Nichtparametrische relative Effekte

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

vorgelegt von
Sebastian Domhof
aus Schwerte / Ruhr

Göttingen 2001

D 7
Referent: Prof. Dr. Edgar Brunner
Korreferent: Prof. Dr. Manfred Denker
Tag der mündlichen Prüfung: 2. Mai 2001

Danksagung

Mein besonderer Dank gilt Herrn Prof. Dr. Edgar Brunner, der mich bei der Anfertigung dieser Arbeit mit wertvollen Ratschlägen und Hinweisen begleitete. Außerdem danke ich ihm für die Bereitstellung der Arbeitsmittel in der Abteilung Medizinische Statistik und die Ermöglichung meines Forschungsaufenthaltes an der Indiana University, Bloomington, der entscheidend zur Entstehung dieser Arbeit beigetragen hat.

Für die Erstellung der Gutachten danke ich Herrn Prof. Dr. Edgar Brunner und Herrn Prof. Dr. Manfred Denker.

Für das sorgfältige Lesen der Arbeit bedanke ich mich bei Herrn Alexander Siemer, Frau Bettina Kulle und ganz besonders bei Frau Nicole Grützner.

Inhaltsverzeichnis

Kapitel 1

Einleitung

1.1 Motivation

Die parametrische Statistik nimmt an, dass die Daten aus einem Versuch oder einer Studie einem wahrscheinlichkeitstheoretischen Modell aus einer Klasse von Modellen genügen, die durch endlich viele reelle Zahlen beschrieben werden können. Diese endlich vielen Zahlen werden Parameter genannt und besitzen häufig anschauliche und leicht verständliche Interpretationen. Als klassisches Beispiel seien hier die Parameter μ und σ^2 der Normalverteilung genannt, die zur Beschreibung der Lage und Streuung dieser Verteilung dienen. Aus den erhobenen Daten werden mit formalen Methoden Rückschlüsse auf die Parameter des Modells gezogen. Zu diesen Methoden zählen Punktschätzer und Konfidenzbereiche sowie Tests von Hypothesen. Eine grundlegende Schwierigkeit bei der Verwendung von Verfahren aus der parametrischen Statistik ist die Wahl der parametrischen Modellklasse. Bei der Wahl einer ungeeigneten Klasse verlieren im Allgemeinen die Parameter ihre Bedeutung und die schließenden Methoden ihre Gültigkeit. Deshalb erfordert die parametrische Statistik viel Sachkenntnis über den Versuch und einen breiten Überblick über möglicherweise geeignete parametrische Modellklassen und die zugehörigen Auswertungsmethoden. Aber auch wenn diese Sachkenntnis und der entsprechende Überblick zur Verfügung stehen, ist es möglich, dass keine der bekannten parametrischen Modellklassen zur Beschreibung der Verteilungen geeignet ist. In diesem Fall bietet die nichtparametrische Statistik eine Alternative.

Die nichtparametrische Statistik geht nicht davon aus, dass sich die Verteilungen der Daten durch endlich viele Parameter beschreiben lassen. Sie lässt eine größere Klasse von Verteilungen zu, zum Beispiel die Klasse aller stetigen Verteilungen. Die Entscheidung, ob eine solche Klasse geeignet ist, kann meist auch ohne tiefergehende Sachkenntnis über den Versuch getroffen werden. Das Gegenstück zu den Parametern aus der parametrischen Statistik bilden in der nichtparametrischen Statistik die sogenannten statistischen Funktionale. Diese Funktionale sind Abbildungen von den Verteilungsklassen in endlichdimensionale Räume. Ihre Aufgabe ist es, die umfangreiche Information, die in den Verteilungen enthalten ist, auf wesentliche Aspekte zu reduzieren. Sie sollten – wie die Parameter – eine anschauliche Interpretation haben, zum Beispiel als Lage- oder Streumaß. Bekannte Beispiele für Lagemaße sind der Erwartungswert $\mathrm{E} : F \mapsto \int x \, dF(x)$ und der Median $\mathrm{Med} : F \mapsto F^{-1}(1/2)$. Wie in der parametrischen Statistik können als formale Methoden, um aus Beobachtungen Rückschlüsse auf die Funktionale zu ziehen, Punktschätzer, Konfidenzbereiche und Tests konstruiert werden.

Die nichtparametrische Statistik geht jedoch nicht immer den Weg über Funktionale, sondern befasst sich in einigen Fällen auch direkt mit den Verteilungen. Insbesondere existiert eine Vielzahl von klassischen Tests, die als Hypothesen die Gleichheit von Verteilungen stellen. Bekannte Tests für die Hypothese, dass die Verteilungsfunktion F einer Stichprobe mit einer vorgegebenen Verteilungsfunktion F_0 übereinstimmt, sind der χ^2-Anpassungstest von Pearson (1900) sowie die Ein-Stichproben-Tests von Cramér und von Mises (Cramér, 1928; von Mises, 1931) und von Kolmogorov und Smirnov (Kolmogorov, 1933; Smirnov, 1939b). Um die Hypothese der Gleichheit der Verteilungsfunktionen F_1 und F_2 von zwei unverbundenen Stichproben zu testen, können die Zwei-Stichproben-Versionen des Cramér-von Mises-Tests (Lehmann, 1951) und des Kolmogorov-Smirnov-Tests (Smirnov, 1939a), aber auch der Run-Test von Wald und Wolfowitz (1940) und der Wilcoxon-Mann-Whitney-Test (Wilcoxon, 1945; Mann und Whitney, 1947) verwendet werden. Verallgemeinerungen auf das Mehr-Stichproben-Problem mit der Hypothese $F_1 = \cdots = F_a$ wurden für den Cramér-von Mises- sowie für den Kolmogorov-Smirnov-Test von Kiefer (1959) vorgeschlagen. Für den Wilcoxon-Mann-Whitney-Test wurde eine entsprechende Verallgemeinerung von Kruskal und Wallis (Kruskal, 1952; Kruskal und Wallis, 1952, 1953) angegeben.

Die Verwendung statistischer Tests wurde jedoch seit der Kritik von Berkson (Berkson, 1939; Berkson und Geary, 1941, Teil I; Berkson, 1942, 1943) immer wieder in Frage gestellt (Rothman und Greenland, 1998, Kapitel 12). Diese Kritik beinhaltet aus Sicht des Autors im Wesentlichen folgende Punkte:

- Die Beschränkung der Fragestellung auf die Entscheidung zwischen Hypothese und Alternative entspricht in den wenigsten Fällen der Fragestellung, die einer Studie zugrunde liegt. Denn die weitaus meisten Fragestellungen, die anhand von Studien untersucht werden sollen, sind quantitativer und nicht qualitativer Art. Eine Entscheidung allein zwischen Hypothese und Alternative ist nicht spezifisch genug. Auch die zusätzliche Angabe eines p-Wertes hilft nicht bei der Beantwortung der quantitativen Fragestellung.

  Besonders fragwürdig ist die Verwendung von Punkthypothesen, wie zum Beispiel bei den oben genannten Tests auf Gleichheit von Verteilungen: Eine Punkthypothese kann bereits ohne die Erhebung von Daten ausgeschlossen werden, da in der Praxis Parameter, Funktionale oder Verteilungen einer Gleichung niemals exakt genügen.

- Selbst wenn die Entscheidung zwischen Hypothese und Alternative bei der Beantwortung der Fragestellung einer Studie hilfreich sein sollte, ist die Entscheidung, die mittels eines statistischen Tests gefällt wird, in den seltensten Fällen sinnvoll. Denn bei einem Test ist lediglich die Wahrscheinlichkeit für den Fehler erster Art nach oben beschränkt. Diese Schranke liegt zudem meist bei historisch begründeten 5% (Freedman et al., 1998, Kapitel 29, Abschnitt 1), hat aber keine Begründung in der Fragestellung des Versuchs.

