

## 1. Kapitel

# Modelle zur Beschreibung statistischer Zusammenhänge in der psychologischen Forschung

*Helfried Moosbrugger*

### *1. Einführung und Überblick*

In diesem Kapitel geben wir eine Einführung in Bereiche der angewandten Statistik, die bereits um die Jahrhundertwende betreten wurden (siehe z.B. Pearson 1896, Spearman 1904), aber bis heute in einer intensiven Entwicklung stehen, deren Ende noch nicht abzusehen ist. Wir behandeln in Form eines integrierenden Überblicks Theorie und Anwendung von Modellen zur Beschreibung statistischer Zusammenhänge.

Beobachtung und Beschreibung von Zusammenhängen haben in der psychologischen Forschung zumindest zwei Funktionen: zum einen entsteht dadurch ein Anlaß, induktiv die Entwicklung inhaltlicher Theorien zu betreiben, weil beobachtete statistische Zusammenhänge einer theoretischen Erklärung bedürfen; zum anderen sind beobachtete statistische Zusammenhänge gewissermaßen die empirischen Gegebenheiten, an denen deduktiv die Theorien auf ihre Gültigkeit hin überprüft werden. Oft sind dabei die empirischen Beobachtungen selbst schon auf theoretische Vorannahmen gegründet: Stellt man beispielsweise bei einer empirischen Untersuchung eine Korrelation zwischen „Intelligenz“ und „Schulleistung“ fest, so hat man vorher implizit im Sinne von Spearman (1904, 1926) die theoretische Entscheidung getroffen, daß „Intelligenz“ eine einheitliche latente Variable sei, die mit einem psychometrischen Test gemessen werden kann. Dieses Vorgehen ist keineswegs unstrittig, weil von anderen Autoren (z.B. Thurstone 1938, Guilford 1967) in Frage gestellt wird, daß es sich bei Intelligenz um eine einheitliche Eigenschaft handle, welche mit dem Intelligenzquotienten als Kennzahl darstellbar ist. Entsprechendes gilt auch für „Schulleistung“. Unter diesem Blickwinkel müssen selbst „empirische“ beobachtete Zusammenhänge bereits als „theoriebehaftet“ (vgl. dazu auch Holzkamp, 1972, S. 81) angesehen werden.

Während wir uns im folgenden Abschnitt mit Modellen befassen, welche die statistischen Zusammenhänge manifester, d.h. direkt beobachtbarer Variablen beschreiben, beschäftigen wir uns im darauffolgenden Abschnitt mit Situationen, in denen beobachtbare Zusammenhänge von manifesten Variablen auf nicht direkt beobachtbare „latente“ Variablen zurückgeführt werden. Solche latenten Variablen sind häufig Begriffe, welche in psychologischen Theorien verwendet werden. „Intelligenz“ z.B. kann nicht als direkt beobachtbar aufgefaßt werden. Beobachtbar sind vielmehr nur Verhaltensweisen, die auf unterschiedliche Ausprägung der Intelligenz schließen lassen. Ebenso verhält es sich mit anderen „Persönlichkeitseigenschaften“. Daher geben wir im dritten Abschnitt eine Einführung in die wesentlichen Grundgedanken der Theorie latenter Variablen, welche insofern die formale Grundlage vieler psychologischer Begriffe bildet, als sie angibt, wie latente Eigenschaften mit beobachtbarem Verhalten verknüpft werden können. Exemplarisch behandeln wir den Fall einer latenten Variablen, welche die Abhängigkeiten zwischen den beobachtbaren Variablen erklärt in dem Sinn, daß deren Abhängigkeiten verschwinden, wenn die gemeinsame latente Variable auf einem festen Wert gehalten wird. Beobachtete Variablen mit dieser Eigenschaft heißen kongenerisch (Jöreskog, 1971, Steyer & Moosbrugger, 1979).

Die formale Gemeinsamkeit aller in diesem Kapitel behandelten statistischen Modelle besteht darin, daß sie alle mit dem Konzept der bedingten Erwartung formuliert werden können, für die im Anhang C einige Rechenregeln zusammengestellt sind. Die bedingte Erwartung  $E(y \mid x)$  beispielsweise ist eine stochastische Variable oder Zufallsvariable<sup>1)</sup>, deren Werte  $E(y \mid x=x)$  die Erwartungswerte der abhängigen stochastischen Variablen  $y$  unter der Bedingung sind, daß die „bedingende“ Variable  $x$  einen bestimmten Wert  $x$  angenommen hat. Im Grunde beschreiben die hier behandelten Modelle immer nur in einer mathematischen Formulierung unsere Vorstellung, welche Werte der Variablen  $y$  bei gegebenen Werten der Variablen  $x$  zu erwarten sind. Diese einfache Grundidee kann zu recht komplizierten Modellen mit mehreren  $x$ - und  $y$ -Variablen ausgebaut werden.

Im Abschnitt 2 dieses Kapitels stellen wir statistische Modelle mit manifesten Variablen dar, die alle mit der Gleichung

$$(1.1) \quad E(y \mid x_1, x_2, \dots, x_Q) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_Q x_Q$$

formuliert werden können, wobei  $y$  sowie die  $x_q$ ,  $q \in \{1, 2, \dots, Q\}$  stochastische Variablen und die Parameter  $\beta_0$  und  $\beta_q$  reellwertige Konstanten sind. Statt

<sup>1)</sup> Wir gebrauchen „stochastische Variable“ synonym zu „Zufallsvariable“ (auf dem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$ ), worunter man eine  $\mathcal{A}$ -meßbare Abbildung von  $\Omega$  nach  $\mathbb{R}$ , der Menge der reellen Zahlen versteht, wobei  $\Omega$  die Menge der Elementarereignisse,  $\mathcal{A}$  die (Ereignis-)  $\sigma$ -Algebra und  $P$  ein Wahrscheinlichkeitsmaß ist (vgl. Bauer, 1974, s. 137).

$E(y | x_1, x_2, \dots, x_Q)$  schreiben wir auch  $E(y | \mathbf{x})$  mit dem  $Q$ -variaten stochastischen Vektor<sup>2)</sup>  $\mathbf{x} = (x_1, x_2, \dots, x_Q)$ .

Wie wir zeigen werden, liegt Gleichung 1.1 einer Reihe von Analyseverfahren zugrunde, welche sich nur hinsichtlich der Skalenqualität der beteiligten Variablen unterscheiden: In der Regressionsanalyse sind sowohl  $y$  als auch die  $x_q$  quantitative Variablen, und  $E(y | \mathbf{x})$  ist somit die Regressionsgerade, -ebene oder -hyperebene, je nachdem, wie viele Variablen in  $\mathbf{x}$  zusammengefaßt sind. In der Varianzanalyse ist  $y$  quantitativ, aber die  $x_q$  sind qualitative zweistufige Indikatorvariablen, welche die Gruppenzugehörigkeit anzeigen; die Werte von  $E(y | \mathbf{x})$  sind dann Erwartungswerte der abhängigen Variablen in den Gruppen. In der Diskriminanzanalyse mit zwei Gruppen ist  $y$  eine Indikatorvariable für die Gruppenzugehörigkeit, wohingegen die  $x_q$  quantitativ, aber auf ein bestimmtes Intervall beschränkt sind, wenn man auch dort die Gleichung 1.1 zugrunde legt. Kodiert man die Zugehörigkeit zu den beiden Gruppen mit Null und Eins, so hat  $E(y | \mathbf{x})$  die bedingten Wahrscheinlichkeiten  $P(y=1 | \mathbf{x}=\mathbf{x})$  als Werte, nämlich die Wahrscheinlichkeiten bei gegebenen Werten der Variablen  $x_q$ , der Gruppe „1“ anzugehören.

In der Kontingenzanalyse schließlich sind sowohl  $y$  als auch  $x_q$  qualitative Indikatorvariablen, und die Kodierung von  $y$  mit Null und Eins führt wieder dazu, daß  $E(y | \mathbf{x})$  bedingte Wahrscheinlichkeiten als Werte enthält, aber diesmal unter einer Gruppenzugehörigkeit und nicht, wie bei der Diskriminanzanalyse, unter der Ausprägung quantitativer Variablen  $x_q$ .

Man kann schon hier erkennen, daß alle genannten Verfahren auf der gleichen Modellannahme basieren und sich lediglich durch das Skalenniveau der beteiligten Variablen unterscheiden.

Im Abschnitt 3 behandeln wir ebenfalls Modelle, denen eine Gleichung der Form der Gleichung 1.1 zugrundeliegt, jedoch handelt es sich dort bei der unabhängigen bedingenden Variablen nicht um eine direkt beobachtbare, sondern um eine latente Variable, die wir mit dem kleinen griechischen Buchstaben  $\xi$  bezeichnen. Der Übersichtlichkeit halber beschränken wir uns in jenem dritten Abschnitt auf eine latente Variable. Dafür ist es aber notwendig, zugleich mehrere abhängige Variablen  $y_p$  zu betrachten. Den meisten Modellen jenes Abschnitts liegt die Gleichung

$$(1.2) \quad E(y_p | \xi) = \lambda_{p0} + \lambda_p \xi \quad p \in \{1, 2, \dots, P\}$$

zugrunde, wobei  $\lambda_{p0}$  wiederum eine reellwertige Skalenkonstante ist und  $\lambda_p$  ein reellwertiger Parameter, der ebenso wie die Konstanten  $\beta_q$  (siehe Gleichung 1.1) die Enge des statistischen Zusammenhangs zwischen den betreffenden Variablen angibt.

<sup>2)</sup> Ein  $Q$ -variater stochastischer Vektor (oder Zufallsvektor) ist eine  $d$ -meßbare Abbildung von  $\Omega$  nach  $\mathbb{R}^Q$  (vgl. auch Fußnote 1).

Nach der Entwicklung einer allgemeinen Theorie latenter Variablen werden wir zeigen, daß auch Gleichung 1.2 verschiedenen, eng miteinander verwandten Modellen mit latenten Variablen zugrunde liegt, welche sich nur dadurch unterscheiden, daß die beteiligten Variablen quantitativ bzw. qualitativ sind: In der Faktorenanalyse mit einem Faktor, welche formal als Analogon zur Regressionsanalyse aufgefaßt werden kann, sind sowohl die  $y_p$  als auch  $\xi$  quantitativ, und  $E(y_p|\xi)$  entspricht einer Regressionsgeraden. Sind die  $y_p$  wieder Indikatorvariablen für jeweils zwei Gruppen, wohingegen  $\xi$  quantitativ, aber auf ein bestimmtes Intervall beschränkt ist, so erhalten wir Lazarsfeld's „*linear traceline model*“, wobei  $E(y_p|\xi)$  die „traceline“ bezeichnet und wieder bedingte Wahrscheinlichkeiten  $P(y_p=1|\xi=\xi)$  als Werte aufweist, wenn man als Kodierung für die  $y_p$  Null und Eins wählt. In formaler Hinsicht entspricht dieses Modell daher der Diskriminanzanalyse. Sind die  $y_p$  quantitativ und ist  $\xi$  eine Variable, welche die Zugehörigkeit zu einer von zwei latenten Gruppen anzeigt, so erhalten wir Gibsons „*latent profile model*“. Die bedingte Erwartung  $E(y_p|\xi)$  enthält dabei die Erwartungswerte der  $y_p$  innerhalb der jeweiligen latenten Gruppe. Formal ist dieses Modell analog zur Varianzanalyse. Sind schließlich alle beteiligten Variablen qualitativ mit der Kodierung Null und Eins, so finden wir, daß Gleichung 1.2 das „*latent class model*“ von Lazarsfeld beschreibt, und  $E(y_p|\xi)$  hat wieder die bedingten Wahrscheinlichkeiten  $P(y_p=1|\xi = \xi)$  als Werte. Dieses Modell entspricht also in formaler Hinsicht die Kontingenzanalyse. Auch „logistische Testmodelle“ wie z.B. das „*Rasch-Modell*“, und das „*klassische latent-additive Testmodell*“ von Moosbrugger & Müller, welche keine unmittelbare Entsprechung bei den Modellen mit manifesten Variablen haben, werden wir als Anwendungsfälle der Theorie latenter Variablen referieren.

Bei allen Gemeinsamkeiten, die in formaler Hinsicht zwischen den Modellen der zweiten und denjenigen des dritten Abschnitts bestehen, darf man jedoch nicht die Unterschiedlichkeit übersehen, die in ihren Zielen bestehen. Während die Modelle des zweiten Abschnitts den Zweck haben, statistische Zusammenhänge zu beschreiben, dienen die Modelle des dritten Abschnitts gewissermaßen der Bildung von psychologischen Begriffen. Die jeweiligen latenten Variablen  $\xi$  lassen sich nämlich als psychologische Konstrukte (s. z.B. Herrmann, 1973 oder MacCorquodale & Meehl, 1948) interpretieren.

Die Modelle latenter Variablen können als eine Art der Erklärung von statistischen Zusammenhängen aufgefaßt werden, da die Zusammenhänge zwischen den abhängigen Variablen  $y_p$ , sofern sie „kongenerisch“ sind, bei jeder Ausprägung von  $\xi$  verschwinden. In diesem Sinn kann man sagen, daß die latente Variable  $\xi$  dem Modell nach die statistischen Abhängigkeiten zwischen den beobachtbaren Variablen  $y_p$  erklärt. Eine grundlegende Behandlung von Modellen zur Erklärung statistischer Zusammenhänge nimmt Steyer (1982) in diesem Band vor.

## 2. Modelle mit manifesten Variablen

### 2.1 Einleitung und Überblick

Im vorliegenden Abschnitt 2 behandeln wir die Regressions-, Varianz-, Diskriminanz- und Kontingenzanalyse. Dabei zeigen wir, daß die Analyseverfahren in formaler Hinsicht grundlegende Gemeinsamkeiten aufweisen. Die verschiedenen Namen dieser Verfahren sind eher historisch zu verstehen, ebenso wie ihre getrennte Behandlung in den Lehrbüchern der angewandten Statistik. Die Worte „Varianz-“, „Regressionsanalyse“ etc. sind eher als Kennzeichnung für die Fragestellungen aufzufassen, bei denen diese Verfahren angewandt werden. Alle in diesem Abschnitt behandelten Modelle präzisieren Vorstellungen über statistische Zusammenhänge direkt beobachtbarer oder „manifesten“ Variablen.

Statistische Zusammenhänge können erstens zur Prädiktion verwendet werden, was ein vorrangiges Ziel in der „*Regressionanalyse*“ ist. Nachdem man in einer Stichprobe das „Zusammenauftreten“ zweier Variablen beobachtet hat, kann man in anderen Stichproben den Wert der einen Variablen aus dem Wert der anderen vorhersagen. Die Stärke statistischer Zusammenhänge kann zweitens aber auch für sich genommen von Interesse sein. Sie wird in der Korrelationsanalyse untersucht. Drittens sind statistische Zusammenhänge zwischen Gruppen einerseits - welche als qualitative Variablen kodierbar sind - und quantitativen Variablen andererseits ein Ausdruck von Unterschieden zwischen den Gruppen hinsichtlich der Mittelwerte der quantitativen Variablen. Die Untersuchung von Mittelwerteunterschieden ist unter den Begriffen „*t-Test*“, „*Varianzanalyse*“ und „*Mittelwertvergleiche*“ bekannt. Viertens kann man statistische Zusammenhänge zwischen Gruppen und quantitativen Variablen zur Vorhersage der Gruppenzugehörigkeit nutzen, sofern der Wert der quantitativen Variablen bekannt ist. Untersuchungsmethoden mit diesem Ziel heißen „*Diskriminanzanalyse*“. Liegen nur qualitative Variablen vor, so kann man fünftens auftretende Häufigkeiten von Merkmalen mit der „*Kontingenzanalyse*“ auf statistische Abhängigkeit hin untersuchen.

Wir zeigen, daß alle fünf Verfahren zur Beschreibung von Zusammenhängen zwischen manifesten (d.h. beobachtbaren) Variablen auf der gleichen formalen Theorie basieren, die wir in Abschnitt 2.2 darstellen. Dabei gehen wir nicht auf Stichprobenprobleme wie Schätzverfahren und Hypothesenbeurteilung (z.B. durch Signifikanztests) ein. Diese sind Gegenstand der nachfolgenden Kapitel dieses Bandes.

## 2.2 Eine formale Theorie zur Beschreibung statistischer Zusammenhänge

### 2.2.1 Die Grundannahme

Die in diesem Abschnitt behandelte Klasse von Modellen zur Beschreibung des statistischen Zusammenhangs zwischen einer stochastischen Variablen  $y$  und mehreren stochastischen Variablen  $x_q$ ,  $q \in \{1, 2, \dots, Q\}$  basieren auf folgender linearen Gleichung

$$(2.1) \quad E(y \mid x_1, x_2, \dots, x_Q) = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \dots + \beta_Q \cdot x_Q$$

oder in Vektorschreibweise

$$E(y \mid \mathbf{x}) = \beta_0 + \boldsymbol{\beta}'\mathbf{x},$$

wobei wir voraussetzen, daß alle beteiligten stochastischen Variablen auf dem gleichen Wahrscheinlichkeitsraum (s. Fußnoten 1 und 2) definiert sind und eine endliche Erwartung haben. Gleichung 2.1 besagt (siehe auch Anhang C), daß die bedingte Erwartung von  $y$  bei gegebenen  $x_1, x_2, \dots, x_Q$  eine Summe von  $\beta_0$  und einer Linearkombination der  $x_q$  ist. Die Variable  $y$  bezeichnet man häufig als abhängige Variable oder Kriteriumsvariable, die Variablen  $x_q$  als unabhängige Variablen oder Prädiktorvariablen<sup>3</sup>).  $\beta_0$  ist eine Skalenkonstante, die angibt, welchen bedingten Erwartungswert die abhängige Variable  $y$  hat, wenn alle unabhängigen Variablen  $x_q$  den Wert Null haben. Die reellwertigen Koeffizienten  $\beta_q$ ,  $q \in \{1, 2, \dots, Q\}$  sind genau dann gleich Null, wenn kein durch Gleichung 2.1 beschriebener statistischer Zusammenhang zwischen  $y$  und  $x_q$  besteht. Die  $\beta_q$  wachsen in ihrem Absolutbetrag mit der „Enge“ des statistischen Zusammenhangs zwischen  $y$  und  $x_q$  bis zu einem Höchstbetrag, der von den Varianzen der beteiligten Variablen abhängig ist. Man beachte jedoch, daß diese vereinfachende Interpretation ihre Gültigkeit verliert, wenn „Suppressorvariablen“ (vgl. z.B. Darlington, 1968, S. 164, Fischer, 1974, S. 75, Conger & Jackson, 1972, Conger 1974) vorliegen.

### 2.2.2 Die Residualvariable

Statistische Zusammenhänge sind in der Regel nicht deterministisch. Zwischen der bedingten Erwartung  $E(y \mid \mathbf{x})$  und  $y$  selbst wird daher meist ein Unterschied bestehen. Diese Differenz, nämlich

---

<sup>3</sup>) Moosbrugger (1980) hat Gleichung 2.1 für nichtlineare multiple Variablenzusammenhänge mit Wechselwirkungen erweitert. Die bedingte Erwartung  $E(y \mid \mathbf{x}) = \beta_0 + \boldsymbol{\beta}'\mathbf{z}$  ist dann eine Linearkombination der  $R \geq Q$  unabhängigen Variablen  $z_{r,r} \in \{1, \dots, R\}$ , welche beliebige Funktionen der bedingenden Variablen  $x_q$  sein können.

$$(2.2) \quad \varepsilon = y - E(y|\mathbf{x})$$

bezeichnen wir als Fehler- oder Residualvariable von  $y$ . Für die abhängige Variable  $y$  folgt dann aus den Gleichungen 2.1 und 2.2

$$(2.3) \quad y = \beta_0 + \boldsymbol{\beta}'\mathbf{x} + \varepsilon.$$

Für die Residualvariable von  $y$  gilt allgemein

$$(2.4) \quad E(\varepsilon|\mathbf{x}) = 0,$$

(siehe Anhang C, Regeln 4, 5 und 10), was bedeutet, daß die bedingten Erwartungswerte der Residualvariablen bei allen Ausprägungskombinationen der  $x_q$  gleich Null sind. Ohne diese Eigenschaft würde die stochastische Variable  $E(y|\mathbf{x})$  als Werte nicht die bedingten Erwartungswerte für  $y$  enthalten.

Eine weitere Eigenschaft der Residualvariablen ist, daß auch ihr unbedingter Erwartungswert

$$(2.5) \quad E(\varepsilon) = E[y - E(y|\mathbf{x})] = 0$$

gleich Null ist (siehe Anhang C, Regel 9), und daß die Kovarianzen zwischen allen unabhängigen Variablen  $x_q$  und der Residualvariablen  $\varepsilon$

$$(2.6) \quad C(\mathbf{x}, \varepsilon) = 0$$

gleich Null sind (siehe Anhang C, Regel 11). Man beachte, daß mit Gleichung 2.6 weder gefordert ist, daß die bedingte Fehlervarianz  $V(\varepsilon|\mathbf{x}=\mathbf{x})$  für alle Ausprägungskombinationen der  $x_q$  gleich ist, noch irgendeine spezielle Verteilungsannahme getroffen wird.

### 2.2.3 Kovarianzmodellgleichungen und Parameteridentifikation

Nachdem wir in Gleichung 2.1 ein statistisches Modell formuliert haben, stellt sich nun die Frage, ob die darin enthaltenen Parameter - in unserem Fall sind das  $\beta_0$  und die Komponenten von  $\boldsymbol{\beta}$  - eindeutig aus beobachtbaren Verteilungskennwerten der beteiligten stochastischen Variablen bestimmt (identifiziert) werden können. Im Fall eines Modells vom Typ der Gleichung 2.1 reicht die Betrachtung der Modellstruktur für den Kovarianzvektor  $C(x,y)$  sowie für die Varianzen  $V(y)$  und  $V(\varepsilon)$  aus.

