Die Hoeffding-Zerlegung in Linearen Modellen

Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultäten
der Georg-August-Universität zu Göttingen

vorgelegt von
Arne Bathke
aus Hamburg-Harburg

Göttingen 2000

D7
Referent: Professor Dr. Manfred Denker
Korreferent: Professor Dr. Edgar Brunner
Tag der mündlichen Prüfung: 02.11.2000

Inhaltsverzeichnis

Einleitung

Hauptgegenstand dieser Arbeit ist eine vereinheitlichende statistische Modellbildung unter Verwendung der Hoeffding-Zerlegung.
Führt man verschiedene Modelle in der Statistik, wie etwa Lineare Modelle, verallgemeinerte Lineare Modelle und Nichtparametrische Modelle, auf ihre mathematisch wesentlichen Bestandteile zurück, so entdeckt man starke Gemeinsamkeiten, die einen vereinheitlichenden Zugang zu diesen Modellen nahelegen. Dazu ist es allerdings zunächst notwendig, für den zentralen Begriff des Faktors ein klares Konzept zu entwickeln, das zum einen mathematisch solide ist, andererseits aber auch in Einklang mit der intuitiven Vorstellung des Statistikers im Hinblick auf Anwendungen steht.
In der gängigen statistischen Literatur über Lineare Modelle ist „Faktor" ein viel verwendeter Standardbegriff. Allerdings finden sich längst nicht in allen Büchern Definitionen oder Umschreibungen dieses Begriffs. Und die Umschreibungen, auf die man stößt, offenbaren, dass es keine allgemeingültige oder allgemein akzeptierte Definition gibt. Die der Literatur entnommenen Definitionen für den Begriff lassen sich grob in drei Gruppen zusammenfassen, die sich darin unterscheiden, inwieweit ein Zusammenhang zwischen Faktoren und Zielgröße angesprochen und formalisiert wird.

Wir führen das Konzept einer dieser drei Gruppen systematisch durch und definieren Faktoren als stochastische Argumente einer Funktion, deren Funktionswert die Zielgröße oder mit ihr zusammenhängende Größen in einem statistischen Modell darstellt und die als Argument noch eine weitere Zufallsvariable besitzt, nämlich den Versuchsfehler. Derartige Funktionen können mit Hilfe der Hoeffding-Zerlegung additiv zerlegt werden. Anhand der Komponenten der Zerlegung lassen sich unmittelbar Hypothesen formulieren und ein Vergleich der Hypothesen verschiedener Modelle, die dadurch charakterisiert sind, ob die Funktionswerte z.B. beobachtbare Zielgrößen oder Verteilungsfunktionen sind, ist möglich.
Von einigen Autoren (z.B. TJUR (1984, 1991), BAILEY (1984), HELLAND (1998)) werden Faktoren als Partitionen des Raumes der Beobachtungseinheiten angesehen und dies wird auf unterschiedliche Arten formalisiert. Obwohl ein solches Vorgehen zunächst kaum etwas mit unserem Ansatz gemeinsam hat, finden sich doch bei der Durchführung beider Konzepte Entsprechungen. Die bei uns auftretende Unkorreliertheit von Zufallsvariablen als Komponenten der Zerlegung korrespondiert z.B. mit der Orthogonalität bestimmter Räume bei HELLAND (1998). Allerdings denken wir, dass der von uns verwendete Formalismus leichter zu handhaben ist und dass ein Verständnis von Faktoren als Einflussgrößen auf die Zielvariable eher der intuitiven Vorstellung entspricht. Außerdem lassen sich so ohne zusätzliche aufwendige Konstruktionen feste und zufällige Faktoren sowie Kovariablen in ein gemeinsames Konzept bringen.

Das bei uns durchgehend angewendete Prinzip der additiven Zerlegung, das für die Gültigkeit des Linearen Modells im Wesentlichen nur noch um die Forderung nach fehlender Wechselwirkung zwischen Faktoren und Fehlerterm ergänzt werden muss, lässt sich, gemeinsam mit dem von uns verfolgten Konzept für den Begriff „Faktor“, nicht nur für die klassischen Linearen Modelle anwenden. AKRITAS & ARNOLD (1994) haben durch die Formulierung nichtparametrischer Hypothesen mit Hilfe von additiv zerlegten Verteilungsfunktionen der Entwicklung der Nichtparametrischen Statistik neue Anstöße gegeben, auch wenn die so formulierten Hypothesen zunächst schwierig zu interpretieren sind. Die Zerlegung von Verteilungsfunktionen lässt sich ebenso in das in dieser Arbeit vorgestellte allgemeine Konzept einbetten, wie etwa die Zerlegung von transformierten Zielgrößen, Parametern oder Verteilungsquantilen.

Im Hinblick auf die Herleitung gewisser Teststatistiken bietet eine additive Zerlegung sogar technische Vorteile bzw. eine größere Transparenz, da sie mit der Aufspaltung zufälliger quadratischer Formen in separat betrachtbare Summanden einhergeht. Wir haben auf diese Art und Weise für die Varianzanalyse bei nicht normalverteilten Fehlern, einer großen Anzahl an Faktorstufen und einer kleinen Zahl an Versuchswiederholungen pro Faktorstufe ein asymptotisches Resultat hergeleitet, dass eine wesentliche Verallgemeinerung eines Ergebnisses von BOOS & BROWNIE (1995) darstellt, indem es Designs mit beliebig vielen Faktoren behandelt.
Die ergänzenden Resultate zu der Arbeit von BOOS & BROWNIE, die AKRITAS & ARNOLD (2000) kürzlich vorgestellt haben und die im Wesentlichen die Verteilungen der Teststatistik unter Alternative betreffen, haben die Autoren ebenfalls über das Hilfsmittel einer additiven Zerlegung der quadratischen Formen erzielt, was die Vorteilhaftigkeit einer Betrachtung dieser Zerlegung untermauert.
Im Rahmen dieser Arbeit ist die Zerlegung jedoch nicht bloßes technisches Hilfmittel, sondern generelles Konzept zur Formulierung und Formalisierung statistischer Modelle. Es ist natürlich kein unerwünschter Nebeneffekt, wenn sich dadurch an manchen Stellen Rechenvorteile ergeben.

Im Einzelnen ist die Arbeit wie folgt gegliedert:
In Kapitel 1 geben wir zunächst einen Überblick über die Verwendung des Faktorbegriffs in der statistischen Literatur. Dabei gruppieren wir die gefundenen Zitate, wie oben erwähnt, in drei verschiedene Klassen.
Im Sinne der Definitionen einer dieser Gruppen formulieren wir zu Beginn des 2. Kapitels verschiedene Modelle und führen anschließend die Hoeffding-Zerlegung für diese Modelle durch. Ein Beweis des Satzes über die Hoeffding-Zerlegung wird auch angegeben.
Im darauffolgenden Abschnitt beschäftigen wir uns mit der Problematik der Unterscheidung zwischen verschiedenen Typen von Faktoren (insbesondere fest und zufällig) und formulieren einige Annahmen, die die verschiedenen Modelle (Lineares Modell, verallgemeinertes Lineares Modell, Nichtparametrisches Modell) voneinander abgrenzen.
Da wir bei der Formulierung der Modelle auch transformierte Zufallsvariablen zulassen, ist ein kurzer Überblick über Transformationen von Zielgrößen eingefügt (Abschnitt 2.4).
Der nächste Abschnitt ist den Beziehungen zwischen den Hypothesen verschiedener Modelle gewidmet. Die Aussagen, die wir dort über Implikationen zwischen Hypothesen in allgemeiner Form beweisen bzw. durch Gegenbeispiele widerlegen, sind kohärent mit von BRUNNER & PURI (1996), OELERICH (1998) und SCHWARZ (1999) für die von ihnen betrachteten Modelle erzielten Resultaten. Dabei erlaubt uns der in dieser Arbeit eingeführte vereinheitlichende Zugang zu statistischen Modellen eine größtenteils viel allgemeinere Formulierung der Bezie-

hungen zwischen den Hypothesen.
Kapitel 2 schließt mit einer kurzen Auflistung einiger möglicher konkreter Modelle, die sich in das hier vorgestellte Konzept der Modellbildung integrieren lassen.
Im 3. Kapitel wenden wir uns der Modellierung statistischer Versuchsdesigns zu. Hier werden wiederholte Realisierungen eines statistischen Modells betrachtet sowie die Kovarianzstruktur der Zielgrößen unter bestimmten Bedingungen angegeben.
Eine Umsetzung der theoretischen Formulierungen aus dem 2. und 3. Kapitel erfolgt in Kapitel 4. Anhand von Beispielen demonstrieren wir die Anwendbarkeit der statistischen Theorie für verschiedene Modelle (Lineares Modell, Lineares Modell ohne Normalverteilungsannahme, Nichtparametrisches Modell). Dabei stellen wir unseren jeweils aus demselben einheitlichen Konzept entwickelten konkreten Modellgleichungen klassische Formulierungen der Literatur gegenüber. An einigen Stellen führen wir die Herleitung von Teststatistiken weiter aus, um die Effekte der sich aus der Hoeffding-Zerlegung ergebenden Zerlegung quadratischer Formen zu illustrieren.
In Kapitel 5 wird schließlich das spezielle Design des Linearen Modells ohne Normalverteilungsannahme betrachtet und es werden asymptotische Resultate für den Fall l (= Anzahl der Faktorstufen) $\to \infty$, n (= Anzahl der Versuchswiederholungen pro Faktorstufe) fest, hergeleitet. Dabei kann das Modell beliebig viele Faktoren enthalten. Wir gehen auch auf den in den bisherigen Arbeiten zu diesem Thema nicht betrachteten Fall der Heteroskedastizität ein; ein weiterer Abschnitt ist Rangstatistiken gewidmet. Das finite Verhalten der untersuchten Teststatistik wird abschließend anhand von Computersimulationen analysiert.
Einige technische Resultate haben wir in den Anhang ausgegliedert. Auf S.118 sind einige weniger geläufige Symbole, die in dieser Arbeit zur Verwendung kommen, aufgelistet.

Abschließend möchte ich mich bei Herrn Prof. M. Denker für die bemerkenswerte Unterstützung bei dieser Arbeit bedanken. Herrn Prof. E. Brunner danke ich für die Bereitschaft, das Korreferat zu übernehmen. Ein weiterer Dank geht an Prof. S. Portnoy, der mir während meines Forschungsaufenthaltes an der University of Illinois at Urbana-Champaign wertvolle Anregungen gegeben hat.

Kapitel 1

Der Faktorbegriff in der statistischen Literatur

In diesem Abschnitt stellen wir Literaturzitate, die den Begriff „Faktor" in statistischen Modellen umschreiben, zusammen. Schaut man in die statistische Literatur, so hat man im Allgemeinen zunächst Schwierigkeiten, exakte Definitionen für diesen Begriff zu finden – auch wenn in vielen weitergehenden Büchern über Statistik z.B. „mehrfaktorielle Modelle" behandelt werden.
In einigen Büchern über Lineare Modelle wird der Begriff „Faktor" überhaupt nicht erläutert. Dazu sei hier lediglich das Standardwerk von ARNOLD (1981) erwähnt.
Im Weiteren betrachten wir Werke, in denen eine Definition oder Umschreibung des Begriffes vorgenommen wird. Sie lassen sich grob in drei Gruppen einteilen, wenn als Kriterium formuliert wird, inwieweit die gefundenen Definitionen den Zusammenhang zwischen Faktoren und Zielgröße herausstellen.
Die erste Gruppe bilden Definitionen, bei denen die Zuordnung der Faktorstufen zu den experimentellen Einheiten im Vordergrund steht. In der zweiten und dritten Gruppe wird eine Beziehung zwischen Faktoren und Zielgröße angesprochen. Dabei wird dies in den Umschreibungen aus Gruppe 2 nicht mathematisch formuliert; die Definitionen aus Gruppe 3 stellen dagegen einen funktionalen Zusammenhang zwischen Faktoren und Zielgröße deutlich heraus.
Obwohl die Konzepte, wie ein Faktor aufzufassen ist, in der ersten und dritten Gruppe der von uns vorgenommenen Einteilung grundlegend verschieden sind, ergeben sich doch verblüffende Parallelen, die deutlich werden, wenn man die Resultate von HELLAND (1998) (Gruppe 1) mit unseren Ergebnissen (Gruppe 3), insbesondere aus Abschnitt 2.2, vergleicht. HELLAND erhält z.B. über den Zugang der Partition des Stichprobenraumes und über orthogonale Projektionen ein äquivalentes Resultat zu der von uns formulierten Aussage (2.1) auf S.25. Dabei werden in der vorliegenden Arbeit Faktoren als Zufallsvariablen aufgefasst, die als Argumente in eine Funktion eingehen. Diese Funktion wird mittels der Hoeffding-Zerlegung linear zerlegt.
Im Folgenden werden die zunächst unterschiedlichen Herangehensweisen an den Faktorbegriff, die in der statistischen Literatur existieren, anhand von Originalzitaten aufgezeigt. Die Zitate sind innerhalb der jeweiligen Gruppen im Wesentlichen chronologisch geordnet.

1. Umschreibungen, die nicht direkt auf den Zusammenhang zwischen Faktor und Zielgröße eingehen, sondern bei denen die Zuordnung der Faktorstufen zu den experimentellen Ein-

heiten im Vordergrund steht. Dabei geben wir zunächst solche Zitate wieder, die keine formale mathematische Sprache verwenden:

> „We shall call each basic treatment a *factor* and the number of possible forms of a factor the number of *levels* for that factor. A particular combination of one level from each factor determines a *treatment*. The experiment as a whole is called a *factorial experiment* if all, or nearly all, factor combinations are of interest."
> COX (1958, S.91f)

> „The term *factor* will be used interchangeably with the terms treatment and experimental variable. More specifically, a factor is a series of related treatments or related classifications. The related treatments making up a factor constitute the *levels* of that factor. [...] Basically a factor is a qualitative variable; in special cases it becomes a quantitative variable."
> WINER (1971, S.311)

> „Die Faktoren haben bei der Varianzanalyse qualitativen Charakter. Zu einem Faktor gehört eine Anzahl verschiedener möglicher Zustände, die [...] als Behandlungen, Klassen oder Stufen bezeichnet werden. Es wird angenommen, daß jedes Individuum bezüglich jedes vorkommenden Faktors genau einer Behandlung bzw. Klasse zugeordnet werden kann."
> AHRENS&LÄUTER (1974, S.1)

> „If the treatments in an experiment consist of all possible combinations of a set of underlying factors, it is called a *factorial experiment*. [...] Each defining variable of the classification [...] is called a *factor* in the experiments. Each value that a factor can take, which defines an individual cell of the marginal classification by that factor, is called a *level* of that factor."
> KENDALL&STUART (1976, S.158f)

> „*Factor (or independent variable)*: A set of treatments or some other effect that is to be evaluated by the research work; or, groupings of experimental units into subsets of the universe for comparison."
> MARKS (1982, S.206)

> „Suppose that two or more different kinds of treatments are of interest. Each kind of treatment is called a factor. Each factor is of interest at some number of different levels. [...] In a factorial design the treatments are taken to be all possible combinations of the levels of the different factors."
> CHRISTENSEN (1987, S.156)

> „The data we examine with analysis of variance involve a *response* variable and structural components called *factors*, which identify the observations under study as collected from specific subgroups or under particular circumstances."
> SINGER (1991, S.50)

TJUR (1984,1991) hat wohl als erster die Zuordnung der Faktorstufen zu den experimentellen Einheiten mathematisch formalisiert. Leider wird bislang von Autoren, die *Faktor* im Sinne „Zuordnung der Faktorstufen zu den experimentellen Einheiten" – also wie TJUR – verstehen, seine exakte Definition zu selten verwendet, obwohl dies dann zu mehr mathematischer Klarheit für den Begriff „Faktor" führen könnte:

> „The finite set I indexing the observations is referred to as the set of *experimental units*. A *factor* is formally defined as a mapping $\phi_F : I \rightarrow F$ from I to another finite set F of *factor levels* or *labels*. [...] The representation of a factor as an asignment of levels to experimental units corresponds to the way factors are usually represented in a computer program."
> TJUR (1991, S.274)

Von dieser Formalisierung hat sich HELLAND (1998) inspirieren lassen. Im Unterschied zu TJUR betrachtet er allerdings Faktoren als Partitionen des Stichprobenraumes und nicht als Funktionen:

> „Assume an infinite population Ω. For each unit in the population we may make observations y, which then will be random variables upon a probability space $(\Omega, P, \mathcal{E})$. In its simplest form, a *factor* can be thought of as a finite measurable partition of this space: $\mathcal{F} = \{\Omega_f^{\mathcal{F}} : f = 1, 2, \ldots, |\mathcal{F}|\}$, or, equivalently, the (σ-)algebra generated by this partition. [...] Thus in general we define a factor as any σ-algebra $\mathcal{F} \subseteq \mathcal{E}$."
> HELLAND (1998, S.4)

Die Interpretation eines Faktors als Partition einer Menge von experimentellen Einheiten wurde schon von WHITE (1975) in Worte gefasst, allerdings nicht mathematisch formalisiert. Außerdem versteht WHITE einen Faktor nur als „natürliche" Zerlegung, wodurch seine Darstellung nicht klarer wird:

> „A factor is understood to be a natural classification of the set of experimental units and not an arbitrary or random partition, such as 'treatments' or 'attempts at a treatment' induced by an experiment."
> WHITE (1975, S.560)

BAILEY (1984) hat schließlich aus Anlass von TJURS Auseinandersetzung mit dem Thema verschiedene äquivalente Formalisierungen zusammengestellt, mit denen Faktoren im obigen Sinne beschrieben werden können. Die von TJUR und HELLAND angegebenen Definitionen lassen sich dort auch einordnen:

> „(1) a *function* ϕ from I to a set F; (2) a *partition* of I into the subsets $\phi^{-1}(f)$ for $f \in F$; (3) an *equivalence relation* $\sim$ on I, where $i \sim j$ if and only if $\phi(i) = \phi(j)$; (4) a *subset* S of $I \times I$, where $(i, j) \in S$ if and only if $i \sim j$; (5) a *matrix* $\mathbf{R}$ on $I \times I$, where $R_{ij} = 1$ if $i \sim j$ and $R_{ij} = 0$ otherwise."
> BAILEY (1984, S.66)

2. Umschreibungen, die die Beziehung zwischen Faktor und Zielgröße ansprechen, ohne dies konkreter zu formalisieren. Dabei gibt es offenbar auch unterschiedliche Ansichten darüber, ob Faktoren kontrollierbare oder nicht kontrollierbare Größen sind (vgl. dazu die Zitate aus HUITSON (1971) und RETZLAFF ET AL. (1978)):

   > „A common problem in research is investigating the effects of each of a number of variables, or *factors* as they are called, on some response Y."
   > SNEDECOR&COCHRAN (1967, S.339)

   > „Versuche werden geplant, um die Wirkung bestimmter Einflußgrößen auf die Zielgröße abzuschätzen. Diese Einflußgrößen seien als *Faktoren* bezeichnet. Wir verstehen hierunter einmal *verschiedene Verfahren*, sodann auch *Stufenwerte einer Einflußgröße*, z.B. der Temperatur."
   > SACHS (1968, S.570)

   > „In any experiment, there are always a large number of external conditions over which the experimenter has either no control or which are too difficult to control. [...] Some of these uncontrolled conditions will affect the results of the experiment. Such external conditions are usually called *factors*."
   > HUITSON (1971, S.1)

   > „In many experiments the main objective is to determine the effect of one or more external conditions on some response variable, X, of basic or primary interest. [...] The external conditions, according to which data are classified, are called *factors*. The number of different classes or categories of a factor to be investigated is called the number of *levels* of the factor."
   > CHOI (1978, S.191)

   > „Als einen Faktor bezeichnet man eine *unabhängige, willkürlich einstellbare Größe*, die vermutlich einen Einfluß auf das Ergebnis eines Versuchs hat."
   > RETZLAFF ET AL. (1978, S.7)

   > „Das Ziel der Varianzanalyse besteht darin, die Wirkung gewisser Einflußgrößen, genannt Faktoren, auf das Untersuchungsmaterial zu studieren. Die Wirkung, genannt Effekt, eines einzelnen Faktors oder einer Kombination von Faktoren wird für gewisse Stufen der Faktoren untersucht."
   > HUMAK (1984, S.250)

   > „In experimental design, the *dependent variable* is called the *response variable*, and the *independent variables* are called *factors*. [...] The different *levels* of a factor (or combination of factors) are called *treatments*."
   > BOWERMAN&O'CONNELL (1990, S.730)

   > „Man unterscheidet Variablen, die beeinflußt werden, die sogenannten *Zielgrößen*, und solche, die beeinflussen. Diese werden weiter aufgeteilt in beobachtbare Variablen, also Variablen, die gemessen werden können und die als *Einflußgrößen* oder *Faktoren* bezeichnet werden, und in nicht beobachtbare

Variablen, die zum Teil als *Störgrößen* gelten, aber auch wichtige Faktoren sein können und dann als *latent* bezeichnet werden.“
FAHRMEIR ET AL. (1999, S.11)

In den beiden folgenden Zitaten werden Faktoren explizit als nicht stochastische Größen definiert:

„*Factors* [...] are what can be essentially controlled and, with some justification, assumed to be nonstochastic, and a *response* is an outcome in some sense. [...] The relationship between the factors and the response [...] can be regarded, from a certain viewpoint, as a generalization of the classical cause and effect relationship. In the general setup, the responses are supposed to have a multivariate probability distribution whose nature or characterization is supposed to depend on the factors [...].“
GNANADESIKAN&SRIVASTAVA (1971, S.1)

„Eine sehr häufige und wichtige Aufgabe für die experimentelle Forschung ist die Untersuchung der Abhängigkeit eines zufallsbeeinflußten Resultats, der sogenannten Zielgröße, von einer Anzahl kontrollierbarer, nicht zufälliger Faktoren, deren Werte man wählen und z.B. einstellen kann. Man nennt diese Faktoren Einflußgrößen oder Einflußfaktoren.“
BANDEMER ET AL. (1973, S.9)

3. Umschreibungen, die die Zielgröße als eine Funktion der Faktoren verstehen lassen. Einen Sonderfall bildet dabei die unmittelbar nachfolgende Definition von SEARLE (1971), bei der Faktoren in einem etwas engeren Sinne betrachtet werden:

„The word *factor* denotes what has hertofore been called a variable. [...] This use of factor in place of variable emphasizes that what is being called a factor cannot be measured precisely by cardinal values: the word *variable* is reserved for that which can be so measured.“
SEARLE (1971, S.140)

SEARLE versteht unter *Faktoren* lediglich nominale oder ordinale Größen, also solche, die nicht „exakt gemessen“ und höchstens angeordnet werden können. Andere Einflußgrößen werden bei ihm *Variablen* genannt. Als Beispiele für Faktoren in diesem Sinne führt er (SEARLE (1971, S.137 ff)) Beruf, Schulabschluß, Religion und Nationalität (in den Sozialwissenschaften), Pflanzenart, Dünger und Bodentyp (in den Agrarwissenschaften), sowie Rohstoff, Bearbeitungsmethode und Fabrikanlage (in industriellen Prozessen) an. Desweiteren bezeichnet SEARLE Größen, die mehr oder weniger willkürlich klassiert wurden, als Faktoren (z.B. jährliches Einkommen).

Die nun folgenden Definitionen unterscheiden, im Gegensatz zu SEARLE, nicht zwischen *Faktor* und *Variable*.
In einem Teil von ihnen wird explizit erwähnt, dass die Zielgröße eine Funktion nicht nur der Faktoren, sondern auch eventueller weiterer Größen (Versuchsfehler) ist (vgl. SCHACH & SCHÄFER (1978) und ATKINSON & DONEV (1992)).

„<The *independent variable* is> the variable that is under the control of the experimenter. The terms independent variable and treatment will be used interchangeably. [...] The *dependent variable* reflects any effects associated with manipulation of the independent variable. [...] The treatment levels should be chosen so as to cover a sufficiently wide range to detect effects of the independent variable if real effects exist. In addition, the number and spacing of the levels should be sufficient to define the shape of the function relating the independent and dependent variables.“
KIRK (1968, S.4ff)

„A single equation or set of equivalent equations is desired for each dependent variable that relates its values to the corresponding conditions or *levels* of the independent variables. Statisticians usually call dependent variables *responses* and independent variables *factors*.“
DANIEL&WOOD (1971, S.5)

„Bei quantitativen wissenschaftlichen Untersuchungen hat man häufig die Vorstellung, daß gewisse Faktoren, welche die Werte $x_1, x_2, \dots, x_k$ annehmen mögen, ein Ergebnis y beeinflussen. Kann man die Einflußfaktoren systematisch variieren, und wird das Ergebnis von diesen Faktoren eindeutig bestimmt, dann ist es im Prinzip möglich, die Abhängigkeit des y-Werts von $x_1, x_2, \dots, x_k$, d.h. die Funktion $y = f(x_1, x_2, \dots, x_k)$, beliebig genau zu ermitteln. [...] Bei der Durchführung von Experimenten findet man aber meistens, daß y außer von $x_1, x_2, \dots, x_k$ auch von gewissen weiteren Einflüssen abhängt, z.B. von einem Meßfehler, von gewissen nicht beobachteten oder nicht beobachtbaren Werten $x_{k+1}, x_{k+2}, \dots$ weiterer Faktoren, von ‚zufälligen ‘Eigenschaften der Untersuchungseinheit, etc.“
SCHACH&SCHÄFER (1978, S.1)

„Much experimentation deals with the problem of obtaining information about the functional relationship between variables. [...] The experimenter may postulate a relationship of the following form: $\eta = f(z, \theta)$ [...]. We usually term η the *response variable* and think of it as the response to the variables z_j which are referred to as the *factors* of the experiment. The components of θ are unknown constants linking the effects of the z_j to η and are called parameters.“
GUTTMANN (1982, S.128)

GUTTMANN formuliert zwar auch einen funktionalen Zusammenhang. Allerdings ist dies bei ihm eher auf regressionsanalytische Modelle zu beziehen. Die einfaktorielle Varianzanalyse wird bei ihm - wie auch bei SEARLE (1971) - als Spezialfall einer Regression mit Hilfe von Dummy-Variablen hergeleitet.

„[...] We attempt to represent the variables $y_1, y_2, \dots, y_p$ as linear combinations of a few random variables $f_1, f_2, \dots, f_m$ $(m < p)$ called *factors*.“
RENCHER (1995, S.445)

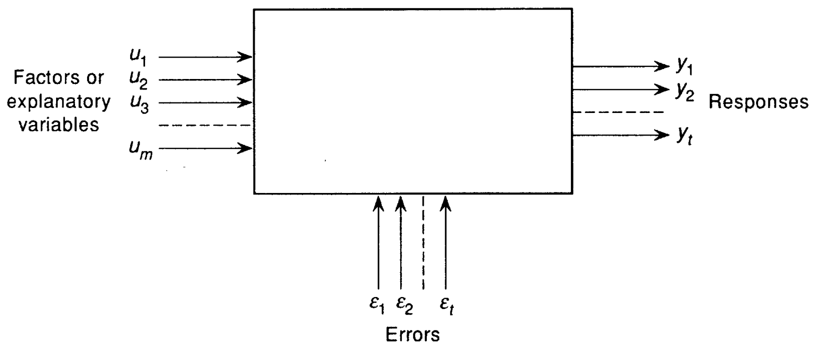

Abbildung 1: Aus ATKINSON&DONEV (1992, S.11). Die Bildunterschrift dort lautet: „Schematic representation of an experiment. The relationship between the factors u and the response y is obscured by the presence of error ε. The values of u are to be chosen by the experimenter who observes y but not ε."

> „A single trial consists of measuring the values of the t response, or output, variables $y_1, \dots, y_t$. These values are believed to depend upon the values of the m factors or explanatory variables $u_1, \dots, u_m$. However, the relationship is obscured by the presence of unobservable random errors $\varepsilon_1, \dots, \varepsilon_t$."
> ATKINSON&DONEV (1992, S.11)

Die Gegenüberstellung der Literaturzitate macht deutlich, dass es keine einheitliche Auffassung dazu gibt, wie der Begriff „Faktor" mathematisch zu verstehen ist. Dieser etwas unbefriedigende Tatbestand bei einem für die Statistik so fundamentalen Begriff ist zunächst erstaunlich und hat sicher auch dazu geführt, dass BAILEY (1984) in seiner Diskussion zu einer Veröffentlichung von TJUR (1984) die oben zitierte Zusammenfassung (s.S.11) geschrieben hat. Allerdings bieten die von ihm genannten Definitionen nur dann eine akzeptable Formalisierung für den Begriff „Faktor", wenn man ihn als Art und Weise der Gruppierung der experimentellen Einheiten auffassen will. Einen Faktor stattdessen als Argument einer Funktion anzusehen, deren Ergebnis die Zielvariable ist, ist direkt und intuitiv plausibel, scheint weniger gekünstelt zu sein – und flexibler für die Modellbildung.
Trotz dieser beiden grundsätzlich verschiedenen Herangehensweisen, die in den Zitaten aus Gruppe 1 und Gruppe 3 deutlich werden (die Umschreibungen aus der zweiten Gruppe lassen sich auch als eine Art Vorstufe zu den Definitionen aus der dritten Gruppe ansehen), ergeben sich interessanterweise bei der Weiterführung der jeweilige Konzepte die oben erwähnten Parallelen (s. S.9).
Der Versuch, die in der Statistik beliebten Linearen Modelle auf ihre mathematisch relevanten Teile zurückzuführen, kann also, selbst bei ganz unterschiedlichen Formalisierungen, zu denselben wesentlichen Resultaten führen. Dies lässt vermuten, dass die große Uneinigkeit, die die statistische Literatur auf den ersten Blick zum Thema *Faktor* offenbart, gar nicht so sehr auf grundsätzlich unterschiedlichen Vorstellungen von der Bedeutung dieses Begriffes beruht, als vielmehr auf der Schwierigkeit, den Faktorbegriff in ein formales und schlüssiges mathe-

matisches Konzept einzubetten, die allerdings zum Teil zu etwas unglücklichen oder unklaren Definitionen geführt hat.

In der vorliegenen Arbeit wird der Begriff des Faktors im Sinne der Definitionen aus Gruppe 3 verstanden. Dieser Zugang zu Faktoren als Argumenten einer Funktion wird im folgenden Abschnitt genauer ausgeführt. Die Formalisierung wird dabei so allgemein gehalten, dass nicht nur das spezielle Verfahren der Varianzanalyse, sondern auch z.B. nichtparametrische Verfahren in das Konzept integriert sind.

Kapitel 2

Statistische Modellbildung

2.1 Formulierung der Modelle

Im Sinne des dieser Arbeit zugrunde liegenden Konzeptes, Faktoren als Argumente einer Funktion aufzufassen, definieren wir nun die Modellgleichung eines statistischen Modells – zunächst in allgemeiner Form. Anschließend geben wir einige Spezialfälle dieser Definition an. Im darauffolgenden Abschnitt betrachten wir die Hoeffding-Zerlegung und wenden sie auf Funktionen an, die statistische Modelle beschreiben.

Definition 2.1 (statistisches Modell, Modellgleichung) *Sei* $X = (a^{(1)}, \dots , a^{(m)}, \varepsilon)$ *ein* $\mathbb{R}^{m+1}$*-wertiger Zufallsvektor mit unabhängigen Komponenten und* $G = G_1 \otimes \dots \otimes G_m \otimes G_\varepsilon$ *die Verteilung des Vektors* X*. Sei* B *ein Banachraum und* $\eta : \mathbb{R}^{m+1} \to B$ *eine Abbildung.*
Dann heißt (η, X) *statistisches Modell mit Modellgleichung*

$$\text{(M)} \qquad U = \eta(a^{(1)}, \dots , a^{(m)}, \varepsilon) = \eta(X) \, ,$$

wenn gilt:

1. *Die Funktion* η *ist a priori unbekannt.*
2. *Die Zufallsvariablen* $a^{(1)}, \dots , a^{(m)}$ *(bzw. ihre ganzzahligen Teile) sind beobachtbar.*
3. *Die Zufallsvariable* ε *ist nicht beobachtbar.*

In diesem Modell werden $a^{(1)}, \dots , a^{(m)}$ Faktoren *und* ε Versuchsfehler *genannt.*

Die Modellierung der Art des Einflusses der Faktoren in Form einer Funktion mit den Faktoren als Argumenten entspricht der intuitiven Vorstellung, die in den oben zitierten Umschreibungen (Gruppe 3) deutlich wird.

Auf die Unabhängigkeit der Faktoren kann hier nicht verzichtet werden, um wesentliche Resultate ableiten zu können. Natürlich lassen sich Anwendungsbeispiele finden, bei denen gewisse Einflussgrößen nicht unabhängig sind. Es fällt in den Bereich der Faktorenanalyse, aus einer gegebenen Menge von Faktoren solche auszuwählen oder zu konstruieren, die voneinander unabhängig sind. Wir werden diese Problematik hier nicht weiter verfolgen.

Nun werden einige Spezialfälle von Definition 2.1 betrachtet. Zunächst formulieren wir die Gleichung so, dass beobachtete Zufallsvariablen direkt modelliert werden.

Definition 2.2 (direkte Modellgleichung) *Es gelte (M) mit $B = \mathbb{R}$ und den Bezeichnungen $U = Y$ und $\eta = \phi$.*

(M1) $$Y = \phi(a^{(1)}, \dots, a^{(m)}, \varepsilon)$$

heißt direkte Modellgleichung eines statistischen Modells*, wenn die Zufallsvariable Y beobachtbar ist.*
Y wird Zielgröße *genannt.*

Von Interesse ist die konkrete Gestalt der unbekannten Funktion ϕ, die die Faktoren und den Versuchsfehler miteinander verknüpft.

Modellgleichung (M1) lässt sich zunächst in folgender Weise verallgemeinern:

Definition 2.3 (transformierte Modellgleichung) *Es gelte (M) mit $B = \mathbb{R}$ und den Notationen $U = Z$ und $\eta = \tilde{\phi}$.*

(M1a) $$Z = \tilde{\phi}(a^{(1)}, \dots, a^{(m)}, \varepsilon)$$

heißt transformierte Modellgleichung eines statistischen Modells*, wenn die Zufallsvariable Y beobachtbar ist und $\tilde{\phi} = q \circ \phi$ mit einer bekannten Funktion q und einer unbekannten Funktion ϕ ist.*
In diesem Modell heißt Z transformierte Zielgröße.

Der Unterschied zwischen (M1) und (M1a) besteht lediglich darin, dass auf $\phi(a^{(1)}, \dots, a^{(m)}, \varepsilon)$ noch eine Funktion q angewandt wird. Offenbar ist Definition 2.2 ein Spezialfall von Definition 2.3.
Bei (M1) und (M1a) wird die Zielgröße direkt modelliert. Es ist aber auch denkbar, lediglich die Verteilung der Zielgröße modellmäßig zu erfassen. Dies führt zu den folgenden Definitionen:

Definition 2.4 (Modellgleichung mit Hilfe von Verteilungsfunktionen) *Es gelte (M) mit einem Funktionenraum B, so dass die Verteilungsfunktionen auf $\mathbb{R}$ in B enthalten sind. Weiter sei $U = F$ und $\eta = \psi$.*

(M2) $$F(t) = \psi(a^{(1)}, \dots, a^{(m)})(t)$$

heißt Modellgleichung eines statistischen Modells mit Hilfe von Verteilungsfunktionen*, falls F die a priori unbekannte Verteilungsfunktion einer beobachtbaren Zufallsvariable Y ist.*
Y wird Zielgröße *genannt.*

Da die Verteilungsfunktion F von den Faktoren $a^{(1)}, \ldots, a^{(m)}$ abhängt, ist sie zufällig.
Zur Einbeziehung des Versuchsfehlers in die Modellgleichung vgl. die Anmerkung auf S.26.
Wir formulieren noch eine weitere Variante der Modellgleichung, bei der Parameter der Verteilung modelliert werden:

Definition 2.5 (Modellgleichung mit Hilfe von Parametern) *Es gelte (M) mit $B = \mathbb{R}$ und den Bezeichnungen $U = g(p)$ und $\eta = \rho$.*

$$\text{(M3)} \qquad g(p) = \rho(a^{(1)}, \ldots, a^{(m)})$$

heißt Modellgleichung eines statistischen Modells mit Hilfe von Parametern*, falls F die a priori unbekannte Verteilungsfunktion einer beobachtbaren Zufallsvariable Y ist und $F = F_p$ Element einer parametrisierten Familie $\{F_p, p \in \mathbf{P}\}$ von Verteilungen; außerdem ist $g : \mathbf{P} \longrightarrow \mathbb{R}$ bijektiv und bekannt. Y wird* Zielgröße*, g* Linkfunktion *genannt.*

Die Betrachtung von Verteilungsfunktionen legt es nahe, Funktionale dieser Verteilungsfunktionen zu analysieren. Eine Zerlegung, die ohne die Annahme einer parametrischen Familie von Verteilungen und ohne Momentenvoraussetzungen auskommt, ist die Quantil-Zerlegung, die im folgenden formuliert wird. Von besonderem Interesse ist dabei der Spezialfall der Median-Zerlegung, also der Fall $\alpha = \frac{1}{2}$.

Definition 2.6 (Modellgleichung mit Hilfe von Verteilungsquantilen) *Sei $\alpha \in]0, 1[$. Es gelte (M) mit $B = \mathbb{R}$, und es seien $U = Q(\alpha)$ und $\eta = \tau^\alpha$.*

$$\text{(M4)} \qquad Q(\alpha) = \tau^\alpha(a^{(1)}, \ldots, a^{(m)})$$

heißt Modellgleichung eines statistischen Modells mit Hilfe von Verteilungsquantilen*, falls $Q(\alpha)$ das α-Quantil der a priori unbekannten Verteilungsfunktion einer beobachtbaren Zufallsvariable Y ist.*
Y wird Zielgröße *genannt.*

In den Definitionen 2.2 und 2.3, in denen die Zielgröße bzw. die transformierte Zielgröße direkt modelliert werden, erscheint auch die Zufallsvariable ε als Argument der unbekannten Funktion ϕ bzw. $\tilde{\phi} = q \circ \phi$. Dagegen haben die Funktionen ψ und ρ aus den Definitionen 2.4, 2.5 und 2.6 lediglich die Faktoren $a^{(1)}, \ldots, a^{(m)}$ als Argumente. Allerdings wird in den letzteren Definitionen auch nur die Verteilung der Zielgröße modelliert, und diese Verteilung sollte sinnvollerweise nur von den Werten abhängen, die die Faktoren annehmen, während die Zielgröße zusätzlich von der unbeobachtbaren Zufallsvariable „Versuchsfehler" abhängt (vgl. hierzu auch die Anmerkung auf S. 26).

Modellbildung über die Gleichung (M1) führt zu den Linearen Modellen, über (M2) zu Nichtparametrischen Verfahren und über (M3) zu den sogenannten Verallgemeinerten Linearen Modellen. Diese Begriffe werden im weiteren noch genauer definiert und abgegrenzt werden.

Die in diesem Abschnitt eingeführte Terminologie soll nun anhand eines einfachen konkreten (und realen) Beispiels illustriert werden.

Beispiel:
Die Wirkung von zwei Düngemitteln A und B auf den Ertrag an Tomaten soll untersucht werden. Dazu wird eine Reihe mit Tomatenpflanzen in 11 Bereiche aufgeteilt, denen jeweils zufällig eine der beiden Düngemethoden zugeordnet wird. Die zugehörigen Daten (aus BOX ET AL. (1978, Table 4.1)) sind in Tabelle 2 auf Seite 66 angegeben.
Bei einer Modellierung des Versuchs betrachtet man einen Faktor $a = a^{(1)}$, der die Werte 1 (Düngemittel A) und 2 (Düngemittel B) annehmen kann; die Verteilung G_1 dieses Faktors ist die diskrete Gleichverteilung auf $\{1, 2\} \subset \mathbb{R}$. Es ist $X = (a, \varepsilon)$, dabei ist die Zufallsvariable a, die die Düngemittel spezifiziert, beobachtbar, der Versuchsfehler ε ist dagegen nicht beobachtbar.
Wird die Zielgröße (Tomatenertrag in Pfund) direkt modelliert, so lautet das Modell gemäß (M1) $Y = \phi(a, \varepsilon)$. Die beobachtbare Zufallsvariable „Ertrag“ lässt sich also darstellen als Funktion des verwendeten Düngers und eines nicht beobachtbaren Fehlerterms, in den z.B. Störfaktoren wie Boden- und Wetterverhältnisse eingehen.
Möglicherweise sind für den Versuchsplaner andere Maßeinheiten von Interesse, so dass die Zielgröße mit einer festen und bekannten Funktion q transformiert wird. Dann lässt sich Modellgleichung (M1a) verwenden, also $Z = q(Y) = q(\phi(a, \varepsilon)) = \tilde{\phi}(a, \varepsilon)$.
Bei einer Betrachtung der Verteilungsfunktion der Zielgröße ist es natürlich, anzunehmen, dass der Ertrag nach Anwendung von Düngemittel A eine Verteilung mit Verteilungsfunktion F_A besitzt. Diese unbekannte Verteilungsfunktion beinhaltet zufällige Ertragsschwankungen, z.B. durch Umwelteinflüsse und andere Einflussfaktoren, die nicht explizit im Versuch modelliert werden. Analog sei F_B die Verteilungsfunktion für den Ertrag bei Dünger B. Eine statistische Modellgleichung, die diese Überlegungen mathematisiert, lässt sich gemäß (M2) als $F(t) = \psi(a)(t)$ angeben. Dabei ist $\psi(a) = \begin{cases} F_A, & a = 1 \\ F_B, & a = 2 \end{cases}$.
Falls weitere Informationen oder Annahmen über die Verteilungsklasse vorliegen, der die Verteilungsfunktionen F_A und F_B entstammen, lässt sich die Dimension des bei einer Modellierung der Verteilungsfunktionen zunächst unendlich-dimensionalen Parameterraumes deutlich reduzieren. Wird beispielsweise Normalverteilung der Tomatenerträge unterstellt, so ist der Parameterraum zweidimensional.
Eine Modellierung der Verteilungsquantile, also eindimensionaler Größen, ist auch dann möglich, wenn keinerlei Annahmen über die in Frage kommenden Verteilungen getroffen werden. Das Modell lautet dann für unser Beispiel $Q(\alpha) = \tau^{\alpha}(a)$. Dabei bezeichnet $Q(\alpha)$ das α-Quantil der Verteilungsfunktion für den Tomatenertrag, $\tau^{\frac{1}{2}}(1)$ ist z.B. der Median der Verteilung bei Anwendung von Dünger A.

2.2 Zerlegungen

Um Einflüsse der Faktoren auf die Zielgröße im Einzelnen untersuchen zu können, werden additive Zerlegungen durchgeführt. Diese gehen auf HOEFFDING (1948) zurück und werden bisweilen in der Literatur auch als H- bzw. Hoeffding-Zerlegungen bezeichnet. Wir führen im folgenden Satz einige der grundlegenden Resultate dazu mit Beweisen an:

Satz 2.7 (Hoeffding-Zerlegung)
Eine von n unabhängigen Zufallsvariablen $X_1 \sim F_1, \ldots, X_n \sim F_n$ abhängige, reellwertige Funktion
$\eta(X_1, \ldots, X_n)$ *mit* $\mathrm{E}|\eta| < \infty$ *kann linear zerlegt werden in*

$$\eta(X_1, \ldots, X_n) = \eta^0 + \sum_{j=1}^{n} \sum_{\{\nu_1,\ldots,\nu_j\} \subset \{1,\ldots,n\}} \eta^{(j)}_{\nu_1,\ldots,\nu_j}(X_{\nu_1}, \ldots, X_{\nu_j}) \, .$$

Dabei gilt:

1. $\forall j = 1, \ldots, n \; \forall c = 1, \ldots, j-1 \; \forall \{\nu_{i_1}, \ldots, \nu_{i_c}\} \subset \{\nu_1, \ldots, \nu_j\}$*:*

   $$\mathrm{E}(\eta^{(j)}_{\nu_1,\ldots,\nu_j}(X_{\nu_1}, \ldots, X_{\nu_c}, X_{\nu_{c+1}}, \ldots, X_{\nu_j}) | X_{\nu_1} = x_{\nu_1}, \ldots, X_{\nu_c} = x_{\nu_c}) = 0 \, .$$

   (zur Vereinfachung der Notation sei o.E.d.A. $\{\nu_{i_1}, \ldots, \nu_{i_c}\} = \{\nu_1, \ldots, \nu_c\}$ *angenommen)*

2. $\mathrm{E}(\eta^{(j)}_{\nu_1,\ldots,\nu_j}(X_{\nu_1}, \ldots, X_{\nu_j})) = 0 \, .$

3. *Für* $j \neq k$ *und beliebige Mengen* $\{\nu_1, \ldots, \nu_j\} \subset \{1, \ldots, n\}, \; \{\mu_1, \ldots, \mu_k\} \subset \{1, \ldots, n\}$ *gilt: Die Funktionen* $\eta^{(j)}_{\nu_1,\ldots,\nu_j}$ *und* $\eta^{(k)}_{\mu_1,\ldots,\mu_k}$ *sind unkorreliert.*

4. *Für* $\{1, \ldots, n\} \supset \{\nu_1, \ldots, \nu_j\} \neq \{\mu_1, \ldots, \mu_j\} \subset \{1, \ldots, n\}$ *gilt ebenfalls: Die Funktionen* $\eta^{(j)}_{\nu_1,\ldots,\nu_j}$ *und* $\eta^{(j)}_{\mu_1,\ldots,\mu_j}$ *sind unkorreliert.*

Beweis: Definiere rekursiv:

$$\begin{aligned}
\eta^0 &:= \mathrm{E}(\eta(X_1, \ldots, X_n)) \, , \\
\eta^{(1)}_{\nu_1}(x_{\nu_1}) &:= \mathrm{E}(\eta(X_1, \ldots, X_n) | X_{\nu_1} = x_{\nu_1}) - \eta^0 \, , \\
\eta^{(c)}_{\nu_1,\ldots,\nu_c}(x_{\nu_1}, \ldots, x_{\nu_c}) &:= \mathrm{E}(\eta(X_1, \ldots, X_n) | X_{\nu_1} = x_{\nu_1}, \ldots, X_{\nu_c} = x_{\nu_c}) \\
&\quad - \sum_{j=1}^{c-1} \sum_{\{i_1,\ldots,i_j\} \subset \{\nu_1,\ldots,\nu_c\}} \eta^{(j)}_{i_1,\ldots,i_j}(x_{i_1}, \ldots, x_{i_j}) - \eta^0 \\
&\quad (2 \leq c \leq n) \, .
\end{aligned}$$

Dann gilt zunächst aufgrund der rekursiven Definition für jede Wahl von $x_1, \ldots, x_n$:

$$\eta(x_1, \ldots, x_n) = \mathrm{E}(\eta(X_1, \ldots, X_n) | X_1 = x_1, \ldots, X_n = x_n)$$

$$= \eta^{(n)}(x_1, \ldots, x_n) + \sum_{j=1}^{n-1} \sum_{\{\nu_1,\ldots,\nu_j\} \subset \{1,\ldots,n\}} \eta^{(j)}_{\nu_1,\ldots,\nu_j}(x_{\nu_1}, \ldots, x_{\nu_j}) + \eta^0 .$$

Weiterhin gilt:

$$\eta^0 = \int \ldots \int \eta(u_1, \ldots, u_n)\, dF_1(u_1) \ldots dF_n(u_n) ,$$

$$\begin{aligned} \eta^{(1)}_{\nu_1}(x_1) &= \int \ldots \int \eta(x_1, u_2, \ldots, u_n)\, dF_2(u_2) \ldots dF_n(u_n) - \eta^0 \\ &= \int \ldots \int \eta(u_1, \ldots, u_n)\, (dG_{x_1}(u_1) - dF_1(u_1)) dF_2(u_2) \ldots dF_n(u_n) \end{aligned}$$

(dabei und im weiteren Verlauf des Beweises bezeichne G_x die Verteilungsfunktion des Dirac-Maßes auf dem Punkt x) ,

$$\begin{aligned} &\eta^{(c)}_{\nu_1,\ldots,\nu_c}(x_1, \ldots, x_c) \\ &\quad = \int \ldots \int \eta(u_1, \ldots, u_n)\, (dG_{x_1}(u_1) - dF_1(u_1)) \ldots (dG_{x_c}(u_c) - dF_c(u_c)) \times \\ &\qquad\qquad \times\ dF_{c+1}(u_{c+1}) \ldots dF_n(u_n) . \end{aligned}$$

1. Zur weiteren Vereinfachung sei o.E.d.A. angenommen, dass gelte:

   $$\{\nu_1, \ldots, \nu_j\} = \{1, \ldots, j\} \text{ und } \{\nu_{i_1}, \ldots, \nu_{i_c}\} = \{1, \ldots, c\} .$$

   Dann ist

$$\begin{aligned} &\eta^{(j)}_{1,\ldots,j}(x_1, \ldots, x_j) \\ = &\int \ldots \int \eta(u_1, \ldots, u_n)\, (dG_{x_1}(u_1) - dF_1(u_1)) \ldots (dG_{x_j}(u_j) - dF_j(u_j)) \times \\ &\qquad \times\ dF_{j+1}(u_{j+1}) \ldots dF_n(u_n) \\ = &\int \ldots \int \eta(u_1, \ldots, x_j, \ldots, u_n)\, (dG_{x_1}(u_1) - dF_1(u_1)) \ldots (dG_{x_{j-1}}(u_{j-1}) - dF_{j-1}(u_{j-1})) \times \\ &\qquad \times\ dF_{j+1}(u_{j+1}) \ldots dF_n(u_n) \\ &- \int \ldots \int \eta(u_1, \ldots, u_n)\, (dG_{x_1}(u_1) - dF_1(u_1)) \ldots (dG_{x_{j-1}}(u_{j-1}) - dF_{j-1}(u_{j-1})) \times \\ &\qquad \times\ dG_{x_j}(u_j)(dF_{j+1}(u_{j+1}) \ldots dF_n(u_n) \\ &+ \int \ldots \int \eta(u_1, \ldots, u_n)\, (dG_{x_1}(u_1) - dF_1(u_1)) \ldots (dG_{x_{j-1}}(u_{j-1}) - dF_{j-1}(u_{j-1})) \times \\ &\qquad \times\ (dG_{x_j}(u_j) - dF_j(u_j))(dF_{j+1}(u_{j+1}) \ldots dF_n(u_n) \\ = &\int \ldots \int \eta(u_1, \ldots, x_j, \ldots, u_n)\, (dG_{x_1}(u_1) - dF_1(u_1)) \ldots (dG_{x_{j-1}}(u_{j-1}) - dF_{j-1}(u_{j-1})) \times \\ &\qquad \times\ dF_{j+1}(u_{j+1}) \ldots dF_n(u_n) \end{aligned}$$

$$
\begin{aligned}
& - \int \dots \int \eta(u_1, \dots, u_n)\,(dG_{x_1}(u_1) - dF_1(u_1)) \dots (dG_{x_{j-1}}(u_{j-1}) - dF_{j-1}(u_{j-1})) \times \\
& \qquad\qquad \times\ (dF_j(u_j) \dots dF_n(u_n) \\
= \ & \int \dots \int \eta(u_1, \dots, x_j, \dots, u_n)\,(dG_{x_1}(u_1) - dF_1(u_1)) \dots (dG_{x_{j-1}}(u_{j-1}) - dF_{j-1}(u_{j-1})) \times \\
& \qquad\qquad \times\ dF_{j+1}(u_{j+1}) \dots dF_n(u_n) \\
& - \eta^{(j-1)}_{1,\dots,j-1}(x_1, \dots, x_{j-1})
\end{aligned}
$$

$$
\begin{aligned}
\Longrightarrow\ & \mathrm{E}(\eta^{(j)}_{1,\dots,j}(X_1, \dots, X_{\nu_{j-1}}, X_j)|X_1 = x_1, \dots, X_{j-1} = x_{j-1}) \\
& \qquad = \eta^{(j-1)}_{1,\dots,j-1}(x_1, \dots, x_{j-1}) - \eta^{(j-1)}_{1,\dots,j-1}(x_1, \dots, x_{j-1}) \ = \ 0
\end{aligned}
$$

$$
\Longrightarrow\ \forall c = 1, \dots, j-1:\ \mathrm{E}(\eta^{(j)}_{1,\dots,j}(X_1, \dots, X_c, X_{c+1}, \dots, X_j)|X_1 = x_1, \dots, X_c = x_c) = 0\,.
$$

2. Diese Aussage folgt nun wegen

$$
\begin{aligned}
& \mathrm{E}(\eta^{(j)}_{1,\dots,j}(X_1, \dots, X_j)) \\
= \ & \mathrm{E}\left(\mathrm{E}(\eta^{(j)}_{1,\dots,j}(X_1, \dots, X_c, X_{c+1}, \dots, X_j)|X_1 = x_1, \dots, X_c = x_c)\right) \\
= \ & \mathrm{E}(0) \ = \ 0\,.
\end{aligned}
$$

3. Sei o.E.d.A. $j < k$ und $\{\nu_1, \dots, \nu_j\} \subset \{1, \dots, n\},\ \{\mu_1, \dots, \mu_k\} \subset \{1, \dots, n\}$.
   Nach 2. bleibt noch zu zeigen:

$$
\mathrm{E}(\eta^{(j)}_{\nu_1,\dots,\nu_j}(X_{\nu_1}, \dots, X_{\nu_j}) \cdot \eta^{(k)}_{\mu_1,\dots,\mu_k}(X_{\mu_1}, \dots, X_{\mu_k})) = 0\,.
$$

   Wegen $j < k$ gilt: $\exists \mu_\kappa:\ \mu_\kappa \notin \{\nu_1, \dots, \nu_j\}$. Damit ist

$$
\begin{aligned}
& \mathrm{E}(\eta^{(j)}_{\nu_1,\dots,\nu_j}(X_{\nu_1}, \dots, X_{\nu_j}) \cdot \eta^{(k)}_{\mu_1,\dots,\mu_k}(X_{\mu_1}, \dots, X_{\mu_\kappa}, \dots, X_{\mu_k})) \\
& = \mathrm{E}\left(\mathrm{E}(\eta^{(j)}_{\nu_1,\dots,\nu_j}(X_{\nu_1}, \dots, X_{\nu_j}) \cdot \eta^{(k)}_{\mu_1,\dots,\mu_k}(X_{\mu_1}, \dots, X_{\mu_\kappa}, \dots, X_{\mu_k})|X_{\nu_1} = x_{\nu_1}, \dots, X_{\nu_j} = x_{\nu_j})\right) \\
& = \mathrm{E}\left(\eta^{(j)}_{\nu_1,\dots,\nu_j}(X_{\nu_1}, \dots, X_{\nu_j}) \cdot \mathrm{E}(\eta^{(k)}_{\mu_1,\dots,\mu_k}(X_{\mu_1}, \dots, X_{\mu_\kappa}, \dots, X_{\mu_k})|X_{\nu_1} = x_{\nu_1}, \dots, X_{\nu_j} = x_{\nu_j})\right) \\
& \qquad = 0 \qquad \text{(mit 1.)}\,.
\end{aligned}
$$

4. Wegen $\{\nu_1, \dots, \nu_j\} \neq \{\mu_1, \dots, \mu_j\}$ gilt: $\exists \mu_\kappa:\ \mu_\kappa \notin \{\nu_1, \dots, \nu_j\}$.
   Es lässt sich somit weiter argumentieren wie bei 3.

