

FU Berlin, A. Iseler,
Aus dem Skript Statistik II, Fassung SS 2003

Achtung: Seitenzahlen sind gegenüber der vervielfältigten Fassung verändert!
Inhaltsverzeichnis am Ende dieses Auszugs.

II. Die Mehrweg-Varianzanalyse

1) Ein Experiment zur Zweiweg-Varianzanalyse

Es soll untersucht werden, welchen Einfluß die Arbeitsplatzgestaltung auf die Arbeit von Männern und Frauen hat. Eine bestimmte Arbeit wird unter drei Bedingungen durchgeführt:

Bedingung:

"nüchtern": Es handelt sich um saubere, aber nüchterne Arbeitsplätze.

"hübsch": Durch geschmackvolle Raumgestaltung (Vorhänge, Tapeten etc.) soll für ein angenehmes Arbeitsklima gesorgt werden.

"altmodisch": Traditionelle Fabrikatmosphäre.

Unter diesen Bedingungen arbeiten Männer und Frauen getrennt. Gemessen wird als Erfolgsquote die "Zahl der erfolgreich bearbeiteten Stücke (x)". Es ergeben sich die in Tabelle 16 zusammengestellten Meßwerte (x).

| Frauen | | | Männer | | |
|------------------|--------|------------|------------------|--------|------------|
| Arbeitsbedingung | | | Arbeitsbedingung | | |
| nüchtern | hübsch | altmodisch | nüchtern | hübsch | altmodisch |
| 142 | 179 | 120 | 173 | 140 | 119 |
| 149 | 172 | 131 | 164 | 150 | 127 |
| 136 | 169 | 139 | 180 | 153 | 126 |
| 146 | 177 | 136 | 172 | 161 | 139 |
| 141 | 168 | 128 | 174 | 138 | 133 |
| 144 | 159 | 121 | 161 | 131 | 113 |
| 152 | 167 | 125 | 183 | 142 | 125 |
| 140 | 164 | 135 | 168 | 146 | 134 |
| 147 | 157 | 117 | 179 | 144 | 128 |
| 143 | 164 | 127 | 184 | 151 | 124 |
| 149 | 172 | 125 | 179 | 153 | 118 |
| 148 | 168 | 132 | 162 | 141 | 125 |
| 156 | 168 | | 165 | 139 | 127 |
| 151 | 161 | | 184 | 138 | 126 |
| 146 | 175 | | 179 | 154 | 130 |
| 159 | | | 156 | 147 | 122 |
| 161 | | | 164 | 145 | |
| 154 | | | 173 | 150 | |
| | | | 168 | 153 | |
| | | | 184 | 145 | |
| | | | 169 | | |
| | | | 176 | | |
| | | | 166 | | |
| | | | 165 | | |
| 2664 | 2520 | 1536 | 4128 | 2920 | 2016 |

Tabelle 16: Ausgangsdaten (Unter dem Strich: Gruppensummen.)

2) "Faktoren" und "Stufen"

a) Begriffe

Wir haben insgesamt sechs Gruppen von Vpn. Es besteht aber ein wesentlicher Unterschied gegenüber der Einweg-Varianzanalyse: Die Gruppen unterscheiden sich unter zwei Aspekten: Nach der Arbeitsbedingung und nach dem Geschlecht.

Allgemein nennt man solche Aspekte, nach denen sich Gruppen von Meßwerten unterscheiden, auch Faktoren. In unserem Fall hätten wir also die zwei Faktoren "Arbeitsbedingung" und "Geschlecht". Die verschiedenen Ausprägungsstufen der einzelnen

Faktoren nennt man auch "Stufen" des Faktors (auch "Niveaus", engl. "levels"). Der Faktor "Geschlecht" hat die Stufen "männlich" und "weiblich".

Welche Stufen hat der Faktor Arbeitsbedingung?

Die Stufen des Faktors "Arbeitsbedingung" sind "hübsch", "nüchtern" und "altmodisch".

Wieviele Faktoren und welche Stufen hatten wir in unserem Experiment zur Einweg-Varianzanalyse?

Wir hatten einen Faktor ("Vorinformation") mit den drei Stufen "überdurchschnittlich", "durchschnittlich" und "unterdurchschnittlich".

Eine Einweg-Varianzanalyse ist dann angebracht, wenn sich die Gruppen nur in einem Faktor unterscheiden. In unserem neuen Experiment wollen wir dagegen die Auswirkungen der beiden Faktoren getrennt betrachten.

Wir analysieren in zwei Richtungen. Daher spricht man auch von einer Zweiweg-Varianzanalyse.

In Tabelle @17 sind die Stichprobengrößen der verschiedenen Gruppen in unserem Experiment zusammengestellt.

| | Frauen | Männer |
|------------|---------------|---------------|
| nüchtern | $N_{11} = 18$ | $N_{12} = 24$ |
| hübsch | $N_{21} = 15$ | $N_{22} = 20$ |
| altmodisch | $N_{31} = 12$ | $N_{32} = 16$ |

Tabelle 17: Stichprobengrößen

b) Die Unterscheidung von Zeilen und Spalten

Die ganze Struktur eines solchen Experiments wird durchsichtiger, wenn wir in Tabellen den einen Faktor mit den Zeilen (Reihen), den anderen Faktor mit den Spalten der Tabelle in Zusammenhang bringen. Den Stufen des einen Faktors (hier: Arbeitsbedingungen) entsprechen die verschiedenen Zeilen der Tabelle. Man nennt diesen Faktor dann den "Zeilenfaktor". Entsprechend nennt man den Faktor, dessen Stufen die Spalten der Tabelle entsprechen, den "Spaltenfaktor" (hier: Geschlecht). Natürlich ist es willkürlich, welchen Faktor man zum Zeilenfaktor und welchen man zum Spaltenfaktor macht. Man sollte sich aber einmal festlegen und in allen weiteren Tabellen eine entsprechende Unterteilung beibehalten.

c) Die Verwendung der Indices in der Zweiweg-Varianzanalyse

Wir sehen in Tabelle 17, daß die verschiedenen "Zellen" (das sind die Kreuzungsfelder der verschiedenen Zeilen und Spalten) nicht einfach durchnummeriert sind. Die verschiedenen N tragen Doppelindices. So ist N_{11} (lies: "N eins eins") die Stichprobengröße in der Zelle, die die

Kreuzung der 1. Zeile und der 1. Spalte ist, also in der Gruppe der Frauen, die unter Arbeitsbedingung "nüchtern" auftreten. Der erste Index bezeichnet die Zeile, der zweite die Spalte der Tabelle 17. Diese Festlegung ist erforderlich, um N_{12} und N_{21} voneinander unterscheiden zu können.

Nehmen wir die verschiedenen Zellen einer Zeile zusammen und berechnen eine Gesamtstichprobengröße für diese Zeile, so setzen wir anstelle des Spaltenindex (also des zweiten) einen Punkt. $N_{1.}$ wäre also das Gesamt- N der ersten Zeile. Entsprechend bedeutet es eine Zusammenfassung der verschiedenen Zellen einer Spalte, wenn wir den Zeilenindex (also den ersten) durch einen Punkt ersetzen. So wäre $N_{.1}$ das Gesamt- N der ersten Spalte.

Was würde $N_{2.}$ bedeuten?

Gesamt- N in der 2. Zeile, also $N_{21} + N_{22}$

Was würde $N_{.2}$ bedeuten?

Gesamt- N in der 2. Spalte, also $N_{12} + N_{22} + N_{32}$

Was würde $N_{..}$ bedeuten?

Gesamte Stichprobengröße nach Zusammenfassung der Zellen aus allen Zeilen und Spalten, wofür wir auch einfach N schreiben.

Entsprechend verfahren wir später auch beim Durchschnitt (vgl. Tabelle 18): \bar{x}_{11} ist der Durchschnitt der Zelle "11", also der Durchschnitt aus allen Frauen, die unter der Bedingung "nüchtern" arbeiten. $\bar{x}_{1.}$ ist der Durchschnitt der Leistung in Zeile 1, $\bar{x}_{.1}$ der Durchschnitt der Leistung in Spalte 1 und $\bar{x}_{..}$ der Gesamtdurchschnitt, für den wir auch einfach \bar{x} schreiben.

3) Effekte in der Zweiweg-Varianzanalyse

a) Die Durchschnitte

Aus den Daten unseres Experiments ergeben sich die folgenden Durchschnitte:

| | Frauen | Männer | Gesamt |
|------------|-------------------------|-------------------------|-------------------------|
| nüchtern | $\bar{x}_{11} = 148$ | $\bar{x}_{12} = 172$ | $\bar{x}_{1.} = 161.71$ |
| hübsch | $\bar{x}_{21} = 168$ | $\bar{x}_{22} = 146$ | $\bar{x}_{2.} = 155.43$ |
| altmodisch | $\bar{x}_{31} = 128$ | $\bar{x}_{32} = 126$ | $\bar{x}_{3.} = 126.86$ |
| Gesamt | $\bar{x}_{.1} = 149.33$ | $\bar{x}_{.2} = 151.07$ | $\bar{x} = 150.32$ |

Tabelle 18

Die Frauen, die unter Arbeitsbedingung "nüchtern" getestet wurden, erzielten also durchschnittlich 148 Punkte, usw.