Die Gleichungen 2.3 und 2.6 implizieren die folgenden Modellstrukturgleichungen

a) für den Kovarianzvektor von  $x_q$  und  $y$

$$(2.7) \quad C(\mathbf{x}, y) = C[\mathbf{x}, (\beta_0 + \boldsymbol{\beta}'\mathbf{x} + \varepsilon)] = C(\mathbf{x}, \mathbf{x})\boldsymbol{\beta}$$

wobei  $C(x, x)$  die Kovarianzmatrix der  $x_q$  bezeichnet (siehe Anhang B, Regeln 1 und 2), und

b) für die Varianz von  $y$

$$(2.8) \quad V(y) = C[(\beta_0 + \boldsymbol{\beta}'\mathbf{x} + \varepsilon), (\beta_0 + \boldsymbol{\beta}'\mathbf{x} + \varepsilon)] = \boldsymbol{\beta}'C(\mathbf{x}, \mathbf{x})\boldsymbol{\beta} + V(\varepsilon)$$

(siehe Anhang B, Regeln 1 und 2).

Unter der zusätzlichen Annahme, daß  $C(x, x)$  invertierbar ist, erhalten wir die Identifikationsgleichung für die Koeffizienten  $\beta_q$  aus Gleichung 2.7

$$(2.9) \quad \boldsymbol{\beta} = C(\mathbf{x}, \mathbf{x})^{-1} C(\mathbf{x}, y)$$

und jene für die Varianz der Residualvariablen  $\varepsilon$  aus den Gleichungen 2.8 und 2.9

$$(2.10) \quad V(\varepsilon) = V(y) - \boldsymbol{\beta}'C(\mathbf{x}, \mathbf{x})\boldsymbol{\beta}.$$

Wenn die Parameter  $\beta_q$  und die Residualvarianz  $V(\varepsilon)$  eindeutig bestimmt sind, läßt sich auch die unbekannte Skalenkonstante  $\beta_0$  leicht berechnen: Aus den Gleichungen 2.3 und 2.5 folgt nämlich nach Umformen (siehe Anhang A, Regel 2)

$$(2.11) \quad \beta_0 = E(y) - \boldsymbol{\beta}'E(\mathbf{x})$$

als Differenz des Erwartungswertes von  $y$  und den mit den Koeffizienten  $\beta_q$  gewichteten Erwartungswerten der  $x_q$ .

Lassen sich wie hier alle unbekanntes Modellparameter aus den Erwartungswerten, Varianzen und Kovarianzen der beobachtbaren Variablen eindeutig bestimmen, so bezeichnet man das Modell als identifiziert. Man beachte, daß es sich bei den Gleichungen 2.9 bis 2.11 um Identifikationsgleichungen handelt. Schätzgleichungen, in denen analoge Stichprobenkennwerte verwendet werden, finden sich z.B. in Moosbrugger (1978, S. 56 und 61). über andere Schätzmethoden informieren z.B. Späth (1973, 1974), Harter (1974a, b, 1975a, b, c), Hoerl & Kennard (1970) oder Humak (1977).

Wenn  $C(x, x)$  nicht invertierbar ist, kann eine allgemeine Lösung für  $\boldsymbol{\beta}$  immer durch

$$(2.12) \quad \boldsymbol{\beta} = C(\mathbf{x}, \mathbf{x})^- C(\mathbf{x}, y)$$

gefunden werden, wobei  $C(x, x)^-$  die verallgemeinerte Inverse (vgl. z.B. Rao & Mitra, 1971) der Kovarianzmatrix  $C(x, x)$  ist. Eine solche Lösung wird im

allgemeinen jedoch erst durch zusätzliche Nebenbedingungen eindeutig (siehe z.B. Moosbrugger & Steyer (1982) in diesem Band).

#### 2.2.4 Determinierte Varianz, multiple Korrelation und Determinationskoeffizient

Wir bezeichnen mit  $V [E(y | x)]$  die durch  $x$  determinierte Varianz von  $y$ . Unter Verwendung von Gleichung 2.1 erhalten wir die Beziehung

$$(2.13) \quad V [E(y | x)] = C(\mathbf{\beta}'x, \mathbf{\beta}'x) = \mathbf{\beta}'C(x, x) \mathbf{\beta}.$$

Setzen wir Gleichung 2.13 in Gleichung 2.10 ein, so sehen wir, daß die determinierte Varianz die Differenz von Gesamtvarianz und Residualvarianz ist:

$$(2.14) \quad V [E(y | x)] = V(y) - V(\varepsilon).$$

Dividieren wir Gleichung 2.14 durch  $V(y)$ , so erhalten wir unter Verwendung von Gleichung 2.13 und unter der Voraussetzung  $V(y) \neq 0$  den durch  $x$  determinierten Varianzanteil von  $y$  :

$$(2.15) \quad \frac{V [E(y | x)]}{V(y)} = \frac{V(y) - V(\varepsilon)}{V(y)} = \frac{\mathbf{\beta}'C(x, x) \mathbf{\beta}}{V(y)} =: R^2(y, x).$$

Dieser Kennwert heißt multipler Determinationskoeffizient, die positive Wurzel daraus multiple Korrelation (vgl. auch Moosbrugger, 1978, S. 65). Die Werte von  $R^2(y, x)$  liegen zwischen Null und Eins, wobei  $R^2(y, x)$  den Wert Null annimmt, wenn  $\mathbf{\beta}$  bzw.  $C(x, y)$  der Nullvektor ist.

#### 2.2.5 Multivariate Verallgemeinerung

Liegen  $P$  abhängige Variablen vor, so können wir die Grundannahme 2.1 durch deren multivariate Verallgemeinerung

$$(2.16) \quad E(\mathbf{y} | \mathbf{x}) = \beta_0 + \mathbf{B}'\mathbf{x}$$

ersetzen. Entsprechend dem Vorgehen beim univariaten ( $P = 1$ ) Fall erhalten wir hier die Modellgleichung für die Kovarianzmatrizen

$$(2.17) \quad C(\mathbf{x}, \mathbf{y}) = C(\mathbf{x}, \mathbf{x}) \mathbf{B}$$

und unter der Voraussetzung der Invertierbarkeit von  $C(\mathbf{x}, \mathbf{x})$ , die Identifikationsgleichungen

$$(2.18) \quad \mathbf{B} = C(\mathbf{x}, \mathbf{x})^{-1} C(\mathbf{x}, \mathbf{y})$$

und

$$(2.19) \quad \beta_0 = E(\mathbf{y}) - \mathbf{B}'E(\mathbf{x}).$$

Andernfalls gilt wieder die allgemeine Lösung

$$(2.20) \quad \mathbf{B} = C(\mathbf{x}, \mathbf{x})^{-1} C(\mathbf{x}, \mathbf{y}),$$

die erst durch zusätzlich Nebenbedingungen eindeutig wird.

Als multivariate Verallgemeinerung des multiplen Determinationskoeffizienten kann der multivariate Determinationskoeffizient

$$(2.21) \quad R^2(\mathbf{y}, \mathbf{x}) = \frac{\text{Spur} \{C[E(\mathbf{y} | \mathbf{x}), E(\mathbf{y} | \mathbf{x})] C(\mathbf{y}, \mathbf{y})^{-1}\}}{P}$$

angesehen werden (vgl. auch Cramer & Nicewander, 1979), dessen Werte zwischen Null und Eins liegen. Dabei ist

$$P = \text{Spur} [C(\mathbf{y}, \mathbf{y}) C(\mathbf{y}, \mathbf{y})^{-1}] = \text{Spur } \mathbf{I}$$

gleich der Anzahl der abhängigen Variablen. Näheres siehe Moosbrugger & Steyer (1982) in diesem Band.

### 2.2.6 Zusammenfassende Bemerkungen

Die in diesem Abschnitt dargestellte Theorie beruht auf einer einzigen Annahme, nämlich der Gleichung 2.1. Daraus läßt sich die Identifikationsgleichung für den dabei auftretenden Parametervektor  $\beta$  ableiten, falls  $C(x, x)$  invertierbar ist. Andernfalls müssen zusätzliche Nebenbedingungen über die Parameter in  $\beta$  formuliert werden, um zu einer eindeutigen Lösung für  $\beta$  aus den Kovarianzmatrizen  $C(x, x)$  und  $C(x, y)$  zu gelangen. Analoges gilt für den multivariaten Fall. Mit dem multiplen - bzw. multivariaten Determinationskoeffizienten kann angegeben werden, wie eng der statistische Zusammenhang zwischen  $y$  (bzw.  $\mathbf{y}$ ) und  $x$  ist. In den nachfolgenden Abschnitten wollen wir verdeutlichen, daß die beschriebenen Konzepte in verschiedensten Situationen der psychologischen Forschung ihre Anwendung finden.

## 2.3 Anwendungen der formalen Theorie.

In den folgenden Unterkapiteln wollen wir zeigen, um welche Spezialfälle der in Kapitel 2.2 dargestellten allgemeinen Theorie es sich bei der Regressions-, Korrelations-, Varianz-, Diskriminanz- und Kontingenzanalyse handelt. Dabei werden auch die Zusammenhänge zwischen den genannten Analyseverfahren deutlich.

### 2.3.1 Regressions- und Korrelationsanalyse

Die Begriffe „Regression“ und „Korrelation“ entstanden im letzten Jahrhundert. Explizit wurden „Regression“ und „Co-Relation“ von Galton zumindest seit 1889 benutzt (Galton 1889). Galtons erste, aber noch nicht als solche bezeichnete „Regressionslinie“ vom Jahre 1877 zeigt die Abhängigkeit des Durchmessers von Blatterbsensamen von jenen der zugehörigen Mutterpflanzen. Mit der Größe des Samendurchmessers der Mutterpflanzen steigt jener der Tochterpflanzen an, bei letzteren jedoch in geringerem Ausmaß. Diese Beobachtung, daß die Durchmesser von Blatterbsensamen der Tochterpflanzen dazu tendieren, zur mittleren Größe des Samendurchmessers der Vorfahren zurückzukehren, bezeichnete Galton 1889 als „Regression“. Regressions- und Korrelationsmethoden wurden dann von Pearson (1896) sowie Yule (1897) weiter entwickelt. Doch bereits lange vor Galton wurden „Regressionsanalysen“ durchgeführt, freilich ohne daß dabei dieser Begriff Verwendung fand, so z.B. auf geophysikalischem Gebiet von Adrain (1818) oder von den Wiener Astronomen Kreil (1831) und Littrow (1818, 1833), welche bereits genaue Berechnungsformeln der „Regressionskoeffizienten“ nach der Minimum-Quadrat-Methode angaben. (Ausführlichere geschichtliche Angaben machen z.B. Weiling, 1972, 1973 und David, 1978).

In der Regressionsanalyse handelt es sich bei  $y$  und den  $x_q$  in der Regel um quantitative stochastische Variablen. Sind z.B. die Meßwerte einer  $V_p$  in den Variablen  $x_q$  gegeben und sind die Koeffizienten  $\beta_q$  in der Gleichung 2.1 bekannt, so läßt sich der Wert  $y$  der Variablen  $y$  mit einem bestimmten Fehler  $\varepsilon$  voraussagen. Jeder für einen festen Vektor  $\mathbf{x}$  vorhergesagte Erwartungswert  $E(y | \mathbf{x} = \mathbf{x})$  ist eine Realisation der bedingten Erwartung  $E(y | \mathbf{x})$ , welche im Fall kontinuierlicher Variablen und linearer Beziehungen zwischen  $y$  und den  $x_q$  eine Ebene bzw. eine Hyperebene ist und Regressionsebene heißt. Für zwei unabhängige Variablen beschreibt  $E(y | x_1, x_2)$  eine Ebene, welche zur Prädiktion von  $y$  aus  $x_1$  und  $x_2$  benutzt werden kann (siehe Abbildung 2.1). Dabei ist  $\beta_0$  der Ordinatenabschnitt, also die Höhe, in der die Regressionsebene die  $y$ -Achse schneidet,  $\beta_1$  die Steigung der Ebene entlang der  $x_1$ -Achse und  $\beta_2$  die Steigung entlang der  $x_2$ -Achse.

Das Ausmaß der Vorhersagbarkeit von  $y$  hängt von der „Enge“ des linearen statistischen Zusammenhanges ab, die wir mit dem multiplen Determinationskoeffizienten angeben können (siehe Gleichung 2.15).

Im Fall eines multiplen Regressionsmodells mit zwei Prädiktoren und Interaktion,

$$(2.22) \quad E(y | x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2$$

ist die bedingte Erwartung  $E(y | x_1, x_2)$  eine gekrümmte Fläche (siehe Abbildung 2.2), deren Steigung entlang der  $x_1$ -Achse auf der Stufe  $x_2=0$  der Parame-

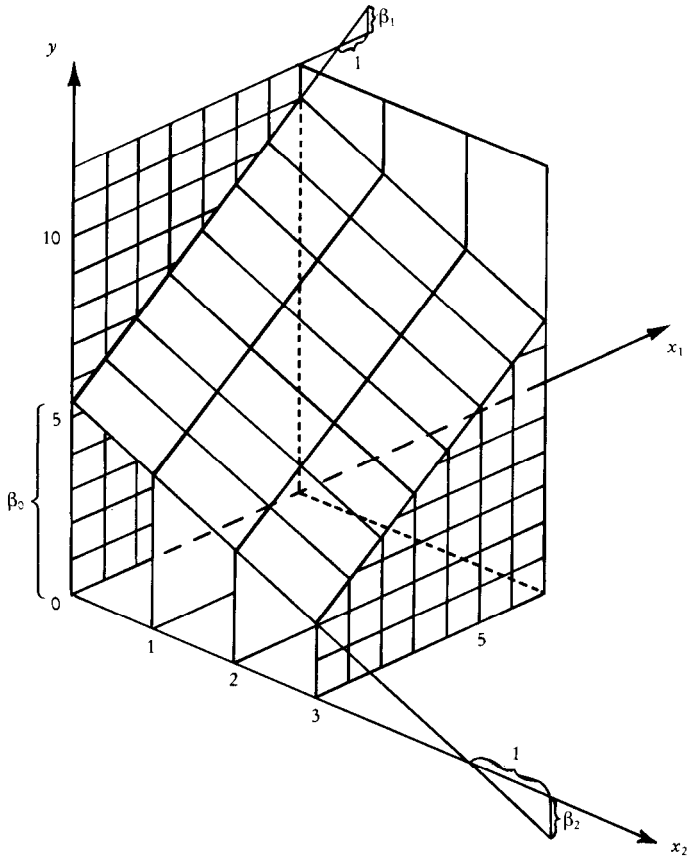


Abb. 2.1: Regressionsmodell mit zwei Prädiktoren

$$E(y | x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

mit

$$\beta_0 = 5.5$$

$$\beta_1 = 0.8$$

$$\beta_2 = -1.1$$

ter  $\beta_1$  angibt und entlang der  $x_2$ -Achse auf der Stufe  $x_1 = 0$  der Parameter  $\beta_2$ . Der Parameter  $\beta_3$  gibt jenen Anteil der bedingten Erwartung an, welcher auf bestimmte Kombinationen von Merkmalsausprägungen in den Variablen  $x_1$  und  $x_2$  zurückgeht: Von Stufe zu Stufe der Variablen  $x_2$  (bzw.  $x_1$ ) ist die Steigung der Variablen  $x_1$  (bzw.  $x_2$ ) um  $\beta_3$  verschieden (moderiert). Somit ist  $\beta_3$  genau der Unterschied zwischen den Steigungen der Regressionsgeraden von  $y$  auf  $x_1$  (bzw.  $x_2$ ) in jeweils zwei um eine Einheit entfernt liegenden Ausprägungen der Variablen  $x_2$  (bzw.  $x_1$ ). Daher wird  $x_2$  (bzw.  $x_1$ ) als Moderatorvariable

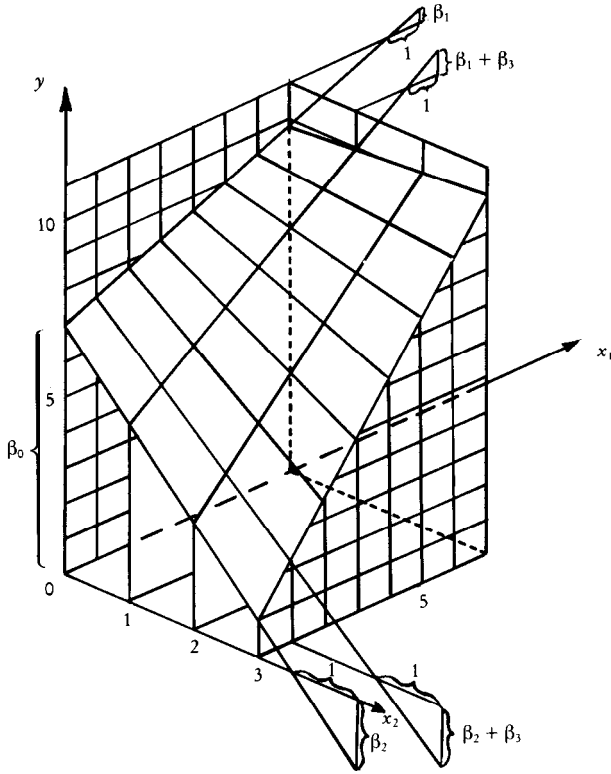


Abb. 2.2: Regressionsmodell mit zwei Prädiktoren und Interaktion

$$E(y | x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2$$

mit

- $\beta_0 = 7.0$
- $\beta_1 = 0.4$
- $\beta_2 = -2.0$
- $\beta_3 = 0.3$

bezeichnet, wenn  $\beta_3 \neq 0$  gilt (vgl. Bartussek, 1970, S. 11, Moosbrugger, 1981 a, S. 220).

Haben wir eine Stichprobe von Werten der Variablen  $y$  und  $x_q$  vorliegen, so können wir für multiple Regressionsmodelle unter der Normalverteilungsannahme der Residualvariablen Hypothesen prüfen, indem wir berechnen, wie wahrscheinlich es ist, daß alle oder einzelne  $\beta_q$  in der Population gleich Null sind (siehe z.B. Moosbrugger, 1978, S. 72-77), oder daß sie bzw. Linearkombinationen aus ihnen andere hypothetische feste Werte annehmen (siehe z.B. Timm, 1975, S. 177ff. oder Moosbrugger & Steyer, 1982, in diesem Band).

Weiterführende Literatur: Die Regressionsanalyse gehört zu den Themen, die von den meisten Lehrbüchern der angewandten Statistik behandelt werden. Ausführlichere Einführungen in die Regressionsanalyse geben Moosbrugger (1978), sowie Gaensslen & Schubö (1973), deren Darstellung Matrixalgebra vermeidet. Des weiteren ist auch Kerlinger & Pedhazur (1973) zur Einführung geeignet, außerdem Edwards (1976) sowie Draper & Smith (1966). Auf die Bedeutung von komplexen Regressionsmodellen mit Moderatorvariablen für die differentielle Validität psychologischer Tests geht z.B. Moosbrugger (1981 a) ein.

### 2.3.2 Varianzanalyse

Die Varianzanalyse wurde von Sir R. A. Fisher (1925) für die Auswertung von landwirtschaftswissenschaftlichen Experimenten entwickelt, bei denen es z.B. um die Effekte von Düngemitteln auf den Ertrag ging. Nach dem zweiten Weltkrieg hat dieses Verfahren jedoch auch in der Psychologie verstärkt Anwendung gefunden, wobei in der Regel die Wirkungen experimenteller Manipulationen überprüft werden. Heute zählt sie zu den Standardverfahren der psychologischen Forschungsmethoden. Zur Geschichte der Varianzanalyse siehe Weiling (1973).

Während bei varianzanalytischen Modellen mit Zufallsparametern (random-model, vgl. z.B. Hays, 1973, Kap. 13) die  $\beta_q$  in Gleichung 2.1 stochastische Variablen wären, können Varianz- und kovarianzanalytische Modelle mit festen Parametern (fixed-model, vgl. z.B. Hays, 1973, Kap. 12) als Spezialfalle der Gleichung 2.1 angesehen werden. Wir beschränken uns daher auf Modelle mit festen Parametern. Bei  $y$  handelt es sich dann wieder um eine quantitative stochastische Variable, bei den  $x_q$  jedoch in der Varianzanalyse ausschließlich um qualitative Indikatorvariablen (synonym auch: *Kodier-* oder *Dummyvariablen*), deren Werte die Gruppenzugehörigkeit anzeigen, und in der *Kovarianzanalyse* (siehe z.B. Moosbrugger, 1978, Kap. 15) um qualitative und quantitative Variablen.

In der einfaktoriellen Varianzanalyse mit zwei Stufen beispielsweise liegen als Werte der abhängigen Variablen  $y$  eine Reihe quantitativer Meßwerte vor, sowie die Information, welcher von zwei Gruppen ein Wert  $y$  jeweils angehört. Von Interesse ist dann der Mittelwerteunterschied von  $y$  in beiden Gruppen (z.B. Behandlungs- und Kontrollgruppe). Kodieren wir die Zugehörigkeit zur ersten Gruppe mit  $x = 1$  und die zur zweiten Gruppe mit  $x = -1$ , so ergibt sich das Datenschema der Abbildung 2.3.<sup>4)</sup>

---

<sup>4)</sup> Man beachte dabei, daß eine Kodierung der Gruppenzugehörigkeit mit 1 und 0 ebenfalls zu gut interpretierbaren Parametern führen würde (vgl. auch Rochel, 1981). Indikationsfragen verschiedener Kodierungsverfahren wurden von Schermelleh (1979) untersucht.

Gruppe $\mathcal{Y}$		$x$
1	$y_{11}$	1
	$y_{12}$	1
	.	.
	.	.
	$y_{1N_1}$	1
2	$y_{21}$	-1
	$y_{22}$	-1
	.	.
	.	.
	$y_{2N_2}$	-1

Abb. 2.3: Datenschema eines einfaktoriellen varianzanalytischen Modells mit zwei Stufen,  $\mathbf{EO}, \mathbf{y} = \beta_0 + \mathbf{B} \mathbf{x}$ . Die abhängige Variable ist  $y$ , die an  $N_1$  bzw.  $N_2$  Beobachtungseinheiten in den beiden Gruppen erhoben wird. Die Indikatorvariable  $x$  enthält die Information über die Gruppenzugehörigkeit der Beobachtungseinheiten.

In der *zweifaktoriellen* Varianzanalyse mit je zwei Stufen (siehe Abbildung 2.4) beispielsweise liegen die Werte der abhängigen Variablen  $y$  vor sowie die Information, aus welcher von vier Gruppen ein Wert  $y$  jeweils stammt. Von Interesse sind lineare oder nichtlineare statistische Zusammenhänge zwischen den  $x_q$  und  $y$ , welche sich hier als Unterschiede der Erwartungswerte von  $y$  in den vier Gruppen zeigen.

