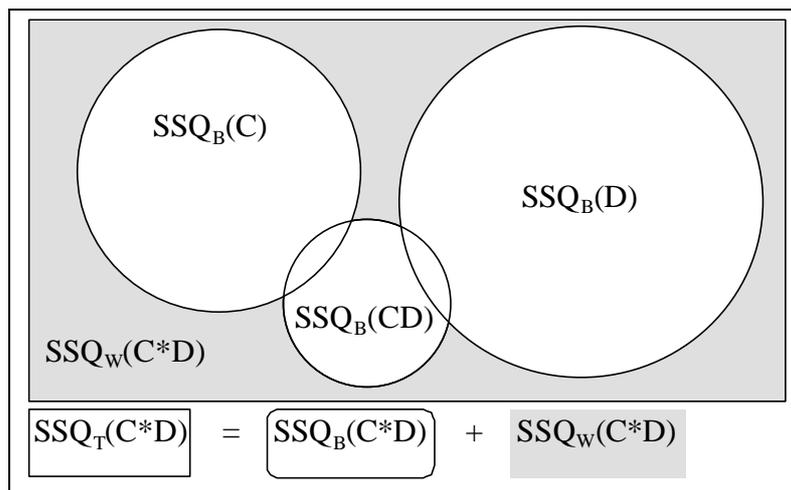


In dieser Grafik sind die einzelnen Komponenten von  $SSQ_{\text{Between}}(A*B)$  voneinander getrennt; sie überlappen nicht.

Im balancierten Design gilt:  $\text{Korr}(A,B) = 0$ . Das balancierte Design ist also ein Spezialfall des orthogonalen Designs. Auch das folgende nicht balancierte Design ist orthogonal:

N		C=1	C=2	C=3	
	D=1	3	6	9	18
	D=2	6	12	18	36
		6	6	6	54

Es gilt nämlich ebenfalls:  $\text{Korr}(C,D) = 0$ . Die Zerlegung von  $SSQ(C*D)$  ist aber nicht mehr eindeutig, obwohl wegen  $\text{Korr}(C,D) = 0$  die Anteile  $SSQ_{\text{Between}}(C)$  und  $SSQ_{\text{Between}}(D)$  weiterhin nicht überlappen:



Die Schnittmengen von  $SSQ_{\text{Between}}(C)$  und  $SSQ_{\text{Between}}(D)$  mit dem Interaktionseffekt  $SSQ_{\text{Between}}(CD)$  sind nicht ohne weiteres eindeutig zuzuordnen; die Zerlegung von  $SSQ_{\text{Between}}(C*D)$  ist somit nicht mehr eindeutig.

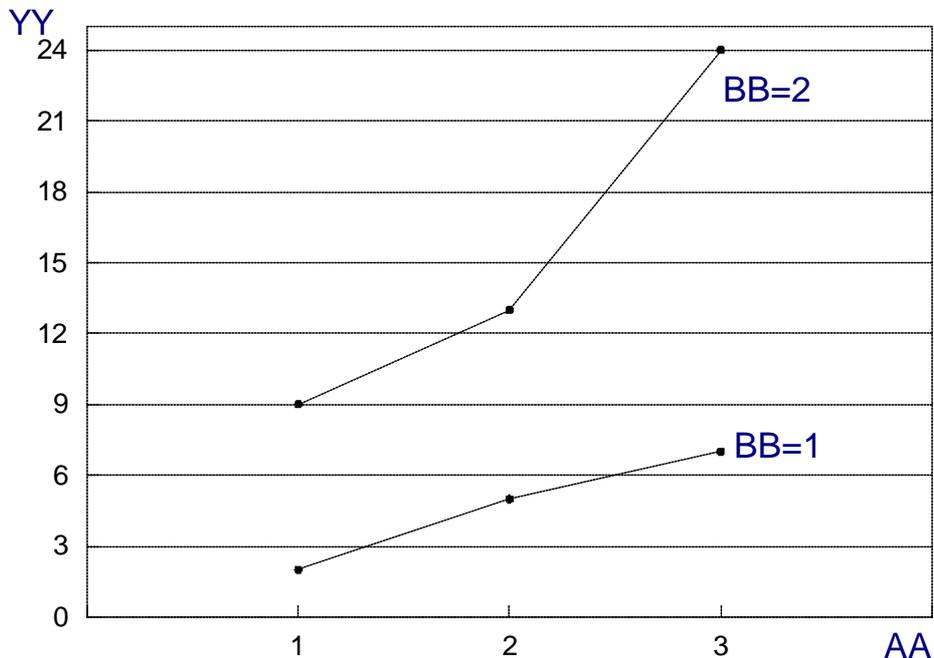
Um den allgemeinen Fall eines nicht-orthogonalen Designs zu analysieren, erweitern wir unseren Beispieldatensatz:

AA	BB	YY
1	1	1
1	1	2
1	1	3
2	1	4
2	1	5
3	1	6
3	1	7
1	2	8
1	2	9
1	2	10
2	2	11
2	2	12
2	2	13
2	2	14
3	2	15
3	2	26
3	2	27
3	2	28

Für AA und BB ergeben sich folgende Zellbesetzungszahlen:

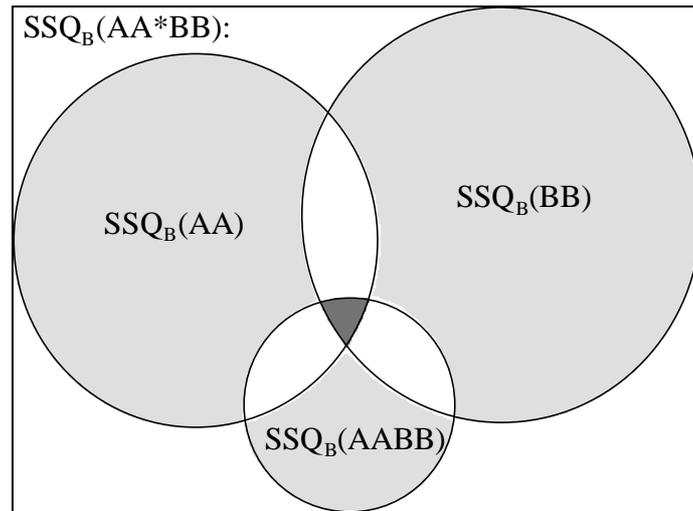
N	AA=1	AA=2	AA=3	
BB=1	3	2	2	7
BB=2	3	4	4	11
	6	6	6	18