Faßt man die Männer aus allen drei Arbeitsbedingungen zusammen zu einer Stichprobe, so ergibt sich hier ein Durchschnitt von 151.07 Punkten. Faßt man umgekehrt die Männer und Frauen unter der Bedingung "hübsch" zusammen, so ergibt sich ein Durchschnitt von 155.43

Punkten.

Warum ist dieser Durchschnitt nicht das Mittel von 168 und 146 (das wäre ja 157)?

Weil der Durchschnitt einer zusammengefaßten Gruppe normalerweise nur dann das Mittel der Einzeldurchschnitte ist, wenn die Einzelgruppen gleich groß sind.

Faßt man schließlich alle sechs Gruppen zusammen, so ergibt sich ein Durchschnitt von 150.32.

b) Effekte des Zeilenfaktors

Wir können nun - genau wie bei der Einweg-Varianzanalyse - Effekte schätzen. Dort war ja jeder Effekt definiert als die Abweichung eines zu einer bestimmten Stufe gehörenden Populationsmittels μ_g vom Populationsmittel μ :

$$a_g = \mu_g - \mu$$

Wir schätzen die Effekte, indem wir die in einer Untersuchung erzielten Durchschnitte als Schätzungen für die entsprechenden Populationsmittel einsetzen:

$$\hat{a}_g = \bar{x}_g - \bar{x}$$

Ganz entsprechend verfahren wir hier.

Wie würde man z.B. a_h , also den Effekt der Arbeitsbedingung "hübsch", schätzen und verbal beschreiben?

$$\hat{a}_h = \bar{x}_h - \bar{x} = 155.43 - 150.32 = +5.11$$

In Worten: Wir schätzen, daß der Effekt der Arbeitsbedingung "hübsch" in einer Hebung der Erfolgsquote um durchschnittlich 5.11 Einheiten besteht. (Das Wort durchschnittlich fügen wir ein, um eventuellen differentiellen Effekten Rechnung zu tragen.)

Schätzen und beschreiben Sie in gleicher Weise den Effekt der beiden anderen Arbeitsbedingungen!

$$\hat{a}_n = \bar{x}_n - \bar{x} = 161.71 - 150.32 = 11.39$$

Wir schätzen, daß der Effekt der Arbeitsbedingung "nüchtern" in einer Hebung der Erfolgsquote um durchschnittlich 11.39 Einheiten besteht.

$$\hat{a}_a = \bar{x}_a - \bar{x} = 126.86 - 150.32 = -23.46$$

Wir schätzen, daß der Effekt der Arbeitsbedingung "altmodisch" in einer Senkung (negatives Vorzeichen) der Erfolgsquote um durchschnittlich 23.46 Einheiten besteht.

Diese Effekte des Zeilenfaktors nennt man auch einfach die Zeileneffekte. Eine allgemeine Formel für die Schätzung eines Zeileneffekts wäre:

$$\hat{a}_r = \text{Zeilennittel} - \text{Gesamtmittel} = \bar{x}_r - \bar{x}$$

Dabei wird (in diesem Skript) als "allgemeiner Index" für die Zeile (Reihe) der Buchstabe r verwendet. Dies geschieht (obschon das Wort Zeile in solchem Zusammenhang geläufiger ist), damit wir später den Index z frei haben, um Zellen zu bezeichnen.

Die obigen Werte sind natürlich nur Schätzungen. Den wahren Effekt würde man erhalten, wenn man statt der Stichprobendurchschnitte die entsprechenden Populationsdurchschnitte

einsetzen würde:

$$a_T = \mu_T - \mu$$

c) Effekte des Spaltenfaktors

Um die Effekte des Spaltenfaktors (die Spalteneffekte) deutlich von denen des Zeilenfaktors unterscheiden zu können, verwenden wir hier den Buchstaben b.

Wie würden Sie den Effekt des Geschlechts "weiblich" schätzen und in Worte fassen?

$$\hat{b}_w = \bar{x}_w - \bar{x} = 149.33 - 150.32 = -0.99$$

Der Effekt des Geschlechts weiblich besteht schätzungsweise in einer Senkung der Erfolgsquote um durchschnittlich 0.99 Punkte.

Dasselbe für männlich?

$$\hat{b}_m = \bar{x}_m - \bar{x} = 151.07 - 150.32 = +0.75$$

Den Effekt des Geschlechts männlich schätzen wir als Hebung um durchschnittlich 0.75 Punkte.

Könnten Sie eine allgemeine Formel für den Effekt der Spalte s und seine Schätzung angeben?

$$b_s = \mu_{.s} - \mu$$

$$\hat{b}_s = \bar{x}_{.s} - \bar{x}$$

d) Wechselwirkung

In Abschnitt c haben wir nur sehr schwache Effekte des Faktors "Geschlecht" festgestellt. Wir berücksichtigten dabei aber nur Gesamtdurchschnitte von Männern und Frauen. Betrachten wir aber die Tabelle der Durchschnitte (Tabelle 18), dann sehen wir eine weitere Auswirkung des Geschlechts, die sich in den Gesamtdurchschnitten nicht niederschlägt: Die Auswirkungen des Faktors "Geschlecht" sind auf den verschiedenen Stufen des Faktors "Arbeitsbedingung" verschieden:

Unter nüchternen Arbeitsbedingungen sind die Männer besser, unter hübschen Arbeitsbedingungen die Frauen, während unter altmodischen Bedingungen beide fast gleich gut sind. Man spricht hier von einer Wechselwirkung.

Allgemein spricht man in der Varianzanalyse von einer Wechselwirkung, wenn die Effekte des einen Faktors auf den verschiedenen Stufen des anderen Faktors nicht gleich sind.

Wir könnten den gleichen Sachverhalt in unserem Experiment auch andersherum betrachten: Auf den verschiedenen Stufen des Faktors "Geschlecht" ist die Auswirkung des Faktors

"Arbeitsbedingung" verschieden:

Bei den Männern ist die Bedingung "nüchtern" die günstigste, bei den Frauen die Bedingung "hübsch".

Allgemein: Jede Wechselwirkung läßt sich aus der Perspektive beider Faktoren betrachten.

Nehmen wir einmal an, wir hätten im Faktor "Arbeitsbedingung" nur die zwei Stufen "hübsch" und "altmodisch" gehabt und die gleichen Durchschnitte erzielt.

Wäre dann noch eine Wechselwirkung vorhanden?

Zunächst würde man versucht sein, die Existenz einer Wechselwirkung zu leugnen:

Unter beiden Arbeitsbedingungen sind die Frauen besser als die Männer. Aber das Ausmaß dieses Unterschieds ist verschieden groß. Die Effekte des Faktors "Geschlecht" haben also auf beiden Stufen des Faktors "Arbeitsbedingung" zwar das gleiche Vorzeichen (+ für Frauen; - für Männer), aber einen verschiedenen Betrag. Es würde also auch hier eine Wechselwirkung vorliegen. - Wie schon erwähnt, kann man jede Wechselwirkung aus der Perspektive beider Faktoren betrachten.

Wie sähe die zweite Perspektive hier aus?

Bei beiden Geschlechtern ist zwar die Arbeitsbedingung "altmodisch" ungünstiger als die Arbeitsbedingung "hübsch", aber das Ausmaß dieses Effekts ist verschieden groß.

Allgemein: Von einer Wechselwirkung spricht man auch dann, wenn sich die Effekte des einen Faktors auf den verschiedenen Stufen des anderen Faktors nicht dem Vorzeichen, sondern nur dem Betrag nach unterscheiden.

Wir wollen zu unserem Versuchsplan mit den drei Stufen des Faktors "Arbeitsbedingung" zurückkehren und uns fragen, wie die Tabelle aussähe, wenn es keine Wechselwirkung gäbe. Dann müßte es *einen* Spalteneffekt b_w und *einen* Spalteneffekt b_m geben, die man auf allen Stufen des Faktors "Arbeitsbedingung" zu dem jeweiligen Zeilendurchschnitt addiert, um zum Zellendurchschnitt zu kommen. Die Zeilendurchschnitte ergeben sich aber, indem man den jeweiligen Zeileneffekt zum Gesamtdurchschnitt \bar{x} addiert. (Überzeugen Sie sich davon anhand der Definitionsformel für \hat{a}_r !)

Zusammengefaßt ergäbe sich:

$$\mu_z = \mu + a_r + b_s$$

mit

μ_z = Populationsmittel, das zur Zelle z gehört

In Worten: Wenn es in der Population keine Wechselwirkung gäbe, würde sich jedes Zellenmittel ergeben, indem man den zugehörigen Zeileneffekt und den zugehörigen Spalteneffekt zum Gesamtmittel addiert.