□

Korollar 2.8 *Sei $\eta(X_1, \ldots, X_n)$ gegeben wie in Satz 2.7 und in der dortigen Form zerlegt. Dann gilt:*

$$\mathrm{E}(\eta(X_1,\ldots,X_n)|X_{\nu_1}=x_{\nu_1},\ldots,X_{\nu_k}=x_{\nu_k}) = \eta^0 + \sum_{j=1}^{k} \sum_{\{i_1,\ldots,i_j\}\subset\{\nu_1,\ldots,\nu_k\}} \eta^{(j)}_{i_1,\ldots,i_j}(x_{i_1},\ldots,x_{i_j}) \,.$$

Beweis:

$$\begin{aligned}
&\mathrm{E}(\eta(X_1,\ldots,X_n)|X_{\nu_1}=x_{\nu_1},\ldots,X_{\nu_k}=x_{\nu_k})\\
&= \mathrm{E}\left(\eta^0 + \sum_{j=1}^{n} \sum_{\{\nu_1,\ldots,\nu_j\}\subset\{1,\ldots,n\}} \eta^{(j)}_{\nu_1,\ldots,\nu_j}(X_{\nu_1},\ldots,X_{\nu_j}) \middle| X_{\nu_1}=x_{\nu_1},\ldots,X_{\nu_k}=x_{\nu_k}\right)\\
&= \eta^0 + \sum_{j=1}^{k} \sum_{\{i_1,\ldots,i_j\}\subset\{\nu_1,\ldots,\nu_k\}} \mathrm{E}(\eta^{(j)}_{i_1,\ldots,i_j}(X_{i_1},\ldots,X_{i_j})|X_{\nu_1}=x_{\nu_1},\ldots,X_{\nu_k}=x_{\nu_k})\\
&\quad + \sum_{j=k+1}^{n} \sum_{\{i_1,\ldots,i_j\}\supset\{\nu_1,\ldots,\nu_k\}} \underbrace{\mathrm{E}(\eta^{(j)}_{i_1,\ldots,i_j}(X_{i_1},\ldots,X_{i_j})|X_{\nu_1}=x_{\nu_1},\ldots,X_{\nu_k}=x_{\nu_k})}_{=0 \text{ (wegen Satz 2.7 (1))}}\\
&\quad + \sum_{j=1}^{n-k} \sum_{\{i_1,\ldots,i_j\}\cap\{\nu_1,\ldots,\nu_k\}=\emptyset} \underbrace{\mathrm{E}(\eta^{(j)}_{i_1,\ldots,i_j}(X_{i_1},\ldots,X_{i_j})|X_{\nu_1}=x_{\nu_1},\ldots,X_{\nu_k}=x_{\nu_k})}_{=0 \text{ (wegen Satz 2.7 (2))}}\\
&\quad + \sum_{j=1}^{n-1} \sum_{\substack{\{i_1,\ldots,i_j\}\cap\{\nu_1,\ldots,\nu_k\}\neq\emptyset\\ \{i_1,\ldots,i_j\}\not\supset\{\nu_1,\ldots,\nu_k\}\\ \{i_1,\ldots,i_j\}\not\subset\{\nu_1,\ldots,\nu_k\}}} \underbrace{\mathrm{E}(\eta^{(j)}_{i_1,\ldots,i_j}(X_{i_1},\ldots,X_{i_j})|X_{\nu_1}=x_{\nu_1},\ldots,X_{\nu_k}=x_{\nu_k})}_{=0 \text{ (wegen Satz 2.7 (1))}}\\
&= \eta^0 + \sum_{j=1}^{k} \sum_{\{i_1,\ldots,i_j\}\subset\{\nu_1,\ldots,\nu_k\}} \eta^{(j)}_{i_1,\ldots,i_j}(x_{i_1},\ldots,x_{i_j}) \,.
\end{aligned}$$

$\square$

Nun gibt es prinzipiell fünf Möglichkeiten, die Hoeffding-Zerlegung durchzuführen, je nachdem, ob ϕ, $\tilde{\phi} = q \circ \phi$, ψ, ρ oder τ^α additiv zerlegt wird. Diese Möglichkeiten werden im Folgenden dargestellt.

1. Additive Zerlegung der Funktion $\phi : \mathbb{R}^{m+1} \to \mathbb{R}$:

$$\begin{aligned}
& \phi(x_1,\ldots,x_m,x_{m+1})\\
= \ & \phi_\mu + \phi_{A^{(1)}}(x_1) + \ldots + \phi_{A^{(m)}}(x_m)\\
& +\phi_{A^{(1)},A^{(2)}}(x_1,x_2) + \ldots + \phi_{A^{(m-1)},A^{(m)}}(x_{m-1},x_m)\\
& + \ldots + \phi_{A^{(1)},\ldots,A^{(m)}}(x_1,\ldots,x_m) + \phi_\varepsilon(x_{m+1})\\
& +\phi_{A^{(1)},\varepsilon}(x_1,x_{m+1}) + \ldots + \phi_{A^{(m)},\varepsilon}(x_m,x_{m+1}) + \ldots\\
& +\phi_{A^{(1)},\ldots,A^{(m)},\varepsilon}(x_1,\ldots,x_m,x_{m+1})
\end{aligned}$$

$$= \quad \phi_\mu + \sum_{k=1}^{m} \sum_{\{i_1,\dots,i_k\}\subset\{1,\dots,m\}} \phi_{A^{(i_1)},\dots,A^{(i_k)}}(x_{i_1},\dots,x_{i_k})$$
$$+ \sum_{k=1}^{m} \sum_{\{i_1,\dots,i_k\}\subset\{1,\dots,m\}} \phi_{A^{(i_1)},\dots,A^{(i_k)},\varepsilon}(x_{i_1},\dots,x_{i_k},x_{m+1}) + \phi_\varepsilon(x_{m+1}) \, .$$

Die Aussagen 1. und 2. aus Satz 2.7 lauten hierbei im Einzelnen:

(R1) $\quad \mathrm{E}(\phi_\varepsilon(\varepsilon)) = \int \phi_\varepsilon(x) dG_\varepsilon(x) = 0 \, .$

(R2) $\quad \forall i: \ \mathrm{E}_{A^{(i)}}(\phi_{A^{(i)}}(a^{(i)})) = \int \phi_{A^{(i)}}(x_i) dG_i(x_i) = 0 \, .$

(R3) $\quad \forall \{\nu_1,\dots,\nu_j\} \subset \{1,\dots,m\} \ \forall \{\nu_{i_1},\dots,\nu_{i_c}\} \subsetneqq \{\nu_1,\dots,\nu_j\}:$
$\mathrm{E}(\phi_{A^{(\nu_1)},\dots,A^{(\nu_j)}}(a^{\nu_1},\dots,a^{\nu_j}) | a^{\nu_{i_1}} = \alpha_1,\dots,a^{\nu_{i_c}} = \alpha_c) = 0 \, .$
Dabei kann $\{\nu_{i_1},\dots,\nu_{i_c}\} = \emptyset$ sein.

(R4) $\quad \forall \{\nu_1,\dots,\nu_j\} \subset \{1,\dots,m+1\} \ \forall \{\nu_{i_1},\dots,\nu_{i_c}\} \subsetneqq \{\nu_1,\dots,\nu_j\}:$
$\mathrm{E}(\phi_{A^{(\nu_1)},\dots,A^{(\nu_j)},\varepsilon}(a^{\nu_1},\dots,a^{\nu_j},\varepsilon) | a^{\nu_{i_1}} = \alpha_1,\dots,a^{\nu_{i_c}} = \alpha_c) = 0 \, .$
Dabei kann $\{\nu_{i_1},\dots,\nu_{i_c}\} = \emptyset$ sein, und $a^{(m+1)}$ wird mit ε identifiziert.

Definition 2.9 *Die Aussagen (R1) bis (R4) werden als* Reparametrisierungsbedingungen *bezeichnet.*

Es ergibt sich nun z.B. (vgl. dazu den Beginn des Beweises von Satz 2.7):

$$\phi_\mu = \int \dots \int \phi(x_1,\dots,x_m,x_{m+1}) \, dG_1(x_1)\dots dG_m(x_m) dG_\varepsilon(x_{m+1}) \, ,$$
$$\forall j: \ \phi_{A^{(j)}}(x_j)$$
$$= \int \phi(a^{(1)},\dots,x_j,\dots,a^{(m)},\varepsilon) dG_{1,\dots,j-1,j+1,m,\varepsilon}(a^{(1)},\dots,a^{(j-1)},a^{(j+1)},\dots,a^{(m)},\varepsilon) - \phi_\mu \, .$$

Die Anwendung von Korollar 2.8 liefert uns schließlich noch folgendes Resultat:

$$\mathrm{E}(Y | a^{(\nu_1)} = x_{\nu_1},\dots,a^{(\nu_k)} = x_{\nu_k})$$
$$= \phi_\mu + \sum_{j=1}^{k} \sum_{\{i_1,\dots,i_j\}\subset\{\nu_1,\dots,\nu_k\}} \phi_{A^{(i_1)},\dots,A^{(i_j)}}(x_{i_1},\dots,x_{i_j}) \, . \tag{2.1}$$

Insbesondere gilt also

$$\mathrm{E}(Y | a^{(1)} = x_1,\dots,a^{(m)} = x_m)$$
$$= \phi_\mu + \sum_{j=1}^{m} \sum_{\{i_1,\dots,i_j\}\subset\{1,\dots,m\}} \phi_{A^{(i_1)},\dots,A^{(i_j)}}(x_{i_1},\dots,x_{i_j}) \, . \tag{2.2}$$

2. Additive Zerlegung der Funktion $\tilde{\phi} = q \circ \phi : \mathbb{R}^{m+1} \to \mathbb{R}$:
Vollkommen analog zur Zerlegung von ϕ ergibt sich hier:

$$\begin{aligned}
& \tilde{\phi}(x_1, \dots, x_m, x_{m+1}) \\
= \;& \tilde{\phi}_\mu + \sum_{k=1}^{m} \sum_{\{i_1,\dots,i_k\}\subset\{1,\dots,m\}} \tilde{\phi}_{A^{(i_1)},\dots,A^{(i_k)}}(x_{i_1}, \dots, x_{i_k}) \\
& \quad + \sum_{k=1}^{m} \sum_{\{i_1,\dots,i_k\}\subset\{1,\dots,m\}} \tilde{\phi}_{A^{(i_1)},\dots,A^{(i_k)},\varepsilon}(x_{i_1}, \dots, x_{i_k}, x_{m+1}) + \tilde{\phi}_\varepsilon(x_{m+1}) .
\end{aligned}$$

Man erhält analoge Aussagen zu (R1) bis (R4) und (1).

3. Additive Zerlegung der Abbildung $\psi : \mathbb{R}^m \to B$:
Die Zerlegung von ψ wird punktweise durchgeführt, d.h. $\forall t \in \mathbb{R}$:

$$\begin{aligned}
& \psi(x_1, \dots, x_m)(t) \\
= \;& \psi_\mu(t) + \sum_{k=1}^{m} \sum_{\{i_1,\dots,i_k\}\subset\{1,\dots,m\}} \psi_{A^{(i_1)},\dots,A^{(i_k)}}(x_{i_1}, \dots, x_{i_k})(t) .
\end{aligned}$$

Da eine solche Zerlegung für jedes $t \in \mathbb{R}$ möglich ist, können die Komponenten $\psi_{A^{(i_1)},\dots,A^{(i_k)}}(x_{i_1}, \dots, x_{i_k})$ selbst wieder als Funktionen angesehen werden.
Die Aussagen (R2) und (R3) haben hier punktweise Analoga.

Eine additive Zerlegung von Verteilungsfunktionen findet man bereits bei AKRITAS & ARNOLD (1994). AKRITAS & ARNOLD haben vorgeschlagen, mit Hilfe dieser Zerlegung Hypothesen zu beschreiben. Explizit formulieren sie das zweifaktorielle Nichtparametrische Modell als $F_{ij}(u) = M(u) + A_i(u) + B_j(u) + C_{ij}(u)$. Dabei ist F_{ij} die Verteilungsfunktion der Zielvariablen Y_{ijk} und die Funktionen M, A_i, B_j und C_{ij} sind analog zu den von uns angegebenen Funktionen definiert, so dass auch die entsprechenden Reparametrisierungsbedingungen gelten. Offenbar fügt sich ihr Zugang zu Nichtparametrischen Modellen nahtlos in das allgemeine Konzept einer Zerlegung geeigneter Funktionen ein.
Anmerkung: Der Satz über die Hoeffding-Zerlegung gilt auch in der folgenden, allgemeineren Form für Abbildungen in einen Banachraum:

Satz 2.10 *Sei B ein reeller, separabler Banachraum mit Norm $||\cdot||$. Seien $X_1, \dots, X_n$ unabhängige reellwertige Zufallsvariablen. Sei $\eta : \mathbb{R}^n \to B$ eine Bochner-integrierbare Abbildung mit* $\mathrm{E}||\eta|| < \infty$. *Dann gibt es Abbildungen*
$\eta^0,\ \eta^{(j)}_{\nu_1,\dots,\nu_j} : \mathbb{R}^j \to B,\ j = 1, \dots, n;\ \{\nu_1, \dots, \nu_j\} \subset \{1, \dots, n\}$*, mit*

$$\eta(X_1, \dots, X_n) = \eta^0 + \sum_{j=1}^{n} \sum_{\{\nu_1,\dots,\nu_j\}\subset\{1,\dots,n\}} \eta^{(j)}_{\nu_1,\dots,\nu_j}(X_{\nu_1}, \dots, X_{\nu_j}) .$$

Die Aussagen 1. und 2. in Satz 2.7 gelten hier genauso (siehe z.B. BOROVSKIKH *(1996)).*

Mit Satz 2.10 lässt sich der Zugang zur Modellgleichung über Verteilungsfunktionen auch wie folgt wählen: Betrachte $F = \psi(a^{(1)}, \ldots, a^{(m)}, \varepsilon)$ mit unabhängigen, reellwertigen Zufallsvariablen $a^{(1)}, \ldots, a^{(m)}, \varepsilon$ und $\eta : \mathbb{R}^{m+1} \to B$. Dabei enthalte der Banachraum B den Raum der Verteilungsfunktionen auf $\mathbb{R}$.
Zunächst zerlege man η in der in Satz 2.10 angegebenen Weise. Durch Integration über ε und die „zufälligen Faktoren" (s.u.) $a^{(i_1)}, \ldots, a^{(i_r)}$ ergibt sich dann eine Zerlegung, die äquivalent zur oben angegebenen (punktweisen) Zerlegung von ψ ist.
Es ist also prinzipiell möglich, den Versuchsfehlerterm zumindest in der Modellgleichung zu berücksichtigen. Spätestens bei der Schätzung der Verteilungsfunktionen macht es jedoch keinen Sinn, ihn mit aufzunehmen. Daher werden wir auch im folgenden Verteilungsfunktionen immer als Funktionen annehmen, die lediglich von den Faktoren abhängen.

4. Additive Zerlegung von $\rho : \mathbb{R}^m \to \mathbb{R}$:
Analog zu 3. erhält man:

$$\rho(x_1, \ldots, x_m) \;=\; \rho_\mu + \sum_{k=1}^{m} \sum_{\{i_1,\ldots,i_k\}\subset\{1,\ldots,m\}} \rho_{A^{(i_1)},\ldots,A^{(i_k)}}(x_{i_1}, \ldots, x_{i_k}) \,.$$

Die Aussagen (R2) und (R3) gelten ebenfalls analog. Auch bei der Zerlegung von ρ werden im Allgemeinen nur feste Faktoren betrachtet.

5. Additive Zerlegung von $\tau^\alpha : \mathbb{R}^m \to \mathbb{R}$:
Analog zu 3. und 4. erhält man:

$$\tau^\alpha(x_1, \ldots, x_m) \;=\; \tau^\alpha_\mu + \sum_{k=1}^{m} \sum_{\{i_1,\ldots,i_k\}\subset\{1,\ldots,m\}} \tau^\alpha_{A^{(i_1)},\ldots,A^{(i_k)}}(x_{i_1}, \ldots, x_{i_k}) \,.$$

Die Aussagen (R2) und (R3) gelten wieder analog. Bei der Zerlegung von τ werden wiederum gewöhnlich nur feste Faktoren betrachtet.

Definition 2.11 *In den genannten fünf Zerlegungen heißen ϕ_μ (bzw. $\tilde{\phi}_\mu, \rho_\mu, \psi_\mu, \tau_\mu$)* Gesamtmittel, *$\phi_\varepsilon$ (bzw. $\tilde{\phi}_\varepsilon$)* Residuum (Fehler)*, die Funktionen $\phi_{A^{(i)}}(x_i)$ (bzw. $\tilde{\phi}_{A^{(i)}}$, $\rho_{A^{(i)}}$, $\psi_{A^{(i)}}$, $\tau_{A^{(i)}}$), die von genau einem Faktor als Argument abhängen,* Haupteffekte*,*
die Funktionen $\phi_{A^{(i_1)},\ldots,A^{(i_k)}}(x_{i_1}, \ldots, x_{i_k})$ (bzw. $\tilde{\phi}_{A^{(i_1)},\ldots,A^{(i_k)}}$, $\rho_{A^{(i_1)},\ldots,A^{(i_k)}}$, $\psi_{A^{(i_1)},\ldots,A^{(i_k)}}$, $\tau_{A^{(i_1)},\ldots,A^{(i_k)}}$) mit card$\{i_1, \ldots, i_k\} \geq 2$ Wechselwirkungen*,*
und die Funktionen $\phi_{A^{(i_1)},\ldots,A^{(i_k)},\varepsilon}(x_{i_1}, \ldots, x_{i_k}, x_{m+1})$ (bzw. $\tilde{\phi}_{A^{(i_1)},\ldots,A^{(i_k)},\varepsilon}$) Wechselwirkungen mit dem Residuum.

Die Zerlegung einer Zufallsvariablen, die eine Funktion von n unabhängigen Zufallsvariablen ist, in „Haupteffekte", „Wechselwirkungen" usw., ist also nicht auf den klassischen Fall der ANOVA beschränkt, sondern lässt sich auch unter allgemeineren Rahmenbedingungen durchführen, bei denen nicht die Zielgrößen selbst, sondern mit ihnen korrespondierende Größen zerlegt werden.

Die große Bedeutung, die eine derartige Zerlegung allerdings gerade im Bereich der Varianzanalyse erlangt hat, wird dadurch deutlich, dass sie bisweilen als „ANOVA decomposition“ oder „ANOVA type decomposition“ bezeichnet wird (vgl. z.B. EFRON & STEIN (1981) und KARLIN & RINOTT (1982), die diesen Ausdruck im Zusammenhang mit Jackknife-Varianzschätzern bei der H-Zerlegung einer Zufallsvariablen $S(X_1, \ldots, X_n)$ verwenden).
Einen anderen Zugang zu dieser Thematik wählt TAKEMURA (1983). Mit Hilfe der Bezeichnungsweisen aus der Tensoranalysis und der Multilinearen Algebra führt er die Analogie zwischen ANOVA und Hoeffding-Zerlegung auf die Orthogonalität von Unterräumen eines Vektorraums zurück. Seine zunächst etwas abstrakte Herangehensweise vermeidet gewisse Nachteile bei der herkömmlichen Verwendung von Kronecker-Produkten in der Statistik (dadurch, dass Kronecker-Produkte von Matrizen als partitionierte Matrizen höherer Dimension aufgefasst werden, ensteht z.B. das Problem der Sortierung von Indizes und es gehen Symmetrieeigenschaften verloren). Allerdings erkennt TAKEMURA nicht, dass die Haupteffekte und Wechselwirkungen in der Varianzanalyse tatsächlich bedingte Erwartungen sind und nicht nur Analogien zwischen Effekten und bedingten Erwartungen bestehen. Seine Arbeit ist darüber hinaus auf den Fall der ANOVA beschränkt, die bei uns mit der Zerlegung der Funktion ϕ korrespondiert.
Ebenfalls für den Fall der Varianzanalyse leitet HELLAND (1998) eine eindeutige Zerlegung der Beobachtung y her, die auf orthogonalen Projektionen basiert. Er erhält auch ein analoges Resultat zu der von uns formulierten Aussage (2.1) auf S.25.

Anmerkung: Die Tatsache, dass die sich hier ergebenden Zerlegungen immer additiv sind, mag zunächst restriktiv erscheinen. Gilt beispielsweise $Y = a^{(1)} \cdot \ldots \cdot a^{(m)} \cdot \varepsilon$, so wird eine additive Zerlegung zwar möglich, aber unangemessen sein. Ein sinnvoller additiver Zusammenhang zwischen Faktoren und Zielgröße besteht erst nach einer Transformation:

$$\log(Y) = \log(a^{(1)}) + \ldots + \log(a^{(m)}) + \log(\varepsilon) \ .$$

Allerdings gehen wir bei der Modellbildung auch davon aus, dass die Funktion ϕ, die die Faktoren und den Versuchsfehler verknüpft, vollständig unbekannt ist. Falls konkrete Vermutungen über die Gestalt von ϕ vorliegen, lassen sich bisweilen Transformationen von Zielgröße und Faktoren finden, durch die Additivität ohne Wechselwirkungen erzeugt wird. Im einfachen Fall, wenn z.B. $\phi(x, y)$ eine Funktion in zwei Argumenten ist, gibt darüber der folgende Satz (s. SCHEFFÉ (1959, S.95f)) umfassend Auskunft – einschließlich eines Verfahrens zur Bestimmung der transformierenden Funktion:

Satz 2.12 *Für eine gegebene Funktion $\phi(x, y) \in C^2(\mathbb{R})$ existieren Funktionen $q(\phi),\ g(x)$ und $h(y) \in C^2(\mathbb{R})$ mit*

$$q(\phi(x, y)) = g(x) + h(y) \qquad \textit{und} \qquad q'(\phi) > 0$$

genau dann, wenn folgendes gilt:

$$\exists \ \textit{integrierbares} \ w : \frac{\partial^2 \phi(x, y)}{\partial x \partial y} = w(\phi) \cdot \frac{\partial \phi(x, y)}{\partial x} \cdot \frac{\partial \phi(x, y)}{\partial y} \ .$$

Die Funktionen $q,\ g$ und h erhält man dann durch

$$q(\phi) = c_1 \int \exp\Big(- \int w(\phi)d\phi\Big)d\phi + c_2$$

($c_1 > 0$ und c_2 sind dabei Konstanten),

$$g(x) = \int \frac{\partial\phi(x,y)}{\partial x} q'(\phi)dx + c_3$$

($\frac{\partial\phi(x,y)}{\partial x} q'(\phi)$ ist eine Funktion in x und c_3 ist eine Konstante)

und $\qquad h(y) = q(\phi) - g(x).$

Beweis: Ein Beweis befindet sich bei SCHEFFÉ (a.a.O.). □

In Kapitel 2.4 werden wir die Thematik der Transformationen noch einmal aufgreifen.

Anmerkung: Wenn man sich mit additiven Zerlegungen von Zufallsvariablen und anderen Funktionen beschäftigt, ist es naheliegend, auch die Fourier-Transformierten (charakteristischen Funktionen) der entsprechenden Verteilungen zu betrachten. Für unabhängige Zufallsvariablen $X_1, \ldots, X_n$, deren Verteilungen die Fourier-Transformierten $h_{X_1}, \ldots, h_{X_n}$ besitzen, gilt bekanntlich:

$$\log(h_{X_1+\ldots+X_n}) = \log(h_{X_1}) + \ldots + \log(h_{X_n})\ .$$

Eine additive Zerlegung in unabhängige Zufallsvariablen korrespondiert also mit einer additiven Zerlegung der logarithmierten charakteristischen Funktionen.
Die Bedeutung dieser Aussage für die hier verwendeten Zerlegungen wird nachfolgend beispielhaft für den Fall genau eines Faktors illustriert:

Sei $Y = \phi(a,\varepsilon) = \phi_\mu + \phi_A(a) + \phi_\varepsilon(\varepsilon)$ mit $(a,\varepsilon) \sim G_A \times G_\varepsilon$. Es gilt dann für die charakteristische Funktion h_Y von Y:

$$\log(h_Y(t)) = \log(h_{\phi_\mu}(t)) + \log(h_{\phi_{A(a)}}(t)) + \log(h_{\phi_{\varepsilon(\varepsilon)}}(t))\ .$$

Dabei sind $h_Y(t) = \int \exp(\mathrm{i}t\phi(x_1,x_2))\ dG_A(x_1)dG_\varepsilon(x_2)$, $h_{\phi_\mu}(t) = \exp(\mathrm{i}t\phi_\mu)$,
$h_{\phi_{A(a)}}(t) = \int \exp(\mathrm{i}t\phi_A(x))\ dG_A(x)$ und $h_{\phi_{\varepsilon(\varepsilon)}}(t) = \int \exp(\mathrm{i}t\phi_\varepsilon(x))\ dG_\varepsilon(x)$.

Analoge Zerlegungen ergeben sich bei mehreren Faktoren. Wir werden im Folgenden allerdings die hier angedeutete Möglichkeit der Zerlegung charakteristischer Funktionen nicht vertiefen.

2.3 Faktoren

Es ist in statistischen Modellen üblich, zwischen sogenannten *festen Faktoren*, *zufälligen Faktoren* und *Kovariablen* zu unterscheiden. Dabei ist, wie schon bei der Definition des Begriffes *Faktor*, die exakte Abgrenzung dieser Bezeichnungen in der Literatur nicht ganz einheitlich - oder überhaupt nicht vorhanden. Es besteht allerdings inzwischen ein gewisses Einvernehmen in der Literatur über Lineare Modelle, *wie* mit festen oder zufälligen Faktoren umzugehen ist, wenn sie erst einmal als solche erkannt sind. Bei Anwendungen ist jedoch die Frage, *ob* ein Faktor fest oder zufällig ist, häufig der Intuition des mehr oder weniger erfahrenen Statistikers überlassen.
SCHEFFÉ (1956) äußert sogar die Ansicht, dass es nicht den Daten inhärent ist, ob Faktoren als fest oder zufällig zu betrachten sind, sondern dass der Zweck der Analyse entscheidend ist:

> „It is conceivable that for two different purposes the same data might be analyzed according to two different models in which the same main effects are regarded as fixed or as random effects."
> SCHEFFÉ (1956, S.254f)

BENNINGTON & THAYNE (1994) sind der Frage nachgegangen, ob die Klassifizierung *fest* oder *zufällig* in Umweltstudien sorgfältig vorgenommen wird. Sie haben unter anderem festgestellt, dass sehr häufig (in 63% der von ihnen untersuchten Artikel, in denen ANOVA mit mindestens zwei Faktoren verwendet wird) nicht erwähnt wird, ob betrachtete Faktoren *fest* oder *zufällig* sind, und dass nicht klar ist, ob die jeweiligen Autoren diesen Unterschied überhaupt kennen.

Dass es für die Erwartungswerte bestimmter Quadratischer Formen einen Unterschied macht, ob man einen Faktor als *fest* oder *zufällig* annimmt, ist allerdings schon spätestens seit den 50er Jahren bekannt. Und eben dieser Unterschied führt dazu, dass statistische Auswertungen und somit auch die Resultate und damit verknüpfte Entscheidungen wesentlich davon abhängen können, welche Annahmen über die Faktoren gemacht worden sind.

Die Klassifizierung eines bestimmten Faktors A als fest oder zufällig betrifft auch die Bedeutung der übrigen Faktoren und die unmittelbare Definition ihrer Effekte, da diese im Allgemeinen als Mittel über die Stufen des Faktors A aufgefasst werden. Die Entscheidung über *fest* oder *zufällig* impliziert eine Entscheidung darüber, ob die Mittelung lediglich über die im Experiment verwendeten Stufen (bei einem festen Faktor) oder über eine größere Menge von Faktorstufen vorzunehmen ist, aus der für das Experiment nur ein Teil (bei einem zufälligen Faktor) ausgewählt wurde (vgl. SCHEFFÉ (1956)).

In diesem Abschnitt werden nun zunächst einige für die historische Entwicklung und das heutige Verständnis der Begriffe *fest* und *zufällig* relevante Definitionen aus der Literatur aufgeführt. Danach stellen wir wesentliche Eigenschaften von zufälligen und festen Faktoren sowie Kovariablen zusammen, die kohärent mit der hier verwendeten grundlegenden Definition eines Faktors (s. Kapitel 2.1) sind.

20 Jahre nachdem FISHER in den 1920er Jahren den Begriff der Varianzanalyse geprägt hatte (er bezeichnete die Varianzanalyse selbst später als „not a mathematical theorem, but rather a convenient method of arranging the arithmetic", vgl. FISHER (1934, S.52)), wurde es

Zeit, die zugrundeliegenden Voraussetzungen und Annahmen noch einmal klar zu formulieren, denn die meisten Bücher, die in der Zwischenzeit zu dem Thema erschienen waren, waren eher Rechenanleitungen zur praktischen Durchführung einer Varianzanalyse als „Lehrbücher". Modellvoraussetzungen spielten in ihnen nur eine untergeordnete Rolle. Diese Zusammenfassung wichtiger Annahmen der Varianzanalyse hat EISENHART (1947) unternommen. Dabei hat er auch eine Unterscheidung zwischen zwei Modellen getroffen, die heute als *Modell mit festen Faktoren* und *Modell mit zufälligen Faktoren* bezeichnet werden.
Das Modell mit festen Faktoren (Modell I, „Detection and Estimation of Fixed (Constant) Relations Among the Means of Sub-Sets of the Universe of Objects Concerned") dient laut EISENHART (1947, S.3f) dem Testen, ob tatsächliche Differenzen zwischen Behandlungs- (und anderen) Mittelwerten vorhanden sind. Dagegen soll das Modell mit zufälligen Faktoren (Modell II, „Detection and Estimation of Components of (Random) Variation Associated with a Composite Population") die Existenz von Varianzkomponenten aufdecken, die der zufälligen Abweichung charakteristischer Eigenschaften von Individuen von jeweiligen Mittelwerten in der „Population" aller Individuen desselben generischen Typs zuzuschreiben sind.

Und wie entscheidet der Anwender, welches Modell angemessen ist?
Dazu schreibt EISENHART:

> „The answer depends in part [...], upon how the observations were obtained; on the extent to which the experimental procedure employed sampled the respective variables at random."
> EISENHART (1947, S.19)

Schwierigkeiten treten allerdings bei den Faktoren „Ort" und „Zeit" auf:

> „Are the fields on which the tests were conducted a random sample of the county, or of the state, etc.? Are the years in which the tests were conducted a random sample of years?"
> EISENHART (1947, S.20)

Als Entscheidungshilfen, ob Effekte fest oder zufällig sind, formuliert EISENHART eine „Verallgemeinerungsregel" und eine „Wiederholungsregel":

> „(1) Are the conclusions to be confined to the things actually studied [...]; to the immediate sources of these things [...]; or expanded to apply to more general populations [...]?
> (2) In complete repetitions of the experiment would the same things be studied again [...]; would new samples be drawn from the identical sources [...]; or would new samples be drawn from the more general populations [...]?"
> EISENHART (1947, S.20)

WILK & KEMPTHORNE (1955) liefern nicht nur einen mathematischen Zugang zu festen und zufälligen Faktoren, bei dem der Prozess der Randomisierung eine zentrale Rolle spielt, sondern sie ermöglichen auch das Betrachten von Faktoren, die weder fest noch zufällig sind und als

„semi-random“ bezeichnet werden.
Dazu betrachten sie ein zweifaktorielles Design mit Faktoren $\mathcal{A}$ und $\mathcal{B}$, die A bzw. B Stufen haben. Das allgemeine Modell lautet dann

$$Y_{ijk} = \mu + a_i + b_j + (ab)_{ij} + e_k \ .$$

(Ein weiterer Term n_{ijk} wird von den Autoren zunächst mit ins Modell aufgenommen, aber für die Herleitung der Varianzanalysetabelle der Einfachheit halber gleich Null gesetzt.)
Im Experiment werden a Stufen von $\mathcal{A}$ und b Stufen von $\mathcal{B}$ zufällig gewählt. Die Formulierung des Modells für das Experiment geschieht mit Hilfe der als „Cornfields Trick“ aus der Sampling Theory bekannten Methode der Einführung von Dummy-Variablen (vgl. u.a. CORNFIELD (1944)).
Gilt $A = a, B = b$, so wird das Modell als „fixed model“ bezeichnet, im Fall $A \gg a, B \gg b$ („$\gg$“ bedeute hier „viel größer als“) liegt ein „random model“ vor, bei $A = a, B \gg b$ haben wir ein „mixed model“ usw.
WILK & KEMPTHORNE berechnen im zweifaktoriellen Modell mit $r \cdot a \cdot b$ experimentellen Einheiten, die den $a \cdot b$ ausgewählten Faktorstufen randomisiert zugeteilt werden, die Erwartungswerte der mittleren Quadratsummen für die verschiedenen Faktoren wie in der folgenden Tabelle angegeben.

Faktor	Erwartungswert
A	$\sigma^2_\varepsilon + \frac{B-b}{B} r\sigma^2_{ab} + rb\sigma^2_a$
B	$\sigma^2_\varepsilon + \frac{A-a}{A} r\sigma^2_{ab} + ra\sigma^2_b$
AB	$\sigma^2_\varepsilon + r\sigma^2_{ab}$
ε	σ^2_ε

Tabelle 1: Erwartungswerte der mittleren Quadratsummen im zweifaktoriellen Design nach WILK & KEMPTHORNE (1955, S.1152).

Heutzutage wird diese Tabelle gewöhnlich nur für die Extremfälle $A = a$ oder $A = \infty$ bzw. $B = b$ oder $B = \infty$ verwendet, in denen die Koeffizienten $\frac{B-b}{B}$ bzw. $\frac{A-a}{A}$ verschwinden oder gleich Eins sind. Sie bietet aber auch eine Alternative für die nicht seltenen Situationen, in denen die Zuordnung *fest* oder *zufällig* nicht so klar ist bzw. z.B. weder $A = a$ noch $A = \infty$ gilt. Allerdings müssen dann zum Testen eines Faktoreffekts Linearkombinationen quadratischer Formen für den Nenner gebildet werden.

Ein wesentlicher Unterschied des Zugangs von WILK & KEMPTHORNE zum heute verbreiteten Ansatz ist, dass die beiden Autoren einen großen Bogen um die Annahme der Normalverteilung machen. Gewöhnlich wird heutzutage ein zufälliger Faktor mit einer normalverteilten Zufallsvariablen identifiziert, deren Varianzkomponente zu schätzen ist. Die von WILK & KEMPTHORNE angegebenen Ausdrücke σ^2_a, σ^2_b und σ^2_{ab} sind aber eher als Varianzen einer auf den Faktorstufen gleichverteilten Zufallsvariablen zu interpretieren:

$$\sigma^2_a = \frac{1}{A-1}\sum_{i=1}^{A} a_i^2, \quad \sigma^2_b = \frac{1}{B-1}\sum_{i=1}^{B} b_i^2, \quad \sigma^2_{ab} = \frac{1}{(A-1)(B-1)}\sum_{ij}(ab)^2_{ij} \ .$$

Ebenso verhält es sich mit dem Fehlerterm (*unit error*) e_k, der näher erläutert wird als

> „due to failure of different experimental units to yield identically under the same conditions of treatment [...].
> The unit errors must be regarded as fixed quantities associated with particular experimental units.“
> WILK & KEMPTHORNE (1955, S.1148)

Andere Fehlerquellen werden als vernachlässigbar angesehen, und eine Versuchswiederholung mit denselben experimentellen Einheiten würde zu denselben Fehlertermen führen. Ob dies plausibel ist, scheint fraglich; der Ansatz von WILK & KEMPTHORNE ist auch in der Welt der Statistik kontrovers diskutiert worden.
Jedenfalls ist es innerhalb ihrer Prämissen konsistent, die Varianz des Fehlers als

$$\sigma_\varepsilon^2 = \frac{1}{abr-1}\sum_{k=1}^{abr} e_k^2$$

zu definieren.

RIDER, HARTER & LUM (1956) verwenden ebenfalls die Bezeichnungsweise *semi-random factor*. Sie definieren, sehr ähnlich wie WILK & KEMPTHORNE:

> „If the levels of a factor which are included in an experiment are considered to constitute the entire population of levels of that factor, the factor is said to be *fixed*. If, on the other hand, the levels included in the experiment represent a random sample from an infinite population of such levels, the factor is said to be *random*. If the levels included in the experiment represent a random sample from a larger (but finite) population, the factor is said to be *semi-random*.“
> RIDER, HARTER & LUM (1956, S.13)

Im Allgemeinen wird in den Standardwerken der heutigen statistischen Literatur ein zufälliger Faktor A durch eine Zufallsvariable repräsentiert, die $\mathrm{N}(0, \sigma_A^2)$-verteilt ist (vgl. z.B. KIRK (1968), SEARLE (1971) oder ARNOLD (1981)). Die Modellgleichung eines zweifaktoriellen Designs lautet zunächst

$$Y_{ijk} = \mu + a_i + b_j + (ab)_{ij} + \varepsilon_{ijk} \ .$$

Im Unterschied zu WILK & KEMPTHORNE (1955) wird allerdings bezüglich der Verteilungen angenommen, dass $\varepsilon_{ijk} \overset{u.i.v.}{\sim} \mathrm{N}(0, \sigma_\varepsilon^2)$ gilt, und darüber hinaus für einen zufälligen Faktor A mit Stufen $a_i,\ i = 1, \ldots, a$ auch $a_i \overset{u.i.v.}{\sim} \mathrm{N}(0, \sigma_A^2)$.

Die Frage, nach welchen Kriterien ein Faktor als *fest* oder *zufällig* zu bezeichnen ist, wird von KIRK im wesentlichen mit der „Verallgemeinerungsregel“ und der „Wiederholungsregel“ beantwortet :

> „<The *fixed-effects* model> is appropriate for experiments in which all treatment levels about which inferences are to be drawn are included in the experiment. If the experiment were replicated, the same treatment levels would be included in the replication. Under these conditions, conclusions drawn from the experiment apply only to the actual k treatment levels in the experiment. If the k treatment levels included in the experiment are a random sample from a much larger population of K levels, the model is a *random-effects* model.“
> KIRK (1968, S.51)

KIRK lässt es allerdings offen, auf welche Art und Weise aus einer Zufallsstichprobe aus einer endlichen Grundgesamtheit der Größe K eine normalverteilte Zufallsvariable konstruiert wird.

SEARLE verwendet ebenfalls das *Verallgemeinerungskriterium*: Für einen festen Faktor liege das Interesse desjenigen, der den Versuch durchführt, lediglich auf den im Experiment verwendeten Stufen des Faktors, und nicht auf irgendwelchen anderen Faktorstufen (SEARLE (1971, S.377f). Wie zahlreiche andere Autoren auch, führt er einleuchtende Beispiele an, um den Unterschied zwischen *festem* und *zufälligem* Fall deutlich zu machen. Beim Beispiel eines zufälligen Faktors handelt es sich dabei um die Versuchseinheiten, in dem Fall Mäuse:

> „But in the mouse experiment each mouse is merely a sample (of one) from a population of female mice. Nothing important has conditioned our choosing any one mouse over another, and we have no specific interest in the difference between any one of our 4 mice and any other of them. Interest does lie, however, in the extent to which maternal ability varies throughout the population of mice, and to this end our model is directed.“
> SEARLE (1971, S.378)

Hieraus lässt sich eine Art „Vergleichsregel“ interpretieren, die wir weiter unten präzisieren werden.
Dass die Grenzen zwischen fest und zufällig nicht klar gesetzt sind, macht SEARLE ebenfalls anhand von Beispielen klar. Der Anschaulichkeit halber seien hier Auszüge zitiert:

> „[...] With the technicians, for example, the situation might have been not that each one came and went as a random sample of employees, so to speak, but that all were available and we wanted to assess differences between these three specific technicians. In that case the technicians effects [...] would be fixed effects, not random. Similarly with the herd effects [...]. Analyses of data of such situations usually involve hundreds of herds that are considered a random sample from some larger population of herds. But were the situation to be one of analyzing just a few herds, five or six say, wherein the sole interest lay in just those herds, then herds effect [...] would more appropriately be fixed and not random. Thus we see that the situation to which a model applies is the deciding factor in determining whether the effects of a factor are fixed or random.“
> SEARLE (1971, S.382)

Das schon von EISENHART artikulierte Problem der Einordnung des Faktors „Jahr“ wird auch von SEARLE aufgegriffen:

> „Are the effects of years on yield to be considered fixed or random? The years themselves are unlikely to be random, for they will probably be a group of consecutive years over which the data have been gathered or the experiments run. But the effects on yield may reasonably be considered random – unless, perhaps, one is interested in comparing specific years for some purpose.“
> SEARLE (1971, S.382)

Dass eine befriedigende Lösung des Entscheidungsproblems *fest* oder *zufällig* bis heute nicht gefunden ist, zeigt die immer wieder aufflackernde und zum Teil kontrovers geführte Diskussion in der Literatur (vgl. dazu z.B. NELDER (1977) und die anschließende Diskussion – NELDER selbst identifiziert feste und zufällige Effekte mit vollständigem und unvollständigem Erheben von Stichproben).

COX (1984) hat einen weiteren Vorschlag der Unterscheidung zwischen verschiedenen Faktoren gemacht. Er differenziert zwischen *treatment variables* (deren Stufen vom Experimentator festgelegt und den experimentellen Einheiten durch Randomisierung zugeordnet werden), *intrinsic variables* (die Eigenschaften der Individuen und Umwelteinflüsse modellieren und außerhalb der Kontrolle des Experimentators liegen) und *nonspecific variables* (deren Stufen nicht eindeutig charakterisiert sind). COX schreibt dazu, dass sich das Interesse bei einer Auswertung auf die *treatment variables* beschränkt, die den festen Faktoren herkömmlichen Verständnisses entsprechen. Weiterhin werden Wechselwirkungen mit den *nonspecific variables* - also den als „zufällig“ etablierten Faktoren - als Teil des Versuchsfehlers aufgefasst.
Die Nomenklatur von COX hat sich allerdings nicht durchsetzen können. Es scheint auch etwas restriktiv, dass lediglich die Einflüsse fester Faktoren von Interesse sein sollen.

Im Handbuch zu Mixed Models des wohl einflussreichsten Statistik-Programmpaketes SAS beschränken sich die Autoren dort, wo sie feste und zufällige Faktoren erläutern, im Wesentlichen auf einleuchtende Beispiele und die Regeln, die schon in ähnlicher Weise von EISENHART (1947) formuliert wurden:

> „Factor effects are either fixed or random depending on how levels of factors that appear in the study are selected. An effect is called *fixed* if the levels in the study represent all possible levels of the factor, or at least all levels about which inference is to be made. [...] Factor effects are *random* if the levels of the factor that are used in the study represent only a random sample of a larger set of potential levels. The factor effects corresponding to the larger set of levels constitute a population of effects with a probability distribution.“
> LITTELL ET AL. (1996, S.ix)

Im Folgenden formulieren wir Abgrenzungen zwischen *festen*, *zufälligen* und *kovariaten* Faktoren, die in das angegebene generelle Konzept von Faktoren eingebettet sind. Dabei geben wir auch die aus Sicht des Anwenders hilfreichen Regeln zur *Auswahl*, *Wiederholung*, *Verallgemeinerung* und zum *Vergleich* an. Da Faktoren als Zufallsvariablen aufgefasst werden, ist eine naheliegende Frage die, welche Bildräume und welche Verteilungen sie besitzen sollten.

1. Ein Faktor $a^{(f)}$ heißt *fester Faktor*, wenn gilt:

   (a) Sei G_f die Randverteilung von $a^{(f)}$. Dann ist G_f die diskrete Gleichverteilung auf $E^{(f)} = \{e_1^{(f)}, \ldots, e_{l_f}^{(f)}\}$, d.h. $\forall x \in E^{(f)} : P(a^{(f)} = x) = \frac{1}{l_f}$.
   Als Konvention zur Vereinfachung der Notation sei $\forall i \in \{1, \ldots, l_f\} : e_i^{(f)} = i$.

   (b)
   $$\begin{aligned} \phi_{A^{(f)}} : E^{(f)} &\rightarrow \chi_f = \{\alpha_1^{(f)}, \ldots, \alpha_{l_f}^{(f)}\} . \\ e_j^{(f)} &\mapsto \alpha_j^{(f)} . \end{aligned}$$

   Der Zielbereich χ_f der Abbildung $\phi_{A^{(f)}}$ ist eine diskrete, unbekannte Teilmenge aus $\mathbb{R}$, und $\phi_{A^{(f)}}$ ist punktweise definiert, lässt sich somit nicht in geschlossener algebraischer Form darstellen.
   Der Funktionenraum $\Phi_{A^{(f)}}$, aus dem $\phi_{A^{(f)}}$ stammt, ist l_f-dimensional.

   (c)
   $$\begin{aligned} \psi_{A^{(f)}} : E^{(f)} &\rightarrow \{F_1^{(f)}, \ldots, F_{l_f}^{(f)}\} . \\ e_j^{(f)} &\mapsto F_j^{(f)} . \end{aligned}$$

   Der Zielbereich der Abbildung $\psi_{A^{(f)}}$ ist eine diskrete, unbekannte Menge von Funktionen $F_j^{(f)} : \mathbb{R} \rightarrow [0, 1]$, $j = 1, \ldots, l_f$. Es handelt sich dabei allerdings nicht notwendig um Verteilungsfunktionen.

   (d) *Auswahlregel*: Es werden im Versuch *alle* Faktorstufen $e_1^{(f)}, \ldots, e_{l_f}^{(f)}$ verwendet.

   (e) *Wiederholungsregel*: Bei einer eventuellen Versuchswiederholung würden genau dieselben Faktorstufen $e_1^{(f)}, \ldots, e_{l_f}^{(f)}$ verwendet werden.

   (f) *Verallgemeinerungsregel*: Die Aussagen, die aufgrund eines Versuchs mit festen Faktorstufen gemacht werden, gelten nur für die im Versuch verwendeten Faktorstufen $e_1^{(f)}, \ldots, e_{l_f}^{(f)}$.

   (g) *Vergleichsregel:* Aus Sicht des Anwenders ist es von Interesse, Werte von ϕ an verschiedenen Stellen zu vergleichen bzw. Differenzen von Funktionswerten zu betrachten, z.B. $\phi_{A^{(f)}}(e_1^{(f)}) - \phi_{A^{(f)}}(e_2^{(f)})$.

   (h) Beispiele für feste Faktoren sind fest definierte Behandlungen oder das Geschlecht von Versuchspersonen in der Medizin, Düngemethoden in der Landwirtschaft oder Maschinentypen bei der Produktionskontrolle.
   Feste Faktoren sind aber auch Größen, die zunächst eine stetige Natur besitzen (z.B. Temperatur, Druck), die aber im Versuch lediglich an einer festgelegten Anzahl von definierten Punkten betrachtet werden (z.B. 20°C, 30°C, 40°C).

2. Ein Faktor $a^{(r)}$ heißt *zufälliger Faktor*, wenn gilt:

   (a) Sei G_r die Randverteilung von $a^{(r)}$. G_r ist die stetige Gleichverteilung auf $E^{(r)} =]0, M]$, d.h. es gilt insbesondere $\forall x \in \{1, \ldots, M\} : P(\lceil a^{(r)} \rceil = x) = \frac{1}{M}$.
   Die Menge $\{1, \ldots, M\}$ wird dabei als Gesamtmenge der theoretisch verfügbaren Versuchseinheiten (subjects) interpretiert, aus der in einem statistischen Experiment eine zufällige Auswahl gezogen wird.
   Die Versuchseinheiten besitzen jedoch eine komplexe Struktur und befinden sich nicht in exakt reproduzierbaren Zuständen. Diese werden durch die Elemente eines stetigen Intervalls der Länge 1 charakterisiert, auf dem wiederum Gleichverteilung besteht (vgl. dazu die nachfolgende Skizze).

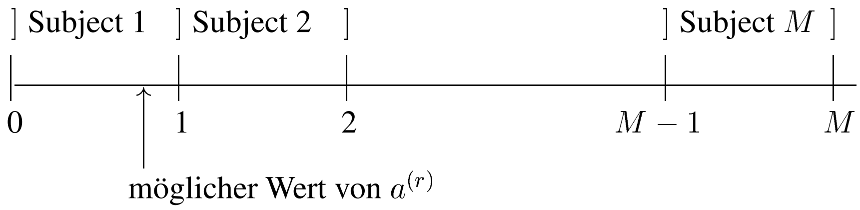

Abbildung 2: Veranschaulichung des Wertebereichs eines zufälligen Faktors.

   Es ist hier unabdingbar, für einen zufälligen Faktor $a^{(r)}$ zu fordern, dass G_r eine stetige Verteilung ist. Nur so ist gewährleistet, dass $\phi_{A^{(r)}}(a^{(r)})$ eine Normalverteilung besitzen kann. Dies ist ein wesentlicher Unterschied zu dem Sampling Theory-Ansatz von WILK & KEMPTHORNE (1955).

   (b)

$$\phi_{A^{(r)}} : E^{(r)} \to \chi_r = \mathbb{R} \,.$$

   Für die Zufallsvariablen $\phi_{A^{(r)}}(a_\iota^{(r)}),\ \iota \in I$, werden i.a. spezielle Verteilungsannahmen gemacht.

   (c) *Auswahlregel*: Die Auswahl aus der Gesamtmenge $\{1, \ldots, M\}$ ist im Versuch *zufällig*.

   (d) *Wiederholungsregel*: Bei einer eventuellen Versuchswiederholung würde erneut eine Stichprobe unter der Verteilung G_r gezogen.

   (e) *Verallgemeinerungsregel*: Die Aussagen, die aufgrund eines Versuchs mit zufälligen Faktorstufen gemacht werden, beziehen sich nicht nur auf die im Versuch verwendeten Versuchseinheiten, sondern auf die gesamte Menge $\{1, \ldots, M\}$.

   (f) *Vergleichsregel:* Aus Sicht des Anwenders macht es keinen Sinn, Differenzen von Funktionswerten von ϕ zu betrachten.

   (g) Zufällige Faktoren sind typischerweise die Versuchspersonen oder andere Versuchseinheiten (subjects) wie Blutproben bei medizinischen und Felder bei landwirtschaftlichen Untersuchungen.

3. Ein Faktor $a^{(k)}$ wird als *Kovariable (kovariater Faktor)* bezeichnet, wenn gilt:

   (a) $a^{(k)}$ besitzt eine im Allgemeinen nicht näher spezifizierte Verteilung G_k auf einer Teilmenge $E^{(k)}$ von $\mathbb{R}$.

   (b)
   $$\phi_{A^{(k)}} : E^{(k)} \to \chi_k = \mathbb{R} \, .$$

   Im Unterschied zu festem und zufälligem Faktor wird bei Kovariablen unterstellt, dass $\phi_{A^{(k)}}$ eine algebraisch formulierbare Form hat, im allgemeinen sogar Linearität, also $\phi_{A^{(k)}}(x_k) = \xi_0 + \xi_1 \cdot x_k$ mit unbekannten Koeffizienten ξ_o, ξ_1.
   Damit ist im Falle der Kovariablen die Klasse möglicher Funktionen zur Beschreibung der Einflüsse erheblich mehr eingeschränkt als bei festem oder zufälligem Faktor. Im (linearen) Beispiel ist $\dim(\Phi_{A^{(k)}}) = 2$.

   (c) *Wiederholungsregel*: Bei einer eventuellen Versuchswiederholung würde erneut zufällig aus $E^{(k)}$ gezogen.

   (d) *Verallgemeinerungsregel*: Die Aussagen die aufgrund eines Versuchs mit kovariaten Faktorstufen gemacht werden, beziehen sich auf die Grundgesamtheit $E^{(k)}$.

   (e) *Vergleichsregel:* Aus Sicht des Anwenders ist es von Interesse, die Funktionswerte von ϕ an verschiedenen Stellen zu vergleichen.

   (f) Beispiele für Kovariablen sind Größen, die neben der Zielvariablen an einer Versuchseinheit erhoben werden. Wenn bei einem medizinischen Versuch z.B. ein bestimmter Blutwert die Zielvariable ist, können andere Blutwerte oder skalierbare Befunde als Kovariable betrachtet werden.

Wir modellieren feste Faktoren als diskrete Zufallsvariablen und zufällige Faktoren als stetige Zufallsvariablen. Im Unterschied zu WILK & KEMPTHORNE (1955) lässt sich hier also nicht ohne Weiteres eine Zwischengruppe von zum Teil festen, zum Teil zufälligen Faktoren („semi-random“) definieren, die tatsächlich eine Mischung der beiden sehr unterschiedlichen Konzepte *fest* und *zufällig* darstellt.
Allerdings bedeutet dies nicht, dass das Problem des Auftretens von Faktoren, die sich zwar grundsätzlich wie feste Faktoren verhalten, aber deren Faktorstufen dennoch nur unvollständig betrachtet werden, nicht lösbar ist. Das oben definierte Konzept eines festen Faktors lässt sich noch um die Komponente *vollständige* oder *unvollständige* Betrachtung der Faktorstufen erweitern. Im Falle eines festen Faktors mit unvollständigen Faktorstufen lässt sich dann gerade auf die Methoden von WILK & KEMPTHORNE (1955) zurückgreifen. Der wesentliche Unterschied zwischen einem festen Faktor mit unvollständigen Faktorstufen und einem zufälligen Faktor besteht darin, dass bei ersterem keine Normalverteilungsannahme über die Faktoreffekte besteht.

Bei festen Faktoren und Kovariablen sind jeweils die Argumente $a^{(j)}$ beobachtbar, bei zufälligen Faktoren die ganzzahligen $\lceil a^{(r)} \rceil$, jedoch nicht die Funktionswerte $\phi_{A^{(j)}}(a^{(j)})$, $\tilde{\phi}_{A^{(j)}}(a^{(j)})$ oder $\rho_{A^{(j)}}(a^{(j)})$ bzw. die Funktionen $\psi_{A^{(j)}}(a^{(j)})$.

Das Ziel der statistischen Analyse ist nun jeweils eine Schätzung der Funktionen $\phi_{A^{(j)}} \in \Phi_{A^{(j)}}$ (bzw. $\tilde{\phi}_{A^{(j)}} \in \tilde{\Phi}_{A^{(j)}}$, $\rho_{A^{(j)}} \in \varrho_{A^{(j)}}$ oder $\psi_{A^{(j)}} \in \Psi_{A^{(j)}}$). Die Betrachtung wird jedoch im Allgemeinen eingeschränkt auf die Fragestellung, ob eine oder mehrere dieser Funktionen identisch verschwinden. Dies führt zu den in Abschnitt 2.4 formulierten Hypothesen.