Falls also die Hypothese durch einen statistischen Test nicht verworfen wird, hat man durch den Test keine Information bezüglich der Entscheidung zwischen Hypothese und Alternative gewonnen. Dieses unbefriedigende Ergebnis ist nur schwierig zu vermitteln und führt deshalb immer wieder zu dem Missverständnis, die Nichtablehnung der Hypothese sei ein Hinweis auf ihre Gültigkeit.

- Die meisten Fragen der Wissenschaft werden durch mehr als eine Studie bearbeitet. Zur Zusammenführung der Ergebnisse von mehreren Studien sind die jeweiligen Testergebnisse aber selbst dann ungeeignet, wenn die Hypothesen und Alternativen der verschiedenen Studien übereinstimmen: Sie enthalten nicht mehr genügend Information, um sie sinnvoll zu kombinieren (Rothman und Greenland, 1998, Kapitel 32).

Die wiedergegebene Kritik an statistischen Tests spricht also im Allgemeinen für die Verwendung von Punktschätzern und Konfidenzbereichen (Rothman und Greenland, 1998, Kapitel 12) und speziell in der nichtparametrischen Statistik somit für die Verwendung von Funktionalen (siehe auch Savage, 1957).

Die meisten der oben genannten statistischen Tests beruhen auf abstrakten Abstandsmaßen zwischen Verteilungsfunktionen, die keine anschauliche Interpretation zulassen. Der Kolmogorov-Smirnov-Test zum Beispiel verwendet eine Teststatistik, die mit Hilfe des Kolmogorov-Abstandes für Verteilungsfuntionen

$$d_K(F_1, F_2) = \sup\{|F_1(x) - F_2(x)| : x \in \mathbb{R}\}$$

definiert wird.

Lohnenswert erscheint dagegen die Auseinandersetzung mit dem Wilcoxon-Mann-Whitney-Test: Obwohl auch bei diesem Test die Hypothesen direkt in den Verteilungsfunktionen gestellt werden, kann die zugehörige Teststatistik nach geeigneter Transformation als Schätzer für ein Funktional mit anschaulicher Interpretation aufgefasst werden. Dies wurde erst über ein Jahrzehnt nach der Einführung des Tests durch Wilcoxon (1945) von Birnbaum (1956) bemerkt. In Formelschreibweise lautet das Funktional, das wir in dieser Einleitung als Wilcoxon-Mann-Whitney-Effekt bezeichnen,

$$(F_1, F_2) \mapsto \int F_1 \, dF_2 = \mathrm{P}(X_1 < X_2) + \frac{1}{2}\,\mathrm{P}(X_1 = X_2),$$

wobei $X_1 \sim F_1$ und $X_2 \sim F_2$ unabhängige Beobachtungen aus den beiden Versuchspopulationen sind. Der Wilcoxon-Mann-Whitney-Effekt ist ein Maß dafür, wie stark Beobachtungen mit der Verteilungsfunktion F_1 zu größeren Werten neigen als Beobachtungen mit der Verteilungsfunktion F_2. Eine wichtige Eigenschaft dieses Funktionals ist, dass es ausschließlich auf der Ordnungsstruktur der möglichen Beobachtungen beruht. Es kann also immer verwendet werden, wenn Daten mit mindestens ordinalem Skalenniveau vorliegen. Inzwischen existiert eine große Anzahl an Veröffentlichungen, die sich mit Punktschätzern, Konfidenzbereichen oder auch Tests für den Wilcoxon-Mann-Whitney-Effekt beschäftigen.

Interessant ist der Wilcoxon-Mann-Whitney-Test, wenn man einmal von der obigen grundsätzlichen Kritik absieht, auch noch aus einem anderen Grund: Er besitzt nicht nur die klassische Verallgemeinerung auf das Mehr-Stichproben-Problem von Kruskal und Wallis, sondern hat während des letzten Jahrzehnts weitere natürliche Verallgemeinerungen auf immer umfassendere Klassen von faktoriellen Versuchsplänen erfahren (Akritas und Arnold, 1994; Akritas et al., 1997; Akritas und Brunner, 1997; Brunner et al., 1999). Wie der Wilcoxon-Mann-Whitney- und der Kruskal-Wallis-Test stellen auch diese Verallgemeinerungen ihre Hypothesen direkt in den Verteilungsfunktionen. Sowohl für den Kruskal-Wallis-Test als auch für die Verallgemeinerungen besteht eine Beziehung zwischen den Teststatistiken und bestimmten Funktionalen, die eine ähnliche Interpretation wie der Wilcoxon-Mann-Whitney-Effekt haben. Im Gegensatz zu diesem haben sie aber den schwerwiegenden Nachteil, von den Stichprobenumfängen abzuhängen. Diese sind jedoch für die Fragestellung einer Studie irrelevant.

Die vorliegende Arbeit beschäftigt sich mit von den Stichprobenumfängen unabhängigen Verallgemeinerungen des Wilcoxon-Mann-Whitney-Effekts für die sehr allgemeine Klasse von faktoriellen Versuchsplänen, die von Brunner et al. (1999) betrachtet wurden. Diese Funktionale werden hier als relative Effekte bezeichnet. Sie sind insbesondere unter dem für die nichtparametrische Statistik wichtigen Gesichtspunkt der möglichst allgemeinen Verwendbarkeit interessant. Denn zum einen sind sie wie der Wilcoxon-Mann-Whitney-Effekt für alle Daten mit mindestens ordinalem Skalenniveau geeignet. Zum anderen können sie für eine sehr große Klasse von faktoriellen Versuchsplänen definiert werden. Ähnliche, ebenfalls von den Stichprobenumfängen unabhängige Funktionale wurden für eine weniger allgemeine Klasse von faktoriellen Versuchsplänen bereits von Kulle (1999) und Siemer (1999) behandelt. Jedoch stellen die dort verwendeten Funktionale keine Verallgemeinerung des Wilcoxon-Mann-Whitney-Effekts dar, da sie sich im Zwei-Stichproben-Fall nicht auf diesen reduzieren. Zudem lag das Gewicht der beiden genannten Arbeiten auf der Konstruktion von Tests für Punkthypothesen, während sich diese Arbeit aufgrund der bereits angeführten Kritik ausschließlich mit den relativen Effekten selbst sowie Punkschätzern und Konfidenzintervallen dafür beschäftigt.

Es bleibt noch zu erwähnen, dass sich die Verwendung jedes der zuletzt genannten Funktionale, also der Funktionale, die in Beziehung zu den Teststatistiken der Verfahren von Akritas und Arnold (1994), Akritas et al. (1997), Akritas und Brunner (1997) und Brunner et al. (1999) stehen und der von Kulle (1999) und Siemer (1999) betrachteten Funktionale sowie der hier behandelten relativen Effekte, in das von Bross (1958) vorgeschlagene Prinzip der Ridit-Analyse einordnen lässt. Allerdings hat Bross in seiner Arbeit in erster Linie die hinter dem Konzept der Ridit-Analyse liegenden Überlegungen geschildert und nur für einen Spezialfall Punktschätzer und Konfidenzintervalle angegeben.

1.2 Aufbau der Arbeit

Im auf diese Einleitung folgenden, zweiten Kapitel werden die Versuchspläne und wahrscheinlichkeitstheoretischen Modelle vorgestellt, mit denen sich diese Arbeit beschäftigt. Als Beispiele und Beleg für die Verwendbarkeit dieser Pläne und Modelle in der Biometrie dienen vier medizinische Studien, auf die im Folgenden immer wieder Bezug genommen wird.