Diese Feststellung läßt sich nicht beweisen. Sie ist als Definition zu verstehen: Eine

Wechselwirkung liegt definitionsgemäß dann vor, wenn es mindestens eine Zelle gibt, bei der die obige Formel für eine Situation ohne Wechselwirkung nicht gilt.

Diese Formel, die für das Populationsmittel μ_z geschrieben ist und die in der Population gültigen Effekte a_r und b_s zugrundelegt, könnten wir analog auf die Verhältnisse in der Stichprobe übertragen:

Wenn es in unseren Daten keine Wechselwirkung gäbe, müßte gelten:

$$\bar{x}_z = \bar{x} + \hat{a}_r + \hat{b}_s$$

Davon ausgehend können wir nun einen numerischen Wert für den Wechselwirkungseffekt angeben, der in einer bestimmten Zelle, also durch die Kombination einer bestimmten Stufe des Zeilenfaktors und einer bestimmten Stufe des Spaltenfaktors wirksam wird.

Wir fragen uns einfach:

Um wieviel liegt der Durchschnitt jeder einzelnen Zelle oberhalb bzw. unterhalb des Wertes, der sich bei Additivität der zugehörigen Spalten- und Zeileneffekte ergeben würde? Hierzu bilden wir die entsprechende Differenz:

$$\hat{c}_z = \bar{x}_z - (\bar{x} + \hat{a}_r + \hat{b}_s)$$

Der nach dieser Formel ermittelte Wert \hat{c}_z ist der geschätzte Wechselwirkungseffekt der zugehörigen Zelle.

Beispiel aus unserem Experiment:

$$\begin{aligned}\hat{c}_{11} &= \bar{x}_{11} - (\bar{x} + \hat{a}_1 + \hat{b}_{,1}) \\ &= 148 - [150.32 + 11.39 + (-0.99)] \\ &= 148 - 160.72 \\ &= -12.72\end{aligned}$$

In Worten: Die Frauen sind unter der Arbeitsbedingung "nüchtern" um 12.72 Punkte schlechter, als es ohne Wechselwirkung zu erwarten wäre.

Übungsaufgabe: Schätzen und verbalisieren Sie die Effekte der Wechselwirkung für weitere Zellen in ähnlicher Weise.

Im Unterschied zu diesen Wechselwirkungseffekten, die sich aus der Kombination einer bestimmten Zeile mit einer bestimmten Spalte ergeben, nennt man die in den vorhergehenden Abschnitten behandelten Effekte der Zeilen und Spalten auch Haupteffekte.

Was mit einer Wechselwirkung gemeint ist, läßt sich auch gut an der Darstellung der Mittelwerte in Abbildung 6 sehen:

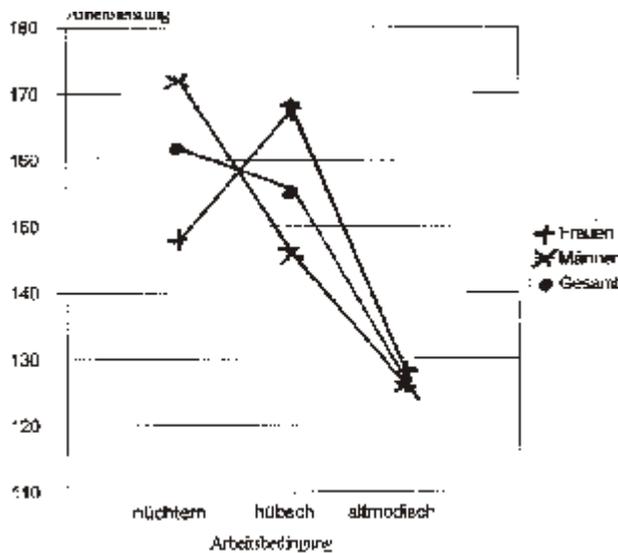


Abb. 6: Wechselwirkungsdiagramm

Die Wechselwirkung besteht - wie wir gesehen haben - darin, daß der Effekt des Geschlechts auf den verschiedenen Stufen des Faktors "Arbeitsbedingung" unterschiedlich ist. In unserer Darstellung zeigt sich das darin, daß die Kurven nicht parallel verlaufen. Wenn nämlich der Effekt des Geschlechts unter allen Arbeitsbedingungen gleich wäre, dann müßte die Kurve der Männer überall um den gleichen Betrag oberhalb der "Gesamt"-Kurve verlaufen. Beide müßten also parallel zur "Gesamt"-Kurve und damit auch parallel zueinander verlaufen.

Allgemein: Was bei einer Wechselwirkung genau vorliegt, kann man sich verdeutlichen, indem man eine solche graphische Darstellung anlegt und feststellt, in welcher Weise die Kurven von einem parallelen Verlauf abweichen.

Wir hatten vorhin festgestellt, daß eine Wechselwirkung auch dann noch vorläge, wenn wir nur die Arbeitsbedingungen "hübsch" und "altmodisch" gehabt hätten und die gleichen Durchschnitte erhalten hätten. Die Besonderheit bestand darin, daß hier zwar unter beiden Arbeitsbedingungen die Frauen besser und andererseits bei beiden Geschlechtern die Leistung unter Arbeitsbedingung "hübsch" besser als unter "altmodisch" ist; aber der Betrag dieser Effekte ist jeweils verschieden. Auch diese Art von Wechselwirkung zeigt sich in der graphischen Darstellung: Von "hübsch" nach "altmodisch" fallen zwar alle Kurven ab, verlaufen aber nicht parallel.

Es ist wichtig drei verschiedene Arten von Effekten zu unterscheiden:

1. Den Effekt, der insgesamt mit einer Stufe eines Faktors verbunden ist (sogenannter

Haupteffekt).

Beispiele: Der Effekt der Arbeitsbedingung "nüchtern" (S. @).

Der Effekts des Geschlechts "weiblich" (S. @).

Ein solcher Effekt gibt an, um welchen Betrag der zu dieser Stufe gehörende Durchschnitt nach Zusammenfassung aller Stufen des anderen Faktors vom Gesamtdurchschnitt abweicht.

2. Der Effekt, der mit einer Stufe eines Faktors auf einer bestimmten Stufe eines anderen Faktors verbunden ist.

Beispiele: Der Effekt der Arbeitsbedingung "nüchtern" bei den Männern beträgt 19.93 Punkte. Auf der Stufe "männlich" liegt nämlich der Durchschnitt unter Arbeitsbedingung "nüchtern" (172) um 19.93 Punkte über dem Gesamtdurchschnitt dieser Stufe (151.07).

Der Effekt des Geschlechts "weiblich" unter der Arbeitsbedingung "hübsch" beträgt 12.57. Auf der Stufe "hübsch" liegt nämlich der Durchschnitt der Frauen (168) um 12.57 Punkte über dem Gesamtdurchschnitt dieser Stufe (155.43).

Effekte dieser Art haben in der Varianzanalyse keinen besonderen Namen. Wir haben sie hier nur eingeführt, weil sich an der Unterschiedlichkeit solcher Effekte am ehesten aufzeigen läßt, was eine Wechselwirkung ist. Solche Effekte selbst sind aber noch keine Wechselwirkungseffekte. Diese Wechselwirkungseffekte sind folgendermaßen definiert:

3. Effekte, die in der Abweichung eines Zellendurchschnitts von demjenigen Wert bestehen, der aufgrund der Haupteffekte (d.h. aufgrund der Zeilen- und Spalten-Effekte) allein zu erwarten wäre (*Wechselwirkungseffekt einer Zelle*).

Beispiel (vgl. Tabelle 18): Für die Zelle "Frauen" unter Arbeitsbedingung "nüchtern" würden wir allein aufgrund der Haupteffekte der Arbeitsbedingung "nüchtern" und des Geschlechts "weiblich" einen Durchschnitt $\bar{x} + \hat{a}_n + \hat{b}_w = 150 + 11.39 + (-0.99) = 160.72$ erwarten. Der tatsächliche Zellendurchschnitt liegt aber bei 148 und damit um 12.72 Punkte unterhalb des Werts, der aufgrund der Haupteffekte zu erwarten wäre. Also beträgt der Wechselwirkungseffekt für diese Zelle -12.72.

Das kann man auch folgendermaßen formulieren: Der Wechselwirkungseffekt ist derjenige Effekt, der von einer Kombination zweier Stufen (z.B. von der Kombination der Arbeitsbedingung "nüchtern" mit dem Geschlecht "weiblich") zusätzlich zu den zugehörigen "Haupteffekten" ausgeht.

<< *Allgemeiner Begriff der Wechselwirkung*

Der hier behandelte Begriff der Wechselwirkung in der Varianzanalyse ist ein Spezialfall eines allgemeinen Begriffs der Wechselwirkung. Allgemein spricht man von einer Wechselwirkung, wenn die Beziehung zwischen zwei oder mehreren Variablen sich ändert in Abhängigkeit von den Werten, die eine dritte Variable annimmt.