Damit gilt  $\text{Korr}(AA, BB) \neq 0$ , d.h., das Design ist nicht orthogonal und damit auch nicht balanciert. Darüberhinaus enthält YY weiterhin einen Interaktionseffekt, wie aus dem folgenden Interaktionsdiagramm abzulesen ist:



Korrelation:  $\text{PHI}_{AA\ BB} = 0.16$

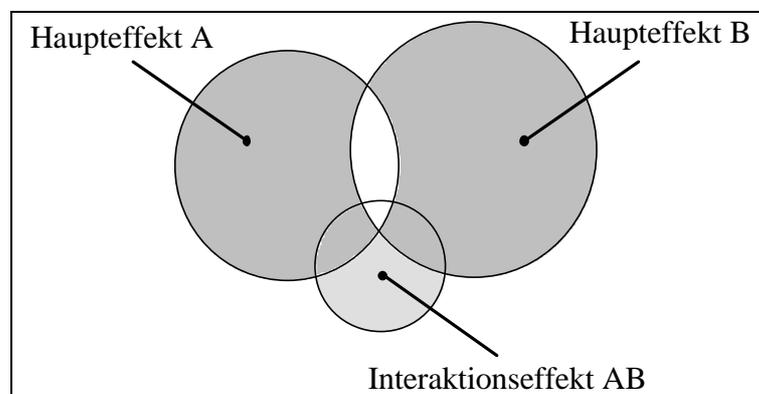
Die Korrelation zwischen AA und BB impliziert, daß nun auch  $\text{SSQ}_{\text{Between}}(AA)$  und  $\text{SSQ}_{\text{Between}}(BB)$  überlappen:



Nur die hellgrauen Flächen können eindeutig einer Komponente zugeordnet werden. Die Überlappungsbereiche sind mehrdeutig zuordnbar - der dunkelgraue dreiseitige Überlappungsbereich kann sogar jeder der drei Komponenten zugeordnet werden.

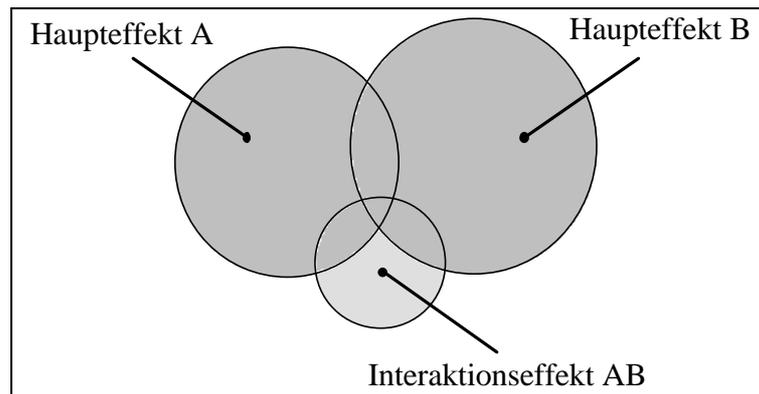
Das allgemeine Zuordnungsproblem lässt sich nur konventionell lösen, indem zusätzliche Regeln aufgestellt werden, die natürlich vor dem Hintergrund des speziellen Forschungsproblems gerechtfertigt werden müssen. Wir unterscheiden die drei folgenden, in der ANOVA-Prozedur realisierten Optionen:

1) **Experimentelles Verfahren** (method experimental):



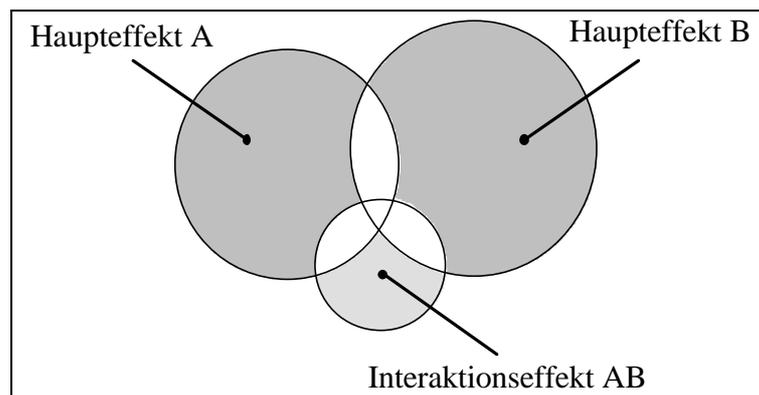
Die Schnittmenge zwischen den Haupteffekten wird nicht zugeordnet. Die Schnittmengen der Haupteffekte mit dem Interaktionseffekt werden den Haupteffekten zugerechnet.

2) **Hierarchisches Verfahren** (method hierarchical):



Es wird eine Hierarchie zwischen den Faktoren etabliert, indem die Schnittmenge zwischen den Haupteffekten dem zuerst in der Anweisung stehenden Faktor zugerechnet wird. In diesem Fall wurde von uns A vor B festgesetzt.

3) **Regressionsanalytisches Verfahren** (method regression; SPSS-Voreinstellung):



Keiner der Überlappungsbereiche wird zugeordnet; übrig bleiben die partiellen Effekte der Regressionsanalyse (vgl. 2.5).