		Faktor A	
		Stufe 1	Stufe 2
Faktor B	Stufe 1	1. Gruppe	3. Gruppe
	Stufe 2	2. Gruppe	4. Gruppe

Abb. 2.4: Zweifaktorieller vollständig gekreuzter Versuchsplan mit 2 Stufen pro Faktor

Wir kodieren die zwei Stufen des Faktors A (siehe Abbildung 2.4) durch die Kodiervariable  $x_1$  mit den Werten  $x_1 = 1$  für die Stufe 1 und  $x_1 = -1$  für Stufe

2, und die zwei Stufen des Faktors B durch die Kodiervariable  $x_2$  mit den Werten  $x_2 = 1$  für die Stufe 1 und  $x_2 = -1$  für Stufe 2. Liegen mehr als zwei Stufen pro Faktor vor, benötigt man auch mehr als eine Indikatorvariable pro Faktor. Man beachte, daß die Art der Kodierung zwar die Interpretation der Parameter bestimmt, aber keinen Einfluß auf die Hypothesenbeurteilung hat (siehe z.B. Moosbrugger, 1978, Moosbrugger & Steyer (1982) in diesem Band, Rochel, 1981, Steyer, 1979 oder Wottawa, 1974).

Nach dem in Abbildung 2.4 schematisierten kreuzfaktoriellen Versuchsplan ergibt sich für die vier Subpopulationen das in Abbildung 2.5 dargestellte Datenschema:

Gruppe	$y$	$x_1$	$x_2$	$x_1 \cdot x_2$
1	$y_{11}$	1	1	1
	$y_{12}$	1	1	1
	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$y_{1N_1}$	1	1	1
2	$y_{21}$	1	-1	-1
	$y_{22}$	1	-1	-1
	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$y_{2N_2}$	1	-1	-1
3	$y_{31}$	-1	1	-1
	$y_{32}$	-1	1	-1
	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$y_{3N_3}$	-1	1	-1
4	$y_{41}$	-1	-1	1
	$y_{42}$	-1	-1	1
	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$y_{4N_4}$	-1	-1	1

Abb. 2.5: Varianzanalytisches Datenschema zum Versuchsplan in Abbildung 2.4.

In der ersten Spalte ist die abhängige Variable  $y$  notiert, die an  $N_g$  Beobachtungseinheiten in der Gruppe  $g$ ,  $g \in \{1,2,3,4\}$ , erhoben wird. Die Spalten 2 und 3 zeigen die Indikatorvariablen  $x_1$  und  $x_2$  für die Stufen der Faktoren A und B. Die Variable  $x_1 \cdot x_2$  in Spalte 4 ist ein Indikator für bestimmte Faktorstufenkombinationen. Sie dient der Erfassung von Wechselwirkungen (Interaktionen, s.u.). Wird keine Interaktion angenommen, so entfällt die Spalte  $x_1 \cdot x_2$ .

Wie in der Regressionsanalyse liegen uns mit den Datenschemata der Abbildungen 2.3 und 2.5 Werte von stochastischen Variablen vor, deren statistischer Zusammenhang mit Gleichung 2.1 beschrieben werden kann.

Wir untersuchen im folgenden am zweifaktoriellen Design, welche Interpretation den Parametern  $\beta_q$  in diesem Kontext zukommt, zunächst ohne, danach mit Annahme einer Interaktion (Wechselwirkung). Ein statistischer Zusammenhang zwischen  $y$  und den  $x_q$  *ohne Interaktion* liegt dann vor, wenn Mittelwerteunterschiede zwischen den Gruppen derart bestehen, daß die Differenz der Erwartungswerte zwischen den Stufen des Faktors A auf jeder der Stufen des Faktors B gleich groß ist und umgekehrt.

In welcher Beziehung die Parameter  $\beta_0$  und  $\beta_q$  der Gleichung

$$(2.23) \quad E(y | x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

mit den Erwartungswerten von  $y$  in den Gruppen stehen, sehen wir, wenn wir die bedingten Erwartungswerte von  $y$  für die vier Gruppen bilden:

$$(2.24) \quad E(y_1) := E(y | x_1 = 1, x_2 = 1) = \beta_0 + \beta_1 + \beta_2$$

$$(2.25) \quad E(y_2) := E(y | x_1 = 1, x_2 = -1) = \beta_0 + \beta_1 - \beta_2$$

$$(2.26) \quad E(y_3) := E(y | x_1 = -1, x_2 = 1) = \beta_0 - \beta_1 + \beta_2$$

$$(2.27) \quad E(y_4) := E(y | x_1 = -1, x_2 = -1) = \beta_0 - \beta_1 - \beta_2.$$

Gewöhnlich werden die Gruppen als gleich groß angenommen, was bedeutet, daß die Wahrscheinlichkeiten für die Faktorstufen  $P(x_1 = 1) = P(x_1 = -1) = 1/2$  und  $P(x_2 = 1) = P(x_2 = -1) = 1/2$  gilt. Dies stellt sicher, daß die unabhängigen Variablen  $x_1$  und  $x_2$  zueinander orthogonal sind, d.h. nicht miteinander korrelieren, so daß deren Einflüsse auf die abhängige Variable  $y$  voneinander unabhängig geschätzt werden können. Ist dies nicht der Fall, handelt es sich um eine nonorthogonale Varianzanalyse (siehe z.B. Steyer, 1979).

Unter den obigen Annahmen der Gleichwahrscheinlichkeit für die Faktorausprägungen finden wir Null als Erwartungswert der Indikatorvariablen:

$$(2.28) \quad E(x_1) = P(x_1=1) \cdot 1 + P(x_1 = -1) \cdot -1 = 0$$

und

$$(2.29) \quad E(x_2) = P(x_2=1) \cdot 1 + P(x_2 = -1) \cdot -1 = 0.$$

Bilden wir den Erwartungswert von Gleichung 2.23, so erhalten wir dann (vgl. Anhang C, Regel 1)

$$(2.30) \quad \beta_0 = E(y) - \beta_1 E(x_1) - \beta_2 E(x_2) = E(y).$$

Demnach ist  $\beta_0$  unter der genannten Voraussetzung gleich dem Erwartungswert von  $y$ . Einsetzen von Gleichung 2.30 in die Gleichungen 2.24 bis 2.27 führt zu den Lösungen für

$$(2.31) \quad \beta_1 = \frac{E(y_1) + E(y_2)}{2} - E(y) = E(y) - \frac{E(y_3) + E(y_4)}{2}$$

und

$$(2.32) \quad \beta_2 = \frac{E(y_1) + E(y_3)}{2} - E(y) = E(y) - \frac{E(y_2) + E(y_4)}{2}.$$

Wir sehen, daß  $\beta_1$  den „Haupteffekt von Faktor A“, (vgl. z.B. Hays, 1973, S. 494), oder, bezogen auf Abbildung 2.4, den „Spalteneffekt“ darstellt, wohingegen  $\beta_2$  der „Haupteffekt von Faktor B“, oder bezogen auf Abbildung 2.4, der „Zeileneffekt“ ist<sup>5</sup>). Abbildung 2.6 veranschaulicht die Beziehungen zwischen den Parametern und den Erwartungswerten in den vier Gruppen:

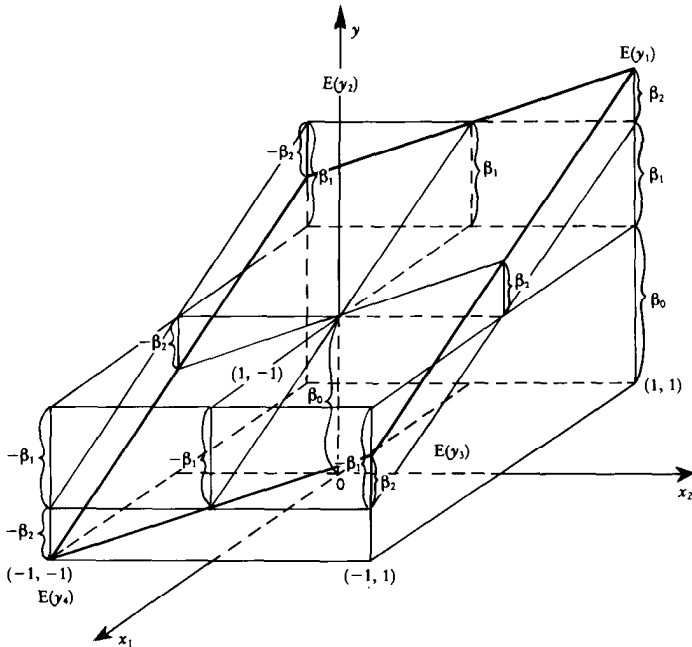


Abb. 2.6: Veranschaulichung der Erwartungswerte von  $y$  in den 4 Gruppen eines zweifaktoriellen varianzanalytischen Modells gemäß den Gleichungen 2.24 bis 2.27 mit  $\beta_0 = 3$ ,  $\beta_1 = 2$  und  $\beta_2 = 1$ .

<sup>5</sup> Die weitverbreitete Bezeichnung „Effekt“ impliziert eine kausale Interpretation, welche man strenggenommen nur im Kontext eines kausalen Modells (siehe Steyer, 1982, in diesem Band) vornehmen sollte.

Ist der statistische Zusammenhang zwischen den bedingenden Variablen  $x_q$  und der abhängigen Variablen  $y$  nichtlinear in den bedingenden Variablen<sup>6)</sup>, so moderieren (verändern) die Stufen des Faktors B den Effekt des Faktors A oder umgekehrt (vgl. auch Kap. 2.3.1). Einen solchen Moderatoreffekt, der in der Varianzanalyse Wechselwirkungs- oder Interaktionseffekt heißt, können wir mit einer weiteren Indikatorvariablen  $x_1 \cdot x_2$  für die Faktorstufenkombinationen und dem Parameter  $\beta_3$  erfassen. Die Spalte 4 in Abbildung 2.5 (s.O.) zeigt die entsprechenden Werte von  $x_1$ ;  $x_2$ . Ein Haupteffekt stellt dann den durchschnittlichen Effekt eines Faktors gemittelt über die Stufen des anderen Faktors dar. Für die bedingte Erwartung von  $y$  unter  $x$  gilt die folgende Gleichung

$$(2.33) \quad E(y | x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2.$$

Zur Untersuchung, in welcher Beziehung in diesem Fall die Parameter  $\beta_0$  und  $\beta_q$  der Gleichung 2.33 mit den Gruppenerwartungswerten stehen, bilden wir wieder die bedingten Erwartungswerte von  $y$  für die vier Gruppen:

$$(2.34) \quad E(y_1) = E(y | x_1 = 1, x_2 = 1) = \beta_0 + \beta_1 + \beta_2 + \beta_3$$

$$(2.35) \quad E(y_2) = E(y | x_1 = 1, x_2 = -1) = \beta_0 + \beta_1 - \beta_2 - \beta_3$$

$$(2.36) \quad E(y_3) = E(y | x_1 = -1, x_2 = 1) = \beta_0 - \beta_1 + \beta_2 - \beta_3$$

$$(2.37) \quad E(y_4) = E(y | x_1 = -1, x_2 = -1) = \beta_0 - \beta_1 - \beta_2 + \beta_3.$$

Bei gleichgroßen Subpopulationen gelten außer den Gleichungen 2.28 und 2.29 auch  $P(x_1 \cdot x_2 = 1) = P(x_1 \cdot x_2 = -1) = 1/2$  und  $E(x_1 \cdot x_2) = 0$ , was wegen  $E(x_1) = E(x_2) = 0$  wiederum (siehe Gleichung 2.30) zu

$$(2.38) \quad \beta_0 = E(y)$$

führt.

Einsetzen von Gleichung 2.38 in die Gleichungen 2.34 bis 2.37 ergibt unveränderte Ausdrücke für  $\beta_1$  und  $\beta_2$  (siehe Gleichungen 2.31 und 2.32) und zusätzlich eine Gleichung für  $\beta_3$ :

$$(2.39) \quad \beta_3 = \frac{E(y_1) + E(y_4)}{2} - E(y) = E(y) - \frac{E(y_2) + E(y_3)}{2}.$$

Der Interaktionsparameter, nämlich der Koeffizient  $\beta_3$  der Variablen  $x_1 \cdot x_2$  ist genau dann Null, wenn der Erwartungswert von  $y$  gleich dem arithmetischen

<sup>6)</sup> Zur Klassifikation von linearen vs. nichtlinearen Modellen hinsichtlich der Parameter  $\beta_q$  bzw. linearen vs. nichtlinearen Modellen hinsichtlich den bedingenden Variablen  $x_q$  siehe Moosbrugger, 1980, S. 5-12.

Mittel der Erwartungswerte in den Gruppen 1 und 4 bzw. 2 und 3 ist. Dies ist der Fall, wenn sich die Erwartungswerte von  $y$  in den vier Gruppen allein aus dem Gesamterwartungswert  $E(y) = \beta_0$  und den Parametern  $\beta_1$  und  $\beta_2$  darstellen lassen. Die Größe von  $\beta_3$  wächst mit der Nichtlinearität, nämlich mit der Abweichung des Durchschnitts von je zwei, bezogen auf die Abbildung 2.4 „diagonalen“, Gruppenerwartungswerten vom Gesamterwartungswert. Der Veranschaulichung dient Abbildung 2.7:

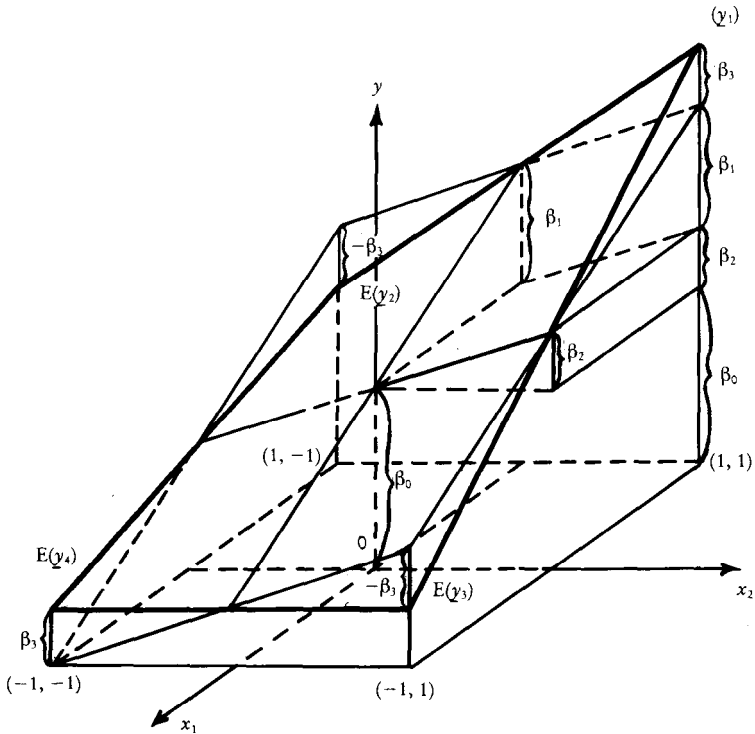


Abb. 2.7: Veranschaulichung der Erwartungswerte von  $y$  in den 4 Subpopulationen eines zweifaktoriellen varianzanalytischen Modells mit Wechselwirkung gemäß den Gleichungen 2.34 bis 2.37 mit  $\beta_0 = 3$ ,  $\beta_1 = 2$ ,  $\beta_2 = 1$  und  $\beta_3 = 1$  im Unterschied zu  $E(y; x_1, x_2)$  in Abbildung 2.6. Der Haupteffekt  $\beta_1 = 2$  gibt hier nur den durchschnittlichen Effekt von Faktor A an, welcher durch die Stufen von Faktor B moderiert wird. Die Steigung (der Effekt) auf der Stufe  $B_1$  beträgt  $(\beta_1 + \beta_3) = 3$ , auf der Stufe  $B_2$  hingegen  $(\beta_1 - \beta_3) = 1$ . Die analoge Aussage gilt für Faktor B: auf der Stufe  $A_1$  beträgt der Effekt  $(\beta_2 + \beta_3) = 2$ , auf der Stufe  $A_2$  hingegen  $(\beta_2 - \beta_3) = 0$ ; den Durchschnitt zeigt der Haupteffekt  $\beta_2 = 1$  an.

Ebenso wie bei der Regressionsanalyse können wir in der Varianzanalyse unter der Normalverteilungsannahme Hypothesen prüfen, indem wir berechnen, wie wahrscheinlich es ist, daß alle oder einzelne  $\beta_q$  in der Population gleich Null sind (siehe z.B. Moosbrugger, 1978, S. 72-77), oder daß Linearkombinationen der  $\beta_q$  hypothetische feste Werte annehmen (siehe z.B. Timm, 1975, S. 177ff. oder Moosbrugger & Steyer, 1982, in diesem Band), falls wir eine Stichprobe von Werten der Variablen  $y$  und  $x_q$  vorliegen haben.

Weiterführende Literatur: Auch die Varianzanalyse gehört zu den Standardthemen der Lehrbücher für angewandte Statistik. Außer den bereits unter der Regressionsanalyse (s.o.) genannten Bücher seien Searle (1971), Timm (1975), Bock (1975) und Finn (1974) aufgeführt. Als ausgesprochenes Standardwerk gilt Winer (1971<sup>2</sup>), der für sehr viele Designs Rechenformeln angibt. Als weitere deutschsprachige Bücher seien Bandemer et. al. (1977), Bortz (1977) und Rasch (1976) genannt. Eine Einführung in die multivariate Varianzanalyse geben Moosbrugger & Steyer (1982) in diesem Band.

### 2.3.3 Diskriminanzanalyse

Auch die Diskriminanzanalyse wurde von R. A. Fisher (1936) entwickelt. Sein Anwendungsbeispiel war die Zuordnung von Irispflanzen zu einer von zwei Arten auf der Basis der Länge und Breite der Kelch- und Blütenblätter (Anderson, 1978, S. 628). Die typische Fragestellung der Diskriminanzanalyse in der Psychologie ist die der Klassifikation, nämlich, zu welcher von bestimmten (z.B. Krankheits-)Gruppen eine Person gehört, von der bestimmte Informationen  $x$  (z.B. ein oder mehrere Testwerte) vorliegen.

Formal gesehen ist die Diskriminanzanalyse eigentlich eine Umkehrung der Varianzanalyse. Ist es nach einer Varianzanalyse mit bedeutsamen Mittelwertunterschieden möglich, von der Kenntnis der Gruppenzugehörigkeit der qualitativen Variablen auf die durchschnittliche Ausprägung der quantitativen Variablen zu schließen, so ermöglicht in Umkehrung dieses Vorgangs die sogenannte Diskriminanzfunktion einen Schluß, in welche der Gruppen eine bestimmte Wertekombination der quantitativen Variablen am ehesten klassifiziert werden kann.

Die Grundgedanken lassen sich am leichtesten für den Fall zweier Gruppen darstellen. Hierbei ist zu beachten, daß wir, um die formale Verträglichkeit mit der Gleichung 2.1 herbeizuführen, gegenüber der Varianzanalyse die Bedeutung von  $x$  und  $y$  vertauschen. Nun sei  $y$  eine Indikatorvariable, wobei  $y = 1$  Zugehörigkeit zur Gruppe 1 und  $y = 0$  Zugehörigkeit zur Gruppe 2 bedeuten soll. Die quantitativen Informationen seien in den kontinuierlichen Variablen  $x_q$  gesammelt. Da wir in unserer bisherigen Erörterung nur eine abhängige Variable gegenüber mehreren unabhängigen zugelassen haben, ist mit einer Kodiervariablen  $y$  eine Unterscheidung von nur zwei Gruppen mög-

lich. Mehrere Gruppen erfordern ein multivariates Modell (siehe z.B. Bock, 1975, S. 403), um mit mehreren Indikatorvariablen  $y_p$  mehr als zwei Gruppen kodieren zu können.

In der Diskriminanzanalyse heißt die bedingte Erwartung  $E(y|\mathbf{x})$  auch Diskriminanzfunktion. Wird diese wiederum als linear angenommen, so gilt weiterhin Gleichung 2.1, doch mit der oben erwähnten Vertauschung von  $x$  und  $y$  geht auch eine Veränderung der Bedeutung der Parameter  $\beta_q$  einher, welche wir kurz untersuchen wollen:

Die Interpretation der Skalenkonstanten  $\beta_0$  wird erleichtert, wenn wir die Erwartungswerte aller kontinuierlichen Variablen  $x_q$ ,  $q \in \{1, 2, \dots, Q\}$ , gleich Null setzen. Dann ist

$$(2.40) \quad E(y) = E[E(y|\mathbf{x})] = \beta_0 + \beta_1 E(x_1) + \dots + \beta_Q E(x_Q) = \beta_0.$$

Da die Gruppenvariable  $y$  nur die Werte  $y = 1$  und  $y = 0$  annehmen kann, gilt für ihren Erwartungswert und folglich für  $\beta_0$

$$(2.41) \quad \beta_0 = E(y) = P(y=1) \cdot 1 + P(y=0) \cdot 0 = P(y=1).$$

Die Skalenkonstante gibt also die Wahrscheinlichkeit an, der mit Eins kodierten Gruppe anzugehören, Folglich könnte bei unbekanntem Wert der Variablen  $x_q$ , ebenso wie bei  $\beta = 0$  eine Vorhersage (Diskrimination) der Gruppenzugehörigkeit *nur* nach den Gruppenwahrscheinlichkeiten erfolgen. Sind beide Gruppen zudem gleich wahrscheinlich, so ist eine begründete Entscheidung zwischen beiden Gruppen nicht mehr möglich, weil  $\beta_0 = 1/2$  wird.

Die Vorhersagbarkeit wächst hingegen mit der „Enge“ des Zusammenhangs zwischen  $y$  und den Variablen  $x_q$ , d.h., sofern wir von unveränderten Varianzen der beteiligten Variablen ausgehen, mit wachsender Höhe der Parameter  $\beta_q$ . Die Diskriminanzfunktion kann bei der hier vorliegenden zweiwertigen Gruppenvariablen  $y$  wegen

$$(2.42) \quad E(y|\mathbf{x}) = P(y=1|\mathbf{x}) \cdot 1 + P(y=0|\mathbf{x}) \cdot 0 = P(y=1|\mathbf{x})$$

als bedingte Wahrscheinlichkeit, der Gruppe 1 anzugehören, interpretiert werden. Liegen die Parameter  $\beta_q \neq 0$  fest, so erhöhen die Absolutbeträge der Differenzen im Vektor  $x-E(x)$  die Sicherheit der Gruppenzuordnung.