Um einige Modelle genauer abgrenzen zu können, betrachten wir die folgenden Annahmen:

- $\phi_{A^{(1)},\varepsilon} \equiv \ldots \equiv \phi_{A^{(m)},\varepsilon} \equiv \ldots \equiv \phi_{A^{(1)},\ldots,A^{(m)},\varepsilon} \equiv 0$ bzw.
  $\tilde{\phi}_{A^{(1)},\varepsilon} \equiv \ldots \equiv \tilde{\phi}_{A^{(m)},\varepsilon} \equiv \ldots \equiv \tilde{\phi}_{A^{(1)},\ldots,A^{(m)},\varepsilon} \equiv 0.$ (A1)
  Diese Annahme ist wesentlich für das Verständnis und die Gültigkeit Linearer Modelle. Der Satz über die Hoeffding-Zerlegung sagt zwar aus, dass die dort angegebene additive Zerlegung unter den aufgeführten Bedingungen immer möglich ist. Entscheidend für die der Varianzanalyse zugundeliegenden Linearen Modelle ist jedoch die additive Isolierung des Fehlerterms bzw. die fehlende Wechselwirkung zwischen Fehlerterm und Faktoren. Satz 2.12 auf S.28 ist in diesem Zusammenhang direkt anwendbar, falls untersucht werden soll, ob ein gegebenes einfaktorielles Modell in ein solches transformiert werden kann, in dem keine Wechselwirkung zwischen Faktor und Fehlerterm existiert.

- Die Familie $(\phi_{A^{(1)}}(a^{(1)}), \ldots, \phi_{A^{(m)}}(a^{(m)}), \phi_{A^{(1)},A^{(2)}}(a^{(1)}, a^{(2)}), \ldots, \phi_{A^{(m-1)},A^{(m)}}(a^{(m-1)}a^{(m)}), \ldots, \phi_{A^{(1)},\ldots,A^{(m)}}(a^{(1)}, \ldots, a^{(m)}))$ von Zufallsvariablen ist unabhängig (bzw. die entsprechend mit $\tilde{\phi}$ definierte Familie). Aus der Hoeffding-Zerlegung folgt bereits die Unkorreliertheit der angegebenen Familien von Zufallsvariablen. (A2)

- $\phi_\varepsilon(\varepsilon)$ ist unabhängig von der Familie $(\phi_{A^{(1)}}(a^{(1)}), \ldots, \phi_{A^{(m)}}(a^{(m)}), \phi_{A^{(1)},A^{(2)}}(a^{(1)}, a^{(2)}), \ldots, \phi_{A^{(m-1)},A^{(m)}}(a^{(m-1)}a^{(m)}), \ldots, \phi_{A^{(1)},\ldots,A^{(m)}}(a^{(1)}, \ldots, a^{(m)}))$ (bzw. analog für $\tilde{\phi}_\varepsilon(\varepsilon)$). (A2a)

- $\phi_\varepsilon(\varepsilon) \sim \mathrm{N}(0, \sigma^2_\varepsilon)$ bzw. $\tilde{\phi}_\varepsilon(\varepsilon) \sim \mathrm{N}(0, \sigma^2_\varepsilon)$. (A3)

- Eine Abschwächung von Annahme (A3) ist es, nicht mehr von gleichen Varianzen σ^2_ε auszugehen, sondern lediglich $\phi_\varepsilon(\varepsilon) \sim \mathrm{N}(0, \sigma^2_\varepsilon(a^{(1)}, \ldots, a^{(m)}))$ zu verlangen. D.h., die Varianz des Fehlerterms kann von der Faktorstufenkombination abhängen.
  Das bedeutet insbesondere, dass Annahme (A2a) nicht erfüllt sein muss. (A3a)

- Bisweilen wird die Annahme (A3) auch abgeschwächt zu $\phi_\varepsilon(\varepsilon) \sim G$ oder $\tilde{\phi}_\varepsilon(\varepsilon) \sim G$, wobei G eine symmetrische Verteilung mit Erwartungswert 0 ist. (A3b)

- Sei $\mathcal{M} := \{a^{(1)}, \ldots, a^{(m)}\}$ die Menge aller Faktoren.
  Außerdem seien $\mathcal{M}_f := \{a^{(f_1)}, \ldots, a^{(f_{m_f})}\} \subset \mathcal{M}$, $\mathcal{M}_r := \{a^{(r_1)}, \ldots, a^{(r_{m_r})}\} \subset \mathcal{M}$ und $\mathcal{M}_k := \{a^{(k_1)}, \ldots, a^{(k_{m_k})}\} \subset \mathcal{M}$ die disjunkten Mengen der festen, zufälligen und kovariaten Faktoren, mit $\mathcal{M}_f \cup \mathcal{M}_r \cup \mathcal{M}_k = \mathcal{M}$.
  Dann gilt weiter:

  - $\forall a^{(j)} \in \mathcal{M}_r : \phi_{A^{(j)}}(a^{(j)}) \sim \mathrm{N}(0, \sigma^2_{A^{(j)}})$, bzw. noch allgemeiner

– $\forall \{i_1, \dots, i_k\} \subset \{1, \dots, m\}$ mit $\{a^{(i_1)}, \dots, a^{(i_k)}\} \cap \mathcal{M}_r \neq \emptyset$:
$\phi_{A^{(i_1)},\dots,A^{(i_k)}}(a^{(i_1)}, \dots, a^{(i_k)}) \sim \mathrm{N}(0, \sigma^2_{A^{(i_1)},\dots,A^{(i_k)}})$. (A4)

– $\forall a^{(k)} \in \mathcal{M}_k : \phi_{A^{(k)}}(a^{(k)}) = \xi_0^{(k)} + \xi_1^{(k)} \cdot a^{(k)}$ (Linearität) . (A5)

Definition 2.13 (Modelle)

1. *Ein statistisches Modell, bei dem ϕ zerlegt wird und in dem die Annahmen (A1), (A2), (A2a), (A3), (A4) und (A5) gelten, heißt* (homoskedastisches) Lineares Modell (Linear Model).

2. *Gilt statt (A3) und (A2a) die relaxierte Annahme (A3a), so spricht man von einem* heteroskedastischen Linearen Modell.

3. *Ein statistisches Modell, bei dem ϕ oder $\tilde{\phi}$ zerlegt wird und in dem die Annahmen (A1), (A2) und (A3b) gelten, heißt* Lineares Modell ohne Normalverteilungsannahme.

4. *Ein statistisches Modell, bei dem ψ zerlegt wird, heißt* allgemeines Nichtparametrisches Modell.

5. *Ein statistisches Modell, bei dem ρ zerlegt wird und in dem eine zu (A2) analoge Annahme für ρ gilt, heißt* verallgemeinertes Lineares Modell (Generalized Linear Model, GLiM).

Anmerkung: Die Aussagen (R1) bis (R4) (s. S.25) ergeben sich in dieser Arbeit als einfache Folgerungen aus dem Satz über die Hoeffding-Zerlegung. Da der gewöhnlich in Lehrbüchern gewählte Zugang zu Linearen Modellen anders ist, müssen die Reparametrisierungsbedingungen dort als Modellbedingungen formuliert werden.
Dabei taucht häufig das Problem auf, dass zum einen die hier als $\phi_{A^{(r)}}(a^{(r)})$ bezeichneten Zufallsvariablen, die den Effekt eines zufälligen Faktors beschreiben, als unabhängig angenommen werden, sie sich aber andererseits zu Null summieren sollen. Wenn sie sich zu Null summieren, können sie natürlich nicht unabhängig sein, und dieser Widerspruch wird denn auch, z.B. bei NELDER (1977,1994), scharf kritisiert.
Im von uns formulierten Modell besteht dieser Widerspruch nicht. Die Reparametrisierungsbedingung (R2), die aus Satz 2.7 folgt, lautet für einen zufälligen Faktor $a^{(r)}$:

$$\mathrm{E}_{A^{(r)}}(\phi_{A^{(r)}}(a^{(r)})) = \int \phi_{A^{(r)}}(x_r) dG_r(x_r) = 0 \qquad \text{(vgl. S.25)} .$$

Es wird also nicht von den im Experiment (zufällig) ausgewählten Ausprägungen des zufälligen Faktors gefordert, dass sie sich zu Null summieren - und trotzdem unabhängig sind, sondern lediglich ihr Erwartungswert ist Null, und dies ist für zentral normalverteilte Zufallsvariablen auch sinnvoll.

Ein ähnliches Problem taucht bei der Wechselwirkung zwischen einem zufälligen und einem festen Faktor auf (vgl. wiederum NELDER (1977,1994) sowie MCLEAN, SANDERS & STROUP (1991)). Die gewöhnlich in der Literatur angegebene Reparametrisierungsbedingung bedeutet für die (bei uns als $\phi_{A^{(f)},A^{(r)}}(a^{(f)}, a^{(r)})$ bezeichnete) Wechselwirkung zwischen einem festen

Faktor $a^{(f)}$ und einem zufälligen Faktor $a^{(r)}$, dass sie sich sowohl bei fixiertem festen Faktor und Summation über die Stufen des zufälligen Faktors als auch bei fixiertem zufälligem Faktor und Summation über die Stufen des festen Faktors zu Null summiert. Ersteres lautet in unserem Modell – wiederum als Folgerung aus Satz 2.7 (vgl. (R3) auf S.25) –

$$\mathrm{E}_{A^{(r)}}(\phi_{A^{(f)},A^{(r)}}(a^{(f)}, a^{(r)})|a^{(f)} = x_f) = \int \phi_{A^{(f)},A^{(r)}}(x_f, x_r) dG_r(x_r) = 0 \,,$$

und bedeutet, analog zur gerade erwähnten Reparametrisierungsbedingung (R2) für einen zufälligen Faktor, keine Kollision mit der Unabhängigkeitsforderung.
Die andere Reparametrisierungsbedingung,

$$\mathrm{E}_{A^{(f)}}(\phi_{A^{(f)},A^{(r)}}(a^{(f)}, a^{(r)})|(a^{(r)} = x_r) = 0 \,,$$

bedeutet ohnehin eine Summation diskreter Zufallsvariablen, da $\phi_{A^{(f)},A^{(r)}} : E^{(f)} \times E^{(r)} \longrightarrow \mathbb{R}$, mit stetigem Wahrscheinlichkeitsraum auf $E^{(r)}$ und diskretem Wahrscheinlichkeitsraum auf $E^{(f)}$. Die Zufallsvariable $a^{(f)} \mapsto \phi_{A^{(f)},A^{(r)}}(a^{(f)}, x_r)$ mit festem $x_r \in E^{(r)}$ kann also nicht mehr normalverteilt sein. Damit ist auch eine Diskussion darüber, inwiefern unabhängige, normalverteilte Zufallsvariablen sich zu Null summieren können, in diesem Fall obsolet.

2.4 Transformationen

In der Anmerkung auf S.28 in Kapitel 2.2 haben wir bereits die Problematik angesprochen, die besteht, wenn Teilinformationen über die Gestalt von ϕ vorliegen, die eine additive Zerlegung nicht sinnvoll erscheinen lassen. Derartige Informationen können z.B. aus theoretischen Überlegungen, durch vorangegangene Studien oder durch Erfahrungen mit ähnlichen Datensätzen kommen. In einem solchen Fall kann es dann angemessen sein, die Informationen zu nutzen, um die Daten in einer Weise zu transformieren, die eine additive Zerlegung der transformierten Daten plausibler macht. Darüber hinaus bestehen gerade im (homoskedastischen) Linearen Modell zahlreiche Annahmen, deren Gültigkeit Voraussetzung für die theoretische Herleitung der entsprechenden Teststatistiken ist, und deren Verletzung bedeutet, dass bestimmte Testprozeduren zu verfälschten Ergebnissen führen. Die wesentlichen restriktiven Annahmen sind dabei die Homoskedastizität (gleiche Varianz des Fehlerterms in allen Faktorstufenkombinationen) und die Normalverteilung des Fehlerterms, somit also insbesondere Symmetrie dieser Verteilung.
SCHEFFÉ (1959, S. 331ff) weist darauf hin, dass eine Abweichung von der Form der Normalverteilung nur eine geringe Auswirkung auf statistische Inferenz über Mittelwerte hat. Heteroskedastizität ist bei gleichen Zellbesetzungen nicht von großer Bedeutung, verursacht aber Probleme im Fall ungleicher Zellbesetzungen.
Auch ARNOLD (1981, S.58ff) diskutiert diese Annahmen. Er bezeichnet die Voraussetzung gleicher Varianzen als besonders störend. Dagegen misst er einer Verletzung der Normalverteilungsannahme keine allzu große Bedeutung bei. Ein Grund dafür ist die asymptotische Validität varianzanalytischer Verfahren auch ohne die Voraussetzung normalverteilter Fehler.
Zahlreiche Arbeiten wurden veröffentlicht, die sich mit der Frage beschäftigen, wie vorliegende Daten am geschicktesten zu transformieren sind, um eine Gültigkeit der genannten Annahmen für die transformierten Daten zu erreichen. Systematische Untersuchungen bestimmter Methoden zur Transformation finden sich schon u.a. bei BARTLETT (1936, 1947) und CURTISS (1943). Zum Klassiker geworden ist allerdings die oft zitierte Arbeit von BOX & COX (1964), in der eine parametrisierte Familie von Transformationen vorgestellt wird, für die sich der Name *Box-Cox-Transformationen* eingebürgert hat. Eine Zusammenfassung wesentlicher Ideen zur Transformation, insbesondere für den Fall der Regression, findet man in CARROL & RUPPERT (1988).
Es scheint aber, dass mit dem Aufkommen robuster und nichtparametrischer Verfahren, deren Theorie in den letzten Jahren große Fortschritte gemacht hat, die Analyse von Transformationen nicht mehr als so wesentlich angesehen wird und daher aus dem Hauptaugenmerk der Forschung verdrängt ist. Dennoch sind nachfolgend kurz einige der klassischen Methoden zur Transformation zusammengestellt.

1. Erzeugung von Homoskedastizität
   Statt Annahme (A3) sei lediglich Annahme (A3a) erfüllt (d.h. unterschiedliche Varianzen des Fehlerterms für verschiedene Faktorstufenkombinationen, vgl. S.39).
   Die Varianzen Var$(Y|a^{(1)} = x_1, \dots , a^{(m)} = x_m) =: \sigma^2_{x_1,\dots,x_m}$ seien dabei als Funktionen der Erwartungswerte E$(Y|a^{(1)} = x_1, \dots , a^{(m)} = x_m) =: \mu_{x_1,\dots,x_m}$ darstellbar, d.h.

$$\exists g : \mathbb{R} \to \mathbb{R}^+ \; \forall (x_1, \dots , x_m) \in E^{(1)} \times \dots \times E^{(m)} : \sigma^2_{x_1,\dots,x_m} = g(\mu_{x_1,\dots,x_m}) \cdot \sigma^2$$

mit festem σ^2.
Ziel ist es nun, eine Transformation der Zielvariable Y zu finden, die Homoskedastizität erzeugt, d.h. zu finden ist eine Funktion q, so dass Var($q \circ Y | a^{(1)} = x_1, \ldots, a^{(m)} = x_m) =: \sigma^2_{x_1,\ldots,x_m}(q)$ nicht mehr abhängig von $x_1, \ldots, x_m$ ist. Mittels Taylor-Entwicklung von $\sigma^2_{x_1,\ldots,x_m}(q)$ lässt sich folgendes zeigen (vgl. CURTISS (1943), BARTLETT (1947)):

$$\sigma^2_{x_1,\ldots,x_m}(q) \approx q'(\mu_{x_1,\ldots,x_m})^2 \cdot (g(\mu_{x_1,\ldots,x_m}))^2 \cdot \sigma^2 \, .$$

Näherungsweise Homoskedastizität lässt sich also durch Wahl von q als unbestimmtes Integral von $1/g(y)$ erreichen. Natürlich ist die Approximation nur „gut" in der Nähe der Stelle, um die entwickelt wird. Schon TIPPETT (1934) kommentiert dazu: „This derivation is not mathematically sound, and the result is only justified if on application it is found to be satisfactory."
Im Spezialfall eines linearen Zusammenhanges zwischen Erwartungswert und Varianz ergeben sich gleiche Varianzen in den verschiedenen Faktorstufenkombinationen durch eine lineare Transformation.

2. Erzeugung von Symmetrie
Wünschenswert, gemäß den Annahmen (A3), (A3a) und (A3b) (vgl. S.39), ist im Linearen Modell stets eine symmetrische Verteilung des Fehlerterms und somit eine symmetrische Verteilung von $Y | a^{(1)} = x_1, \ldots, a^{(m)} = x_m$. Ist die Verteilung von Y rechtsschief (Schiefe > 0), so wird die Schiefe durch eine konkave, strikt monotone Transformation reduziert, eine konvexe Transformation reduziert Linksschiefe.

3. Box-Cox-Transformationen
BOX & COX (1964) betrachten (u.a.) eine parametrisierte Familie von Transformationen

$$q_\lambda : \; y \mapsto q_\lambda(y) \quad \text{mit} \quad q_\lambda(y) = \begin{cases} \frac{y^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \log(y), & \lambda = 0 \end{cases} \, .$$

Dies entspricht den Transformationen, die man gemäß 1 (Erzeugung von Homoskedastizität) für $g(\mu) = \mu^{1-\lambda}$ erhält. $\lambda = 1$ bedeutet eine lediglich lineare Transformation, die die Hypothesen im Linearen Modell nicht verändert (siehe Abschnitt 2.5), $\lambda = 0$ entspricht dem verbreiteten Logarithmieren von Daten, $\lambda = 1/2$ dem Betrachten der Quadratwurzeln. Für $\lambda > 1$ erhält man konvexe, für $\lambda < 1$ konkave Transformationen. Somit ist es, gemäß 2 (Erzeugung von Symmetrie), innerhalb der betrachteten Familie von Transformationen sinnvoll, rechtsschiefe Daten logarithmisch zu transformieren.
Eine in gewisser Weise optimale Box-Cox-Transformation wird im Allgemeinen über einen Maximum-Likelihood-Ansatz bestimmt. Der Maximum-Likelihood-Schätzer $\hat{\lambda}$ für λ kann dabei über ein rekursives Verfahren berechnet werden. Für praktische Zwecke werden häufig zu $\hat{\lambda}$ benachbarte „runde" Werte für die Transformation verwendet (z.B. $-1, -\frac{1}{2}, 0, \frac{1}{2}, 1, 2$).
ANDREWS (1971) hat darüber hinaus einen Test vorgestellt, mit dem auf einen vorgegebenen Parameter $\lambda = \lambda_0$ getestet werden kann.
Nicht ganz unproblematisch ist es, die transformierten Daten so zu analysieren, als wenn

der geschätzte Parameter $\hat{\lambda}$ nicht eine Zufallsvariable, sondern wahr wäre. Wird die Variabilität von $\hat{\lambda}$ vernachlässigt, so führt dies dazu, dass eine zu hohe Präzision der Auswertung im bezüglich $\hat{\lambda}$ transformierten Modell unterstellt wird. Der dadurch entstehende Fehler hält sich allerdings in maßvollen Grenzen, wie CARROLL & RUPPERT (1984) gezeigt haben (s. auch SEBER & WILD (1988, S.74f)).

4. Transformationen von Proportionen
   Für Zielgrößen, die Anteile darstellen und somit Werte im Bereich $]0,1[$ annehmen, wird häufig die logit-Transformation $q(y) = \log(\frac{y}{1-y})$ gewählt. Eine allgemeinere, parametrisierte Klasse von Transformationen erhält man durch

$$q_\lambda(y) = \begin{cases} y^\lambda - (1-y)^\lambda, & \lambda \neq 0 \\ \log(\frac{y}{1-y}), & \lambda = 0 \end{cases} \quad .$$

2.5 Hypothesen

In diesem Abschnitt vergleichen wir Hypothesen verschiedener Modelle. Zu Beginn von Kapitel 2 wurden unterschiedliche Modellgleichungen aufgestellt, die zwar alle auf der allgemeinen Modellgleichung (M) (s. S.17) als grundlegendem Konzept beruhen, aber bei denen verschiedenen Dinge modelliert werden (die Zielgröße; eine transformierte Zielgröße; ihre Verteilungsfunktion; ein Parameter, der die Verteilung der Zielgröße determiniert; Verteilungsquantile). Es ist naheliegend, zu fragen, inwieweit diese verschiedenen Modelle und die mit ihnen verbundenen jeweiligen Zerlegungen sowie die auf den Zerlegungen basierenden Hypothesen vergleichbar sind. Wünschenswert wären Aussagen darüber, ob – unter der simultanen Annahme zweier verschiedener Modelle – die Gültigkeit bestimmter Hypothesen im einen Modell impliziert, dass entsprechende Hypothesen auch im anderen Modell gelten.
In der Tat lassen sich derartige Aussagen beweisen. Wir haben im Folgenden die Relationen zwischen den Hypothesen der verschiedenen Modelle möglichst umfassend dargestellt, d.h. gültige Implikationen bewiesen und für nicht gültige Implikationen Gegenbeispiele angegeben. Eine Zusammenfassung der Resultate ist in Abbildung 3 auf S.60 gegeben.
Einige der nachfolgenden Aussagen lassen sich mit elementaren Methoden beweisen. Für die Analyse der Relation zwischen den Zerlegungen, die auf Modellgleichung (M1) (direkte Modellierung der Zielgröße) und (M2) (Modellierung der Verteilungsfunkton der Zielgröße) beruhen, war zunächst der Beweis einer allgemeinen maßtheoretischen Aussage (Satz A.6) nötig. Für Faktoren mit diskreter Verteilung ist diese Aussage trivial, aber im Sinne einer möglichst allgemeingültigen Darstellung, die z.B. auch zufällige Faktoren einschließt, haben wir die Aussage für Faktoren mit beliebigen Verteilungen formuliert und bewiesen.

Aussagen über Implikationen zwischen in verschiedenen Modellen formulierten Hypothesen haben Auswirkungen auf ein logisch konsistentes Vorgehen bei Testentscheidungen. Falls die Gültigkeit einer bestimmten Hypothese in Modell M_1 die Gültigkeit einer entsprechenden Hypothese in Modell M_2 impliziert – formal sei dies ausgedrückt als $H_0^{M_1}(x) \Longrightarrow H_0^{M_2}(x)$, wir werden dann $H_0^{M_1}(x)$ als *feiner* bzw. $H_0^{M_2}(x)$ als *gröber* bezeichnen – so bedeutet die Testentscheidung „Ablehnung von $H_0^{M_2}(x)$“, dass der Tester sich bei formallogischem Vorgehen auch für „Ablehnung von $H_0^{M_1}(x)$“ entscheiden muss. Ein expliziter Test auf die Gültigkeit von $H_0^{M_1}(x)$ braucht dann nicht mehr durchgeführt zu werden. Wir nehmen beispielsweise für das Anwendungsbeispiel, das auf den Seiten 20 und 66 genauer erläutert wird, an, dass der Versuch in geeigneter Weise durch ein Lineares Modell gemäß (M1), (A1) und (A2) (vgl. S.18 und S.39), aber auch durch ein Nichtparametrisches Modell (vgl. S.40) beschrieben werden kann und formulieren die Hypothese, die Wahl des Düngemittels habe keinen Einfluss auf den Tomatenertrag – mathematisch beschrieben als $H_0^{\phi}(a)$ im Linearen Modell und als $H_0^{\psi}(a)$ im Nichtparametrischen Modell. Lemma 2.20 liefert uns dann das Resultat $H_0^{\psi}(a) \Longrightarrow H_0^{\phi}(a)$, mit Korollar 2.23 erhalten wir sogar $H_0^{\psi}(a) \Longleftrightarrow H_0^{\phi}(a)$, d.h. beide Hypothesen sind äquivalent bzw. sind gleich fein. Damit ist es ausreichend, *entweder* einen t-Test für $H_0^{\phi}(a)$ *oder* einen Wilcoxon-Mann-Whitney-Test für $H_0^{\psi}(a)$ durchzuführen. Die Testentscheidung für einen der Tests bedingt automatisch die für den anderen Test.
Natürlich birgt dieses Vorgehen die Gefahr, dass der Tester sich unter mehreren möglichen Tests denjenigen heraussucht, dessen Ergebnis am nächsten am gewünschten Resultat liegt. Ein vernünftiges und verantwortungsvolles Vorgehen – unter Berücksichtigung der angesproche-

nen sachlogischen Überlegungen – sollte darin bestehen, sich unter mehreren geeigneten Tests (also solchen, die unter den gegebenen Bedingungen das Niveau einhalten bzw. deren Modellvoraussetzungen erfüllt sind, was in jedem Fall einer Überprüfung bedarf) wenn möglich (!) denjenigen auszusuchen, der die größtmögliche Macht hat, der also Alternativen mit maximaler Wahrscheinlichkeit aufdeckt. Die Entscheidung, die dieser Test liefert, gilt dann für die entsprechende Hypothese in *allen* Modellen, deren Voraussetzungen gültig sind.

Die Tatsache, dass sich in das in dieser Arbeit formulierte allgemeine Konzept sowohl gewöhnliche Lineare Modelle, als auch verallgemeinerte Lineare Modelle und Nichtparametrische Methoden einbetten lassen, lässt erwarten, dass die in den letzten Jahren gefundenen Relationen zwischen den Hypothesen der drei genannten Modelle hier bestätigt werden. BRUNNER & PURI (1996) haben Implikationen zwischen den Hypothesen des gewöhnlichen Linearen Modells und des Nichtparametrischen Modells bewiesen, während sich OELERICH (1998) und SCHWARZ (1999) mit entsprechenden Zusammenhängen zwischen verallgemeinerten Linearen Modellen und dem Nichtparametrischen Modell beschäftigt haben. Dabei hat OELERICH als verallgemeinertes Lineares Modell das sogenannte Proportional Odds Modell untersucht, SCHWARZ dagegen das logistische Regressionsmodell. Die von den genannten Autoren erzielten wesentlichen Resultate, die auch mittels der in dieser Arbeit betrachteten Hypothesen dargestellt werden, werden hier nicht nur bestätigt, sondern lassen sich zum Teil in noch allgemeinerer Form formulieren. Insbesondere leiten wir hier alle Ergebnisse für Modelle mit beliebig vielen Faktoren her.

Wir geben nun einen kurzen Überblick über die von uns betrachteten Hypothesen. Generell sind statistische Hypothesen Aussagen über unbekannte Parameter oder Funktionen, die im Allgemeinen gut interpretierbar sind. Anhand von Beobachtungen, also Realisierungen von Zufallsvariablen, sollen Entscheidungen darüber getroffen werden, ob eine Hypothese plausibel ist oder nicht. Dazu werden statistische Testfunktionen verwendet.

Mit Hilfe der durch die Zerlegungen definierten Funktionen formulieren wir die folgenden Hypothesen:

Hypothesen über die (Haupteffekte der) Faktoren:
$H_0^{\phi}(a^{(i)}) : \phi_{A^{(i)}} \equiv 0$
(bzw. $H_0^{\tilde{\phi}}(a^{(i)}) : \tilde{\phi}_{A^{(i)}} \equiv 0$, $H_0^{\rho}(a^{(i)}) : \rho_{A^{(i)}} \equiv 0$, $H_0^{\tau^\alpha}(a^{(i)}) : \tau^\alpha_{A^{(i)}} \equiv 0$
oder $H_0^{\psi}(a^{(i)}) : \psi_{A^{(i)}} \equiv 0$).

Hypothesen über die bedingten Effekte der Faktoren:
$H_0^{\phi}(a^{(i)}|a^{(j)}) : \phi_{A^{(i)}} + \phi_{A^{(i)},A^{(j)}} \equiv 0$
(bzw. $H_0^{\tilde{\phi}}(a^{(i)}|a^{(j)}) : \tilde{\phi}_{A^{(i)}} + \tilde{\phi}_{A^{(i)},A^{(j)}} \equiv 0$, $H_0^{\rho}(a^{(i)}|a^{(j)}) : \rho_{A^{(i)}} + \rho_{A^{(i)},A^{(j)}} \equiv 0$,
$H_0^{\tau^\alpha}(a^{(i)}|a^{(j)}) : {\tau^\alpha}_{A^{(i)}} + \rho_{A^{(i)},A^{(j)}} \equiv 0$ oder $H_0^{\psi}(a^{(i)}|a^{(j)}) : \psi_{A^{(i)}} + \psi_{A^{(i)},A^{(j)}} \equiv 0$).

Hypothesen über die Wechselwirkungen:
$H_0^{\phi}(a^{(i_1)} \dots a^{(i_k)}) : \phi_{A^{(i_1)},\dots,A^{(i_k)}} \equiv 0$
(bzw. $H_0^{\tilde{\phi}}(a^{(i_1)} \dots a^{(i_k)}) : \tilde{\phi}_{A^{(i_1)},\dots,A^{(i_k)}} \equiv 0$, $H_0^{\rho}(a^{(i_1)} \dots a^{(i_k)}) : \rho_{A^{(i_1)},\dots,A^{(i_k)}} \equiv 0$,
$H_0^{\tau^\alpha}(a^{(i_1)} \dots a^{(i_k)}) : {\tau^\alpha}_{A^{(i_1)},\dots,A^{(i_k)}} \equiv 0$ oder $H_0^{\psi}(a^{(i_1)} \dots a^{(i_k)}) : \psi_{A^{(i_1)},\dots,A^{(i_k)}} \equiv 0$).

Allgemeinere Hypothesen über zufällige Faktoren:
In einem Modell, in dem die Annahmen (A3) und (A4) erfüllt sind, lassen sich folgende allge-

meinere Hypothesen über die zufälligen Faktoren formulieren (s. SCHEFFÉ (1959, S.227)):
$H_0^{\phi}(\sigma_{A^{(i)}}) : \sigma^2_{A^{(i)}} \leq \theta_0 \cdot \sigma^2_{\varepsilon}$ mit fest vorgegebenem $\theta_0 \geq 0$.
Offenbar entspricht $H_0^{\phi}(a^{(i)})$ dem Spezialfall $\theta_0 = 0$.
Wir werden auf diese Form der Hypothesenbildung allerdings nicht weiter eingehen.

Um zu analysieren, ob die jeweiligen Hypothesen sich je nach Wahl des Modells, also insbesondere je nach Zerlegung von ϕ, $\tilde{\phi}$, ψ, τ oder ρ, unterscheiden, betrachten wir ein statistisches Modell, in dem die Annahmen (A1) und (A2) (s. S.39) gelten und führen gleichzeitig Zerlegungen für die Funktionen ϕ, $\tilde{\phi}$, ψ, τ und ρ durch.

Für die Zerlegungen und die hier formulierten und bewiesenen Aussagen über Hypothesen in den verschiedenen Modellen ist es nicht wesentlich, ob die Faktoren $a^{(1)}, \ldots, a^{(m)}$ fest oder zufällig sind.
Es gelte also:

$$\begin{aligned}
Y &= \phi(a^{(1)}, \ldots, a^{(m)}, \varepsilon) \\
&= \phi_\mu + \sum_{k=1}^{m} \sum_{\{i_1,\ldots,i_k\}\subset\{1,\ldots,m\}} \phi_{A^{(i_1)},\ldots,A^{(i_k)}}(a^{(i_1)}, \ldots, a^{(i_k)}) + \phi_\varepsilon(\varepsilon)\ , \\
Z &= q \circ \phi(a^{(1)}, \ldots, a^{(m)}, \varepsilon) \\
&= \tilde{\phi}_\mu + \sum_{k=1}^{m} \sum_{\{i_1,\ldots,i_k\}\subset\{1,\ldots,m\}} \tilde{\phi}_{A^{(i_1)},\ldots,A^{(i_k)}}(a^{(i_1)}, \ldots, a^{(i_k)}) + \tilde{\phi}_\varepsilon(\varepsilon) \\
&\text{und } Y \sim F \text{ mit} \\
F(t) &= \psi(a^{(1)}, \ldots, a^{(m)})(t) \\
&= \psi_\mu(t) + \sum_{k=1}^{m} \sum_{\{i_1,\ldots,i_k\}\subset\{1,\ldots,m\}} \psi_{A^{(i_1)},\ldots,A^{(i_k)}}(a^{(i_1)}, \ldots, a^{(i_k)})(t) \quad \text{sowie} \\
Q(\alpha) &= \tau^\alpha(a^{(1)}, \ldots, a^{(m)}) \\
&= \tau^\alpha_\mu + \sum_{k=1}^{m} \sum_{\{i_1,\ldots,i_k\}\subset\{1,\ldots,m\}} \tau^\alpha_{A^{(i_1)},\ldots,A^{(i_k)}}(a^{(i_1)}, \ldots, a^{(i_k)}) \quad (\alpha \in]0,1[)\ .
\end{aligned}$$

Außerdem entstamme $F = F_p$ einer parametrisierten Familie $\{F_p, p \in \mathbf{P}\}$ von Verteilungen, und es gelte

$$\begin{aligned}
g(p) &= \rho(a^{(1)}, \ldots, a^{(m)}) \\
&= \rho_\mu + \sum_{k=1}^{m} \sum_{\{i_1,\ldots,i_k\}\subset\{1,\ldots,m\}} \rho_{A^{(i_1)},\ldots,A^{(i_k)}}(a^{(i_1)}, \ldots, a^{(i_k)})\ .
\end{aligned}$$

Zunächst beweisen wir eine allgemeine Aussage über den Zusammenhang zwischen der Zerlegung von ϕ und der von ψ.

Satz 2.14 *Im zu Beginn dieses Abschnittes definierten Modell gelten folgende Beziehungen zwischen den Zerlegungen von ϕ und ψ:*

a) $\phi_\mu = \int t \, d\psi_\mu(t)$.

b) $\forall \{\nu_1, \dots, \nu_k\} \subset \{1, \dots, m\}$:
$\phi_{A^{(\nu_1)},\dots,A^{(\nu_k)}}(x_{\nu_1}, \dots, x_{\nu_k}) = \int t \, d\psi_{A^{(\nu_1)},\dots,A^{(\nu_k)}}(x_{\nu_1}, \dots, x_{\nu_k})(t)$.

Beweis:

a)
$$\begin{aligned}
& \int t \, d\psi_\mu(t) \\
= & \int t \, d\left[\mathrm{E}(\psi(a^{(1)} \dots a^{(m)})(t))\right] \\
= & \int t \, d\left[\int \dots \int \psi(x_1 \dots x_m)(t) \, dG_1(x_1) \dots dG_m(x_m)\right] \\
= & \int \dots \int \int t \, d\left[\psi(x_1 \dots x_m)(t)\right] \, dG_1(x_1) \dots dG_m(x_m) \quad \text{(mit Satz A.6)} \\
= & \int \dots \int \int t \, dF^{Y|a^{(1)}=x_1,\dots,a^{(m)}=x_m}(t) \, dG_1(x_1) \dots dG_m(x_m) \\
= & \int \dots \int \mathrm{E}(Y|a^{(1)} = x_1, \dots, a^{(m)} = x_m) \, dG_1(x_1) \dots dG_m(x_m) \\
= & \ \mathrm{E}(Y) \ = \ \phi_\mu \, .
\end{aligned}$$

b) Für $\{\nu_1, \dots, \nu_k\} \subset \{1, \dots, m\}$ gilt mit Korollar 2.8, angewandt auf die Zerlegung von ϕ:

$$\mathrm{E}(Y|a^{(\nu_1)} = x_{\nu_1}, \dots, a^{(\nu_k)} = x_{\nu_k}) = \phi_\mu + \sum_{\{i_1,\dots,i_j\}\subset\{\nu_1,\dots,\nu_k\}} \phi_{A^{(i_1)},\dots,A^{(i_j)}}(x_{i_1}, \dots, x_{i_j}) \, .$$

Andererseits ist

$$\begin{aligned}
& \mathrm{E}(Y|a^{(\nu_1)} = x_{\nu_1}, \dots, a^{(\nu_k)} = x_{\nu_k}) \\
= & \int t \, dF^{Y|a^{(\nu_1)}=x_{\nu_1},\dots,a^{(\nu_k)}=x_{\nu_k}}(t) \\
= & \int t \, d\left[\mathrm{E}(\psi(a^{(1)} \dots a^{(m)})(t)|a^{(\nu_1)} = x_{\nu_1}, \dots, a^{(\nu_k)} = x_{\nu_k})\right] \\
= & \int t \, d\Big[\psi_\mu(t) + \sum_{\{i_1,\dots,i_j\}\subset\{\nu_1,\dots,\nu_k\}} \psi_{A^{(i_1)},\dots,A^{(i_j)}}(x_{i_1}, \dots, x_{i_j})(t)\Big] \, .
\end{aligned}$$

Für die letzte Gleichheit wird Korollar 2.8 angewandt auf die Zerlegung von ψ.
Insgesamt folgt

$$\begin{aligned}
&\phi_\mu + \sum_{\{i_1,\dots,i_j\}\subset\{\nu_1,\dots,\nu_k\}} \phi_{A^{(i_1)},\dots,A^{(i_j)}}(x_{(i_1)},\dots,x_{(i_j)}) \\
&= \int t\, d\Big[\psi_\mu(t) + \sum_{\{i_1,\dots,i_j\}\subset\{\nu_1,\dots,\nu_k\}} \psi_{A^{(i_1)},\dots,A^{(i_j)}}(x_{i_1},\dots,x_{i_j})(t)\Big]\,. \qquad (2.3)
\end{aligned}$$

Die Aussage b) lässt sich nun durch Induktion über k beweisen. Dabei entspricht a) dem Induktionsanfang ($k = 0$).
Sei die Aussage also für alle $k' < k$ bewiesen. Dann gilt:

$$\begin{aligned}
& \int t\, d\psi_{A^{(\nu_1)},\dots,A^{(\nu_k)}}(x_{\nu_1},\dots,x_{\nu_k})(t) \\
= \; & \int t\, d\Big[\Big(\psi_\mu(t) + \sum_{\{i_1,\dots,i_j\}\subset\{\nu_1,\dots,\nu_k\}} \psi_{A^{(i_1)},\dots,A^{(i_j)}}(x_{i_1},\dots,x_{i_j})(t)\Big) \\
& -\Big(\psi_\mu(t) + \sum_{\{i_1,\dots,i_j\}\subsetneq\{\nu_1,\dots,\nu_k\}} \psi_{A^{(i_1)},\dots,A^{(i_j)}}(x_{i_1},\dots,x_{i_j})(t)\Big)\Big] \\
= \; & \int t\, d\Big[\psi_\mu(t) + \sum_{\{i_1,\dots,i_j\}\subset\{\nu_1,\dots,\nu_k\}} \psi_{A^{(i_1)},\dots,A^{(i_j)}}(x_{i_1},\dots,x_{i_j})(t)\Big] \\
& - \int t\, d\Big[\psi_\mu(t) + \sum_{\{i_1,\dots,i_j\}\subsetneq\{\nu_1,\dots,\nu_k\}} \psi_{A^{(i_1)},\dots,A^{(i_j)}}(x_{i_1},\dots,x_{i_j})(t)\Big] \\
= \; & \int t\, d\Big[\psi_\mu(t) + \sum_{\{i_1,\dots,i_j\}\subset\{\nu_1,\dots,\nu_k\}} \psi_{A^{(i_1)},\dots,A^{(i_j)}}(x_{i_1},\dots,x_{i_j})(t)\Big] \\
& -\Big(\int t\, d\psi_\mu(t) + \sum_{\{i_1,\dots,i_j\}\subsetneq\{\nu_1,\dots,\nu_k\}} \int t\, d\psi_{A^{(i_1)},\dots,A^{(i_j)}}(x_{i_1},\dots,x_{i_j})(t)\Big) \\
= \; & \Big(\phi_\mu + \sum_{\{i_1,\dots,i_j\}\subset\{\nu_1,\dots,\nu_k\}} \phi_{A^{(i_1)},\dots,A^{(i_j)}}(x_{i_1},\dots,x_{i_j})\Big) \\
& -\Big(\phi_\mu + \sum_{\{i_1,\dots,i_j\}\subsetneq\{\nu_1,\dots,\nu_k\}} \phi_{A^{(i_1)},\dots,A^{(i_j)}}(x_{i_1},\dots,x_{i_j})\Big) \\
& \qquad \text{(mit (2) und der Induktionsvoraussetzung)} \\
= \; & \phi_{A^{(\nu_1)},\dots,A^{(\nu_k)}}(x_{\nu_1},\dots,x_{\nu_k})\,.
\end{aligned}$$

□

Satz 2.15 *Zwischen den Zerlegungen von $\tilde{\phi}$ und ψ gelten folgende Beziehungen:*

a) $\tilde{\phi}_\mu = \int q \circ t \, d\psi_\mu(t)$.

b) $\forall\{\nu_1, \ldots, \nu_k\} \subset \{1, \ldots, m\}$:
$\tilde{\phi}_{A^{(\nu_1)},\ldots,A^{(\nu_k)}}(x_{\nu_1}, \ldots, x_{\nu_k}) = \int q \circ t \, d\psi_{A^{(\nu_1)},\ldots,A^{(\nu_k)}}(x_{\nu_1}, \ldots, x_{\nu_k})(t)$.

Beweis: Der Beweis verläuft analog zu dem von Satz 2.14, unter Beachtung von $Z = q \circ Y$.
□

Folgender trivialer Zusammenhang besteht, bei Linearität von q, zwischen den Zerlegungen von ϕ und $\tilde{\phi} = q \circ \phi$:

Lemma 2.16 *Sei q linear in dem Sinne, dass gilt: $q(t) = u_0 + u_1 \cdot t$ mit $u_0, u_1 \in \mathbb{R}$. Dann gilt:*

a) $\tilde{\phi}_\mu = u_0 + u_1 \cdot \phi_\mu$.

b) $\forall\{\nu_1, \ldots, \nu_k\} \subset \{1, \ldots, m\}$:
$\tilde{\phi}_{A^{(\nu_1)},\ldots,A^{(\nu_k)}}(x_{\nu_1}, \ldots, x_{\nu_k}) = u_1 \cdot \phi_{A^{(\nu_1)},\ldots,A^{(\nu_k)}}(x_{\nu_1}, \ldots, x_{\nu_k})$.

Beweis:

a) $Z = u_0 + u_1 \cdot Y \Rightarrow \tilde{\phi}_\mu = \mathrm{E}(Z) = u_0 + u_1 \cdot \mathrm{E}(Y) = u_0 + u_1 \cdot \phi_\mu$.

b) Korollar 2.8 liefert zunächst

$$\begin{aligned}
& \tilde{\phi}_\mu + \sum_{\{i_1,\ldots,i_j\} \subset \{\nu_1,\ldots,\nu_k\}} \tilde{\phi}_{A^{(i_1)},\ldots,A^{(i_j)}}(x_{i_1}, \ldots, x_{i_j}) \\
= \; & \mathrm{E}(Z | a^{(\nu_1)} = x_{\nu_1}, \ldots, a^{(\nu_k)} = x_{\nu_k}) \\
= \; & u_0 + u_1 \cdot \mathrm{E}(Y | a^{(\nu_1)} = x_{\nu_1}, \ldots, a^{(\nu_k)} = x_{\nu_k}) \\
= \; & u_0 + u_1 \cdot \phi_\mu + u_1 \cdot \sum_{\{i_1,\ldots,i_j\} \subset \{\nu_1,\ldots,\nu_k\}} \phi_{A^{(i_1)},\ldots,A^{(i_j)}}(x_{i_1}, \ldots, x_{i_j}) .
\end{aligned}$$

Die Aussage ergibt sich nun mit einem ähnlichen induktiven Argument wie in Satz 2.14.

□

Nun formulieren und beweisen wir Implikationen zwischen den verschiedenen Hypothesen. Dazu definieren wir zunächst die Begriffe *feiner* und *gröber* für Hypothesen.

Definition 2.17 *Sei $x \subset \{a^{(1)}, \ldots, a^{(m)}\}$ und seien $M_1, M_2 \in \{\phi, \hat{\phi}, \psi, \tau, \rho\}$. $H_0^{M_1}(x)$ heißt* feiner *als $H_0^{M_2}(x)$ (bzw. $H_0^{M_2}(x)$* gröber *als $H_0^{M_1}(x)$), falls $H_0^{M_1}(x) \Longrightarrow H_0^{M_2}(x)$ gilt. Entsprechend heißt das durch M_1 festgelegte statistische Modell* feiner *als das zu M_2 gehörige, falls gilt: $\forall x \subset \{a^{(1)}, \ldots, a^{(m)}\} : H_0^{M_1}(x) \Longrightarrow H_0^{M_2}(x)$.*

Die Nichtgültigkeit einer Hypothese $H_0^{M_2}(x)$ impliziert, dass alle noch feineren Hypothesen ebenfalls nicht gültig sind. Im Sinne einer formallogischen Vorgehensweise beim Testen von Hypothesen, gemäß der zu Beginn dieses Abschnittes gemachten Anmerkungen, genügt es also, zum Verwerfen zunächst die groben Hypothesen zu betrachten. Falls eine grobe Hypothese jedoch nicht abgelehnt wird, können trotzdem möglicherweise feinere Hypothesen verworfen werden, die dann im nächsten Schritt zu untersuchen sind. Ein in diesem Zusammenhang auftretender Diskussionspunkt wäre der, ob bei einem solchen sequentiellen Vorgehen das Testniveau zu adjustieren ist, falls man sich bei Testentscheidungen zum einen auf dieselbe Datenbasis stützt, aber zum anderen ein völlig anderes Modell zu Grunde legt.

Das anschließende Lemma bildet einen Sonderfall, da in ihm verschiedene Hypothesen innerhalb eines Modells betrachtet werden.

Lemma 2.18 *Für $\{i, j\} \subset \{1, \dots, m\}$ beliebig gilt:*

a) $H_0^{\phi}(a^{(i)}|a^{(j)}) \Longrightarrow H_0^{\phi}(a^{(i)})$ (und ebenso für $H_0^{\psi},\ H_0^{\rho},\ H_0^{\tau^\alpha}$ und $H_0^{\tilde{\phi}}$) .

b) $H_0^{\phi}(a^{(i)}|a^{(j)}) \Longrightarrow H_0^{\phi}(a^{(i)}, a^{(j)})$ (und ebenso für $H_0^{\psi},\ H_0^{\rho},\ H_0^{\tau^\alpha}$ und $H_0^{\tilde{\phi}}$) .

Beweis:

a) Sei $\phi_{A^{(i)}} + \phi_{A^{(i)},A^{(j)}} \equiv 0$. Dann gilt für beliebiges $x_i \in E^{(i)}$:

$$\begin{aligned}
0 &= \int (\phi_{A^{(i)}}(x_i) + \phi_{A^{(i)},A^{(j)}}(x_i, x_j))\, dG_j(x_j) \\
&= \int \phi_{A^{(i)}}(x_i)\, dG_j(x_j) + \int \phi_{A^{(i)},A^{(j)}}(x_i, x_j)\, dG_j(x_j) \\
&= \phi_{A^{(i)}}(x_i) + 0 \ = \ \phi_{A^{(i)}}(x_i) \ .
\end{aligned}$$

Somit ist $\phi_{A^{(i)}} \equiv 0$. (Analog beweist man die Aussage für $H_0^{\psi},\ H_0^{\rho},\ H_0^{\tau^\alpha}$ und $H_0^{\tilde{\phi}}$).

b) Mit $\phi_{A^{(i)}} + \phi_{A^{(i)},A^{(j)}} \equiv 0$ und dem daraus in a) gefolgerten Resultat $\phi_{A^{(i)}} \equiv 0$ gilt sofort auch $\phi_{A^{(i)},A^{(j)}} \equiv 0$. (Analog ergibt sich die Behauptung für $H_0^{\psi},\ H_0^{\rho},\ H_0^{\tau^\alpha}$ und $H_0^{\tilde{\phi}}$).

□

Lemma 2.19 *Bei Linearität von q im Sinne von $q(t) = u_0 + u_1 \cdot t$ mit $u_0, u_1 \in \mathbb{R}$, gilt für die Hypothesen über die Faktoren und über die Wechselwirkungen:*

$$H_0^{\phi}(a^{(i_1)} \dots a^{(i_k)}) \Longleftrightarrow H_0^{\tilde{\phi}}(a^{(i_1)} \dots a^{(i_k)}) \quad (\{i_1, \dots, i_k\} \subset \{1, \dots, m\} \ \textit{beliebig}) \ .$$

Beweis: Die Aussage folgt unmittelbar mit Lemma 2.16. $\square$

Anmerkung: Falls q nicht linear ist, gilt die Behauptung aus Lemma 2.19 nicht, wie folgendes einfache Beispiel zeigt:
Sei $Y = \phi(a^{(1)}, a^{(2)}, \varepsilon)$ mit $E^{(1)} = \{x_1^{(1)}, x_2^{(1)}\}$ und $E^{(2)} = \{x_1^{(2)}, x_2^{(2)}\}$ sowie
$P(a^{(1)} = x_1^{(1)}) = P(a^{(1)} = x_2^{(1)}) = \frac{1}{2} = P(a^{(2)} = x_1^{(2)}) = P(a^{(2)} = x_2^{(2)})$.
Weiterhin sei $Z = q \circ Y$ mit $q(t) = t^2$. Die Verteilung von Y sei der Einfachheit halber durch die folgenden Einpunktverteilungen gegeben:

$$\begin{aligned}
a^{(1)} = x_1^{(1)}, a^{(2)} = x_1^{(2)} \quad &\Rightarrow \quad Y = 1, \\
a^{(1)} = x_2^{(1)}, a^{(2)} = x_1^{(2)} \quad &\Rightarrow \quad Y = 2, \\
a^{(1)} = x_1^{(1)}, a^{(2)} = x_2^{(2)} \quad &\Rightarrow \quad Y = 4, \\
a^{(1)} = x_2^{(1)}, a^{(2)} = x_2^{(2)} \quad &\Rightarrow \quad Y = 3.
\end{aligned}$$

In der Zerlegung von ϕ ergibt sich dann

$$\begin{aligned}
Y \quad = \quad & \phi_\mu + \phi_{A^{(1)}}(a^{(1)}) + \phi_{A^{(2)}}(a^{(2)}) + \phi_{A^{(1)},A^{(2)}}(a^{(1)}, a^{(2)}) + \phi_\varepsilon(\varepsilon) \quad \text{mit} \\
& \phi_\mu = 2.5, \\
& \phi_{A^{(1)}} \equiv 0, \\
& \phi_{A^{(2)}}(x_1^{(2)}) = -1,\ \phi_{A^{(2)}}(x_2^{(2)}) = 1, \\
& \phi_{A^{(1)},A^{(2)}}(x_1^{(1)}, x_1^{(1)}) = -0.5,\ \phi_{A^{(1)},A^{(2)}}(x_2^{(1)}, x_1^{(1)}) = 0.5, \\
& \phi_{A^{(1)},A^{(2)}}(x_1^{(1)}, x_2^{(1)}) = 0.5,\ \phi_{A^{(1)},A^{(2)}}(x_2^{(1)}, x_2^{(1)}) = -0.5 \quad \text{und} \\
& \phi_\varepsilon(\varepsilon) \equiv 0\,.
\end{aligned}$$

Somit gilt $H_0^\phi(a^{(1)}) : \phi_{A^{(1)}} \equiv 0$.
In der Zerlegung von $\tilde{\phi}$ erhält man dagegen

$$\begin{aligned}
& \tilde{\phi}_\mu = 7.5, \\
& \tilde{\phi}_{A^{(1)}}(x_1^{(1)}) = 1, \tilde{\phi}_{A^{(1)}}(x_1^{(1)}) = -1, \\
& \tilde{\phi}_{A^{(2)}}(x_1^{(2)}) = -5,\ \tilde{\phi}_{A^{(2)}}(x_2^{(2)}) = 5, \\
& \tilde{\phi}_{A^{(1)},A^{(2)}}(x_1^{(1)}, x_1^{(1)}) = -2.5,\ \tilde{\phi}_{A^{(1)},A^{(2)}}(x_2^{(1)}, x_1^{(1)}) = 2.5, \\
& \tilde{\phi}_{A^{(1)},A^{(2)}}(x_1^{(1)}, x_2^{(1)}) = 2.5,\ \tilde{\phi}_{A^{(1)},A^{(2)}}(x_2^{(1)}, x_2^{(1)}) = -2.5 \quad \text{und} \\
& \tilde{\phi}_\varepsilon(\varepsilon) \equiv 0\,,
\end{aligned}$$

so dass $H_0^{\tilde{\phi}}(a^{(1)})$ nicht gilt.

Lemma 2.20 *Für beliebige $\{i_1, \ldots, i_k\} \subset \{1, \ldots, m\}$ gilt:*

a) $H_0^{\psi}(a^{(i_1)}, \ldots, a^{(i_k)}) \Longrightarrow H_0^{\phi}(a^{(i_1)}, \ldots, a^{(i_k)})$.

b) $H_0^{\psi}(a^{(i_1)}, \ldots, a^{(i_k)}) \Longrightarrow H_0^{\tilde{\phi}}(a^{(i_1)}, \ldots, a^{(i_k)})$.

Das auf der Zerlegung von ψ basierende Modell ist also feiner als die durch ϕ und $\tilde{\phi}$ festgelegten Modelle.

Beweis: Die Behauptung ist eine unmittelbare Folgerung aus Satz 2.14 bzw. Satz 2.15:

a) Sei $\{i_1, \ldots, i_k\} \subset \{1, \ldots, m\}$ beliebig und sei $\psi_{A^{(i_1)}, \ldots, A^{(i_k)}} \equiv 0$.
Dann gilt für $(x_{i_1}, \ldots, x_{i_k}) \in E^{(i_1)} \times \ldots \times E^{(i_k)}$ beliebig:

$$\begin{aligned} \phi_{A^{(i_1)}, \ldots, A^{(i_k)}}(x_{i_1}, \ldots, x_{i_k}) &= \int t \, d\psi_{A^{(i_1)}, \ldots, A^{(i_k)}}(x_{i_1}, \ldots, x_{i_k})(t) \quad \text{(mit Satz 2.14)} \\ &= 0 \,. \end{aligned}$$

b) Ergibt sich analog.