Im dritten Kapitel werden die eigentlichen Gegenstände dieser Arbeit, die relativen Effekte, eingeführt und ausführlich erläutert. Die Kapitel 4 und 5 befassen sich zunächst mit der Konstruktion von Punkt- und Intervallschätzern für die relativen Effekte. Darüber hinaus enthalten sie eine eingehende Untersuchung der finiten und asymptotischen Eigenschaften der konstruierten Schätzer.

Das abschließende Kapitel 6 trägt einen etwas spezielleren Charakter als die anderen Kapitel: Für eine im Vergleich zum Rest der Arbeit deutlich eingeschränkte Klasse von Versuchsplänen wird untersucht, inwiefern zusätzliche Information in Form von Kovariablen oder Ausgangswerten verwendet werden kann, um die Genauigkeit der Punkt- und Intervallschätzer zu erhöhen.

Da die vorliegende Dissertation in erster Linie eine biometrische Arbeit ist, treten die mathematischen Beweise für die Eigenschaften der diskutierten Methoden in den Hintergrund. Sie befinden sich deshalb in den meisten Fällen im separaten Anhang A. Um den Lesefluss nicht zu stören, wurden darüber hinaus die Anleitungen zu den SAS-Makros `RE_CI` und `SIMPLE_RE_CI`, mit denen die vorgestellten statistischen Verfahren praktisch umgesetzt werden können, sowie die Tabellen mit Simulationsergebnissen in den Anhängen B und C platziert.

Im Anschluss an den Anhang findet der Leser noch eine kurze Zusammenstellung der verwendeten Notation und das Literaturverzeichnis.

Auch wenn die Mathematik in dieser Arbeit nicht um ihrer selbst willen betrieben wird, so ist sie doch als Sprache zur präzisen Formulierung der entwickelten statistischen Methoden und der dazugehörigen Eigenschaften unentbehrlich. Der Leser sollte deshalb mit den grundlegenden Begriffen und Resultaten aus der Maß- und Wahrscheinlichkeitstheorie sowie der mathematischen Statistik vertraut sein, wie man sie zum Beispiel in den Lehrbüchern von Billingsley (1995) und Shao (1999) findet. Dies wird insbesondere für das Verständnis von Anhang A hilfreich sein, in dem grundlegende mathematische Regeln und Sätze zumeist stillschweigend benutzt werden.

Kapitel 2

Versuchspläne und Modelle

Im ersten Abschnitt dieses Kapitels beschreiben wir die Klasse von Versuchsplänen, die sogenannten faktoriellen Versuchspläne, für die in den folgenden Kapiteln Auswertungsverfahren entwickelt werden. Um Verfahren aus der schließenden Statistik zu ermöglichen, geben wir im zweiten Abschnitt entsprechende wahrscheinlichkeitstheoretische Modelle an. Sowohl die Klasse von Versuchsplänen als auch die Modelle sind der Arbeit von Brunner et al. (1999) entnommen.

2.1 Faktorielle Versuchspläne

Wir verwenden in dieser Arbeit Versuchpläne, bei denen die Versuchseinheiten verschiedenen Gruppen angehören. Die Anzahl der Gruppen bezeichnen wir mit a und die Anzahl der Versuchseinheiten in der i-ten Gruppe mit n_i. Die Beobachtungen an den einzelnen Versuchseinheiten wiederum kommen unter t verschiedenen Bedingungen zustande, wobei es möglich ist, dass bei den verschiedenen Versuchseinheiten unter den verschiedenen Bedingungen unterschiedlich viele Beobachtungen vorliegen. In der Praxis werden die Bedingungen häufig die Zeitpunkte sein, zu denen die Beobachtungen gemacht werden. In diesem Fall spricht man von longitudinalen Daten.

Alle Beobachtungen an einer Versuchseinheit unter einer Bedingung fassen wir jeweils zu Vektoren

$$\boldsymbol{X}_{iks} = (X_{iks1}, \ldots, X_{iksm_{iks}})'$$

zusammen, wobei der Index $i \in \{1, \ldots, a\}$ die i-te Gruppe , der Index $k \in \{1, \ldots, n_i\}$ die k-te Versuchseinheit in der i-ten Gruppe und der Index $s \in \{1, \ldots, t\}$ die s-te Bedingung bezeichnet. Als Werte für die Dimensionen m_{iks} der Vektoren sind alle natürlichen Zahlen, aber auch die Null möglich, falls Beobachtungen fehlen. Als weitere Schreibweisen vereinbaren wir

$$\boldsymbol{X}_{ik} = (\boldsymbol{X}'_{ik1}, \ldots, \boldsymbol{X}'_{ikt})', \qquad \boldsymbol{m}_{ik} = (m_{ik1}, \ldots, m_{ikt})', \qquad n = \sum_{i=1}^{a} n_i,$$

$$\lambda_{iks} = \begin{cases} 1 & : m_{iks} > 0, \\ 0 & : m_{iks} = 0, \end{cases} \qquad \lambda_{i\cdot s} = \sum_{k=1}^{n_i} \lambda_{iks}.$$

Somit fasst der Vektor $\boldsymbol{X}_{ik}$ alle Beobachtungen an der k-ten Versuchseinheit in der i-ten Gruppe zusammen, $\boldsymbol{m}_{ik}$ ist der Vektor der Dimensionen der Teilvektoren $\boldsymbol{X}_{iks}$ von $\boldsymbol{X}_{ik}$, und n bezeichnet die Gesamtanzahl der Versuchseinheiten. Die Variable λ_{iks} zeigt an, ob an der k-ten Versuchseinheit in der i-ten Gruppe unter der s-ten Bedingung überhaupt Beobachtungen vorliegen und $\lambda_{i \cdot s}$ bezeichnet die Anzahl der Versucheinheiten in der i-ten Gruppe, für die mindestens eine Beobachtung unter der s-ten Bedingung vorhanden ist.

Bei allen Beobachtungen $X_{iks\ell}$ muss es sich um Beobachtungen der gleichen Größe handeln, und die Werte dieser Größe müssen sich als reelle Zahl darstellen lassen. Damit sind auch geordnet kategorielle Daten eingeschlossen, wenn die verschiedenen Kategorien als reelle Zahlen aufgefasst werden.

Die verschiedenen Gruppen von Versuchseinheiten können als Stufen eines Faktors verstanden werden, dessen Einfluss auf die Beobachtungen untersucht werden soll. Dasselbe gilt für die verschiedenen Bedingungen, wobei diese allerdings unter den Versuchseinheiten verschachtelt sind. Tatsächlich umfasst die betrachtete Klasse von Versuchsplänen sogar faktorielle Versuchspläne mit einer beliebigen Anzahl von Faktoren in beliebiger Anordnung, wenn den Gruppen und Bedingungen eine entsprechende Struktur unterlegt wird. Um solchen Versuchsplänen gerecht zu werden, müsste allerdings eine aufwendigere Schreibweise mit weiteren Indizes vereinbart werden, worauf hier aus Sorge um die Lesbarkeit verzichtet wird.

2.1.1 Beispiele

In diesem Unterabschnitt werden exemplarisch vier Versuche vorgestellt. Dabei geht es zum einen darum, die Verwendung der oben eingeführten Schreibweise anhand von konkreten Beispielen vorzuführen. Zum anderen soll aber auch angedeutet werden, dass die vorgestellte Klasse von Versuchsplänen viele in der Praxis vorkommende Pläne umfasst. In den folgenden Abschnitten und Kapiteln werden wir immer wieder auf die hier beschriebenen Versuche zurückkommen.