Ein weiteres Beispiel für eine Wechselwirkung - außerhalb der Varianzanalyse: Empirische Untersuchungen haben ergeben, daß (zumindest in einigen Schultypen) die korrelative Beziehung zwischen dem Interesse der Schüler und ihren Leistungen um so enger ist, je weniger starr,

pedantisch ("rigide") die Schüler sind.⁹⁵ Die Beziehung zwischen den Variablen "Interesse" und "Leistung" ändert sich also in Abhängigkeit von Werten, die die Variable "Rigidität" annimmt.

>>

e) Allgemeines Modell der Zweiweg-Varianzanalyse

Wir haben bisher definiert:

$$\mu_r - \mu = a_r$$

$$\mu_s - \mu = b_s$$

$$\mu_z - (\mu + a_r + b_s) = c_z$$

Außerdem gibt es auch in der Zweiweg-Varianzanalyse einen "Fehler". Er ist auch hier definiert als die Abweichung eines Meßwerts vom zugehörigen Gruppen- (also Zellen-) Mittel:

$$x_i - \mu_z = e_i$$

Setzen wir nun für a_r und b_s in der dritten Gleichung die in der ersten bzw. zweiten Gleichung angegebenen Ausdrücke ein und addieren dann die vier Gleichungen, so erhalten wir auf der linken Seite:

$$\mu_r - \mu + \mu_s - \mu + \mu_z - (\mu + \mu_r - \mu + \mu_s - \mu) + x_i - \mu_z$$

Wenn wir diesen Ausdruck ausrechnen (wobei sich viele Glieder gegenseitig aufheben) und daneben die Summe der rechten Seite unserer vier Gleichungen schreiben, so erhalten wir:

$$x_i - \mu = a_r + b_s + c_z + e_i$$

In Worten: Die Abweichung eines Meßwerts x_i vom Gesamtmittel μ setzt sich aus vier Komponenten zusammen: Dem Effekt der zugehörigen Zeile (Reihe), dem Effekt der zugehörigen Spalte, dem Wechselwirkungseffekt der Zelle und dem "Fehler". Diese Abweichungen der Meßwerte x_i vom Gesamtmittel μ machen aber die Varianz der x -Werte aus (die Gesamtvarianz ist ja die durchschnittliche quadrierte Abweichung der Meßwerte vom Gesamtmittel).

Entsprechend lösen wir in der Zweiweg-Varianzanalyse dann auch die Varianz in vier Komponenten auf: In die Varianz zwischen den Zeilen, die Varianz zwischen den Spalten, die auf Wechselwirkung zurückzuführende Varianz und die Varianz innerhalb der Zellen ("Fehlervarianz").

Addiert man auf beiden Seiten der letzten Gleichung den Wert μ , so erhält man:

$$x_i = \mu + a_{r(i)} + b_{s(i)} + c_{z(i)} + e_i.$$

⁹⁵Psychologische Interpretation: Die allgemeine Erwartung, daß Interesse und Leistung eng zusammenhängen, trifft bei rigiden Schülern weniger zu, weil sie (vielleicht aufgrund eines starken "Über-Ich"?) ihren Arbeitseinsatz weniger von Interessen abhängig machen und weil sie andererseits weniger dazu tendieren, ihre Interessen zu verändern, wenn Leistungen und Interessen nicht übereinstimmen - eben weil sie starr, rigide sind.

Dabei sind $r(i)$, $s(i)$ und $z(i)$ die Reihe, die Spalte und die Zelle, zu denen der Meßwert x_i gehört.

Genau so wie in der Einweg-Varianzanalyse ist auch hier die Modellgleichung keine Zusatzvoraussetzung für die Tests; sie ergibt sich aus der Definition der Effekte. (Wollen Sie es als Übung selbst herleiten?)

4) Die Zusatzvoraussetzungen der Zweiweg-Varianzanalyse

a) Unabhängige Stichproben

(Darauf wird in einem späteren Abschnitt genauer eingegangen).

b) Normalverteilung und Varianzhomogenität

Die den verschiedenen Zellen entsprechenden Populationverteilungen müssen alle *Normalverteilungen* sein und *gleiche Varianz* haben. Die Normalitätsannahme und die Homogenitätsannahme verlieren an Gewicht bei großen bzw. gleich großen Stichproben.

c) Orthogonalität

Die sogenannte *Orthogonalität*. Zur Erläuterung des damit Gemeinten betrachten wir die Tabelle 17. Es gilt die Proportion:

$$N_{11} : N_{12} = N_{21} : N_{22} = N_{31} : N_{32} = N_{.1} : N_{.2}$$

In Zahlen:

$$18 : 24 = 15 : 20 = 12 : 16 = 45 : 60$$

Allgemein muß gelten:

$$N_{11} : N_{12} : N_{13} : \dots = N_{21} : N_{22} : N_{23} : \dots = N_{31} : N_{32} : N_{33} : \dots = \text{usw.} = N_{.1} : N_{.2} : N_{.3} : \dots$$

In Worte gefaßt: Innerhalb jeder Zeile muß das Verhältnis der Besetzungshäufigkeiten der Zellen gleich sein.

Bisher haben wir den Begriff der Orthogonalität auf einen Vergleich der Häufigkeitsverhältnisse innerhalb der verschiedenen Zeilen gestützt. Genauso kann man aber auch die Häufigkeitsverhältnisse innerhalb der Spalten zugrundelegen. In unserem Beispiel gilt

$$N_{11} : N_{21} : N_{31} = N_{12} : N_{22} : N_{32} = N_{1.} : N_{2.} : N_{3.},$$

bzw. in Zahlen

$$18 : 15 : 12 = 24 : 20 : 16 = 42 : 35 : 28.$$

In Worte gefaßt: Auch innerhalb jeder Spalte muß das Verhältnis der Besetzungshäufigkeiten der Zellen gleich sein.

Es genügt aber, diese Proportionalität in einer der beiden Richtungen zu überprüfen; denn wenn sie in der einen Richtung zutrifft, gilt sie auch für die andere.

Eine derartige Proportionalität von Häufigkeiten kennen wir schon als Nullhypothese des χ^2 -Tests für stochastische Unabhängigkeit. Der Unterschied läßt sich folgendermaßen charakterisieren:

- Beim χ^2 -Test stellen die Häufigkeiten, für die lt. H_0 diese Proportionalität gelten soll, die eigentlichen Daten dar. Ob die Proportionalität tatsächlich zutrifft, ist die eigentliche empirische Frage.
- In der Zweiweg-Varianzanalyse ist die Proportionalität als Zusatzvoraussetzung durch den Versuchsplan zu gewährleisten. Die empirische Fragestellung bezieht sich dann auf die Effekte der Zugehörigkeit zu den Gruppen auf Meßwerte in einer völlig andere Variable (in unserem Fall die in Tabelle @ angegebenen Leistungsmaße).

Wenn man diesen Unterschied einmal klargestellt hat, kann man die Orthogonalitäts-Forderung auch folgendermaßen formulieren:

Allgemeine Beschreibung: Die Zusatzvoraussetzung der Orthogonalität verlangt, daß für die Besetzungshäufigkeiten der Zellen die gleichen Proportionalitäten gelten wie für die f_e -Werte im χ^2 -Test für stochastische Unabhängigkeit.

Auch eine verkürzte Formeldarstellung ergibt sich aus diesem Vergleich. Die allgemeine Formel für f_e -Werte (vgl. @) läßt sich folgendermaßen übertragen: Für jede Zeile (Reihe) r und jede Spalte s muß für die Besetzungshäufigkeit N_{rs} der von dieser Zeile und dieser Spalte gebildeten Zelle gelten:

$$N_{rs} = N_{r.} \cdot N_{.s} / N.$$

<<

Eine Begründung für die Orthogonalitäts-Forderung ergibt sich aus Interpretationsschwierigkeiten in Fällen, in denen die Orthogonalität verletzt ist. Nehmen wir z.B. an, wir hätten unser bisher betrachtetes Experiment ohne die Arbeitsbedingung "altmodisch" und mit folgenden Zell-Besetzungen durchgeführt:

$$N_{11} = 10; N_{12} = 10; N_{21} = 10; N_{22} = 40.$$

Falls wir die gleichen Zellendurchschnitte erhalten hätten (vgl. Tabelle 19), dann ließen sich die Gesamtdurchschnitte für Zeilen und Spalten aus den Zellendurchschnitten und den N-Werten der Zellen nach der Formel für Durchschnitte zusammengefaßter Gruppen (S. @) berechnen:

| | Frauen | Männer | Gesamt |
|----------|----------------------|------------------------|------------------------|
| nüchtern | $\bar{x}_{11} = 148$ | $\bar{x}_{12} = 172$ | $\bar{x}_{1.} = 160$ |
| hübsch | $\bar{x}_{21} = 168$ | $\bar{x}_{22} = 146$ | $\bar{x}_{2.} = 150.4$ |
| gesamt | $\bar{x}_{.1} = 158$ | $\bar{x}_{.2} = 151.2$ | $\bar{x} = 153.1$ |

Tabelle 19: Zellenmittelwerte des "nicht-orthogonalen Versuchsplans"

Übungsaufgabe: Überprüfen Sie die Richtigkeit!