□

Anmerkung: Die Relation aus Lemma 2.20 lässt sich nicht umkehren, d.h. es gilt im Allgemeinen nicht:

$$H_0^{\phi}(a^{(1)}, a^{(2)}) \Longrightarrow H_0^{\psi}(a^{(1)}, a^{(2)}) \,.$$

Dazu betrachten wir folgendes Gegenbeispiel:
Sei $Y = \phi(a^{(1)}, a^{(2)}, \varepsilon)$ und $Y \sim F$ mit $F(t) = \psi(a^{(1)}, a^{(2)})(t)$.
Weiterhin seien $E^{(1)} = \{x_1^{(1)}, x_2^{(1)}\}$ und $E^{(2)} = \{x_1^{(2)}, x_2^{(2)}\}$ mit
$P(a^{(1)} = x_1^{(1)}) = P(a^{(1)} = x_2^{(1)}) = \frac{1}{2} = P(a^{(2)} = x_1^{(2)}) = P(a^{(2)} = x_2^{(2)})$.
Die Verteilung von Y sei gegeben durch

$$\begin{aligned} \psi(x_1^{(1)}, x_1^{(2)})(t) &= (t - 0.5) \cdot 1\!\!1_{]0.5,1.5]}(t), \\ \psi(x_2^{(1)}, x_1^{(2)})(t) &= (t - 1.5) \cdot 1\!\!1_{]1.5,2.5]}(t), \\ \psi(x_1^{(1)}, x_2^{(2)})(t) &= (t - 2.5) \cdot 1\!\!1_{]2.5,3.5]}(t), \\ \psi(x_2^{(1)}, x_2^{(2)})(t) &= (t - 3.5) \cdot 1\!\!1_{]3.5,4.5]}(t). \end{aligned}$$

Dann berechnen sich die Erwartungswerte $E(Y|a^{(1)} = x_i^{(1)}, a^{(2)} = x_j^{(2)})$ zu

$$\begin{aligned} E(Y|a^{(1)} = x_1^{(1)}, a^{(2)} = x_1^{(2)}) &= 1, \\ E(Y|a^{(1)} = x_2^{(1)}, a^{(2)} = x_1^{(2)}) &= 2, \\ E(Y|a^{(1)} = x_1^{(1)}, a^{(2)} = x_2^{(2)}) &= 3, \end{aligned}$$

$$E(Y|a^{(1)}=x_2^{(1)}, a^{(2)}=x_2^{(2)}) \;=\; 4,$$

und es ergibt sich in der Zerlegung

$$\begin{aligned}
Y \;=\; & \phi_\mu + \phi_{A^{(1)}}(a^{(1)}) + \phi_{A^{(2)}}(a^{(2)}) + \phi_{A^{(1)},A^{(2)}}(a^{(1)}, a^{(2)}) + \phi_\varepsilon(\varepsilon) \quad \text{mit} \\
& \phi_\mu = 2.5, \\
& \phi_{A^{(1)}}(x_1^{(1)}) = -0.5,\ \phi_{A^{(1)}}(x_2^{(1)}) = 0.5, \\
& \phi_{A^{(2)}}(x_1^{(2)}) = -1,\ \phi_{A^{(2)}}(x_2^{(2)}) = 1, \\
& \phi_{A^{(1)},A^{(2)}} \equiv 0, \quad \text{und} \\
& \phi_\varepsilon(\varepsilon) \sim \mathrm{R}[-0.5, 0.5] \quad \text{(Rechteckverteilung auf dem Intervall } [-0.5, 0.5] \text{) .}
\end{aligned}$$

Somit gilt die Hypothese $H_0^\phi(a^{(1)}, a^{(2)}) : \phi_{A^{(1)},A^{(2)}} \equiv 0$.
Die Hypothese $H_0^\psi(a^{(1)}, a^{(2)}) : \psi_{A^{(1)},A^{(2)}} \equiv 0$ gilt jedoch nicht, denn es sind

$$\begin{aligned}
\psi_\mu(t) =& \frac{1}{4}(t-0.5)\cdot \mathbb{1}_{]0.5,4.5]}(t), \\
\psi_{A^{(1)}}(x_1^{(1)})(t) =& \frac{1}{4}(t-0.5)\cdot \mathbb{1}_{]0.5,1.5]}(t) - \frac{1}{4}(t-2.5)\cdot \mathbb{1}_{]1.5,2.5]}(t) + \\
& \frac{1}{4}(t-2.5)\cdot \mathbb{1}_{]2.5,3.5]}(t) - \frac{1}{4}(t-4.5)\cdot \mathbb{1}_{]3.5,4.5]}(t), \\
\psi_{A^{(2)}}(x_1^{(2)})(t) =& \frac{1}{4}(t-0.5)\cdot \mathbb{1}_{]0.5,2.5]}(t) - \frac{1}{4}(t-4.5)\cdot \mathbb{1}_{]2.5,4.5]}(t)
\end{aligned}$$

und damit z.B.

$$\begin{aligned}
& \psi_{A^{(1)},A^{(2)}}(x_1^{(1)}, x_1^{(2)})(t) \\
=\; & \psi(x_1^{(1)}, x_1^{(2)})(t) - \psi_\mu(t) - \psi_{A^{(1)}}(x_1^{(1)})(t) - \psi_{A^{(2)}}(x_1^{(2)})(t) \\
=\; & \frac{1}{4}(t-0.5)\cdot \mathbb{1}_{]0.5,1.5]}(t) - \frac{1}{4}(t-2.5)\cdot \mathbb{1}_{]1.5,3.5]}(t) + \frac{1}{4}(t-4.5)\cdot \mathbb{1}_{]3.5,4.5]}(t)\ ,
\end{aligned}$$

also insbesondere $\psi_{A^{(1)},A^{(2)}} \not\equiv 0$.

Lemma 2.21 *Im Modell ohne Wechselwirkungen sind die Hypothesen H_0^ψ und H_0^ρ über die Faktoren äquivalent, d.h.:*

$$H_0^\psi(a^{(i)}) \Longleftrightarrow H_0^\rho(a^{(i)}) \quad (i \in \{1, \ldots, m\} \textit{ beliebig})\ .$$

Beweis: Seien $(x_1, \ldots, x_m) \in E^{(1)} \times \cdots \times E^{(m)}$ und $x_i' \in E^{(i)}$ beliebig. Sei weiterhin

$$\begin{aligned}
& F^{Y|a^{(1)}=x_1,\ldots,a^{(m)}=x_m} = F_{p_1} \quad \text{mit} \\
& p_1 = g^{-1}(\rho_\mu + \sum_{k=1}^m \rho_{A^{(k)}}(x_k)) \quad \text{und}
\end{aligned}$$

$$F^{Y|a^{(1)}=x_1,\dots,a^{(i-1)}=x_{i-1},a^{(i)}=x_i',a^{(i+1)}=x_{i+1},\dots,a^{(m)}=x_m} = F_{p_2} \quad \text{mit}$$
$$p_2 = g^{-1}\Big(\rho_\mu + \sum_{\substack{k=1\\k\neq i}}^{m} \rho_{A^{(k)}}(x_k) + \rho_{A^{(i)}}(x_i')\Big) \ .$$

Außerdem sei

$$F^{Y|a^{(1)}=x_1,\dots,a^{(m)}=x_m}(t) = \psi_\mu(t) + \sum_{k=1}^{m} \psi_{A^{(k)}}(x_k)(t) \quad \text{und}$$
$$F^{Y|a^{(1)}=x_1,\dots,a^{(i-1)}=x_{i-1},a^{(i)}=x_i',a^{(i+1)}=x_{i+1},\dots,a^{(m)}=x_m}(t) = \psi_\mu(t) + \sum_{\substack{k=1\\k\neq i}}^{m} \psi_{A^{(k)}}(x_k)(t) + \psi_{A^{(i)}}(x_i')(t) \ .$$

Dann gilt:

„$\Longrightarrow$“: Aus $\psi_{A^{(i)}} \equiv 0$ folgt:

$$\begin{aligned}
& F^{Y|a^{(1)}=x_1,\dots,a^{(m)}=x_m}(t) - F^{Y|a^{(1)}=x_1,\dots,a^{(i-1)}=x_{i-1},a^{(i)}=x_i',a^{(i+1)}=x_{i+1},\dots,a^{(m)}=x_m}(t) \\
& = \psi_{A^{(i)}}(x_i)(t) - \psi_{A^{(i)}}(x_i')(t) = 0 \\
\Rightarrow\ & F_{p_1} - F_{p_2} \ \Rightarrow\ p_1 = p_2 \\
\Rightarrow\ & \rho_\mu + \sum_{k=1}^{m} \rho_{A^{(k)}}(x_k) = \rho_\mu + \sum_{\substack{k=1\\k\neq i}}^{m} \rho_{A^{(k)}}(x_k) + \rho_{A^{(i)}}(x_i') \\
& \quad \text{(wegen der Bijektivität von } g) \\
\Rightarrow\ & \rho_{A^{(i)}}(x_i) = \rho_{A^{(i)}}(x_i') \\
\Rightarrow\ & \rho_{A^{(i)}} \equiv c \quad \text{(wegen beliebiger Wahl von } x_i, x_i') \\
\Rightarrow\ & \rho_{A^{(i)}} \equiv 0 \quad (\text{wegen } \mathrm{E}(\rho_{A^{(i)}}(a^{(i)})) = 0) \ .
\end{aligned}$$

„$\Longleftarrow$“: Aus $\rho_{A^{(i)}} \equiv 0$ folgt:

$$\begin{aligned}
& \rho_\mu + \sum_{k=1}^{m} \rho_{A^{(k)}}(x_k) \ = \ \rho_\mu + \sum_{\substack{k=1\\k\neq i}}^{m} \rho_{A^{(k)}}(x_k) + \rho_{A^{(i)}}(x_i') \\
\Rightarrow\ & p_1 = p_2 \\
\Rightarrow\ & F_{p_1} = F_{p_2} \\
\Rightarrow\ & \psi_\mu(t) + \sum_{k=1}^{m} \psi_{A^{(k)}}(x_k)(t) = \psi_\mu(t) + \sum_{\substack{k=1\\k\neq i}}^{m} \psi_{A^{(k)}}(x_k)(t) + \psi_{A^{(i)}}(x_i')(t) \\
\Rightarrow\ & \psi_{A^{(i)}}(x_i) = \psi_{A^{(i)}}(x_i') \ \Rightarrow\ \forall t : \psi_{A^{(i)}}(.)(t) \equiv c \\
\Rightarrow\ & \forall t : \psi_{A^{(i)}}(.)(t) \equiv 0 \quad (\text{wegen } \mathrm{E}(\psi_{A^{(i)}}(a^{(i)})(t)) = 0) \\
\Rightarrow\ & \psi_{A^{(i)}} \equiv 0 \ .
\end{aligned}$$

□

Anmerkung: Die Behauptung aus Lemma 2.21 gilt nicht im Modell mit Wechselwirkungen, wie folgendes Gegenbeispiel zeigt:
Sei $Y = \phi(a^{(1)}, a^{(2)}, \varepsilon)$ und $Y \sim F$ mit $F(t) = \psi(a^{(1)}, a^{(2)})(t)$.
Weiterhin seien $E^{(1)} = \{x_1^{(1)}, x_2^{(1)}\}$ und $E^{(2)} = \{x_1^{(2)}, x_2^{(2)}\}$ mit
$P(a^{(1)} = x_1^{(1)}) = P(a^{(1)} = x_2^{(1)}) = \frac{1}{2} = P(a^{(2)} = x_1^{(2)}) = P(a^{(2)} = x_2^{(2)})$.
Es gelte $Y \sim F_p$ mit $\log(\frac{p}{1-p}) = \rho(a^{(1)}, a^{(2)})$.
Die Verteilung von Y sei gegeben durch

$$\begin{aligned}
a^{(1)} = x_1^{(1)}, a^{(2)} = x_1^{(2)} \quad &\Rightarrow \quad Y \sim B(0.2) \quad \text{(Bernoulli-Verteilung)}\,,\\
a^{(1)} = x_1^{(1)}, a^{(2)} = x_2^{(2)} \quad &\Rightarrow \quad Y \sim B(0.6),\\
a^{(1)} = x_2^{(1)}, a^{(2)} = x_1^{(2)} \quad &\Rightarrow \quad Y \sim B(0.4),\\
a^{(1)} = x_2^{(1)}, a^{(2)} = x_2^{(2)} \quad &\Rightarrow \quad Y \sim B(0.4).
\end{aligned}$$

Dann ergibt sich in der Zerlegung von ψ für $t \in [0, 1[$ (für $t < 0$ ist $\psi(a^{(1)}, a^{(2)})(t) \equiv 0$ und für $t \geq 1$ ist $\psi(a^{(1)}, a^{(2)})(t) \equiv 1$, diese Fälle brauchen also nicht weiter betrachtet zu werden):

$$\begin{aligned}
&\psi_\mu(t) = 0.6,\\
&\psi_{A^{(1)}}(x_1^{(1)})(t) = 0, \psi_{A^{(1)}}(x_2^{(1)})(t) = 0,\\
&\psi_{A^{(2)}}(x_1^{(2)})(t) = 0.1, \psi_{A^{(2)}}(x_2^{(2)})(t) = -0.1,\\
&\psi_{A^{(1)},A^{(2)}}(x_1^{(1)}, x_1^{(2)})(t) = 0.1, \psi_{A^{(1)},A^{(2)}}(x_1^{(1)}, x_2^{(2)})(t) = -0.1,\\
&\psi_{A^{(1)},A^{(2)}}(x_2^{(1)}, x_1^{(2)}(t)) = -0.1, \psi_{A^{(1)},A^{(2)}}(x_2^{(1)}, x_2^{(2)})(t) = 0.1\ .
\end{aligned}$$

Somit gilt $H_0^\psi(a^{(1)}) : \psi_{A^{(1)}} \equiv 0$.
Mit $g(p) = \log(\frac{p}{1-p})$ ist

$$\begin{aligned}
a^{(1)} = x_1^{(1)}, a^{(2)} = x_1^{(2)} \quad &\Rightarrow \quad g(p) = \log(1/4) \approx -1.386,\\
a^{(1)} = x_1^{(1)}, a^{(2)} = x_2^{(2)} \quad &\Rightarrow \quad g(p) = \log(3/2) \approx 0.405,\\
a^{(1)} = x_2^{(1)}, a^{(2)} = x_1^{(2)} \quad &\Rightarrow \quad g(p) = \log(2/3) \approx -0.405,\\
a^{(1)} = x_2^{(1)}, a^{(2)} = x_2^{(2)} \quad &\Rightarrow \quad g(p) = \log(2/3) \approx -0.405,
\end{aligned}$$

und die Zerlegung von ρ liefert

$$\begin{aligned}
&\rho_\mu = -0.448,\\
&\rho_{A^{(1)}}(x_1^{(1)}) = -0.043, \rho_{A^{(1)}}(x_2^{(1)}) = 0.043,\\
&\rho_{A^{(2)}}(x_1^{(2)}) = -0.448, \rho_{A^{(2)}}(x_2^{(2)}) = 0.448,\\
&\rho_{A^{(1)},A^{(2)}}(x_1^{(1)}, x_1^{(2)}) = -0.448, \rho_{A^{(1)},A^{(2)}}(x_1^{(1)}, x_2^{(2)}) = 0.448,\\
&\rho_{A^{(1)},A^{(2)}}(x_2^{(1)}, x_1^{(2)}) = 0.448, \rho_{A^{(1)},A^{(2)}}(x_2^{(1)}, x_2^{(2)}) = -0.448\ .
\end{aligned}$$

Damit ist die Hypothese $H_0^\rho(a^{(1)}) : \rho_{A^{(1)}} \equiv 0$ nicht gültig.

Lemma 2.22 *Im Modell ohne Wechselwirkungen im Sinne der Zerlegung von ϕ gilt:*

$$H_0^\rho(a^{(i)}) \overset{(1)}{\Longleftrightarrow} H_0^\psi(a^{(i)}) \overset{(2)}{\Longleftrightarrow} H_0^\phi(a^{(i)}) \quad (i \in \{1, \ldots, m\} \textit{ beliebig}) .$$

Beweis: Wegen der Lemmata 2.20 und 2.21 bleibt lediglich die Aussage „$\Longleftarrow$" aus (2) zu zeigen:
Das Modell ohne Wechselwirkungen im Sinne der Zerlegung von ϕ lässt sich schreiben als

$$Y = \phi_\mu + \sum_{k=1}^{m} \phi_{A^{(k)}}(a^{(k)}) + \phi_\varepsilon(\varepsilon) .$$

Es gilt nun für $(x_1, \ldots, x_m) \in E^{(1)} \times \cdots \times E^{(m)}$ und $x'_i \in E^{(i)}$ beliebig:

$$\begin{aligned}
\phi_{A^{(i)}} \equiv 0 \quad &\Rightarrow \quad Y = \phi_\mu + \sum_{\substack{k=1 \\ k\neq i}}^{m} \phi_{A^{(k)}}(a^{(k)}) + \phi_\varepsilon(\varepsilon) \\
&\Rightarrow \quad F^{Y|a^{(1)}=x_1,\ldots,a^{(m)}=x_m}(t) \\
&\qquad\qquad = F^{Y|a^{(1)}=x_1,\ldots,a^{(i-1)}=x_{i-1},a^{(i)}=x'_i,a^{(i+1)}=x_{i+1},\ldots,a^{(m)}=x_m}(t) \\
&\Rightarrow \quad \psi_{A^{(i)}}(x_i) = \psi_{A^{(i)}}(x'_i) \\
&\Rightarrow \quad \forall t : \psi_{A^{(i)}}(.)(t) \equiv c \\
&\Rightarrow \quad \forall t : \psi_{A^{(i)}}(.)(t) \equiv 0 \\
&\Rightarrow \quad \psi_{A^{(i)}} \equiv 0 .
\end{aligned}$$

□

Korollar 2.23 *Bei nur einem Faktor sind die Hypothesen $H_0^\rho(a)$, $H_0^\psi(a)$ und $H_0^\phi(a)$ äquivalent.*

Beweis: Bei nur einem Faktor existieren keine Wechselwirkungen. Somit ergibt sich die Behauptung mit Lemma 2.22. □

Lemma 2.24 *Sei $a = (a^{(1)}, \ldots, a^{(m)})$, also $Y = \phi(a, \varepsilon)$ und $\phi_a = \int \phi(a, \varepsilon)\, dG_\varepsilon(\varepsilon)$. Seien weiter $F,\ F^{Y|a=x}$ rechtsstetige Verteilungsfunktionen mit α-Quantilen*
$Q(\alpha) = \inf\{x : F(x) \geq \alpha\}$ bzw. $Q_Y(\alpha) = \inf\{x : F_Y(x) \geq \alpha\}$.

a) Seien $F,\ F_Y$ strikt isoton und sei $a = x \in E^{(1)} \times \cdots \times E^{(m)}$ fest vorgegeben. Dann gilt:

$$\forall t : F_Y(t) = F(t - \phi_a) \Longleftrightarrow \forall \alpha \in]0, 1[: Q_Y(\alpha) = Q(\alpha) + \phi_a . \tag{2.4}$$

b) Sei $Y \sim F_Y$ und $Z = q(Y)$ mit isotonem, bijektivem q. Dann gilt für $\alpha \in]0, 1[$ beliebig:

$$Q_Z(\alpha) = q(Q_Y(\alpha)) .$$

c) Sei $Y \sim F_Y$ mit symmetrischer Verteilung F_Y und $E|Y| < \infty$. Dann ist $\phi_a = \tau^{0.5}(a)$.

Beweis:

a) „$\Longrightarrow$": Sei $\forall t: F_Y(t) = F(t - \phi_a)$ und $\alpha \in]0,1[$ beliebig. Dann ist

$$\begin{aligned} Q_Y(\alpha) &= \inf\{x : F_Y(x) \geq \alpha\} = \inf\{x : F(x - \phi_a) \geq \alpha\} = \inf\{x + \phi_a : F(x) \geq \alpha\} \\ &= \inf\{x : F(x) \geq \alpha\} + \phi_a = Q(\alpha) + \phi_a \,. \end{aligned}$$

„$\Longleftarrow$": Sei $\forall \alpha \in]0,1[$: $Q_Y(\alpha) = Q(\alpha) + \phi_a$. Sei t beliebig und α so, dass $Q_Y(\alpha) = t$ (existiert wegen der strikten Isotonie). Dann ist $F_Y(t) = \alpha$ (wegen der Rechtsstetigkeit von F_Y). Mit $Q(\alpha) = t - \phi_a$ folgt analog $F(t - \phi_a) = \alpha$.

b)

$$\begin{aligned} Q_Z(\alpha) &= \inf\{x : F_Z(x) \geq \alpha\} = \inf\{x : F_{q(Y)}(x) \geq \alpha\} \\ &= \inf\{x : P(q(Y) \leq x) \geq \alpha\} = \inf\{x : P(Y \leq q^{-1}(x)) \geq \alpha\} \\ &= \inf\{q(x) : P(Y \leq x) \geq \alpha\} = \inf\{q(x) : F_Y(x) \geq \alpha\} \\ &= q(\inf\{x : F_Y(x) \geq \alpha\}) = q(Q_Y(\alpha)) \,. \end{aligned}$$

c) Klar. □

Bedingung (2.4) bedeutet, dass die betrachteten Verteilungen F_Y einer Lokationsfamilie entstammen, in der es lediglich einen durch die Faktoren bestimmten Verschiebungsparameter gibt. Dies gilt z.B. im Linearen Modell. In einem solchen Fall spielt es offenbar für die Hypothesen keine Rolle, ob ϕ, ψ oder τ^α zerlegt wird.

Lemma 2.25 *Es gelte eine der äquivalenten Bedingungen aus (2.4). Sei $\alpha \in]0,1[$ beliebig. Dann sind die Hypothesen H_0^ϕ, H_0^ψ und $H_0^{\tau^\alpha}$ über die Faktoren und über die Wechselwirkungen äquivalent, d.h. für beliebige $\{i_1, \ldots, i_k\} \subset \{1, \ldots, m\}$ gilt:*

$$H_0^\phi(a^{(i_1)}, \ldots, a^{(i_k)}) \Longleftrightarrow H_0^\psi(a^{(i_1)}, \ldots, a^{(i_k)}) \Longleftrightarrow H_0^{\tau^\alpha}(a^{(i_1)}, \ldots, a^{(i_k)}) \,.$$

Beweis: Klar. □

Anmerkung: In (2.4) ist für die Implikation „$\Longleftarrow$" wesentlich, dass die Gleichheit $Q_Y(\alpha) = Q(\alpha) + \phi_a$ nicht nur für ein $\alpha \in]0,1[$ gilt.

Anmerkung: Falls keine der äquivalenten Bedingungen aus (2.4) erfüllt ist, sind die Hypothesen H_0^ϕ, H_0^ψ und $H_0^{\tau^\alpha}$ im Allgemeinen nicht äquivalent. Als Gegenbeispiel betrachte man $Y = \phi(a, \varepsilon),\ Q(\alpha) = \tau^\alpha(a),\ F_Y(t) = \psi(a)$ mit $E = \{1, 2\}$ und $P(a = 1) = P(a = 2) = \frac{1}{2}$. Die Verteilung von Y sei gegeben durch

$$a = 1 \Longrightarrow Y = \frac{1}{2}, \quad a = 2 \Longrightarrow Y \sim \mathrm{R}[0,1] \,.$$

Dann ist $Y = \phi_\mu + \phi_A(a) + \phi_\varepsilon(\varepsilon)$ mit $\phi_\mu = \frac{1}{2}$ und $\phi_A \equiv 0$. Andererseits ist $Q_Y(\frac{1}{4}) = \tau_\mu^{1/4} + \tau_A^{1/4}(a)$ mit $\tau_\mu^{1/4} = \frac{3}{8},\ \tau_A^{1/4}(1) = \frac{1}{8},\ \tau_A^{1/4}(2) = -\frac{1}{8}$. Außerdem ist $F_Y(\frac{1}{4}) = \psi_\mu(\frac{1}{4}) + \psi_A(a)(\frac{1}{4})$ mit $\psi_\mu(\frac{1}{4}) = \frac{1}{8},\ \psi_A(1)(\frac{1}{4}) = -\frac{1}{8},\ \psi_A(2)(\frac{1}{4}) = \frac{1}{8}$. Es gilt also $H_0^\phi(a)$, aber nicht $H_0^{\tau^\alpha}(a)$ und auch nicht $H_0^\psi(a)$.

Lemma 2.26 *Im Modell ohne Wechselwirkungen gilt für die Hypothesen über die Faktoren:*

$$H_0^{\psi}(a^{(i)}) \Longrightarrow H_0^{\tau^\alpha}(a^{(i)}) \qquad (\alpha \in]0,1[\textit{ beliebig}, \quad i \in \{1,\dots,m\} \textit{ beliebig}) .$$

Beweis: Es sei

$$F^{Y|a^{(1)}=x_1,\dots,a^{(m)}=x_m}(t) = \psi_\mu(t) + \sum_{k=1}^{m} \psi_{A^{(k)}}(x_k)(t) \qquad \text{und} \qquad \psi_{A^{(i)}} \equiv 0 .$$

Dann folgt für $x_i, x_{i'} \in E^{(i)}$ beliebig und $x_k \in E^{(k)}$ beliebig für $k \neq i$:

$$\begin{aligned} &F^{Y|a^{(1)}=x_1,\dots,a^{(m)}=x_m}(t) - F^{Y|a^{(1)}=x_1,\dots,a^{(i-1)}=x_{i-1},a^{(i)}=x_i',a^{(i+1)}=x_{i+1},\dots,a^{(m)}=x_m}(t) \\ &= \psi_{A^{(i)}}(x_i)(t) - \psi_{A^{(i)}}(x_i')(t) = 0 . \end{aligned}$$

Sei nun $Q(\alpha)^{|a^{(1)}=x_1,\dots,a^{(m)}=x_m}$ das α-Quantil von $F^{Y|a^{(1)}=x_1,\dots,a^{(m)}=x_m}$. Es gilt

$$\begin{aligned} &Q(\alpha)^{|a^{(1)}=x_1,\dots,a^{(m)}=x_m} = Q(\alpha)^{|a^{(1)}=x_1,\dots,a^{(i-1)}=x_{i-1},a^{(i)}=x_i',a^{(i+1)}=x_{i+1},\dots,a^{(m)}=x_m} \\ \Longrightarrow &\tau_\mu^\alpha + \sum_{k=1}^{m} \tau_{A^{(k)}}^\alpha(x_k) \ = \ \tau_\mu^\alpha + \sum_{\substack{k=1\\k\neq i}}^{m} \tau_{A^{(k)}}^\alpha(x_k) + \tau_{A^{(i)}}^\alpha(x_i') \\ \Longrightarrow &\tau_{A^{(i)}}^\alpha(x_i) \ = \ \tau_{A^{(i)}}^\alpha(x_i') \\ \Longrightarrow &\tau_{A^{(i)}}^\alpha \equiv c = 0 . \end{aligned}$$

□

Anmerkung: Die Behauptung aus Lemma 2.26 lässt sich nicht umkehren, d.h. es gilt im Allgemeinen nicht

$$H_0^{\tau^\alpha}(a^{(i)}) \Longrightarrow H_0^{\psi}(a^{(i)}) .$$

Dies ist, z.B. für $\alpha = \frac{1}{2}$, offensichtlich, wenn man symmetrische Verteilungen mit gleichem Symmetriezentrum, aber unterschiedlichen Streuungsparametern betrachtet.

Anmerkung: Die Behauptung aus Lemma 2.26 wird falsch, wenn man nicht die Annahme fehlender Wechselwirkungen fordert. Dazu betrachte man das folgende Gegenbeispiel:
Sei $Y = \phi(a^{(1)}, a^{(2)}, \varepsilon)$ und $Y \sim F$ mit $F(t) = \psi(a^{(1)}, a^{(2)})(t)$ und $Q_Y(\alpha) = \tau^\alpha(a^{(1)}, a^{(2)})$.
Es sei $E^{(1)} = \{x_1^{(1)}, x_2^{(1)}\}$ und $E^{(2)} = \{x_1^{(2)}, x_2^{(2)}\}$ mit
$P(a^{(1)} = x_1^{(1)}) = P(a^{(1)} = x_2^{(1)}) = \frac{1}{2} = P(a^{(2)} = x_1^{(2)}) = P(a^{(2)} = x_2^{(2)})$.
Die Verteilungsfunktionen seien im Einzelnen gegeben durch

$$\begin{aligned} \psi(x_1^{(1)}, x_1^{(2)})(t) &= 0.3 \cdot t \cdot \mathbb{1}_{]0,1]}(t) + (0.3 + 0.07 \cdot (t-1)) \cdot \mathbb{1}_{]1,11]}(t), \\ \psi(x_2^{(1)}, x_1^{(2)})(t) &= 0.7 \cdot t \cdot \mathbb{1}_{]0,1]}(t) + (0.7 + 0.03 \cdot (t-1)) \cdot \mathbb{1}_{]1,11]}(t), \\ \psi(x_1^{(1)}, x_2^{(2)})(t) &= 0.4 \cdot t \cdot \mathbb{1}_{]0,1]}(t) + (0.4 + 0.06 \cdot (t-1)) \cdot \mathbb{1}_{]1,11]}(t), \\ \psi(x_2^{(1)}, x_2^{(2)})(t) &= 0.6 \cdot t \cdot \mathbb{1}_{]0,1]}(t) + (0.6 + 0.04 \cdot (t-1)) \cdot \mathbb{1}_{]1,11]}(t). \end{aligned}$$

Es ist dann $\psi_\mu(t) = 0.5 \cdot t \cdot \mathbb{1}_{]0,1]}(t) + (0.5 + 0.05 \cdot (t-1)) \cdot \mathbb{1}_{]1,11]}(t)$ und $\psi_{A^{(2)}} \equiv 0$, aber für den Median ($\alpha = \frac{1}{2}$) gilt:

$$
\begin{aligned}
\tau^{1/2}(x_1^{(1)}, x_1^{(2)})(t) &= \frac{27}{7}\,, \qquad & \tau^{1/2}(x_2^{(1)}, x_1^{(2)})(t) &= \frac{5}{7}\,, \\
\tau^{1/2}(x_1^{(1)}, x_2^{(2)})(t) &= \frac{16}{6}\,, \qquad & \tau^{1/2}(x_2^{(1)}, x_2^{(2)})(t) &= \frac{5}{6}\,,
\end{aligned}
$$

und daher $\tau_\mu^{1/2} = \frac{8}{7} + \frac{7}{8}$ sowie $\tau^{1/2}_{A^{(2)}}(x_1^{(2)}) = \frac{8}{7} - \frac{7}{8}, \tau^{1/2}_{A^{(2)}}(x_2^{(2)}) = \frac{7}{8} - \frac{8}{7}$, also insbesondere $\tau^{1/2}_{A^{(2)}} \not\equiv 0$.

Zusammenfassend lassen sich die Implikationen zwischen den verschiedenen Hypothesen als Diagramm wie folgt darstellen:

$$
\begin{tikzcd}
H_0^{\tilde\phi} \arrow[rr, harpoon] \arrow[dr] & & H_0^{\phi} \arrow[dl, dashed] \arrow[dr, dashed, leftrightarrow] & \\
& H_0^{\psi} \arrow[ul, Rightarrow] \arrow[ur, Rightarrow] \arrow[rr, dashed, leftrightarrow, bend left=10] \arrow[rr, bend right=10] \arrow[d, leftrightarrow] & & H_0^{\tau_\alpha} \\
& H_0^{\rho} & &
\end{tikzcd}
$$

Legende:	$\Longrightarrow$	Implikation gilt immer
	$\longrightarrow$	Implikation gilt im Modell ohne Wechselwirkungen
	$\dashrightarrow$	Implikation gilt im Lokationsmodell gemäß (2.4)
	$\rightharpoonup$	Implikation gilt bei linearem q

Abbildung 3: Implikationen zwischen den Hypothesen.

Das Modell, in dem gemäß (M2) Verteilungsfunktionen zerlegt werden, ist also feiner als die bei einer direkten Modellierung der Zielgröße oder der Betrachtung einer transformierten Zielgröße entstehenden Modelle. Falls keine Wechselwirkungen existieren, sind alle betrachteten Modelle äquivalent, abgesehen von der Quantil-Zerlegung für festes α, die gröber als die anderen Modelle ist.

2.6 Beispiele für Modelle

In diesem Abschnitt werden beispielhaft einige konkrete statistische Modelle aufgeführt, auf die die hergeleiteten allgemeinen Aussagen anwendbar sind.

Diese kleine Auswahl von ein- und zweifaktoriellen Modellen soll sowohl die Analogien in den Modellformulierungen illustrieren, die darauf beruhen, dass alle genannten Modelle sich in das gleiche mathematische Konzept einbetten lassen, als auch die Beschreibung der Reparametrisierungsbedingungen und Hypothesen verdeutlichen.

Eine weitaus ausführlichere Darstellung von Beispielen, einschließlich der Teststatistiken und eines Vergleichs mit entsprechenden „klassischen“ Modellformulierungen, die man der statistischen Literatur entnehmen kann, folgt in Kapitel 4.

Zu den Annahmen (A1)-(A5) vgl. jeweils die Seiten 39f.

2.6.1 Lineares Modell

Ein fester Faktor
MODELLGLEICHUNG
Unter den Annahmen (A1),(A2) und (A3) gilt:

$$Y = \phi(a, \varepsilon) = \phi_\mu + \phi_A(a) + \phi_\varepsilon(\varepsilon)$$

mit $\phi_\varepsilon(\varepsilon) \sim \mathrm{N}(0, \sigma_\varepsilon^2)$ und $\mathrm{E}_A(\phi_A(a)) = \int \phi_A(x) dG_A(x) = 0$.
$\phi_A(a)$ und $\phi_\varepsilon(\varepsilon)$ sind unabhängig.
HYPOTHESEN
Hypothese über den Faktor: $H_0^\phi(a) : \phi_A \equiv 0$.

Ein zufälliger Faktor
MODELLGLEICHUNG
Unter den Annahmen (A1),(A2), (A3) und (A4) gilt:

$$Y = \phi(a, \varepsilon) = \phi_\mu + \phi_A(a) + \phi_\varepsilon(\varepsilon)$$

mit $\phi_\varepsilon(\varepsilon) \sim \mathrm{N}(0, \sigma_\varepsilon^2)$ und $\phi_A(a) \sim \mathrm{N}(0, \sigma_A^2)$.
$\phi_A(a)$ und $\phi_\varepsilon(\varepsilon)$ sind unabhängig.
HYPOTHESEN
Hypothese über den Faktor: $H_0^\phi(a) : \phi_A \equiv 0$.

Eine Kovariable
MODELLGLEICHUNG
Unter den Annahmen (A1),(A2), (A3) und (A5) gilt:

$$Y = \phi(a, \varepsilon) = \phi_\mu + \xi + \xi_1 \cdot a + \phi_\varepsilon(\varepsilon) =: \xi_0 + \xi_1 \cdot a + \phi_\varepsilon(\varepsilon)$$

mit $\phi_\varepsilon(\varepsilon) \sim \mathrm{N}(0, \sigma_\varepsilon^2)$.
a und $\phi_\varepsilon(\varepsilon)$ sind unabhängig.
Aus $\mathrm{E}_A(\xi + \xi_1 \cdot a) = \int(\xi + \xi_1 \cdot x)dG_A(x) = \xi + \xi_1 \int x dG_A(x) = 0$ folgt $\xi = -\frac{\int x dG_A(x)}{\xi_1}$.
HYPOTHESEN
$H_0^{\xi_0} : \xi_0 = 0$ und $H_0^{\xi_1} : \xi_1 = 0$.

Zwei feste Faktoren
MODELLGLEICHUNG
Unter den Annahmen (A1),(A2) und (A3) gilt:

$$Y = \phi(a^{(1)}, a^{(2)}, \varepsilon) = \phi_\mu + \phi_{A^{(1)}}(a^{(1)}) + \phi_{A^{(2)}}(a^{(2)}) + \phi_{A^{(1)},A^{(2)}}(a^{(1)}, a^{(2)}) + \phi_\varepsilon(\varepsilon)$$

mit $\phi_\varepsilon(\varepsilon) \sim \mathrm{N}(0, \sigma_\varepsilon^2),\ \mathrm{E}_{A^{(i)}}(\phi_{A^{(i)}}(a^{(i)})) = \int \phi_{A^{(i)}}(x)dG_{A^{(i)}}(x) = 0,\ i = 1, 2,$
$\forall x \in E^{(1)} : \mathrm{E}(\phi_{A^{(1)},A^{(2)}}(a^{(1)}, a^{(2)})|a^{(1)} = x) = 0$ und
$\forall x \in E^{(2)} : \mathrm{E}(\phi_{A^{(1)},A^{(2)}}(a^{(1)}, a^{(2)})|a^{(2)} = x) = 0$.
$\phi_{A^{(1)}}(a^{(1)}),\ \phi_{A^{(2)}}(a^{(2)}),\ \phi_{A^{(1)},A^{(2)}}(a^{(1)}, a^{(2)})$ und $\phi_\varepsilon(\varepsilon)$ sind unabhängig.
HYPOTHESEN
Hypothese über den ersten Faktor: $H_0^\phi(a^{(1)}) : \phi_{A^{(1)}} \equiv 0$.
Hypothese über den zweiten Faktor: $H_0^\phi(a^{(2)}) : \phi_{A^{(2)}} \equiv 0$.
Hypothese über die Wechselwirkung: $H_0^\phi(a^{(1)}, a^{(2)}) : \phi_{A^{(1)},A^{(2)}} \equiv 0$.

Zwei zufällige Faktoren
MODELLGLEICHUNG
Unter den Annahmen (A1),(A2), (A3) und (A4) gilt:

$$Y = \phi(a^{(1)}, a^{(2)}, \varepsilon) = \phi_\mu + \phi_{A^{(1)}}(a^{(1)}) + \phi_{A^{(2)}}(a^{(2)}) + \phi_{A^{(1)},A^{(2)}}(a^{(1)}, a^{(2)}) + \phi_\varepsilon(\varepsilon)$$

mit $\phi_\varepsilon(\varepsilon) \sim \mathrm{N}(0, \sigma_\varepsilon^2),\ \phi_{A^{(i)}}(a^{(i)}) \sim \mathrm{N}(0, \sigma^2_{A^{(i)}}),\ i = 1, 2,$
und $\phi_{A^{(1)},A^{(2)}}(a^{(1)}, a^{(2)}) \sim \mathrm{N}(0, \sigma^2_{A^{(1)},A^{(2)}})$.
$\phi_{A^{(1)}}(a^{(1)}),\ \phi_{A^{(2)}}(a^{(2)}),\ \phi_{A^{(1)},A^{(2)}}(a^{(1)}, a^{(2)})$ und $\phi_\varepsilon(\varepsilon)$ sind unabhängig.
HYPOTHESEN
Hypothese über den ersten Faktor: $H_0^\phi(a^{(1)}) : \phi_{A^{(1)}} \equiv 0$.
Hypothese über den zweiten Faktor: $H_0^\phi(a^{(2)}) : \phi_{A^{(2)}} \equiv 0$.
Hypothese über die Wechselwirkung: $H_0^\phi(a^{(1)}, a^{(2)}) : \phi_{A^{(1)},A^{(2)}} \equiv 0$.

Ein fester und ein zufälliger Faktor
MODELLGLEICHUNG
Unter den Annahmen (A1),(A2), (A3) und (A4) gilt:

$$Y = \phi(a^{(1)}, a^{(2)}, \varepsilon) = \phi_\mu + \phi_{A^{(1)}}(a^{(1)}) + \phi_{A^{(2)}}(a^{(2)}) + \phi_{A^{(1)},A^{(2)}}(a^{(1)}, a^{(2)}) + \phi_\varepsilon(\varepsilon)$$

mit $\phi_\varepsilon(\varepsilon) \sim \mathrm{N}(0, \sigma_\varepsilon^2),\ \mathrm{E}_{A^{(1)}}(\phi_{A^{(1)}}(a^{(1)})) = \int \phi_{A^{(1)}}(x)dG_{A^{(1)}}(x) = 0,\ \phi_{A^{(2)}}(a^{(2)}) \sim \mathrm{N}(0, \sigma^2_{A^{(2)}})$
und $\phi_{A^{(1)},A^{(2)}}(a^{(1)}, a^{(2)}) \sim \mathrm{N}(0, \sigma^2_{A^{(1)},A^{(2)}})$.

$\phi_{A^{(1)}}(a^{(1)}),\ \phi_{A^{(2)}}(a^{(2)}),\ \phi_{A^{(1)},A^{(2)}}(a^{(1)},a^{(2)})$ und $\phi_\varepsilon(\varepsilon)$ sind unabhängig.
HYPOTHESEN
Hypothese über den ersten Faktor: $H_0^\phi(a^{(1)}) : \phi_{A^{(1)}} \equiv 0$.
Hypothese über den zweiten Faktor: $H_0^\phi(a^{(2)}) : \phi_{A^{(2)}} \equiv 0$.
Hypothese über die Wechselwirkung: $H_0^\phi(a^{(1)},a^{(2)}) : \phi_{A^{(1)},A^{(2)}} \equiv 0$.

2.6.2 Allgemeines Nichtparametrisches Modell

Ein fester Faktor
MODELLGLEICHUNG

$$F(t) = \psi(a)(t) = \psi_\mu(t) + \psi_A(a)(t)$$

mit $\forall t : \mathrm{E}_A(\psi_A(a)(t)) = \int \psi_A(x)(t)dG_A(x) = 0.$
HYPOTHESEN
Hypothese über den Faktor: $H_0^\psi(a) : \psi_A \equiv 0$.

Zwei feste Faktoren
MODELLGLEICHUNG

$$F(t) = \psi(a)(t) = \psi_\mu(t) + \psi_{A^{(1)}}(a^{(1)})(t) + \psi_{A^{(2)}}(a^{(2)})(t) + \psi_{A^{(1)},A^{(2)}}(a^{(1)},a^{(2)})(t)$$

mit $\forall t :\ \mathrm{E}_{A^{(i)}}(\psi_{A^{(i)}}(a^{(i)})) = \int \psi_{A^{(i)}}(x)(t)dG_{A^{(i)}}(x) = 0,\ i = 1,2,$
$\forall t \forall x \in E^{(1)} :\ \mathrm{E}(\psi_{A^{(1)},A^{(2)}}(a^{(1)},a^{(2)})(t)|a^{(1)} = x) = 0$ und
$\forall t \forall x \in E^{(2)} :\ \mathrm{E}(\psi_{A^{(1)},A^{(2)}}(a^{(1)},a^{(2)})(t)|a^{(2)} = x) = 0$.
$\psi_{A^{(1)}}(a^{(1)}),\ \psi_{A^{(2)}}(a^{(2)})$ und $\psi_{A^{(1)},A^{(2)}}(a^{(1)},a^{(2)})$ sind unabhängig.
HYPOTHESEN
Hypothese über den ersten Faktor: $H_0^\psi(a^{(1)}) : \psi_{A^{(1)}} \equiv 0$.
Hypothese über den zweiten Faktor: $H_0^\psi(a^{(2)}) : \psi_{A^{(2)}} \equiv 0$.
Hypothese über die Wechselwirkung: $H_0^\psi(a^{(1)},a^{(2)}) : \psi_{A^{(1)},A^{(2)}} \equiv 0$.

2.6.3 Verallgemeinertes Lineares Modell

Ein fester Faktor
MODELLGLEICHUNG

$$g(p) = \rho(a) = \rho_\mu + \rho_A(a)$$

mit $\mathrm{E}_A(\rho_A(a)) = \int \rho_A(x)dG_A(x) = 0.$
HYPOTHESEN
Hypothese über den Faktor: $H_0^\rho(a) : \rho_A \equiv 0$.

Zwei feste Faktoren

MODELLGLEICHUNG

Unter einer zu (A2) analogen Annahme gilt:

$$g(p) = \rho(a^{(1)}, a^{(2)}) = \rho_\mu + \rho_{A^{(1)}}(a^{(1)}) + \rho_{A^{(2)}}(a^{(2)}) + \rho_{A^{(1)},A^{(2)}}(a^{(1)}, a^{(2)})$$

mit $\mathrm{E}_{A^{(i)}}(\rho_{A^{(i)}}(a^{(i)})) = \int \rho_{A^{(i)}}(x) dG_{A^{(i)}}(x) = 0,\ i = 1, 2,$
$\forall x \in E^{(1)}$: $\mathrm{E}(\rho_{A^{(1)},A^{(2)}}(a^{(1)}, a^{(2)}) | a^{(1)} = x) = 0$ und
$\forall x \in E^{(2)}$: $\mathrm{E}(\rho_{A^{(1)},A^{(2)}}(a^{(1)}, a^{(2)}) | a^{(2)} = x) = 0$.
$\rho_{A^{(1)}}(a^{(1)}),\ \rho_{A^{(2)}}(a^{(2)})$ und $\rho_{A^{(1)},A^{(2)}}(a^{(1)}, a^{(2)})$ sind unabhängig.

HYPOTHESEN

Hypothese über den ersten Faktor: $H_0^\rho(a^{(1)}) : \rho_{A^{(1)}} \equiv 0$.
Hypothese über den zweiten Faktor: $H_0^\rho(a^{(2)}) : \rho_{A^{(2)}} \equiv 0$.
Hypothese über die Wechselwirkung: $H_0^\rho(a^{(1)}, a^{(2)}) : \rho_{A^{(1)},A^{(2)}} \equiv 0$.

Kapitel 3

Statistische Experimente

3.1 Modellierung Statistischer Designs

Im vorigen Kapitel haben wir statistische Modelle allgemein definiert und Aussagen über Faktoren und Hypothesen getroffen. Darauf aufbauend wenden wir uns nun, um die Anwendbarkeit dieser Modellformulierungen zu zeigen, statistischen Versuchsdesigns zu. In ihnen werden wiederholte Beobachtungen in Form von Zufallsvariablen, die den Bedingungen der statistischen Modelle genügen, modelliert. Dazu betrachten wir einen $\mathbb{R}^{m+1}$-wertigen stochastischen Prozeß $(\Omega, \mathcal{A}, P, (X_\iota)_{\iota \in I})$ mit einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ und einer darauf definierten Familie von Zufallsvektoren $X_\iota = (a_\iota^{(1)}, \ldots, a_\iota^{(m)}, \varepsilon_\iota),\ \iota \in I$, deren $m+1$ Komponenten unabhängig sind. Allerdings sind die Folgen $(a_\iota^{(i)})_{\iota \in I}$ für festes $i \in \{1, \ldots, m\}$ nicht notwendig unabhängig.
Außerdem betrachten wir durch Funktionen ϕ und q induzierte stochastische Prozesse $(Y_\iota)_{\iota \in I} = (\phi(X_\iota))_{\iota \in I}$ und $(Z_\iota)_{\iota \in I} = (q(Y_\iota))_{\iota \in I}$.
$E = E^{(1)} \times \ldots \times E^{(m)} \times E^{(\varepsilon)} \subset \mathbb{R}^{m+1}$ sei der Zustandsraum des stochastischen Prozesses $(X_\iota)_{\iota \in I}$. Für $j \in \{1, \ldots, m\}$ gelte: Die Zufallsvariablen $a_\iota^{(j)},\ \iota \in I$, seien identisch verteilt mit Verteilungsfunktion $G_j(t)$. Die Verteilungsfunktion der ebenfalls identisch verteilten $\varepsilon_\iota,\ \iota \in I$, wird mit $G_\varepsilon(t)$ bezeichnet. Die Indexmenge I ist eine endliche Teilmenge $\{1, \ldots, N\}$ aus $\mathbb{N}$.
Die Verteilungsfunktion von Y_ι wird mit $F_\iota(t)$ bezeichnet. Es gelte

$$F_\iota(t) = \psi(a_\iota^{(1)}, \ldots, a_\iota^{(m)})\,(t)\ ,\ \iota \in I\ .$$

ψ ist dabei eine Abbildung von $\mathbb{R}^{m+1}$ in einen Funktionenraum B, so dass die Verteilungsfunktionen auf $\mathbb{R}$ in B enthalten sind, d.h. $\psi : \mathbb{R}^{m+1} \to B$.
τ_ι^α sei das α-Quantil der Verteilung F_ι.
Besitzen die $Y_\iota,\ \iota \in I$, Verteilungen aus einer durch einen Parameter $p \in \mathbf{P}$ parametrisierten Familie von Verteilungen, so bezeichne p_ι den Parameter der Verteilung von $Y_\iota,\ \iota \in I$.
Wie im vorigen Kapitel nennen wir die Y_ι Zielgrößen, $a_\iota^{(1)}, \ldots, a_\iota^{(m)}$ Faktoren und ε_ι Versuchsfehler.

Die Modellgleichungen, durch die die statistischen Modelle definiert werden, gelten hier für jedes Element Y_ι des stochastischen Prozesses. Dabei werden bei den statistischen Designs sinnvolle Vorgaben bezüglich der Werte gemacht, die gewisse Faktoren annehmen. Der Prozeß wird also dahingehend gesteuert, daß z.B. nicht $a_1^{(1)} = \ldots = a_N^{(1)} = x$ gilt, sondern die Ausprägungen des Faktors $a^{(1)}$ entsprechend dessen unterstellter theoretischer Verteilung variieren. Damit ist klar, daß dann z.B. die Zufallsvariablen $a_1^{(1)}, \ldots, a_N^{(1)}$ nicht unabhängig sind. Eine derartige Unabhängigkeit ist aber auch im Allgemeinen gar nicht von Interesse.

Das Versuchsdesign selbst ist schließlich dadurch definiert, welche Kombinationen der Faktorstufen in wievielen Wiederholungen auftreten. Klassische Beispiele sind dabei das vollständig gekreuzte Design (vgl. Abschnitt 4.1.5) und der Split-Plot (vgl. Abschnitt 4.1.6).

Zur Veranschaulichung der Bezeichnungen betrachten wir das konkrete Beispiel, das schon auf S.20 beschrieben wurde (Wirkung zweier Düngemittel auf den Tomatenertrag). Die Rohdaten (vgl. BOX ET AL. (1978, Table 4.1) oder HAND ET AL. (1994, S.316)) liegen in folgender Form vor:

Positionsnummer	1	2	3	4	5	6	7	8	9	10	11
Düngemittel	A	A	B	B	A	B	B	B	A	A	B
Ertrag [Pfund]	29.9	11.4	26.6	23.7	25.3	28.5	14.2	17.9	16.5	21.1	24.3

Tabelle 2 *Datensatz: Tomatenertrag unter Verwendung zweier verschiedener Düngemittel. Aus* BOX ET AL. (1978).

Beim Modellieren des statistischen Designs indizieren wir die betrachteten Prozesse mit den angegebenen Positionsnummern, es ist also $I = \{1, \ldots, 11\}$. Damit ist $(a_\iota)_{\iota \in I} = (a_1, \ldots, a_{11}) = (1, 1, 2, 2, 1, 2, 2, 2, 1, 1, 2)$ und $(Y_\iota)_{\iota \in I} = (Y_1, \ldots, Y_{11}) = (29.9, \ldots, 24.3)$. Der Faktor $a = a^{(1)}$ nimmt die beiden Werte 1 und 2 an, also $E^{(1)} = \{1, 2\}$; und $(a_\iota)_{\iota \in I}$ ist beobachtbar, genau wie $(Y_\iota)_{\iota \in I}$. Modellgleichung (M1) bedeutet hier angewendet: $Y_\iota = \phi(a_\iota, \varepsilon_\iota),\ \iota = 1, \ldots, 11$. Dabei ist der Versuchsfehler ε_ι nicht beobachtbar und ϕ unbekannt.

Aus den in einem statistischen Design beobachteten Zielvariablen $Y_\iota = \phi(a_\iota^{(1)}, \ldots, a_\iota^{(m)}, \varepsilon_\iota)$ und den Faktoren $a_\iota^{(1)}, \ldots, a_\iota^{(m)}$ werden Schätzer für die unbekannten Faktoreinflüsse $\phi_{A^{(f)}}$ der festen Faktoren, die Wechselwirkungen $\phi_{A^{(f_1)}, \ldots, A^{(f_k)}}$ zwischen den festen Faktoren, ggf. die Koeffizienten der Kovariablen und Schätzer für das Gesamtmittel konstruiert. Dabei wird im Allgemeinen angestrebt, die Summe der quadratischen Abweichungen zwischen den tatsächlich beobachteten Y_ι und den Schätzern

$$\hat{Y}_\iota = \hat{\phi}(a_\iota^{(1)}, \ldots, a_\iota^{(m)}, \varepsilon_\iota) = \hat{\phi}_\mu + \sum_{k=1}^{m} \sum_{\{i_1, \ldots, i_k\} \subset \{1, \ldots, m\}} \hat{\phi}_{A^{(i_1)}, \ldots, A^{(i_k)}}(a_\iota^{i_1}, \ldots, a_\iota^{i_k})$$

zu minimieren, analog geschieht dies für die Schätzer der transformierten Zielvariablen Z_ι, der Verteilungsfunktionen F_ι und der Parameter p_ι.
Über die exakte oder asymptotische Verteilung dieser Schätzer unter gewissen Hypothesen erhält man Tests auf die jeweiligen Hypothesen.

Tests sind Abbildungen aus dem Stichprobenraum in die Menge $\{0, 1\}$. Die konkrete Gestalt der Testfunktionen werden wir jedoch im Folgenden bei den Beispielen nicht angeben, da sie durch die angegebenen Teststatistiken und ihre jeweiligen Verteilungen vollständig determiniert sind.

Für manche Zwecke ist es vorteilhaft, die betrachteten stochastischen Prozesse in Vektoren bzw. Matrizen zusammenzufassen.
Dazu seien $\tilde{\mathbf{X}} := (X'_1, \ldots, X'_N)'$, $\tilde{\mathbf{Y}} := (Y_1, \ldots, Y_N)'$ und $\tilde{\mathbf{Z}} := (Z_1, \ldots, Z_N)'$.
Außerdem definieren wir eine Ordnungsrelation „$\prec$" für die Zufallsvariablen X_ι, Y_ι und Z_ι:

$$\begin{aligned} X_{\iota_1} \prec X_{\iota_2} \quad :\Longleftrightarrow \quad & a_{\iota_1}^{(1)} < a_{\iota_2}^{(1)} \\ & \vee\ \exists j \in \{1, \ldots, m\}\ \forall i < j : a_{\iota_1}^{(i)} = a_{\iota_2}^{(i)},\ a_{\iota_1}^{(j)} < a_{\iota_2}^{(j)} \\ & \vee\ \forall j \in \{1, \ldots, m\} : a_{\iota_1}^{(j)} = a_{\iota_2}^{(j)},\ \iota_1 < \iota_2\,. \end{aligned}$$

Für Y_ι und Z_ι sei die Relation vollkommen analog definiert.
Mithilfe dieser Ordnungsrelation konstruieren wir „sortierte Prozesse", bei denen die vorkommenden Zufallsvektoren $X_1, \ldots, X_N$ und die Zufallsvariablen $Y_1, \ldots, Y_N$ sowie $Z_1, \ldots, Z_N$ nach den Werten sortiert sind, die die Faktoren $a_\iota^{(1)}, \ldots, a_\iota^{(m)}$ annehmen.
Die sortierten Prozesse bezeichnen wir mit X_ι^s, Y_ι^s und Z_ι^s. Sie seien also so definiert, daß für sie gilt:

$$\forall \iota_1, \iota_2 :\ \iota_1 < \iota_2 \Longrightarrow X_{\iota_1}^s \prec X_{\iota_2}^s,\ Y_{\iota_1}^s \prec Y_{\iota_2}^s,\ Z_{\iota_1}^s \prec Z_{\iota_2}^s\,.$$

Auch diese Prozesse werden wieder zu Vektoren zusammengefaßt, die wir nun mit $\mathbf{X}$, $\mathbf{Y}$ und $\mathbf{Z}$ bezeichnen.