Panik-Skala-Studie I

Der hier mit Panik-Skala-Studie I bezeichnete Versuch ist Teil einer von Bandelow et al. (1998) durchgeführten Studie. Im Rahmen dieser Studie wurden 16 Patienten mit Panikstörung und Agoraphobie über einen Zeitraum von acht Wochen mit dem Antidepressivum Imipramin behandelt. Zu Beginn der Behandlung und danach im Abstand von je zwei Wochen wurde die *Clinical Global Impression (CGI)* auf einer Skala mit dem Wertebereich $\{2, 3, \ldots, 8\}$ gemessen. Dabei entspricht der Wert 2 einem gesunden und der Wert 8 einem extrem schwer kranken Patienten. Die Daten des Versuchs sind in Tabelle 2.1 dargestellt.

In der Panik-Skala-Studie I sind die Versuchseinheiten, also die Patienten, nicht in mehrere Gruppen aufgeteilt ($a = 1$). Die verschiedenen Bedingungen, unter denen die Beobachtungen gemacht werden, sind die $t = 5$ Zeitpunkte. An jedem der $n = 16$ Patienten wird zu jedem Zeitpunkt genau eine Beobachtung gemacht: Für alle Paare (k, s) gilt also $m_{1ks} = 1$ und somit $\boldsymbol{m}_{1k} = (1, 1, 1, 1, 1)'$. Die CGI-Scores

Tabelle 2.1: CGI-Scores der Patienten aus der Panik-Skala-Studie I.

	Woche						Woche				
Patient	0	2	4	6	8	Patient	0	2	4	6	8
1	8	6	5	5	4	9	5	4	3	3	2
2	8	6	5	4	2	10	8	6	5	5	4
3	6	5	5	4	2	11	7	6	5	4	2
4	6	6	6	5	5	12	6	5	5	4	2
5	7	6	6	6	6	13	6	6	6	5	5
6	8	7	3	2	2	14	8	6	6	6	6
7	7	6	7	3	3	15	8	7	4	2	2
8	6	4	5	3	3	16	7	6	7	3	3

eines Patienten entsprechen hier den Vektoren

$$\boldsymbol{X}_{1k} = (\boldsymbol{X}'_{1k1}, \ldots, \boldsymbol{X}'_{1k5})' = (X_{1k11}, \ldots, X_{1k51})'.$$

Der Index i, der für die Gruppen der Versuchseinheiten steht, hat hier immer den Wert 1, so dass man ihn weglassen würde, wenn man nur diese eine Studie beschreiben wollte. Er wurde hier lediglich beibehalten, um für alle Beispiele eine einheitliche Notation zu verwenden.

Panik-Skala-Studie II

Die Daten der Panik-Skala-Studie II entstammen wie die des vorangehenden Unterabschnitts der Arbeit von Bandelow et al. (1998). Diese Studie umfasst 37 Patienten mit Panikstörung, davon 24 mit und 13 ohne Agoraphobie. Wie in der Panik-Skala-Studie I wurden die Patienten acht Wochen lang mit dem Antidepressivum Imipramin behandelt. Jedoch wurden keine CGI-Scores, sondern die Schwere der Panikstörung auf der P&A-Skala (Bandelow, 1997) gemessen, die die ganzen Zahlen von 0 bis 52 umfasst. Dabei bedeuten kleine Werte eine schwach ausgeprägte und große Werte eine stark ausgeprägte Panikstörung. Ein weiterer Unterschied zur Panik-Skala-Studie I liegt darin, dass hier bei einigen Patienten Werte fehlen, zum Teil auf Grund von Therapieabbrüchen. Die Versuchsdaten sind in Tabelle 2.2 zusammengestellt. Die fehlenden Werte sind in dieser Tabelle durch einen Punkt dargestellt.

Die $n = 37$ Versuchseinheiten der Panik-Skala-Studie II sind in $a = 2$ Gruppen eingeteilt: In die Gruppe von $n_1 = 24$ Patienten mit Agoraphobie und die Gruppe von $n_2 = 13$ Patienten ohne Agoraphobie. Die Beobachtungen werden unter $t = 5$ Bedingungen gemacht, die den Zeitpunkten entsprechen. Jedoch liegen nicht für jeden Patienten zu jedem Zeitpunkt Beobachtungen vor: Für einige Tripel (i, k, s) gilt $m_{iks} = 1$, für andere aber $m_{iks} = 0$. Für den sechsten Patienten aus der ersten Gruppe gilt zum Beispiel

$$\boldsymbol{m}_{16} = (m_{161}, m_{162}, m_{163}, m_{164}, m_{165})' = (1, 1, 0, 0, 0)'.$$

Tabelle 2.2: P&A-Scores der Patienten aus der Panik-Skala-Studie II.

mit Agoraphobie						ohne Agoraphobie					
	Woche						Woche				
Patient	0	2	4	6	8	Patient	0	2	4	6	8
1	42	35	22	16	9	1	21	14	11	6	0
2	36	29	21	13	1	2	29	29	25	21	19
3	43	38	38	31	33	3	16	10	7	.	6
4	37	16	11	6	5	4	21	14	11	6	0
5	24	27	8	14	6	5	29	29	25	21	19
6	27	30	.	.	.	6	20	0	2	.	.
7	31	5	10	2	0	7	15	10	6	2	15
8	22	17	6	1	1	8	31	20	13	16	16
9	28	29	.	.	.	9	24	6	6	4	9
10	39	36	22	16	9	10	25	9	11	14	13
11	32	29	21	13	1	11	19	38	30	.	.
12	43	38	38	33	33	12	14	8	9	0	0
13	37	16	11	6	5	13	26	14	9	0	.
14	24	27	8	14	6						
15	34	26	25	35	24						
16	35	30	17	19	.						
17	30	48	22	22	26						
18	32	18	28	36	37						
19	26	5	6	2	2						
20	27	37	22	34	19						
21	21	23	16	.	.						
22	37	27	.	11	20						
23	26	26	14	31	31						
24	45	42	18	2	0						

Dementsprechend ist

$$\boldsymbol{X}_{16} = (\boldsymbol{X}'_{161}, \boldsymbol{X}'_{162}, \boldsymbol{X}'_{163}, \boldsymbol{X}'_{164}, \boldsymbol{X}'_{165})' = (27, 30, \cdot, \cdot, \cdot)'$$

ein Vektor mit drei leeren Komponenten.

Schulterschmerz-Studie

Die Daten der Schulterschmerz-Studie sind der Arbeit von Lumley (1996) entnommen. Sie umfasst 41 Patienten an denen eine laparoskopische Operation durchgeführt wurde. Bei 22 zufällig ausgewählten Patienten wurde nach der Operation die Luft nach einem speziellen Verfahren abgesaugt (Behandlung Y), die restlichen 19 Patienten dienten als Kontrolle (Behandlung N). Zu sechs Zeitpunkten nach der Operation haben die Patienten den als Nebenwirkung dieser Operation häufig auftretenden Schmerz in der Schulterspitze auf einer Skala von 1 (niedrig) bis 5 (sehr stark) bewertet. Die Versuchsdaten sind in Tabelle 2.3 wiedergegeben.

Tabelle 2.3: Schmerz-Scores der Patienten aus der Schulterschmerz-Studie.