Wir sehen: Die "Gesamt"-Durchschnitte haben sich geändert, da sich die "relativen Stichprobengrößen" geändert haben und da die dritte Arbeitsbedingung ("altmodisch") fehlt. Entsprechend haben sich auch die Effekte geändert.

Allgemein: Die Höhe der Effekte hängt vom Verhältnis der Stichprobengrößen in den Zellen ab und davon, welche weiteren Stufen der einzelnen Faktoren vertreten sind.

Für uns ist dabei aber eines vor allem interessant: Wir finden jetzt Unterschiede zwischen den Zeilen und zwischen den Spalten, die vorher nicht in diesem Ausmaß da waren: Unter Arbeitsbedingung "hübsch" wird bemerkenswert schlechter gearbeitet als unter Arbeitsbedingung "nüchtern", und die Männer sind im Durchschnitt schlechter als die Frauen. Aber es gibt die Möglichkeit, daß einer der beiden Effekte nur eine sekundäre Konsequenz des anderen ist: Es wäre denkbar, daß unter Arbeitsbedingung "hübsch" nur deshalb schlechter gearbeitet wird, weil 40 der 50 Vpn, die unter dieser Bedingung getestet wurde, Männer sind (die laut Spaltendurchschnitt schlechter sind), während unter Arbeitsbedingung "nüchtern" gleich viele Männer wie Frauen arbeiten. Anders formuliert: Der unter Arbeitsbedingung "hübsch" besonders hohe Anteil von Männern (80% gegenüber 50%) hat den Durchschnitt dieser Arbeitsbedingung gesenkt. Es wäre denkbar, daß sonst gar kein Unterschied zwischen den Zeilen bestünde.

Daneben wäre auch die folgende Interpretation denkbar: In Wirklichkeit bestehen keine Unterschiede zwischen Männern und Frauen. Von den 50 Männern haben aber 40 unter der - laut Zeilendurchschnitt - ungünstigeren Arbeitsbedingung "hübsch" gearbeitet. Also: Die Männer sind vielleicht nur deshalb im Durchschnitt schlechter, weil sie vor allem unter der ungünstigeren Arbeitsbedingung antraten.

Schließlich besteht eine dritte Möglichkeit:

In Wirklichkeit bestehen keine Zeilen- und keine Spalteneffekte, sondern nur die Wechselwirkung. Da aber die Arbeitsbedingung "hübsch" für die Männer die ungünstigere ist und da gerade diese Zelle besonders stark besetzt (mit 40 Vpn im Gegensatz zu je 10 Vpn bei den übrigen Zellen), wird der Zeilendurchschnitt von "hübsch" und der Spaltendurchschnitt der Männer künstlich gesenkt.

Allgemein: Ist die Orthogonalitätsvoraussetzung verletzt, so kann es vorkommen, daß der eine Effekt die Existenz (oder auch das Fehlen) des anderen vortäuscht. Es läßt sich mathematisch beweisen, daß so etwas nicht vorkommt, wenn die Orthogonalitätsvoraussetzung erfüllt ist.

Es gibt spezielle Verfahren, mit denen man eine Zweiweg-Varianzanalyse auch bei Verletzung der Orthogonalitätsvoraussetzung durchführen kann. Die Ergebnisse derartiger Varianzanalysen sind aber schwer zu interpretieren. Deshalb ist es besser, schon im Stadium der

Versuchsplanung für Orthogonalität zu sorgen, soweit das möglich ist.

>>

d) Eine letzte Voraussetzung

<<

Eine letzte Voraussetzung der Zweiweg-Varianzanalyse in der hier vorgestellten Form ist es, daß es sich nicht um ein sogenanntes "Modell mit Zufallsfaktoren" und auch nicht um ein "gemischtes Modell" handelt. Was darunter zu verstehen ist, wird in einem späteren Abschnitt (9 cb) behandelt.

>>

5) Grundlagen und Durchführung der Berechnungen in der Zweiweg-Varianzanalyse.

a) Die Zerlegung der Abweichungen vom Gesamtdurchschnitt

Wir können die 6 Gruppen (Zellen) unseres Versuchsplans zunächst einmal so betrachten, als handele es sich um 6 Gruppen in einer Einweg- Varianzanalyse. Dann läßt sich jede Abweichung eines Meßwerts vom Gesamtdurchschnitt zerlegen in eine Abweichung des Meßwerts vom zugehörigen Zellendurchschnitt und in die Abweichung des Zellendurchschnitts vom Gesamtdurchschnitt. Für unsere erste Vp heiße das z.B.:

$$\begin{array}{rclcl} 142 - 150.32 & = & (142 - 148) & + & (148 - 150.32) \\ \text{Abweichung eines} & & \text{Abweichung des Meßwerts} & & \text{Abweichung des} \\ \text{Meßwerts vom} & = & \text{vom zugehörigen} & + & \text{Zellendurchschnitts vom} \\ \text{Gesamtdurchschnitt} & & \text{Zellendurchschnitt} & & \text{Gesamtdurchschnitt} \end{array}$$

Das wesentlich Neue bei der Zweiwegvarianzanalyse ist nun, daß wir die Abweichung eines Zellendurchschnitts vom Gesamtdurchschnitt weiter zerlegen können in 3 Komponenten:

- Die Abweichung des zugehörigen Zeilendurchschnitts vom Gesamtdurchschnitt (und das ist die Schätzung des Zeileneffekts),
- die Abweichung des zugehörigen Spaltendurchschnitts vom Gesamtdurchschnitt (=Schätzung des Spalteneffekts) und in
- die Abweichung des Zellendurchschnitts von dem Wert, der ohne Wechselwirkung allein aufgrund des zugehörigen Zeilen- und Spalteneffekts zu erwarten wäre (=Schätzung des Wechselwirkungseffekts).

Da wir auf diese Komponenten im folgenden häufiger zurückgreifen, ist es umständlich, immer von "Schätzung des ...-Effekts" zu sprechen. Andererseits sind aber diese Effekte aufgrund der wahren Durchschnitte definiert und nicht aufgrund der Stichprobendurchschnitte. Um Verwechslungen zu vermeiden, wollen wir vereinbaren: Wenn wir den Namen eines Effekts in

Apostrophe (') einschließen, dann ist jeweils der aufgrund der Stichprobendurchschnitte geschätzte Effekt gemeint. Und um auch für die Abweichung eines Zellendurchschnitts vom Gesamtdurchschnitt einen kurzen Namen zu haben, nennen wir sie sinngemäß 'Zelleneffekt'.

Dann können wir die obige Zerlegung des 'Zelleneffekts' z.B. für die Gruppe "Frauen unter Arbeitsbedingung nüchtern" folgendermaßen hinschreiben:

$$(148-150.32) = (161.71-150.32) + (149.33-150.32) + (148-160.72)$$

$$\text{'Zelleneffekt'} = \text{'Zeileneffekt'} + \text{'Spalteneffekt'} + \text{'Wechselwirkungseffekt'}$$

Dabei ist 160.72 der Wert, der sich ohne Wechselwirkung allein aufgrund des 'Zeileneffekts' und des 'Spalteneffekts' als Durchschnitt für Frauen unter Arbeitsbedingung nüchtern ergeben müßte. Die Berechnung dieses Werts (vgl. S. @) zeigt, warum diese Zerlegung gültig ist: Er ergibt sich dadurch, daß zum Gesamtdurchschnitt der 'Zeileneffekt' und der 'Spalteneffekt' addiert wird. Die obige Gleichung läßt sich dann auch folgendermaßen schreiben:

$$\begin{aligned} &\text{Zellendurchschnitt} - \text{Gesamtdurchschnitt} \\ &= \text{'Zeileneffekt'} + \text{'Spalteneffekt'} \\ &+ (\text{Zellendurchschnitt} - (\text{Gesamtdurchschnitt} + \text{'Zeileneff.'} + \text{'Spalteneff.'})) \end{aligned}$$

Und daß das nicht nur in unserem Fall, sondern ganz allgemein zutrifft, zeigt sich sofort, wenn man die Klammern in der letzten Zeile auflöst.

b) Die Zerlegung der Abweichungsquadratsumme

Zunächst können wir wieder - wie bei der Einweg-Varianzanalyse - die Quadratsumme der Abweichungen aller Meßwerte vom Gesamtdurchschnitt zerlegen in eine Quadratsumme innerhalb der Gruppe und eine Quadratsumme zwischen den Gruppen (Zellen). Letztere beruht auf den 'Zelleneffekten' und kann deshalb - ähnlich wie die 'Zelleneffekte' in 3 Komponenten zerlegt werden:

- Die Quadratsumme für Zeileneffekte ergibt sich, indem für jeden Meßwert der zugehörige 'Zeileneffekt' quadriert und aufsummiert wird.
- Die Quadratsumme für Spalteneffekte ergibt sich, indem für jeden Meßwert der zugehörige 'Spalteneffekt' quadriert und aufsummiert wird.
- Die Quadratsumme für Wechselwirkungseffekte ergibt sich, indem für jeden Meßwert der zugehörige 'Wechselwirkungseffekt' quadriert und aufsummiert wird.