Bei der Annahme einer linearen Diskriminanzfunktion ist zu beachten, daß der Wertebereich jeder quantitativen Variablen  $x_q$ ,  $q \in \{1, 2, \dots, Q\}$ , deren zugehöriger Koeffizient  $\beta_q \neq 0$  ist, auf ein endliches Intervall beschränkt sein muß. Andernfalls ergäbe sich ein Widerspruch zu der Tatsache, daß die Werte von  $E(y|\mathbf{x})$  wegen der Kodierung mit Null und Eins als bedingte Wahrscheinlichkeiten zu interpretieren sind. Demnach läßt sich beispielsweise die Annahme der multivariaten Normalverteilung der  $x_q$  innerhalb der beiden Gruppen

nicht in Einklang bringen mit der Annahme, daß  $E(y \mid \mathbf{x})$  eine lineare Funktion im Sinne der Gleichung 2.1 ist, falls mindestens einer der Koeffizienten  $\beta_{q,q} \in \{1, 2, \dots, Q\}$ , ungleich Null ist.

Weiterführende Literatur: Eine einführende Darstellung in die Diskriminanzanalyse findet man z.B. bei Tatsuoka (1971) oder Marinell (1977). Bock (1975) zeigt darüber hinaus auch, wie man Nutzenüberlegungen mit der Diskriminanzanalyse verbinden kann, was für Zuordnungsprobleme insofern von Bedeutung ist, als verschiedene Arten von falschen Zuordnungen verschieden große Schäden anrichten können.

### 2.3.4 Kontingenzanalyse

Von einer Kontingenzanalyse sprechen wir, wenn der statistische Zusammenhang zwischen qualitativen Variablen untersucht wird. Während wir im Abschnitt über die Varianzanalyse die Gruppenzugehörigkeit mit Variablen kodiert haben, deren Werte nur 1 und -1 sein konnten, so verwenden wir hier wie in der Diskriminanzanalyse Indikatorvariablen, die nur die Werte 1 und 0 annehmen. Dies führt zu besonders gut interpretierbaren Parametern.

Sind sowohl die unabhängigen Variablen  $x_q$  als auch die abhängige Variable  $y$  Kodiervariablen für jeweils zwei Kategorien, so reduziert sich die Untersuchung der Kovarianzen nach dem Modell der Gleichung 2.1 auf die Analyse der Wahrscheinlichkeit des Auftretens von Merkmalskombinationen. Exemplarisch betrachten wir hier den Fall einer  $2 \times 2 \times 2$  Kontingenzanalyse (siehe Abbildung 2.8) sowohl ohne als auch mit Interaktion.

Zur Interpretation der Parameter im kontingenzanalytischen Modell ohne Interaktion

$$(2.43) \quad E(y \mid x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

untersuchen wir wieder die Erwartungswerte von  $y$  bei gegebenem  $\mathbf{x} = \mathbf{x}$ . Nach der obigen Gleichung ergibt die Betrachtung des Falles  $x_1 = 0, x_2 = 0$

$$(2.44) \quad \beta_0 = E(y \mid x_1 = 0, x_2 = 0) = P(y = 1 \mid x_1 = 0, x_2 = 0),$$

die des Falles  $x_1 = 1, x_2 = 0$

$$(2.45) \quad \begin{aligned} \beta_1 &= E(y \mid x_1 = 1, x_2 = 0) - \beta_0 = \\ &= P(y = 1 \mid x_1 = 1, x_2 = 0) - P(y = 1 \mid x_1 = 0, x_2 = 0) \end{aligned}$$

und die des Falles  $x_1 = 0, x_2 = 1$

$$(2.46) \quad \begin{aligned} \beta_2 &= E(y \mid x_1 = 0, x_2 = 1) - \beta_0 = \\ &= P(y = 1 \mid x_1 = 0, x_2 = 1) - P(y = 1 \mid x_1 = 0, x_2 = 0). \end{aligned}$$

Gruppe	y	x <sub>1</sub>	x <sub>2</sub>	x <sub>1</sub> · x <sub>2</sub>
1	0	0	0	0
2	1	0	0	0
3	0	1	0	0
4	1	1	0	0
5	0	0	1	0
6	1	0	1	0
7	0	1	1	1
8	1	1	1	1

Abb. 2.8: Kodierung einer dreidimensionalen  $2 \times 2 \times 2$  Kontingenztabelle für ein Modell mit Interaktion. Bei einer Kontingenzanalyse sind alle beteiligten Variablen qualitativ. Die unabhängigen Variablen  $x_1$  und  $x_2$  enthalten Information über die Gruppenzugehörigkeit einer Beobachtungseinheit. Bei einem Modell ohne Interaktion entfällt die Spalte für  $x_1 \cdot x_2$ . Eine Merkmalskombination liegt mit den Werten der Variablen  $y$ ,  $x_1$  und  $x_2$  fest.

Der Parameter  $\beta_0$  ist also in einer  $2 \times 2 \times 2$  Kontingenzanalyse mit der Kodierung Null und Eins gleich der bedingten Wahrscheinlichkeit für „ $y = 1$ “ bei gegebenem  $x_1 = x_2 = 0$  (siehe Gleichung 2.44). Die Parameter  $\beta_1$  ist nach Gleichung 2.45 die Differenz der bedingten Wahrscheinlichkeiten von  $y = 1$  bei gegebenem  $x_1 = 1$  bzw.  $x_1 = 0$  auf der Stufe  $x_2 = 0$ , und der Parameter  $\beta_2$  ist nach Gleichung 2.46 die Differenz der bedingten Wahrscheinlichkeiten von  $y = 1$  bei gegebenem  $x_2 = 1$  bzw.  $x_2 = 0$  auf der Stufe  $x_1 = 0$ . Diese Interpretationen gelten sowohl für den Fall ohne als auch mit Interaktion.

Legen wir einen Zusammenhang zwischen  $y$  und den  $x_i$  zugrunde, bei dem eine Interaktion besteht, so können wir die Ausprägungskombinationen der Variablen  $x_1$  und  $x_2$  durch die Variable  $x_1 \cdot x_2$  erfassen (siehe die Spalte für  $x_1 \cdot x_2$  der Abbildung 2.8)

$$(2.47) \quad E(y | x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 \cdot x_2.$$

Als Identifikationsgleichungen für die Parameter  $\beta_1$  bis  $\beta_3$  dient auch hier die Gleichung 2.9, für  $\beta_0$  die Gleichung 2.11.

Die Interpretation der Parameter  $\beta_0$ ,  $\beta_1$  und  $\beta_2$  bleibt unverändert (siehe Gleichungen 2.44 bis 2.46); die für  $\beta_3$  ergibt sich aus der Gleichung 2.47 wie folgt:

$$(2.48) \quad \begin{aligned} \beta_3 &= E(y \mid x_1 = 1, x_2 = 1) - \beta_0 - \beta_1 - \beta_2 = \\ &= P(y = 1 \mid x_1 = 1, x_2 = 1) - P(y = 1 \mid x_1 = 1, x_2 = 0) - \\ &\quad - [P(y = 1 \mid x_1 = 0, x_2 = 1) - P(y = 1 \mid x_1 = 0, x_2 = 0)]. \end{aligned}$$

Eine Interaktion  $\beta_3 \neq 0$  liegt vor, wenn die Differenz der bedingten Wahrscheinlichkeiten  $P(y = 1 \mid x_2 = 1)$  und  $P(y = 1 \mid x_2 = 0)$  auf den beiden Stufen von  $x_1$  nicht gleich groß ist.

Als Beispiel für eine Kontingenzanalyse mit extremer Interaktion behandelt Moosbrugger (1980) das sogenannte Meehl'sche Paradoxon (Meehl, 1950).

Liegen für die Wahrscheinlichkeiten, einer Untergruppe anzugehören, Schätzungen vor, so können wir unter Beachtung entsprechender Voraussetzung (vgl. z.B. Bortz, 1977, S. 202) mit Signifikanztests ( $\chi^2$ -Tests) überprüfen, wie wahrscheinlich es ist, daß alle oder einzelne Parameter  $\beta_q$  in der Population gleich Null sind oder einen anderen hypothetischen festen Wert haben.

*Weiterführende Literatur:* Zur Kontingenzanalyse oder Analyse qualitativer Daten gibt es eine Vielzahl von unterschiedlichen Ansätzen, welche z.T. (wie etwa die „loglinearen Modelle“ oder die „Konfigurationsfrequenzanalyse“) keine Unterscheidung zwischen abhängigen und unabhängigen Variablen treffen. Die wichtigsten Ansätze scheinen uns durch folgende Autoren vertreten zu sein: Bishop, Fienberg & Holland (1975), Cox (1970), Goodman (1978), Krauth & Lienert (1973).

### 3. Modelle mit latenten Variablen

#### 3.1 Einleitung

Im Alltag erklärt man „gleichartiges“ Verhalten einer Person in verschiedenen Situationen oft, indem man ihr eine Eigenschaft zuschreibt, die für dieses Verhalten verantwortlich ist. Verhält sich jemand beispielsweise effektiv bei der Lösung von Problemen verschiedener Art, so schreibt man dies seiner hohen Intelligenz zu. Diese alltägliche Vorgehensweise kann man durch mathematische Modelle präzisieren. Das Prinzip dabei ist, daß man die Abhängigkeiten manifester oder beobachtbarer Variablen (das entspricht dem Verhalten in verschiedenen Situationen) durch eine oder mehrere latente, d.h. nicht

direkt beobachtbare Variablen (das entspricht der Eigenschaft) erklärt. Mit „Erklären“ meint man dabei, daß die Abhängigkeiten zwischen beobachtbaren Variablen verschwinden, wenn eine dahinterliegende latente Variable gefunden wird. Die latenten Variablen lassen sich dann als hypothetische Konstrukte (vgl. Herrmann, 1973 oder Mac Corquodale & Meehl, 1948) interpretieren.

Die latenten Eigenschaften können somit als Ursache für die Variation und Kovariation des beobachteten Verhaltens betrachtet werden (vgl. z.B. Bentler 1980). Beobachtbare Variablen, deren Kovariation auf eine gemeinsame latente Variable zurückzuführen ist, heißen auch *kongenerisch* (vgl. z.B. Jöreskog, 1971). Abbildung 3.1 zeigt ein einfaches Modell kongenerischer Variablen:

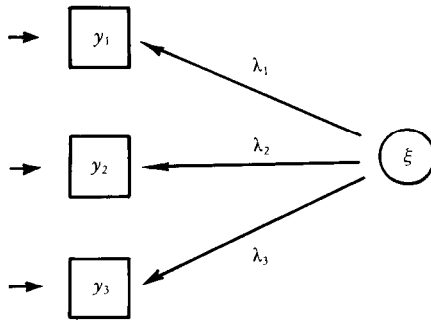


Abb. 3.1: Modell mit einer latenten Variablen  $\xi$  und drei beobachtbaren Variablen  $y_p$ , bei denen es sich in psychometrischen Anwendungen um drei Tests oder auch drei Testitems handeln kann. Die manifestigen Variablen  $y_1$ ,  $y_2$  und  $y_3$  hängen mit den Gewichten  $\lambda_1$ ,  $\lambda_2$  und  $\lambda_3$  von der gemeinsamen latenten Variablen  $\xi$  sowie jeweils von den spezifischen Residualvariablen ab, welche durch unbeschriftete Pfeile gekennzeichnet sind.

In Kapitel 3.2 werden wir eine allgemeine *Theorie latenter Variablen* darstellen, auf deren Basis es möglich sein wird, eine Anzahl von Modellen, welche auf den ersten Blick durchaus verschieden erscheinen, als Spezialfälle dieser Theorie vorzustellen, weil sie sich nur durch die Art der beobachtbaren und latenten Variablen unterscheiden. Bei der linearen Faktorenanalyse<sup>7)</sup> mit einem Faktor sind sowohl die manifesten Variablen als auch die latente Variable (= Faktor) kontinuierlich. Beim „linear tralceline model“ (siehe z.B. Lazarsfeld & Henry, 1968, S. 206 oder Fischer, 1974, S. 160) sind die manifesten Variablen zweiwertig und die latente Variable kontinuierlich. Beim „latent

<sup>7)</sup> Damit ist die Faktorenanalyse im Sinne von Spearman (1904, 1926), Thurstone (1947) und Jöreskog (1966) gemeint und nicht die Hauptkomponentenanalyse nach Pearson (1901) und Hotelling (1933). Die Unterschiedlichkeit beider Ansätze beschreiben z.B. Lawley & Maxwell, 1971, S. 2-4).

*profile model*“ (siehe z.B. Gibson, 1959) sind die beobachtbaren Variablen kontinuierlich und die latente Variable repräsentiert nur die Zugehörigkeit zu einer latenten Klasse. Beim „*latent class model*“ (siehe z.B. Lazarsfeld & Henry, 1968, S. 46) sind die manifesten Variablen zweiwertig und die latente Variable repräsentiert ebenfalls nur die Zugehörigkeit zu den Klassen. Schließlich kann auch das „klassische latent-additive Testmodell“ (siehe Moosbrugger & Müller, 1980, 1981, 1982) in diesem Kontext abgehandelt werden.

Während allen vorgenannten Modellen eine lineare Variablencharakteristik zugrundeliegt, werden wir auch auf andere Gruppen von Modellen, nämlich solche mit polynomialer Variablencharakteristik (Lazarsfeld, 1950, McDonald 1967) sowie solche mit logistischer Variablencharakteristik (Rasch 1960, Birnbaum 1968) eingehen, die sich in formaler Hinsicht nur geringfügig von den linearen Modellen unterscheiden.

Die genannten Modelle können auch für mehrere latente Variablen verallgemeinert werden. Zur Einführung der wesentlichen Konzepte reicht jedoch der Fall einer einzigen latenten Variablen aus, welche die Abhängigkeiten zwischen den beobachtbaren Variablen erklärt. Auch in diesem dritten Abschnitt ist es nicht unser Ziel, einen umfassenden Überblick über Anwendungsmöglichkeiten, Probleme der Parameterschätzung und Hypothesenbeurteilung zu geben, sondern vielmehr die gemeinsamen theoretischen Grundlagen und Beziehungen herauszuarbeiten.

## 3.2 Eine allgemeine Theorie latenter Variablen

In diesem Abschnitt stellen wir zunächst die Annahmen der Theorie latenter Variablen und die wichtigsten Folgerungen zusammen. Danach untersuchen wir, wann und wie sich die unbekannt Parameter aus den statistischen Kennwerten manifester Variablen, nämlich aus Erwartungswerten, Varianzen und Kovarianzen, eindeutig bestimmen lassen.

### 3.2.1 Die Grundannahmen

Die formale Theorie latenter Variablen basiert auf zwei Grundannahmen, nämlich auf der bedingten stochastischen Unabhängigkeit der manifesten Variablen und auf der Art der Variablencharakteristik.

#### 3.2.1.1 Bedingte Unabhängigkeit

Der Grundgedanke von Modellen zur Erklärung der Abhängigkeiten beobachtbarer oder manifester Variablen  $y_p$ ,  $p \in \{1, 2, \dots, P\}$ , durch eine latente

Variable besteht darin, daß für jeden festen Wert der latenten Variablen  $\xi$  keine Abhängigkeit mehr zwischen den manifesten Variablen  $y_p$  besteht. Ein Beispiel soll dies erläutern: Handelt es sich bei den  $y_p$  um psychologische Testvariablen, welche miteinander korrelieren, so wird per Modellannahme erwartet, daß die Korrelationen verschwinden, wenn man die Testvariablen  $V_{pn}$  vorlegt, welche gleiche Ausprägungen der latenten Variablen  $\xi$  aufweisen. Für diese  $V_{pn}$  sollen die Testvariablen dann unabhängig sein.

Formal läßt sich dieser Gedanke wie folgt formulieren: Für alle nichtleeren Teilmengen  $J$  der Indexmenge  $\{1,2,\dots,p,\dots,P\}$  gelte (vgl. Anderson 1959, S. 11, Lord & Novick 1968, S. 543)

$$(3.1) \quad E\left(\prod_{p \in J} y_p \mid \xi\right) = \prod_{p \in J} E(y_p \mid \xi),$$

wobei  $\prod$  das Produktzeichen ist. Nach dieser Gleichung ist die bedingte Erwartung des Produkts der Variablen  $y_p$  gleich dem Produkt der bedingten Erwartungen der Variablen  $y_p$ . Zufallsvariablen  $y_p$ , welche diese Gleichung erfüllen, heißen bedingt linear stochastisch unabhängig.

Diese Bedingung sei für zwei Variablen  $y_p$ ,  $p \in \{1,2\}$  mit den Werten  $y_p = 1$  und  $y_p = 0$  veranschaulicht. Die bedingte Erwartung ist dann als bedingte Wahrscheinlichkeit interpretierbar, da  $E(y_p \mid \xi) = 1 \cdot P(y_p = 1 \mid \xi) + 0 \cdot P(y_p = 0 \mid \xi) = P(y_p = 1 \mid \xi)$  gilt. In Gleichung 3.1 erkennen wir dann das bekannte Multiplikationstheorem der Wahrscheinlichkeitsrechnung (vgl. z.B. Bortz, 1977, S. 71), welches besagt, daß zwei Ereignisse dann bedingt stochastisch unabhängig sind, wenn ihre Verbundwahrscheinlichkeit  $P(y_1 = 1 \mid \xi, (y_2 = 1 \mid \xi)) = P(y_1 = 1 \mid \xi) \cdot P(y_2 = 1 \mid \xi)$  als Produkt der bedingten Einzelwahrscheinlichkeiten darstellbar ist.

In der allgemeinen Theorie latenter Variablen machen wir ausdrücklich keine Verteilungsannahmen. Für die beteiligten Zufallsvariablen setzen wir jedoch voraus, daß sie eine endliche Erwartung haben.

### 3.2.1.2 Variablencharakteristische Funktion (VC-Funktion)

Die bedingte Erwartung  $E(y_p \mid \xi)$  nennen wir im vorliegenden Zusammenhang auch variablencharakteristische Funktion, abgekürzt VC-Funktion<sup>8</sup>). Sie gibt für jeden Wert  $\xi$  der latenten Variablen  $\xi$  an, welchen (bedingten) Erwartungswert  $E(y_p, \mid \xi = \xi)$  die manifeste Variable  $y_p$ ,  $p \in \{1,2,\dots,P\}$ , hat. Die Verschiedenheit der hier behandelten Modelle zur Erklärung der Abhängigkeit

<sup>8</sup>) Andere Bezeichnungen sind: Variablen-, Itemcharakteristik, itemcharakteristische Funktion).

manifester Variablen durch eine latente Variable besteht darin, von welchem Typ die variablencharakteristische Funktion  $E(y_p | \xi)$  ist.

a) Bei *linearer VC-Funktion* legt man den Typ von  $E(y_p | \xi)$  durch die Gleichung

$$(3.2) \quad E(y_p | \xi) = \lambda_{p0} + \lambda_p \xi$$

fest. Nach Gleichung 3.2 ist die bedingte Erwartung der manifesten Variablen  $y_p$  eine lineare Funktion der latenten Variablen  $\xi$ . In Abbildung 3.2 ist eine lineare VC-Funktion für den speziellen Fall kontinuierlicher latenter und manifester Variablen dargestellt. Man beachte, daß bei einer diskreten latenten Variablen  $\xi$  (siehe z.B. Abb. 3.7) auch die Variablencharakteristik  $E(y_p | \xi)$  diskret ist.