Behandlung Y							Behandlung N						
	Zeitpunkt							Zeitpunkt					
Patient	1	2	3	4	5	6	Patient	1	2	3	4	5	6
1	1	1	1	1	1	1	1	5	2	3	5	5	4
2	3	2	1	1	1	1	2	1	5	3	4	5	3
3	3	2	2	2	1	1	3	4	4	4	4	1	1
4	1	1	1	1	1	1	4	4	4	4	4	4	3
5	1	1	1	1	1	1	5	2	3	4	3	3	2
6	1	2	1	1	1	1	6	3	4	3	3	3	2
7	1	3	2	1	1	1	7	3	3	4	4	4	3
8	2	2	1	1	1	1	8	1	1	1	1	1	1
9	1	1	1	1	1	1	9	1	1	1	1	1	1
10	3	1	1	1	1	1	10	1	5	5	5	4	3
11	1	1	1	1	1	1	11	1	3	2	2	1	1
12	2	1	1	1	1	2	12	2	2	3	4	2	2
13	1	2	2	2	2	2	13	2	2	1	3	3	2
14	3	1	1	1	3	3	14	1	1	1	1	1	1
15	2	1	1	1	1	1	15	1	1	1	1	1	1
16	1	1	1	1	1	1	16	5	5	5	4	3	3
17	1	1	1	1	1	1	17	3	3	3	3	1	1
18	2	1	1	1	1	1	18	5	4	4	4	2	2
19	4	4	2	4	2	2	19	1	3	3	3	2	1
20	4	4	4	2	1	1							
21	1	1	1	2	1	1							
22	1	1	1	2	1	2							

Die $n = 41$ Teilnehmer dieser Studie sind in $a = 2$ Gruppen mit $n_1 = 22$ und $n_2 = 19$ Patienten eingeteilt. Es liegen pro Person $t = 6$ Messwerte zu verschiedenen Zeitpunkten vor, so dass alle m_{iks} den Wert 1 haben. Alle Beobachtungen an einer Versuchseinheit fassen wir in Gestalt von Vektoren zusammen:

$$\boldsymbol{X}_{ik} = (\boldsymbol{X}'_{ik1}, \ldots, \boldsymbol{X}'_{ik6})' = (X_{ik11}, \ldots, X_{ik61})'.$$

Wasser-Irrgarten-Test

In einem Tierversuch zur Prüfung der Teratogenität eines Stoffes wurden 144 Wistar-Ratten verwendet, wobei davon jeweils zwei zu einem von insgesamt 72 Muttertieren gehören. Die Muttertiere erhielten die Substanz während der Tragezeit in drei verschieden hohen Dosen oder ein Placebo. Aus den 72 Muttertieren wurden für die Placebogruppe und die Gruppe mit der geringsten Dosis jeweils 17 und für die beiden Gruppen mit der mittleren und der höchsten Dosierung jeweils 19 Tiere zufällig ausgewählt. Aus den Kindern der Muttertiere wurden wiederum zufällig je zwei Jungtiere herausgegriffen. Am ersten Versuchstag sollte jedes der 144 Versuchstiere sechs Mal im Abstand von jeweils einer Stunde den Ausgang aus einem Wasser-Irrgarten finden. Den Tieren standen dabei jeweils 150 Sekunden zur Verfügung.

Tabelle 2.4: Scores für die Jungtiere im Wasser-Irrgarten-Test.

Placebo			Dosis 1			Dosis 2			Dosis 3		
	Tag			Tag			Tag			Tag	
Mutter	1	7	Mutter	1	7	Mutter	1	7	Mutter	1	7
1	13	21	1	14	21	1	19	18	1	9	18
	11	21		21	21		21	21		20	14
2	14	13	2	20	21	2	12	21	2	15	11
	20	21		11	20		21	21		9	5
3	17	21	3	21	21	3	16	20	3	13	21
	14	21		13	21		5	14		20	11
4	7	20	4	20	20	4	20	21	4	21	21
	0	16		21	20		20	21		8	21
5	15	21	5	1	15	5	19	0	5	21	21
	16	17		0	18		20	21		18	21
6	20	21	6	20	17	6	8	18	6	12	19
	15	21		20	19		2	21		11	15
7	14	18	7	14	18	7	19	19	7	14	21
	18	21		19	21		19	21		0	18
8	21	21	8	21	19	8	10	21	8	19	9
	14	20		15	12		17	15		18	21
9	7	19	9	18	15	9	21	17	9	11	3
	11	6		15	13		21	20		9	14
10	20	21	10	21	21	10	11	21	10	18	21
	21	21		11	18		20	21		8	21
11	20	21	11	20	21	11	18	0	11	11	7
	5	20		18	21		6	21		6	21
12	12	19	12	6	21	12	6	18	12	13	21
	13	15		11	21		15	21		6	15
13	1	19	13	21	21	13	15	21	13	1	21
	17	21		18	21		19	21		5	12
14	18	17	14	21	16	14	13	20	14	12	21
	21	15		18	21		4	6		6	0
15	16	21	15	14	21	15	16	21	15	13	19
	18	18		21	21		19	19		15	18
16	11	20	16	21	21	16	0	6	16	10	14
	21	21		15	20		0	5		14	11
17	20	14	17	8	13	17	3	19	17	12	19
	19	21		11	0		17	19		12	16
						18	11	21	18	19	13
							19	21		14	21
						19	19	16	19	18	21
							21	19		21	21

Daraus wurde für jedes Tier durch Addition ein Score errechnet: Das Finden des Ausgangs wurde beim ersten Mal mit einem Punkt bewertet, beim zweiten Mal mit zwei Punkten und bei den folgenden Malen mit jeweils einem Punkt mehr, so dass die Scores einen Wertebereich von 0 bis 21 haben. Die Versuche wurden nach sechs Tagen wiederholt und auf dieselbe Art und Weise wurde ein zweiter Score berechnet. Die Scores aller Wistar-Ratten befinden sich in Tabelle 2.4.

In dieser Studie gibt es $a = 4$ Gruppen von Versuchseinheiten. Dabei ist zu beachten, dass unter den Versuchseinheiten hier die Muttertiere und nicht die Jungtiere zu verstehen sind. Für jedes Muttertier gibt es unter $t = 2$ Bedingungen (= Tagen) jeweils zwei Beobachtungen: je eine an jedem der beiden Jungtiere am ersten und am siebten Versuchstag. Demnach gilt für alle Paare (i, k)

$$\boldsymbol{m}_{ik} = (m_{ik1}, m_{ik2})' = (2, 2)'.$$

Für jeden Versuchstag und jedes Muttertier liegt also ein Paar

$$\boldsymbol{X}_{iks} = (X_{iks1}, X_{iks2})'$$

von Werten vor. Diese werden wiederum zu Vektoren

$$\boldsymbol{X}_{ik} = (\boldsymbol{X}'_{ik1}, \boldsymbol{X}'_{ik2})' = (X_{ik11}, X_{ik12}, X_{ik21}, X_{ik22})'$$

zusammengefasst.

2.1.2 Bedingungen an die Stichprobenumfänge

Einige der noch folgenden mathematischen Ergebnisse aus den Kapiteln über Punktschätzer und Konfidenzintervalle sind asymptotischer Natur. Das heißt, dass sie Aussagen über das Verhalten von Punktschätzern und Konfidenzintervallen für eine Folge von wachsenden Stichprobenumfängen $(n_1, \ldots, n_a)$ treffen. Eine Schreibweise, die eine korrekte Darstellung einer solchen Folge von Stichprobenumfängen erlauben würde, käme ohne einen zusätzlichen Index für die Symbole n_i und $\lambda_{i \cdot s}$ zur Kennzeichnung der Folgenglieder nicht aus. Angesichts der ohnehin schon großen Anzahl von Indizes verzichten wir zugunsten der Lesbarkeit auf diesen weiteren Index.

In Bezug auf den Versuchsplan sind asymptotische Resultate für Folgen von wachsenden Stichprobenumfängen insofern zu berücksichtigen, als die asymptotischen Eigenschaften der statistischen Verfahren nur bei hinreichend großen Stichprobenumfängen als in ausreichendem Maß erfüllt angesehen werden dürfen. Darüber hinaus sind bei den Resultaten noch von Fall zu Fall unterschiedliche Teile der folgenden drei Annahmen über die Folgen von Stichprobenumfängen zu treffen.