Alle drei Verfahren stimmen darin überein, daß der Varianzanteil des Interaktionseffekts AB residual bestimmt wird, d.h. nur die Anteile, die eindeutig weder von A noch von B erklärt werden können, werden AB zugeordnet. Bestimmt man auch die Haupteffekte strikt residual, so erhält man die regressionsanalytische Methode. In den beiden anderen Verfahren gilt eine implizite Hierarchie, welche die Haupteffekte über die Interaktionseffekte stellt. Problematisch ist aber weiterhin die Zuordnung des Überlappungsbereiches von A und B. Hier legt das hierarchische Verfahren aufgrund einer (inhaltlich zu rechtfertigenden!) Vorentscheidung eine eindeutige Hierarchie fest. Im experimentellen Verfahren wird dieser Überlappungsbereich bei der direkten Zuweisung auf die einzelnen Haupteffekte ausgespart, erscheint aber in der Summe der beiden Haupteffekte (in der Rubrik `main effects`; vgl. 2.4).

## 2.4 Der allgemeine Fall eines nicht-orthogonalen Designs

Bevor wir die oben angesprochenen Methoden der Varianzzerlegung an unserem Beispiel durchrechnen, ermitteln wir zunächst die jeweilige Erklärkraft von AA und BB in Bezug auf YY mit Hilfe von einfaktoriellen Varianzanalysen. Wir erhalten die folgenden Ergebnisse:

Oneway: AA ® YY		erklärte Varianz
SSQ <sub>B</sub> (AA)	497,3	42,0%
SSQ <sub>W</sub> (AA)	687,2	
SSQ <sub>T</sub> (AA)	1184,5	

Oneway: BB ® YY		erklärte Varianz
SSQ <sub>B</sub> (BB)	588,3	49,7%
SSQ <sub>W</sub> (BB)	596,2	
SSQ <sub>T</sub> (BB)	1184,5	

### 2.4.1 Das experimentelle Verfahren

Der Aufruf der mehrfaktoriellen Varianzanalyse mit SPSS erfolgt analog zu 2.1. Unter Optionen → Methode wird die jeweilige Methode ausgewählt (hier experimental). Zu beachten ist, daß wir (wegen des Interaktionseffektes) auf die Ausgabe der MCA verzichten.

-> anova variables yy by aa(1,3) bb(1,2) / method experim.

\* \* \* A N A L Y S I S O F V A R I A N C E \* \* \*

YY  
by AA  
BB

EXPERIMENTAL sums of squares  
Covariates entered FIRST

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig of F
Main Effects	972,593	3	324,198	32,420	,000
AA	384,275	2	192,138	19,214	,000
BB	475,260	1	475,260	47,526	,000
2-Way Interactions	91,907	2	45,953	4,595	,033
AA BB	91,907	2	45,953	4,595	,033
Explained	1064,500	5	212,900	21,290	,000
Residual	120,000	12	10,000		
Total	1184,500	17	69,676		

18 cases were processed.  
0 cases (,0 pct) were missing.

## 2.4.2 Das hierarchische Verfahren

-> anova variables yy by aa(1,3) bb(1,2) / method hierarchical.

\* \* \* A N A L Y S I S O F V A R I A N C E \* \* \*

YY  
by AA  
BB

HIERARCHICAL sums of squares  
Covariates entered FIRST

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig of F
Main Effects	972,593	3	324,198	32,420	,000
AA	497,333	2	248,667	24,867	,000
BB	475,260	1	475,260	47,526	,000
2-Way Interactions	91,907	2	45,953	4,595	,033
AA BB	91,907	2	45,953	4,595	,033
Explained	1064,500	5	212,900	21,290	,000
Residual	120,000	12	10,000		
Total	1184,500	17	69,676		

18 cases were processed.  
0 cases (,0 pct) were missing.

### 2.4.3 Das regressionsanalytische Verfahren

-> anova variables yy by aa(1,3) bb(1,2) / method unique.

\* \* \* A N A L Y S I S O F V A R I A N C E \* \* \*

YY  
by AA  
BB

UNIQUE sums of squares  
All effects entered simultaneously

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig of F
Main Effects	877,804	3	292,601	29,260	,000
AA	277,560	2	138,780	13,878	,001
BB	487,500	1	487,500	48,750	,000
2-Way Interactions	91,907	2	45,953	4,595	,033
AA BB	91,907	2	45,953	4,595	,033
Explained	1064,500	5	212,900	21,290	,000
Residual	120,000	12	10,000		
Total	1184,500	17	69,676		

18 cases were processed.  
0 cases (,0 pct) were missing.

Zur vertiefenden Interpretation der jeweiligen Varianzzerlegungen arbeiten wir zunächst die Parallele zwischen Varianzanalyse und dem allgemeinen linearen Modell mit den entsprechenden Dummy-Variablen heraus.

### 2.5 Die mehrfaktorielle Varianzanalyse als Spezialfall des allgemeinen linearen Modells

In diesem Abschnitt demonstrieren wir, wie die Varianzzerlegung der mehrfaktoriellen Varianzanalyse als Spezialfall der linearen Regression verstanden werden kann. Dazu werden die Faktoren durch sog. Dummy-Variablen repräsentiert. Allgemein gilt, daß jede Varianzkomponente durch genau so viele Dummy-Variablen dargestellt wird, wie sie Freiheitsgrade hat. Wir verwenden im folgenden die sog. Effekt-Codierung, wobei jeweils die letzte Ausprägung eines Faktors als Referenzkategorie gewählt wird. Für den Fall der mehrfaktoriellen Varianzanalyse mit den Faktoren AA und BB definieren wir die folgenden Dummy-Variablen:

Ursprungsvariable	Dummy-Variablen		Ursprungsvariable	Dummy-Variablen
AA	AA1	AA2	BB	BB1
1	1	0	1	1
2	0	1	2	-1
3	-1	-1		

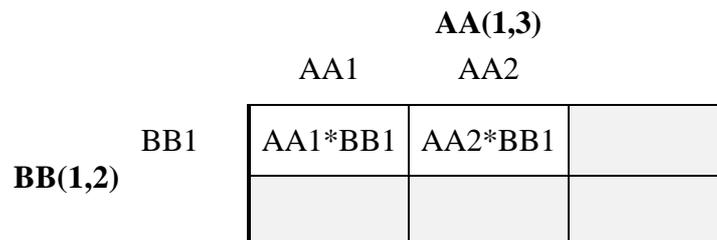
AA = 3 dient als Referenzkategorie von AA, die Dummies AA1 und AA2 repräsentieren den Haupteffekt von AA. BB = 2 dient als Referenzkategorie von BB; die Dummy-Variable BB1 repräsentiert den Haupteffekt von BB.