Wir werden im Weiteren auch gelegentlich, um die Schreibweise zu vereinfachen, die in der Statistik üblichen Multiindizes zur Beschreibung von Faktorstufenkombinationen verwenden. Dazu wird jedem $\iota \in I = \{1, \ldots, N\}$ das $m+1$-Tupel $(\tilde{a}_\iota^{(1)}, \ldots, \tilde{a}_\iota^{(m)}, n_\iota)$ zugeordnet. Es gelte dabei

$$\tilde{a}_\iota^{(i)} = \begin{cases} a_\iota^{(i)}, & \text{falls } a^{(i)} \in \mathcal{M}_f \cup \mathcal{M}_k \quad \text{(fester oder kovariater Faktor)} \\ \lceil a_\iota^{(i)} \rceil, & \text{falls } a^{(i)} \in \mathcal{M}_r \quad \text{(zufälliger Faktor)}, \end{cases}$$

$$n_\iota = 1 + \text{card}\{\iota' \in I : \iota' < \iota, \forall a^{(i)} \in \mathcal{M}_f \cup \mathcal{M}_r : \tilde{a}_\iota^{(i)} = \tilde{a}_{\iota'}^{(i)}\}\,.$$

Anstelle von Y_ι schreiben wir dann, insbesondere bei konkreteren Beispielen, auch $Y_{\tilde{a}_\iota^{(1)}, \ldots, \tilde{a}_\iota^{(m)}, n_\iota}$.
Weiterhin machen wir Gebrauch von der ebenfalls in der statistischen Literatur üblichen abkürzenden Schreibweise für Summen und Mittelwerte. Dabei bedeute, im Falle $\mathcal{M} = \mathcal{M}_f \cup \mathcal{M}_r$, ein Punkt anstelle eines oder mehrerer Indizes $\tilde{a}_\iota^{(i_1)}, \tilde{a}_\iota^{(i_2)}, \ldots$ innerhalb des Multiindexes die Summation über sämtliche Y_ι, bei denen die restlichen Indizes der Faktoren übereinstimmen. Beispielsweise ist

$$Y_{\cdot, \tilde{a}_\iota^{(2)}, \ldots, \tilde{a}_\iota^{(m)}} = \sum_{\substack{\iota' \in I \\ \forall i \neq 1 : \tilde{a}_\iota^{(i)} = \tilde{a}_{\iota'}^{(i)}}} Y_{\iota'}\,.$$

Mittelwertbildung wird durch zusätzliches Überstreichen gekennzeichnet, z.B.

$$\bar{Y}_{\cdot, \tilde{a}_\iota^{(2)}, \ldots, \tilde{a}_\iota^{(m)}} = \text{card}\{\iota' \in I : \forall i \neq 1 : \tilde{a}_\iota^{(i)} = \tilde{a}_{\iota'}^{(i)}\}^{-1} \cdot Y_{\cdot, \tilde{a}_\iota^{(2)}, \ldots, \tilde{a}_\iota^{(m)}}\,.$$

3.2 Kovarianzstruktur

Die Kovarianzstruktur der Zufallsvariablen Y_ι (bzw. Z_ι) ist, je nach betrachtetem Modell, unterschiedlich, und mehr oder weniger explizit angebbar. In den nachfolgenden Lemmata werden dazu einige Resultate formuliert.

Lemma 3.1 *Seien $Y_\iota, \iota = 1, \dots, N$, die Zufallsvariablen in einem statistischen Design, das mit einem durch Modellgleichung (M1) und die Annahmen (A1), (A2), (A2a) und (A3a) (s. S.39) gegebenen Modell korrespondiert (Lineares Modell).*
Dann gilt für $\iota, \iota' \in \{1, \dots, N\}$:

$$\begin{aligned}
& \mathrm{Cov}(Y_\iota, Y_{\iota'}) \\
= & \sum_{k=1}^{m} \sum_{\{i_1,\dots,i_k\}\subset\{1,\dots,m\}} \mathrm{Cov}\big(\phi_{A^{(i_1)},\dots,A^{(i_k)}}(a_\iota^{(i_1)},\dots,a_\iota^{(i_k)}), \phi_{A^{(i_1)},\dots,A^{(i_k)}}(a_{\iota'}^{(i_1)},\dots,a_{\iota'}^{(i_k)})\big) \\
& + \begin{cases} \sigma_\varepsilon^2(a_\iota^{(1)},\dots,a_\iota^{(m)}), & \iota = \iota' \\ 0, & \text{sonst.} \end{cases}
\end{aligned}$$

Beweis:

$$\begin{aligned}
& \mathrm{Cov}(Y_\iota, Y_{\iota'}) \\
= & \mathrm{Cov}\big(\phi(a_\iota^{(1)},\dots,a_\iota^{(m)},\varepsilon_\iota), \phi(a_{\iota'}^{(1)},\dots,a_{\iota'}^{(m)},\varepsilon_{\iota'})\big) \\
= & \mathrm{Cov}\big(\phi_\mu + \sum_{k=1}^{m} \sum_{\{i_1,\dots,i_k\}\subset\{1,\dots,m\}} \phi_{A^{(i_1)},\dots,A^{(i_k)}}(a_\iota^{(i_1)},\dots,a_\iota^{(i_k)}) + \phi_\varepsilon(\varepsilon_\iota), \\
& \quad \phi_\mu + \sum_{k=1}^{m} \sum_{\{i_1,\dots,i_k\}\subset\{1,\dots,m\}} \phi_{A^{(i_1)},\dots,A^{(i_k)}}(a_{\iota'}^{(i_1)},\dots,a_{\iota'}^{(i_k)}) + \phi_\varepsilon(\varepsilon_{\iota'})\big) \\
= & \sum_{k=1}^{m} \sum_{\{i_1,\dots,i_k\}\subset\{1,\dots,m\}} \mathrm{Cov}\big(\phi_{A^{(i_1)},\dots,A^{(i_k)}}(a_\iota^{(i_1)},\dots,a_\iota^{(i_k)}), \phi_{A^{(i_1)},\dots,A^{(i_k)}}(a_{\iota'}^{(i_1)},\dots,a_{\iota'}^{(i_k)})\big) \\
& + \mathrm{Cov}\big(\phi_\varepsilon(\varepsilon_\iota), \phi_\varepsilon(\varepsilon_{\iota'})\big) .
\end{aligned}$$

□

Lemma 3.2 *Seien die Annahmen aus Lemma 3.1 erfüllt. Außerdem gebe es nur feste Faktoren, d.h. $\mathcal{M} = \mathcal{M}_f$.*
Dann gilt für $\iota, \iota' \in \{1, \dots, N\}$:

$$\mathrm{Cov}(Y_\iota, Y_{\iota'} | a_\iota^{(1)} = x_1, \ldots, a_\iota^{(m)} = x_m, a_{\iota'}^{(1)} = x'_1, \ldots, a_{\iota'}^{(m)} = x'_m) = \begin{cases} \sigma_\varepsilon^2(x_1, \ldots, x_m), & \iota = \iota' \\ 0, & \text{sonst.} \end{cases}$$

Beweis: Folgt aus Lemma 3.1. □

Lemma 3.3 *Seien die Annahmen aus Lemma 3.1 und zusätzlich Annahme (A4) erfüllt. Es gebe außerdem nur zufällige Faktoren, d.h. $\mathcal{M} = \mathcal{M}_r$.*
Dann gilt für $\iota, \iota' \in \{1, \ldots, N\}$:

$$\begin{aligned} \mathrm{Cov}(Y_\iota, Y_{\iota'}) &= \sum_{k=1}^{m} \sum_{\{i_1,\ldots,i_k\} \subset \{1,\ldots,m\}} \begin{cases} \sigma^2_{A^{(i_1)},\ldots,A^{(i_k)}}, & (a_\iota^{(i_1)}, \ldots, a_\iota^{(i_k)}) = (a_{\iota'}^{(i_1)}, \ldots, a_{\iota'}^{(i_k)}) \\ 0, & \text{sonst} \end{cases} \\ &+ \begin{cases} \sigma_\varepsilon^2(a_\iota^{(1)}, \ldots, a_\iota^{(m)}), & \iota = \iota' \\ 0, & \text{sonst.} \end{cases} \end{aligned}$$

Beweis: Folgt aus Lemma 3.1. □

Lemma 3.4 *Seien die Annahmen aus Lemma 3.1 und zusätzlich Annahme (A4) erfüllt. Sei weiterhin $\mathcal{M} = \mathcal{M}_f \cup \mathcal{M}_r$ (nur feste und zufällige Faktoren), und sei $\mathcal{M}_f = \{a^{(f_1)}, \ldots, a^{(f_{m_f})}\}$.*
Dann gilt für $\iota, \iota' \in \{1, \ldots, N\}$:

$$\begin{aligned} &\mathrm{Cov}(Y_\iota, Y_{\iota'} | a_\iota^{(f_1)} = x_{f_1}, \ldots, a_\iota^{(f_{m_f})} = x_{f_{m_f}}, a_{\iota'}^{(f_1)} = x'_{f_1}, \ldots, a_{\iota'}^{(f_{m_f})} = x'_{f_{m_f}}) \\ = &\sum_{k=1}^{m} \sum_{\substack{\{i_1,\ldots,i_k\} \subset \{1,\ldots,m\} \\ \{i_1,\ldots,i_k\} \cap \mathcal{M}_r \neq \emptyset}} \begin{cases} \sigma^2_{A^{(i_1)},\ldots,A^{(i_k)}}, & (a_\iota^{(i_1)}, \ldots, a_\iota^{(i_k)}) = (a_{\iota'}^{(i_1)}, \ldots, a_{\iota'}^{(i_k)}) \\ 0, & \text{sonst} \end{cases} \\ &+ \begin{cases} \sigma_\varepsilon^2(a_\iota^{(1)}, \ldots, a_\iota^{(m)}), & \iota = \iota' \\ 0, & \text{sonst.} \end{cases} \end{aligned}$$

Beweis: Folgt aus Lemma 3.1. □

Lemma 3.5 *Seien die Annahmen aus Lemma 3.1 und zusätzlich Annahme (A4) erfüllt. Sei $\mathcal{M} = \mathcal{M}_f \cup \mathcal{M}_r \cup \mathcal{M}_k$ und $\mathcal{M}_f = \{a^{(f_1)}, \ldots, a^{(f_{m_f})}\}$.*
Dann gilt für $\iota, \iota' \in \{1, \ldots, N\}$:

$$
\begin{aligned}
&\mathrm{Cov}(Y_\iota, Y_{\iota'} | a_\iota^{(f_1)} = x_{f_1}, \dots, a_\iota^{(f_{m_f})} = x_{f_{m_f}}, a_{\iota'}^{(f_1)} = x'_{f_1}, \dots, a_{\iota'}^{(f_{m_f})} = x'_{f_{m_f}}) \\
= &\sum_{k=1}^{m} \sum_{\substack{\{i_1,\dots,i_k\} \subset \{1,\dots,m\} \\ \{i_1,\dots,i_k\} \cap \mathcal{M}_r \neq \emptyset}} \begin{cases} \sigma^2_{A^{(i_1)},\dots,A^{(i_k)}}, & (a_\iota^{(i_1)}, \dots, a_\iota^{(i_k)}) = (a_{\iota'}^{(i_1)}, \dots, a_{\iota'}^{(i_k)}) \\ 0, & \text{sonst} \end{cases} \\
&+ \sum_{k=1}^{m} \sum_{\{i_1,\dots,i_k\} \subset \mathcal{M}_k} \mathrm{Cov}\big(\phi_{A^{(i_1)},\dots,A^{(i_k)}}(a_\iota^{(i_1)}, \dots, a_\iota^{(i_k)}), \phi_{A^{(i_1)},\dots,A^{(i_k)}}(a_{\iota'}^{(i_1)}, \dots, a_{\iota'}^{(i_k)})\big) \\
&+ \begin{cases} \sigma^2_\varepsilon(a_\iota^{(1)}, \dots, a_\iota^{(m)}), & \iota = \iota' \\ 0, & \text{sonst.} \end{cases}
\end{aligned}
$$

Beweis: Folgt aus Lemma 3.1. □

Kapitel 4

Beispiele

Anhand der nun folgenden Beispiele soll exemplarisch die Anwendung (und Anwendbarkeit) der in dieser Arbeit vorgestellten allgemeinen Modellformulierung demonstriert werden. Dabei beschränken wir uns aus Platzgründen auf einige wenige Fälle, die so weit ausgeführt werden, dass klar wird, wie das Prinzip in anderen Situationen angewendet werden kann.
Namentlich handelt es sich bei den betrachteten Beispielen im Linearen Modell um den Ein-Stichproben-Fall sowie einige ein- und zweifaktorielle Designs und eine Anmerkung über das austauschbare Lineare Modell. Im Linearen Modell ohne Normalverteilung sowie im allgemeinen Nichtparametrischen Modell betrachten wir Ein-Stichproben-Tests und ein Design mit einem festen Faktor.
Darüber hinaus lassen sich in das in dieser Arbeit formulierte allgemeine Modell aber auch jeweils höherfaktorielle Designs sowie verallgemeinerte Lineare Modelle und gewisse robuste Verfahren einbetten.
Im Linearen Modell mit festen Faktoren ergeben sich hier Vorteile bei der Herleitung der Teststatistik, da die von uns betrachtete Zerlegung einen klareren Zugang zu den auftretenden quadratischen Formen darstellt. Somit lässt sich auch für den asymptotischen Fall l (= Anzahl der Faktorstufen) $\to \infty$, n (= Anzahl der Versuchswiederholungen pro Faktorstufe) fest, ein sehr allgemeines Resultat herleiten.
In einer gerade erschienenen Veröffentlichung, die etwas andere Aspekte derselben asymptotischen Fragestellung behandelt, haben die Autoren ihr wesentliches Resultat nur mit Hilfe einer Zerlegung erzielen können, die sehr ähnlich zu derjenigen ist, die wir in dieser Arbeit als grundlegende Methode etabliert haben (vgl. AKRITAS & ARNOLD (2000)).

Bei den angegebenen Versuchsplänen haben wir jeweils das Modell sowohl in der in dieser Arbeit durchgängig verwendeten Schreibweise als auch in einer Form, die der Standardliteratur entnommen wurde, formuliert. Dabei haben wir auch die Reparametrisierungsbedingungen angegeben, die sich bei uns als Folgerung aus Satz 2.7 ergeben. Die aufgeführten Resultate über die Kovarianzstruktur ergeben sich aus den Lemmata in Kapitel 3.2. Korollar 2.8 liefert die Grundlage für die angegebenen erwarteten Werte der Zielvariablen.
Die Modellgleichungen (M1), (M1a) und (M2), auf die wiederholt verwiesen wird, sind auf den Seiten 17f zu finden, die Annahmen (A1) bis (A5) stehen auf den Seiten 39f.

4.1 Lineares Modell

Das Lineare Modell sei zunächst durch Modellgleichung (M1) und die Annahmen (A1), (A2), (A2a), (A3), (A4) und (A5) gegeben, denen die Zufallsvariablen $Y_\iota,\ \iota \in I$, genügen.
Die $\varepsilon_\iota,\ \iota \in I$, seien unabhängig.

4.1.1 Ein-Stichproben-Tests

Wir betrachten eine Stichprobe der Größe N.

MODELL, ZERLEGUNG
Es gibt keinen Faktor, also können wir die Zielvariablen darstellen als

$$Y_\iota = \phi(\varepsilon_\iota) = \phi_\mu + \phi_\varepsilon(\varepsilon_\iota),\ \iota = 1, \dots, N. \tag{4.1}$$

In Vektorschreibweise lautet diese Gleichung

$$\mathbf{Y} = \boldsymbol{\phi}_\mu + \boldsymbol{\phi}_\varepsilon\ .$$

Es gilt die Reparametrisierungsbedingung (R1):

$$\mathrm{E}(\phi_\varepsilon(\varepsilon_\iota)) = \int \phi_\varepsilon(x) dG_\varepsilon(x) = 0\ . \tag{4.2}$$

Somit folgt

$$\phi_\mu = \int \phi(x)\ dG_\varepsilon(x)\ .$$

KLASSISCHE FORMULIERUNG DES MODELLS
Die klassische Formulierung des Ein-Stichproben-Falls ist gegeben durch

$$Y_\iota = \mu + \varepsilon_\iota,\ \iota = 1, \dots, N \quad \text{mit } \mathrm{E}(\varepsilon_\iota) = 0.$$

Dies ist äquivalent zu den von uns angegeben Modellgleichungen (4.1) und (4.2).

ERWARTETE WERTE UND KOVARIANZSTRUKTUR
Es gilt: $\mathrm{E}(Y_\iota) = \phi_\mu, \iota = 1, \dots, N$, und $\mathrm{Cov}(Y_\iota, Y_{\iota'}) = \begin{cases} \sigma_\varepsilon^2, & \iota = \iota' \\ 0, & \text{sonst.} \end{cases}$

HYPOTHESE
$H_0^\phi(\mu) : \phi_\mu = \mu_0$ mit $\mu_0 \in \mathbb{R}$ fest.

TESTSTATISTIK

Satz 4.1 (Gaußtest, Ein-Stichproben-t-Test) *Sei $Y_\iota = \phi(\varepsilon_\iota) = \phi_\mu + \phi_\varepsilon(\varepsilon_\iota),\ \iota = 1, \dots, N$, eine Stichprobe der Größe N, die der Modellgleichung (M1) und Annahme (A3) genügt.*
Dann ist, unter der Hypothese $H_0^\phi(\mu) : \phi_\mu = \mu_0$,

a) die Teststatistik $T_N^G = \sqrt{N}\frac{\bar{Y}-\mu_0}{\sigma_\varepsilon}$ standardnormalverteilt,

b) die Teststatistik $T_N^t = \sqrt{N}\frac{\bar{Y}-\mu_0}{\hat{\sigma}_\varepsilon}$ mit $\hat{\sigma}_\varepsilon^2 = \frac{1}{N-1}\sum\limits_{\iota=1}^{N}(Y_\iota - \bar{Y})^2$ t_{N-1}-verteilt.

4.1.2 Ein fester Faktor

Wir betrachten ein balanciertes Design mit einem Faktor a, der l Stufen besitzt. Es ist dann $N = l \cdot n$ mit einem $n \in \mathbb{N}$. Der stochastische Prozess $(a_\iota)_{\iota \in I},\ I = \{1, \dots, N\}$, nimmt also jeden der Werte aus $E = \{1, \dots, l\}$ genau n-mal an.

Bei diesem Design geben wir beispielhaft ausführliche Herleitungen der Teststatistik an, um zu demonstrieren, in welcher Weise die Hoeffding-Zerlegung dabei verwendet werden kann.

MODELL, ZERLEGUNG
Die Zielvariablen sind

$$Y_\iota = Y_{\tilde{a}_\iota, n_\iota} = \phi(a_\iota, \varepsilon_\iota) = \phi_\mu + \phi_A(a_\iota) + \phi_{A,\varepsilon}(a_\iota, \varepsilon_\iota) + \phi_\varepsilon(\varepsilon_\iota),\ \iota = 1, \dots, N.$$

Mithilfe des in Abschnitt 3.1 definierten sortierten Vektors $\mathbf{Y}$ und analog definierter Vektoren $\boldsymbol{\phi}_\mu,\ \boldsymbol{\phi}_A,\ \boldsymbol{\phi}_{A,\varepsilon}$ und $\boldsymbol{\phi}_\varepsilon$ lässt sich dies auch vektoriell schreiben als

$$\mathbf{Y} = \boldsymbol{\phi}_\mu + \boldsymbol{\phi}_A + \boldsymbol{\phi}_{A,\varepsilon} + \boldsymbol{\phi}_\varepsilon\ .$$

Unter den Annahmen (A1), (A2) und (A3) gilt:

$$Y_\iota = \phi_\mu + \phi_A(a_\iota) + \phi_\varepsilon(\varepsilon_\iota) \quad \text{mit unabhängigen Zufallsvariablen } \phi_A(a_\iota) \text{ und } \phi_\varepsilon(\varepsilon_\iota)\ ,$$
$$\text{dabei ist } \ \phi_\varepsilon(\varepsilon_\iota) \sim \mathrm{N}(0, \sigma_\varepsilon^2)\ .$$

Außerdem gelten die folgenden Reparametrisierungsbedingungen (R1) und (R2):

$$(a) \qquad \mathrm{E}(\phi_\varepsilon(\varepsilon_\iota)) = \int \phi_\varepsilon(x) dG_\varepsilon(x) = 0\ ,$$

$$(b) \qquad \mathrm{E}_A(\phi_A(a_\iota)) = \int \phi_A(x) dG_A(x) = \frac{1}{l} \sum_{j=1}^{l} \alpha_j = 0\ .$$

Wir erhalten somit

$$\begin{aligned} \phi_\mu &= \int\int \phi(x_1, x_2)\ dG_a(x_1) dG_\varepsilon(x_2) \quad \text{und} \\ \phi_A(x_1) &= \int \phi(x_1, x_2)\ dG_\varepsilon(x_2) - \phi_\mu\ . \end{aligned}$$

KLASSISCHE FORM
Bei ARNOLD (1981) lautet die Formulierung des Designs

$$\begin{aligned} Y_{ij} &= \mu + \alpha_i + \varepsilon_{ij}, \quad i = 1, \dots, l,\ j = 1, \dots, n \\ & \text{mit } \sum_i \alpha_i = 0 \\ & \text{und } \varepsilon_{ij} \sim \mathrm{N}(0, \sigma_\varepsilon^2)\ . \end{aligned} \tag{4.3}$$

Bei Gültigkeit der Annahmen (A1) und (A2) ist dies äquivalent zu dem von uns angegebenen Modell. Allerdings brauchen wir die Bedingung (4.3) nicht zu fordern, da sich (R1) direkt aus der Zerlegung ergibt.

ERWARTETE WERTE UND KOVARIANZSTRUKTUR
Es gilt: $\mathrm{E}(Y_\iota | a_\iota = e_i) = \phi_\mu + \phi_A(e_i), \iota = 1, \ldots, N$, und
$\mathrm{Cov}(Y_\iota, Y_{\iota'} | a_\iota = e_i, a_{\iota'} = e_i') = \mathrm{Cov}(\phi_\varepsilon(\varepsilon_\iota), \phi_\varepsilon(\varepsilon_{\iota'})) = \begin{cases} \sigma_\varepsilon^2, & \iota = \iota' \\ 0, & \text{sonst.} \end{cases}$

HYPOTHESE
$H_0^\phi(a) : \phi_A \equiv 0$.

HERLEITUNG EINER TESTSTATISTIK
Unter der Hypothese $H_0^\phi(a)$ und der Annahme (A3) gilt:

$$Y_\iota = \phi_\mu + \phi_\varepsilon(\varepsilon_\iota) \overset{u.i.v.}{\sim} \mathrm{N}(\phi_\mu, \sigma_\varepsilon^2) \ (\iota \in I) \ ,$$

bzw. mit der oben verwendeten vektoriellen Schreibweise:

$$\mathbf{Y} \sim \mathrm{N}(\boldsymbol{\phi}_\mu, \sigma_\varepsilon \cdot \mathbf{I}_N) \ .$$

Eine Teststatistik für die Hypothese $H_0^\phi(a)$ erhält man nun folgendermaßen:

$$\begin{aligned} Q_Z &= \mathbf{Y}'(\mathbf{P}_l \otimes \frac{1}{n}\mathbf{J}_n)\mathbf{Y} \ = \ \sum_{j=1}^{l} n \cdot \hat{\phi}_A(e_j)^2 \\ &= \sum_{j=1}^{l} n \cdot (\bar{Y}(e_j) - \bar{Y})^2 \sim \sigma_\varepsilon^2 \cdot \chi_{l-1}^2 \ , \quad \text{d.h. } \mathrm{E}\big(\frac{1}{l-1} Q_Z\big) = \sigma_\varepsilon^2 \ , \\ &\quad \text{(mit Satz A.2)} \\ Q_N &= \mathbf{Y}'(\mathbf{I}_l \otimes \mathbf{P}_n)\mathbf{Y} \\ &= \sum_{j=1}^{l} \sum_{\substack{\iota=1 \\ a_\iota = e_j}}^{N} (Y_\iota - \bar{Y}(e_j))^2 \sim \sigma_\varepsilon^2 \cdot \chi_{N-l}^2 \ , \quad \text{d.h. } \mathrm{E}\big(\frac{1}{N-l} Q_N\big) = \sigma_\varepsilon^2 \ . \\ &\quad \text{(mit Satz A.2)} \end{aligned}$$

Wegen $(\mathbf{P}_l \otimes \frac{1}{n}\mathbf{J}_n) \cdot (\sigma_\varepsilon \cdot \mathbf{I}_N) \cdot (\mathbf{I}_l \otimes \mathbf{P}_n) = 0$ sind Q_Z und Q_N mit Satz A.5 unabhängig, so dass gilt:

$$F = \frac{\frac{1}{l-1} Q_Z}{\frac{1}{N-l} Q_N} \sim F_{l-1, N-l} \ .$$

Wir formulieren diese Aussage als Satz:

Satz 4.2 (F-Test bei einem festen Faktor) *Seien $Y_\iota = \phi(a_\iota, \varepsilon_\iota)$, $\iota = 1, \ldots, N$, die Zielvariablen in einem balancierten Linearen Modell gemäß (M1) mit einem festen Faktor a mit l Stufen*

$e_1, \ldots, e_l$. Außerdem gelten die Annahmen (A1), (A2) und (A3). Dann ist, unter der Hypothese $H_0^\phi(a) : \phi_A \equiv 0$, die Teststatistik

$$F = \frac{\frac{1}{l-1} Q_Z}{\frac{1}{N-l} Q_N} \quad \textit{mit } Q_Z = \sum_{j=1}^{l} n \cdot (\bar{Y}(e_j) - \bar{Y})^2,\ Q_N = \sum_{j=1}^{l} \sum_{\substack{\iota=1 \\ a_\iota = e_j}}^{N} (Y_\iota - \bar{Y}(e_j))^2$$

verteilt nach $F_{l-1,N-l}$.
Dabei sei $\bar{Y}(e_j) := \frac{1}{n} \sum\limits_{\substack{\iota=1 \\ a_\iota = e_j}}^{N} Y_\iota$ und $\bar{Y} := \frac{1}{l} \sum\limits_{j=1}^{l} \bar{Y}(e_j)$.
Mit den auf S.67 eingeführten vereinfachenden Schreibweisen sind

$$Q_Z = \sum_{j=1}^{l} n \cdot (\bar{Y}_{j\cdot} - \bar{Y}_{\cdot\cdot})^2,\ Q_N = \sum_{j=1}^{l} \sum_{k=1}^{N} (Y_{j,k} - \bar{Y}_{j\cdot})^2 .$$

ALTERNATIVE HERLEITUNG DER TESTSTATISTIK
Die Quadratformen Q_Z und Q_N lassen sich wegen der Zerlegung von $\mathbf{Y}$ ($\mathbf{Y} = \boldsymbol{\phi}_\mu + \boldsymbol{\phi}_\varepsilon$) unter Hypothese auch folgendermaßen schreiben:

$$\begin{aligned}
Q_Z &= \mathbf{Y}'(\mathbf{P}_l \otimes \frac{1}{n}\mathbf{J}_n)\mathbf{Y} \ = \ (\boldsymbol{\phi}_\mu + \boldsymbol{\phi}_\varepsilon)'(\mathbf{P}_l \otimes \frac{1}{n}\mathbf{J}_n)(\boldsymbol{\phi}_\mu + \boldsymbol{\phi}_\varepsilon) \\
&= \underbrace{\boldsymbol{\phi}_\mu{}'(\mathbf{P}_l \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\mu}_{=0} + \underbrace{2\boldsymbol{\phi}_\mu{}'(\mathbf{P}_l \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon}_{=0} + \boldsymbol{\phi}_\varepsilon{}'(\mathbf{P}_l \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon \\
&= \boldsymbol{\phi}_\varepsilon{}'(\mathbf{P}_l \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon \ , \\
Q_N &= \mathbf{Y}'(\mathbf{I}_l \otimes \mathbf{P}_n)\mathbf{Y} \ = \ (\boldsymbol{\phi}_\mu + \boldsymbol{\phi}_\varepsilon)'(\mathbf{I}_l \otimes \mathbf{P}_n)(\boldsymbol{\phi}_\mu + \boldsymbol{\phi}_\varepsilon) \\
&= \underbrace{\boldsymbol{\phi}_\mu{}'(\mathbf{I}_l \otimes \mathbf{P}_n)\boldsymbol{\phi}_\mu}_{=0} + \underbrace{2\boldsymbol{\phi}_\mu{}'(\mathbf{I}_l \otimes \mathbf{P}_n)\boldsymbol{\phi}_\varepsilon}_{=0} + \boldsymbol{\phi}_\varepsilon{}'(\mathbf{I}_l \otimes \mathbf{P}_n)\boldsymbol{\phi}_\varepsilon \\
&= \boldsymbol{\phi}_\varepsilon{}'(\mathbf{I}_l \otimes \mathbf{P}_n)\boldsymbol{\phi}_\varepsilon \ .
\end{aligned}$$

Mit $\boldsymbol{\phi}_\varepsilon \sim \mathrm{N}(\mathbf{0}, \sigma_\varepsilon^2 \cdot \mathbf{I}_N)$ und Satz A.2 ist $Q_Z \sim \sigma_\varepsilon^2 \cdot \chi^2_{l-1}$ und $Q_N \sim \sigma_\varepsilon^2 \cdot \chi^2_{N-l}$.
Die Unabhängigkeit von Q_Z und Q_N folgt wieder aus Satz A.5.

VERHALTEN DER TESTSTATISTIK UNTER ALTERNATIVE
Unter Alternative ist $Y_\iota = \phi_\mu + \phi_A(a_\iota) + \phi_\varepsilon(\varepsilon_\iota)$ mit $\phi_A \not\equiv 0$ bzw. $\mathbf{Y} = \boldsymbol{\phi}_\mu + \boldsymbol{\phi}_A + \boldsymbol{\phi}_\varepsilon$.
Damit ist dann

$$\begin{aligned}
Q_Z &= \mathbf{Y}'(\mathbf{P}_l \otimes \frac{1}{n}\mathbf{J}_n)\mathbf{Y} \ = \ (\boldsymbol{\phi}_\mu + \boldsymbol{\phi}_A + \boldsymbol{\phi}_\varepsilon)'(\mathbf{P}_l \otimes \frac{1}{n}\mathbf{J}_n)(\boldsymbol{\phi}_\mu + \boldsymbol{\phi}_A + \boldsymbol{\phi}_\varepsilon) \\
&= \underbrace{\boldsymbol{\phi}_\mu{}'(\mathbf{P}_l \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\mu}_{=0} + \underbrace{2\boldsymbol{\phi}_\mu{}'(\mathbf{P}_l \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_A}_{=0} + \underbrace{2\boldsymbol{\phi}_\mu{}'(\mathbf{P}_l \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon}_{=0} \\
&\quad + \underbrace{\boldsymbol{\phi}_A{}'(\mathbf{P}_l \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_A}_{=:A} + \underbrace{2\boldsymbol{\phi}_A{}'(\mathbf{P}_l \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon}_{=:B} + \underbrace{\boldsymbol{\phi}_\varepsilon{}'(\mathbf{P}_l \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon}_{=:C} \ .
\end{aligned}$$

Es gilt:

$$\begin{aligned}
A &= \sum_{j=1}^{l} n \cdot \phi_A(e_j)^2 > 0 \text{ (wegen } \phi_A \not\equiv 0) , \\
B &= \sum_{j=1}^{l} n \cdot \phi_A(e_j) \cdot \big(\frac{1}{n} \sum_{\substack{\iota=1 \\ a_\iota = e_j}}^{N} \phi_\varepsilon(\varepsilon_\iota)\big) \text{ mit } \mathrm{E}(B) = 0 = \mathrm{E}(B|a_1 = e_{i_1}, \ldots , a_N = e_{i_N}) , \\
C &\sim \sigma_\varepsilon^2 \cdot \chi_{l-1}^2 .
\end{aligned}$$

In ähnlicher Weise ergibt sich für Q_N:

$$\begin{aligned}
Q_N &= \mathbf{Y}'(\mathbf{I}_l \otimes \mathbf{P}_n)\mathbf{Y} \; = \; (\boldsymbol{\phi}_\mu + \boldsymbol{\phi}_A + \boldsymbol{\phi}_\varepsilon)'(\mathbf{I}_l \otimes \mathbf{P}_n)(\boldsymbol{\phi}_\mu + \boldsymbol{\phi}_A + \boldsymbol{\phi}_\varepsilon) \\
&= \underbrace{\boldsymbol{\phi}_\mu{}'(\mathbf{I}_l \otimes \mathbf{P}_n)\boldsymbol{\phi}_\mu}_{=0} + \underbrace{2\boldsymbol{\phi}_\mu{}'(\mathbf{I}_l \otimes \mathbf{P}_n)\boldsymbol{\phi}_A}_{=0} + \underbrace{2\boldsymbol{\phi}_\mu{}'(\mathbf{I}_l \otimes \mathbf{P}_n)\boldsymbol{\phi}_\varepsilon}_{=0} \\
&\quad + \underbrace{\boldsymbol{\phi}_A{}'(\mathbf{I}_l \otimes \mathbf{P}_n)\boldsymbol{\phi}_A}_{=0} + \underbrace{2\boldsymbol{\phi}_A{}'(\mathbf{I}_l \otimes \mathbf{P}_n)\boldsymbol{\phi}_\varepsilon}_{=0} + \boldsymbol{\phi}_\varepsilon{}'(\mathbf{I}_l \otimes \mathbf{P}_n)\boldsymbol{\phi}_\varepsilon \\
&= \boldsymbol{\phi}_\varepsilon{}'(\mathbf{I}_l \otimes \mathbf{P}_n)\boldsymbol{\phi}_\varepsilon .
\end{aligned}$$

Q_N bleibt also unter Alternative gleich, d.h. $Q_N \sim \sigma_\varepsilon^2 \cdot \chi_{N-l}^2$.
Die auf den Faktor bedingten Erwartungswerte der durch die Freiheitsgrade dividierten Quadratformen Q_Z und Q_N lassen sich direkt ablesen als:

$$\begin{aligned}
\mathrm{E}(\frac{1}{l-1} \cdot Q_Z|a_1 = e_{i_1}, \ldots , a_N = e_{i_N}) &= \sigma_\varepsilon^2 + \frac{1}{l-1} \sum_{j=1}^{l} n \cdot \phi_A(e_j)^2 \\
&= \sigma_\varepsilon^2 + n \cdot \underbrace{\frac{1}{l-1} \sum_{j=1}^{l} \alpha_j^2}_{=:\sigma_\alpha^2} \; = \; \sigma_\varepsilon^2 + n \cdot \sigma_\alpha^2 , \\
\mathrm{E}(\frac{1}{l-1} \cdot Q_N|a_1 = e_{i_1}, \ldots , a_N = e_{i_N}) &= \sigma_\varepsilon^2 .
\end{aligned}$$

Die Hypothese $H_0^\phi(a) : \phi_A \equiv 0$ impliziert die Gleichheit der beiden bedingten Erwartungswerte, da dann $\sigma_\alpha^2 = 0$ gilt. Die Verteilung von Q_Z unter Hypothese ergibt sich wieder, wie oben, als $Q_Z \sim \sigma_\varepsilon^2 \cdot \chi_{l-1}^2$, da unter Hypothese die beiden Ausdrücke A und B gleich Null sind.
Die hergeleiteten Aussagen lassen sich verallgemeinern auf ein unbalanciertes Design:

Satz 4.3 (F-Test bei einem festen Faktor, unbalanciertes Design)
Seien $Y_\iota = \phi(a_\iota, \varepsilon_\iota)$, $\iota = 1, \ldots , N$, die Zielvariablen in einem unbalancierten Linearen Modell gemäß (M1) mit einem festen Faktor a mit l Stufen $e_1, \ldots , e_l$. Die Stufe e_j werde n_j-mal angenommen, d.h. $\mathrm{card}\{\iota : a_\iota = e_j\} = n_j$. Außerdem gelten die Annahmen (A1), (A2) und (A3). Dann ist, unter der Hypothese $H_0^\phi(a) : \phi_A \equiv 0$, die Teststatistik

$$F = \frac{\frac{1}{l-1} Q_Z}{\frac{1}{N-l} Q_N} \quad \textit{mit } \; Q_Z = \sum_{j=1}^{l} n_j \cdot (\bar{Y}(e_j) - \bar{Y})^2 = \sum_{j=1}^{l} n_j \cdot (\bar{Y}_{j\cdot} - \bar{Y}_{\cdot\cdot})^2$$

$$\textit{und}\ \ Q_N = \sum_{j=1}^{l} \sum_{\substack{\iota=1 \\ a_\iota = e_j}}^{N} (Y_\iota - \bar{Y}(e_j))^2 = \sum_{j=1}^{l} \sum_{k=1}^{n_j} (Y_{j,k} - \bar{Y}_{j\cdot})^2$$

verteilt nach $F_{l-1,N-l}$.

4.1.3 Ein zufälliger Faktor

Wir betrachten - analog zu 4.1.2 - ein balanciertes Design mit einem zufälligen Faktor a, der l Stufen besitzt. Das bedeutet, dass aus der Menge $E \cap \mathbb{N} = \{1, \dots, M\}$ eine l-elementige Teilmenge $\tilde{E}$ ausgewählt wird. Es ist $N = l \cdot n$ mit einem $n \in \mathbb{N}$, so dass der stochastische Prozess $(\lceil a_\iota \rceil)_{\iota \in I},\ I = \{1, \dots, N\}$ jeden der Werte aus $\tilde{E}$ genau n-mal annimmt.
Dem Charakter eines zufälligen Faktors gemäß lassen sich die genauen Ausprägungen des Prozesses $(a_\iota)_{\iota \in I}$ nicht kontrollieren. Es gilt lediglich: $\lceil a_\iota \rceil = \lceil a_{\iota'} \rceil \Rightarrow a_\iota = a_{\iota'}$.

MODELL, ZERLEGUNG

$$\begin{aligned} Y_\iota &= \phi(a_\iota, \varepsilon_\iota) = \phi_\mu + \phi_A(a_\iota) + \phi_{A,\varepsilon}(a_\iota, \varepsilon_\iota) + \phi_\varepsilon(\varepsilon_\iota),\ \iota = 1, \dots, N, \quad \text{bzw.} \\ \mathbf{Y} &= \boldsymbol{\phi}_\mu + \boldsymbol{\phi}_A + \boldsymbol{\phi}_{A,\varepsilon} + \boldsymbol{\phi}_\varepsilon\ . \end{aligned}$$

Unter den Annahmen (A1) und (A2) gilt:

$$Y_\iota = \phi_\mu + \phi_A(a_\iota) + \phi_\varepsilon(\varepsilon_\iota) \quad \text{mit unabhängigen Zufallsvariablen } \phi_\varepsilon(\varepsilon_\iota),\ \iota = 1, \dots, N, \tag{4.4}$$

$$\text{und:} \quad a_\iota \neq a_{\iota'} \Rightarrow \phi_A(a_\iota) \text{ und } \phi_A(a_{\iota'}) \text{ sind unabhängig.} \tag{4.5}$$

Die beiden Familien $(\phi_A(a_\iota),\ \iota \in I)$ und $(\phi_\varepsilon(\varepsilon_\iota),\ \iota \in I)$ sind voneinander unabhängig.
Mit (A3) und (A4) gilt weiterhin

$$\phi_A(a_\iota) \sim \mathrm{N}(0, \sigma_A^2),\ \phi_\varepsilon(\varepsilon_\iota) \sim \mathrm{N}(0, \sigma_\varepsilon^2),\ \iota \in I\ . \tag{4.6}$$

In Vektorschreibweise bedeutet dies mit (4.4) und (4.5):

$$\boldsymbol{\phi}_A \sim \mathrm{N}(\mathbf{0}, \sigma_A^2 \cdot \mathbf{I}_l \otimes \mathbf{J}_n) \quad \text{und} \quad \boldsymbol{\phi}_\varepsilon \sim \mathrm{N}(\mathbf{0}, \sigma_\varepsilon^2 \cdot \mathbf{I}_N)\ .$$

Die Gültigkeit der folgenden Reparametrisierungsbedingungen ist wegen (4.6) klar und braucht nicht extra gefordert zu werden:

$$\begin{aligned} &(a) \quad \mathrm{E}(\phi_\varepsilon(\varepsilon_\iota)) = \int \phi_\varepsilon(x) dG_\varepsilon(x) = 0\,, \\ &(b) \quad \mathrm{E}_A(\phi_A(a_\iota)) = \int \phi_A(x) dG_A(x) = 0\,. \end{aligned}$$

Wir erhalten

$$\phi_\mu = \int\int \phi(x_1, x_2)\ dG_a(x_1) dG_\varepsilon(x_2) \quad \text{und}$$

$$\phi_A(x_1) \;=\; \int \phi(x_1,x_2)\,dG_\varepsilon(x_2) - \phi_\mu\,.$$

KLASSISCHE FORM

Bei ARNOLD (1981) lautet die Darstellung der Zielvariablen

$$\begin{aligned} Y_{ij} \;&=\; \mu + a_i + \varepsilon_{ij}, \quad i = 1,\dots,l,\ j = 1,\dots,n \\ &\quad \text{mit unabhängigen } a_i \sim \mathrm{N}(0,\sigma_a^2) \text{ und } \varepsilon_{ij} \sim \mathrm{N}(0,\sigma_\varepsilon^2)\,. \end{aligned}$$

Dies ist äquivalent zu dem von uns angegebenen Modell.

ERWARTETE WERTE UND KOVARIANZSTRUKTUR

Es gilt: $\mathrm{E}(Y_\iota) = \phi_\mu, \iota = 1,\dots,N$, und

$$\mathrm{Cov}(Y_\iota, Y_{\iota'}) = \mathrm{Cov}(\phi_A(a_\iota),\phi_A(a_{\iota'})) + \mathrm{Cov}(\phi_\varepsilon(\varepsilon_\iota),\phi_\varepsilon(\varepsilon_{\iota'})) = \begin{cases} \sigma_A^2, & a_\iota = a_{\iota'} \\ 0, & \text{sonst} \end{cases} + \begin{cases} \sigma_\varepsilon^2, & \iota = \iota' \\ 0, & \text{sonst.} \end{cases}$$

HYPOTHESE

$H_0^\phi(a) : \phi_A \equiv 0\,.$

TESTSTATISTIK

Wir betrachten dieselben Quadratformen wie in 4.1.2. Die vorgenommene Zerlegung bietet hier allerdings keinen technischen Vorteil, da die Verteilung der Teststatistik am Einfachsten über eine Spektralzerlegung der Kovarianzmatrix von $\mathbf{Y}$ zu berechnen ist:

$$\begin{aligned} Q_Z = \mathbf{Y}'(\mathbf{P}_l \otimes \tfrac{1}{n}\mathbf{J}_n)\mathbf{Y} \;&=\; (\boldsymbol{\phi}_\mu + \boldsymbol{\phi}_A + \boldsymbol{\phi}_\varepsilon)'(\mathbf{P}_l \otimes \tfrac{1}{n}\mathbf{J}_n)(\boldsymbol{\phi}_\mu + \boldsymbol{\phi}_A + \boldsymbol{\phi}_\varepsilon) \\ &= \underbrace{\boldsymbol{\phi}_\mu{}'(\mathbf{P}_l \otimes \tfrac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\mu}_{=0} + \underbrace{2\boldsymbol{\phi}_\mu{}'(\mathbf{P}_l \otimes \tfrac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_A}_{=0} + \underbrace{2\boldsymbol{\phi}_\mu{}'(\mathbf{P}_l \otimes \tfrac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon}_{=0} \\ &\quad + (\boldsymbol{\phi}_A + \boldsymbol{\phi}_\varepsilon)'(\mathbf{P}_l \otimes \tfrac{1}{n}\mathbf{J}_n)(\boldsymbol{\phi}_A + \boldsymbol{\phi}_\varepsilon)\,. \end{aligned}$$

Wegen

$$\begin{aligned} \mathrm{Var}(\boldsymbol{\phi}_A + \boldsymbol{\phi}_\varepsilon) &= \sigma_A^2 \cdot \mathbf{I}_l \otimes \mathbf{J}_n + \sigma_\varepsilon^2 \cdot \mathbf{I}_l \otimes \mathbf{I}_n \\ &= (n\cdot\sigma_A^2 + \sigma_\varepsilon^2)\cdot \tfrac{1}{l}\mathbf{J}_l \otimes \tfrac{1}{n}\mathbf{J}_n + (n\cdot\sigma_A^2 + \sigma_\varepsilon^2)\cdot \mathbf{P}_l \otimes \tfrac{1}{n}\mathbf{J}_n + \sigma_\varepsilon^2 \cdot \mathbf{I}_l \otimes \mathbf{P}_n \\ \text{gilt} \quad & (\mathbf{P}_l \otimes \tfrac{1}{n}\mathbf{J}_n)\cdot \mathrm{Var}(\boldsymbol{\phi}_A + \boldsymbol{\phi}_\varepsilon) = (n\cdot\sigma_A^2 + \sigma_\varepsilon^2)\cdot \mathbf{P}_l \otimes \tfrac{1}{n}\mathbf{J}_n \\ \text{und daher} \quad & Q_Z \sim (n\cdot\sigma_A^2 + \sigma_\varepsilon^2)\cdot\chi^2_{l-1}\,. \end{aligned}$$

Analog erhält man $Q_N \sim \sigma_\varepsilon^2 \cdot \chi^2_{N-l}$.

Satz 4.4 (F-Test bei einem zufälligen Faktor) *Seien $Y_\iota = \phi(a_\iota,\varepsilon_\iota),\ \iota = 1,\dots,N$, die Zielvariablen in einem balancierten Linearen Modell gemäß (M1) mit einem zufälligen Faktor a, so dass $(\lceil a_\iota \rceil)_{\iota\in I},\ I = \{1,\dots,N\}$ jeden der Werte aus $\tilde{E} = \{1,\dots,l\}$ genau n-mal annimmt.*

Außerdem gelten die Annahmen (A1), (A2), (A3) und (A4). Dann ist, unter der Hypothese $H_0^{\phi}(a) : \phi_A \equiv 0$, die Teststatistik

$$F = \frac{\frac{1}{l-1}Q_Z}{\frac{1}{N-l}Q_N} \quad \textit{mit} \quad Q_Z = \sum_{j=1}^{l} n \cdot (\bar{Y}(e_j) - \bar{Y})^2 = \sum_{j=1}^{l} n \cdot (\bar{Y}_{j\cdot} - \bar{Y}_{\cdot\cdot})^2$$

$$\textit{und} \ \ Q_N = \sum_{j=1}^{l} \sum_{\substack{\iota=1 \\ a_\iota = e_j}}^{N} (Y_\iota - \bar{Y}(e_j))^2 = \sum_{j=1}^{l} n \cdot (\bar{Y}_{j\cdot} - \bar{Y}_{\cdot\cdot})^2 = \sum_{j=1}^{l} \sum_{k=1}^{N} (Y_{j,k} - \bar{Y}_{j\cdot})^2$$

verteilt nach $F_{l-1,N-l}$.

4.1.4 Eine Kovariable

Wir betrachten eine Stichprobe der Größe N und nehmen hier an, dass der kovariate Faktor a genau N Stufen besitzt, die aus dem M-elementigen Zustandsraum E stammen, und die zu einer Menge $\tilde{E} \subset E$ zusammengefasst werden.

MODELL, ZERLEGUNG
Die Zielvariablen sind

$$Y_\iota = \phi(a_\iota, \varepsilon_\iota) \ = \ \phi_\mu + \phi_A(a_\iota) + \phi_{A,\varepsilon}(a_\iota, \varepsilon_\iota) + \phi_\varepsilon(\varepsilon_\iota) \ ,$$

bzw. in vektorieller Schreibweise

$$\mathbf{Y} = \boldsymbol{\phi}_\mu + \boldsymbol{\phi}_A + \boldsymbol{\phi}_{A,\varepsilon} + \boldsymbol{\phi}_\varepsilon \ .$$

Unter den Annahmen (A1), (A2) und (A5) gilt:

$$Y_\iota = \xi_0 + \xi_1 a_\iota + \phi_\varepsilon(\varepsilon_\iota) \quad \text{mit unabhängigen Zufallsvariablen } a_\iota \text{ und } \phi_\varepsilon(\varepsilon_\iota).$$

Es gelten die folgenden Reparametrisierungsbedingungen:

$$(a) \quad \mathrm{E}(\phi_\varepsilon(\varepsilon_\iota)) = \int \phi_\varepsilon(x) dG_\varepsilon(x) = 0 \ ,$$

$$(b) \quad \mathrm{E}_A(a_\iota) = \int x dG_A(x) = 0 \ .$$

KLASSISCHE FORM
Das lineare Regressionsmodell mit einer Regressorvariable wird bei SEBER (1977) als

$$Y_i = \beta_0 + \beta_1 x_{i1} + \varepsilon_i, \ i = 1, \dots, n,$$

bzw. $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ formuliert.

Indem wir $\mathbf{X} = (\mathbb{1}_N \vdots \mathbf{a})$ mit $\mathbf{a} = (a_1, \dots, a_N)'$ und $\boldsymbol{\xi} = (\xi_0, \xi_1)'$ setzen, erhalten wir eine dazu äquivalente Vektordarstellung $\mathbf{Y} = \mathbf{X}\boldsymbol{\xi} + \boldsymbol{\phi}_\varepsilon$.

ERWARTETE WERTE UND KOVARIANZSTRUKTUR
Es gilt: $\mathrm{E}(Y_\iota|a_\iota = e_i) = \phi_\mu + \phi_A(e_i), \iota = 1, \ldots, N$, und
$\mathrm{Cov}(Y_\iota, Y_{\iota'}|a_\iota = e_i, a_{\iota'} = e_i') = \mathrm{Cov}(\phi_\varepsilon(\varepsilon_\iota), \phi_\varepsilon(\varepsilon_{\iota'})) = \begin{cases} \sigma_\varepsilon^2, & \iota = \iota' \\ 0, & \text{sonst.} \end{cases}$

HYPOTHESEN
Als Hypothesen werden formuliert:
$H_0^\phi(\xi_0) : \xi_0 = 0$,
$H_0^\phi(\xi_1) : \xi_1 = 0$.
TESTSTATISTIK
Wir geben hier nur das Resultat an:

Satz 4.5 *Sei $Y_\iota = \xi_0 + \xi_1 a_\iota + \phi_\varepsilon(\varepsilon_\iota),\ \iota = 1,\ \ldots, N$, mit $\phi_\varepsilon(\varepsilon_\iota) \sim \mathrm{N}(0, \sigma_\varepsilon^2)$ ein lineares Regressionsmodell gemäß (M1) unter den Annahmen (A1), (A2), (A3) und (A5). Dann gilt, bedingt auf $\mathbf{a} = (a_1, \ldots, a_N)'$:*

$$\textit{Unter } H_0 : \xi_0 = 0 \textit{ ist } \frac{(N-2)\hat{\xi}_0^2 N S_a}{\sum_{\iota=1}^N a_\iota^2 (S_Y - \hat{\xi}_1^2 S_a)} \textit{ verteilt nach } F(1, N-2),$$

$$\textit{unter } H_0 : \xi_1 = 0 \textit{ ist } \frac{(N-2)\hat{\xi}_1^2 S_a}{S_Y - \hat{\xi}_1^2 S_a} \textit{ verteilt nach } F(1, N-2).$$

$$\textit{Dabei seien } S_a = \sum_{\iota=1}^N (a_\iota - \bar{a}_.)^2,\ S_Y = \sum_{\iota=1}^N (Y_\iota - \bar{Y}_.)^2,\ S_{aY} = \sum_{\iota=1}^N (a_\iota - \bar{a}_.)(Y_\iota - \bar{Y}_.),$$

$$\hat{\xi}_1 = \frac{S_{aY}}{S_a}, \textit{ und } \hat{\xi}_0 = \bar{Y}_. - \hat{\xi}_1 \bar{a}_. \ .$$

4.1.5 Zwei feste Faktoren (vollständig gekreuzt)

Wir betrachten hier ein balanciertes Design mit zwei Faktoren $a^{(1)}$ und $a^{(2)}$, die l_1 bzw. l_2 Stufen besitzen, außerdem gebe es n Wiederholungen pro Faktorstufenkombination. Es sei $N = l_1 \cdot l_2 \cdot n$. Der stochastische Prozess $(a_\iota^{(1)}, a_\iota^{(2)})_{\iota \in I},\ I = \{1, \ldots, N\}$, nimmt also jeden Wert aus $E = E^{(1)} \times E^{(2)} = \{1, \ldots, l_1\} \times \{1, \ldots, l_2\}$ genau n-mal an.

MODELL, ZERLEGUNG
Die Zielvariablen sind

$$\begin{aligned} Y_\iota =& Y_{a_\iota^{(1)}, a_\iota^{(2)}, n_\iota} = \phi(a_\iota^{(1)}, a_\iota^{(2)}, \varepsilon_\iota) \\ =& \phi_\mu + \phi_{A^{(1)}}(a_\iota^{(1)}) + \phi_{A^{(2)}}(a_\iota^{(2)}) + \phi_{A^{(1)},A^{(2)}}(a_\iota^{(1)}, a_\iota^{(2)}) + \phi_{A^{(1)},\varepsilon}(a_\iota^{(1)}, \varepsilon_\iota) \\ &+ \phi_{A^{(2)},\varepsilon}(a_\iota^{(2)}, \varepsilon_\iota) + \phi_{A^{(1)},A^{(2)},\varepsilon}(a_\iota^{(1)}, a_\iota^{(2)}, \varepsilon_\iota) + \phi_\varepsilon(\varepsilon_\iota),\ \iota = 1, \ldots, N. \end{aligned}$$

In Vektorschreibweise lautet dies

$$\mathbf{Y} = \boldsymbol{\phi}_\mu + \boldsymbol{\phi}_{A^{(1)}} + \boldsymbol{\phi}_{A^{(2)}} + \boldsymbol{\phi}_{A^{(1)},A^{(2)}} + \boldsymbol{\phi}_{A^{(1)},\varepsilon} + \boldsymbol{\phi}_{A^{(2)},\varepsilon} + \boldsymbol{\phi}_{A^{(1)},A^{(2)},\varepsilon} + \boldsymbol{\phi}_\varepsilon \ .$$

Unter den Annahmen (A1), (A2) und (A3) gilt:

$$Y_\iota = \phi_\mu + \phi_{A^{(1)}}(a_\iota^{(1)}) + \phi_{A^{(2)}}(a_\iota^{(2)}) + \phi_{A^{(1)},A^{(2)}}(a_\iota^{(1)}, a_\iota^{(2)}) + \phi_\varepsilon(\varepsilon_\iota)$$

mit unabhängigen Zufallsvariablen $\phi_{A^{(1)}}(a_\iota^{(1)}),\ \phi_{A^{(2)}}(a_\iota^{(2)}),\ \phi_{A^{(1)},A^{(2)}}(a_\iota^{(1)}, a_\iota^{(2)})$ und $\phi_\varepsilon(\varepsilon_\iota)$, dabei ist $\phi_\varepsilon(\varepsilon_\iota) \sim \mathrm{N}(0, \sigma_\varepsilon^2)$.