(A1) $\lambda_{i \cdot s} \to \infty$ für alle Paare (i, s)

(A2) $n/\lambda_{i \cdot s} \leq n_0 < \infty$ für alle Paare (i, s)

(A3) $m_{iks} \leq m_0 < \infty$ für alle Tripel (i, k, s)

Diese Annahmen bedeuten für einen Versuchsplan mit endlichen Stichprobenumfängen, dass zu prüfen ist, ob

(A1) in jeder Gruppe unter jeder Bedingung hinreichend viele Versuchseinheiten mit mindestens einer Beobachtung vorhanden sind,

(A2) der Versuchsplan in Bezug auf die Zahlen von Beobachtungen an verschiedenen Versucheinheiten in den einzelnen Gruppen und unter den verschiedenen Bedingungen hinreichend balanciert ist,

(A3) die Zahlen von Beobachtungen unter gleichen Bedingungen an den einzelnen Versucheinheiten nicht zu groß sind.

Die Frage, wann Stichprobenumfänge hinreichend groß sind oder ein Versuch ausreichend balanciert ist, wird durch asymptotische Aussagen nicht beantwortet. Sie ist aber für die Praxis von größter Bedeutung. Während es – zumindest bei experimentellen Studien – meist möglich sein wird, vollständig balancierte Versuchspläne anzustreben, ist die Forderung nach möglichst großen Stichprobenumfängen in Anbetracht der stets vorhandenen äußeren Zwänge, wie begrenzten Mitteln oder ethischen Bedenken, wenig hilfreich.

Besonders erstrebenswert ist es, bereits vor der Versuchsdurchführung Informationen über benötigte Stichprobenumfänge zu erlangen, um diese während der Versuchsplanung berücksichtigen zu können. Dazu müssen aber auch vor dem Versuch wenigstens ungefähre Vorstellungen über das dem Versuch zugrunde liegende wahrscheinlichkeitstheoretische Modell vorhanden sein. Dann nämlich können einige Modelle ausgewählt werden, die innerhalb des gesamten in Betracht bezogenen Bereichs von Modellen hinreichend dicht liegen. Für diese Modelle kann stellvertretend geprüft werden, wann die Stichprobenumfänge groß genug sind.

Nichtparametrische Verfahren werden jedoch meist dann verwendet, wenn die Klasse von möglichen wahrscheinlichkeitstheoretischen Modellen sehr groß ist. Dies erschwert die Bestimmung hinreichend großer Stichprobenumfänge vor der Versuchsdurchführung mit der oben beschriebenen Vorgehensweise erheblich. Aus diesem Grund muss man sich unter Umständen damit zufriedengeben, erst im Anschluss an den Versuch auf Grundlage der Ergebnisse schätzen zu können, ob die Stichprobenumfänge hinreichend groß waren.

2.2 Wahrscheinlichkeitstheoretische Modelle

Diese Arbeit handelt von schließender Statistik: Die Entstehung der Beobachtungen wird durch einen Zufallsmechanismus modelliert und auf Grundlage der Beobachtungen werden Rückschlüsse auf diesen Zufallsmechanismus gezogen. Die hier getroffenen Annahmen über den Zufallsmechanismus, die auch als wahrscheinlichkeitstheoretisches Modell bezeichnet werden, sind – entsprechend den in der Einleitung beschriebenen Grundgedanken der nichtparametischen Statistik – sehr allgemein gehalten und einfach zu prüfen. Dennoch sollte auch die Verwendung eines solchen Modells nicht als selbstverständlich angesehen werden. Im Zweifel sollte man sich auf die Verfahren der beschreibenden Statistik beschränken, die ohne wahrscheinlichkeitstheoretische Modelle auskommen (vgl. Freedman et al., 1998, Kapitel 29, Abschnitt 4).

Das hier verwendete Modell umfasst drei Annahmen:

- Die den Versuchseinheiten entsprechenden Vektoren von Beobachtungen $\boldsymbol{X}_{ik}$ sind Zufallsvektoren, deren Verteilungen nur von den Gruppen der Versuchseinheiten und den Dimensionsvektoren $\boldsymbol{m}_{ik}$ abhängen.
- Die Randverteilungen der Zufallsvariablen $X_{iks\ell}$ hängen nur von den Gruppen und Bedingungen, also nur von den Indizes i und s ab.
- Alle Zufallsvektoren $\boldsymbol{X}_{ik}$ sind unabhängig.

Diese Annahmen lassen sich wie folgt kurz darstellen:

$$\boldsymbol{X}_{ik} \text{ u. i. v. } \sim F_{i\boldsymbol{m}_{ik}}, \; X_{iks\ell} \sim F_{is}.$$

Ein solches Modell wird häufig dann verwendet, wenn die Versuchseinheiten in den verschiedenen Gruppen als zufällige Stichproben aus hinreichend großen Grundgesamtheiten aufgefasst werden können. Ziel des Versuchs ist dann, Informationen über die Verteilungen der beobachteten Größe in den Grundgesamtheiten unter den verschiedenen Bedingungen zu erhalten, die durch die Verteilungsfunktionen F_{is} dargestellt werden.

Im Hinblick auf Schreibweisen, die später noch eingeführt werden, fassen wir die lexikographisch nach den Indexpaaren (i, s) geordneten Randverteilungen in Form eines Vektors

$$\boldsymbol{F} = (F_{11}, \ldots, F_{1t}, \ldots, F_{a1}, \ldots, F_{at})'$$

zusammen.

2.2.1 Zufällige Messwiederholungsanzahlen

In den meisten praktischen Fällen wird im Rahmen der Versuchsplanung angestrebt, für alle Gruppen, Versuchseinheiten und Bedingungen gleiche Anzahlen an Messwiederholungen m_{iks} zu erhalten. Trotzdem ist es sinnvoll, auch den Fall unterschiedlich vieler Messwiederholungen in Betracht zu ziehen, denn nicht immer gelingt es, alle geplanten Beobachtungen auch tatsächlich vorzunehmen. Insbesondere der Fall fehlender Werte ($m_{iks} = 0$) ist von großer praktischer Bedeutung, was unter anderem daran abzulesen ist, dass zu diesem Thema bereits mehrere Monographien erschienen sind (Dodge, 1985; Little und Rubin, 1987; Rubin, 1987; Schafer, 1997).

Wie viele der angestrebten Messwiederholungen tatsächlich durchgeführt werden können, wird als zufällig angesehen, weil dies nicht schon im Rahmen der Versuchsplanung festgelegt wird. Um ein vollständiges wahrscheinlichkeitstheoretisches Modell für die Versuchsdaten zu erhalten, ist dieser Zufallsmechanismus in Beziehung zu dem oben beschriebenen Zufallsmechanismus zu setzen, der die Verteilung der Beobachtungen $X_{iks\ell}$ bestimmt. Dies ist implizit bereits geschehen, denn die Verteilungsfunktionen $F_{i\boldsymbol{m}_{ik}}$ hängen durch die gewählte Schreibweise von den Zufallsvektoren $\boldsymbol{m}_{ik}$ ab und können somit nur als bedingte Verteilungsfunktionen aufgefasst werden.

Die oben getroffene Annahme, dass die Verteilungsfunktionen F_{is} der Zufallsvariablen $X_{iks\ell}$ unabhängig von den Anzahlen an Messwiederholungen m_{iks} sind, ist nur dann gerechtfertigt, wenn der Ausfall einer geplanten Beobachtung unabhängig davon ist, welcher Wert beobachtet worden wäre. Diese Unabhängigkeit entspricht in der von Rubin (1976) eingeführten Terminologie für die Klassifikation von zufälligen Mechanismen, die zu fehlenden Werten führen, der Annahme, dass die Werte *vollständig zufällig* fehlen. Nur bei vollständig zufällig fehlenden Werten können die beobachteten Werte als repräsentativ für die Werte in der zu Grunde liegenden Gesamtheit angesehen werden.