Durch Multiplikation der Haupteffektdummies erhält man die Dummies, die die jeweiligen Interaktionseffekte repräsentieren:

$$AABB11 := AA1 \times BB1$$

$$AABB21 := AA2 * BB1$$

Die folgende Grafik macht den Zusammenhang zwischen dem Faktor und den zugeordneten Dummies klar und gibt darüber hinaus über die mit dem jeweiligen Effekt verbundenen Freiheitsgrade Auskunft (identisch mit der Zahl der zugeordneten Dummies in der Regressionsgleichung):



Eine multiple Regression mit allen gebildeten Dummy-Variablen in SPSS liefert den folgenden Output:

```

* * * * * M U L T I P L E   R E G R E S S I O N   * * * * *

Listwise Deletion of Missing Data

Equation Number 1   Dependent Variable..   YY

Block Number  1.  Method:  Enter
  AA1      AA2      BB1      AABB11   AABB21

Variable(s) Entered on Step Number
  1..      AABB21
  2..      BB1
  3..      AA1
  4..      AABB11
  5..      AA2
    
```

Als abhängige Variable (Dependent Variable) haben wir YY angegeben, unabhängige Variablen sind unsere fünf Dummy-Variablen AA1, AA2, BB1, AABB11 und AABB21. Von besonderem Interesse ist nun die erklärte Varianz:

Multiple R                   ,94799  
 R Square                    ,89869  
 Adjusted R Square         ,85648  
 Standard Error            3,16228

Analysis of Variance			
	DF	Sum of Squares	Mean Square
Regression	5	1064,50000	212,90000
Residual	12	120,00000	10,00000

F =            21,29000            Signif F =   ,0000

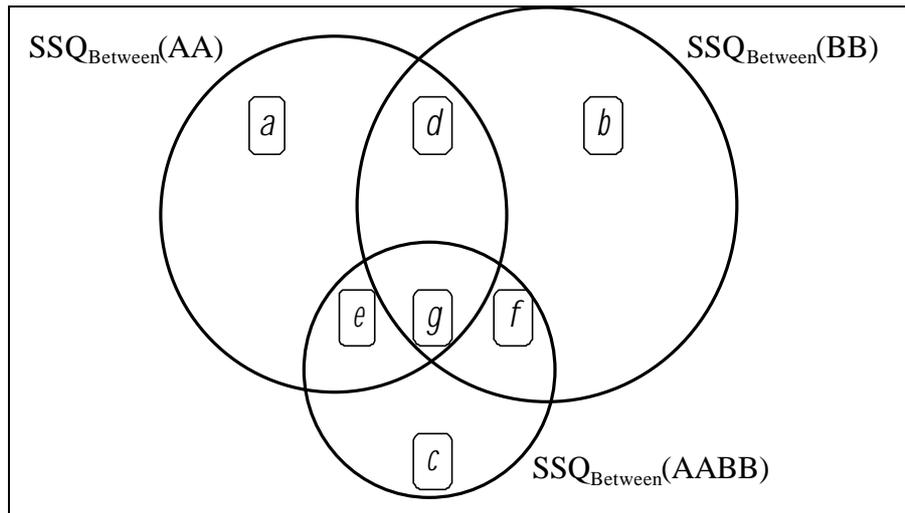
$SSQ_{\text{Regression}}$  (Sum of Squares) beträgt 1064,50 bei 5 Freiheitsgraden und  $SSQ_{\text{Residual}}$  120,00 bei 12 Freiheitsgraden. Diese Werte stimmen notwendigerweise mit den entsprechenden (identischen!) Werten für die experimentelle, die hierarchische und die regressionsanalytische Methode, nämlich  $SSQ_{\text{Between}}(\text{AA}*\text{BB})$  [ $\equiv SSQ_{\text{Explained}}$ ] und  $SSQ_{\text{Within}}(\text{AA}*\text{BB})$  [ $\equiv SSQ_{\text{Residual}}$ ], überein (vgl. 2.4.1, 2.4.2 und 2.4.3). Die erklärte Varianz  $R^2 = SSQ_{\text{Regression}}/(SSQ_{\text{Regression}}+SSQ_{\text{Residual}})$  beträgt 0,899 (R Square). Auf den weiteren Output (die einzelnen Effekte) soll an dieser Stelle nicht eingegangen werden.

Stattdessen sollen die unterschiedlichen Formen der Zerlegung von  $SSQ_{\text{Regression}}$  [ $\equiv SSQ_{\text{Between}}(\text{AA}*\text{BB})$ ] erläutert werden. Dazu werden 6 weitere Regressionsanalysen mit YY als abhängiger Variable und unterschiedlichen Konstellationen der Dummy-Variablen als unabhängige Variablen gerechnet. Die Ergebnisse aller Regressionsanalysen werden in der folgenden Tabelle zusammengefaßt:

Regr. Nr.	verwendete Dummy-Variablen					Ergebnisse		Felder in der Grafik
	AA1	AA2	BB1	AABB11	AABB21	R <sup>2</sup>	SSQ <sub>Regression</sub> <sup>n</sup>	
0	X	X	X	X	X	0,899	1064,50	<i>a+b+c+d+e+f+g</i>
1	X	X				0,420	497,33	<i>a+d+e+g</i>
2			X			0,497	588,32	<i>b+d+f+g</i>
3				X	X	0,158	186,70	<i>c+e+f+g</i>
4	X	X	X			0,821	972,59	<i>a+b+d+e+f+g</i>
5	X	X		X	X	0,487	577,00	<i>a+c+d+e+f+g</i>
6			X	X	X	0,664	786,94	<i>b+c+d+e+f+g</i>