Diese 3 Quadratsummen zusammen ergeben die Quadratsumme zwischen den Gruppen. Als

Formel:

$$QS_{zw} = QS_{Zeilen} + QS_{Spalten} + QS_{Wechselwirkung}$$

<< Daß das so funktioniert, kann man ganz ähnlich wie bei der Zerlegung der totalen Quadratsumme in der Einwegvarianzanalyse zeigen. Wenn man für einen einzelnen Wert die Summe

'Zeileneffekt' + 'Spalteneffekt' + 'Wechselwirkungseffekt'

quadriert, dann ergeben sich nach der Formel

$$(a+b+c)^2 = a^2 + b^2 + c^2 + 2ab + 2ac + 2bc$$

außer den Quadraten der 'Effekte' auch doppelte Produkte, die sich aber bei einer Summation über alle Vpn aufheben. Was bei einer Summation über alle Vpn übrigbleibt (nach dem Muster der letzten 2 Zeilen von Tabelle V3, Seite @), sind die Summen der quadrierten 'Effekte', wobei natürlich für jede Vp die für sie zutreffenden 'Effekte' zu verrechnen sind. Und diese Summen sind die o.g. Quadratsummen.

>>

c) Vereinfachte Berechnung der Quadratsummen.

<<

Da im Grunde alles läuft wie bei der Einwegvarianzanalyse, gibt es auch vereinfachte Berechnungsmöglichkeiten für die Quadratsummen.

Wir bilden die folgenden Ausdrücke:

RQ, die "rohe Quadratsumme", ist wieder die Summe aller quadrierten Rohwerte. In unserem Beispiel:

$$RQ = 142^2 + 149^2 + \dots + 122^2 = 2409572.00$$

Zu jeder Zelle, jeder Zeile, jeder Spalte und für alle Meßwerte zusammen ("gesamt") bildet man ein Korrekturglied. Aus den bereits bei der Einweg-Varianzanalyse genannten Gründen berechnet man diese Korrekturglieder, indem man die entsprechende Meßwertsumme quadriert und dieses Quadrat durch die Zahl der zugrundeliegenden Meßwerte dividiert. Am einfachsten ist es, wenn man zunächst die Meßwertsummen für die Zellen berechnet (vgl. Tabelle S. @) und dann die Meßwertsummen für Zeilen, Spalten und "Gesamt" durch Zusammenfassung bildet.

Aus diesen Korrekturgliedern bildet man Summen, indem man "Korrekturglieder gleicher Art" zusammenfaßt:

SCR ist die Summe der zu den Zeilen ("Reihen") gebildeten Korrekturglieder:

$$SCR = (2664+4128)^2:42 + (2520+2920)^2:35 + (1536+2016)^2:28 = 2394491.47$$

SCS ist die Summe der zu den Spalten gebildeten Korrekturglieder:

$$SCS = (2664+2520+1536)^2:45 + (4128+2920+2016)^2:60 = 2372788.27$$

SCZ ist die Summe der zu den Zellen gebildeten Korrekturglieder:

$$SCZ = 2664^2:18 + 4128^2:24 + 2520^2:15 + 2920^2:20 + 1536^2:12 + 2016^2:16 = 2404592.00$$

C ist wieder das aufgrund der Gesamtsumme aller Meßwerte gebildete Korrekturglied:

$$C = (2664+4128+2520+2920+1536+2016)^2:105 = 2372711.01$$

Aus diesen verschiedenen Ausdrücken können wir nun die Quadratsummen berechnen:

$$QS_{tot} = RQ - C = 2409572.00 - 2372711.01 = 36860.99$$

QS_{tot} wird zerlegt in

$$QS_{in} = RQ - SCZ = 2409572.00 - 2404592.00 = 4980.00$$

$$QS_{zw} = SCZ - C = 2400592.00 - 2372711.01 = 31880.99$$

QS_{zw} wird weiter zerlegt in

$$QS_{Zeilen} = SCR - C = 2394491.47 - 2372711.01 = 21780.46$$

$$QS_{Spalten} = SCS - C = 2372788.27 - 2372711.01 = 77.26$$

$$\begin{aligned} QS_{Wechselw.} &= SCZ - SCR - SCS + C \\ &= 2400592.00 - 2394491.47 - 2372788.27 + 2372711.01 \\ &= 10023.27 \end{aligned}$$

Als Rechenprobe kann man überprüfen, ob

$$QS_{tot} = QS_{in} + QS_{zw}$$

und

$$QS_{zw} = QS_{Zeilen} + QS_{Spalten} + QS_{Wechselw.}$$

Man sollte sich aber darüber klar sein, daß Rechen- oder Rundungsfehler bei der Berechnung der den Quadratsummen zugrundeliegenden Ausdrücke RQ, SCR, SCS, SCZ und C von dieser Rechenprobe nicht bemerkt werden.

>>

d) Die Freiheitsgrade der Varianzquellen.

Wie in der Einweg-Variananalyse ist auch hier jeder Varianzquelle eine Freiheitsgradzahl zugeordnet. In den folgenden Formeln ist R die Zahl der Zeilen ("Reihen"), S die Zahl der Spalten und Z die Zahl der Zellen.

$$df_{tot} = N - 1 = 105 - 1 = 104$$

$$df_{in} = N - Z = 105 - 6 = 99$$

$$\begin{aligned}df_{zw} &= Z - 1 = 6 - 1 = 5 \\df_{Zeilen} &= R - 1 = 3 - 1 = 2 \\df_{Spalten} &= S - 1 = 2 - 1 = 1 \\df_{Wechselw.} &= (R-1) \cdot (S-1) = (3-1) \cdot (2-1)=2\end{aligned}$$

Natürlich müssen wieder entsprechende Zerlegungen gelten wie für die Quadratsummen (Rechenprobe):

$$\begin{aligned}df_{tot} &= df_{in} + df_{zw} \\ \text{und} \\ df_{zw} &= df_{Zeilen} + df_{Spalten} + df_{Wechselw.}\end{aligned}$$

Natürlich verläuft auch die Begründung für die Freiheitsgradzahlen ganz ähnlich wie in der Einwegvarianzanalyse. Für df_{tot} , df_{in} und df_{zw} können wir die Begründung für die entsprechenden Varianzquellen der Einwegvarianzanalyse unmittelbar übernehmen.

Für df_{Zeilen} können wir die Begründung für df_{zw} aus der Einweg- Varianzanalyse ebenfalls sinngemäß übernehmen. Nach den in Abschnitt FI4e entwickelten Prinzipien könnten wir aus den Zeilendurchschnitten den Gesamtdurchschnitt berechnen, und nach den gleichen Prinzipien gilt dann auch die Restriktion: Die mit den Vpn-Zahlen pro Zeile multiplizierten Zeileneffekte müssen zusammen 0 ergeben. Also sind nur R-1 Zeileneffekte "frei beweglich"; sind sie vorgegeben, dann ist der letzte festgelegt.

Entsprechendes gilt auch für die Spalten.

Für die Freiheitsgrade der Wechselwirkung ($df_{Wechselw.}$) können wir zunächst postulieren, daß sie zusammen mit df_{Zeilen} und $df_{Spalten}$ die Freiheitsgrade zwischen den Zellen (df_{zw}) ergeben müssen, und wie man leicht nachprüfen kann, tun sie dies dann, wenn sie nach der obigen Formel berechnet werden.

<< Wer es noch genauer wissen will, kann es auch aufgrund von Restriktionen überprüfen. Für die 'Wechselwirkungseffekte' gelten zwei Restriktionen - eine Zeilenrestriktion und eine Spaltenrestriktion.

Die Zeilenrestriktion lautet: Innerhalb einer jeden Zeile müssen die mit der jeweiligen Gruppengröße multiplizierten 'Wechselwirkungseffekte' der verschiedenen Zellen zusammen 0 ergeben. (Wem es wichtig ist, das nachzuprüfen, kann es versuchen oder den Beweis nachlesen, der diesem Abschnitt als Anhang beigefügt ist.)

Die Spaltenrestriktion lautet entsprechend: Innerhalb einer jeden Spalte müssen die mit der jeweiligen Gruppengröße multiplizierten 'Wechselwirkungseffekte' der verschiedenen Zellen zusammen 0 ergeben. (Der Beweis würde ganz entsprechend aussehen wie für die Zeilenrestriktion.)

Beide Restriktionen zusammen kann man aber auch folgendermaßen formulieren: Innerhalb der ersten Zeilen (außer der letzten) sind nur die 'Wechselwirkungseffekte' bis auf den letzten "frei beweglich"; der letzte in jeder Zeile ist dann wegen der Zeilenrestriktion durch die vorangehenden festgelegt. Und die 'Wechselwirkungseffekte' der letzten Zeile sind wegen der Spaltenrestriktion festgelegt. Da wir aber insgesamt R Zeilen mit jeweils S 'Wechselwirkungseffekten' haben,

können wir auch sagen: In R-1 Zeilen sind S-1 'Wechselwirkungseffekte' frei beweglich. Und daher ist die Freiheitsgradzahl für Wechselwirkung (R-1) · (S-1).