2.2.2 Erläuterungen anhand der Beispiele

Im Folgenden werden die in diesem Abschnitt formulierten Annahmen anhand der Beispiele aus Unterabschnitt 2.1.1 diskutiert.

Panik-Skala-Studie I

Wenn wir annehmen, dass die 16 Patienten zufällig aus einer hinreichend großen Population herausgegriffen wurden, können die Vektoren $\boldsymbol{X}_{1k}$ als unabhängige Zufallsvektoren mit gleicher multivariater Verteilung aufgefasst werden. Daraus ergibt sich die Gültigkeit des formulierten Modells.

Da alle m_{1ks} den Wert 1 haben, erübrigt sich die Betrachtung zufälliger Stichprobenumfänge.

Panik-Skala-Studie II

Hier gehen wir zunächst davon aus, dass sowohl die Patienten mit Agoraphobie als auch die Patienten ohne Agoraphobie zufällig aus hinreichend großen Populationen ausgewählt wurden. Damit können wir die um die unbekannten fehlenden Werte ergänzten Vektoren $\widetilde{\boldsymbol{X}}_{ik}$ als unabhängige Zufallsvektoren mit den multivariaten Verteilungsfunktionen F_1 in der ersten Gruppe und F_2 in der zweiten Gruppe ansehen.

Falls der Zufallsmechanismus, der die fehlenden Werte verursacht, unabhängig von den Zufallsvektoren $\widetilde{\boldsymbol{X}}_{ik}$ ist (was hier durchaus in Frage zu stellen ist), so stimmen die Verteilungen der tatsächlich beobachteten Werte X_{iks} mit denen der Komponenten von $\widetilde{\boldsymbol{X}}_{ik}$ überein und hängen deshalb nur von den Indizes i und k ab.

Schulterschmerz-Studie

Auch bei dieser Studie treffen wir die Annahme, dass alle teilnehmenden Patienten eine zufällige Stichprobe aus einer hinreichend großen Obermenge von Patienten sind und sehen deshalb die Vektoren $\boldsymbol{X}_{ik}$ als unabhängige Zufallsvektoren mit den multivariaten Verteilungsfunktionen F_1 für die Behandlung Y beziehungsweise F_2 für die Behandlung N an. Damit ergibt sich die Gültigkeit des formulierten Modells.

Da alle m_{iks} den Wert 1 haben, erübrigt sich die Betrachtung zufälliger Stichprobenumfänge.

Wasser-Irrgarten-Test

Wie bei den drei anderen Studien gehen wir davon aus, dass die Versuchseinheiten, hier also die Muttertiere, zufällig aus einer größeren Population ausgewählt wurden. Deshalb können wir die Vektoren $\boldsymbol{X}_{ik}$ als unabhängige Zufallsvektoren mit den multivariaten Verteilungsfunktionen F_i, $i = 1, \ldots, 4$, auffassen. Aber erst dadurch, dass von jeder Mutter *zufällig* zwei Jungtiere ausgewählt wurden, kann die Annahme gerechtfertigt werden, dass die Komponenten der Vektoren

$$\boldsymbol{X}_{iks} = (X_{iks1}, X_{iks2})'$$

für alle Paare (i, s) die gleiche Randverteilungsfunktion F_{is} haben.

Da alle m_{iks} den Wert 2 haben, entfällt die Betrachtung zufälliger Stichprobenumfänge.

Kapitel 3

Relative Effekte

In diesem Kapitel definieren wir die relativen Effekte als statistische Funktionale für die im vorangehenden Kapitel vorgestellten Versuchspläne und wahrscheinlichkeitstheoretischen Modelle. Ihre Eigenschaften und die daraus resultierenden Interpretationen werden ausführlich diskutiert.

Um die Darstellung so einfach wie möglich zu halten, behandeln wir relative Effekte in den ersten beiden Abschnitten dieses Kapitels zunächst für die einfachen Versuchspläne mit zwei unverbundenen oder zwei verbundenen Stichproben. Erst im dritten Abschnitt verallgemeinern wir die Definition relativer Effekte auf die allgemeine Klasse von Versuchsplänen aus Abschnitt 2.1.

3.1 Zwei unverbundene Stichproben

Für den einfachen Versuchsplan mit zwei unverbundenen Stichproben ($a = 2$ Gruppen, $t = 1$ Bedingung, alle $m_{ik1} = 1$) wählen wir zunächst eine gegenüber dem vorangehenden Kapitel vereinfachte Schreibweise:

$$X_{ik} \sim F_i,\ i = 1, 2,\ k = 1, \ldots, n_i.$$

Als relative Effekte definieren wir für diesen Versuchsplan die Linearkombinationen von Wahrscheinlichkeiten

$$p_1 = \mathrm{P}(X_{21} < X_{11}) + \frac{1}{2}\,\mathrm{P}(X_{21} = X_{11}), \qquad (3.1)$$

$$p_2 = \mathrm{P}(X_{11} < X_{21}) + \frac{1}{2}\,\mathrm{P}(X_{11} = X_{21}). \qquad (3.2)$$

Wir nennen p_i den relativen Effekt der i-ten Gruppe.

In den beiden folgenden Unterabschnitten wird zunächst die geschichtliche Entwicklung geschildert, die zur Definition der relativen Effekte in obiger Form führte. Danach werden einige wichtige Eigenschaften und die Interpretation der relativen Effekte diskutiert.

3.1.1 Geschichtliche Entwicklung

In diesem Unterabschnitt verwenden wir immer die oben eingeführten Schreibweisen für die Zufallsvariablen und die relativen Effekte, auch wenn wir Teile von Veröffentlichungen wiedergeben, in denen andere Notationen benutzt werden.

Ausgangspunkt für die spätere Definition relativer Effekte ist die Arbeit von Wilcoxon (1945), in der ein statistischer Test für die Hypothese $F_1 = F_2$ und den Spezialfall $n_1 = n_2$ angegeben wird. Die Teststatistik des Wilcoxon-Tests ist die Rangsumme aus der zweiten Stichprobe

$$R_{2\cdot} = \sum_{k=1}^{n_2} R_{2k}.$$

Dabei steht R_{ik} für den Mittelrang (im folgenden kurz mit Rang bezeichnet) der Beobachtung X_{ik} unter allen $n_1 + n_2$ Beobachtungen. Wie Kruskal (1957) berichtet, war Wilcoxon zwar nicht der erste, der einen solchen Test veröffentlichte, jedoch scheinen die früheren Arbeiten von Deuchler (1914) und Ottaviani (1939) keine breite Beachtung gefunden zu haben.

Der Test von Wilcoxon wurde von Mann und Whitney (1947) unter Benutzung der äquivalenten Teststatistik

$$U = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_{2\cdot}$$

auf den Fall $n_1 \neq n_2$ verallgemeinert. Jedoch beschränkten sich Mann und Whitney dabei auf stetige Verteilungsfunktionen F_1 und F_2. In ihrer Arbeit findet sich in einer Nebenrechnung auf Seite 58 der Ausdruck

$$\mathrm{P}(X_{11} > X_{21}),$$

der bei stetigen Verteilungsfunktionen identisch mit dem relativen Effekt p_1 ist. Dieser Ausdruck wird aber von Mann und Whitney nicht weiter diskutiert.