Es gelten die Reparametrisierungsbedingungen (R1), (R2) und (R3), das heißt:

$$(a)\quad \mathrm{E}(\phi_\varepsilon(\varepsilon_\iota)) = \int \phi_\varepsilon(x) dG_\varepsilon(x) = 0\,,$$

$$(b)\quad \mathrm{E}_{A^{(1)}}(\phi_{A^{(1)}}(a_\iota^{(1)})) = \int \phi_{A^{(1)}}(x) dG_{A^{(1)}}(x) = \frac{1}{l_1}\sum_{i=1}^{l_1} \phi_{A^{(1)}}(i) = 0\,,$$

$$(c)\quad \mathrm{E}_{A^{(2)}}(\phi_{A^{(2)}}(a_\iota^{(2)})) = \int \phi_{A^{(2)}}(x) dG_{A^{(2)}}(x) = \frac{1}{l_2}\sum_{i=1}^{l_2} \phi_{A^{(2)}}(i) = 0\,,$$

$$(d)\quad \mathrm{E}_{A^{(1)}}(\phi_{A^{(1)},A^{(2)}}(a_\iota^{(1)}, a_\iota^{(2)}) | a_\iota^{(2)} = i_2) = \int \phi_{A^{(1)},A^{(2)}}(x, i_2) dG_{A^{(1)}}(x)$$
$$= \frac{1}{l_1}\sum_{i_1=1}^{l_1} \phi_{A^{(1)},A^{(2)}}(i_1, i_2) = 0\,,$$

$$(e)\quad \mathrm{E}_{A^{(2)}}(\phi_{A^{(1)},A^{(2)}}(a_\iota^{(1)}, a_\iota^{(2)}) | a_\iota^{(1)} = i_1) = \int \phi_{A^{(1)},A^{(2)}}(i_1, x) dG_{A^{(2)}}(x)$$
$$= \frac{1}{l_2}\sum_{i_2=1}^{l_2} \phi_{A^{(1)},A^{(2)}}(i_1, i_2) = 0\,.$$

KLASSISCHE FORM
ARNOLD (1981) formuliert das Design als

$$Y_{ijk} \sim \mathrm{N}(\mu_{ij}, \sigma^2), \quad i = 1, \dots, l_1,\ j = 1, \dots, l_2,\ k = 1 \dots, n, \quad \text{mit}$$
$$\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} \quad \text{und den zu fordernden Bedingungen}$$
$$\sum_i \alpha_i = 0,\ \sum_j \beta_j = 0,\ \sum_i \gamma_{ij} = 0,\ \sum_j \gamma_{ij} = 0\,. \tag{4.7}$$

Dies ist wiederum, bei Gültigkeit der Annahmen (A1), (A2) und (A3), äquivalent zu dem oben von uns angegebenen Modell. (4.7) entspricht dabei den unter (b)-(e) angegebenen Reparametrisierungsbedingungen, die sich bei uns direkt aus der Zerlegung ergeben.

ERWARTETE WERTE UND KOVARIANZSTRUKTUR
Es gilt:

$$\mathrm{E}(Y_\iota | a_\iota^{(1)} = i_1, a_\iota^{(2)} = i_2) = \phi_\mu + \phi_{A^{(1)}}(i_1) + \phi_{A^{(2)}}(i_2) + \phi_{A^{(1)},A^{(2)}}(i_1, i_2),\ \iota = 1, \dots, N, \text{ und}$$

$$\mathrm{Cov}(Y_\iota, Y_{\iota'} | a_\iota^{(1)} = i_1, a_\iota^{(2)} = i_2, a_{\iota'}^{(1)} = i_1', a_{\iota'}^{(2)} = i_2') = \mathrm{Cov}(\phi_\varepsilon(\varepsilon_\iota), \phi_\varepsilon(\varepsilon_{\iota'})) = \begin{cases} \sigma_\varepsilon^2, & \iota = \iota' \\ 0, & \text{sonst.} \end{cases}$$

HYPOTHESEN
Es werden die folgenden Hypothesen formuliert:
Hypothese über den ersten Faktor: $H_0^\phi(a^{(1)}) : \phi_{A^{(1)}} \equiv 0$.
Hypothese über den zweiten Faktor: $H_0^\phi(a^{(2)}) : \phi_{A^{(2)}} \equiv 0$.
Hypothese über die Wechselwirkung zwischen beiden Faktoren: $H_0^\phi(a^{(1)}, a^{(2)}) : \phi_{A^{(1)},A^{(2)}} \equiv 0$.

TESTSTATISTIK
Seien

$$Q_{Z_1} = \mathbf{Y}'(\mathbf{P}_{l_1} \otimes \frac{1}{l_2}\mathbf{J}_{l_2} \otimes \frac{1}{n}\mathbf{J}_n)\mathbf{Y}, \quad Q_{Z_2} = \mathbf{Y}'(\frac{1}{l_1}\mathbf{J}_{l_1} \otimes \mathbf{P}_{l_2} \otimes \frac{1}{n}\mathbf{J}_n)\mathbf{Y},$$
$$Q_{Z_{1,2}} = \mathbf{Y}'(\mathbf{P}_{l_1} \otimes \mathbf{P}_{l_2} \otimes \frac{1}{n}\mathbf{J}_n)\mathbf{Y} \text{ und } Q_N = \mathbf{Y}'(\mathbf{I}_{l_1} \otimes \mathbf{I}_{l_2} \otimes \mathbf{P}_n)\mathbf{Y} .$$

Wegen der Zerlegung von $\mathbf{Y}$ ist

$$\begin{aligned}
Q_{Z_1} &= (\boldsymbol{\phi}_\mu + \boldsymbol{\phi}_{A^{(1)}} + \boldsymbol{\phi}_{A^{(2)}} + \boldsymbol{\phi}_{A^{(1)},A^{(2)}} + \boldsymbol{\phi}_\varepsilon)'(\mathbf{P}_{l_1} \otimes \frac{1}{l_2}\mathbf{J}_{l_2} \otimes \frac{1}{n}\mathbf{J}_n) \times \\
&\qquad\qquad (\boldsymbol{\phi}_\mu + \boldsymbol{\phi}_{A^{(1)}} + \boldsymbol{\phi}_{A^{(2)}} + \boldsymbol{\phi}_{A^{(1)},A^{(2)}} + \boldsymbol{\phi}_\varepsilon) \\
&= \boldsymbol{\phi}'_{A^{(1)}}(\mathbf{P}_{l_1} \otimes \frac{1}{l_2}\mathbf{J}_{l_2} \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_{A^{(1)}} + 2\boldsymbol{\phi}'_{A^{(1)}}(\mathbf{P}_{l_1} \otimes \frac{1}{l_2}\mathbf{J}_{l_2} \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon \\
&\quad + \boldsymbol{\phi}'_\varepsilon(\mathbf{P}_{l_1} \otimes \frac{1}{l_2}\mathbf{J}_{l_2} \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon \\
&\quad \text{und entsprechend} \\
Q_{Z_2} &= \boldsymbol{\phi}'_{A^{(2)}}(\frac{1}{l_1}\mathbf{J}_{l_1} \otimes \mathbf{P}_{l_2} \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_{A^{(2)}} + 2\boldsymbol{\phi}'_{A^{(2)}}(\frac{1}{l_1}\mathbf{J}_{l_1} \otimes \mathbf{P}_{l_2} \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon \\
&\quad + \boldsymbol{\phi}'_\varepsilon(\frac{1}{l_1}\mathbf{J}_{l_1} \otimes \mathbf{P}_{l_2} \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon , \\
Q_{Z_{1,2}} &= \boldsymbol{\phi}'_{A^{(1)},A^{(2)}}(\mathbf{P}_{l_1} \otimes \mathbf{P}_{l_2} \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_{A^{(1)},A^{(2)}} + 2\boldsymbol{\phi}'_{A^{(1)},A^{(2)}}(\mathbf{P}_{l_1} \otimes \mathbf{P}_{l_2} \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon \\
&\quad + \boldsymbol{\phi}'_\varepsilon(\mathbf{P}_{l_1} \otimes \mathbf{P}_{l_2} \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon \qquad \text{sowie} \\
Q_N &= \boldsymbol{\phi}'_\varepsilon(\mathbf{I}_{l_1} \otimes \mathbf{I}_{l_2} \otimes \mathbf{P}_n)\boldsymbol{\phi}_\varepsilon .
\end{aligned}$$

Unter den jeweiligen Hypothesen vereinfachen sich diese Quadratformen, und wir erhalten die folgenden Verteilungsaussagen:

$$\begin{gathered}
\text{Unter } H_0^\phi(a^{(1)}) : \phi_{A^{(1)}} \equiv 0 \text{ ist } Q_{Z_1} = \boldsymbol{\phi}'_\varepsilon(\mathbf{P}_{l_1} \otimes \frac{1}{l_2}\mathbf{J}_{l_2} \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon \sim \sigma_\varepsilon^2 \cdot \chi^2_{l_1-1} , \\
\text{unter } H_0^\phi(a^{(2)}) : \phi_{A^{(2)}} \equiv 0 \text{ ist } Q_{Z_2} = \boldsymbol{\phi}'_\varepsilon(\frac{1}{l_1}\mathbf{J}_{l_1} \otimes \mathbf{P}_{l_2} \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon \sim \sigma_\varepsilon^2 \cdot \chi^2_{l_2-1} , \\
\text{unter } H_0^\phi(a^{(1)}, a^{(2)}) : \phi_{A^{(1)},A^{(2)}} \equiv 0 \text{ ist } Q_{Z_{1,2}} = \boldsymbol{\phi}'_\varepsilon(\mathbf{P}_{l_1} \otimes \mathbf{P}_{l_2} \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon \sim \sigma_\varepsilon^2 \cdot \chi^2_{(l_1-1)(l_2-1)} , \\
\text{außerdem ist } Q_N \sim \sigma_\varepsilon^2 \cdot \chi^2_{l_1 l_2 (n-1)} .
\end{gathered}$$

Dabei verwenden wir $\boldsymbol{\phi}_\varepsilon \sim \mathrm{N}(\mathbf{0}, \sigma_\varepsilon^2 \cdot \mathbf{I}_N)$ und Satz A.2. Unter Verwendung von Satz A.5 erhält man schließlich F-verteilte Teststatistiken:

$$\text{Unter } H_0^\phi(a^{(1)}) : \phi_{A^{(1)}} \equiv 0 \quad \text{gilt :} \qquad F_1 = \frac{\frac{1}{l_1-1} Q_{Z_1}}{\frac{1}{l_1 l_2 (n-1)} Q_N} \sim F_{l_1-1, l_1 l_2 (n-1)} \ , \tag{4.8}$$

$$\text{unter } H_0^\phi(a^{(2)}) : \phi_{A^{(2)}} \equiv 0 \quad \text{gilt :} \qquad F_2 = \frac{\frac{1}{l_2-1} Q_{Z_2}}{\frac{1}{l_1 l_2 (n-1)} Q_N} \sim F_{l_2-1, l_1 l_2 (n-1)} \ , \tag{4.9}$$

und unter $H_0^\phi(a^{(1)}, a^{(2)}) : \phi_{A^{(1)}, A^{(2)}} \equiv 0 \quad$ gilt :

$$F_{1,2} = \frac{\frac{1}{(l_1-1)(l_2-1)} Q_{Z_{1,2}}}{\frac{1}{l_1 l_2 (n-1)} Q_N} \sim F_{(l_1-1)(l_2-1), l_1 l_2 (n-1)} \ . \tag{4.10}$$

Wir formulieren auch dieses Resultat als Satz:

Satz 4.6 (F-Test bei zwei festen Faktoren)
Seien $Y_\iota = \phi(a_\iota^{(1)}, a_\iota^{(2)}, \varepsilon_\iota)$, $\iota = 1, \ldots, N$, die Zielvariablen in einem balancierten Linearen Modell gemäß (M1) mit zwei festen Faktoren $a^{(1)}$ und $a^{(2)}$ mit Stufen $j_1 = 1, \ldots, l_1$ bzw. $j_2 = 1, \ldots, l_2$. Außerdem gelten die Annahmen (A1), (A2) und (A3). Dann gelten, unter den jeweiligen Hypothesen, die Verteilungsaussagen (4.8), (4.9) und (4.10).

4.1.6 Split Plot

Es handelt sich beim Split-Plot um ein Design mit zwei festen Faktoren $a^{(1)}$ und $a^{(2)}$ sowie einem zufälligen Faktor $a^{(3)}$, dabei haben die Faktoren l_1, l_2 bzw. l_3 Stufen (im Experiment). Der Faktor $a^{(3)}$ ist unter $a^{(1)}$ verschachtelt, d.h.:

$$\forall \iota, \iota' \in \{1, \ldots, N\} : a_\iota^{(1)} \neq a_{\iota'}^{(1)} \Longrightarrow a_\iota^{(3)} \neq a_{\iota'}^{(3)} \ . \tag{4.11}$$

Wir nehmen außerdem an, dass es ein $n \in \mathbb{N}$ gibt mit

$$\forall i \in \{1, \ldots, l_1\} : \mathrm{card}\{a_\iota^{(3)} : a_\iota^{(1)} = i\} = n \ , \tag{4.12}$$

das heißt, zu jeder Faktorstufe von $a_\iota^{(1)}$ gibt es genau n Faktorstufen von $a_\iota^{(3)}$. Weiterhin sind $a^{(2)}$ und $a^{(3)}$ vollständig gekreuzt. Der stochastische Prozess $(a_\iota^{(2)}, a_\iota^{(3)})_{\iota \in I}$, $I = \{1, \ldots, N\}$, nimmt also jeden Wert aus $E^{(2)} \times E^{(3)} = \{1, \ldots, l_2\} \times \{1, \ldots, l_3\}$ an, und zwar je genau einmal. Die Gesamtzahl der Beobachtungen in dem Design ist dann $N = l_2 \cdot l_3 = l_1 \cdot l_2 \cdot n$.

MODELL, ZERLEGUNG
Unter den Annahmen (A1), (A2), (A3) und (A4) lassen sich die Zielvariablen darstellen als

$$\begin{aligned} Y_\iota =& Y_{\tilde{a}_\iota^{(1)}, \tilde{a}_\iota^{(2)}, \tilde{a}_\iota^{(3)}} = \phi(a_\iota^{(1)}, a_\iota^{(2)}, a_\iota^{(3)}, \varepsilon_\iota) \\ =& \phi_\mu + \phi_{A^{(1)}}(a_\iota^{(1)}) + \phi_{A^{(2)}}(a_\iota^{(2)}) + \phi_{A^{(3)}}(a_\iota^{(3)}) + \phi_{A^{(1)},A^{(2)}}(a_\iota^{(1)}, a_\iota^{(2)}) \\ &+ \phi_{A^{(1)},A^{(3)}}(a_\iota^{(1)}, a_\iota^{(3)}) + \phi_{A^{(2)},A^{(3)}}(a_\iota^{(2)}, a_\iota^{(3)}) + \phi_{A^{(1)},A^{(2)},A^{(3)}}(a_\iota^{(1)}, a_\iota^{(2)}, a_\iota^{(3)}) \end{aligned}$$

$$+ \phi_\varepsilon(\varepsilon_\iota),\ \iota = 1, \ldots, N.$$

Die verschachtelte (hierarchische) Struktur im Design führt dazu, dass es nicht zweckmäßig ist, die Funktionen $\phi_{A^{(1)},A^{(3)}}$ und $\phi_{A^{(1)},A^{(2)},A^{(3)}}$ in das Modell mit aufzunehmen, sie werden also jeweils als identisch 0 angenommen. Im um diese Funktionen bereinigten Modell sind die Zielvariablen

$$\begin{aligned} Y_\iota =& \phi_\mu + \phi_{A^{(1)}}(a_\iota^{(1)}) + \phi_{A^{(2)}}(a_\iota^{(2)}) + \phi_{A^{(3)}}(a_\iota^{(3)}) + \phi_{A^{(1)},A^{(2)}}(a_\iota^{(1)}, a_\iota^{(2)}) \\ &+ \phi_{A^{(2)},A^{(3)}}(a_\iota^{(2)}, a_\iota^{(3)}) + \phi_\varepsilon(\varepsilon_\iota),\ \iota = 1, \ldots, N, \end{aligned} \tag{4.13}$$

bzw. in Vektorschreibweise

$$\mathbf{Y} = \boldsymbol{\phi}_\mu + \boldsymbol{\phi}_{A^{(1)}} + \boldsymbol{\phi}_{A^{(2)}} + \boldsymbol{\phi}_{A^{(3)}} + \boldsymbol{\phi}_{A^{(1)},A^{(2)}} + \boldsymbol{\phi}_{A^{(2)},A^{(3)}} + \boldsymbol{\phi}_\varepsilon \ .$$

Die angegebenen Zufallsvariablen sind unabhängig und es gilt

$$\begin{aligned} &\phi_{A^{(3)}}(a_\iota^{(3)}) \sim \mathrm{N}(0, \sigma^2_{A^{(3)}}), \quad \phi_{A^{(2)},A^{(3)}}(a_\iota^{(2)}, a_\iota^{(3)}) \sim \mathrm{N}(0, \sigma^2_{A^{(2)},A^{(3)}}), \quad \phi_\varepsilon(\varepsilon_\iota) \sim \mathrm{N}(0, \sigma^2_\varepsilon) \\ &\text{bzw.} \quad \boldsymbol{\phi}_{A^{(3)}} \sim \mathrm{N}(\mathbf{0}, \sigma^2_{A^{(3)}} \cdot \mathbf{I}_{l_1} \otimes \mathbf{J}_{l_2} \otimes \mathbf{I}_n), \quad \boldsymbol{\phi}_{A^{(2)},A^{(3)}} \sim \mathrm{N}(\mathbf{0}, \sigma^2_{A^{(2)},A^{(3)}} \cdot \mathbf{I}_N) \\ &\text{und} \quad \boldsymbol{\phi}_\varepsilon \sim \mathrm{N}(\mathbf{0}, \sigma^2_\varepsilon \cdot \mathbf{I}_N) \ . \end{aligned}$$

Die Reparametrisierungsbedingungen (R1), (R2) und (R3) lauten hier:

$$(a) \quad \mathrm{E}(\phi_\varepsilon(\varepsilon_\iota)) = \int \phi_\varepsilon(x) dG_\varepsilon(x) = 0\,,$$

$$(b) \quad \mathrm{E}_{A^{(1)}}(\phi_{A^{(1)}}(a_\iota^{(1)})) = \int \phi_{A^{(1)}}(x) dG_{A^{(1)}}(x) = \frac{1}{l_1} \sum_{i=1}^{l_1} \phi_{A^{(1)}}(i) = 0\,,$$

$$(c) \quad \mathrm{E}_{A^{(2)}}(\phi_{A^{(2)}}(a_\iota^{(2)})) = \int \phi_{A^{(2)}}(x) dG_{A^{(2)}}(x) = \frac{1}{l_2} \sum_{i=1}^{l_2} \phi_{A^{(2)}}(i) = 0\,,$$

$$(d) \quad \mathrm{E}_{A^{(3)}}(\phi_{A^{(3)}}(a_\iota^{(3)})) = \int \phi_{A^{(3)}}(x) dG_{A^{(3)}}(x) = 0\,,$$

$$\begin{aligned} (e) \quad & \mathrm{E}_{A^{(1)}}(\phi_{A^{(1)},A^{(2)}}(a_\iota^{(1)}, a_\iota^{(2)}) | a_\iota^{(2)} = i_2) = \int \phi_{A^{(1)},A^{(2)}}(x, i_2) dG_{A^{(1)}}(x) \\ & = \frac{1}{l_1} \sum_{i_1=1}^{l_1} \phi_{A^{(1)},A^{(2)}}(i_1, i_2) = 0\,, \end{aligned}$$

$$\begin{aligned} (f) \quad & \mathrm{E}_{A^{(2)}}(\phi_{A^{(1)},A^{(2)}}(a_\iota^{(1)}, a_\iota^{(2)}) | a_\iota^{(1)} = i_1) = \int \phi_{A^{(1)},A^{(2)}}(i_1, x) dG_{A^{(2)}}(x) \\ & = \frac{1}{l_2} \sum_{i_2=1}^{l_2} \phi_{A^{(1)},A^{(2)}}(i_1, i_2) = 0\,, \end{aligned}$$

$$\begin{aligned} (g) \quad & \mathrm{E}_{A^{(2)}}(\phi_{A^{(2)},A^{(3)}}(a_\iota^{(2)}, a_\iota^{(3)}) | a_\iota^{(3)} = i_3) = \int \phi_{A^{(2)},A^{(3)}}(x, i_3) dG_{A^{(2)}}(x) \\ & = \frac{1}{l_2} \sum_{i_2=1}^{l_2} \phi_{A^{(2)},A^{(3)}}(i_2, i_3) = 0\,, \end{aligned}$$

$$(h)\quad \mathrm{E}_{A^{(3)}}(\phi_{A^{(2)},A^{(3)}}(a_\iota^{(2)},a_\iota^{(3)})|a_\iota^{(2)}=i_2)=\int \phi_{A^{(2)},A^{(3)}}(i_2,x)dG_{A^{(3)}}(x)=0\,.$$

KLASSISCHE FORM

KIRK (1968) formuliert den Split-Plot als

$$Y_{ijk}=\mu+\alpha_i+\pi_{k(i)}+\beta_j+\alpha\beta_{ij}+\beta\pi_{jk(i)}+\varepsilon_{ijk}\ ,$$

wobei er erwähnt, dass hier $\beta\pi_{jk(i)}$ und ε_{ijk} nicht getrennt voneinander geschätzt werden können. Damit entspricht das Modell bei KIRK dem von uns oben angegebenen bereinigten Modell (4.13).

ERWARTETE WERTE UND KOVARIANZSTRUKTUR

Es gilt:

$$\begin{aligned}
&\mathrm{E}(Y_\iota|a_\iota^{(1)}=i_1,a_\iota^{(2)}=i_2)=\phi_\mu+\phi_{A^{(1)}}(i_1)+\phi_{A^{(2)}}(i_2)\\
&\qquad\qquad\qquad+\phi_{A^{(1)},A^{(2)}}(i_1,i_2),\ \iota=1,\dots,N,\\
\text{und}\quad &\mathrm{Cov}(Y_\iota,Y_{\iota'}|a_\iota^{(1)}=i_1,a_\iota^{(2)}=i_2,a_{\iota'}^{(1)}=i_1',a_{\iota'}^{(2)}=i_2')\\
&=\mathrm{Cov}(\phi_{A^{(3)}}(a_\iota^{(3)}),\phi_{A^{(3)}}(a_{\iota'}^{(3)}))+\mathrm{Cov}(\phi_{A^{(2)},A^{(3)}}(a_\iota^{(2)},a_\iota^{(3)}),\phi_{A^{(2)},A^{(3)}}(a_{\iota'}^{(2)},a_{\iota'}^{(3)}))\\
&\quad+\mathrm{Cov}(\phi_\varepsilon(\varepsilon_\iota),\phi_\varepsilon(\varepsilon_{\iota'}))\\
&=\begin{cases}\sigma^2_{A^{(3)}}, & a_\iota^{(3)}=a_{\iota'}^{(3)}\\ 0, & \text{sonst}\end{cases}+\begin{cases}\sigma^2_{A^{(2)},A^{(3)}}, & \iota=\iota'\\ 0, & \text{sonst}\end{cases}+\begin{cases}\sigma^2_\varepsilon, & \iota=\iota'\\ 0, & \text{sonst.}\end{cases}
\end{aligned}$$

HYPOTHESEN

Es werden Hypothesen anhand der beiden festen Faktoren formuliert:
Hypothese über den ersten Faktor: $H_0^\phi(a^{(1)}):\phi_{A^{(1)}}\equiv 0$.
Hypothese über den zweiten Faktor: $H_0^\phi(a^{(2)}):\phi_{A^{(2)}}\equiv 0$.
Hypothese über die Wechselwirkung zwischen beiden Faktoren: $H_0^\phi(a^{(1)},a^{(2)}):\phi_{A^{(1)},A^{(2)}}\equiv 0$.

TESTSTATISTIK

Seien

$$\begin{aligned}
&Q_{Z_1}=\mathbf{Y}'(\mathbf{P}_{l_1}\otimes\frac{1}{l_2}\mathbf{J}_{l_2}\otimes\frac{1}{n}\mathbf{J}_n)\mathbf{Y}, \qquad Q_{Z_2}=\mathbf{Y}'(\frac{1}{l_1}\mathbf{J}_{l_1}\otimes\mathbf{P}_{l_2}\otimes\frac{1}{n}\mathbf{J}_n)\mathbf{Y},\\
&Q_{Z_{1,2}}=\mathbf{Y}'(\mathbf{P}_{l_1}\otimes\mathbf{P}_{l_2}\otimes\frac{1}{n}\mathbf{J}_n)\mathbf{Y}, \qquad Q_N=\mathbf{Y}'(\mathbf{I}_{l_1}\otimes\frac{1}{l_2}\mathbf{J}_{l_2}\otimes\mathbf{P}_n)\mathbf{Y}\\
&\text{und } Q_{N_{1,2}}=\mathbf{Y}'(\mathbf{I}_{l_1}\otimes\mathbf{P}_{l_2}\otimes\mathbf{P}_n)\mathbf{Y}\,.
\end{aligned}$$

Wegen der Zerlegung von $\mathbf{Y}$ gilt

$$\begin{aligned}
Q_{Z_1}&=(\boldsymbol{\phi}_\mu+\boldsymbol{\phi}_{A^{(1)}}+\boldsymbol{\phi}_{A^{(2)}}+\boldsymbol{\phi}_{A^{(3)}}+\boldsymbol{\phi}_{A^{(1)},A^{(2)}}+\boldsymbol{\phi}_{A^{(2)},A^{(3)}}+\boldsymbol{\phi}_\varepsilon)'(\mathbf{P}_{l_1}\otimes\frac{1}{l_2}\mathbf{J}_{l_2}\otimes\frac{1}{n}\mathbf{J}_n)\times\\
&\qquad(\boldsymbol{\phi}_\mu+\boldsymbol{\phi}_{A^{(1)}}+\boldsymbol{\phi}_{A^{(2)}}+\boldsymbol{\phi}_{A^{(3)}}+\boldsymbol{\phi}_{A^{(1)},A^{(2)}}+\boldsymbol{\phi}_{A^{(2)},A^{(3)}}+\boldsymbol{\phi}_\varepsilon)\\
&=(\boldsymbol{\phi}_{A^{(1)}}+\boldsymbol{\phi}_{A^{(3)}}+\boldsymbol{\phi}_{A^{(2)},A^{(3)}}+\boldsymbol{\phi}_\varepsilon)'(\mathbf{P}_{l_1}\otimes\frac{1}{l_2}\mathbf{J}_{l_2}\otimes\frac{1}{n}\mathbf{J}_n)(\boldsymbol{\phi}_{A^{(1)}}+\boldsymbol{\phi}_{A^{(3)}}+\boldsymbol{\phi}_{A^{(2)},A^{(3)}}+\boldsymbol{\phi}_\varepsilon)\,.
\end{aligned}$$

Unter $H_0^\phi(a^{(1)}) : \phi_{A^{(1)}} \equiv 0$ ist somit

$$Q_{Z_1} = (\boldsymbol{\phi}_{A^{(3)}} + \boldsymbol{\phi}_{A^{(2)},A^{(3)}} + \boldsymbol{\phi}_\varepsilon)'(\mathbf{P}_{l_1} \otimes \frac{1}{l_2}\mathbf{J}_{l_2} \otimes \frac{1}{n}\mathbf{J}_n)(\boldsymbol{\phi}_{A^{(3)}} + \boldsymbol{\phi}_{A^{(2)},A^{(3)}} + \boldsymbol{\phi}_\varepsilon)\,.$$

Aufgrund von

$$\begin{aligned}
&\mathrm{Var}(\boldsymbol{\phi}_{A^{(3)}} + \boldsymbol{\phi}_{A^{(2)},A^{(3)}} + \boldsymbol{\phi}_\varepsilon)\\
&= \sigma^2_{A^{(3)}} \cdot \mathbf{I}_{l_1} \otimes \mathbf{J}_{l_2} \otimes \mathbf{I}_n + (\sigma^2_{A^{(2)},A^{(3)}} + \sigma^2_\varepsilon) \cdot \mathbf{I}_N\\
&= (l_2 \cdot \sigma^2_{A^{(3)}} + \sigma^2_{A^{(2)},A^{(3)}} + \sigma^2_\varepsilon) \cdot \frac{1}{l_1}\mathbf{J}_{l_1} \otimes \frac{1}{l_2}\mathbf{J}_{l_2} \otimes \mathbf{I}_n\\
&\quad + (l_2 \cdot \sigma^2_{A^{(3)}} + \sigma^2_{A^{(2)},A^{(3)}} + \sigma^2_\varepsilon) \cdot \mathbf{P}_{l_1} \otimes \frac{1}{l_2}\mathbf{J}_{l_2} \otimes \mathbf{I}_n\\
&\quad + (\sigma^2_{A^{(2)},A^{(3)}} + \sigma^2_\varepsilon) \cdot \mathbf{I}_{l_1} \otimes \mathbf{P}_{l_2} \otimes \mathbf{I}_n\\
\text{gilt}\quad &(\mathbf{P}_{l_1} \otimes \frac{1}{l_2}\mathbf{J}_{l_2} \otimes \frac{1}{n}\mathbf{J}_n) \cdot \mathrm{Var}(\boldsymbol{\phi}_{A^{(3)}} + \boldsymbol{\phi}_{A^{(2)},A^{(3)}} + \boldsymbol{\phi}_\varepsilon)\\
&= (l_2 \cdot \sigma^2_{A^{(3)}} + \sigma^2_{A^{(2)},A^{(3)}} + \sigma^2_\varepsilon) \cdot \mathbf{P}_{l_1} \otimes \frac{1}{l_2}\mathbf{J}_{l_2} \otimes \mathbf{I}_n\\
&\text{und damit folgt}\quad Q_{Z_1} \sim (l_2 \cdot \sigma^2_{A^{(3)}} + \sigma^2_{A^{(2)},A^{(3)}} + \sigma^2_\varepsilon) \cdot \chi^2_{l_1-1}\,.
\end{aligned}$$

In ähnlicher Weise können die Verteilungen der anderen Quadratformen bestimmt werden. Es ergibt sich das im nachfolgenden Satz zusammengefasste Resultat:

Satz 4.7 (F-Test im Split-Plot)
Seien $Y_\iota = \phi(a_\iota^{(1)}, a_\iota^{(2)}, a_\iota^{(3)}, \varepsilon_\iota)$, $\iota = 1, \ldots, N$, die Zielvariablen in einem Split-Plot unter den Voraussetzungen des Linearen Modells gemäß (M1). Die Faktoren $a^{(1)}$ und $a^{(2)}$ mit Stufen $j_1 = 1, \ldots, l_1$ bzw. $j_2 = 1, \ldots, l_2$ seien fest, der Faktor $a^{(3)}$ sei zufällig und unter $a^{(1)}$ balanciert verschachtelt, d.h. es gelten (4.11) und (4.12). Außerdem gelten die Annahmen (A1), (A2), (A3) und (A4). Dann gelten, mit den in diesem Abschnitt verwendeten Bezeichnungen für die quadratischen Formen, die folgenden Verteilungsaussagen:

Unter $H_0^\phi(a^{(1)}) : \phi_{A^{(1)}} \equiv 0$ ist $$F_1 = \frac{\frac{1}{l_1-1}Q_{Z_1}}{\frac{1}{l_1(n-1)}Q_N} \sim F_{l_1-1,l_1(n-1)}\,,$$

unter $H_0^\phi(a^{(2)}) : \phi_{A^{(2)}} \equiv 0$ ist $$F_2 = \frac{\frac{1}{l_2-1}Q_{Z_2}}{\frac{1}{l_1(n-1)}Q_N} \sim F_{l_2-1,l_1(n-1)}\,,$$

und unter $H_0^\phi(a^{(1)}, a^{(2)}) : \phi_{A^{(1)},A^{(2)}} \equiv 0$ ist

$$F_{1,2} = \frac{\frac{1}{(l_1-1)(l_2-1)}Q_{Z_{1,2}}}{\frac{1}{l_1(l_2-1)(n-1)}Q_{N_{1,2}}} \sim F_{(l_1-1)(l_2-1),l_1(l_2-1)(n-1)}\,.$$

4.1.7 Austauschbares Lineares Modell

Im Unterschied zum bisher betrachteten Linearen Modell gelte hier für die ε_ι, $\iota \in I$, lediglich, dass sie eine austauschbare Verteilung besitzen. Dies bedeutet, dass für jede Permutation

$\pi : I \to I$ der Vektor $(\varepsilon_{\pi(1)}, \ldots, \varepsilon_{\pi(N)})$ dieselbe Verteilung hat wie $(\varepsilon_1, \ldots, \varepsilon_N)$. Falls die ε_ι sogar unabhängig sind, wie zu Beginn des Kapitels 4.1 gefordert, sind sie insbesondere auch austauschbar, die Umkehrung gilt jedoch nicht. Es zeigt sich, dass unter Annahme der Austauschbarkeit die Verfahren der ein- und zweifaktoriellen Varianzanalyse für feste Faktoren anwendbar sind wie unter der stärkeren Annahme der unabhängigen Fehlerterme (vgl. ARNOLD (1981, S.233ff)).

4.2 Lineares Modell ohne Normalverteilungsannahme

Wir legen wieder im Wesentlichen das zu Beginn von Kapitel 4.1 beschriebene Lineare Modell zugrunde, lassen aber die Forderung der normalverteilten Fehlerterme fallen und beschränken uns hier auf den Fall $\mathcal{M} = \mathcal{M}_f$. Es gelten also die Annahmen (A1), (A2), (A2a) und (A3b). Da wir hier auch transformierte Zielvariablen betrachten, ist Modellgleichung (M1a) maßgeblich.

4.2.1 Ein-Stichproben-Tests

Wir betrachten eine Stichprobe der Größe N.

MODELL, ZERLEGUNG
Die transformierten Zielvariablen sind

$$Z_\iota = \tilde{\phi}(\varepsilon_\iota) = \tilde{\phi}_\mu + \tilde{\phi}_\varepsilon(\varepsilon_\iota) \text{ bzw. } \mathbf{Z} = \tilde{\boldsymbol{\phi}}_\mu + \tilde{\boldsymbol{\phi}}_\varepsilon \,.$$

Es gilt $\mathrm{E}(\tilde{\phi}_\varepsilon(\varepsilon_\iota)) = \int \tilde{\phi}_\varepsilon(x) dG_\varepsilon(x) = 0$ und $\tilde{\phi}_\mu = \int \tilde{\phi}(x)\, dG_\varepsilon(x)$.

HYPOTHESE
$H_0^{\tilde{\phi}}(\mu) : \tilde{\phi}_\mu = \mu_0$ mit $\mu_0 \in \mathbb{R}$ fest.

TESTSTATISTIK
Im Unterschied zum Linearen Modell gilt hier lediglich die Annahme (A3b), also $\tilde{\phi}_\varepsilon(\varepsilon_\iota) \overset{u.i.v.}{\sim} G$. Unter den Voraussetzungen für den zentralen Grenzwertsatz erhält man dennoch die asymptotische Verteilungsaussage

$$T_N = \sqrt{N} \frac{\frac{1}{N} \sum_{\iota=1}^{N} Z_\iota - \tilde{\phi}_\mu}{\sigma_{\tilde{\phi}_\varepsilon}} \,\dot{\sim}\, \mathrm{N}(0,1) \qquad (\text{mit } \sigma^2_{\tilde{\phi}_\varepsilon} = \mathrm{Var}(\tilde{\phi}_\varepsilon(\varepsilon_1)))\,.$$

SPEZIALFALL
Wählt man $\tilde{\phi} = q \circ \phi$ mit $q(x) = \begin{cases} 0, & x \le c \\ 1, & x > c \end{cases}$, so ist die Hypothese $H_0^{\tilde{\phi}}(\mu) : \tilde{\phi}_\mu = \frac{1}{2}$ äquivalent zu $P(\phi(\varepsilon_\iota) \le c) = P(\phi(\varepsilon_\iota) > c)$.
Da unter Hypothese $Z_\iota \overset{u.i.v.}{\sim} B(p),\ \iota = 1, \ldots, N$, mit $p = P(\phi(\varepsilon_\iota) > c) = \frac{1}{2}$ gilt, ist $T_N^{V_e} = \sum_{\iota=1}^{N} Z_\iota$ binomialverteilt mit den Parametern N und $\frac{1}{2}$.

Mit dem zentralen Grenzwertsatz folgt dann aber unter Hypothese:

$$T_N^{V_a} = \sqrt{N}\frac{\frac{1}{N}\sum\limits_{\iota=1}^{N} Z_\iota - \mu_0}{\sigma_{\tilde{\phi}_\varepsilon}} = \sqrt{N}\frac{\frac{1}{N}\sum\limits_{\iota=1}^{N} Z_\iota - \frac{1}{2}}{\frac{1}{2}} = \frac{2}{\sqrt{N}}(\mathrm{card}\{\iota : Y_\iota > c\} - \frac{N}{2}) \mathrel{\dot{\sim}} \mathrm{N}(0,1)\ .$$

Wir erhalten zusammengefasst also folgende Aussage:

Satz 4.8 (Vorzeichen-Test) *Sei $Y_\iota = \phi(\varepsilon_\iota) = \phi_\mu + \phi_\varepsilon(\varepsilon_\iota),\ \iota = 1, \ldots, N$, eine Stichprobe der Größe N, die der Annahme (A3b) genügt.*
Dann ist, unter der Hypothese $H_0^{\tilde{\phi}} : P(Y_\iota \le c) = P(Y_\iota > c) = \frac{1}{2}$

a) die Teststatistik $T_N^{V_e} = \sum\limits_{\iota=1}^{N} 1\!\!1_{\{Y_\iota > c\}} = \mathrm{card}\{\iota : Y_\iota > c\}$ binomialverteilt mit den Parametern N und $\frac{1}{2}$, sowie

b) die Teststatistik $T_N^{V_a} = \frac{2}{\sqrt{N}}(\mathrm{card}\{\iota : Y_\iota > c\} - \frac{N}{2})$ asymptotisch $\mathrm{N}(0,1)$-verteilt.

Der durch $T_N^{V_e}$ bzw. $T_N^{V_a}$ beschriebene Test wird als exakter bzw. asymptotischer Vorzeichentest bezeichnet.

4.2.2 Ein fester Faktor

Wir betrachten ein balanciertes Design mit einem Faktor a, der l Stufen besitzt. Es ist dann $N = l \cdot n$ mit einem $n \in \mathbb{N}$. Der stochastische Prozeß $(a_\iota)_{\iota \in I},\ I = \{1, \ldots, N\}$, nimmt also jeden der Werte aus $E = \{e_1, \ldots, e_l\}$ genau n-mal an.

MODELL, ZERLEGUNG
Die transformierten Zielvariablen sind

$$Z_\iota = \tilde{\phi}(a_\iota, \varepsilon_\iota) = \tilde{\phi}_\mu + \tilde{\phi}_A(a_\iota) + \tilde{\phi}_{A,\varepsilon}(a_\iota, \varepsilon_\iota) + \tilde{\phi}_\varepsilon(\varepsilon_\iota)\ .$$

Mithilfe des in Abschnitt 3.1 definierten sortierten Vektors **Z** und analog definierter Vektoren $\tilde{\boldsymbol{\phi}}_\mu,\ \tilde{\boldsymbol{\phi}}_A,\ \tilde{\boldsymbol{\phi}}_{A,\varepsilon}$ und $\tilde{\boldsymbol{\phi}}_\varepsilon$ läßt sich dies auch vektoriell schreiben als

$$\mathbf{Z} = \tilde{\boldsymbol{\phi}}_\mu + \tilde{\boldsymbol{\phi}}_A + \tilde{\boldsymbol{\phi}}_{A,\varepsilon} + \tilde{\boldsymbol{\phi}}_\varepsilon\ .$$

Unter den Annahmen (A1), (A2) gilt:

$$Z_\iota = \tilde{\phi}_\mu + \tilde{\phi}_A(a_\iota) + \tilde{\phi}_\varepsilon(\varepsilon_\iota) \quad \text{mit unabhängigen Zufallsvariablen } \tilde{\phi}_A(a_\iota) \text{ und } \tilde{\phi}_\varepsilon(\varepsilon_\iota)\ .$$

Außerdem gelten die folgenden Reparametrisierungsbedingungen:

$$(a) \qquad \mathrm{E}(\tilde{\phi}_\varepsilon(\varepsilon)) = \int \tilde{\phi}_\varepsilon(x)dG_\varepsilon(x) = 0\ .$$

$$(b) \qquad \mathrm{E}_A(\tilde{\phi}_A(a_\iota)) = \int \tilde{\phi}_A(x) dG_A(x) = \frac{1}{l}\sum_{j=1}^{l} q(\alpha_j) = 0 \,.$$

Wir erhalten also

$$\begin{aligned} \tilde{\phi}_\mu &= \int\int \tilde{\phi}(x_1, x_2)\, dG_a(x_1) dG_\varepsilon(x_2) \quad \text{und} \\ \tilde{\phi}_A(x_1) &= \int \tilde{\phi}(x_1, x_2)\, dG_\varepsilon(x_2) - \tilde{\phi}_\mu \,. \end{aligned}$$

HYPOTHESE
$H_0^{\tilde{\phi}}(a) : \tilde{\phi}_A \equiv 0$.

Bis zu diesem Punkt herrscht völlige Analogie zum gewöhnlichen Linearen Modell. Im Unterschied dazu sind allerdings die transformierten Zielvariablen nicht mehr normalverteilt, sondern es gilt lediglich: $\tilde{\phi}_\varepsilon(\varepsilon_\iota) \overset{u.i.v.}{\sim} G$, wobei G eine Verteilungsfunktion mit Erwartungswert 0 und Varianz σ_ε^2 ist. Da die Verteilung G ansonsten nicht näher spezifiziert ist, lassen sich Verteilungsaussagen für die im gewöhnlichen Linearen Modell verwendeten F-Test-Statistiken nur asymptotisch machen. Dazu wird angenommen, dass entweder die Anzahl n der Beobachtungen pro Faktorstufe oder die Anzahl l der Faktorstufen gegen unendlich geht.
Für $n \to \infty$, l fest, sind die Resultate über die Grenzverteilungen, die auch in der hier benutzen Modellformulierung gültig sind, bereits seit längerem bekannt (vgl. z.B. die Darstellung bei ARNOLD (1981)).
Im Fall $l \to \infty$, n fest, haben wir (vgl. BATHKE (2000)) eine von BOOS und BROWNIE veröffentlichte Aussage (BOOS & BROWNIE (1995), BROWNIE & BOOS (1994)) für das einfaktorielle Design auf Versuchspläne mit beliebig vielen Faktoren verallgemeinert. Im Beweis, der zusammen mit einer ausführlicheren Darstellung des Falles $l \to \infty$, n fest, in Kapitel 5 zu finden ist, haben wir dazu Kroneckerprodukte und eine spezielle Zerlegung einer Projektionsmatrix verwendet.

TESTSTATISTIKEN

Satz 4.9 (asymptotischer F-Test bei einem festen Faktor, $n \to \infty$, l fest)
Seien $Z_\iota = \tilde{\phi}(a_\iota, \varepsilon_\iota), \iota = 1, \ldots, N$, die Zielvariablen in einem balancierten Modell gemäß (M1a) mit einem festen Faktor a mit l Stufen. Außerdem gelten die Annahmen (A1), (A2) und (A3b). Die $\tilde{\phi}_\varepsilon(\varepsilon_\iota)$, $\iota = 1, \ldots, N$, seien unabhängig mit $\mathrm{Var}(\tilde{\phi}_\varepsilon(\varepsilon_\iota)) = \sigma_\varepsilon^2 < \infty$. Dann ist, unter der Hypothese $H_0^{\tilde{\phi}}(a) : \tilde{\phi}_A \equiv 0$, die Teststatistik $\frac{Q_Z}{\frac{1}{N-l}Q_N}$ mit Q_Z, Q_N wie in Satz 4.2, asymptotisch (l fest, $n \to \infty$) verteilt nach χ_{l-1}^2.

Beweis: Der Nenner ist für $n \to \infty$ ein stark konsistenter Varianzschätzer.
Für den Zähler ergibt sich die χ^2-Verteilung als unmittelbare Anwendung von Theorem 1 aus ROTAR (1973): die Klasse $\mathcal{F}_{0,\sigma_\varepsilon^2}$ aller Verteilungen mit Erwartungswert 0 und Varianz σ_ε^2 ist Invarianzklasse bezüglich der Grenzverteilung von $\frac{1}{l-1}Q_Z$ (für $n \to \infty$). □

Satz 4.10 (asymptotischer F-Test bei einem festen Faktor, $l \to \infty$, n fest)
Sei das Modell gegeben wie in Satz 4.9. Dann gilt, unter der Hypothese $H_0^{\tilde{\phi}}(a) : \tilde{\phi}_A \equiv 0$, für $F = \frac{\frac{1}{l-1}Q_Z}{\frac{1}{N-l}Q_N}$ mit Q_Z, Q_N wie in Satz 4.2: $\sqrt{l}(F-1)$ ist asymptotisch (n fest, $l \to \infty$) verteilt nach $\mathrm{N}(0, \frac{2n}{n-1})$.

Beweis: Die Methoden aus ROTAR (1973) und der darauf aufbauenden Literatur führen hier nicht zum gewünschten Resultat, da sich mit ihnen lediglich ergibt, dass $\frac{1}{l-1}Q_Z$ für $l \to \infty$ gegen σ_ε^2 konvergiert.
Ein vollständiger Beweis der Aussage findet sich im Beweis zu Satz 5.2. In Kapitel 5 wird ein entsprechendes Resultat sogar für ein Design mit beliebig vielen Faktoren bewiesen. □

Anmerkung: Satz 4.10 wird nicht in der Weise angewendet, dass man die mit dem Faktor $\sqrt{l}$ aufgeblähte und asymptotisch normalverteilte F-Statistik verwendet. Vielmehr gestattet das Resultat des Satzes, nämlich die Invarianz von $\sqrt{l}(F-1)$ unter beliebigen Verteilungen mit $\mathrm{Var}(\tilde{\phi}_\varepsilon(\varepsilon_\iota)) = \sigma_\varepsilon^2 < \infty$, die direkte Verwendung der F-Statistik und der ensprechenden F-Quantile, auch bei nicht normalverteilten Daten, sofern l groß genug ist (vgl. dazu Tabelle 2 auf S. 110).

4.3 Allgemeines Nichtparametrisches Modell

Für das allgemeine Nichtparametrische Modell sei im Folgenden stets Modellgleichung (M2) maßgeblich.

4.3.1 Ein-Stichproben-Tests

Wir betrachten eine Stichprobe der Größe N.

MODELL, ZERLEGUNG
Hier wird die Verteilungsfunktion von Y_ι betrachtet. Es ist $F_\iota(t) = \psi_\mu(t) =: F(t)$, $\iota = 1, \ldots, N$.

SCHÄTZER
Einen Schätzer für die Funktion $F(t)$ erhalten wir durch die empirische Verteilungsfunktion

$$\hat{\psi}_\mu(t) := \hat{F}(t) := \frac{1}{N} \sum_{\iota=1}^{N} c(t - Y_\iota) \; .$$

Dabei und im Folgenden bezeichne immer $c(\cdot)$ die „Zählfunktion“ $c(x) = \begin{cases} 1, & x > 0 \\ \frac{1}{2}, & x = 0 \\ 0, & x < 0 \end{cases}$.

HYPOTHESEN UND TESTSTATISTIKEN
Da kein Faktor vorliegt, lassen sich auch keine Hypothesen der Form $H_0^\psi(A) : \psi_A \equiv 0$ testen.

Die Verteilungsfunktion, die im allgemeinen nichtparametrischen Modell betrachtet wird, ist jedoch ein derartig mächtiges Instrument, dass auf ihr leicht Hypothesen aufgebaut werden können.

1. Es bezeichne $\mathrm{Med}(F)$ den Median der Verteilung F.
   Unter der Hypothese $H_0^{\psi,Med}$: $\mathrm{Med}(F) = m$ gilt: $P(Y_\iota > m) = P(Y_\iota < m)$ bzw. $P(Y_\iota > m | Y_\iota \neq m) = \frac{1}{2}$.
   Damit ist unter Hypothese $T_N = \sum\limits_{\iota=1}^{N} \mathrm{I\!I}_{\{Y_\iota > m\}} = \mathrm{card}\{\iota : Y_\iota > m\}$ binomialverteilt mit Parametern N_0 und $\frac{1}{2}$, dabei ist $N_0 = \mathrm{card}\{\iota : Y_\iota \neq m\}$. Bei stetiger Verteilung F ist $N_0 = N$. Mit dem zentralen Grenzwertsatz folgt nun unter Hypothese:
   $$\sqrt{N_0}\frac{\frac{1}{N_0}\sum\limits_{\iota=1}^{N} \mathrm{I\!I}_{\{Y_\iota > m\}} - \frac{1}{2}}{\frac{1}{2}} = \frac{2}{\sqrt{N_0}}(\mathrm{card}\{\iota : Y_\iota > m\} - \frac{N_0}{2}) \stackrel{\cdot}{\sim} \mathrm{N}(0,1)\,.$$
   Es ergibt sich also eine Variante zum in Satz 4.8 genannten Vorzeichentest.

2. Falls F als symmetrisch zu einem Symmetriezentrum s vorausgesetzt wird, ist $\mathrm{Med}(F) = s$, und es gilt wegen $F(s+x) = 1 - F(s-x)$:
   $$\frac{1}{2} = \int (1 - F(s-x))\, dF(s+x) = 1 - \int F(s-x)\, dF(s+x)\,.$$
   Unter der Hypothese $H_0^{\psi,Sym}$: $s = 0$ ist somit $\int F(-x)\, dF(x) = \frac{1}{2}$.
   Um eine Teststatistik zu erhalten, wird $\int F(-x)\, dF(x)$ durch $\int \hat{F}(-x)\, d\hat{F}(x)$, also mithilfe der empirischen Verteilungsfunktion geschätzt.
   Es ergibt sich
   $$\begin{aligned}&\int \hat{F}(-x)\, d\hat{F}(x)\\ &\quad = \frac{1}{N}\sum_{\iota=1}^{N} \hat{F}(-Y_\iota) = \frac{1}{N}\sum_{\iota=1}^{N}\frac{1}{N}\sum_{\iota'=1}^{N} c(-Y_\iota - Y_{\iota'})\\ &\quad = 1 - \frac{1}{N^2}\mathrm{card}\{(\iota,\iota') : Y_\iota + Y_{\iota'} > 0\} - \frac{1}{2N^2}\mathrm{card}\{(\iota,\iota') : Y_\iota + Y_{\iota'} = 0\}\,.\end{aligned}$$
   Durch Normierung erhält man hieraus den Wilcoxon-Vorzeichen-Rank-Test, dessen Teststatistik ebenfalls asymptotisch standardnormalverteilt ist (vgl. z.B. RANDLES & WOLFE (1979)).

Die beiden genannten Tests ergeben sich übrigens auch als U-Statistiken ersten bzw. zweiten Grades zu den symmetrischen Kernen

$$h(x_1) = \begin{cases} 1, & x_1 > 0 \\ \frac{1}{2}, & x_1 = 0 \\ 0, & x_1 < 0 \end{cases} \quad \text{bzw.} \quad h(x_1,x_2) = \begin{cases} 1, & x_1 + x_2 > 0 \\ \frac{1}{2}, & x_1 + x_2 = 0 \\ 0, & x_1 + x_2 < 0 \end{cases}\,.$$

Beide Hypothesen sind allerdings nicht mehr „nichtparametrisch“ im engeren Sinne, da jeweils ein Parameter, nämlich der Median bzw. das Symmetriezentrum der Verteilung im Modell enthalten ist.

Ein anderer Zugang, der zu den hier aufgeführten Tests führt, besteht in einer Modellformulierung gemäß (M4), also mithilfe von Verteilungsquantilen. Darauf soll hier aber nicht näher eingegangen werden.

4.3.2 Ein fester Faktor

Wir betrachten ein balanciertes Design mit einem Faktor a, der l Stufen besitzt. Es ist dann $N = l \cdot n$ mit einem $n \in \mathbb{N}$. Der stochastische Prozeß $(a_\iota)_{\iota \in I},\ I = \{1, \ldots, N\}$, nimmt also jeden der Werte aus $E = \{e_1, \ldots, e_l\}$ genau n-mal an.

MODELL, ZERLEGUNG
In diesem Modell werden die Verteilungsfunktionen von Y_ι betrachtet. Es ist

$$F_\iota(t) = \psi(a_\iota)(t) = \psi_\mu(t) + \psi_A(a_\iota)(t) \,.$$

Hierbei wird implizit unterstellt, daß gilt: $a_{\iota_1} = a_{\iota_2} \Rightarrow F_{\iota_1} = F_{\iota_2}$. Somit wären in einem Gleichungssystem der Form

$$\begin{aligned} F_1(t) &= \psi_\mu(t) + \psi_A(a_1)(t) \\ &\vdots \\ F_N(t) &= \psi_\mu(t) + \psi_A(a_N)(t) \end{aligned}$$

etliche Gleichungen redundant, denn a_ι nimmt nur die l verschiedenen Werte $e_1^{(f)}, \ldots, e_l^{(f)}$ an. Wir können das Gleichungssystem also reduzieren auf die wesentlichen Gleichungen

$$\begin{aligned} F_1^s(t) &= \psi_\mu(t) + \psi_A(e_1^{(f)})(t) \\ &\vdots \\ F_l^s(t) &= \psi_\mu(t) + \psi_A(e_l^{(f)})(t) \,. \end{aligned}$$

Dieses System erhält ein vektorielles Äquivalent

$$\mathbf{F}(t) = \boldsymbol{\psi}_\mu(t) + \boldsymbol{\psi}_A(t)$$

mit $\mathbf{F} = (F_1^s, \ldots, F_l^s)',\ \boldsymbol{\psi}_\mu = (\underbrace{\psi_\mu, \ldots, \psi_\mu}_{\iota\text{-mal}})',\ \boldsymbol{\psi}_A = (\psi_A(e_1^{(f)}), \ldots, \psi_A(e_l^{(f)}))' \,.$

Es gilt die punktweise definierte Reparametrisierungsbedingung

$$\mathrm{E}_A(\psi_A(a_\iota)(t)) \;=\; \int \psi_A(x)(t)dG_A(x) \;=\; \frac{1}{l}\sum_{j=1}^{l} \psi_A(e_j^{(f)})(t)$$

$$= \frac{1}{l}\sum_{j=1}^{l} F_j^s(t) - \psi_\mu(t) \; = \; 0 \; .$$

Damit erhalten wir:

$$\begin{aligned} \psi_\mu(t) \; &= \; \int \psi(x)(t) dG_A(x) \; = \; \frac{1}{l}\sum_{j=1}^{l} F_j^s(t) \; =: \; H(t)\,, \qquad (4.14) \\ \psi_A(e_j^{(f)})(t) \; &= \; \psi(e_j^{(f)})(t) - \psi_\mu(t) \; = \; F_j^s(t) - H(t)\; . \end{aligned}$$

HYPOTHESEN
Hypothese über den Faktor: $H_0^\psi(a) : \psi_A \equiv 0$.
Unter der Hypothese $H_0^\psi(a) \, : \, \psi_A \, \equiv \, 0$ gilt für alle $j \, = \, 1, \ldots, l \, : \;\; F_j(t) \, = \, H(t)$. Damit ist $\forall j : p_j := \int H(t)\, dF_j(t) = \frac{1}{2}$.
Eine schwächere Hypothese ließe sich also formulieren als $H_0^{\psi,p}(a) : \forall j : p_j = \frac{1}{2}$.