>>

<< Anhang zu @d.: Beweis der Zeilenrestriktion für 'Wechselwirkungseffekte'.

Es ist zu zeigen, daß innerhalb jeder Zeile die 'Wechselwirkungseffekte' nach Multiplikation mit der jeweiligen Gruppengröße zusammen 0 ergeben müssen. Sehen wir es uns für die erste Zeile unseres Versuchsplans an. Nach der Definition eines 'Wechselwirkungseffekts' gilt zunächst:

$$N_{11} \cdot \hat{c}_{11} = N_{11} \cdot (\bar{x}_{11} - (\bar{x} + \hat{a}_{\text{nüchtern}} + \hat{b}_{\text{Frauen}}))$$

$$N_{11} \cdot \hat{c}_{12} = N_{12} \cdot (\bar{x}_{12} - (\bar{x} + \hat{a}_{\text{nüchtern}} + \hat{b}_{\text{Männer}}))$$

Addieren wir beide Gleichungen, dann haben wir auf der linken Seite die Summe der mit den jeweiligen Gruppengrößen multiplizierten 'Wechselwirkungseffekte' - und die soll also 0 sein. Das sehen wir, wenn wir auch die rechten Seiten addieren. Um bei der Addition der rechten Seiten zusammengehörige Glieder besser zusammenfassen zu können, machen wir uns klar:

1. Nach unseren Regeln über Durchschnitte von zusammengefaßten Gruppen (vgl. Abschnitt FI4e) muß gelten:

$$N_{11} \cdot \bar{x}_{11} + N_{12} \cdot \bar{x}_{12} = (N_{11} + N_{12}) \cdot \bar{x}_1.$$

2. Nach der Definition eines 'Zeileneffekts' muß gelten

$$\bar{x} + \hat{a}_{\text{nüchtern}} = \bar{x}_1.$$

3. Nach dem, was wir in Abschnitt FI4e über die Restriktion der 'Effekte' in zusammengefaßten Gruppen gesagt haben, muß gelten:

$$N_{.1} \cdot \hat{b}_{\text{Frauen}} + N_{.2} \cdot \hat{b}_{\text{Männer}} = 0$$

Multiplizieren wir diese Gleichung mit N_{11} und dividieren sie durch $N_{.1}$, dann erhalten wir:

$$N_{11} \cdot \hat{b}_{\text{Frauen}} + (N_{11} \cdot N_{.2} : N_{.1}) \cdot \hat{b}_{\text{Männer}} = 0$$

Nun ist aber nach der Orthogonalitätsvoraussetzung (vgl. S. @)

$$N_{11} \cdot N_{12} = N_{.1} \cdot N_{.2}$$

oder - was dasselbe bedeutet (Multiplikation beider Seiten mit $N_{12} \cdot N_{.2} : N_{.1}$) -

$$N_{11} \cdot N_{.2} : N_{.1} = N_{12}$$

Setzen wir das in die zweite Gleichung von "3."@ ein, dann erhalten wir:

$$N_{11} \cdot \hat{b}_{\text{Frauen}} + N_{12} \cdot \hat{b}_{\text{Männer}} = 0$$

Wenn wir nun also die rechten Seiten der vor 1. stehenden zwei Gleichungen addieren und dabei von den bisher gefundenen Beziehungen Gebrauch machen, um zusammenpassende Ausdrücke zusammenzufassen, dann ergibt sich:

$$(N_{11}+N_{12})\cdot\bar{x}_1 - (N_{11}+N_{12})\cdot\bar{x}_1 - (N_{11}\cdot\hat{b}_{\text{Frauen}} + N_{12}\cdot\hat{b}_{\text{Männer}})$$

und das ist (unter Berücksichtigung der in 3. gefundenen Gleichung) 0, wie zu beweisen war.

>>

e) Die Signifikanztests in der Zweiweg-Varianzanalyse.

Mit einer Zweiweg-Varianzanalyse können wir 3 Nullhypothesen überprüfen:

- a. Die den Zeilendurchschnitten entsprechenden Populationsmittel sind alle gleich. Anders formuliert: In der Population sind alle Zeileneffekte 0.
- b. Die den Spaltendurchschnitten entsprechenden Populationsmittel sind alle gleich. Anders formuliert: In der Population sind alle Spalteneffekte 0.
- c. In der Population existieren keine Wechselwirkungen; die Wechselwirkungseffekte in der Population sind also alle 0.

Für jede dieser Nullhypothesen wird ein eigener F-Test als Signifikanztest durchgeführt. Wie in der Einweg-Varianzanalyse wird dabei die zu der jeweiligen Nullhypothese gehörende Quadratsumme durch ihre Freiheitsgradzahl dividiert. Das resultierende Mittelquadrat wird durch MQ_{in} dividiert, und der daraus resultierende F-Wert ist wieder "seinem Wesen nach einseitig" zu interpretieren.

Die Begründung ist ebenfalls dieselbe wie bei der Einweg-Varianzanalyse: MQ_{in} ist immer eine erwartungstreue Schätzung der lt. Homogenitätsvoraussetzung in allen Zellen gleichen Populationsvarianz. Trifft eine der Nullhypothesen zu, dann ist auch das dazu gehörende Mittelquadrat eine erwartungstreue Schätzung dieser Populationsvarianz, und daher ist der Quotient aus diesem Mittelquadrat und MQ_{in} in diesem Fall F-verteilt. Trifft dagegen eine Nullhypothese nicht zu, dann nimmt das entsprechende Mittelquadrat größere Werte an, weil auch noch die entsprechende "wahre Varianz" (zwischen den Zeilen oder zwischen den Spalten oder Wechselwirkungsvarianz) in das Mittelquadrat eingeht, und zwar umso mehr, je größer die Stichproben. Daher ergeben sich bei Ungültigkeit einer Nullhypothese für den F- Wert, der aus dem zu dieser Hypothese gehörenden Mittelquadrat berechnet wurde, größere Werte, als per Zufall zu erwarten ist. (Und die Werte in der F-Tabelle geben ja an, was - unter dem entsprechenden Signifikanzniveau - per Zufall zu erwarten ist.)

<< Genauer lassen sich mit ähnlichen Methoden, wie sie in Abschnitt @FI8 für die Einweg-Varianzanalyse demonstriert wurden, die folgenden Gleichungen für die Erwartungswerte der Quadratsummen und Mittelquadrate in der Zweiweg-Varianzanalyse herleiten:

$$\begin{aligned}
 E(QS_{\text{Zeilen}}) &= (R-1) \cdot \sigma^2 + \sum_r N_r \cdot a_r^2 & E(MQ_{\text{Zeilen}}) &= \sigma^2 + \sum_r N_r \cdot a_r^2 / (R-1) \\
 E(QS_{\text{Spalten}}) &= (S-1) \cdot \sigma^2 + \sum_s N_s \cdot b_s^2 & E(MQ_{\text{Spalten}}) &= \sigma^2 + \sum_s N_s \cdot b_s^2 / (S-1) \\
 E(QS_{\text{Ww.}}) &= (R-1) \cdot (S-1) \cdot \sigma^2 + \sum_z N_z \cdot c_z^2 & E(MQ_{\text{Ww.}}) &= \sigma^2 + \sum_z N_z \cdot c_z^2 / ((R-1) \cdot (S-1)) \\
 E(QS_{\text{in}}) &= (N-R \cdot S) \cdot \sigma^2 & E(MQ_{\text{in}}) &= \sigma^2
 \end{aligned}$$

Dabei ist...

σ^2 = die (als identisch vorausgesetzte) Populationsvarianz innerhalb der Zellen.

N = Gesamtzahl aller Meßwerte (Vpn) in allen Gruppen zusammen.

R, S = Zahl der Zeilen ("Reihen") bzw. Spalten.

N_r, N_s, N_z = Zahl der Meßwerte in Zeile r bzw. Spalte s bzw. Zelle z.

a_r, b_s, c_z : Wahrer Haupt-Effekt der Zeile r bzw. Spalte s bzw. Wechselwirkungs-Effekt der Zelle z.

Wer die Herleitungen in Abschnitt @FI8 verstanden hat, mag selbst entscheiden, ob er die Übertragung des Ansatzes auf den Beweis der obigen Gleichungen selbst versuchen will. Ein Hinweis dazu: Die Formeln

$$E(RQ) = N \cdot \sigma^2 + \sum_g N_g \cdot \mu_g^2$$

und

$$E(C) = \sigma^2 + N \cdot \mu^2$$

für die Erwartungswerte der rohen Quadratsumme RQ und des Gesamt-Korrekturglieds C ergeben sich aus den gleichen Überlegungen wie bei der Einweg-Varianzanalyse. Die Herleitung der Formel

$$E(C_g) = \sigma^2 + N_g \cdot \mu_g^2$$

für das Korrekturglied der Gruppe g läßt sich ebenfalls sinngemäß übertragen, indem man Zeilen, Spalten und Zellen als "Gruppen" betrachtet. Unmittelbar einleuchtend ist das bei Zellen. Bei Zeilen und Spalten ist zu beachten, wie sich das wahre Zeilen- bzw. Spalten-Mittel aus den Zellenmitteln zusammensetzt, und dieser Wert ist dann als μ_g in die obige Formel für den Erwartungswert $E(C_g)$ einzusetzen.