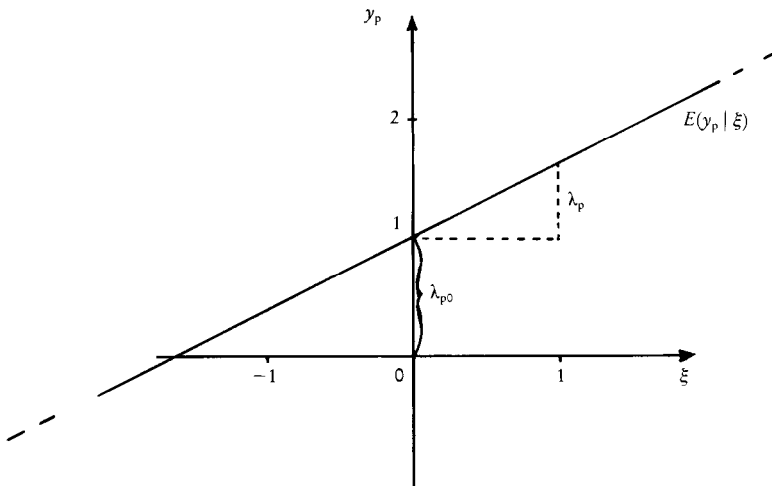


Abb. 3.2: Lineare VC-Funktion bei kontinuierlichen manifesten Variablen und einer kontinuierlichen latenten Variablen. Die Gerade  $E(y_p | \xi)$  gibt für jeden Wert  $\xi$  der latenten Variablen  $\xi$  den für  $y_p$  erwarteten Wert an. In diesem Fall wurden willkürlich  $\lambda_{p0} = 1.0$  und  $\lambda_p = 0.6$  angenommen.

b) Bei *polynomialer VC-Funktion* legt man den Typ von  $E(y_p | \xi)$  durch die Gleichung

$$(3.3) \quad E(y_p | \xi) = \lambda_{p0} + \lambda_{p1}\xi + \lambda_{p2}\xi^2 + \dots$$

fest. Die bedingte Erwartung der manifesten Variablen  $y_p$  ist dann eine kurvilinearere Funktion der latenten Variablen  $\xi$ , wie sie in Abb. 3.3 für Polynome 2. Ordnung dargestellt ist:

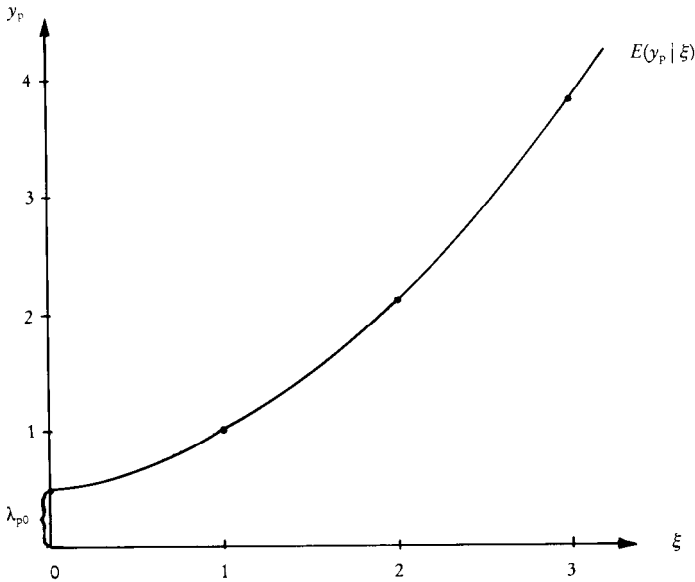


Abb. 3.3: Nichtlineare VC-Funktion bei kontinuierlichen manifesten Variablen und einer kontinuierlichen latenten Variablen in Form eines Polynoms 2. Ordnung, hier willkürlich mit  $\lambda_{p0} = 0.5$ ,  $\lambda_{p1} = 0.2$  und  $\lambda_{p2} = 0.3$ . Die Parabel  $E(y_p | \xi)$  gibt für den festen Wert  $\xi$  der latenten Variablen  $\xi$  den für  $y_p$  erwarteten Wert an.

c) Bei logistischer VC-Funktion legt man den Typ von  $E(y_p | \xi)$  durch die Gleichung

$$(3.4) \quad E(y_p | \xi) = \frac{\exp(\lambda_{p0} + \lambda_p \xi)}{1 + \exp(\lambda_{p0} + \lambda_p \xi)}$$

fest, wobei  $\exp(\lambda_{p0} + \lambda_p \xi) = e^{\lambda_{p0}} + \lambda_p \xi$  gilt mit  $e \cong 2.718$ , der Euler'schen Zahl. Die bedingte Erwartung der manifesten Variablen  $y_p$  ist dann eine logistische Funktion der latenten Variablen  $\xi$ . Der Zusammenhang ist in Abbildung 3.4, S. 31, dargestellt:

### 3.2.3 Die Residualvariablen

Bei Modellen mit einer latenten Variablen  $\xi$  ist der Zusammenhang zwischen  $\xi$  und den manifesten Variablen  $y_p$  nicht deterministisch, d.h. zwischen den  $y_p$ ,  $p \in \{1, 2, \dots, P\}$ , und deren bedingten Erwartungen  $E(y_p | \xi)$  wird in der Regel eine Differenz

$$(3.5) \quad \varepsilon_p := y_p - E(y_p | \xi),$$

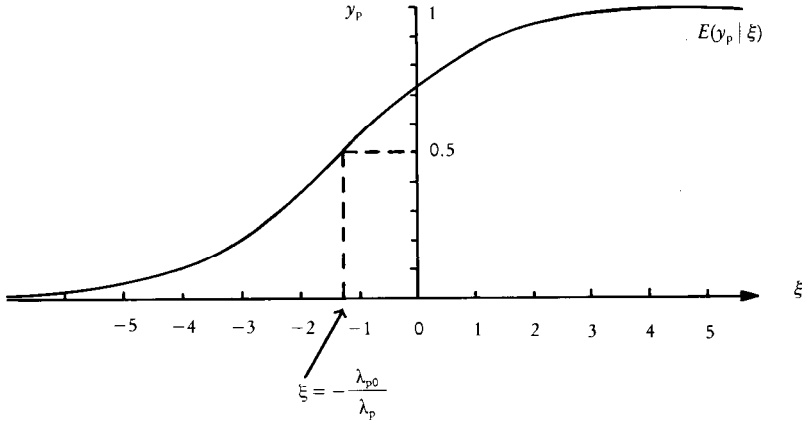


Abb. 3.4: Logistische VC-Funktion, hier willkürlich mit den Parametern  $\lambda_{p0} = 1.0$  und  $\lambda_p = 0.8$ . Der Parameter  $\lambda_p$  beschreibt die „Steilheit“ der VC-Funktion, der Parameter  $\lambda_{p0}$  ihre Lokation. Bei  $\lambda_{p0} + \lambda_p \xi = 0$ , d.h. bei  $\xi = -\lambda_{p0}/\lambda_p$  gilt immer  $E(y_p | \xi) = 1/2$ . Die bedingte Erwartung kann hier nur Werte zwischen Null und Eins annehmen. Auch bei diskreten manifesten Variablen  $y_p$  ist  $E(y_p | \xi)$  kontinuierlich.

bestehen, die wir als Residualvariablen bezeichnen. Diese haben die Eigenschaft, daß ihr bedingter Erwartungswert für alle Ausprägungen von  $\xi$  gleich Null ist:

$$(3.6) \quad E(\varepsilon_p | \xi) = 0$$

(siehe Anhang C, Regel 10). Für die Residualvariablen  $\varepsilon_p$  gelten die weiteren Eigenschaften

$$(3.7) \quad E(\varepsilon_p) = E[E(\varepsilon_p | \xi)] = 0,$$

(siehe Anhang C, Regel 1), und

$$(3.8) \quad C(\xi, \varepsilon_p) = C[\xi, y_p - E(y_p | \xi)] = 0$$

(siehe Anhang C, Regel 11). Nach Gleichung 3.7 ist also auch der unbedingte Erwartungswert der  $\varepsilon_p$  gleich Null, ebenso wie nach Gleichung 3.8 die Kovarianzen zwischen  $\xi$  und den  $\varepsilon_p$  Null sind.

Eine weitere Konsequenz aus der Modellannahme 3.1 ist, daß Residualvariablen  $\varepsilon_p$  paarweise nicht kovariieren

$$(3.9) \quad C(\varepsilon_p, \varepsilon_{p^*}) = 0, \text{ für alle } p, p^* \in \{1, 2, \dots, P\} \text{ mit } p \neq p^*.$$

## 3.2.4 Kovarianzmodellgleichungen und Identifikation

Im folgenden untersuchen wir für lineare VC-Funktionen, welche Folgerungen sich aus den Modellannahmen (Gleichungen 3.1 und 3.2) für die Modellgleichungen der Erwartungswerte, Varianzen und Kovarianzen der beobachtbaren Variablen  $y_p$  ergeben, und anschließend, ob und wie sich die Parameter  $\lambda_{p0}$ ,  $\lambda_p$ ,  $V(\xi)$  und  $V(\varepsilon_p)$  aus den Erwartungswerten, Varianzen und Kovarianzen der  $y_p$  eindeutig bestimmen lassen. Dies ist die Frage der Identifizierbarkeit. In manchen Fällen können, wie wir in Abschnitt 3.3.3 erwähnen, auch Modellgleichungen für Momente höherer Ordnung zur eindeutigen Bestimmung von Modellparametern beitragen.

Nach den Gleichungen 3.2 und 3.5 gilt für  $y_p$

$$(3.10) \quad y_p = \lambda_{p0} + \lambda_p \xi + \varepsilon_p.$$

Daraus und aus den Gleichungen 3.8 und 3.9 folgen die Modellgleichungen für die Kovarianzen der Variablen  $y_p$

$$(3.11) \quad \begin{aligned} C(y_p, y_{p^*}) &= C(\lambda_{p0} + \lambda_p \xi + \varepsilon_p, \lambda_{p^*0} + \lambda_{p^*} \xi + \varepsilon_{p^*}) = \\ &= \lambda_p C(\xi, \xi) \lambda_{p^*} + C(\varepsilon_p, \varepsilon_{p^*}) = \\ &= \begin{cases} \lambda_p \lambda_{p^*} V(\xi), & \text{für } p \neq p^* \\ \lambda_p^2 V(\xi) + V(\varepsilon_p), & \text{für } p = p^* \end{cases} \\ &\text{für alle } p, p^* \in \{1, 2, \dots, p\}. \end{aligned}$$

Für  $p = p^*$  geben die Gleichungen 3.11 die Modellgleichungen für die Varianz der beobachtbaren Variablen  $y_p$  an. Die durch die latente Variable  $\xi$  erklärte Varianz einer manifesten Variablen  $y_p$  ist gleich  $\lambda_p^2 V(\xi)$  und somit gleich  $\lambda_p^2$ , wenn  $\xi$  auf die Varianz  $V(\xi) = 1$  normiert wird.

Liegen drei beobachtbare Variablen  $y_p$  vor, wie in Abbildung 3.1 veranschaulicht, so implizieren die Annahmen 3.1 und 3.2 genau sechs Kovarianzmodellgleichungen (vgl. auch die „accounting equations“ von Lazarsfeld 1959, S. 504), welche die drei bekannten Varianzen und die drei bekannten Kovarianzen der beobachtbaren Variablen  $y_p$  mit den drei unbekanntem Ladungen  $\lambda_p$  und den drei Fehlervarianzen  $V(\varepsilon_p)$  verbinden. Die Gleichungen 3.11 enthalten in diesem Fall genau  $P(P + 1)/2 = 6$  Modellgleichungen für die Kovarianzen, die gerade genügen, um eindeutig die sechs Parameter  $\lambda_1, \lambda_2, \lambda_3, V(\varepsilon_1), V(\varepsilon_2), V(\varepsilon_3)$  zu bestimmen, wenn wir entweder die Varianz der latenten Variablen  $\xi$  festlegen oder einem Parameter  $\lambda_p$  einen bestimmten festen Wert geben:

Wir leiten zunächst aus den Modellgleichungen 3.11 für die Kovarianzen der beobachtbaren Variablen die Identifikationsgleichung für  $\lambda_1$  ab. Dazu bilden wir durch dreimaliges Benutzen von Gleichung 3.11 für  $p \neq p^*$  den Ausdruck

$$(3.12) \quad \frac{C(y_1, y_2) \cdot C(y_1, y_3)}{C(y_2, y_3)} = \frac{\lambda_1 V(\xi) \lambda_2 \cdot \lambda_1 V(\xi) \lambda_3}{\lambda_2 V(\xi) \lambda_3} = \lambda_1^2 V(\xi).$$

Für den Koeffizienten  $\lambda_1$  erhalten wir demnach

$$(3.13) \quad \lambda_1 = \pm \sqrt{\frac{C(y_1, y_2) \cdot C(y_1, y_3)}{C(y_2, y_3) \cdot V(\xi)}}.$$

Wenn wir die Variablen  $y_p$  so definieren, daß sie positiv mit  $\xi$  kovariieren, so hat  $\lambda_1$  immer ein positives Vorzeichen. Da die Kovarianzen der manifesten Variablen  $y_p$  gegebene Größen sind, sieht man aus den Gleichungen 3.12 und 3.13, daß mit der Festlegung der Varianz  $V(\xi)$  der Koeffizient  $\lambda_1$  gegeben ist bzw. umgekehrt mit der Festlegung von  $\lambda_1$  auch  $V(\xi)$ . Da die Numerierung der Variablen  $y_p$  beliebig ist, kann man Gleichung 3.13 entsprechend für jede der Variablen  $y_p$ ,  $p \in \{1, 2, 3\}$  verwenden.

Wenn wir bedenken, daß wegen  $C(y_p, \xi) = C(\lambda_{p0} + \lambda_p \xi + \varepsilon_p, \xi) = \lambda_p V(\xi)$  für die Koeffizienten  $\lambda_p$

$$(3.14) \quad \lambda_p = \frac{C(y_p, \xi)}{V(\xi)}$$

gilt, so sehen wir, daß  $\lambda_p$  um so größer wird, je kleiner die Varianz von  $\xi$  festgelegt wird. Darüber hinaus ist zu beachten, daß  $\lambda_p > \lambda_{p^*}$  nur dann auf einen stärkeren Zusammenhang zwischen  $y_p$  und  $\xi$  gegenüber  $y_{p^*}$  und  $\xi$  hinweist, wenn  $y_p$  und  $y_{p^*}$  gleiche Varianzen haben.

Im Abschnitt 3.3.3 werden wir einen Fall kennenlernen, bei dem  $V(\xi)$  nicht beliebig festzulegen ist. In diesem Fall hat  $V(\xi)$  eine wichtige empirische Bedeutung, und wir werden sehen, daß durch Berücksichtigung der Modellstruktur für die zentralen Momente dritter Ordnung  $C(y_1, y_2, y_3)$  in diesem Fall auch die Varianz  $V(\xi)$  eindeutig bestimmt werden kann.

Aus den Gleichungen 3.11 (für  $p = p^*$ ) erhalten wir auch die Identifikationsgleichungen für die Fehlervarianzen  $V(\varepsilon_p)$ , wenn wir für  $\lambda_p^2 V(\xi)$  den linken Ausdruck aus Gleichung 3.12 einsetzen. Die folgende Gleichung gilt wieder für alle Variablen  $y_p$ , wenn man die Indices entsprechend umnumeriert:

$$(3.15) \quad V(\varepsilon_1) = V(y_1) - \frac{C(y_1, y_2) \cdot C(y_1, y_3)}{C(y_2, y_3)}$$

Sind die Parameter  $\lambda_p$  und  $V(\varepsilon_p)$  erst einmal eindeutig bestimmt, so kann man auch die Skalenkonstanten  $\lambda_{p0}$  unter Verwendung von Gleichung 3.10 aus

$$(3.16) \quad E(y_p) = \lambda_{p0} + \lambda_p E(\xi)$$

eindeutig bestimmen, wenn man  $E(\xi)$  festlegt. Gewöhnlich wird die Normierung  $E(\xi) = 0$  gewählt.

Allgemein erhält man bei einem Modell mit  $P$  kongenerischen Variablen genau  $P(P+1)/2$  Kovarianzgleichungen (siehe Gleichungen 3.11) mit  $2 \cdot P$  unbekannt Parametern, nämlich den  $\lambda_p$  und  $V(\varepsilon_p)$ , und außerdem  $P$  Gleichungen für die Erwartungswerte der  $y_p$  mit weiteren  $P$  unbekannt Parametern, nämlich den Skalenkonstanten  $\lambda_{p0}$ . Liefert ein Modell nur so viele Gleichungen für die Kovarianzen wie zu bestimmende Parameter, so kann das Modell zwar identifizierbar sein, es ist aber empirisch noch nicht überprüfbar, da es zu einer perfekten Anpassung des Modells an die Daten einer Stichprobe führen würde. Erst wenn  $P > 3$  ist, hat man mehr Gleichungen als Unbekannte, und man spricht von Überidentifikation. In diesem Fall können zumindest einige Parameter durch mehrere Gleichungen bestimmt werden und das Modell wird empirisch überprüfbar.

### 3.2.5 Beispiel

Ein fiktives Beispiel soll die oben dargestellte Theorie erläutern. Wir gehen davon aus, daß die folgenden Varianzen, Kovarianzen und Korrelationen zwischen den kongenerischen Variablen  $y_1$ ,  $y_2$  und  $y_3$  für eine bestimmte Population bekannt und in Tabelle 3.1 gesammelt sind. Dabei enthält die untere Dreiecksmatrix die Korrelationen, die obere die Kovarianzen und die Hauptdiagonale die Varianzen:

Tabelle 3.1: Korrelations- (untere Dreiecksmatrix) und Kovarianzmatrix (obere Dreiecksmatrix) fiktiver Daten.

	$y_1$	$y_2$	$y_3$
$y_1$	4.00	0.90	4.20
$y_2$	0.90	0.25	1.20
$y_3$	0.70	0.80	9.00

Die Parameter  $\lambda_p$  und  $V(\varepsilon_p)$ , die unter der Normierung  $V(\xi) = 1$  nach den Gleichungen 3.13 und 3.15 berechnet wurden, sind in Tabelle 3.2 gesammelt.

Wählen wir z.B. eine Personengruppe A mit der Ausprägung  $\xi_A$  in der latenten Variablen und eine Personengruppe B mit der um eine Einheit höheren Ausprägung  $\xi_B = \xi_A + 1$ , so erwarten wir nach Gleichung 3.2 für Personen der

Tabelle 3.2: Modellparameter für das fiktive Zahlenbeispiel der Tabelle 3.1.

	$\lambda_p$	$V(\varepsilon_p)$
$y_1$	1.77	0.87
$y_2$	0.50	0.00
$y_3$	2.37	3.38

Gruppe B verglichen mit Personen der Gruppe A in der manifesten Variablen  $y_1$  eine um 1.77, in  $y_2$  eine um 0.50 und in  $y_3$  eine um 2.37 Einheiten „höhere“ Merkmalsausprägung. Eine vergleichende Reihung der Gewichte  $\lambda_p$  hinsichtlich des Grades der Abhängigkeit der manifesten Variablen  $y_p$  von der latenten Variablen  $\xi$  ist erst möglich, wenn alle beobachteten Variablen auf die gleiche Varianz normiert werden.

Bis auf Rundungsfehler läßt sich mit diesen Daten und den Gleichungen 3.11 die Ausgangskovarianzmatrix exakt reproduzieren. Bei drei Variablen  $y_p$  läßt sich diese perfekte Anpassung des Modells an die Daten mit jeder beliebigen Korrelationsmatrix erzielen, in der alle Eintragungen von Null verschieden sind. Erst bei vier beobachteten Variablen  $y_p$  wäre dies nicht zwangsläufig der Fall, womit das Modell dann empirisch überprüfbar wäre. Dabei ist die inhaltlich wichtigste Frage, ob eine latente Variable zur Erklärung der beobachteten Zusammenhänge ausreicht.

### 3.2.6 Modell paralleler Variablen

Führt man noch restriktivere Annahmen ein, als sie in der formalen Theorie latenter Variablen mit linearer VC-Funktion (Gleichungen 3.1 und 3.2) gemacht wurden, erhält man das Modell paralleler Variablen (vgl. hierzu den Begriff „parallele Tests“ in der klassischen Testtheorie, z.B. in Fischer, 1974, S. 34). Beim Modell paralleler Variablen werden zusätzlich die Gleichungen

$$(3.17) \quad \lambda_{p0} = \lambda_0; \lambda_p^2 = \lambda^2; V(\varepsilon_p) = V(\varepsilon), \text{ für alle } p \in \{1, 2, \dots, P\},$$

angenommen. Parallele Variablen haben demnach gleiche Erwartungswerte (das folgt aus den Gleichungen 3.16 und 3.17), gleiche erklärte Varianzen  $\lambda^2 V(\xi)$  (siehe Gleichungen 3.11 für  $p = p^*$ ) und gleiche Fehlervarianzen. Parallele Variablen sind in diesem Sinn austauschbar, denn sie messen dieselbe latente Eigenschaft auf der gleichen Skala gleich gut. Kongenerische Variablen, die nicht zugleich parallel sind, messen hingegen dieselbe Eigenschaft auf verschiedenen Skalen verschieden gut.

### 3.2.7 Zusammenfassende Bemerkungen

Die in diesem Abschnitt dargestellte allgemeine formale Theorie liegt den in den folgenden Abschnitten zu behandelnden Modellen zugrunde. Sie besteht aus zwei Grundannahmen, der bedingten stochastischen Unabhängigkeit (siehe Gleichung 3.1) und der Art der Variablencharakteristik (siehe Gleichungen 3.2 bis 3.4). Diese beiden Annahmen genügen zur eindeutigen Bestimmung der Modellparameter (siehe für lineare VC-Funktion die Gleichungen 3.13 und 3.15) sofern Varianzen und Kovarianzen von genügend vielen beobachtbaren Variablen vorliegen und die Varianz der latenten Variablen festgelegt wird. Bei restriktiveren Annahmen über die Modellparameter genügen für deren eindeutige Bestimmung bereits zwei beobachtbare kongenerische Variablen, die wir dann auch „parallel“ nennen.

*Weiterführende Literatur:* Die Theorie latenter Variablen ist zunächst nur in Ansätzen behandelt worden (siehe z.B. Anderson 1959, Lord & Novick 1968). Eine monographische Abhandlung liefert Harnerle (1982).

## 3.3 Anwendungen der allgemeinen Theorie latenter Variablen

### 3.3.1 Faktorenanalyse

Als erste Anwendung der im Abschnitt 3.2 dargestellten Theorie behandeln wir die lineare Faktorenanalyse mit einem Faktor (einer latenten Variablen).

Von der Faktorenanalyse in dem hier verwendeten engeren Sinn (Spearman 1904, 1926, Thurstone 1947, Jöreskog 1966) muß die „Hauptkomponentenanalyse“ (Pearson 1901, Hotelling 1933) unterschieden werden. Während es Ziel der Faktorenanalyse ist, die Kovarianzmatrix von kongenerischen Variablen durch eine geringe Anzahl von Faktoren (latenten Variablen) zu erklären, beschränkt sich die Hauptkomponentenmethode darauf, eine Kovarianzmatrix von Variablen durch orthogonale Komponenten, welche auch Faktoren genannt werden, aber dem theoretischen Anspruch latenter Variablen meist nicht genügen, zu beschreiben (vgl. z.B. Lawley & Maxwell, 1971, S. 2-4). Die Hoffnung in der bisherigen Forschung, exploratorisch mit der häufig angewendeten Hauptkomponentenmethode geeignete Theorien „automatisch“ zu finden, konnte sich deshalb nur kaum erfüllen. Doch seit Jöreskog (1967, 1969, 1973) rechentechnische Probleme in der Faktorenanalyse gelöst hat, ist auch die Überprüfung von Theorien mit konfirmatorischen Faktorenanalysen praktisch leicht durchführbar (siehe Sörbom & Jöreskog, 1976, Jöreskog & Sörbom, 1978).

In der Faktorenanalyse werden sowohl die latente Variable  $\xi$  (der Faktor) als auch die beobachtbaren Variablen  $y_p$  als kontinuierlich angenommen. Dann ist

auch die lineare Variablencharakteristik  $E(y_p|\xi)$  kontinuierlich (siehe Abbildung 3.5). Sie kann als Regressionsgerade zwischen der jeweiligen Variablen  $y_p$  und der latenten Variablen  $\xi$  angesehen werden. Die Werte  $\xi$  von  $\xi$  heißen in der Faktorenanalyse „factor scores“.