Die Regressionen 1 bzw. 2 enthalten jeweils nur die zu den Faktoren AA bzw. BB zugeordneten Dummy-Variablen. Sie gleichen einfaktoriellen Varianzanalysen. Daher stimmen die entsprechenden Werte für  $SSQ_{\text{Regression}}$  notwendigerweise mit den Ergebnissen aus 2.4 für  $SSQ_{\text{Between}}(\text{AA})$  (= 497,3) bzw.  $SSQ_{\text{Between}}(\text{BB})$  (= 588,3) überein (vgl. 2.4).

Allgemein gilt, daß sich die Werte für  $SSQ_{\text{Regression}}$  aus den einzelnen Regressionen durch additive Kombination von Teilbereichen der insgesamt erklärten Varianz  $SSQ_{\text{Between}}(\text{AA}*\text{BB})$  berechnen. Zur Erläuterung haben wir den Teilbereichen der Varianzzerlegung in der folgenden Grafik Buchstaben von *a* bis *g* zugeordnet:



Die Regressionen sind zur Vereinfachung von 0 bis 6 durchnummeriert. Zunächst berechnen wir die Bereiche  $a$ ,  $b$ , und  $c$ , die eindeutig den Faktoren AA, BB bzw. dem Interaktionseffekt AABB zugeordnet werden können. Wir erhalten:

Feld	zugeordnet zu ...	Formel (Regressionsnummern)	Rechnung	SSQ <sub>Explained</sub>
$a$	AA	0 - 6	1064,50-786,94	277,56
$b$	BB	0 - 5	1064,50-577,00	487,50
$c$	AABB	0 - 4	1064,50-972,59	91,91

Als nächstes berechnen wir die Bereiche  $(d+g)$ ,  $(e+g)$  und  $(f+g)$ , die jeweils 2 Komponenten gleichzeitig zugeordnet werden können:

Feld	zugeordnet zu ...	Formel (Regressionsnummern)	Rechnung	SSQ <sub>Explained</sub>
$d+g$	AA, BB	1 + 2 - 4	497,33 + 588,32 - 972,59	113,06
$e+g$	AA, AABB	1 + 3 - 5	497,33 + 186,70 - 577,00	107,03
$f+g$	BB, AABB	2 + 3 - 5	588,32 + 186,70 - 786,94	-11,92 <sup>9</sup>

Der dreiseitige Überlappungsbereich  $g$  kann nun wie folgt berechnet werden:

$$g = a + (d + g) + (e + g) - (a + d + e + g) - a = 277,56 + 113,06 + 107,03 - 497,33 = 0,32$$

Für die Felder  $d$ ,  $e$  und  $f$  ergeben sich daraus folgende Werte:

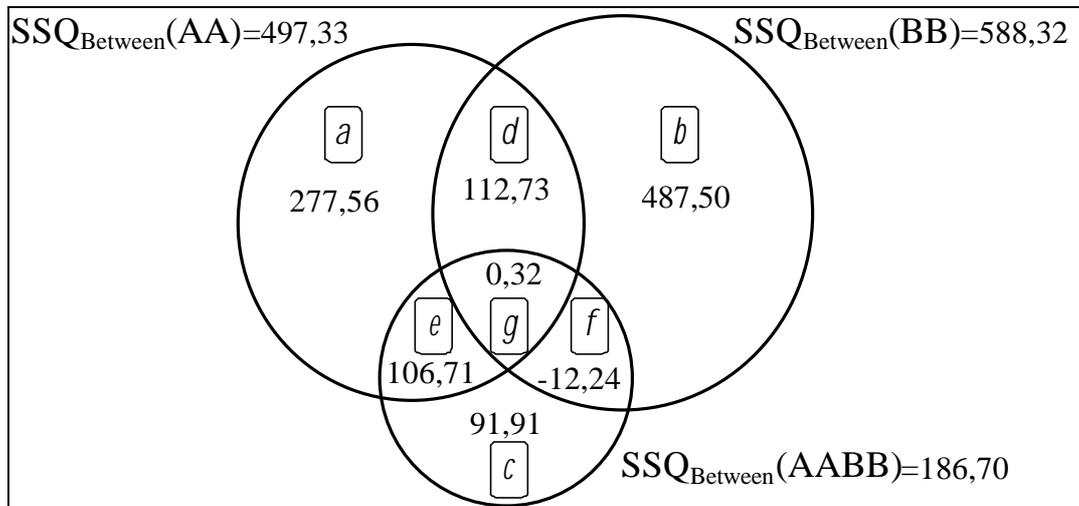
$$d = (d + g) - g = 113,06 - 0,32 = 112,74$$

$$e = (e + g) - g = 107,03 - 0,32 = 106,71$$

<sup>9</sup> Der negative Wert bedeutet hier inhaltlich, daß der partielle Effekt des Faktors B (bei Kontrolle des Interaktionseffektes) stärker ist, als der einfaktorielles Effekt von B. Dies käme in unserer Typologie von Konstellationen der Drittvariablenkontrolle (vgl. Skript 3.1.1.4) einer Verstärkung gleich (der partielle Effekt ist stärker als der bivariate Effekt).

$$f = (f + g) - g = -11,92 - 0,32 = -12,24$$

Zur Verdeutlichung haben wir die Werte in die Grafik übertragen:



Je nach der verwendeten Methode zur Varianzzerlegung werden nun die Haupteffekte und der Interaktionseffekt wie folgt berechnet (vgl. die schematische Darstellung der einzelnen Methoden in 2.3):