>>

Die Ergebnisse der Berechnungen faßt man dann wieder in einer Tabelle zusammen. Für unser Experiment würde sie folgendermaßen aussehen:

| Quelle | QS | df | MQ | F |
|--------------------------|----------|-----|----------|-----------|
| Arbeitsbedingungen | 21780.46 | 2 | 10890.23 | 216.51*** |
| Geschlecht | 77.26 | 1 | 77.26 | 1.54* |
| Arbeitsbed. × Geschlecht | 10023.27 | 2 | 5011.64 | 99.63*** |
| <hr/> | | | | |
| zwischen Zellen | 31880.99 | 5 | | |
| innerhalb der Zellen | 4980.00 | 99 | 50.30 | |
| <hr/> | | | | |
| total | 36860.99 | 104 | | |

* : $p > .1$

***: $p < .01$

Zur Gestaltung dieser Tabelle sehen wir (zusätzlich zu den schon bei der Einweg-Varianzanalyse behandelten Regeln):

1. Die Quellen werden bei der tabellarischen Zusammenfassung möglichst mit einem inhaltlichen Namen benannt ("Arbeitsbedingungen" und "Geschlecht" statt "Zeilen" und "Spalten"). Das erleichtert die Lesbarkeit.

2. Für die Wechselwirkung verbindet man die inhaltlichen Bezeichnungen der Faktoren, um deren Wechselwirkung es geht, mit einem \times -Zeichen (für "mal") und spricht von der "Wechselwirkung Arbeitsbedingungen mal Geschlecht". Wichtig ist das vor allem dann, wenn es - bei Varianzanalysen mit mehr als 2 Faktoren, wie wir sie im nächsten Kapitel kennenlernen - mehrere Wechselwirkungen gibt.

<< Genauer betrachtet, ist das Zeichen \times hier als Symbol für das "cartesische Produkt" zweier Mengen zu verstehen. Das cartesische Produkt der Menge der Arbeitsbedingungen und der Menge der Geschlechter besteht aus den "geordneten Paaren" (nüchtern, Frauen), (nüchtern, Männer), (hübsch, Frauen) usw., also aus allen Paaren, bei denen an erster Stelle eine Arbeitsbedingung und an zweiter Stelle ein Geschlecht steht. Bei der Wechselwirkung geht es ja gerade um die Effekte, die solche Bedingungskombinationen noch über die Effekte der beiden einzelnen Bedingungen hinaus haben.

>>

Nun zu den Ergebnissen. In unserem Experiment haben wir zunächst einen signifikanten Effekt der Arbeitsbedingungen: Die Nullhypothese, daß in der Population alle Arbeitsbedingungen im Durchschnitt zu gleichen Ergebnissen führen, kann mit extrem niedriger Irrtumswahrscheinlichkeit verworfen werden.

Allerdings ist es auch hier wie bei der Einweg-Varianzanalyse: Wenn wir die Nullhypothese verworfen, daß nicht alle Populationsmittel gleich sind, dann heißt das noch nicht, daß jedes von

jedem verschieden ist. Es können immer noch einzelne untereinander gleich sein. (In unserem Fall könnte es z.B. durchaus sein, daß die Populationsmittel für die Arbeitsbedingungen "hübsch" und "nüchtern" gleich sind.) Wenn man genaue Aussagen über den Unterschied zweier bestimmter Zeilenmittel in der Population machen wollte, müßte man wieder post-hoc-Tests anwenden.

Auch für die Wechselwirkung gibt es ein signifikantes F. Es liegt also eine signifikante Wechselwirkung vor.

Auch dazu ist aber eine Einschränkung erforderlich: Je nachdem, wie man die Wechselwirkung betrachtet, lautet die entsprechende Nullhypothese: "Die Effekte der verschiedenen Arbeitsbedingungen sind bei Männern und Frauen gleich", oder: "Die Effekte des Geschlechts sind bei allen Arbeitsbedingungen gleich". Ein signifikantes Ergebnis besagt auch bei der Wechselwirkung lediglich, daß die Nullhypothese nicht in vollem Umfang zutrifft: Nicht alle Effekte der Arbeitsbedingungen sind bei Männern und Frauen gleich - bzw. der Effekt des Geschlechts ist nicht bei allen Arbeitsbedingungen gleich. Es kann aber durchaus sein, daß der Effekt zweier Arbeitsbedingungen bei Männern und Frauen gleich ist - bzw. daß der Effekt des Geschlechts bei 2 Arbeitsbedingungen gleich ist. Auch hier gilt: Wenn man sich mit der globalen Ablehnung der Nullhypothese nicht begnügt, dann muß man post-hoc-Tests rechnen, um jede einzelne Aussage über Unterschiede, die man treffen will, zu überprüfen.

Nicht signifikant ist dagegen der Effekt des Geschlechts. Die Nullhypothese, daß Männer und Frauen gleich gut sind, kann nicht verworfen werden.

Auch dazu ist eine einschränkende Anmerkung erforderlich: Das Ergebnis gilt (soweit ein nicht signifikantes Ergebnis überhaupt eins ist) nur dann, wenn die Häufigkeit, mit der die verschiedenen Arbeitsbedingungen in der Population bei Männern und Frauen auftreten, die gleichen sind, wie in unserem Experiment, also

$$\text{nüchtern} : \text{hübsch} : \text{altmodisch} = 42 : 35 : 28 = 6 : 5 : 4$$

Würde dagegen z.B. in der Population wesentlich häufiger unter Arbeitsbedingung "hübsch" als unter Arbeitsbedingung "nüchtern" gearbeitet, dann wären die Frauen (die ja unter Arbeitsbedingung "hübsch" wesentlich besser sind als die Männer) auch insgesamt besser.

Ein entsprechender Gedanke gilt allgemein: Die Populationsmittel, über die eine Nullhypothese über einen Haupteffekt in der Zweiweg- Varianzanalyse eine Aussage macht, werden immer unter der Voraussetzung betrachtet, daß die Häufigkeiten, mit denen die Stufen des anderen Faktors besetzt sind, repräsentativ für die Population sind, um die es geht. Wenn - wie in unserem Fall - Wechselwirkungen vorliegen, dann kann von diesen Häufigkeiten das Auftreten oder Nicht-Auftreten eines Haupteffekts abhängen.

<< Auch hier sind "halbseitige" Entscheidungsregeln möglich (vgl. S.@).
Beispiel: Hätten wir die Entscheidungsregel formuliert, daß wir die Spalten- H_0 nur dann verwerfen, wenn wir ein genügend hohes F erhalten und wenn außerdem die Frauen einen höheren Durchschnitt erzielen als die Männer, so könnten wir zur Ermittlung der Irrtums-

wahrscheinlichkeit die Überschreitungswahrscheinlichkeit des F -Wertes halbieren.

>>

Inhaltsverzeichnis dieses Auszugs:

| | |
|--------------------------------------------------------------------------------|---------|
| II. Die Mehrweg-Varianzanalyse | - 295 - |
| 1) Ein Experiment zur Zweiweg-Varianzanalyse | - 295 - |
| 2) "Faktoren" und "Stufen" | - 296 - |
| a) Begriffe | - 296 - |
| b) Die Unterscheidung von Zeilen und Spalten | - 297 - |
| c) Die Verwendung der Indices in der Zweiweg-Varianzanalyse | - 297 - |
| 3) Effekte in der Zweiweg-Varianzanalyse | - 298 - |
| a) Die Durchschnitte | - 298 - |
| b) Effekte des Zeilenfaktors | - 299 - |
| c) Effekte des Spaltenfaktors | - 300 - |
| d) Wechselwirkung | - 300 - |
| e) Allgemeines Modell der Zweiweg-Varianzanalyse | - 305 - |
| 4) Die Zusatzvoraussetzungen der Zweiweg-Varianzanalyse | - 306 - |
| a) Unabhängige Stichproben | - 306 - |
| b) Normalverteilung und Varianzhomogenität | - 306 - |
| c) Orthogonalität | - 306 - |
| d) Eine letzte Voraussetzung | - 309 - |
| 5) Grundlagen und Durchführung der Berechnungen in der Zweiweg-Varianzanalyse. | - 309 - |
| a) Die Zerlegung der Abweichungen vom Gesamtdurchschnitt | - 309 - |
| b) Die Zerlegung der Abweichungsquadratsumme | - 310 - |
| c) Vereinfachte Berechnung der Quadratsummen. | - 311 - |
| d) Die Freiheitsgrade der Varianzquellen. | - 312 - |
| e) Die Signifikanztests in der Zweiweg-Varianzanalyse. | - 315 - |