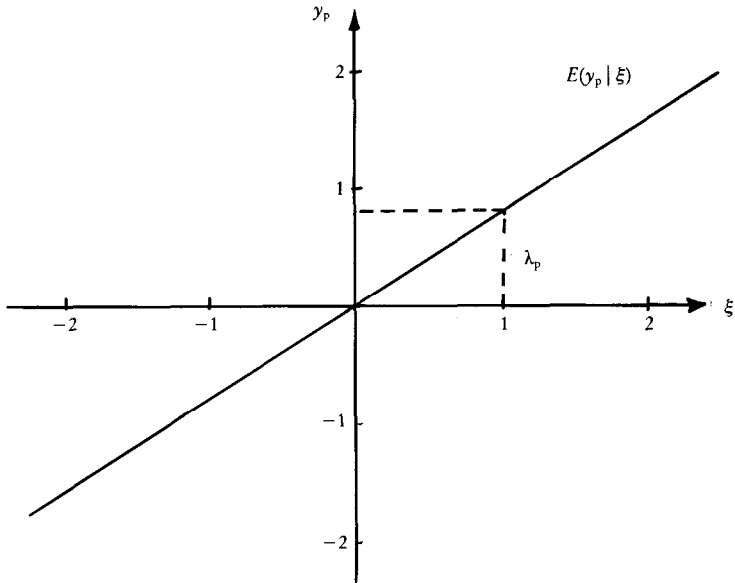


Abb. 3.5: Faktorenanalytisches Modell mit einem Faktor als Spezialfall der allgemeinen Theorie latenter Variablen mit linearer VC-Funktion entsprechend Gleichung 3.2 sowie kontinuierlichen manifesten und kontinuierlicher latenten Variablen. Die Gerade  $E(y_p|\xi)$  gibt für jeden Wert  $\xi$  (factor score) der latenten Variablen  $\xi$  den für  $y_p$  erwarteten Wert an. In der Faktorenanalyse sind die Normierungen  $E(y_p) = E(\xi) = 0$  üblich, wodurch die Skalenkonstanten  $\lambda_{p0} = 0$  sind. Die Ladung  $\lambda_p$  wurde hier willkürlich mit  $\lambda_p = 0.8$  angenommen.

Die Gewichte  $\lambda_p$  werden in der Faktorenanalyse „Ladungen“ der Variablen  $y_p$  auf  $\xi$  genannt. Ist die Varianz  $V(\xi)$  auf Eins normiert, so ist  $\lambda_p^2$  die „Kommunalität“ der Variablen  $y_p$ . Die Gleichungen 3.11 zeigen, daß es sich dabei um die durch  $\xi$  erklärte Varianz von  $y_p$  handelt, also um  $V[E(y_p|\xi)]$ .

Als klassischen Anwendungsfall können wir Spearman's Generalfaktor-Theorie betrachten (Spearman, 1904).

*Weiterführende Literatur:* Als deutschsprachige Lehrbücher seien Pawlik (1968), Revenstorf (1976, 1980) und Überla (1971) genannt. Revenstorf (1980), Anderson & Rubin (1956) sowie Lawley & Maxwell (1971) gehen ausführlicher auf die statistischen Probleme der Faktorenanalyse ein, als die Monographien von Mulaik (1972), Harman (1967) oder Horst (1965), welche ebenso wie die oben angeführten deutschsprachigen Autoren mehr Platz der Hauptkomponentenanalyse widmen. Die nichtlineare Faktorenanalyse (entsprechend Gl. 3.3) wird z.B. von McDonald (1967) behandelt.

### 3.3.2 „Linear traceline model“

In diesem Abschnitt untersuchen wir, welche Besonderheiten sich für die Interpretation einiger Gleichungen und Parameter ergeben, wenn man zusätzlich zu den Modellannahmen der Gleichungen 3.1 und 3.2 festlegt, daß die latente Variable kontinuierlich ist, die manifesten Variablen aber zweiwertig sind.

Lazarsfeld (vgl. Lazarsfeld & Henry, 1968, S. 206) hat ein Modell eingeführt, das von einer linearen Beziehung zwischen der kontinuierlichen latenten Variablen  $\xi$  und den zweiwertigen beobachtbaren Variablen  $y_p$  ausgeht, das „linear traceline model“. Wir zeigen daher in diesem Abschnitt zunächst, daß das Lazarsfeld'sche Modell mit linearer Variablencharakteristik („linear traceline“) ein Spezialfall des allgemeinen Modells latenter Variablen ist, wenn man von zweiwertigen manifesten Variablen und einer kontinuierlichen latenten Variablen ausgeht.

In den Gleichungen 3.1 und 3.2 sind die Annahmen des allgemeinen linearen Modells latenter Variablen formuliert. Die zusätzlichen Annahmen des „linear traceline model“ bestehen in der notwendigen (vgl. dazu Wottawa, 1979, S. 49) Einschränkung von  $\xi$  auf das Intervall  $[\xi_1, \xi_2]$  und in der Festlegung, daß  $\xi$  kontinuierlich ist, die  $y_p$  aber nur die Werte 0 bzw. 1 annehmen können. Wir untersuchen nun, welche Folgerungen sich aus diesen Festlegungen ergeben:

Die erste Besonderheit betrifft die lineare VC-Funktion, für die in diesem Fall gilt:

$$(3.18) \quad E(y_p | \xi) = P(y_p = 1 | \xi) = \lambda_{p0} + \lambda_{p1}\xi.$$

Demnach ist die bedingte Wahrscheinlichkeit, daß  $y_p$  den Wert 1 unter  $\xi = \xi$  annimmt, eine lineare Funktion der latenten Variablen  $\xi$  (siehe Abbildung 3.6). Hier wird deutlich, daß das allgemeine Modell latenter Variablen zu einem „probabilistischen Modell“ wird, wenn  $y_p$  nur zwei Werte annehmen kann, welche mit 0 und 1 kodiert werden (vgl. auch Abschnitt 3.3.4). Der einzige Unterschied zwischen dem faktorenanalytischen und dem probabilistischen Modell in Gleichung 3.18 besteht darin, daß die Werte der bedingten

Erwartung  $E(y_p|\xi)$  im letzteren Fall als bedingte Wahrscheinlichkeiten interpretiert werden können.

Bei psychometrischen Tests, in denen  $\xi$  eine latente Fähigkeit repräsentiert, kann bei der Kodierung  $y_p = 0$  für „Nichtlösung“ und  $y_p = 1$  für „Lösung“ einer Aufgabe  $p$  die Konstante  $\lambda_{p0}$  als Leichtigkeit der Aufgabe  $p$  interpretiert werden. Die unbedingten Erwartungswerte der beobachtbaren Variablen

$$(3.19) \quad E(y_p) = E(\lambda_{p0} + \lambda_p \xi + \varepsilon_p) = \lambda_{p0} + \lambda_p E(\xi),$$

vereinfachen sich nämlich bei der Normierung  $E(\xi) = 0$  zu

$$(3.20) \quad E(y_p) = \lambda_{p0} = P(y_p = 1).$$

Man sieht, daß  $\lambda_{p0}$  hier die Lösungswahrscheinlichkeit der Aufgabe  $p$  darstellt, welche in der klassischen Testtheorie (vgl. Lienert 1969, S. 87, od. Moosbrugger, 1981 b, S. 27) als Aufgabenschwierigkeit bezeichnet wird.

Der Parameter  $\lambda_p$  zeigt die „Trennschärfe“ einer Aufgabe an. Je größer  $\lambda_p$  ist, desto größer der Zusammenhang (siehe Gleichung 3.14) zwischen Fähigkeit und Lösung dieser Aufgabe, und desto weniger hängt das Lösen der Aufgabe von anderen, nicht erfaßten Umständen ab.

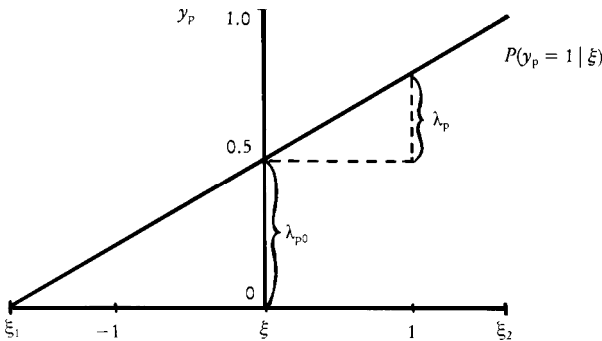


Abb. 3.6: Linear trace-line model. Die bedingte Wahrscheinlichkeit, daß die Variable  $y_p$  den Wert 1 annimmt, ist eine lineare Funktion der Ausprägung der latenten Variablen  $\xi$ . Man sieht, daß die bedingte Erwartung (Wahrscheinlichkeit) auch bei diskreten  $y_p$  kontinuierlich ist. Wurde  $\xi$  normiert mit  $E(\xi) = 0$ , dann ist  $\lambda_{p0}$  die unbedingte Wahrscheinlichkeit für „ $y_p = 1$ “. Bei Fähigkeitstests kann  $\lambda_{p0}$  deshalb als „Leichtigkeit“ interpretiert werden. Der Parameter  $\lambda_p$  gibt den Wert an, um den die Wahrscheinlichkeit für „ $y_p = 1$ “ ansteigt, wenn sich  $\xi$  um eine Einheit erhöht. Zugleich ist  $\lambda_p$  bei  $V(\xi) = 1$  die Kovarianz zwischen  $y_p$  und  $\xi$  (siehe Gleichung 3.14). In diesem Fall wurden willkürlich  $\lambda_{p0} = 0.5$  und  $\lambda_p = 0.3$  angenommen.

*Weiterführende Literatur:* Lazarsfeld & Henry (1968) schreiben an mehreren Stellen ausführlicher über das „linear tracement model“. Auch Fischer (1974, S. 160ff.) geht auf dieses Modell ein. Ein Anwendungsbeispiel findet man in Lazarsfeld (1959). Auch die polynomiale latent structure analysis von Lazarsfeld (1950) kann als Spezialfall der allgemeinen Theorie gemäß Gleichung 3.1 und Gleichung 3.3 dargestellt werden.

### 3.3.3 „Latent profile-“ und „latent class model“

Wir betrachten in diesem Abschnitt den Fall, daß es sich bei der latenten Variablen  $\xi$  um eine zweiwertige Kodiervariable handelt, die nur die Werte „Null“ oder „Eins“ annehmen kann und somit die Zugehörigkeit zu einer von zwei Klassen anzeigt. Die manifesten Variablen  $y_p$ ,  $p \in \{1,2,\dots,P\}$ , hingegen können kontinuierlich oder ebenfalls zweiwertig sein. Gibson (1959) nennt den ersten Fall „latent profile model“, auf den auch Lazarsfeld & Henry (1968, S. 228) sowie McDonald (1967, S. 58) eingehen. Der Fall, daß sowohl die latente Variable als auch die manifesten Variablen dichotom sind, ist unter der Bezeichnung „latent class model“ bekannt (vgl. Lazarsfeld & Henry, 1968, S. 17). Beide Modelle sind Spezialfälle der allgemeinen Theorie, die wir in Abschnitt 3.2 dargestellt haben, d.h. die beiden Modelle sind charakterisiert durch die Gleichungen 3.1 und 3.2 sowie die Festlegungen, daß beim „latent profile model“  $y_p$  kontinuierlich für alle  $p \in \{1,2,\dots,P\}$ , und  $\xi$  zweiwertig mit den Werten 0 oder 1 sind, wohingegen beim „latent class model“ sowohl  $y_p$  als auch  $\xi$  zweiwertig ist mit den Werten 0 oder 1, für alle  $p \in \{1,2,\dots,P\}$ .

Wie man durch Berechnung der bedingten Erwartungswerte sehen kann, vereinfacht sich in diesen Fällen der Parameter  $\lambda_{p0}$  zu

$$(3.21) \quad E(y_p \mid \xi = 0) = E(\lambda_{p0} + \lambda_p \xi + \varepsilon_p \mid \xi = 0) = \lambda_{p0},$$

(siehe Gleichung 3.6 und Anhang A, Regel 3), und im Fall manifester zweiwertiger Variablen gilt

$$(3.22) \quad P(y_p = 1 \mid \xi = 0) = \lambda_{p0}.$$

Demnach ist also  $\lambda_{p0}$  der bedingte Erwartungswert von  $y_p$  bzw. die bedingte Wahrscheinlichkeit für „ $y_p = 1$ “ im Fall  $\xi = 0$ . Auf entsprechende Weise finden wir für  $\lambda_p$  im Fall kontinuierlicher manifester Variablen

$$(3.23) \quad E(y_p \mid \xi = 1) = E(\lambda_{p0} + \lambda_p \xi + \varepsilon_p \mid \xi = 1) = \lambda_{p0} + \lambda_p,$$

(siehe Gleichung 3.5 und Anhang A, Regel 3), und durch Einsetzen der Gleichung 3.21 erhalten wir die Gleichung

$$(3.24) \quad \lambda_p = E(y_p \mid \xi = 1) - E(y_p \mid \xi = 0),$$

die sich im Fall zweiwertiger manifester Variablen vereinfacht zu

$$(3.25) \quad \lambda_p = P(y_p = 1 \mid \xi = 1) - P(y_p = 1 \mid \xi = 0).$$

Nach diesen Gleichungen ist  $\lambda_p$  im latent profile model gleich der Differenz der bedingten Erwartungswerte von  $y_p$  unter „ $\xi = 1$ “ und „ $\xi = 0$ “ (siehe Abb. 3.7) bzw. im latent class model der Differenz der bedingten Wahrscheinlichkeiten für „ $y_p = 1$ “ unter „ $\xi = 1$ “ und „ $\xi = 0$ “ (siehe Abb. 3.8). Ist  $\lambda_p$  gleich Null, so differenzieren die beiden latenten Klassen nicht hinsichtlich den betreffenden manifesten Variablen  $y_p$ .

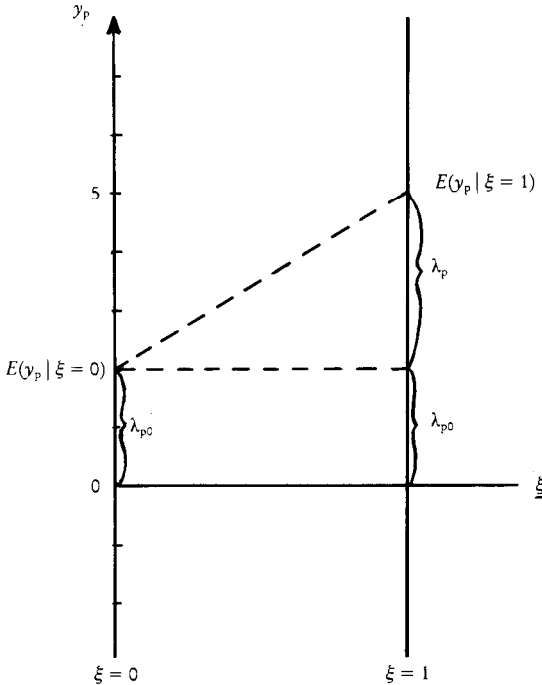


Abb. 3.7: Latent profile model als Spezialfall der allgemeinen Theorie latenter Variablen mit linearer VC-Funktion gemäß Gleichung 3.2 sowie kontinuierlichen manifesten und dichter 0/1 kodierter latenten Variablen. Die bedingte Erwartung hat dann nur zwei Werte, nämlich  $\lambda_{p0}$  (hier 2) auf der Stufe  $\xi = 0$  und  $\lambda_{p0} + \lambda_p$  (hier  $2 + 3 = 5$ ) auf der Stufe  $\xi = 1$ .

Die grundlegende Modellvorstellung der bedingten stochastischen Unabhängigkeit der  $y_p$  (siehe Gleichung 3.1) bedeutet hier, daß innerhalb einer latenten Klasse die manifesten Variablen nicht korrelieren und auch in keiner Abhängigkeit höherer Ordnung stehen (vgl. Beispiel 3.2.5).

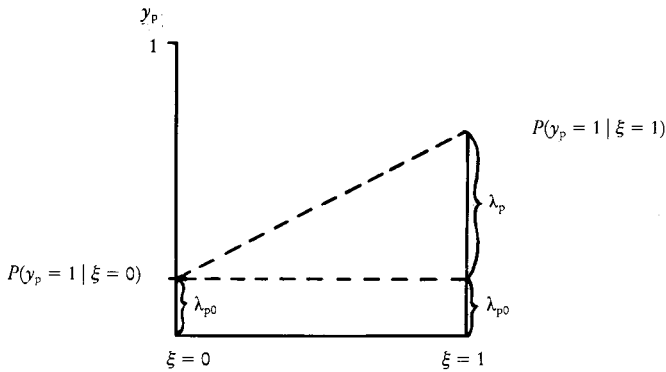


Abb. 3.8: Latent class model als Spezialfall der allgemeinen Theorie latenter Variablen mit linearer VC-Funktion gemäß Gleichung 3.2 sowie dichotomen 0/1 kodierten manifesten und dichotomer 0/1 kodierter latenter Variablen. Die bedingte Erwartung hat dann nur zwei Werte, nämlich die Wahrscheinlichkeiten für  $y_p = 1$  auf der Stufe  $\xi = 0$ , welche nach Gleichung 3.22 gleich  $\lambda_{p0}$  (hier 0.2) ist, sowie die Wahrscheinlichkeit für  $y_p = 1$  auf der Stufe  $\xi = 1$  (hier 0.7). Die Differenz beider Wahrscheinlichkeiten beschreibt nach Gleichung 3.25 der Parameter  $\lambda_p$  (hier 0.5).

Bei einem Modell, bei dem eine dichotome latente Variable mit Null bzw. Eins kodiert wird, kann man die Standardisierung  $E(\xi) = 0$  und  $V(\xi) = 1$  nicht einführen, da diese beiden Kennwerte hier eine besondere Bedeutung haben, denn

$$(3.26) \quad E(\xi) = P(\xi=1)$$

und

$$(3.27) \quad V(\xi) = P(\xi=1) \cdot [1 - P(\xi=1)]$$

(siehe Anhang A, Gleichung A.1).

Die Wahrscheinlichkeit, einer der beiden latenten Klassen anzugehören, kann theoriegeleitet a priori festgelegt oder empirisch ermittelt werden, wenn man die Modellstrukturgleichung für das zentrale Moment dritter Ordnung zur Identifikation der Klassenwahrscheinlichkeiten  $P(\xi=1)$  und  $P(\xi=0)$  benutzt. Näheres siehe Steyer & Moosbrugger, 1979, S. 38-41 sowie Lazarsfeld & Henry, 1968, S. 42. Ein illustratives Beispiel findet man in Lazarsfeld & Henry (1968, S. 27ff.).

Weiterführende Literatur: Das „latent class model“ wird ausführlich in Lazarsfeld & Henry (1968) behandelt. Einen kurzen Überblick gibt Madansky (1978). Mit Problemen der Parameterschätzung befassen sich Goodman (1974, 1979) und Formann (1978, 1980).

### 3.3.4 Logistische Modelle

Bei dem in Abschnitt 3.3.2 dargestellten „linear traseline model“ mußten wir  $\xi$  auf ein bestimmtes Intervall einschränken. Andernfalls könnte die bedingte Erwartung auch Werte größer als Eins oder kleiner als Null annehmen (vgl. auch Wottawa, 1980, S. 49), welche mit der Wahrscheinlichkeitsinterpretation unvereinbar wären. Beim logistischen Testmodell von Birnbaum (1968) und dessen Spezialfall, dem dichotomen logistischen Testmodell von Rasch (1960) wird diesem Problem vorgebeugt, indem außer der bedingten stochastischen Unabhängigkeit (siehe Gleichung 3.1) anstelle der linearen VC-Funktion (siehe Gleichung 3.2) die logistische VC-Funktion (siehe Gleichung 3.4) angenommen wird. Bei dichotomen manifesten Variablen  $y_p$  mit den Werten 0 bzw. 1 werden die Testmodelle „probabilistisch“ und als bedingte Erwartung erhalten wir folgende bedingte Wahrscheinlichkeit (vgl. Gleichung 3.4):

$$(3.27) \quad E(y_p | \xi) = P(y_p = 1 | \xi) = \frac{\exp(\lambda_{p0} + \lambda_p \xi)}{1 + \exp(\lambda_{p0} + \lambda_p \xi)}$$

In der logistischen Modellgleichung 3.27 kommen zwei Parameter vor, die spezifisch für die Variable  $y_p$  sind;  $\lambda_p$  beschreibt die „Steilheit“ der VC-Funktion,  $\lambda_{p0}$  ihre Lokalisation entlang der  $\xi$ -Achse (vgl. Abbildung 3.4). In psychometrischen Anwendungen können wir  $\lambda_p$  als Maß für die Trennschärfe („discriminating power“ sensu Birnbaum 1968, S. 409) einer Variablen auffassen und  $\lambda_{p0}$  als „Leichtigkeit“. Anstelle von  $\lambda_{p0}$  verwendet Birnbaum (1968, S. 409) einen Schwierigkeitsparameter  $\lambda_{p0}^*$ , welcher nach Ausklammern von  $\lambda_p$  in Gleichung 3.27 und Setzen von  $\lambda_{p0}/\lambda_p = -\lambda_{p0}^*$  gefunden wird. Die Modellgleichung 3.27 lautet dann

$$(3.28) \quad E(y_p | \xi) = P(y_p = 1 | \xi) = \frac{\exp[\lambda_p(\xi - \lambda_{p0}^*)]}{1 + \exp[\lambda_p(\xi - \lambda_{p0}^*)]}$$

Der Schwierigkeitsparameter  $\lambda_{p0}^*$  ermöglicht einen direkten Vergleich mit dem Personenparameter  $\xi$ , welcher bei Leistungstests zu besonders einfachen Parameterinterpretationen führt: Bezeichnet nämlich  $\xi$  die „Fähigkeit“ einer Person in der latenten Variablen  $\xi$  und  $\lambda_{p0}^*$  die Schwierigkeit der Aufgabe, so beträgt wegen  $\exp(0) = 1$  die Lösungswahrscheinlichkeit  $P(y_p = 1 | \xi = \lambda_{p0}^*) = 1/2$ , wenn  $\xi = \lambda_{p0}^*$  ist, nämlich genau dann, wenn die Fähigkeit der Person ebenso groß wie die Schwierigkeit der Aufgabe ist. Bei positiv gewählten  $\lambda_p$  gilt  $P(y_p = 1 | \xi) > 1/2$ , wenn  $\xi > \lambda_{p0}^*$  und umgekehrt.