TESTSTATISTIK
Es werden zunächst Schätzer für die Funktionen ψ_μ, F_j^s und $\psi_A(e_j^{(f)})$ konstruiert:

$$\begin{aligned} \hat{\psi}_\mu(t) := \hat{H}(t) \; &:= \; \frac{1}{N}\sum_{\iota=1}^{N} c(t - Y_\iota), \qquad (4.15) \\ \hat{F}_j^s(t) := \frac{1}{n} &\sum_{\substack{\iota=1 \\ a_\iota = e_j^{(f)}}}^{N} c(t - Y_\iota) \quad \text{und} \\ \hat{\psi}_A(e_j^{(f)})(t) := \hat{F}_j^{(f)}(t) - \hat{H}(t)\; . \end{aligned}$$

Aus dem Vektor $\hat{\mathbf{p}} = (\hat{p}_1, \ldots, \hat{p}_l)'$, $\hat{p}_j = \int \hat{H}(t)\, d\hat{F}_j^s(t)$, der geschätzten sogenannten relativen Effekte werden dann quadratische Formen gebildet, die asymptotisch χ^2-verteilt sind oder deren Verteilung sich durch eine χ^2-Verteilung approximieren lässt. Details dazu findet man z.B. bei AKRITAS & BRUNNER (1997) und AKRITAS ET AL. (1997). Dort wird auch die Güte der mit dieser Methode hergeleiteten Tests ausführlich diskutiert.
Die Zerlegung $\mathbf{F} \, = \, \boldsymbol{\psi}_\mu + \boldsymbol{\psi}_A$ spiegelt sich wider in der vektoriellen Schreibweise für die relativen Effekte:

$$\begin{aligned} \mathbf{p} &= \int H \, d\mathbf{F} = \int \psi_\mu \, d(\boldsymbol{\psi}_\mu + \boldsymbol{\psi}_A) \quad \text{bzw.} \\ \hat{\mathbf{p}} &= \int \hat{H} \, d\hat{\mathbf{F}} = \int \hat{\psi}_\mu \, d(\hat{\boldsymbol{\psi}}_\mu + \hat{\boldsymbol{\psi}}_A)\, . \end{aligned}$$

Für höherfaktorielle Designs ergeben sich entsprechende Analoga.
Auf die Herleitung der Teststatistik gehen wir hier nicht im Detail ein, da die additivte Zerlegung dabei keine technischen Vereinfachungen mit sich bringt. Dies liegt daran, dass die quadratischen Formen aus Vektoren der Bauart

$$\sqrt{N}\mathbf{M}\hat{\mathbf{p}} = \sqrt{N}\mathbf{M}\int \hat{\psi}_\mu \, d(\hat{\boldsymbol{\psi}}_\mu + \hat{\boldsymbol{\psi}}_{A^{(1)}} + \hat{\boldsymbol{\psi}}_{A^{(2)}} + \ldots)$$

konstruiert werden, wobei sich die Integratorfunktion, im Gegensatz zum Linearen Modell, weder durch eine geeignete Wahl der Matrix **M** noch unter Hypothese vereinfacht, da sie aus Schätzern zusammengesetzt ist.

Kapitel 5

Varianzanalyse bei einer großen Anzahl an Faktorstufen

In diesem Kapitel betrachten wir etwas ausführlicher das in Abschnitt 4.2 bereits erwähnte varianzanalytische Modell ohne Normalverteilungsannahme. Wir leiten hier ein asymptotisches Resultat für den Fall $l \to \infty$, n fest, her, das eine weitgehende Verallgemeinerung einer von BOOS & BROWNIE (1995) bewiesenen Aussage darstellt.

Im weiteren Verlauf dieses Abschnittes gelten Modellgleichung (M1) (vgl. S.18) und die Annahmen (A1), (A2), (A2a) und (A3b) (vgl. S.39), außerdem sei $\mathcal{M} = \mathcal{M}_f$.

Wir zeigen nachfolgend asymptotische Normalität der ANOVA F-Test-Statistiken zum Testen beliebiger Haupteffekte und Wechselwirkungen unter der Annahme, dass die Anzahl der Faktorstufen gegen unendlich wächst, während die Anzahl der Wiederholungen pro Faktorstufe beschränkt ist.
Dies bedeutet für die Praxis, dass F-Tests auch für nicht normalverteilte Daten angewandt werden können, wenn die Anzahl der Behandlungsstufen groß ist.
Simulationsstudien belegen eine gute Übereinstimmung von tatsächlichem und nominellem Testniveau bereits ab etwa 10 Faktorstufen.

5.1 Einordnung der Thematik

Der Fall l (= Anzahl der Faktorstufen) $\to \infty$, n (= Anzahl der Versuchswiederholungen pro Faktorstufe) fest, tritt z.B. in Screening Versuchen in der agrarwissenschaftlichen Forschung auf. Dort werden viele verschiedene „Behandlungen“ (Pflanzensorten, Zuchtlinien, Pestizide, Herbizide, usw.) betrachtet, gewöhnlich in einem Design mit begrenzter Zahl an Versuchswiederholungen. Die Zielvariable kann ein Rating oder Score sein, deren Verteilung mehr oder weniger offensichtlich nicht die Normalverteilung ist (vgl. BROWNIE & BOOS (1994)).

Ein anderes Anwendungsgebiet ergibt sich bei Goodness-of-Fit Tests in der Item Response Theory. WRIGHT, MEAD & DRUBA (1976) haben einen Zugang zur Modellvalidierung für das Rasch-Modell vorgeschlagen, bei dem eine Varianzanalyse auf den Residuen durchgeführt wird. (Eine detaillierte Beschreibung des Rasch-Modells und eine Zusammenfassung verschiedener

Techniken zur Messung des Goodness-of-Fit findet man z.B. in dem Standardwerk von BAKER (1992).)

Im letzteren Kontext gibt es gewöhnlich viele Items und eine große Zahl an Prüflingen, die in viele „Fähigkeitsgruppen“ gruppiert werden. Beim Überprüfen des Modells stellen die Items und die Gruppen jeweils einen Faktor dar. Die Residuen sind ohne Zweifel nicht normalverteilt, da sie Differenzen zwischen Bernoulli-Variablen und Parameterschätzern sind.

Das asymptotische Verhalten der F-Test-Statistik im Falle $n \to \infty$, l fest, ist erschöpfend untersucht worden. Es ist allgemein bekannt, dass die F-Statistik, normiert bezüglich der Freiheitsgrade, asymptotisch χ^2-verteilt ist (siehe z.B. ARNOLD (1981), das Resultat wird dort mit Hilfe der Huber-Bedingung hergeleitet).
Der Fall $l \to \infty$, n fest, ist dagegen nicht so gut zugänglich, da die allgemeinen Resultate von ROTAR (1973) über die asymptotische Verteilung quadratischer Formen hier nicht mehr anwendbar sind.

Mit einem sehr allgemeinen Ansatz zur Untersuchung von M-Schätzern von Regressionsparametern hat PORTNOY (1984) ein Resultat über asymptotische Normalität der Schätzer in der einfaktoriellen Varianzanalyse mit p Stufen und n Beobachtungen pro Stufe bewiesen, falls $(\log p)^2/n \to 0$. PORTNOY scheint einer der ersten Autoren zu sein, der die Bedingung $p/n \to 0$ abschwächt, so dass die Zahl die Faktorstufen auch größer als die Zahl der Wiederholungen pro Stufe sein kann. Allerdings nimmt auch er an, dass beide Größen gegen unendlich gehen.

BOOS & BROWNIE (1995) bzw. BROWNIE & BOOS (1994) haben sich dem Fall $l \to \infty$, n fest, zugewandt und asymptotische Normalität der standardisierten ANOVA F-Statistik in einfaktoriellen Designs gezeigt. AKRITAS & ARNOLD (2000) haben dieses Resultat kürzlich um die Betrachtung eines zufälligen anstelle eines festen Faktors und die Analyse der Verteilungen unter Alternative ergänzt.

Mit Hilfe von Matrixkalkül und mittels einer speziellen Zerlegung der Projektionsmatrix erhalten wir ein allgemeines Resultat zur asymptotischen Normalität für ein Design mit beliebig vielen Faktoren (s. dazu auch BATHKE (2000)). Das heißt, die normalisierte F-Statistik ist asymptotisch ($l \to \infty$, n fest) normalverteilt, unabhängig von der zugrunde liegenden Verteilung des Fehlerterms $\phi_\varepsilon(\varepsilon)$. Damit ergeben sich die erwähnten Resultate für das einfaktorielle Design als Spezialfälle.

Nachfolgend beweisen wir zunächst die asymptotische Normalität für ein- und zweifaktorielle Designs. Damit wird die Idee demonstriert, wie der Beweis in einem Design mit beliebig vielen Faktoren durchzuführen ist. Das Hauptresultat dieses Kapitels ist daran anschließend in Satz 5.4 formuliert.

Von der ansonsten in dieser Arbeit verwendeten Schreibweise werden wir in diesem Kapitel etwas abweichen, indem wir Faktoren hier zum Teil mit Großbuchstaben bezeichnen. Mit den entsprechenden Kleinbuchstaben kennzeichnen wir die jeweilige Anzahl ihrer Faktorstufen.

5.2 Einfaktorielles Design

Das nachfolgende Lemma wird für die meisten der späteren Sätze benötigt. Das Resultat selbst ist anscheinend wohlbekannt, es wird z.B. in BOOS & BROWNIE (1995) benutzt. Betrachtet man es aus einer Sichtweise des Matrizenkalküls und der quadratischen Formen, so liegt die Wichtigkeit des Lemmas in der Möglichkeit, eine quadratische Form mit abhängigen Summanden in eine quadratische Form mit unabhängigen Summanden zuzüglich eines asymptotisch verschwindenden Ausdrucks zu transformieren.

Eine allgemeinere Formulierung der wesentlichen Bestandteile des Lemmas für ein Design mit m Faktoren erfolgt in Lemma A.7.

Lemma 5.1 *Sei* $\mathbf{Y} = \boldsymbol{\phi}_\mu + \boldsymbol{\phi}_\varepsilon$ *mit* $\mathbf{Y} = (Y_1, \ldots, Y_l)'$, $\boldsymbol{\phi}_\mu = \phi_\mu \cdot \mathbb{1}_l$, $\boldsymbol{\phi}_\varepsilon = (\phi_\varepsilon(\varepsilon_1), \ldots, \phi_\varepsilon(\varepsilon_l))'$ *und unabhängig identisch verteilten Zufallsvariablen* $\phi_\varepsilon(\varepsilon_1), \ldots, \phi_\varepsilon(\varepsilon_l)$.
Es gelte $\mathrm{E}(\phi_\varepsilon(\varepsilon_j)) = 0$, $\mathrm{Var}(\phi_\varepsilon(\varepsilon_j)) = \sigma^2 < \infty$. *Dann existiert ein* U *mit*

$$\frac{1}{l-1}\mathbf{Y}'\mathbf{P}_l\mathbf{Y} = \frac{1}{l}\boldsymbol{\phi}_\varepsilon'\boldsymbol{\phi}_\varepsilon - U \,, \qquad \mathrm{E}(U) = 0 \text{ und } \mathrm{Var}(U) = \frac{2\sigma^4}{l(l-1)} \,.$$

Beweis: Es gilt

$$\begin{aligned}
\frac{1}{l-1}\mathbf{Y}'\mathbf{P}_l\mathbf{Y} &= \frac{1}{l-1}(\boldsymbol{\phi}_\mu + \boldsymbol{\phi}_\varepsilon)'\mathbf{P}_l(\boldsymbol{\phi}_\mu + \boldsymbol{\phi}_\varepsilon) = \frac{1}{l-1}\boldsymbol{\phi}_\varepsilon'\mathbf{P}_l\boldsymbol{\phi}_\varepsilon \\
&= \frac{1}{l}\boldsymbol{\phi}_\varepsilon'\boldsymbol{\phi}_\varepsilon - U \qquad \text{mit} \\
U &= \frac{1}{l(l-1)}\boldsymbol{\phi}_\varepsilon'\mathbf{J}_l\boldsymbol{\phi}_\varepsilon - \frac{1}{l(l-1)}\boldsymbol{\phi}_\varepsilon'\boldsymbol{\phi}_\varepsilon \,.
\end{aligned}$$

Erwartungswert und Varianz von U berechnen sich zu

$$\begin{aligned}
\mathrm{E}(U) &= \mathrm{E}\left(\frac{2}{l(l-1)}\sum_{i<j}\phi_\varepsilon(\varepsilon_i)\phi_\varepsilon(\varepsilon_j)\right) \\
&= \frac{2}{l(l-1)}\sum_{i<j}\mathrm{E}(\phi_\varepsilon(\varepsilon_i)\phi_\varepsilon(\varepsilon_j)) \;=\; 0 \qquad \text{und} \\
\mathrm{Var}(U) = \mathrm{E}(U^2) \;&=\; \frac{4}{l^2(l-1)^2}\mathrm{E}\left(\sum_{i<j}\phi_\varepsilon(\varepsilon_i)\phi_\varepsilon(\varepsilon_j)\right)^2 \\
&= \frac{4}{l^2(l-1)^2}\mathrm{E}\left(\sum_{i_1<j_1}\sum_{i_2<j_2}\phi_\varepsilon(\varepsilon_{i_1})\phi_\varepsilon(\varepsilon_{j_1})\phi_\varepsilon(\varepsilon_{i_2})\phi_\varepsilon(\varepsilon_{j_2})\right) \\
&= \frac{4}{l^2(l-1)^2}\sum_{i<j}\mathrm{E}\left(\phi_\varepsilon(\varepsilon_i)^2\phi_\varepsilon(\varepsilon_j)^2\right) \\
&= \frac{4}{l^2(l-1)^2}\frac{l(l-1)}{2}\sigma^4 \;=\; \frac{2\sigma^4}{l(l-1)} \,.
\end{aligned}$$

□

Der Beweis von Satz 5.2 – die asymptotische Normalität im einfaktoriellen Design – stellt in seiner direkten Verallgemeinerbarkeit durch die konsequente Verwendung von Matrizen einen klareren Zugang verglichen mit BOOS & BROWNIE (1995) dar. In ihm werden bereits die wesentlichen Methoden der Beweise der nachfolgenden allgemeinen Resultate demonstriert.

Die F-Statistik zum Testen des Effekts eines Faktors ist der Quotient zweier quadratischer Formen. Dabei kann die quadratische Form im Zähler in die Summe unabhängig identisch verteilter Zufallsvariablen zuzüglich eines asymptotisch verschwindenden Terms zerlegt werden und erfüllt somit die Bedingungen für die Anwendung des zentralen Grenzwertsatzes. Da der Nenner ein konsistenter Schätzer für die Varianz σ^2 ist, erhält man asymptotische Normalität für den gesamten Ausdruck.

Satz 5.2 *Sei ein einfaktorielles Design gemäß (M1) und den Annahmen (A1), (A2), (A2a) und (A3b) gegeben. Der Faktor A habe die Stufen $1, \ldots, l$ und es gebe n Beobachtungen pro Stufe. Es gelte also $\mathbf{Y} = \boldsymbol{\phi}_\mu + \boldsymbol{\phi}_A + \boldsymbol{\phi}_\varepsilon$ mit $\mathbf{Y} = (Y_{1,1}, \ldots, Y_{1,n}, \ldots, Y_{l,n})'$ sowie $\boldsymbol{\phi}_\mu = \phi_\mu \cdot \mathbb{1}_{ln}$, $\boldsymbol{\phi}_A = (\phi_A(1) \cdot \mathbb{1}_n', \ldots, \phi_A(l) \cdot \mathbb{1}_n')'$ und $\boldsymbol{\phi}_\varepsilon = (\phi_\varepsilon(\varepsilon_{1,1}), \ldots, \phi_\varepsilon(\varepsilon_{l,n}))'$. Die $\phi_\varepsilon(\varepsilon_{i,j})$ seien unabhängig identisch verteilt mit $\mathrm{E}(\phi_\varepsilon(\varepsilon_{i,j})) = 0$ und $\mathrm{Var}(\phi_\varepsilon(\varepsilon_{i,j})) = \sigma^2 < \infty$.*
Sei

$$F = \frac{1}{l-1} Q_A / \frac{1}{l(n-1)} Q_\varepsilon \text{ mit } Q_A = \mathbf{Y}'(\mathbf{P}_l \otimes \frac{1}{n}\mathbf{J}_n)\mathbf{Y} \text{ und } Q_\varepsilon = \mathbf{Y}'(\mathbf{I}_l \otimes \mathbf{P}_n)\mathbf{Y} \,.$$

Dann gilt unter der Hypothese $H_0^\phi(A) : \phi_A \equiv 0$:

$$\sqrt{l}(F-1) \overset{\substack{l\to\infty \\ n \text{ fest}}}{\dot\sim} \mathrm{N}\big(0, \frac{2n}{n-1}\big) \tag{5.1}$$

Beweis:

$$\begin{aligned}
&\sqrt{l}(F-1)\\
&= \sqrt{l}\left(\frac{1}{l-1}Q_A - \frac{1}{l(n-1)}Q_\varepsilon\right) / \frac{1}{l(n-1)}Q_\varepsilon\\
&= \sqrt{l}\left(\frac{1}{l-1}\mathbf{Y}'(\mathbf{P}_l \otimes \frac{1}{n}\mathbf{J}_n)\mathbf{Y} - \frac{1}{l(n-1)}\mathbf{Y}'(\mathbf{I}_l \otimes \mathbf{P}_n)\mathbf{Y}\right) / \frac{1}{l(n-1)}Q_\varepsilon\\
&= \sqrt{l}\left(\frac{1}{l}\boldsymbol{\phi}_\varepsilon'(\mathbf{I}_l \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon - \frac{1}{l(n-1)}\boldsymbol{\phi}_\varepsilon'(\mathbf{I}_l \otimes \mathbf{P}_n)\boldsymbol{\phi}_\varepsilon - U\right) / \frac{1}{l(n-1)}Q_\varepsilon\\
&\qquad \text{mit } \mathrm{E}(U) = 0,\ \mathrm{Var}(U) = \frac{2\sigma^4}{l(l-1)} \quad \text{(mit Lemma A.7)}\\
&\overset{l\to\infty}{\doteq} \frac{1}{\sqrt{l}}\left(\boldsymbol{\phi}_\varepsilon'(\mathbf{I}_l \otimes (\frac{1}{n}\mathbf{J}_n - \frac{1}{n-1}\mathbf{P}_n))\boldsymbol{\phi}_\varepsilon\right) / \sigma^2\\
&\qquad \Big(\text{mit } \frac{1}{l(n-1)}Q_\varepsilon \xrightarrow{f.s.} \sigma^2 \Big)\\
&= \frac{1}{\sqrt{l}}\left(\boldsymbol{\phi}_\varepsilon'(\mathbf{I}_l \otimes (-\frac{1}{n-1}\mathbf{I}_n + \frac{1}{n-1}\mathbf{J}_n))\boldsymbol{\phi}_\varepsilon\right) / \sigma^2
\end{aligned}$$

$$= \frac{1}{\sqrt{l}}\left(\frac{1}{n-1}\boldsymbol{\phi}_\varepsilon'(\mathbf{I}_l \otimes \mathbf{J}_n)\boldsymbol{\phi}_\varepsilon - \frac{1}{n-1}\boldsymbol{\phi}_\varepsilon'(\mathbf{I}_l \otimes \mathbf{I}_n)\boldsymbol{\phi}_\varepsilon\right)/\sigma^2$$
$$= \frac{1}{\sqrt{l}}\sum_{i=1}^{l}\left(\frac{1}{n-1}\boldsymbol{\phi}_{\varepsilon,i}'(\mathbf{J}_n - \mathbf{I}_n)\boldsymbol{\phi}_{\varepsilon,i}\right)/\sigma^2$$
$$\text{mit } \boldsymbol{\phi}_{\varepsilon,i} := (\phi_\varepsilon(\varepsilon_{i,1}), \ldots, \phi_\varepsilon(\varepsilon_{i,n}))'$$

Da die

$$Z_i = \frac{1}{n-1}\boldsymbol{\phi}_{\varepsilon,i}'(\mathbf{J}_n - \mathbf{I}_n)\boldsymbol{\phi}_{\varepsilon,i}, \quad i = 1, \ldots, l,$$

unabhängig und identisch verteilte Zufallsvariablen mit E$(Z_i) = 0$ und Var$(Z_i) = 2\sigma^4\frac{n}{n-1}$ sind, kann der zentrale Grenzwertsatz angewandt werden und man erhält schließlich das Resultat (5.1). □

Für das einfaktorielle Design ist es unter Hypothese nicht maßgeblich, ob der Faktor A fest oder zufällig ist. In diesem Fall spielt also die Annahme $\mathcal{M} = \mathcal{M}_f$ keine Rolle, (5.1) bleibt auch bei $\mathcal{M} = \mathcal{M}_r$ richtig.

5.3 Zweifaktorielles Design

Mit den Methoden, die im Beweis von Satz 5.2 verwendet wurden, wird das Resultat nun auf zweifaktorielle Designs erweitert.
Es ist zu bemerken, dass hier für den Beweis der asymptotischen Normalität der F-Statistik zum Testen der Wechselwirkung zwischen den beiden Faktoren, die Anzahlen der Faktorstufen *beider* Faktoren gegen unendlich streben müssen (Teil c) von Satz 5.3).

Satz 5.3 *Sei ein zweifaktorielles Design gemäß (M1) und den Annahmen (A1), (A2), (A2a) und (A3b) gegeben. Die Faktorstufen der beiden Faktoren A und B seien $i = 1, \ldots, a$ bzw. $j = 1, \ldots, b$ und es gebe n Beobachtungen pro Faktorstufenkombination.*
Somit lässt sich das Design darstellen als $\mathbf{Y} = \boldsymbol{\phi}_\mu + \boldsymbol{\phi}_A + \boldsymbol{\phi}_B + \boldsymbol{\phi}_{A,B} + \boldsymbol{\phi}_\varepsilon$
mit $\mathbf{Y} = (Y_{1,1,1}, \ldots, Y_{a,b,n})'$, $\boldsymbol{\phi}_\mu = \phi_\mu \cdot \mathbb{1}_{abn}$, $\boldsymbol{\phi}_A = (\phi_A(1) \cdot \mathbb{1}_{bn}', \ldots, \phi_A(a) \cdot \mathbb{1}_{bn}')'$,
$\boldsymbol{\phi}_B = (\phi_B(1) \cdot \mathbb{1}_n', \ldots, \phi_B(b) \cdot \mathbb{1}_n')' \cdot \mathbb{1}_a$, $\boldsymbol{\phi}_{A,B} = (\phi_{A,B}(1,1) \cdot \mathbb{1}_n', \ldots, \phi_{A,B}(a,b) \cdot \mathbb{1}_n')'$ und
$\boldsymbol{\phi}_\varepsilon = (\phi_\varepsilon(\varepsilon_{1,1,1}), \ldots, \phi_\varepsilon(\varepsilon_{a,b,n}))'$. Die $\phi_\varepsilon(\varepsilon_{i,j,k})$ seien unabhängig identisch verteilt mit
E$(\phi_\varepsilon(\varepsilon_{i,j,k})) = 0$ und Var$(\phi_\varepsilon(\varepsilon_{i,j,k})) = \sigma^2 < \infty$.
Seien

$$F_A = \frac{1}{a-1}Q_A / \frac{1}{ab(n-1)}Q_\varepsilon, \qquad F_B = \frac{1}{b-1}Q_B / \frac{1}{ab(n-1)}Q_\varepsilon \qquad \textit{und}$$
$$F_{AB} = \frac{1}{(a-1)(b-1)}Q_{AB} / \frac{1}{ab(n-1)}Q_\varepsilon \qquad \textit{mit}$$
$$Q_A = \mathbf{Y}'(\mathbf{P}_a \otimes \frac{1}{b}\mathbf{J}_b \otimes \frac{1}{n}\mathbf{J}_n)\mathbf{Y}, \qquad Q_B = \mathbf{Y}'(\frac{1}{a}\mathbf{J}_a \otimes \mathbf{P}_b \otimes \frac{1}{n}\mathbf{J}_n)\mathbf{Y},$$

$$Q_{AB} = \mathbf{Y}'(\mathbf{P}_a \otimes \mathbf{P}_b \otimes \frac{1}{n}\mathbf{J}_n)\mathbf{Y} \qquad \textit{und} \qquad Q_\varepsilon = \mathbf{Y}'(\mathbf{I}_a \otimes \mathbf{I}_b \otimes \mathbf{P}_n)\mathbf{Y}\ .$$

Dann gilt,

a) *unter der Hypothese* $H_0^\phi(A) : \phi_A \equiv 0 :\ \sqrt{a}(F_A - 1) \overset{\substack{a\to\infty\\ b,n\ fest}}{\dot{\sim}} \mathrm{N}\big(0, 2(1 + \frac{1}{b(n-1)})\big)$,

b) *unter der Hypothese* $H_0^\phi(B) : \phi_B \equiv 0 :\ \sqrt{b}(F_B - 1) \overset{\substack{b\to\infty\\ a,n\ fest}}{\dot{\sim}} \mathrm{N}\big(0, 2(1 + \frac{1}{a(n-1)})\big)$,

c) *unter der Hypothese* $H_0^\phi(A,B) : \phi_{A,B} \equiv 0 :\ \sqrt{ab}(F_{AB} - 1) \overset{\substack{a,b\to\infty\\ n\ fest}}{\dot{\sim}} \mathrm{N}\big(0, \frac{2n}{n-1}\big)$.

Beweis:

a) Sei $\phi_A \equiv 0$. Daraus folgt $(\mathbf{P}_a \otimes \mathbf{J}_b \otimes \mathbf{J}_n)(\boldsymbol{\phi}_\mu + \boldsymbol{\phi}_A + \boldsymbol{\phi}_B + \boldsymbol{\phi}_{A,B}) = 0$. Somit gilt:

$$\begin{aligned}
&\sqrt{a}(F_A - 1)\\
&= \sqrt{a}\left(\frac{1}{a-1}\mathbf{Y}'(\mathbf{P}_a \otimes \frac{1}{b}\mathbf{J}_b \otimes \frac{1}{n}\mathbf{J}_n)\mathbf{Y} - \frac{1}{ab(n-1)}\mathbf{Y}'(\mathbf{I}_a \otimes \mathbf{I}_b \otimes \mathbf{P}_n)\mathbf{Y}\right)\Big/\frac{1}{ab(n-1)}Q_\varepsilon\\
&= \sqrt{a}\left(\frac{1}{a}\boldsymbol{\phi}_\varepsilon'(\mathbf{I}_a \otimes \frac{1}{b}\mathbf{J}_b \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon - \frac{1}{ab(n-1)}\boldsymbol{\phi}_\varepsilon'(\mathbf{I}_a \otimes \mathbf{I}_b \otimes \mathbf{P}_n)\boldsymbol{\phi}_\varepsilon - U\right)\Big/\frac{1}{ab(n-1)}Q_\varepsilon\\
&\qquad \text{mit } \mathrm{E}(U) = 0,\ \mathrm{Var}(U) = \frac{2\sigma^4}{a(a-1)} \quad \text{(mit Lemma A.7)}\\
&\overset{a\to\infty}{\doteq} \frac{1}{\sqrt{a}}\left(\boldsymbol{\phi}_\varepsilon'(\mathbf{I}_a \otimes \frac{1}{b}\mathbf{J}_b \otimes \frac{1}{n}\mathbf{J}_n - \frac{1}{b(n-1)}\mathbf{I}_a \otimes \mathbf{I}_b \otimes \mathbf{P}_n)\boldsymbol{\phi}_\varepsilon\right)\Big/\sigma^2\\
&\qquad \Big(\text{ mit } \frac{1}{ab(n-1)}Q_\varepsilon \xrightarrow{f.s.} \sigma^2\ \Big)\\
&= \frac{1}{\sqrt{a}}\left(\frac{1}{bn}\boldsymbol{\phi}_\varepsilon'(\mathbf{I}_a \otimes (\mathbf{J}_{bn} - \mathbf{I}_{bn} + \frac{1}{n-1}(\mathbf{I}_b \otimes (\mathbf{J}_n - \mathbf{I}_n))))\boldsymbol{\phi}_\varepsilon\right)\Big/\sigma^2\\
&= \frac{1}{\sqrt{a}}\sum_{i=1}^{a}\left(\frac{1}{bn}\boldsymbol{\phi}_{\varepsilon,i}'(\mathbf{J}_{bn} - \mathbf{I}_{bn} + \frac{1}{n-1}\mathbf{I}_b \otimes (\mathbf{J}_n - \mathbf{I}_n))\boldsymbol{\phi}_{\varepsilon,i}\right)\Big/\sigma^2\\
&\qquad \text{mit } \boldsymbol{\phi}_{\varepsilon,i} := (\phi_\varepsilon(\varepsilon_{i,1,1}), \ldots, \phi_\varepsilon(\varepsilon_{i,b,n}))'\ .
\end{aligned}$$

Die Zufallsvariablen

$$Z_i = \frac{1}{bn}\boldsymbol{\phi}_{\varepsilon,i}'(\mathbf{J}_{bn} - \mathbf{I}_{bn} + \frac{1}{n-1}\mathbf{I}_b \otimes (\mathbf{J}_n - \mathbf{I}_n))\boldsymbol{\phi}_{\varepsilon,i}, \quad i = 1, \ldots, a,$$

sind unabhängig und identisch verteilt mit $\mathrm{E}(Z_i) = 0$ und $\mathrm{Var}(Z_i) = 2\sigma^4(1 + \frac{1}{b(n-1)})$ (siehe Lemma A.9). Mit dem zentralen Grenzwertsatz erhält man das oben formulierte Resultat.

b) Analog.

c) Sei $\phi_{A,B} \equiv 0$ und somit $(\mathbf{P}_a \otimes \mathbf{P}_b \otimes \mathbf{J}_n)(\boldsymbol{\phi}_\mu + \boldsymbol{\phi}_A + \boldsymbol{\phi}_B + \boldsymbol{\phi}_{A,B}) = 0$. Es folgt:

$$\begin{aligned}
&\sqrt{ab}(F_{ab} - 1) \\
&= \sqrt{ab}\bigg(\frac{1}{(a-1)(b-1)} \mathbf{Y}'(\mathbf{P}_a \otimes \mathbf{P}_b \otimes \frac{1}{n}\mathbf{J}_n)\mathbf{Y} \\
&\qquad\qquad - \frac{1}{ab(n-1)} \mathbf{Y}'(\mathbf{I}_a \otimes \mathbf{I}_b \otimes \mathbf{P}_n)\mathbf{Y} \bigg) \Big/ \frac{Q_\varepsilon}{ab(n-1)} \\
&= \sqrt{ab}\bigg(\frac{1}{ab} \boldsymbol{\phi}'_\varepsilon(\mathbf{I}_{ab} \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon - \frac{1}{ab(n-1)} \boldsymbol{\phi}'_\varepsilon(\mathbf{I}_{ab} \otimes \mathbf{P}_n)\boldsymbol{\phi}_\varepsilon \\
&\qquad\qquad + U_1 - U_2 - U_3 \bigg) \Big/ \frac{Q_\varepsilon}{ab(n-1)}
\end{aligned}$$

$\Big[$ mit $U_1 = \dfrac{1}{ab(a-1)(b-1)} \boldsymbol{\phi}'_\varepsilon((\mathbf{J}_{ab} - \mathbf{I}_{ab}) \otimes \dfrac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon$,

$U_2 = \dfrac{1}{a(a-1)(b-1)} \boldsymbol{\phi}'_\varepsilon((\mathbf{J}_a - \mathbf{I}_a) \otimes \mathbf{I}_b \otimes \dfrac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon$,

und $U_3 = \dfrac{1}{b(a-1)(b-1)} \boldsymbol{\phi}'_\varepsilon(\mathbf{I}_a \otimes (\mathbf{J}_b - \mathbf{I}_b) \otimes \dfrac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon$.

Folglich gilt $\mathrm{E}(U_i) = 0, i = 1,2,3$, $\mathrm{Var}(U_1) = \dfrac{2\sigma^4(ab-1)}{ab(a-1)^2(b-1)^2}$,

$\mathrm{Var}(U_2) = \dfrac{2\sigma^4 b}{a(a-1)(b-1)^2}$ und $\mathrm{Var}(U_3) = \dfrac{2\sigma^4 a}{b(a-1)^2(b-1)}$ (mit Lemma A.7).$\Big]$

$$\begin{aligned}
&\overset{a,b\to\infty}{\doteq} \frac{1}{\sigma^2\sqrt{ab}} \bigg(\boldsymbol{\phi}'_\varepsilon(\mathbf{I}_{ab} \otimes (\frac{1}{n}\mathbf{J}_n - \frac{1}{n-1}\mathbf{P}_n)\boldsymbol{\phi}_\varepsilon) \bigg) \\
&\qquad \Big(\text{ mit } \frac{1}{ab(n-1)} Q_\varepsilon \xrightarrow{f.s.} \sigma^2 \Big) \\
&= \frac{1}{\sigma^2\sqrt{ab}} \bigg(\frac{1}{n-1} \boldsymbol{\phi}'_\varepsilon(\mathbf{I}_{ab} \otimes (\mathbf{J}_n - \mathbf{I}_n))\boldsymbol{\phi}_\varepsilon \bigg) \\
&= \frac{1}{\sigma^2\sqrt{ab}} \sum_{i,j} \bigg(\frac{1}{n-1} \boldsymbol{\phi}'_{\varepsilon,i,j}(\mathbf{J}_n - \mathbf{I}_n)\boldsymbol{\phi}_{\varepsilon,i,j} \bigg),
\end{aligned}$$

wobei $\boldsymbol{\phi}_{\varepsilon,i,j} := (\phi_\varepsilon(\varepsilon_{i,j,1}), \ldots, \phi_\varepsilon(\varepsilon_{i,j,n}))'$ ist.

Die Zufallsvariablen

$$Z_{ij} = \frac{1}{n-1} \boldsymbol{\phi}'_{\varepsilon,i,j}(\mathbf{J}_n - \mathbf{I}_n)\boldsymbol{\phi}_{\varepsilon,i,j}, \quad i = 1, \ldots, a,\ j = 1, \ldots, b,$$

sind unabhängig und identisch verteilt mit $\mathrm{E}(Z_{ij}) = 0$ und $\mathrm{Var}(Z_{ij}) = 2\sigma^4 \frac{n}{n-1}$. Unter Verwendung des zentralen Grenzwertsatzes erhält man Aussage c). $\square$

Die Idee der Zerlegung des Kronecker-Produktes zweier Projektionsmatrizen $\mathbf{P}_a$ und $\mathbf{P}_b$ kann zu der Zerlegung des Kronecker-Produktes beliebig vieler Matrizen $\mathbf{P}_{a_1}, \ldots, \mathbf{P}_{a_m}$ verallgemeinert werden. Darüber hinaus lässt sich, beim Beweis des Resultates für das m-faktorielle Design im nächsten Abschnitt, die Berechnung der Varianzen der unabhängig identisch verteilten Zufallsvariablen, die zur Anwendung des zentralen Grenzwertsatzes aufsummiert werden, auf den zweifaktoriellen Fall zurückführen.

5.4 m-faktorielles Design

In diesem Abschnitt wird das Resultat über die asymptotische Normalität der F-Statistik in seiner größtmöglichen Allgemeinheit, was die Komplexität des Designs betrifft, formuliert und bewiesen. Wir betrachten dabei eine *beliebige* Anzahl m an Faktoren und die F-Statistiken zum Testen *aller möglichen* Haupteffekte und Wechselwirkungen zwischen diesen Faktoren.

Satz 5.4 *Sei ein m-faktorielles Design gemäß (M1) und den Annahmen (A1), (A2), (A2a) und (A3b) gegeben. Die Faktorstufen der Faktoren $a^{(1)}, a^{(2)}, \ldots$ seien $j_1 = 1, \ldots, a_1,\ j_2 = 1, \ldots, a_2$ usw., es gebe n Beobachtungen pro Faktorstufenkombination. In Vektorschreibweise laute das Versuchsdesign*

$$\mathbf{Y} = \boldsymbol{\phi}_\mu + \sum_{f=1}^{m} \sum_{\{i_1,\ldots,i_f\}\subset\{1,\ldots,m\}} \boldsymbol{\phi}_{A^{(i_1)},\ldots,A^{(i_f)}} + \boldsymbol{\phi}_\varepsilon\,.$$

Dabei seien die Vektoren $\mathbf{Y}$, $\boldsymbol{\phi}_\mu$, $\boldsymbol{\phi}_{A^{(i_1)},\ldots,A^{(i_f)}}$ und $\boldsymbol{\phi}_\varepsilon$ entsprechend Satz 5.3 definiert. Die $\phi_\varepsilon(\varepsilon_{j_1,\ldots,j_m,k})$ seien unabhängig identisch verteilt mit

$$\mathrm{E}(\phi_\varepsilon(\varepsilon_{j_1,\ldots,j_m,k})) = 0 \quad \textit{und} \quad \mathrm{Var}(\phi_\varepsilon(\varepsilon_{j_1,\ldots,j_m,k})) = \sigma^2 < \infty\,.$$

Sei $F = \{f_1, \ldots, f_i\} \subset \{1, \ldots, m\}$ eine beliebige Teilmenge mit $\mathrm{card}(F) = i$ *und sei*

$$M = \bigotimes_{\iota=1}^{m} M_\iota, \quad \textit{mit } M_\iota = \begin{cases} \mathbf{P}_{a_\iota}, & \iota \in F \\ \mathbf{J}_{a_\iota}, & \iota \notin F \end{cases}\,.$$

Es definiere

$$F_{A_F} := \frac{1}{\tilde{a}} Q_{A_F} \Big/ \frac{1}{a(n-1)} Q_\varepsilon\,, \qquad \textit{mit} \quad \tilde{a} := \prod_{\iota\in F}(a_\iota - 1), \qquad a := \prod_{\iota=1}^{m} a_\iota,$$

$$Q_{A_F} := \Big(\prod_{\iota\notin F} a_\iota^{-1}\Big) \cdot \mathbf{Y}' M \mathbf{Y} \qquad \textit{und} \qquad Q_\varepsilon := \mathbf{Y}'\Big(\bigotimes_{\iota=1}^{m} \mathbf{I}_{a_\iota} \otimes \mathbf{P}_n\Big)\mathbf{Y}$$

und es gelte die Hypothese $H_0^\phi(a^{(f_1)}, \ldots, a^{(f_i)}) : \phi_{A^{(f_1)},\ldots,A^{(f_i)}} \equiv 0$, folglich

$$(M \otimes \mathbb{1}_n')\Big(\boldsymbol{\phi}_\mu + \sum_{f=1}^{m} \sum_{\{i_1,\ldots,i_f\}\subset\{1,\ldots,m\}} \boldsymbol{\phi}_{A^{(i_1)},\ldots,A^{(i_f)}}\Big) = 0\,. \tag{5.2}$$

Dann gilt

$$\sqrt{\prod_{\iota\in F} a_\iota}(F_{A_F}-1) \quad \overset{\substack{\min_{\iota\in F} a_\iota\to\infty \\ a_\iota \text{ fest für } (\iota\notin F) \\ n \text{ fest}}}{\dot\sim} \quad \mathrm{N}\big(0, 2(1+\frac{1}{(n-1)(\prod_{\iota\notin F} a_\iota)})\big). \tag{5.3}$$

Beweis: Ohne Einschränkung der Allgemeinheit nehmen wir an, dass $F=\{f_1,\dots,f_i\}=\{1,\dots,i\}$ ist.
Weiterhin definieren wir $a:=\prod_{\iota=1}^{i} a_\iota$, $\tilde a:=\prod_{\iota=1}^{i}(a_\iota-1)$, und $b:=\prod_{\iota=i+1}^{m} a_\iota$.
Dann kann (5.2) umgeschrieben werden als

$$\left(\bigotimes_{\iota=1}^{i}\mathbf{P}_{a_\iota}\otimes\bigotimes_{\iota=i+1}^{m}\mathbf{J}_{a_\iota}\otimes\mathbb{1}'_n\right)\left(\boldsymbol{\phi}_\mu+\sum_{f=1}^{m}\sum_{\{i_1,\dots,i_f\}\subset\{1,\dots,m\}}\boldsymbol{\phi}_{A^{(i_1)},\dots,A^{(i_f)}}\right)=0 \tag{5.4}$$

und wir erhalten das folgende Resultat:

$$\begin{aligned}
&\sqrt{a}(F_{A_F}-1)\\
&=\sqrt{a}\Bigg(\frac{1}{\tilde a}\mathbf{Y}'(\bigotimes_{\iota=1}^{i}\mathbf{P}_{a_\iota}\otimes\bigotimes_{\iota=i+1}^{m}\frac{1}{a_\iota}\mathbf{J}_{a_\iota}\otimes\frac{1}{n}\mathbf{J}_n)\mathbf{Y}\\
&\qquad\qquad -\frac{1}{(n-1)\prod_{\iota=1}^{m}a_\iota}\mathbf{Y}'(\bigotimes_{\iota=1}^{m}\mathbf{I}_{a_\iota}\otimes\mathbf{P}_n)\mathbf{Y}\Bigg)\Big/\frac{Q_\varepsilon}{(n-1)\prod_{\iota=1}^{m}a_\iota}\\
&\overset{(5.4)}{=}\sqrt{a}\Bigg(\frac{1}{a}\boldsymbol{\phi}'_\varepsilon(\bigotimes_{\iota=1}^{i}\mathbf{I}_{a_\iota}\otimes\bigotimes_{\iota=i+1}^{m}\frac{1}{a_\iota}\mathbf{J}_{a_\iota}\otimes\frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon\\
&\qquad\qquad -\frac{1}{(n-1)\prod_{\iota=1}^{m}a_\iota}\boldsymbol{\phi}'_\varepsilon(\bigotimes_{\iota=1}^{m}\mathbf{I}_{a_\iota}\otimes\mathbf{P}_n)\boldsymbol{\phi}_\varepsilon+U\Bigg)\Big/\frac{Q_\varepsilon}{(n-1)\prod_{\iota=1}^{m}a_\iota}\\
&\qquad\text{mit } \mathrm{E}(U)=0,\ \mathrm{Var}(U)\overset{\min_{\iota\le i}a_\iota\to\infty}{\longrightarrow}0\quad\text{(mit Lemma A.12 und Korollar A.14)}\\
&\overset{\min_{\iota\le i}a_\iota\to\infty}{\doteqdot}\frac{1}{\sqrt{a}}\Bigg(\boldsymbol{\phi}'_\varepsilon(\bigotimes_{\iota=1}^{i}\mathbf{I}_{a_\iota}\otimes(\frac{1}{bn}\mathbf{J}_{bn}-\frac{1}{b(n-1)}\mathbf{I}_{bn}+\frac{1}{bn(n-1)}\mathbf{I}_b\otimes\mathbf{J}_n))\boldsymbol{\phi}_\varepsilon\Bigg)\Big/\sigma^2\\
&=\frac{1}{\sigma^2\sqrt{a}}\Bigg(\frac{1}{bn}\boldsymbol{\phi}'_\varepsilon(\mathbf{I}_a\otimes(\mathbf{J}_{bn}-\mathbf{I}_{bn}+\frac{1}{n-1}\mathbf{I}_b\otimes(\mathbf{J}_n-\mathbf{I}_n)))\boldsymbol{\phi}_\varepsilon\Bigg)\\
&=\frac{1}{\sigma^2\sqrt{a}}\sum_{j=1}^{a}\Bigg(\frac{1}{bn}\boldsymbol{\phi}'_{\varepsilon,j}(\mathbf{J}_{bn}-\mathbf{I}_{bn}+\frac{1}{n-1}\mathbf{I}_b\otimes(\mathbf{J}_n-\mathbf{I}_n))\boldsymbol{\phi}_{\varepsilon,j}\Bigg)
\end{aligned}$$

mit $\boldsymbol{\phi}_{\varepsilon,j}=(\phi_\varepsilon(\varepsilon_{j,1,\dots,1,1}),\dots,\phi_\varepsilon(\varepsilon_{j,a_2,\dots,a_m,n}))'$.

Die Zufallsvariablen

$$Z_j=\frac{1}{bn}\boldsymbol{\phi}'_{\varepsilon,j}(\mathbf{J}_{bn}-\mathbf{I}_{bn}+\frac{1}{n-1}\mathbf{I}_b\otimes(\mathbf{J}_n-\mathbf{I}_n))\boldsymbol{\phi}_{\varepsilon,j},\quad j=1,\dots,a_1,$$

sind unabhängig und identisch verteilt mit $\mathrm{E}(Z_j) = 0$ und $\mathrm{Var}(Z_j) = 2\sigma^4(1+\frac{1}{b(n-1)})$ (analog zu den am Ende des Beweises von Satz 5.3a) eingeführten Zufallsvariablen Z_i, daher kann Lemma A.9 mit $b = \prod\limits_{\iota=i+1}^{m} a_\iota$ angewandt werden). Der zentrale Grenzwertsatz führt zu Resultat (5.3). □

5.5 Ungleiche Stichprobenumfänge und Varianzen

In den bisherigen Abschnitten dieses Kapitels haben wir immer unterstellt, dass balancierte Designs vorliegen, d.h. es gibt immer n Beobachtungen für jede Faktorstufenkombination. Außerdem haben wir, wie auch BOOS & BROWNIE (1995) und AKRITAS & ARNOLD (2000), für die Herleitung der asymptotischen Aussagen identisch verteilte, also insbesondere homoskedastische, Versuchsfehler gefordert. Im zweifaktoriellen Design bedeutet das:
$\forall i = 1, \ldots, a\ \forall j = 1, \ldots, b\ \forall k = 1, \ldots, n : \mathrm{Var}(\phi_\varepsilon(\varepsilon_{i,j,k})) = \sigma^2$.
Wir wollen in diesem Abschnitt kurz erläutern, welche Folgen es hat, wenn die beiden erwähnten Forderungen fallen gelassen werden.

Der Fall ungleicher Stichprobenumfänge ist schon von BOOS & BROWNIE (1995) betrachtet worden. Das leider etwas ernüchternde Ergebnis ihrer Untersuchungen ist, dass die asymptotische Varianz der normalisierten F-Statistik im Falle ungleicher Stichprobenumfänge nicht mehr verteilungsunabhängig ist. Dies kann höchstens durch Forderungen an die dritten Momente (Schiefe $= 0$) wieder erkauft werden, aber das wäre natürlich eine starke Einschränkung der betrachtbaren Verteilungsklasse.

Zur Heteroskedastizität haben sich weder BOOS & BROWNIE (1995) bzw. BROWNIE & BOOS (1994) noch AKRITAS & ARNOLD (2000) explizit geäußert. Für eine derartige Betrachtung ist es zunächst notwendig, Lemma 5.1 in geeigneter Weise allgemeiner zu formulieren:

Lemma 5.5 *Sei* $\mathbf{Y} = \boldsymbol{\phi}_\mu + \boldsymbol{\phi}_\varepsilon$ *mit* $\mathbf{Y} = (Y_1, \ldots, Y_l)'$, $\boldsymbol{\phi}_\mu = \phi_\mu \cdot \mathbb{1}_l$, $\boldsymbol{\phi}_\varepsilon = (\phi_\varepsilon(\varepsilon_1), \ldots, \phi_\varepsilon(\varepsilon_l))'$ *und unabhängigen Zufallsvariablen* $\phi_\varepsilon(\varepsilon_1), \ldots, \phi_\varepsilon(\varepsilon_l)$.
Es gelte $\mathrm{E}(\phi_\varepsilon(\varepsilon_j)) = 0$ *und* $\mathrm{Var}(\phi_\varepsilon(\varepsilon_j)) = \sigma_j^2 < \infty$.
Dann existiert ein U *mit*

$$\frac{1}{l-1}\mathbf{Y}'\mathbf{P}_l\mathbf{Y} = \frac{1}{l}\boldsymbol{\phi}_\varepsilon'\boldsymbol{\phi}_\varepsilon - U\ , \qquad \mathrm{E}(U) = 0 \textit{ und } \mathrm{Var}(U) = \frac{2}{l^2(l-1)^2}\sum_{i\neq j}\sigma_i^2\sigma_j^2\ .$$

Beweis: Der Beweis verläuft analog zu dem von Lemma 5.1. □

Es macht wenig Sinn, für einen Faktor mit großer bzw. gegen unendlich strebender Anzahl an Faktorstufen anzunehmen, dass der Versuchsfehler für jede Faktorstufe eine andere Varianz besitzt, da dann die Anzahl der unbekannten Verteilungsparameter ebenso groß wird bzw. unbeschränkt wächst, während für eine mögliche Schätzung der einzelnen Parameter nur eine feste, endliche Anzahl an Beobachtungen zur Verfügung steht. In solchen Fällen bleibt wohl lediglich

das Mittel einer geeigneten Variablentransformation.
Beispielhaft betrachten wir daher ein zweifaktorielles Design mit Faktorstufen $i = 1, \dots, a$ bzw. $j = 1, \dots, b$. Dabei habe der Versuchsfehler eine von j abhängige Varianz σ_j^2 und es gelte $a \to \infty$. Erwartungsgemäß verändert sich die Varianz der weiterhin asymptotisch normalverteilten Teststatistik. Für bestimmte Spezialfälle haben wir diese Ergebnisse explizit angegeben. Der nachfolgende Satz gibt zunächst allgemein Auskunft über die asymptotische Verteilung der F-Statistik im zweifaktoriellen Design mit unterschiedlichen Varianzen.

Satz 5.6 *Sei ein zweifaktorielles Design gemäß (M1) und den Annahmen (A1), (A2), (A2a) und (A3b) gegeben. Die Faktorstufen der beiden Faktoren A und B seien $i = 1, \dots, a$ bzw. $j = 1, \dots, b$ und es gebe n Beobachtungen pro Faktorstufenkombination. Das Design lässt sich also darstellen als $\mathbf{Y} = \boldsymbol{\phi}_\mu + \boldsymbol{\phi}_A + \boldsymbol{\phi}_B + \boldsymbol{\phi}_{A,B} + \boldsymbol{\phi}_\varepsilon$ mit $\mathbf{Y} = (Y_{1,1,1}, \dots, Y_{a,b,n})'$, $\boldsymbol{\phi}_\mu = \phi_\mu \cdot \mathbb{1}_{abn}$, $\boldsymbol{\phi}_A = (\phi_A(1) \cdot \mathbb{1}_{bn}', \dots, \phi_A(a) \cdot \mathbb{1}_{bn}')'$, $\boldsymbol{\phi}_B = (\phi_B(1) \cdot \mathbb{1}_n', \dots, \phi_B(b) \cdot \mathbb{1}_n')' \cdot \mathbb{1}_a$ sowie $\boldsymbol{\phi}_{A,B} = (\phi_{A,B}(1,1) \cdot \mathbb{1}_n', \dots, \phi_{A,B}(a,b) \cdot \mathbb{1}_n')'$ und $\boldsymbol{\phi}_\varepsilon = (\phi_\varepsilon(\varepsilon_{1,1,1}), \dots, \phi_\varepsilon(\varepsilon_{a,b,n}))'$.
Die $\phi_\varepsilon(\varepsilon_{i,j,k})$ seien für festes, gegebenes j unabhängig identisch verteilt mit $\mathrm{E}(\phi_\varepsilon(\varepsilon_{i,j,k})) = 0$ und $\mathrm{Var}(\phi_\varepsilon(\varepsilon_{i,j,k})) = \sigma_j^2 < \infty$.
Sei*

$$F_A = \frac{1}{a-1} Q_A \Big/ \frac{1}{ab(n-1)} Q_\varepsilon \qquad \textit{mit}$$
$$Q_A = \mathbf{Y}'(\mathbf{P}_a \otimes \frac{1}{b}\mathbf{J}_b \otimes \frac{1}{n}\mathbf{J}_n)\mathbf{Y} \qquad \textit{und} \qquad Q_\varepsilon = \mathbf{Y}'(\mathbf{I}_a \otimes \mathbf{I}_b \otimes \mathbf{P}_n)\mathbf{Y}\ .$$

Dann gilt, unter der Hypothese $H_0^\phi(A) : \phi_A \equiv 0$:

$$\sqrt{a}(F_A - 1) \overset{\substack{a\to\infty \\ b,n\ \mathit{fest}}}{\dot{\sim}} \mathrm{N}\Big(0, 2\big(1 + (\sum_{j=1}^b \sigma_j^2)^{-2}(\frac{1}{n-1}\sum_{j=1}^b \sigma_j^4)\big)\Big) \tag{5.5}$$

Beweis: Der Beweis verläuft im Wesentlichen analog zu dem von Satz 5.3a).
Sei $\phi_A \equiv 0$. Daraus folgt $(\mathbf{P}_a \otimes \mathbf{J}_b \otimes \mathbf{J}_n)(\boldsymbol{\phi}_\mu + \boldsymbol{\phi}_A + \boldsymbol{\phi}_B + \boldsymbol{\phi}_{A,B}) = 0$. Somit gilt:

$$\begin{aligned}
&\sqrt{a}(F_A - 1) \\
&= \sqrt{a}\left(\frac{1}{a-1}\mathbf{Y}'(\mathbf{P}_a \otimes \frac{1}{b}\mathbf{J}_b \otimes \frac{1}{n}\mathbf{J}_n)\mathbf{Y} - \frac{1}{ab(n-1)}\mathbf{Y}'(\mathbf{I}_a \otimes \mathbf{I}_b \otimes \mathbf{P}_n)\mathbf{Y}\right) \Big/ \frac{1}{ab(n-1)} Q_\varepsilon \\
&= \sqrt{a}\left(\frac{1}{a}\boldsymbol{\phi}_\varepsilon'(\mathbf{I}_a \otimes \frac{1}{b}\mathbf{J}_b \otimes \frac{1}{n}\mathbf{J}_n)\boldsymbol{\phi}_\varepsilon - \frac{1}{ab(n-1)}\boldsymbol{\phi}_\varepsilon'(\mathbf{I}_a \otimes \mathbf{I}_b \otimes \mathbf{P}_n)\boldsymbol{\phi}_\varepsilon - U\right) \Big/ \frac{1}{ab(n-1)} Q_\varepsilon \\
&\qquad \text{mit } \mathrm{E}(U) = 0,\ \mathrm{Var}(U) = \frac{2}{a^2(a-1)^2}\sum_{j\neq j'} \sigma_j^2\sigma_{j'}^2 \quad \text{(mit Lemma A.8)} \\
&\overset{a\to\infty}{\doteq} \frac{1}{\sqrt{a}}\left(\boldsymbol{\phi}_\varepsilon'(\mathbf{I}_a \otimes \frac{1}{b}\mathbf{J}_b \otimes \frac{1}{n}\mathbf{J}_n - \frac{1}{b(n-1)}\mathbf{I}_a \otimes \mathbf{I}_b \otimes \mathbf{P}_n)\boldsymbol{\phi}_\varepsilon\right) \Big/ \left(\frac{1}{b}\sum_{j=1}^b \sigma_j^2\right) \\
&\qquad \left(\text{ mit } \frac{1}{ab(n-1)} Q_\varepsilon \xrightarrow{f.s.} \frac{1}{b}\sum_{j=1}^b \sigma_j^2 =: \tilde{\sigma}^2\ \right)
\end{aligned}$$

$$= \frac{1}{\sqrt{a}}\left(\frac{1}{bn}\boldsymbol{\phi}'_\varepsilon(\mathbf{I}_a \otimes (\mathbf{J}_{bn} - \mathbf{I}_{bn} + \frac{1}{n-1}(\mathbf{I}_b \otimes (\mathbf{J}_n - \mathbf{I}_n))))\boldsymbol{\phi}_\varepsilon\right)/\tilde{\sigma}^2$$

$$= \frac{1}{\sqrt{a}}\sum_{i=1}^{a}\left(\frac{1}{bn}\boldsymbol{\phi}'_{\varepsilon,i}(\mathbf{J}_{bn} - \mathbf{I}_{bn} + \frac{1}{n-1}\mathbf{I}_b \otimes (\mathbf{J}_n - \mathbf{I}_n))\boldsymbol{\phi}_{\varepsilon,i}\right)/\tilde{\sigma}^2$$

$$\text{mit } \boldsymbol{\phi}_{\varepsilon,i} := (\phi_\varepsilon(\varepsilon_{i,1,1}), \dots, \phi_\varepsilon(\varepsilon_{i,b,n}))'$$

Die Zufallsvariablen

$$Z_i = \frac{1}{bn}\boldsymbol{\phi}'_{\varepsilon,i}(\mathbf{J}_{bn} - \mathbf{I}_{bn} + \frac{1}{n-1}\mathbf{I}_b \otimes (\mathbf{J}_n - \mathbf{I}_n))\boldsymbol{\phi}_{\varepsilon,i}, \quad i = 1, \dots, a,$$

sind unabhängig und identisch verteilt mit $\mathrm{E}(Z_i) = 0$ und $\mathrm{Var}(Z_i) = \frac{2}{b^2}\big(\frac{n}{n-1}\sum_{j=1}^{n}\sigma_j^4 + \sum_{j\neq j'}\sigma_j^2\sigma_{j'}^2\big)$ (siehe Lemma A.10). Daher lässt sich der zentrale Grenzwertsatz anwenden und es ergibt sich schließlich das Resultat (5.5). □

Wesentliches Ergebnis des Satzes ist, dass die asymptotische Varianz von $\sqrt{a}(F_A - 1)$ bei unterschiedlichen Varianzen in den Faktorstufen $j = 1, \dots, b$ von eben diesen Varianzen abhängt und nicht mehr in einer Form angegeben werden kann, in der lediglich die Anzahlen der Faktorstufen eine Rolle spielen.
In einigen Spezialfällen vereinfacht sich die allgemein berechnete asymptotische Varianz

$$V := 2\big(1 + (\sum_{j=1}^{b}\sigma_j^2)^{-2}(\frac{1}{n-1}\sum_{j=1}^{b}\sigma_j^4)\big) \tag{5.6}$$

allerdings erheblich, wie die nachfolgenden Beispiele zeigen:

1. $\sigma_1^2 = \ldots = \sigma_b^2 = \sigma^2 : \qquad V = 2(1 + \frac{1}{(n-1)b})$

   (es ergibt sich für diesen Spezialfall also genau das Resultat, das auch schon in Satz 5.3a) angegeben ist)

2. $\sigma_j^2 = q^{j-1}\sigma^2,\ j = 1, \dots, b : \qquad V = 2(1 + \frac{1}{n-1} \cdot \frac{q^{2b}-1}{q^2-1} \cdot \frac{(q-1)^2}{(q^b-1)^2})$

   (a) $b = 2 : \qquad V = 2(1 + \frac{1}{n-1} \cdot \frac{q^2+1}{(q+1)^2})$

   (b) $q = 2 : \qquad V = 2(1 + \frac{1}{n-1} \cdot \frac{4^b-1}{3(2^b-1)^2})$

3. $\sigma_j^2 = j \cdot \sigma^2,\ j = 1, \dots, b : \qquad V = 2(1 + \frac{1}{n-1} \cdot \frac{2}{3} \cdot \frac{2b+1}{b(b+1)})$

4. $\sigma_1^2 = \ldots = \sigma_k^2 = \sigma^2, \sigma_{k+1}^2 = \ldots = \sigma_b^2 = q \cdot \sigma^2 : \quad V = 2(1 + \frac{1}{n-1} \cdot \frac{k+(b-k)q^2}{(k+(b-k)q)^2})$

   (a) $b = 2k : \qquad V = 2(1 + \frac{1}{n-1} \cdot \frac{1}{k} \cdot \frac{q^2+1}{(q+1)^2})$

Es ist jedoch in den meisten Fällen davon auszugehen, dass Informationen über die Varianzstruktur, die eine Anwendung der expliziten Formeln ermöglichen, nicht vorliegen. Man wird sich daher im Allgemeinen mit der Abschätzung

$$2(1+\frac{1}{(n-1)b}) \leq V \leq 2(1+\frac{1}{n-1})$$

begnügen müssen, die sich direkt aus der allgemeinen Formel (5.6) ergibt. Diese Abschätzung lässt erkennen, dass eine Verletzung der Homoskedastizität in der beschriebenen Weise keine allzu dramatischen Auswirkungen auf die asymptotische Verteilung von $\sqrt{a}(F_A - 1)$ hat.