### Regressionsanalytisches Verfahren:

Effekte	Felder	SSQ
main effects	$a + d + b$	877,80
main effect AA	$a$	277,56
main effect BB	$b$	487,50
2-way-interaction AABB	$c$	91,91

### Hierarchisches Verfahren (mit AA dominant):

Effekte	Felder	SSQ
main effects	$a + b + d + e + f + g$	972,59
main effect AA	$a + d + e + g$	497,33
main effect BB	$b + f$	475,26
2-way-interaction AABB	$c$	91,91

### Experimentelles Verfahren:

Effekte	Felder	SSQ
main effects	$a + b + d + e + f + g$	972,59
main effect AA	$a + e$	384,27
main effect BB	$b + f$	475,26
2-way-interaction AABB	$c$	91,91

Die Werte entsprechen exakt den Ergebnissen der unter 2.4.1, 2.4.2 und 2.4.3 berechneten mehrfaktoriellen Varianzanalysen.

## **2.6 Beispiel aus der Forschungspraxis: die normative Geschlechtsrollenorientierung in Abhängigkeit von Konfession und Erhebungsgebiet**

Im ALLBUS 1996 wurde die normative Geschlechtsrollenorientierung (NGRO) der Befragten mit einer Likertskala gemessen. Zusätzlich liegen im ALLBUS 1996 Angaben über die Konfession der Befragten (hier evangelisch, römisch-katholisch oder konfessionslos) und deren Wohnort (alte oder neue Bundesländer) vor.

Wir untersuchen folgende Hypothesen:

1. *Die NGRO der Westdeutschen ist traditionaler als die der Ostdeutschen.* Das bedeutet, die Westdeutschen sollten auf der Likertskala zur Messung der NGRO einen niedrigeren Mittelwert als die Ostdeutschen aufweisen.
2. *Die NGRO der konfessionslosen Befragten ist liberaler als die der Mitglieder einer der beiden großen christlichen Religionsgemeinschaften.* Konfessionslose sollten also einen höheren Mittelwert auf der NGRO-Skala aufweisen.
3. *In Ostdeutschland gehören mehr Menschen keiner der beiden großen Kirchen an.* Wir gehen davon aus, daß zwischen dem Erhebungsgebiet und der Konfession eine starke Korrelation besteht.

Zunächst wenden wir uns der Überprüfung der dritten Hypothese zu. Sowohl das Erhebungsgebiet als auch die Konfession der Befragten ist nominalskaliert. Wir berechnen deshalb in SPSS eine Kreuztabelle und verwenden als Zusammenhangsmaße  $\chi^2$ ,  $\Phi$  und Cramer's V (vgl. die Beispieldatei zur Tabellenanalyse). Wir erhalten den folgenden (gekürzten) Output:



- - - - - O N E W A Y - - - - -

Variable	NGRO				
By Variable	V3	ERHEBUNGSGEBIET: WEST - OST			
Analysis of Variance					
Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	1	4912,5184	4912,5184	323,1668	,0000
Within Groups	2860	43475,3935	15,2012		
Total	2861	48387,9119			

$SSQ_{\text{Between}}$  beträgt 4913. Der Mittelwertsunterschied zwischen Ost- und Westdeutschen führt also zu einer erklärten Varianz von 10,2%. Die Prüfgröße F ist mit 323,2 hochsignifikant.

Group	Count	Mean	Standard Deviation	Standard Error	95 Pct Conf Int for Mean
ALTE BUN	1888	5,3273	4,0505	,0932	5,1445 TO 5,5102
NEUE BUN	974	8,0924	3,5865	,1149	7,8669 TO 8,3179
Total	2862	6,2683	4,1125	,0769	6,1176 TO 6,4191

Wie in Hypothese 1 postuliert, ist der Mittelwert der Ostdeutschen höher als der Mittelwert der Westdeutschen. Ostdeutsche haben also eine liberalere NGRO als Westdeutsche.

Levene Test for Homogeneity of Variances

Statistic	df1	df2	2-tail Sig.
26,6108	1	2860	,000

Wir müssen jedoch feststellen, daß die Voraussetzungen für den oben durchgeführten F-Test nicht ganz erfüllt sind, da der Levene-Test signifikant unterschiedliche Varianzen in den Gruppen ausweist. Das bedeutet, daß die Schätzung der Prüfgröße F möglicherweise verzerrt ist. Angesichts des sehr großen F-Wertes sollte allerdings selbst eine stärkere Verzerrung inhaltlich nichts an unserem Ergebnis ändern. Wir sehen unsere Hypothese also zunächst als bestätigt an.

Die Untersuchung des Einflusses der Konfession auf die NGRO liefert folgenden Output:

- - - - - O N E W A Y - - - - -

Variable	NGRO				
By Variable	KONFESS	Konfession			
Analysis of Variance					
Source	D.F.	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	2	5030,1997	2515,0999	165,8452	,0000
Within Groups	2859	43357,7122	15,1653		
Total	2861	48387,9119			

Die  $SSQ_{\text{Between}}$  beträgt hier 5030. Das bedeutet, daß die Konfession 10,4% der Varianz der NGRO erklärt. Die Prüfgröße F ist mit einem Betrag von 165,8 ebenfalls hochsignifikant.

Group	Count	Mean	Standard Deviation	Standard Error	95 Pct Conf Int for Mean
konfessi	970	8,0979	3,7822	,1214	7,8596 TO 8,3363
katholis	826	5,0460	3,9310	,1368	4,7775 TO 5,3145
evangeli	1066	5,5507	3,9656	,1215	5,3123 TO 5,7890
Total	2862	6,2683	4,1125	,0769	6,1176 TO 6,4191

Die Konfessionslosen haben mit 8,1 einen höheren Mittelwert auf der NGRO-Skala als die Angehörigen der christlichen Kirchen - Befragte ohne Konfession haben also eine liberalere NGRO.