Ein interessanter Spezialfall des logistischen Testmodells und damit der allgemeinen Theorie latenter Variablen ist das dichotome logistische Testmodell von Rasch (1960), bei dem alle Parameter  $\lambda_p$ ,  $p \in \{1, 2, \dots, P\}$  gleich Eins gesetzt werden. Das bedeutet, daß alle Variablen  $y_p$  die latente Variable  $\xi$  „gleich gut“ messen sollen, aber auf im allgemeinen unterschiedlichem Schwierigkeitsniveau. Demnach verlaufen die VC-Funktionen der verschiedenen  $y_p$  parallel, aber jeweils um die Differenz ihrer  $\lambda_{p0}^*$  entlang der  $\xi$ -Achse verschoben. Abbildung 3.9 illustriert diesen Sachverhalt:

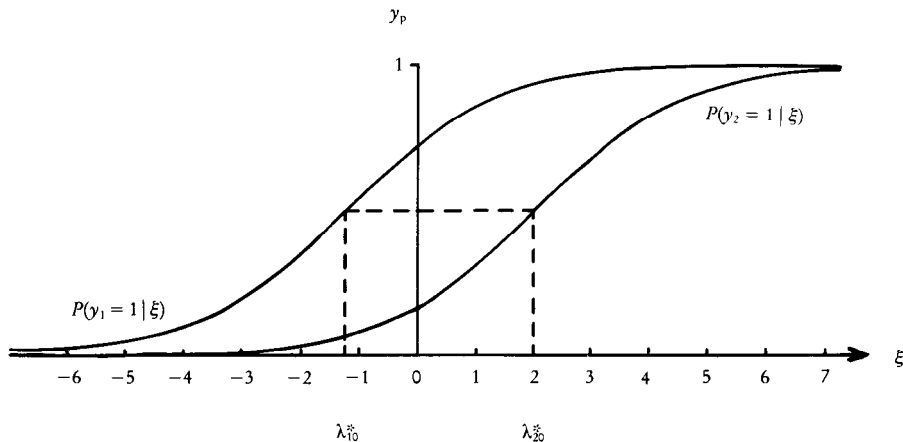


Abb. 3.9: Dichotomes logistisches Testmodell nach Rasch (1960) als Spezialfall der allgemeinen Theorie latenter Variablen mit logistischer VC-Funktion entsprechend Gleichung 3.4 sowie dichotomen 0/1 kodierten manifesten und kontinuierlicher latenten Variablen. Die Kurven  $P(y_p = 1 | \xi)$  geben für jeden Wert  $\xi$  der latenten Variablen  $\xi$  die Wahrscheinlichkeit für die Ausprägung  $y_p = 1$  an; bei psychometrischen Anwendungen ist das die „Lösungswahrscheinlichkeit“ der Aufgabe  $p$ . Die VC-Funktionen der Aufgaben  $p = 1$  und  $p = 2$  haben gleiche „Trennschärfen“ ( $\lambda_1 = \lambda_2 = 1$ ), aber unterschiedliche Schwierigkeiten  $\lambda_{10}^* = -1.25$  und  $\lambda_{20}^* = 2.0$ . Ihre Schwierigkeitsdifferenz beträgt  $\lambda_{20}^* - \lambda_{10}^* = 3.25$ . Für Personen mit  $\xi = \xi_1 = -1.25$  erhalten wir bei Aufgabe 1 die Lösungswahrscheinlichkeit  $1/2$  wie auch für Personen mit  $\xi = \xi_2 = 2.0$  bei Aufgabe 2. Mit der Wahl konstanter  $\lambda_p$  ist auch die Varianz der latenten Variablen festgelegt.

Die bei Modellgültigkeit gleichen  $\lambda_p$  für alle  $p \in \{1, 2, \dots, P\}$  Aufgaben ermöglichen „spezifisch objektive“, d.h. aufgaben- bzw. personenstichprobenunabhängige, Vergleiche (siehe z.B. Fischer, 1978, S. 299) zwischen Personen- bzw. Aufgabenparameter.

*Weiterführende Literatur:* Als Weiterführung in deutscher Sprache können Fischer (1968, 1974), Rost & Spada (1978) sowie Wottawa (1980) angeführt werden. In englischer Sprache sei außer Birnbaum (1968) und Lord & Novick (1968) auch das Heft 2 des Bandes 14 (1977) des Journal of Educational Measurement genannt, in dem man neben einer Arbeit von Hambleton & Cook (1977) mehrere ein- und weiterführende Beiträge zum Modell von Rasch und dem von Birnbaum findet. Ein Übersichtsreferat gibt Fischer (1978).

### 3.3.5 Klassisches latent-additives Testmodell

Ein weiteres psychometrisches Testmodell, welches sich als Erweiterung der klassischen Testtheorie versteht, nämlich das „Klassische latent-additive Testmodell“ (KLA-Modell, Moosbrugger & Müller 1981, 1982, Müller 1980) läßt sich als Spezialfall der allgemeinen Theorie latenter Variablen darstellen: die manifesten Variablen  $y_p$  sind intervallskalierte Reaktionsvariablen, welche durch geeignete streng monotone Transformationen aus beobachtbaren kontinuierlichen Testvariablen  $x_p$  gewonnen werden. Der Zusammenhang zwischen den  $y_p$  und der kontinuierlichen latenten Variablen  $\xi$  wird mit einer linearen VC-Funktion gemäß Gleichung 3.2 beschrieben, wobei für alle Aufgaben  $\lambda_p = 1$  gesetzt wird:

$$(3.29) \quad E(y_p | \xi) = \lambda_{p0} + 1 \cdot \xi.$$

Ähnlich wie im Birnbaum-Modell (vgl. Abschnitt 3.3.4) wird anstelle des Leichtigkeitsparameters ein Schwierigkeitsparameter  $\lambda_{p0}^* = -\lambda_{p0}$  benutzt, welcher wegen

$$(3.30) \quad E(\xi | y_p = 0) = -\lambda_{p0} = \lambda_{p0}^*$$

angibt (siehe Abbildung 3.10), in welchem Abstand vom Nullpunkt die VC-Funktion die  $\xi$ -Achse schneidet. Die KLA-Modellgleichung lautet dann in den hier verwendeten Ausdrücken

$$(3.31) \quad E(y_p | \xi) = \xi - \lambda_{p0}^*.$$

Der Personenparameter  $\xi$  und der Aufgabenparameter  $\lambda_{p0}^*$  können wegen Gleichung 3.30 auf einer Skala verglichen werden.

Gilt  $\xi > \lambda_{p0}^*$  ist also die Merkmalsausprägung  $\xi$  einer Person größer als die „Schwierigkeit“ der Aufgabe  $p$ , so wird eine positive Reaktion erwartet und umgekehrt eine negative Reaktion.

Wegen der bei Modellgültigkeit gleichen  $\lambda_p$  für alle  $p \in \{1, 2, \dots, P\}$  Aufgaben und der dann parallelen VC-Funktionen (vgl. Abbildung 3.10) weist auch das KLA-Modell die von Fischer (1974, S. 407) für das Rasch-Modell (siehe Abschnitt 3.3.4) hervorgehobene Modelleigenschaft „spezifische Objektivität“

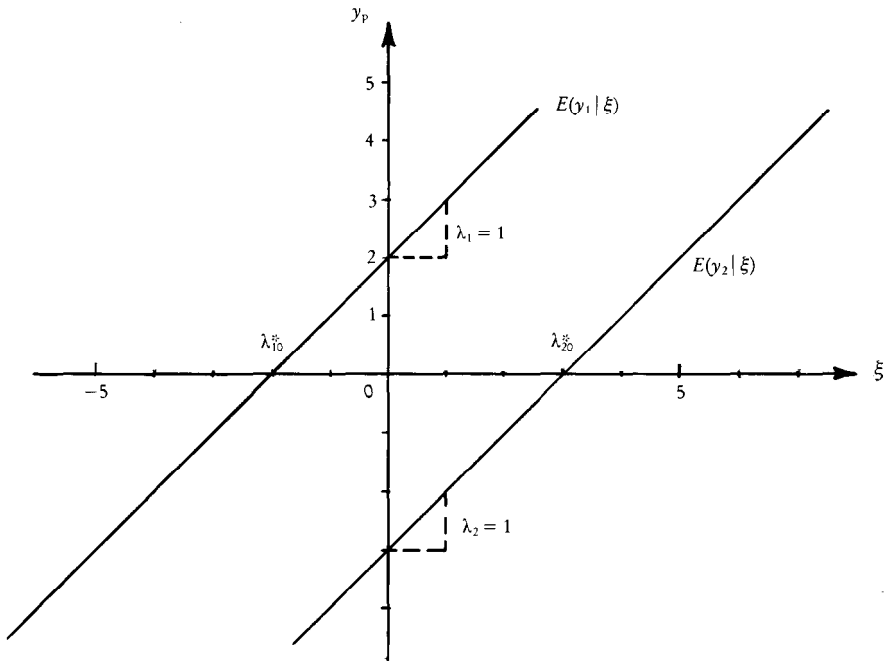


Abb. 3.10: Klassisches latent-additives Testmodell (Moosbrugger & Müller, 1981, 1982) als Spezialfall der allgemeinen Theorie latenter Variablen mit linearer VC-Funktion entsprechend Gleichung 3.2 sowie kontinuierlichen manifesten und kontinuierlicher latenter Variablen. Die Geraden  $E(y_p | \xi)$  gemäß Gleichung 3.31 geben bei psychometrischen Anwendungen für jeden Wert der latenten Variablen  $\xi$  an, welche Reaktion  $y_p$  erwartet wird. In jedem Item werden von jenen Personen positive Reaktionen erwartet, deren latente Merkmalsausprägung größer ist als die entsprechende Itemschwierigkeit, z.B. wenn im 1. Item  $\xi > \lambda_{10}^* = -2$  gilt bzw. im 2. Item  $\xi > \lambda_{20}^* = 3$ . Da die VC-Funktionen aller Aufgaben gleiche Steigungen  $\lambda_p = 1$  aufweisen, sind „spezifisch objektive Vergleiche“ zwischen Personen und Aufgaben möglich.

auf; es erlaubt, Schwierigkeitsunterschiede zwischen Aufgaben mit beliebigen Personen und Merkmalsunterschiede zwischen Personen mit beliebigen Items zu beurteilen.

*Weiterführende Literatur:* Das KLA-Modell wurde von Moosbrugger & Müller (1980) hinsichtlich seiner wesentlichen Eigenschaften mit der klassischen Testtheorie und dem dichotomen logistischen Testmodell (siehe auch Abschnitt 3.3.4) verglichen. Zur Parameterschätzung siehe Müller (1980); über geeignete Computeralgorithmen berichten Müller & Moosbrugger (1980). Mit empirischen Anwendungen befassen sich Moosbrugger (1982), Ruch & Hehl (1982) sowie Tarnai (1981).

### 3.4 Zusammenfassende Bemerkungen

In diesem dritten Abschnitt haben wir eine formale Theorie vorgestellt, die beschreiben kann, wie beobachtbares Verhalten mit psychologischen Begriffen, Eigenschaften oder Konstrukten in Beziehung steht. Erst bei Vorliegen einer Theorie über die konkreten Beziehungen bestimmter psychologischer Konstrukte mit bestimmten beobachtbaren Variablen wird es sinnvoll, von diesen Konstrukten, Eigenschaften oder Begriffen zu sprechen.

Eine mathematisch formulierte Theorie über den Zusammenhang zwischen beobachtbarem Verhalten und den psychologischen Konstrukten hat den Vorteil größerer Präzision und leichterer Überprüfbarkeit.

Die hier behandelte Theorie latenter Variablen basiert auf den Annahmen der bedingten stochastischen Unabhängigkeit und der Art der Variablencharakteristik. Dabei ist es unerheblich, ob die beteiligten Variablen kontinuierlich oder nur zweiwertig sind. Die Einschränkung auf eine einzige latente Variable wurde nur der Einfachheit halber gemacht. Theorien mit mehreren latenten Variablen erfordern jedoch einen größeren mathematischen Aufwand, als wir in diesem Überblick betreiben wollten.

## 4. Anhänge

### 4.1 Anhang A:

*Regeln für Erwartungswerte (vgl. Hays, 1973, S. 871ff.)*

Der *Erwartungswert* einer stochastischen Variablen oder Zufallsvariablen kann als eine mathematisch exakte Formulierung des umgangssprachlichen Begriffes des Durchschnittswertes angesehen werden. Im Fall einer diskreten Zufallsvariablen  $x$  mit den Werten  $x_n$  kann man den Erwartungswert  $E(x)$  folgendermaßen definieren:

$$(A.1) \quad E(x) = \sum_n x_n \cdot P(x = x_n).$$

Der Erwartungswert  $E(x)$  ist demnach die Summe der mit ihren Auftretenswahrscheinlichkeiten gewichteten Werte  $x_n$  der stochastischen Variablen  $x$ . Im Fall, daß die stochastische Variable  $N$  verschiedene Werte hat und jeder dieser Werte mit der Wahrscheinlichkeit  $1/N$  auftritt, erhält die Gleichung A.1 die Form

$$(A.2) \quad E(x) = \sum_{n=1}^N x_n \cdot \frac{1}{N} = \frac{1}{N} \sum_{n=1}^N x_n.$$

Im Fall, daß  $x$  kontinuierlich ist, definiert man  $E(x)$  mit Hilfe des Integrals (siehe Hays, 1973, S. 871 oder Bauer, 1974, S. 138, Steyer 1982, in diesem Band).

Der Erwartungswertvektor eines stochastischen Spaltenvektors  $x = (x_1, \dots, x_Q)'$  ist dann einfach definiert durch

$$(A.3) \quad E(x) := \begin{bmatrix} E(x_1) \\ \vdots \\ E(x_Q) \end{bmatrix}$$

Die wichtigsten Regeln für diese Konzepte sind im folgenden angegeben:

*Regel 1:* Wenn  $\alpha$  ein Vektor konstanter reeller Zahlen ist, dann gilt  
 $E(\alpha) = \alpha$ .

*Regel 2:* Wenn  $\alpha'$  ein Zeilenvektor konstanter reeller Zahlen und  $x$  ein stochastischer Vektor mit dem Erwartungswertvektor  $E(x)$  ist, dann gilt  
 $E(\alpha'x) = \alpha'E(x)$ .

*Regel 3:* Wenn  $\alpha'$  und  $\beta'$  Zeilenvektoren konstanter reeller Zahlen und  $x$  und  $y$  stochastische Vektoren mit den Erwartungswertvektoren  $E(x)$  und  $E(y)$  sind, dann gilt  
 $E(\alpha'x + \beta'y) = \alpha'E(x) + \beta'E(y)$ .

Diese Regeln gelten entsprechend auch für skalare Größen, welche als Vektoren mit einer Komponente aufgefaßt werden können.

## 4.2 Anhang B:

### *Regeln für Varianzen und Kovarianzen*

Die *Kovarianz* zweier stochastischer Variablen oder Zufallsvariablen ist ein Parameter, der die Enge ihres linearen statistischen Zusammenhangs widerspiegelt, d.h. sie gibt an, wie stark die beiden Variablen gemeinsam variieren. Die Kovarianz zweier stochastischer Variablen  $x$  und  $y$ , die wir mit  $C(x,y)$  symbolisieren, ist folgendermaßen definiert:

$$(B.1) \quad C(x,y) := E[[x - E(x)] \cdot [y - E(y)]]$$

Die Verallgemeinerung für mehrere stochastische Variablen ist die *Kovarianzmatrix*. Sind  $x$  bzw.  $y$  stochastische Spaltenvektoren mit  $Q$  bzw.  $P$  Komponenten, so ist deren Kovarianzmatrix  $C(x,y)$  wie folgt definiert:

$$(B.2) \quad C(\mathbf{x}, \mathbf{y}) := E[[\mathbf{x} - E(\mathbf{x})] [\mathbf{y} - E(\mathbf{y})]'] =$$

$$= \begin{bmatrix} C(x_1, y_1) & C(x_1, y_2) & \dots & C(x_1, y_p) & \dots & C(x_1, y_p) \\ C(x_2, y_1) & C(x_2, y_2) & \dots & C(x_2, y_p) & \dots & C(x_2, y_p) \\ \vdots & \vdots & & \vdots & & \vdots \\ C(x_q, y_1) & C(x_q, y_2) & \dots & C(x_q, y_p) & \dots & C(x_q, y_p) \\ \vdots & \vdots & & \vdots & & \vdots \\ C(x_Q, y_1) & C(x_Q, y_2) & \dots & C(x_Q, y_p) & \dots & C(x_Q, y_p) \end{bmatrix}$$

Ein Kovarianzvektor ergibt sich, wenn einer der beiden stochastischen Vektoren nur eine Komponente enthält. Ist z.B.  $Q=1$ , so reduziert sich die Kovarianzmatrix  $C(\mathbf{x}, \mathbf{y})$  auf den Kovarianzvektor

$$(B.3) \quad C(\mathbf{x}, \mathbf{y}) = [C(x, y_1), C(x, y_2), \dots, C(x, y_p), \dots, C(x, y_p)].$$

Ist hingegen  $P = 1$ , so reduziert sie sich auf den Kovarianzvektor

$$(B.4) \quad C(\mathbf{x}, \mathbf{y}) = \begin{bmatrix} C(x_1, y) \\ C(x_2, y) \\ \vdots \\ C(x_q, y) \\ \vdots \\ C(x_Q, y) \end{bmatrix}$$

Man beachte, daß  $C(\mathbf{x}, \mathbf{y}) = C(\mathbf{y}, \mathbf{x})'$  und  $C(\mathbf{y}, \mathbf{x}) = C(\mathbf{x}, \mathbf{y})'$  gelten.

Im folgenden sind einige Regeln für diese Konzepte aufgeführt, die auch für die *Varianz* einer stochastischen Variablen von Bedeutung sind, da diese durch

$$(B.5) \quad V(\mathbf{x}) := C(\mathbf{x}, \mathbf{x})$$

definiert ist. Bei allen stochastischen Variablen setzen wir im folgenden voraus, daß sie einen endlichen Erwartungswert haben.

*Regel 1:* Sind  $\mathbf{a}$  ein Vektor von reellen Konstanten und  $\mathbf{x}$  ein stochastischer Vektor, dann gilt

$$C(\mathbf{a}, \mathbf{x}) = \mathbf{0}.$$

*Regel 2:* Wenn  $A_1, A_2, B_1, B_2$  Matrizen konstanter reeller Zahlen sind und  $x_1, x_2, y_1, y_2$  stochastische Vektoren mit entsprechenden Anzahlen von Komponenten, dann gilt

$$C(A_1 \mathbf{x}_1 + A_2 \mathbf{x}_2, B_1 \mathbf{y}_1 + B_2 \mathbf{y}_2) =$$

$$= A_1 C(\mathbf{x}_1, \mathbf{y}_1) B_1' + A_1 C(\mathbf{x}_1, \mathbf{y}_2) B_2' + A_2 C(\mathbf{x}_2, \mathbf{y}_1) B_1' + A_2 C(\mathbf{x}_2, \mathbf{y}_2) B_2'.$$

Diese Regeln gelten entsprechend auch für skalare Größen, welche als Vektoren mit einer Komponente aufgefaßt werden können.

### 4.3 Anhang C

#### *Regeln für bedingte Erwartungen*

In diesem Anhang sind einige Rechenregeln für bedingte Erwartungen (vgl. Steyer, Moosbrugger & Ullrich 1981) zusammengestellt. Dabei sind  $x$ ,  $y$  und  $z$  stochastische Variablen und  $\mathbf{x}$ ,  $\mathbf{y}$  und  $\mathbf{z}$  stochastische Vektoren auf dem gleichen Wahrscheinlichkeitsraum sowie  $\alpha$ ,  $\beta$  reellwertige Konstanten und  $\boldsymbol{\alpha}$ ,  $\boldsymbol{\beta}$  Vektoren reellwertiger Konstanten.

Die *bedingte Erwartung* von  $y$  unter  $x$ ,

$$E(y | \mathbf{x}) = E(y | x_1, \dots, x_Q)$$

ist eine stochastische Variable auf dem gemeinsamen Wahrscheinlichkeitsraum von  $x$  und  $y$ , deren Werte  $E(y | \mathbf{x}=\mathbf{x})$  die Erwartungswerte von  $y$  unter der Bedingung sind, daß der stochastische Vektor  $x$  einen bestimmten festen Wertevektor  $\mathbf{x}$  angenommen hat (vgl. Feller, 1971<sup>2</sup>, S. 162). Für *bedingte Erwartungswerte*  $E(y, \mathbf{x}=\mathbf{x})$  sind sinngemäß die gleichen Regeln wie für unbedingte Erwartungswerte gültig.

Der Vektor der bedingten Erwartungen ist definiert durch

$$(C.1) \quad E(\mathbf{y} | \mathbf{x}) := [E(y_1 | \mathbf{x}), \dots, E(y_P | \mathbf{x})]'$$

Für *bedingte Erwartungen* gelten folgende *Regeln*, die z.T. der Arbeit von Zimmermann (1975) entnommen sind. Die meisten davon gelten „nur“ mit Wahrscheinlichkeit 1. Auf diese Differenzierung verzichten wir jedoch in diesem mehr an den Praktiker gerichteten Artikel.