Die hier verwendete Methode lässt sich auf höherfaktorielle Designs ausweiten. An dieser Stelle verzichten wir jedoch auf eine entsprechende Herleitung.

5.6 Rangstatistiken

BOOS & BROWNIE (1995) haben für das einfaktorielle Design gezeigt, dass die asymptotische Verteilung der Statistik in (5.1) dieselbe bleibt, wenn man die originalen Zielvariablen durch ihre Ränge ersetzt. Allerdings setzen sie dazu stetige Verteilungen der Zufallsvariablen voraus. Ihr Beweis ist technisch sehr mühselig und wird auch von ihnen selbst als zum Teil „brute force calculation" (BOOS & BROWNIE (1995, S.188)) bezeichnet.
LANKOWSKI (1997) behandelt dasselbe Problem in einem anderen Zusammenhang und bestätigt das Resultat von BOOS & BROWNIE. Dabei benötigt sie nicht die Annahme stetiger Verteilungen und lässt somit auch Bindungen bei den betrachteten Zufallsvariablen zu. Den größten Teil des unvermeidlichen technischen Aufwandes lagert sie in den Beweis eines allgemeinen Resultats über asymptotische Normalität einer Klasse von Rangstatistiken in Gestalt quadratischer Formen aus. Das von ihr betrachtete einfaktorielle Design stellt dann ein mögliches Anwendungsbeispiel dieser allgemeinen Aussage dar.

Wir formulieren im nachfolgenden Satz das Resultat über die asymptotische Verteilung der Rangstatistik im einfaktoriellen Modell und skizzieren den Beweis, der prinzipiell analog zu dem von Satz 5.2 verläuft, aber auch entscheidend Gebrauch von der allgemeinen Aussage (aus LANKOWSKI (1997, Satz 3.5)) über Rangstatistiken in Gestalt quadratischer Formen macht.

Innerhalb dieses Kapitels stellt der nun folgende Satz einen Sonderfall dar, da das betrachtete Modell nicht mehr das Lineare Modell ohne Normalverteilungsannahme ist, sondern das allgemeine nichtparametrische Modell, in dem Hypothesen mittels der Verteilungsfunktionen ausgedrückt werden.

Satz 5.7 *Sei ein einfaktorielles Design im nichtparametrischen Modell gemäß (M2) und der Annahme (A2) gegeben. Der Faktor A habe die Stufen $1, \ldots, l$ und es gebe n Beobachtungen pro Stufe. Also $Y_{i,j} \sim F_i,\ i = 1, \ldots, l,\ j = 1, \ldots, n$, und $\mathbf{F}(t) = \boldsymbol{\psi}_\mu(t) + \boldsymbol{\psi}_A(t)$ mit*
$\mathbf{F} = (F_1, \ldots, F_l)'$, $\boldsymbol{\psi}_\mu = \psi_\mu \cdot \mathbb{1}_l$ und $\boldsymbol{\psi}_A = (\psi_A(1), \ldots, \psi_A(l))'$.
Die $Y_{i,j},\ i = 1, \ldots, l,\ j = 1, \ldots, n$, seien unabhängig.
Sei

$$F^R = \frac{1}{l-1} Q_A^R \Big/ \frac{1}{l(n-1)} Q_\varepsilon^R \text{ mit } Q_A^R = \mathbf{R}'(\mathbf{P}_l \otimes \frac{1}{n}\mathbf{J}_n)\mathbf{R} \text{ und } Q_\varepsilon^R = \mathbf{R}'(\mathbf{I}_l \otimes \mathbf{P}_n)\mathbf{R}\,,$$

dabei sei $\mathbf{R} = (R_{1,1}, \ldots, R_{l,n})'$ *und* $R_{i,j}$ *der (Mittel-)Rang von* $Y_{i,j}$ *unter* $Y_{1,1}, \ldots, Y_{l,n}$. *Dann gilt unter der Hypothese* $H_0^{\psi}(A) : \psi_A \equiv 0$*:*

$$\sqrt{l}(F^R - 1) \overset{\substack{l\to\infty \\ n \text{ fest}}}{\dot{\sim}} \mathrm{N}\big(0, \frac{2n}{n-1}\big) \tag{5.7}$$

Beweisskizze:
Seien H und $\hat{H}$ definiert wie in (4.14) bzw. (4.15). Weiterhin definieren wir

$$\tilde{Y}_{i,j} := H(Y_{i,j}),\ \hat{Y}_{i,j} := \hat{H}(Y_{i,j}),\ \tilde{\mathbf{Y}} := (\tilde{Y}_{1,1}, \ldots, \tilde{Y}_{l,n}),\ \hat{\mathbf{Y}} := (\hat{Y}_{1,1}, \ldots, \hat{Y}_{l,n}),$$
$$\hat{\mathbf{Y}}_\mu := \hat{\mathbf{Y}} - \frac{1}{2} \cdot \mathbb{1}_{ln},\ Q_A := \hat{\mathbf{Y}}'(\mathbf{P}_l \otimes \frac{1}{n}\mathbf{J}_n)\hat{\mathbf{Y}},$$
$$Q_\varepsilon := \hat{\mathbf{Y}}'(\mathbf{I}_l \otimes \mathbf{P}_n)\hat{\mathbf{Y}}, \ \text{ und } \ \tilde{Q}_\varepsilon := \tilde{\mathbf{Y}}'(\mathbf{I}_l \otimes \mathbf{P}_n)\tilde{\mathbf{Y}}\,.$$

Dann ist

$$\begin{aligned}
&\sqrt{l}(F^R - 1) \\
&= \sqrt{l}\left(\frac{1}{l-1}Q_A - \frac{1}{l(n-1)}Q_\varepsilon\right) \Big/ \frac{1}{l(n-1)}Q_\varepsilon \\
&= \sqrt{l}\left(\frac{1}{l-1}\hat{\mathbf{Y}}_\mu'(\mathbf{P}_l \otimes \frac{1}{n}\mathbf{J}_n)\hat{\mathbf{Y}}_\mu - \frac{1}{l(n-1)}\hat{\mathbf{Y}}_\mu'(\mathbf{I}_l \otimes \mathbf{P}_n)\hat{\mathbf{Y}}_\mu\right) \Big/ \frac{1}{l(n-1)}Q_\varepsilon \\
&= \sqrt{l}\left(\frac{1}{l}\hat{\mathbf{Y}}_\mu'(\mathbf{I}_l \otimes \frac{1}{n}\mathbf{J}_n)\hat{\mathbf{Y}}_\mu - \frac{1}{l(n-1)}\hat{\mathbf{Y}}_\mu'(\mathbf{I}_l \otimes \mathbf{P}_n)\hat{\mathbf{Y}}_\mu - U\right) \Big/ \frac{1}{l(n-1)}Q_\varepsilon \\
&\quad \text{mit } \sqrt{l} \cdot U \xrightarrow{p} 0
\end{aligned}$$

(die Herleitung dieser Konvergenzaussage ist allerdings aufwendiger als z.B. die des analogen Resultats im Beweis von Satz 5.2)

$$\begin{aligned}
&\overset{l\to\infty}{\doteq} \frac{1}{\sqrt{l}}\left(\frac{1}{n-1}\hat{\mathbf{Y}}_\mu'(\mathbf{I}_l \otimes (\mathbf{J}_n - \mathbf{I}_n))\hat{\mathbf{Y}}_\mu\right) \Big/ \sigma^2 \\
&\quad \text{mit } \sigma^2 := \mathrm{Var}(H(Y_{1,1}))
\end{aligned}$$

(der Nachweis von $\frac{1}{l(n-1)}Q_\varepsilon \xrightarrow{p} \sigma^2$ erfolgt in zwei Stufen, indem die beiden Aussagen $\frac{1}{l(n-1)}\tilde{Q}_\varepsilon \xrightarrow{p} \sigma^2$ und $\frac{1}{l(n-1)}Q_\varepsilon - \frac{1}{l(n-1)}\tilde{Q}_\varepsilon \xrightarrow{p} 0$ gezeigt werden)

Auf den Ausdruck $T_l = \frac{1}{n-1}\hat{\mathbf{Y}}_\mu'(\mathbf{I}_l \otimes (\mathbf{J}_n - \mathbf{I}_n))\hat{\mathbf{Y}}_\mu$ lässt sich nun Satz 3.5 aus LANKOWSKI (1997) anwenden und es ergibt sich das Resultat (5.7). □

Eine Verallgemeinerung der Aussage aus Satz 5.7 auf höherfaktorielle Designs ist prinzipiell möglich, indem die beiden entscheidenden Ideen der Zerlegung eines Kronecker-Produktes von Projektionsmatrizen (aus BATHKE (2000)) und der asymptotischen Normalität einer Klasse von Rangstatistiken (aus LANKOWSKI (1997)) kombiniert werden. Es ist allerdings zu beachten, dass die Forderung der Homoskedastizität aus Satz 5.4 sich hier in der Weise überträgt, dass die Varianz $\sigma_\iota^2 := \mathrm{Var}(H(Y_\iota))$ für alle $\iota = 1, \ldots, N$ als gleich angenommen wird. Dies ist in Anwendungen schwer zu überprüfen, da H eine unbekannte Funktion und $H(Y_\iota)$ nicht beobachtbar ist.

5.7 Simulationsergebnisse

Um genauere numerische Informationen über die „Robustheit" der F-Statistiken zu erhalten, wurden Computersimulationen mit SAS-IML durchgeführt. Es wurde dabei ein zweifaktorielles Design mit $a = 5, 7, 10, 15, 20$, $b = 4$ und $n = 4$ zugrundegelegt. Dabei seien a bzw. b die Anzahlen der Faktorstufen für den ersten bzw. zweiten Faktor und n die Anzahl der Versuchswiederholungen pro Faktorstufenkombination.

Der Test auf einen Effekt des ersten Faktors wird mit F_A bezeichnet, der für den zweiten Faktor mit F_B und der für die Wechselwirkung mit F_{AB}.

Als simulierte Verteilungen für die Zielgröße wurden die Normal-, Lognormal- und Bernoulli-Verteilung gewählt. Für jede Situation gab es 10.000 Simulationsdurchläufe.

Für die zugrundegelegte Normalverteilung ist die finite Verteilung der F-Statistik exakt eine F-Verteilung. Die für diesen Fall durchgeführten Simulationen dienen nur als nützlicher Vergleich für die Ergebnisse der Simulationen bei anderen zugrundegelegten Verteilungen.

Es zeigt sich, dass die bewiesene asymptotische Invarianz der F-Statistik bezüglich der Verteilung des Versuchsfehlers schon von praktischer Relevanz ist, wenn die Anzahl der Faktorstufen etwa 10 (Lognormalverteilung) oder sogar noch kleiner ist (Bernoulli-Verteilung). Da Robustheit gegenüber der Verletzung der Normalverteilungsannahme im Fall l klein, n groß, ein wohlbekanntes Resultat ist, ist es nicht überraschend, dass das Niveau α schnell für F_B eingehalten wird, wenn die Anzahl der Faktorstufen für den Faktor A wächst.

	Nominelles Niveau α	Zugrundegelegte Verteilung	a 5	7	10	15	20
F_A	90%	Normal	.9033	.9014	.9001	.8953	.8989
		Lognormal	.9123	.9163	.9102	.9131	.9108
		Bernoulli	.9009	.9019	.9057	.9052	.8973
	95%	Normal	.9503	.9492	.952	.9473	.9502
		Lognormal	.9577	.9609	.9545	.9577	.956
		Bernoulli	.9471	.9521	.9499	.9543	.9487
	99%	Normal	.9903	.9909	.9898	.9882	.9896
		Lognormal	.9919	.9927	.9899	.9898	.991
		Bernoulli	.9895	.9898	.9889	.9907	.9898
F_B	90%	Normal	.9016	.9023	.9034	.9006	.9048
		Lognormal	.9149	.9142	.9113	.9063	.9018
		Bernoulli	.8983	.8985	.8969	.8998	.9015
	95%	Normal	.9491	.9481	.9504	.9516	.9516
		Lognormal	.9604	.9608	.9582	.9561	.9526
		Bernoulli	.9466	.949	.9501	.9501	.9475
	99%	Normal	.9886	.9883	.9899	.9903	.9905
		Lognormal	.9923	.9925	.994	.992	.9924
		Bernoulli	.9878	.9902	.9888	.9893	.9905
F_{AB}	90%	Normal	.8974	.9002	.8989	.8974	.8962
		Lognormal	.9171	.9132	.9085	.9129	.902
		Bernoulli	.8916	.9006	.8975	.8988	.8976
	95%	Normal	.9497	.9499	.9539	.9486	.9499
		Lognormal	.9595	.955	.9512	.9547	.9472
		Bernoulli	.9461	.9477	.9496	.9506	.948
	99%	Normal	.9881	.9902	.9907	.9902	.9891
		Lognormal	.9897	.9892	.987	.9878	.984
		Bernoulli	.9896	.9887	.9879	.9894	.9893

Tabelle 3 *Simulierte Testniveaus für die F-Statistik, verglichen mit F-Quantilen, bei zugrundegelegter Normal-, Lognormal- und Bernoulli-Verteilung ($b = 4, n = 4$).*

Anhang A

Technische Resultate

Satz A.1 (Lancaster) *Sei* $\mathbf{X} = (X_1, \dots, X_k)'$ *Zufallsvektor mit* $E(\mathbf{X}) = \boldsymbol{\mu} = (\mu_1, \dots, \mu_k)'$ *und* $\mathbf{V} = Var(\mathbf{X})$. *Ferner sei* $\mathbf{A} = \mathbf{A}'$ *und* $Sp(\cdot)$ *bezeichne die Spur einer Matrix. Dann gilt* $E(\mathbf{X}'\mathbf{A}\mathbf{X}) = Sp(\mathbf{A}\mathbf{V}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$.

In den anschließenden Sätzen A.2 und A.4 sowie Korollar A.3 sei jeweils für eine Matrix $\mathbf{V}$ mit $r(\mathbf{V})$ ihr Rang bezeichnet.

Satz A.2 (Verteilung einer Quadratform) *Sei* $\mathbf{X} \sim N(\mathbf{0}, \mathbf{V})$, $\mathbf{A} = \mathbf{A}'$ *eine* $n \times n$ *Matrix und* $r(\mathbf{V}) = r \leq n$. *Dann gilt:*

$$\mathbf{X}'\mathbf{A}\mathbf{X} \;\sim\; \sum_{i=1}^{n} \lambda_i C_i \,,$$

wobei die C_i *u.i.v.* $\sim \chi_1^2$ *und die* λ_i *die Eigenwerte von* $\mathbf{A}\mathbf{V}$, $i = 1, \dots, n$ *sind.*

Korollar A.3 *Sei* $\mathbf{X} \sim N(\mathbf{0}, \mathbf{V})$, $\mathbf{A} = \mathbf{A}'$ *und* $\mathbf{A}\mathbf{V}$ *idempotent. Dann ist*

$$\mathbf{X}'\mathbf{A}\mathbf{X} \;\sim\; \chi^2_{r(\mathbf{A}\mathbf{V})} \,.$$

Satz A.4 *Sei* $\mathbf{X} \sim N(\boldsymbol{\mu}, \mathbf{V})$ *mit* $|\mathbf{V}| \neq 0$, $\mathbf{A}_{n\times n} = \mathbf{A}'$ *und* $r = r(\mathbf{A}\mathbf{V})$.
Dann ist $\mathbf{X}'\mathbf{A}\mathbf{X} \;\sim\; \chi^2(r, \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu})$ *genau dann, wenn* $\mathbf{A}\mathbf{V}$ *idempotent ist.*

Satz A.5 (Craig und Sakamoto) *Seien* $\mathbf{X} \sim N(\boldsymbol{\mu}, \mathbf{V})$, $\mathbf{A} = \mathbf{A}'$ *positiv semidefinit,* $\mathbf{B} = \mathbf{B}'$ *positiv semidefinit und* $\mathbf{b}$ *ein konstanter Vektor. Dann gilt*

(1) $\mathbf{X}'\mathbf{A}\mathbf{X}$ *und* $\mathbf{X}'\mathbf{B}\mathbf{X}$ *bzw.* $\mathbf{X}'\mathbf{A}\mathbf{X}$ *und* $\mathbf{B}\mathbf{X}$ *sind stochastisch unabhängig, falls* $\mathbf{B}\mathbf{V}\mathbf{A} = \mathbf{0}$ *ist,*

(2) $\mathbf{A}\mathbf{X}$ *und* $\mathbf{b}'\mathbf{X}$ *sind stochastisch unabhängig, falls* $\mathbf{b}'\mathbf{V}\mathbf{A} = \mathbf{0}$ *ist.*

Satz A.6 *Sei $x \in \mathbb{R}^m$ und $G(x)$ die Verteilungsfunktion eines Wahrscheinlichkeitsmaßes auf $(\mathbb{R}^m, \mathcal{B}^m)$. Sei weiter $\psi(x,t)$ für festes $t \in \mathbb{R}$ eine $(\mathbb{R}^m, \mathcal{B}^m) - (\mathbb{R}, \mathcal{B})$-meßbare Funktion und für festes $x \in \mathbb{R}^m$ die (rechtsstetige Version der) Verteilungsfunktion eines Wahrscheinlichkeitsmaßes auf $(\mathbb{R}, \mathcal{B})$. Dann gilt für $\mathcal{B}$-meßbares $f(t) : \mathbb{R} \to \mathbb{C}$ mit f für beliebiges $x \in \mathbb{R}^m$ integrierbar bezüglich $\psi(x,\cdot)$ und $\int f(t)\, d\psi(x,t)$ integrierbar bezüglich $G(x)$:*

$$\int \int f(t)\, d\psi(x,t)\, dG(x) = \int f(t)\, d\Big[\int \psi(x,t)\, dG(x)\Big]\,.$$

Beweis:

a) Für beliebiges x ist $\psi(x,t)$ als Verteilungsfunktion eines Wahrscheinlichkeitsmaßes auf $(\mathbb{R}, \mathcal{B})$ isoton in t. Daraus folgt die Isotonie von $\int \psi(x,t)\, dG(x)$ in t. Somit definiert $\int \psi(x,t)\, dG(x)$ ein Maß auf $(\mathbb{R}, \mathcal{B})$.
Im Falle von Unstetigkeitsstellen von $\int \psi(x,t)\, dG(x)$ sei im Folgenden stets die rechtsstetige Version gemeint.

b) Wir zeigen die Gleichheit der beiden Integrale zunächst für reellwertige Elementarfunktionen f. Sei $f(t) = \mathbb{I}_{[a,b]}(t)$, dann gilt:

$$\begin{aligned}
\int f(t)\, d\Big[\int \psi(x,t)\, dG(x)\Big] \;&=\; \int \mathbb{I}_{[a,b]}(t)\, d\Big[\int \psi(x,t)\, dG(x)\Big] \\
=\;& \int \psi(x,b^-)\, dG(x) - \int \psi(x,a)\, dG(x) + \int \psi(x,b^+)\, dG(x) - \int \psi(x,b^-)\, dG(x) \\
=\;& \int (\psi(x,b^-) - \psi(x,a) + \psi(x,b^+) - \psi(x,b^-))\, dG(x) \\
=\;& \int\int \mathbb{I}_{[a,b]}(t)\, d\psi(x,t)\, dG(x) \;=\; \int\int f(t)\, d\psi(x,t)\, dG(x)\,.
\end{aligned}$$

c) Sei nun f eine reellwertige Treppenfunktion, d.h. $f(t) = \sum\limits_{i=1}^{n} a_i \cdot \mathbb{I}_{[l_i,r_i]}(t)$. Dann folgt:

$$\begin{aligned}
\int f(t)\, d\Big[\int \psi(x,t)\, dG(x)\Big] \;&=\; \int \sum_{i=1}^{n} a_i \cdot \mathbb{I}_{[l_i,r_i]}(t)\, d\Big[\int \psi(x,t)\, dG(x)\Big] \\
=\;& \sum_{i=1}^{n} a_i \cdot \int \mathbb{I}_{[l_i,r_i]}(t)\, d\Big[\int \psi(x,t)\, dG(x)\Big] \\
=\;& \sum_{i=1}^{n} a_i \cdot \int\int \mathbb{I}_{[l_i,r_i]}(t)\, d\psi(x,t)\, dG(x) \quad \text{(mit b))} \\
=\;& \int\int \sum_{i=1}^{n} a_i \cdot \mathbb{I}_{[l_i,r_i]}(t)\, d\psi(x,t)\, dG(x) \;=\; \int\int f(t)\, d\psi(x,t)\, dG(x)\,.
\end{aligned}$$

d) Die Treppenfunktionen bilden einen Stoneschen Vektorverband auf $\mathbb{R}$. Mit dem Satz von Daniell-Stone erhalten wir damit Gleichheit der beiden Integrale für alle $\mathcal{B}$-meßbaren, reellwertigen Funktionen f mit den oben genannten Messbarkeitseigenschaften.
Für komplexwertige Funktionen ergibt sich das Resultat sofort durch Aufspalten in Real- und Imaginärteil.

□

Das nachfolgende Lemma ist eine Verallgemeinerung der wesentlichen Aussage von Lemma 5.1.

Lemma A.7 *Sei* $\boldsymbol{\phi}_\varepsilon = (\phi_\varepsilon(\varepsilon_{1,1,1}), \ldots, \phi_\varepsilon(\varepsilon_{a,b,n}))'$*. Die* $\phi_\varepsilon(\varepsilon_{i,j,k})$ *seien unabhängig identisch verteilt mit* $\mathrm{E}(\phi_\varepsilon(\varepsilon_{i,j,k})) = 0$ *und* $\mathrm{Var}(\phi_\varepsilon(\varepsilon_{i,j,k})) = \sigma^2$.
Sei weiter

$$U = \frac{1}{b(b-1)} \boldsymbol{\phi}'_\varepsilon (\mathbf{I}_a \otimes \mathbf{J}_b \otimes \frac{1}{n}\mathbf{J}_n - \mathbf{I}_a \otimes \mathbf{I}_b \otimes \frac{1}{n}\mathbf{J}_n) \boldsymbol{\phi}_\varepsilon \, .$$

Dann gilt $\mathrm{E}(U) = 0$ *und* $\mathrm{Var}(U) = \frac{2a\sigma^4}{b(b-1)}$.

Beweis: Direkte Rechnung unter Verwendung von Lemma 5.1. □

Für den Fall unterschiedlicher Varianzen, der in Abschnitt 5.5 betrachtet wird, formulieren wir das Resultat in einer Version, die direkt auf die Verwendung in Satz 5.6 zugeschnitten ist.

Lemma A.8 *Sei* $\boldsymbol{\phi}_\varepsilon = (\phi_\varepsilon(\varepsilon_{1,1,1}), \ldots, \phi_\varepsilon(\varepsilon_{a,b,n}))'$*. Die* $\phi_\varepsilon(\varepsilon_{i,j,k})$ *seien für festes* j *unabhängig identisch verteilt mit* $\mathrm{E}(\phi_\varepsilon(\varepsilon_{i,j,k})) = 0$ *und* $\mathrm{Var}(\phi_\varepsilon(\varepsilon_{i,j,k})) = \sigma_j^2$.
Sei weiter

$$U = \frac{1}{a(a-1)} \boldsymbol{\phi}'_\varepsilon (\mathbf{J}_a \otimes \frac{1}{b}\mathbf{J}_b \otimes \frac{1}{n}\mathbf{J}_n - \mathbf{I}_a \otimes \frac{1}{b}\mathbf{J}_b \otimes \frac{1}{n}\mathbf{J}_n) \boldsymbol{\phi}_\varepsilon \, .$$

Dann gilt $\mathrm{E}(U) = 0$ *und* $\mathrm{Var}(U) = \frac{2}{a^2(a-1)^2} \sum_{j \neq j'} \sigma_j^2 \sigma_{j'}^2$.

Beweis: Direkte Rechnung unter Verwendung von Lemma 5.5. □

Das nachfolgende Lemma ist notwendig zur Berechnung der Varianz der unabhängig identisch verteilten Zufallsvariablen, die für die Anwendung des zentralen Grenzwertsatzes im Beweis der Sätze 5.3 und 5.4 aufsummiert werden. Es genügt, die Varianz im zweifaktoriellen Design auszurechnen, weil das Problem der Berechnung im m-faktoriellen Fall auf den zweifaktoriellen Fall zurückgeführt werden kann.

Lemma A.9 *Seien die Bezeichnungen wir in Satz 5.3 und sei Z_i definiert wie am Ende des Beweises von Satz 5.3a). Dann gilt* $\mathrm{E}(Z_i) = 0$ *und* $\mathrm{Var}(Z_i) = 2\sigma^4(1 + \frac{1}{b(n-1)})$.

Beweis: $\mathrm{E}(Z_i) = 0$ ist klar. Damit ist $\mathrm{Var}(Z_i) = \mathrm{E}(Z_i^2)$. Wir zerlegen nun

$$Z_i = \frac{1}{bn}\boldsymbol{\phi}'_{\varepsilon,i}(\mathbf{J}_{bn} - \mathbf{I}_{bn} + \frac{1}{n-1}\mathbf{I}_b \otimes (\mathbf{J}_n - \mathbf{I}_n))\boldsymbol{\phi}_{\varepsilon,i}$$

in $Z_i = A + B$ mit

$$\begin{aligned} A &= \frac{1}{bn}\boldsymbol{\phi}'_{\varepsilon,i}(\mathbf{J}_{bn} - \mathbf{I}_{bn})\boldsymbol{\phi}_{\varepsilon,i} \\ &= \frac{1}{bn} \sum_{(j,k)\neq(j',k')} \phi_\varepsilon(\varepsilon_{i,j,k})\phi_\varepsilon(\varepsilon_{i,j',k'}) \\ &\qquad \text{und} \\ B &= \frac{1}{bn(n-1)}\boldsymbol{\phi}'_{\varepsilon,i}(\mathbf{I}_b \otimes (\mathbf{J}_n - \mathbf{I}_n))\boldsymbol{\phi}_{\varepsilon,i} \\ &= \frac{2}{bn(n-1)} \sum_j \sum_{k<k'} \phi_\varepsilon(\varepsilon_{i,j,k})\phi_\varepsilon(\varepsilon_{i,j,k'}) \ . \end{aligned}$$

Somit reduziert sich das Problem auf die Berechnung von $\mathrm{E}(A^2)$, $\mathrm{E}(AB)$ und $\mathrm{E}(B^2)$.

$$\begin{aligned} \mathrm{E}(A^2) &= \frac{1}{b^2n^2} \sum_{(j_1,k_1)\neq(j_1',k_1')} \sum_{(j_2,k_2)\neq(j_2',k_2')} \mathrm{E}\Big(\phi_\varepsilon(\varepsilon_{i,j_1,k_1})\phi_\varepsilon(\varepsilon_{i,j_1',k_1'})\phi_\varepsilon(\varepsilon_{i,j_2,k_2})\phi_\varepsilon(\varepsilon_{i,j_2',k_2'})\Big) \\ &= \frac{2(bn-1)}{bn}\sigma^4 \ , \end{aligned}$$

$$\begin{aligned} \mathrm{E}(B^2) &= \frac{4}{b^2n^2(n-1)^2} \sum_{j_1} \sum_{k_1<k_1'} \sum_{j_2} \sum_{k_2<k_2'} \mathrm{E}\Big(\phi_\varepsilon(\varepsilon_{i,j_1,k_1})\phi_\varepsilon(\varepsilon_{i,j_1,k_1'})\phi_\varepsilon(\varepsilon_{i,j_2,k_2})\phi_\varepsilon(\varepsilon_{i,j_2,k_2'})\Big) \\ &= \frac{2}{bn(n-1)}\sigma^4 \qquad \text{und} \end{aligned}$$

$$\begin{aligned} \mathrm{E}(AB) &= \frac{2}{b^2n^2(n-1)} \sum_{(j_1,k_1)\neq(j_1',k_1')} \sum_{j_2} \sum_{k_2<k_2'} \mathrm{E}\Big(\phi_\varepsilon(\varepsilon_{i,j_1,k_1})\phi_\varepsilon(\varepsilon_{i,j_1',k_1'})\phi_\varepsilon(\varepsilon_{i,j_2,k_2})\phi_\varepsilon(\varepsilon_{i,j_2,k_2'})\Big) \\ &= \frac{2}{bn}\sigma^4 \ . \end{aligned}$$

Indem diese Ausdrücke zusammengefasst werden, erhält man $\mathrm{Var}(Z_i) = 2\sigma^4(1 + \frac{1}{b(n-1)})$. □

Im Falle unterschiedlicher Varianzen, der in Satz 5.6 betrachtet wird, ergibt sich das folgende Resultat.

Lemma A.10 *Seien die Bezeichnungen wir in Satz 5.6 und sei Z_i definiert wie am Ende des Beweises von Satz 5.6. Dann gilt* $\mathrm{E}(Z_i) = 0$ *und* $\mathrm{Var}(Z_i) = \frac{2}{b^2}\big(\frac{n}{n-1}\sum\limits_{j=1}^{n}\sigma_j^4 + \sum\limits_{j\neq j'}\sigma_j^2\sigma_{j'}^2\big)$.

Beweis: $\mathrm{E}(Z_i) = 0$ ist klar. Damit ist $\mathrm{Var}(Z_i) = \mathrm{E}(Z_i^2)$. Analog zum Beweis von Lemma A.9 zerlegen wir

$$Z_i = \frac{1}{bn}\boldsymbol{\phi}'_{\varepsilon,i}(\mathbf{J}_{bn} - \mathbf{I}_{bn} + \frac{1}{n-1}\mathbf{I}_b \otimes (\mathbf{J}_n - \mathbf{I}_n))\boldsymbol{\phi}_{\varepsilon,i}$$

in $Z_i = A + B$ mit

$$\begin{aligned} A &= \frac{1}{bn}\boldsymbol{\phi}'_{\varepsilon,i}(\mathbf{J}_{bn} - \mathbf{I}_{bn})\boldsymbol{\phi}_{\varepsilon,i} \qquad \text{und} \\ B &= \frac{1}{bn(n-1)}\boldsymbol{\phi}'_{\varepsilon,i}(\mathbf{I}_b \otimes (\mathbf{J}_n - \mathbf{I}_n))\boldsymbol{\phi}_{\varepsilon,i} \, . \end{aligned}$$

Es bleiben wieder die Größen $\mathrm{E}(A^2)$, $\mathrm{E}(AB)$ und $\mathrm{E}(B^2)$ zu berechnen und dafür ergibt sich

$$\begin{aligned} \mathrm{E}(A^2) &= \frac{2}{b^2n^2}\sum_{(j,k)\neq(j',k')}\sigma_j^2\sigma_{j'}^2 = \frac{2}{b^2}\sum_{j\neq j'}\sigma_j^2\sigma_{j'}^2 + \frac{2(n-1)}{b^2n}\sum_{j=1}^{b}\sigma_j^4 \, , \\ \mathrm{E}(B^2) &= \frac{2}{b^2n(n-1)}\sum_{j=1}^{b}\sigma_j^4 \quad \text{und} \quad \mathrm{E}(AB) = \frac{2}{b^2n}\sum_{j=1}^{b}\sigma_j^4 \, . \end{aligned}$$

Damit erhält man schließlich die angegebene Varianz von Z_i. □

Um die quadratischen Formen mit abhängigen Summanden in solche mit unabhängigen Summanden zuzüglich eines asymptotisch verschwindenden Terms zu transformieren, muss ein Kronecker-Produkt von Projektor-Matrizen zerlegt werden. Dies geschieht mit Hilfe der nachfolgenden beiden Lemmata.

Lemma A.11

$$\bigotimes_{\iota=1}^{i}\mathbf{P}_{a_\iota} = \sum_{j=0}^{i}(-1)^j \sum_{\{\nu_1,\ldots,\nu_j\}\subset\{1,\ldots,i\}} M(\nu_1,\ldots,\nu_j),$$

$$\textit{mit} \qquad M(\nu_1,\ldots,\nu_j) = \bigotimes_{\iota=1}^{i} M_\iota \qquad \textit{und} \qquad M_\iota = \begin{cases} a_\iota^{-1}\cdot\mathbf{J}_{a_\iota}, & \iota \in \{\nu_1,\ldots,\nu_j\} \\ \mathbf{I}_{a_\iota}, & \textit{sonst.} \end{cases}$$

Beweis: Direkte Rechnung. □

Lemma A.12

$$\bigotimes_{\iota=1}^{i} \mathbf{P}_{a_\iota} = \frac{\prod\limits_{\iota=1}^{i}(a_\iota - 1)}{\prod\limits_{\iota=1}^{i} a_\iota} \mathbf{I}_{\prod\limits_{\iota=1}^{i} a_\iota} + \sum_{j=1}^{i}(-1)^j \sum_{\{\nu_1,\dots,\nu_j\}\subset\{1,\dots,i\}} (\prod_{k=1}^{j} \nu_k^{-1}) \tilde{M}(\nu_1,\dots,\nu_j),$$

$$\textit{mit} \qquad \tilde{M}(\nu_1,\dots,\nu_j) = \bigotimes_{\iota=1}^{i} \tilde{M}_\iota - \bigotimes_{\iota=1}^{i} \mathbf{I}_{a_\iota}$$

$$\textit{und} \qquad \tilde{M}_\iota = \begin{cases} \mathbf{J}_{a_\iota}, & \iota \in \{\nu_1,\dots,\nu_j\} \\ \mathbf{I}_{a_\iota}, & \textit{sonst.} \end{cases}$$

Beweis: Folgt aus Lemma A.11 und der Beziehung

$$\prod_{\iota=1}^{i} a_\iota = \prod_{\iota=1}^{i}(a_\iota - 1) + \sum_{j=1}^{i-1}(-1)^{i-1-j} \sum_{\{\nu_1,\dots,\nu_j\}\subset\{1,\dots,i\}} \prod_{k=1}^{j} a_{\nu_k} + (-1)^{i-1} .$$

□

Nach der Zerlegung des Kronecker-Produktes von Projektor-Matrizen zeigen das folgende Lemma und das anschließende Korollar, dass für die asymptotische Verteilung der in Satz 5.4 betrachteten quadratischen Form nur der erste Term der in Lemma A.12 angegebenen Zerlegung von Bedeutung ist, da der Rest asymptotisch verschwindet.

Lemma A.13 *Seien die Bezeichnungen wie in Satz 5.4. Sei* $1 \leq j \leq i$, $F = \{\nu_1,\dots,\nu_j\}$ *und*

$$\left(\bigotimes_{\iota=1}^{i} \mathbf{P}_{a_\iota} \otimes \bigotimes_{\iota=i+1}^{m} \mathbf{J}_{a_\iota} \otimes \mathbb{1}_n'\right)\left(\boldsymbol{\phi}_\mu + \sum_{f=1}^{m} \sum_{\{i_1,\dots,i_f\}\subset\{1,\dots,m\}} \boldsymbol{\phi}_{A^{(i_1)},\dots,A^{(i_f)}}\right) = 0 .$$

Sei weiterhin

$$U = \frac{1}{f}\boldsymbol{\phi}_\varepsilon'\left(\tilde{M}_i(\nu_1,\dots,\nu_j) \otimes \bigotimes_{\iota=i+1}^{m} \frac{1}{a_\iota}\mathbf{J}_{a_\iota} \otimes \frac{1}{n}\mathbf{J}_n\right)\boldsymbol{\phi}_\varepsilon,$$

$$\textit{mit} \qquad f = \prod_{\iota=1}^{i}(a_\iota - 1)\prod_{k\in F} a_k,$$

$$\tilde{M}_i(\nu_1,\dots,\nu_j) = \bigotimes_{\iota=1}^{i} \tilde{M}_\iota - \bigotimes_{\iota=1}^{i} \mathbf{I}_{a_\iota} \qquad \textit{und} \qquad \tilde{M}_\iota = \begin{cases} \mathbf{J}_{a_\iota}, & \iota \in F \\ \mathbf{I}_{a_\iota}, & \textit{sonst} \end{cases} .$$

Dann gilt:

$$E(U) = 0, \qquad \mathrm{Var}(U) = \frac{2\sigma^4(\prod\limits_{k\in F} a_k - 1)\prod\limits_{k\in\{1,\dots,i\}-F} a_k}{(\prod\limits_{k\in F} a_k)(\prod\limits_{\iota=1}^{i}(a_\iota - 1)^2)}$$

Beweis: Folgt mit Hilfe von Lemma A.7. □

Korollar A.14 *Seien $U, \;F$ wie in Lemma A.13. Dann gilt:*

$$\operatorname{Var}\left(U\sqrt{\prod_{\iota=1}^{i} a_\iota}\right) \xrightarrow[\;]{\substack{\min\limits_{\iota\in F} a_\iota \to\infty \\ a_\iota \text{ fest für } (\iota\notin F) \\ n \text{ fest}}} 0\;.$$

Notation

Spezielle Matrizen und Vektoren

$\mathbf{I}_n = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$ (als $n \times n$-Matrix),

$\mathbb{1}_n = (\underbrace{1, \ldots, 1}_{n})'$,

$\mathbf{J}_n = \mathbb{1}_n \cdot \mathbb{1}_n'$,

$\mathbf{P}_n = \mathbf{I}_n - \frac{1}{n}\mathbf{J}_n$.

Abkürzende Konvergenzsymbole

$\dot{\sim}$ bedeute „ist asymptotisch verteilt nach“,

$\doteq$ bedeute „ist asymptotisch verteilungsäquivalent mit“.

Sonstige Symbole

$\lceil x \rceil$ sei die eindeutig bestimmte ganze Zahl n mit $n - 1 < x \leq n$,

$\otimes$ sei das Kroneckerprodukt (für Matrizen).

Literatur

AHRENS, H., J. LÄUTER (1974)
Mehrdimensionale Varianzanalyse. *Akademie-Verlag. Berlin.*

AKRITAS, M. G., S. F. ARNOLD (1994)
Fully Nonparametric Hypotheses for Factorial Designs I: Multivariate Repeated Measures Designs. *Journal of the American Statistical Association* **89**, 336–343.

AKRITAS, M. G., S. F. ARNOLD (2000)
Asymptotics for Analysis of Variance When the Number of Levels is Large. *Journal of the American Statistical Association* **95**, 212–226.

AKRITAS, M. G., S. F. ARNOLD, E. BRUNNER (1997)
Nonparametric Hypotheses and Rank Statistics for Unbalanced Factorial Designs. *Journal of the American Statistical Association* **92**, 258–265.

AKRITAS, M. G., E. BRUNNER (1997)
A unified approach to rank tests for mixed models. *Journal of Statistical Planning and Inference* **61**, 249–277.

ANDERSON, T. W. (1984)
An introduction to multivariate statistical analysis. *Wiley. New York.*

ANDREWS, D. (1971)
A note on the selection of data transformations. *Biometrika* **58**,**2**, 249–254.

ARNOLD, S. F. (1981)
The Theory of Linear Models and Multivariate Analysis. *Wiley. New York.*

ATKINSON, A. C., A. N. DONEV (1992)
Optimum Experimental Designs. *Oxford Science Publications.*

BAILEY, R. A. (1984)
Diskussion von T. Tjurs Paper „Analysis of Variance in Orthogonal Designs“. *International Statistical Review* **52**,**1**, 65–77.

BAKER, F. B. (1992)
Item response theory. *Dekker. New York.*

BANDEMER, H., A. BELLMANN, W. JUNG, K. RICHTER (1973)
Optimale Versuchsplanung. *Akademie-Verlag. Berlin.*

BARTLETT, M. S. (1936)
The Square Root Transformation in Analysis of Variance. *Journal of the Royal Statistical Society, Supplement* **3**, 68–78.

BARTLETT, M. S. (1947)
The use of transformations. *Biometrics* **3**, 39–52.

BATHKE, A. (2000)
ANOVA for a large number of treatments. Eingesendet zu *Mathematical Methods of Statistics*.

BENNINGTON, C. C., W. V. THAYNE (1994)
Use and Misuse of Mixed Model Analysis of Variance in Ecological Studies. *Ecology* **75**,**3**, 717–722.

BOOS, D. D., C. BROWNIE (1995)
ANOVA and rank tests when the number of treatments is large. *Statistics & Probability Letters* **23**, 183–191.

BOROVSKICH, YU. V. (1996)
U-Statistics in Banach spaces. *VSP. Utrecht.*

BOWERMAN, B. L., R. T. O'CONNELL (1990)
Linear Statistical Models: An Applied Approach. *Duxbury Press. Belmont.*

BOX, G. E. P., D. R. COX (1964)
An Analysis of Transformations. *Journal of the Royal Statistical Society. B* **26**, 211–252.

BOX, G. E. P., W. G. HUNTER, J. S. HUNTER (1978)
Statistics for experimenters. *Wiley. New York.*

BROWNIE, C., D. D. BOOS (1994)
Type I error robustness of ANOVA and ANOVA on ranks when the number of treatments is large. *Biometrics* **50**, 542-549.

BRUNNER, E., M. DENKER (1994)
Rank statistics under dependent observations and applications to factorial designs. *Journal of Statistical Planning and Inference* **42**, 353–378.

BRUNNER, E., M. L. PURI (1996)
Nonparametric Methods in Design and Analysis of Experiments. *Handbook of Statistics* **13**, 631–703. (Herausgeber S. Ghosh und C. R. Rao.)

CARROLL, R. J., D. RUPPERT (1984)
Power Transformations When Fitting Theoretical Models to Data. *Journal of the American Statistical Association.* **79**, 321–328.

CARROLL, R. J., D. RUPPERT (1988)
Transformation and Weighting in Regression. *Chapman and Hall. New York.*

CHOI, S. C. (1978)
Introductory Applied Statistics in Science. *Prentice Hall. Englewood Cliffs.*

CHRISTENSEN, R. A. (1987)
Plane answers to Complex Questions. The Theory of Linear Models. *Springer Verlag. New York.*

CORNFIELD, J. (1944)
On samples from finite populations. *Journal of the American Statistical Association* **39**, 236–239.

COX, D. R. (1958)
Planning of Experiments. *Wiley. New York.*

COX, D. R (1984)
Interaction. *International Statistical Review.* **52**,**1**, 1–31.

CURTISS, J. H. (1943)
On Transformations used in the Analysis of Variance. *Annals of Mathematical Statistics* **14**, 107–122.

DANIEL, C., F. S. WOOD (1971)
Fitting Equations to Data. *Wiley. New York.*

DENKER, M., N. NEUMANN (1979)
Verteilungskonvergenz in der nichtparametrischen Statistik. *Vorlesungsausarbeitung. Universität Göttingen.*

EISENHART, C. (1947)
The Assumptions Underlying the Analysis of Variance. *Biometrics* **3**,**1**, 1–21.

EFRON, B., C. STEIN (1981)
The Jackknife Estimate of Variance. *The Annals of Statistics* **9**,**3**, 586–596.

FAHRMEIR, L., R. KÜNSTLER, I. PIGEOT, G. TUTZ (1999)
Statistik. *Springer Verlag. Berlin.*

FISHER, R. A. (1934)
Diskussion von J. Wisharts Paper „Statistics in Agricultural Research". *Journal of the Royal Statistical Society, Supplement* **1**, 51–53.

GNANADESIKAN, R., J. N. SRIVASTAVA (1971)
Analysis and Design of Certain Quantitative Multiresponse Experiments. *Pergamon Press. Oxford.*

GUENTHER, W. C. (1964)
Analysis of Variance. *Prentice Hall. Englewood Cliffs.*

GUTTMANN, I. (1982)
Linear Models: An Introduction. *Wiley. New York.*

HAND, D. J., F. DALY, A. D. LUNN, K. J. MCCONWAY, E. OSTROWSKI (1994)
A Handbook of Small Data Sets. *Chapman & Hall. London.*

HELLAND, I. S. (1998)
A Population Approach to Analysis of Variance Models. *Scandinavian Journal of Statistics* **25**,**1**, 3–15.

HOEFFDING, W. (1948)
A class of statistics with asymptotically normal distributions. *Annals of Mathematical Statistics* **19**, 293–325.

HUITSON, A. (1971)
The Analysis of Variance. *Griffin. London.*

K. M. S. HUMAK (1984)
Statistische Methoden der Modellbildung III. *Akademie-Verlag. Berlin.*

KARLIN, S., Y. RINOTT (1982)
Applications of Anova Type Decompositions for Comparisons of Conditional Variance Statistics Including Jackknife Estimates. *The Annals of Statistics* **10**,**2**, 485–501.

KENDALL, M. G., A. STUART (1976)
The Advanced Theory of Statistics. Vol.3: Design and Analysis, and Time-Series. *Griffin. London.*

KIRK, R. E. (1968)
Experimental Design: Procedures for the Behavioral Sciences. *Brooks/Cole. Belmont.*

LANKOWSKI, D. (1997)
Nichtparametrische Tests für einen Random-Faktor. *Diplomarbeit. Universität Göttingen.*

LEE, A. J. (1990)
U-Statistics. Theory and practice. *Dekker. New York.*

LITTELL, R. C., G. A. MILLIKEN, W. W. STROUP, R. D. WOLFINGER (1996)
SAS System for Mixed Models. *SAS Institute Inc. Cary.*

MARKS, R. G. (1982)
Analyzing Research Data. *Lifetime Learning Publications. Belmont.*

MCLEAN, R. A., W. L. SANDERS, W. W. STROUP (1991)
A Unified Approach to Mixed Linear Models. *The American Statistician* **45**,**1**, 54–64.

NELDER, J. A. (1977)
A Reformulation of Linear Models. *Journal of the Royal Statistical Society. A* **140**,**1**, 48–76.

NELDER, J. A. (1994)
The statistics of linear models: back to basics. *Statistics and Computing* **4**, 221–234.

OELERICH, A. (1998)
Teststatistiken zur Analyse ordinaler Daten bei kleinen Stichprobenumfängen. *Diplomarbeit. Universität Göttingen.*

PORTNOY, S. (1984)
Asymptotic behavior of M-estimators of p regression parameters when p^2/n is large. I. Consistency. *The Annals of Statistics* **12, 4**, 1298-1309.

RANDLES, R. H., D. A. WOLFE (1979)
Introduction to the Theory of Nonparametric Statistics. *Wiley. New York.*

RENCHER, A. C. (1995)
Methods of Multivariate Analysis. *Wiley. New York.*

RETZLAFF, G., G. RUST, J. WAIBEL (1978)
Statistische Versuchsplanung. *Verlag Chemie. Weinheim.*

RIDER, P. H., H. L. HARTER, M. D. LUM (1956)
An Elementary Approach to the Analysis of Variance. *Wright Air Development Center Technical Report 56-20.*

ROTAR, V. I. (1973)
Some limit theorems for polynomials of second degree. *Theory of probability and its applications* **18, 3**, 499-507.

SACHS, L. (1968)
Statistische Auswertungsmethoden. *Springer Verlag. Berlin.*

SCHACH, S., T. SCHÄFER (1978)
Regressions- und Varianzanalyse. *Springer Verlag. Berlin.*

SCHEFFÉ, H. (1956)
Alternative Models for the Analysis of Variance. *The Annals of Mathematical Statistics* **27**, 251–271.

SCHEFFÉ, H. (1959)
The Analysis of Variance. *Wiley. New York.*

SCHWARZ, J. (1999)
Vergleich von Verfahren zur Analyse dichotomer Daten. *Diplomarbeit. Universität Göttingen.*

SEARLE, S. R. (1971)
Linear Models. *Wiley. New York.*

SEARLE, S. R. (1997)
Built-In Restrictions on Best Linear Unbiased Predictors (BLUP) of Random Effects in Mixed Models. *The American Statistician* **51**,**1**, 19–21.

SEBER, G. A. F. (1977)
Linear Regression Analysis. *Wiley. New York.*

SEBER, G. A. F., C. J. WILD (1988)
Nonlinear Regression. *Wiley. New York.*

SINGER, J. (1991)
Types of Factors and Their Structural Layouts. *In: Fundamentals of Exploratory Analysis of Variance. Editors: Hoaglin, D.C., F. Mosteller, J.W. Tukey. Wiley. New York.*

SNEDECOR, G. W., W. G. COCHRAN (1967)
Statistical Methods. 6th edition. *Iowa State University Press. Ames.*

TAKEMURA, A. (1983)
Tensor Analysis of ANOVA Decomposition. *Journal of the American Statistical Association.* **78**, 894–900.

TIPPETT, L. H. C. (1934)
Statistical Methods in Textile Research, Part 2, Uses of the binomial and Poisson distributions. *Shirley Institute Memoires* **13**, 35–72.

TJUR, T. (1984)
Analysis of Variance Models in Orthogonal Designs. *International Statistical Review.* **52**,**1**, 33–81.

TJUR, T. (1991)
Analysis of Variance and Design of Experiments. *Scandinavian Journal of Statistics.* **18**,**4**, 273–308.

WHITE, R. F. (1975)
Randomization and the Analysis of Variance. *Biometrics* **31**,**2**, 555–571.

WILK, M. B., O. KEMPTHORNE (1955)
Fixed, Mixed, and Random Models. *Journal of the American Statistical Association.* **50**, 1144–1167.

WINER, B. J. (1971)
Statistical Principles in Experimental Design. 2nd edition. *McGraw-Hill Kogakusha. Tokio.*

WRIGHT, B. D., MEAD, R. J. and R. DRABA (1976)
Detecting and correcting test item bias with a logistic response model. *Research Memorandum No. 22. MESA Psychometric Laboratory, University of Chicago.*

Lebenslauf

9.6.1972	Geboren in Hamburg-Harburg
1978-1982	Christian-Flemes-Grundschule in Völksen
1982-1984	Orientierungsstufe Nord in Springe
1984-1991	Otto-Hahn-Gymnasium in Springe
1991-1992	Zivildienst in der Sozialstation des Deutschen Roten Kreuzes in Springe
1992-1998	Mathematikstudium (mit Nebenfach Betriebswirtschaftslehre) an der Georg-August-Universität in Göttingen
14.10.1994	Vordiplom
15.5.1998	Diplom
seit 1998	Promotionsstudium bei Prof. Dr. M. Denker am Institut für Mathematische Stochastik
6-10/1998	Wissenschaftlicher Angestellter an der Abteilung Medizinische Statistik, Georg-August-Universität Göttingen
seit 8/1998	Wissenschaftlicher Angestellter am Institut für Mathematische Stochastik, Georg-August-Universität Göttingen
8-10/1999	Aufenthalt an der University of Illinois at Urbana-Champaign, USA