Levene Test for Homogeneity of Variances

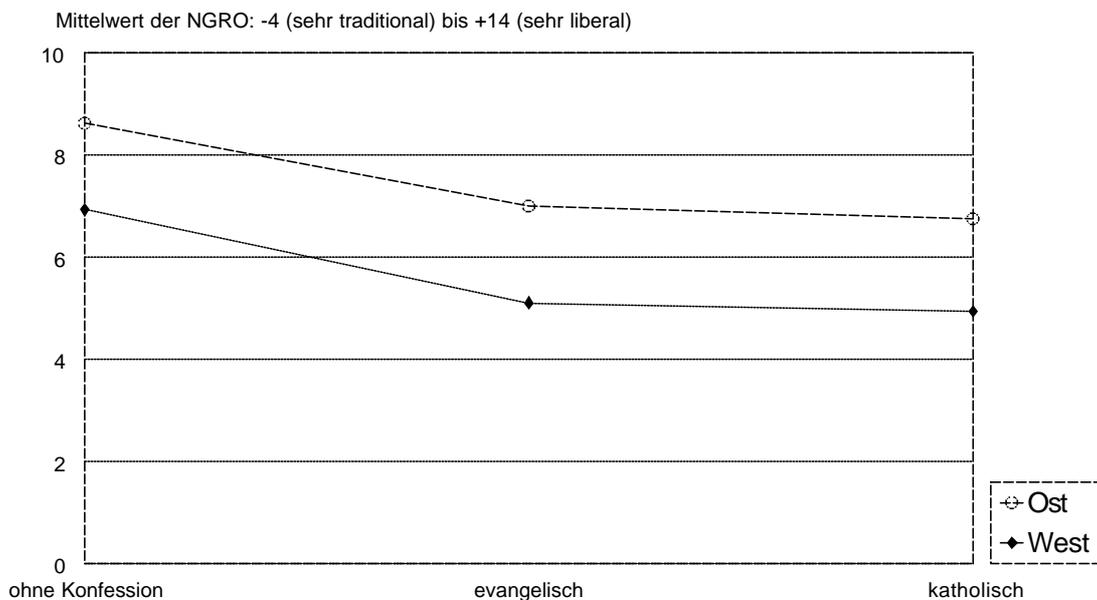
Statistic	df1	df2	2-tail Sig.
2,7995	2	2859	,061

Die Voraussetzungen des F-Testes sind erneut nicht vollständig erfüllt. Der Unterschied der Gruppenvarianzen ist jedoch nur noch auf dem 10%-Niveau signifikant (2-tail Sig. kleiner als 0,10). Auch hier ist der F-Wert sehr groß, so daß wir unsere Hypothese als bestätigt ansehen.

Bevor wir uns der multivariaten Analyse im einzelnen zuwenden, betrachten wir das Interaktionsdiagramm, welches die Mittelwerte der NGRO für die Konfessionsgruppen und getrennt nach Ost- und Westdeutschen enthält:

## Normative Geschlechterorientierung (NGRO) in West- und Ostdeutschland

Mittelwerte getrennt nach der Konfession der Befragten



Es ist deutlich zu sehen, daß Ostdeutsche in jeder Konfessionsgruppe einen höheren NGRO-Mittelwert als Westdeutsche aufweisen. Die Verbindungslinien zwischen den Mittelwerten verlaufen annähernd parallel, was darauf schließen läßt, das kein (oder nur ein sehr kleiner)

Interaktionseffekt vorhanden ist. Die Korrelation der unabhängigen Variablen ist also nicht gleichbedeutend mit der Existenz eines Interaktionseffektes.

Wir berechnen nun zunächst die zweifaktorielle ANOVA nach der experimentellen Methode und erhalten folgenden Output:

```
*** ANALYSIS OF VARIANCE ***
      by      NGRO
      KONFESS Konfession
      V3      ERHEBUNGSGEBIET: WEST - OST

      EXPERIMENTAL sums of squares
      Covariates entered FIRST

Source of Variation          Sum of
                             Squares    DF      Mean
                             Squares    F      Sig
                             of F

Main Effects
  KONFESS                    1518,552    2      759,276    51,691    ,00
  V3                          1400,871    1      1400,871   95,371   ,000

2-Way Interactions
  KONFESS V3                   6,160    2        3,080     ,210    ,811

Explained                    6437,230    5      1287,446   87,649    ,00

Residual                     41950,681  2856     14,689

Total                        48387,912  2861     16,913

3518 cases were processed.
656 cases (18,6 pct) were missing.
```

Wir sehen, daß die Haupteffekte beider unabhängigen Variablen stark abgenommen haben. Sie betragen für die Konfession nur noch 1519 (bivariat 5030) und für das Erhebungsgebiet nur noch 1401 (bivariat 5605). Dieser starke Rückgang bei den partiellen Effekten ist auf die Korrelation zwischen den unabhängigen Variablen zurückzuführen. Die insgesamt von den beiden Variablen erklärte Varianz enthält auch den Überlappungsbereich der beiden Variablen und dieser wird, weil er nicht eindeutig einer Variablen zugeordnet werden kann, bei der experimentellen Methode keiner der Variablen zugeordnet (vgl. 2.3). Er ist jedoch in der insgesamt durch beide Faktoren erklärten Varianz (Main Effects = 6431) enthalten.

Trotz der starken Abnahme der einzelnen Effekte sind trotzdem beide Effekte hochsignifikant. Dagegen ist der zweiseitige Interaktionseffekt mit einem Wert von lediglich 6 nicht signifikant ( $F = 0,210$  und  $\text{Sig of } F = 0,811$ ).

Insgesamt wird durch die beiden Variablen 13,3% der Varianz der NGRO-Skala erklärt. Wir führen nun erneut eine ANOVA mit diesen beiden Variablen durch, allerdings ohne Berücksichtigung des (nicht signifikanten!) Interaktionseffektes. Dazu geben wir unter den Optionen des Dialogfeldes der einfachen mehrfaktoriellen ANOVA keine Wechselwirkungen an. Zusätzlich fordern wir dort die multiple Klassifikationsanalyse (MCA) an.