*Regel 1:* Der Erwartungswertvektor des Vektors der bedingten Erwartung von  $y$  unter  $x$  ist gleich dem Erwartungswertvektor von  $y$ :

$$E[E(\mathbf{y} | \mathbf{x})] = E(\mathbf{y}).$$

*Regel 2:* Die bedingte Erwartung von  $x_q$ ,  $q \in \{1, \dots, Q\}$ , unter  $\mathbf{x} = (x_1, \dots, x_Q)'$  ist gleich  $x_q$ :

$$E(x_q | x_1, \dots, x_Q) = x_q.$$

*Regel 3:* Die bedingte Erwartung eines reellwertigen Konstantenvektors  $\boldsymbol{\alpha}$  unter  $x$  ist gleich  $\boldsymbol{\alpha}$ :

$$E(\boldsymbol{\alpha} | \mathbf{x}) = \boldsymbol{\alpha}.$$

*Regel 4:* Sind  $\boldsymbol{\alpha}$  und  $\boldsymbol{\beta}$  reellwertige Konstantenvektoren sowie  $x$  und  $y$  stochastische Vektoren, so gilt für die bedingte Erwartung von  $\boldsymbol{\alpha}'\mathbf{x} + \boldsymbol{\beta}'\mathbf{y}$  unter  $z$

$$E(\boldsymbol{\alpha}'\mathbf{x} + \boldsymbol{\beta}'\mathbf{y} | z) = \boldsymbol{\alpha}'E(\mathbf{x} | z) + \boldsymbol{\beta}'E(\mathbf{y} | z).$$

*Regel 5:* Alle Komponenten von  $\mathbf{x}^*$  seien auch in  $\mathbf{x}$  enthalten. Dann gilt: Die bedingte Erwartung unter  $\mathbf{x}^*$  der bedingten Erwartung von  $\mathbf{y}$  unter  $\mathbf{x}$  ist gleich der bedingten Erwartung von  $\mathbf{y}$  unter  $\mathbf{x}^*$   
 $E[E(\mathbf{y} | \mathbf{x}) | \mathbf{x}^*] = E(\mathbf{y} | \mathbf{x}^*)$

*Regel 6:* Die bedingte Erwartung des stochastischen Vektors  $\mathbf{x}_q \cdot \mathbf{y}$  unter  $x_1, \dots, x_Q$  ist gleich dem Produkt von  $x_q$  mit der bedingten Erwartung von  $\mathbf{y}$  unter  $x_1, \dots, x_Q$ :  
 $E[x_q \cdot \mathbf{y} | x_1, \dots, x_Q] = x_q \cdot E(\mathbf{y} | x_1, \dots, x_Q)$ .

*Regel 7:* Der Erwartungswert der bedingten Erwartung der stochastischen Variablen

$\prod_{p=1}^P y_p := y_1 \cdot \dots \cdot y_P$  unter  $\mathbf{x}$  ist gleich dem Erwartungswert von  $\prod_{p=1}^P y_p$ :

$$E \left[ E \left( \prod_{p=1}^P y_p | \mathbf{x} \right) \right] = E \left( \prod_{p=1}^P y_p \right).$$

*Regel 8:* Der Erwartungswert des stochastischen Vektors  $x_q \cdot E(\mathbf{y} | \mathbf{x})$  ist gleich dem Erwartungswertvektor von  $x_q \cdot \mathbf{y}$ :  
 $E[x_q \cdot E(\mathbf{y} | x_1, \dots, x_Q)] = E(x_q \cdot \mathbf{y})$

*Regel 9:* Der Erwartungswert des stochastischen Vektors  $\mathbf{y} - E(\mathbf{y} | \mathbf{x})$  ist gleich dem Nullvektor:  
 $E[\mathbf{y} - E(\mathbf{y} | \mathbf{x})] = \mathbf{0}$ .

*Regel 10:* Der Vektor der bedingten Erwartungen des stochastischen Vektors  $[\mathbf{y} - E(\mathbf{y} | \mathbf{x})]$  unter  $\mathbf{x}$  ist gleich dem Nullvektor:  
 $E[[\mathbf{y} - E(\mathbf{y} | \mathbf{x})] | \mathbf{x}] = \mathbf{0}$ .

*Regel 11:* Der Kovarianzvektor von  $x_q$  und  $\mathbf{y} - E(\mathbf{y} | x_1, \dots, x_Q)$  ist gleich dem Nullvektor  
 $C[x_q, \mathbf{y} - E(\mathbf{y} | x_1, \dots, x_Q)] = \mathbf{0}'$ .

*Regel 12:* Der Erwartungswert des Produkts der bedingten Erwartungen zweier Zufallsvariablen  $z$  und  $y$ , beide unter  $\mathbf{x}$ , ist gleich dem Erwartungswert des Produkts der stochastischen Variablen  $z$  mit der bedingten Erwartung von  $y$  unter  $\mathbf{x}$  und umgekehrt.  
 $E[E(z | \mathbf{x}) \cdot E(y | \mathbf{x})] = E[z \cdot E(y | \mathbf{x})] = E[E(z | \mathbf{x}) \cdot y]$ .

*Regel 13:* Wenn  $\alpha$  eine reellwertige Konstante sowie  $x_1, \dots, x_Q$  und  $y$  stochastische Variablen sind, dann gilt:

$$E(y \mid x_1, \dots, x_Q) = \alpha \Rightarrow C(y, x_q) = 0,$$

für jedes  $q \in \{1, 2, \dots, Q\}$ .

Diese Regeln gelten entsprechend auch für skalare Größen, welche als Vektoren mit einer Komponente aufgefaßt werden können.

## Literatur

- Adrain. 1818. Investigation of the figure of the earth and of the gravity in different latitudes. Transactions of the American Philosophical Society 1. 119-135.
- Anderson, T. W. & Rubin, H. 1956. Statistical inference in factor analysis. Proceedings of the 3<sup>rd</sup> Berkeley Symposium on Mathematical Statistics and Probability, 5, 111-150.
- Anderson, T. W. 1959. Some scaling models and estimation procedures in the latent class model. In: Grenander, U. (Ed.): Probability and statistics, The Harold Cramér Volume. New York: Wiley, 9-38.
- Anderson, T. W. 1978. Classification and discrimination. In: Kruskal, W. H. & Tanur, J. M. (Eds): International Encyclopedia of Statistics. New York: Free Press.
- Bandemer, H. et al. 1977. Theorie und Anwendung der optimalen Versuchsplanung I. Berlin: Akademie-Verlag.
- Bartussek, D. 1970. Eine Methode zur Bestimmung von Moderatoreffekten. Diagnostica, 16, 57-76.
- Bauer, H. 1974. Wahrscheinlichkeitstheorie und Grundzüge der Meßtheorie. Berlin: De Gruyter.
- Bentler, P. M. 1980. Multivariate analysis with latent variables: causal modeling. Annual Review of Psychology, 31, 419-456.
- Birnbaum, A. 1968. Some latent trait models and their use in inferring an examinee's ability. In: Lord & Novick (Eds): Statistical theories of mental test scores. London: Addison-Wesley, 377-479.
- Bishop, Y. M. M. & Fienberg, St. E. & Holland, P. W. 1975. Discrete multivariate analysis: Theory and practice. Cambridge (Mass.): MIT-Press.
- Bock, R. D. 1975. Multivariate statistical methods in behavioral research. New York: McGraw-Hill.
- Bortz, J. 1977. Lehrbuch der Statistik. Berlin: Springer.
- Conger, A. J. & Jackson, D. N. 1972. Suppressor variables, prediction and the interpretation of psychological relationship. Educational and Psychological Measurement, 32, 579-599.

- Conger, A. J. 1974. A revised definition for suppressor variables; a guide to their identification and interpretation. *Educational and Psychological Measurement*, 34, 35-46.
- Cox, D. R. 1970. *The analysis of binary data*. London: Methuen.
- Cramer, E. M. & Nicewander, W. A. 1979. Some symmetric, invariant measures of multivariate association. *Psychometrika*, 44, 43-54.
- Darlington, R. B. 1968. Multiple regression in psychological research and practice. *Psychological Bulletin*, 69, 161-182.
- David, F. N. 1978. Galton, Francis: In: Kruskal, W. H. & Tanur, J. M. (Eds): *International Encyclopedia of Statistics*. New York: Free Press.
- Darper, N. R. & Smith, H. 1966. *Applied regression analysis*. New York: Wiley.
- Edwards, A. L. 1976. *An introduction to linear regression and correlation*. San Francisco: Freeman.
- Feller, W. 1971<sup>2</sup>. *An introduction to probability theory and its applications*. New York: Wiley.
- Finn, J. D. 1974. *A general model for multivariate analysis*. New York: Holt, Rinehart & Winston.
- Fischer, G. H. 1968. *Psychologische Testtheorie*. Bern: Huber.
- Fischer, G. H. 1974. *Einführung in die Theorie psychologischer Tests*. Bern: Huber.
- Fischer, G. H. 1978. Probabilistic test models and their applications. *German Journal of Psychology* 2, 298-319.
- Fisher, R. A. 1925. *Statistical methods for research workers*. London: Oliver & Boyd.
- Fisher, R. A. 1936: The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7, 179-188.
- Formann, A. K. 1978. A note on Parameter estimation for Lazarsfeld's latent class analysis. *Psychometrika* 43, 123-126.
- Formann, A. K. 1980. Neuere Verfahren der Parameterschätzung in der Latent-Class-Analyse. *Zeitschrift für Differentielle und Diagnostische Psychologie* 1, 107-116.
- Gaensslen, H. & Schubö, W. 1973. *Einfache und komplexe statistische Analyse*. München: Reinhardt.
- Galton, F. 1877. Typical laws of heredity. *Royal Institution of Great Britain. Proceedings* 8, 282-301.
- Galton, F. 1889. *Natural inheritance*. London, New York: Macmillan.
- Gibson, W. A. 1959. Three multivariate models: Factor analysis, latent structure analysis and latent profile analysis. *Psychometrika* 24, 229-252.
- Goodman, L. A. 1974. Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika* 61, 215-231.
- Goodman, L. A. 1978. *Analyzing qualitative/categorical data*. London: Addison-Wesley.

- Goodman, L. A. 1979. On the estimation of Parameters in latent structure analysis. *Psychometrika* 44, 123-128.
- Guilford, J. P. 1967. *The structure of human intellect*. New York: Mc Graw Hill.
- Hambleton, R. K. & Cook, L. L. 1977. Latent trait models and their use in the analysis of educational test data. *Journal of Educational Measurement*, 14, 75-96.
- Harnerle, A. 1982. *Latent-Trait-Modelle. Die me theoretische und multivariate statistische Analyse kategorialer Reaktionsmodelle*. Weinheim und Basel: Beltz.
- Harman, H. H. 1967. *Modern factor analysis*. Chicago: The University of Chicago Press.
- Harter, H. L. 1974: The method of least squares and some alternatives - Part I, II. *International Statistical Review* 42, 147-174; 235-264.
- Harter, H. L. 1975. The method of least squares and some alternatives - Part III, IV, V. *International Statistical Review* 43, 1-44; 125-190; 269-278.
- Hays, W. L. 1973. *Statistics for the social sciences*. New York: Holt, Rinehart & Winston.
- Herrmann, Th. 1973. *Pers nlichkeitsmerkmale. Bestimmung und Verwendung in der psychologischen Wissenschaft*. Stuttgart: Kohlhammer.
- Hoerl, A. E. & Kennard, R. W. (1970). Ridge regression: biased estimation for nonorthogonal Problems. *Technometrics*, 12, 55-67.
- Holzkamp, K. 1972. *Kritische Psychologie*. Frankfurt am Main: Fischer.
- Horst, P. 1965. *Factor analysis of data matrices*. New York: Holt, Rinehart & Winston.
- Hotelling, H. 1933. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology* 24, 417-441, 498-520.
- Humak, K. M. S. 1977. *Statistische Methoden der Modellbildung*. Berlin: Akademieverlag.
- J reskog, K. G. 1966. Testing a simple structure hypothesis in factor analysis. *Psychometrika* 31, No. 2, 165-178.
- J reskog, K. G. 1967. Some contributions to maximum likelihood factor analysis. *Psychometrika* 32, 443-482.
- J reskog, K. G. 1969. A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika* 34, No. 2, 183-202.
- J reskog, K. G. 1971. Statistical analysis of sets of congeneric tests. *Psychometrika* 40, 395-412.
- J reskog, K. G. 1973. A general method for estimating a linear structural equation System. In: Goldberger, A. S. & Duncan, O. D. (Eds): *Structural equation models in the social sciences*. New York: Seminar-Press.
- J reskog, K. G. & S rbom, D. 1978. LISREL IV. A general Computer Program for estimation of linear structural equation Systems by maximum likelihood methods. University of Uppsala: Departement of Statistics.

- Kerlinger, F. N. & Pedhazur, E. J. 1973. Multiple regression in behavioral research. New York: Holt, Rinehart & Winston.
- Krauth, J. & Lienert, G. A. 1973. Die Konfigurationsfrequenzanalyse. München: Alber.
- Kreil, C. 1831. Sammlung der notwendigsten math. Formeln aus der Algebra, Trigonometrie, Astronomie und Mechanik. Wien: J. B. Wallishausser.
- Kruskal, W. H. & Tanur, J. M. 1978. International Encyclopedia of Statistics. New York: Free Press.
- Lazarsfeld, P. F. 1950. Logical and mathematical foundations of latent structure analysis. In: Stouffer, S. A. et. al.: Studies in social psychology in world war II, Vol. IV, Princeton N. J. Princeton University Press.
- Lazarsfeld, P. F. 1959. Latent structure analysis. In: Koch, S. (Ed.): Psychologie: A study of a science. Vol. 3, New York: McGraw Hill.
- Lazarsfeld, P. F. & Henry, N. W. 1968: Latent structure analysis. Boston, Houghton Mifflin.
- Lawley, D. N. & Maxwell, A. E. 1971. Factor analysis as a statistical method. London: Butterworths.
- Lienert, G. A. 1969<sup>3</sup>. Testaufbau und Testanalyse. Weinheim: Beltz.
- Littrow, J. J. 1818. Über die gerade Aufsteigung der vornehmsten Fixsterne. *Astronomie* 6, 1-26.
- Littrow, J. J. 1833. Die Wahrscheinlichkeitsrechnung in ihrer Anwendung auf das wissenschaftliche und praktische Leben. Wien: Becks Universitäts-Buchhandlung.
- Lord, F. M. & Novick, M. R. 1968. Statistical theories of mental test scores. Reading, Mass.: Addison & Wesley.
- Mac Corquodale, K. & Meehl, P. E. 1948. On a distinction between hypothetical constructs and intervening variables. *Psychological Review*, 95-107.
- Madansky, A. 1978. Latent structure. In: Kruskal, W. A. & Tanur, J. M. (Eds): International Encyclopedia of Statistics. New York: The Free Press.
- Marinell, G. 1977. Multivariate Verfahren. Eine Einführung für Studierende und Praktiker. 1. Aufl. München: Oldenbourg.
- McDonald, R. P. 1967. Nonlinear factor analysis. *Psychometric Monograph* 15.
- Meehl, P. E. 1950. Configural scoring. *Journal of Consulting Psychology*. 14, 165-171.
- Moosbrugger, H. 1978. Multivariate statistische Analyseverfahren. Stuttgart: Kohlhammer.
- Moosbrugger, H. 1980. Lineare statistische Modelle zur Beschreibung linearer und nichtlinearer multipler Variablenzusammenhänge. Arbeiten aus dem Institut für Psychologie der JWG-Universität Frankfurt am Main 1980/1.
- Moosbrugger, H. 1981a. Zur differentiellen Validität bei nichtlinearen Test-Kriterium-

- Zusammenhängen. Zeitschrift für Differentielle und Diagnostische Psychologie, 2, 219-234.
- Moosbrugger, H. 1981b. Testtheorien und Testkonstruktion. Arbeiten aus dem Institut für Psychologie der Johann-Wolfgang-Goethe-Universität Frankfurt am Main 1981/6.
- Moosbrugger, H. 1982. Dimensionalitätsuntersuchungen von FPI-Skalen mit dem KLA-Modell. Zeitschrift für Differentielle und Diagnostische Psychologie 3, Heft 4 (im Druck).
- Moosbrugger, H. & Müller, H. 1980. Das klassische latent-additive Testmodell (KLA-Modell) und seine Beziehungen zur klassischen Testtheorie sowie zum dichotomen logistischen Testmodell von Rasch. Wissenschaftliches Wandposter, ausgestellt anlässlich der Poster-Session am 32. Kongreß der Deutschen Gesellschaft für Psychologie in Zürich 1980. Arbeiten aus dem Institut für Psychologie der JWG-Universität Frankfurt am Main 1980/8.
- Moosbrugger, H. & Müller, H. 1981. Ein klassisches latent-additives Testmodell (KLA-Modell). In: W. Michaelis (Ed.): Bericht über den 32. Kongreß der Deutschen Gesellschaft für Psychologie in Zürich 1980. 483-486. Göttingen: Hogrefe.
- Moosbrugger, H. & Müller, H. 1982. A classical latent additive test model (CLA Model). The German Journal of Psychology, 6, 145-149.
- Moosbrugger, H. & Steyer, R. 1982. Uni- und multivariate Varianzanalyse mit festen Parametern. In: Bredenkamp, J. & Feger, H. (Eds): Enzyklopädie der Psychologie, Serie: Forschungsmethoden der Psychologie, Band 4. Göttingen: Hogrefe.
- Müller, H. 1980. über „probabilistische“ und „klassische“ Testmodelle mit latent-additiver Struktur: Neue Aspekte der klassischen Testtheorie. Jahresarbeit. Institut für Psychologie der JWG-Universität Frankfurt am Main.
- Müller, H. & Moosbrugger, H. 1980. Computerprogramm für Parameterschätzung, Item- und Personenanalyse nach dem KLA-Modell. Arbeiten aus dem Institut für Psychologie der JWG-Universität Frankfurt am Main 1980/9.
- Mulaik, S. A. 1972. The foundations of factor analysis. New York: McGraw Hill.
- Orth, B. 1974. Einführung in die Theorie des Messens. Stuttgart: Kohlhammer.
- Pawlik, K. 1968. Dimensionen des Verhaltens. Eine Einführung in Methodik und Ergebnisse faktorenanalytischer psychologischer Forschung. Bern: Huber.
- Pearson, K. 1896. Mathematical contributions to the theory of evolution. III. Regression, Heredity and Panmixia. Philosophical Transactions 187 A, 253-318.
- Pearson, K. 1901. On lines and planes of closest fit to a system of points in space. Philosophical Magazine, 2, 6<sup>th</sup> Series, 557-572.
- Rao, C. R. & Mitra, S. K. 1971. Generalized inverse of matrices and its applications. New York: Wiley.
- Rasch, D. 1960. Probabilistic models for some intelligence and attainment tests. Kopenhagen: The Danish Institute for Educational Research.

- Rasch, G. 1976. Einführung in die mathematische Statistik. Berlin: VEB Deutscher Verlag der Wissenschaften.
- Revenstorf, D. 1976. Lehrbuch der Faktorenanalyse. Stuttgart: Kohlhammer,
- Revenstorf, D. 1980. Faktorenanalyse. Stuttgart: Kohlhammer.
- Rachel, H. 1981. Planung und Auswertung von Untersuchungen im Rahmen des allgemeinen linearen Modells. Psychologisches Institut, Ruhr-Universität Bochum.
- Rost, J. & Spada, H. 1978. Probabilistische Testtheorie. In: K. J. Klauer (Ed.): Handbuch der Pädagogischen Diagnostik, Bd. 1, Düsseldorf: Schwann.
- Ruch, W. & Hehl, F.-J. 1982. Möglichkeiten der Kontrolle der Stichprobenabhängigkeit KLA-skaliertter Items bei Vorliegen von Retestdaten. Vortrag auf der 24. Tagung experimentell arbeitender Psychologen in Trier 1982.
- Schermelleh, K. 1979. Vergleichende Darstellung verschiedener Kodierungsverfahren im allgemeinen linearen Modell. Unveröffentlichte Vordiplomarbeit. Institut für Psychologie der JWG-Universität Frankfurt am Main.
- Searle, S. R. 1971. Linear models. New York: Wiley.
- Sörbom, D. & Jöreskog, K. G. 1976. COFAMM: Confirmatory factor analysis with model modifikation. User's Guide. Chicago: National Educational Resources.
- Späth, H. 1973. Algorithmen für elementare Ausgleichsmodelle. München: Oldenbourg.
- Späth, H. 1974. Algorithmen für multivariate Ausgleichsmodelle. München: Oldenbourg.
- Spearman, C. E. 1904. General intelligence objectively determined and measured. *American Journal of Psychology*, 15, 201-293.
- Spearman, C. E. 1926. The abilities of man. Their nature and measurement. London: Macmillan.
- Steyer, R. 1979. Untersuchungen zur nonorthogonalen Varianzanalyse. Weinheim: Beltz.
- Steyer, R. 1982. Modelle zur kausalen Erklärung statistischer Zusammenhänge. In: Bredenkamp, J. & Feger, H. (Eds): Enzyklopädie der Psychologie, Serie: Forschungsmethoden der Psychologie, Band 4. Göttingen: Hogrefe.
- Steyer, R. & Moosbrugger, H. 1979. Lineare Modelle: Modelle kongenerischer Variablen. Arbeiten aus dem Institut für Psychologie der JWG-Universität Frankfurt am Main 1979/2.
- Steyer, R., Moosbrugger, H. & Ullrich, K. 1981. Rechenregeln für bedingte Erwartungen. Arbeiten aus dem Institut für Psychologie der JWG-Universität Frankfurt am Main 1981/1.
- Tarnai, Ch. 1981. Anwendbarkeit des KLA-Modells bei nichtlinearen Funktionen zwischen latenten und manifesten Reaktionsvariablen. Vortrag auf der 23. Tagung experimentell arbeitender Psychologen in Berlin 1981.

- Tatsuoka, M. M. 1971. Multivariate analysis. Techniques for educational and psychological research. New York: Wiley.
- Thurstone, L. L. 1938. Primary mental abilities. Chicago: The University of Chicago Press.
- Thurstone, L. L. 1947. Multiple factor analysis. Chicago: University of Chicago Press.
- Timm, N. H. 1975. Multivariate analysis with applications in education and psychology. Belmont, Cal.: Wadsworth.
- Überla, K. 1971. Faktorenanalyse. Heidelberg: Springer.
- Weiling, F. 1972. Geschichtliche Notizen zur Entwicklung der Regressionsanalyse. Biometrische Zeitschrift, 14, 164-177.
- Weiling, F. 1973. Die Varianzanalyse. Eine Übersicht mit historischem Aspekt. Vortrag, 19. Kolloquium der Deutschen Region der Internationalen Biometrischen Gesellschaft, Berlin.
- Winer, B. J. 1971<sup>2</sup>. Statistical principles in experimental design. New York: McGraw Hill.
- Wottawa, H. 1974. Das allgemeine lineare Modell. Ein universelles Auswertungsverfahren. EDV in Medizin und Biologie, 3, 65-73.
- Wottawa, H. 1979. Grundlagen und Probleme von Dimensionen in der Psychologie. Meisenheim am Glan: Hain.
- Wottawa, H. 1980. Grundriß der Testtheorie. München: Juventa.
- Yule, G. U. 1897. On the significance of Bravais' formulae for regression etc in the case of skew correlation. Proceedings of the Royal Society London, 60, 477-489.
- Zimmerman, D. W. 1975. Probability spaces, hilbert spaces, and the axioms of test theory. Psychologika, 40, 395-412.