\* \* \* A N A L Y S I S O F V A R I A N C E \* \* \*

NGRO  
by KONFESS Konfession  
V3 ERHEBUNGSGEBIET: WEST - OST

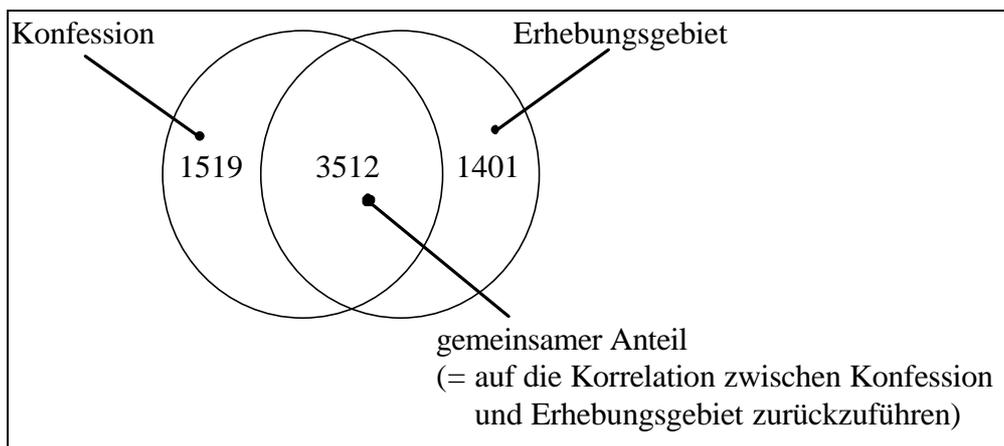
EXPERIMENTAL sums of squares  
Covariates entered FIRST

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig of F
Main Effects	6431,071	3	2143,690	146,023	,00
KONFESS	1518,552	2	759,276	51,720	,00
V3	1400,871	1	1400,871	95,424	,000
Explained	6431,071	3	2143,690	146,023	,00
Residual	41956,841	2858	14,680		
Total	48387,912	2861	16,913		

3518 cases were processed.  
656 cases (18,6 pct) were missing.

Der Output enthält nun nur noch die beiden Haupteffekte (Main Effects). Deren Größe hat sich notwendigerweise nicht verändert, da wir schon oben die experimentelle Methode angewandt haben, die die Überlappungsbereiche mit dem Interaktionseffekt den Haupteffekten zu-rechnet. Geändert hat sich (geringfügig) die  $SSQ_{\text{Between}}$  für alle Effekte gemeinsam (Explained), und zwar genau um den Betrag des nicht mehr berücksichtigten (nicht signifikanten) Interaktionseffektes.

Die Zerlegung der erklärten Varianz veranschaulicht noch einmal die folgende Grafik:



Wir betrachten nun den Output der MCA:

\* \* \* M U L T I P L E C L A S S I F I C A T I O N A N A L Y S I S \* \* \*

	NGRO				
by	KONFESS	Konfession			
	V3	ERHEBUNGSGEBIET: WEST - OST			
Grand Mean =	6,27				
Variable + Category	N	Unadjusted Dev'n	Eta	Adjusted for Independents Dev'n	Beta
KONFESS					
0 Konfessionslos	970	1,83		1,21	
1 evangelisch	1066	-,72		-,54	
2 katholisch	826	-1,22		-,72	
			,32		,21
V3					
1 ALTE BUNDESLAENDER	1888	-,94		-,60	
2 NEUE BUNDESLAENDER	974	1,82		1,17	
			,32		,20
Multiple R Squared					,133
Multiple R					,365

In der Spalte *Unadjusted Dev'n* sind die Abweichungen der Gruppenmittelwerte vom Gesamtmittelwert der NGRO (*Grand Mean*) bei jeweils bivariater Betrachtung aufgeführt. So erzielt die Gruppe der Konfessionslosen einen Mittelwert von  $6,27 + 1,83 = 8,10$ . In den alten Bundesländern liegt der Mittelwert der NGRO bei  $6,27 - 0,94 = 5,33$ . In der Spalte *Eta* wird für als Maß für die bivariate Beziehung  $\eta$  aufgeführt. Bivariat betrachtet erscheinen beide Beziehungen mit einem  $\eta$  von  $0,32$  in etwa gleich stark ( $\eta^2$  entspricht der erklärten Varianz in der bivariaten Analyse, s.o.).

In der Spalte *Adjusted for Independents Dev'n* wird die Abweichung der Gruppenmittelwerte vom Gesamtmittelwert der NGRO unter Kontrolle aller übrigen unabhängigen Variablen der Analyse ausgegeben. Diese Information ist in diesem Fall besonders interessant, da die unabhängigen Variablen eine hohe Korrelation miteinander aufweisen. Berücksichtigt man den Einfluß des Erhebungsgebietes, so liegt z.B. der Mittelwert der Konfessionslosen nur noch um  $1,21$  höher als der *Grand Mean*  $6,27$ . Der Mittelwert der NGRO in den alten Bundesländern beträgt unter Berücksichtigung des Einflusses der Konfession  $6,27 - 0,60 = 5,67$ . *Beta* stellt ein Beziehungsmaß für den partiellen Einfluß der unabhängigen Variablen auf die NGRO dar. Beide Werte für  $\beta$  fallen deutlich niedriger aus als im bivariaten Fall. Der Einfluß der Konfession erweist sich mit  $0,21$  als geringfügig stärker, als der Einfluß des Erhebungsgebietes mit  $0,20$ . Es liegt also für beide Beziehungen eine Abschwächung vor, da die partiellen Beziehungsmaße im multivariaten Fall kleiner ausfallen, als die Beziehungsmaße in der bivariaten Analyse. Der Grund hierfür liegt in der Korrelation zwischen den unabhängigen Variablen.