Größere Aufmerksamkeit wurde den relativen Effekten im Zusammenhang mit dem Wilcoxon-Mann-Whitney-Test von van Dantzig (1951) entgegengebracht: Er stellte fest, dass der einseitige Wilcoxon-Mann-Whitney-Test genau dann konsistent ist, wenn p_1 größer (bzw. kleiner) als $1/2$ ist, und dass die zweiseitige Variante genau im Fall $p_1 \neq 1/2$ konsistent ist. Das erste dieser beiden Ergebnisse wurde unabhängig von van Dantzig auch von Lehmann (1951) angegeben. Sowohl van Dantzig als auch Lehmann beschränken sich allerdings bei ihren Betrachtungen auf stetige Verteilungsfunktionen. Die Verallgemeinerung ihrer Ergebnisse auf nicht stetige Verteilungen wurde von Putter (1955) erarbeitet.

Die erste Arbeit, die die relativen Effekte um ihrer selbst willen und nicht nur wegen ihrer Beziehungen zum Wilcoxon-Mann-Whitney-Test behandelt, stammt von Birnbaum (1956). Das in der Arbeit von Birnbaum betrachtete Anwendungsbeispiel für die relativen Effekte fällt in das Gebiet der Technometrie: Wenn X_{11} die Haltbarkeit eines Produktes ist und X_{21} die Belastung, der das Produkt ausgesetzt ist, so ist der relative Effekt p_1 die Wahrscheinlichkeit, dass das Produkt der Belastung standhält. Da Haltbarkeiten und Belastungen in der Technometrie meist als stetige Größen aufgefasst werden, beschränkt sich Birnbaum auf den Fall stetiger Verteilungsfunktionen F_1 und F_2, so dass in seiner Definition der relativen Effekte der Summand $\mathrm{P}(X_{21} = X_{11})/2$ nicht gebraucht wird.

Die von Birnbaum vorgeschlagene Anwendung war für die Technometrie Auslöser einer Reihe von Veröffentlichungen zu relativen Effekten, zum Teil auch in parametrischen Modellen. So existieren Arbeiten zu normalverteilten Beobachtungen (z. B. Owen et al., 1964), zu exponentialverteilten Beobachtungen (z. B. Tong, 1974; Johnson, 1975; Tong, 1975a) und zu gammaverteilten Beobachtungen (z. B. Woodward und Gray, 1975). Darüber hinaus wurden im Zusammenhang mit relativen Effekten Bayes-Modelle (Enis und Geisser, 1971), semiparametrische Modelle (Ury und Wiggins, 1979) und Kontaminationsmodelle (Harris und Soms, 1983) angewandt. Diese große Anzahl von Modellen, die keinen nichtparametrischen Charakter tragen, unterstreicht die gute Interpretierbarkeit der relativen Effekte. Man erkennt, dass ihre Bedeutung über die eines künstlichen Parameter-Ersatzes in der nichtparametrischen Statistik hinausgeht. Allerdings muss angemerkt werden, dass die von Birnbaum vorgeschlagene technometrische Interpretation der relativen Effekte für zwei Stichproben bei der in in dieser Arbeit (siehe Abschnitt 3.3) vorgeschlagenen Verallgemeinerung der relativen Effekte auf mehrere Stichproben oder faktorielle Versuchspläne verloren geht. Dennoch gibt es auch in der Technometrie Ansätze zu Verallgemeinerungen (Bhattacharyya und Johnson, 1974; Chandra und Owen, 1975), die sich aber von den in dieser Arbeit betrachteten relativen Effekten für mehr als zwei Stichproben grundsätzlich unterscheiden.

Die erste Definition relativer Effekte, die durch Addition von $\mathrm{P}(X_{21} = X_{11})/2$ auch unstetige Verteilungen sinnvoll berücksichtigt, wurde von Bross (1958) in seiner Arbeit zur Ridit-Analyse angegeben. Diese Arbeit enthält neben einer ausführlichen Diskussion einiger grundlegender Eigenschaften der relativen Effekte weitere Anwendungsbeispiele. Diese fallen im Gegensatz zu dem von Birnbaum genannten Beispiel aber nicht in den Bereich der Technometrie, sondern in den Bereich der Biometrie. Für diese Beispiele ist auch die in der vorliegenden Arbeit vorgeschlagene Verallgemeinerung der relativen Effekte auf mehrere Stichproben oder faktorielle Versuchspläne sinnvoll.

In der Folgezeit wurde die Verwendung der von Bross eingeführten relativen Effekte für zwei unverbundene Stichproben weiter diskutiert, aber auch kritisiert. Unter den Autoren finden sich zum Beispiel Schwartz und Wearden (1959), Mantel (1963), Wolfe und Hogg (1971), Mantel (1979), Selvin (1979), Bross (1979), Beder und Heim (1990, 1992), McGraw und Wong (1992), Senn (1997), Chen und Kianifard (2000), Hauck et al. (2000) und Vargha und Delaney (2000).

3.1.2 Eigenschaften und Interpretation

Um bei der späteren Diskussion der Eigenschaften und möglichen Interpretationen der relativen Effekte über einen etwas konkreteren Eindruck von diesen Funktionalen zu verfügen, beginnen wir diesen Unterabschnitt mit der beispielhaften Berechnung der relativen Effekte für Bernoulli- und Normalverteilungen.

Relative Effekte bei Bernoulli-Verteilungen

Die Zufallsvariablen X_{ik} seien Bernoulli-verteilt mit den Erfolgswahrscheinlichkeiten

$$\mathrm{P}(X_{ik} = 1) = q_i = 1 - \mathrm{P}(X_{ik} = 0), \quad i = 1, 2.$$

Als relativer Effekt p_1 ergibt sich

$$\begin{aligned}
p_1 &= \mathrm{P}(X_{21} < X_{11}) + \frac{1}{2}\,\mathrm{P}(X_{21} = X_{11}) \\
&= \mathrm{P}(\{X_{21} = 0\} \cap \{X_{11} = 1\}) + \frac{1}{2}\,\mathrm{P}(\{X_{21} = X_{11} = 0\} \cup \{X_{21} = X_{11} = 1\}) \\
&= (1 - q_2)q_1 + \frac{1}{2}\left((1 - q_1)(1 - q_2) + q_1 q_2\right) \\
&= q_1 - q_1 q_2 + \frac{1}{2} - \frac{q_1}{2} - \frac{q_2}{2} + \frac{q_1 q_2}{2} + \frac{q_1 q_2}{2} \\
&= \frac{q_1 - q_2}{2} + \frac{1}{2}. \qquad (3.3)
\end{aligned}$$

Analog erhalten wir

$$p_2 = \frac{q_2 - q_1}{2} + \frac{1}{2}. \qquad (3.4)$$

Für Bernoulli-Verteilungen sind die relativen Effekte also äquivalent zur Differenz der Erfolgswahrscheinlichkeiten $q_1 - q_2$.

Relative Effekte bei Normalverteilungen

Die Zufallsvariablen X_{ik} seien normalverteilt mit Erwartungwert μ_i und Varianz σ_i^2:

$$X_{ik} \sim \mathrm{N}(\mu_i, \sigma_i^2).$$

Wegen der Stetigkeit der Normalverteilung und wegen

$$\frac{X_{21} - X_{11} + \mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \sim \mathrm{N}(0, 1)$$

gilt für den relativen Effekt p_1

$$\begin{aligned}
p_1 &= \mathrm{P}(X_{21} < X_{11}) + \frac{1}{2}\,\mathrm{P}(X_{21} = X_{11}) \\
&= \mathrm{P}(X_{21} \leq X_{11}) \\
&= \mathrm{P}(X_{21} - X_{11} \leq 0) \\
&= \mathrm{P}(X_{21} - X_{11} + \mu_1 - \mu_2 < \mu_1 - \mu_2) \\
&= \mathrm{P}\left(\frac{X_{21} - X_{11} + \mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \leq \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) \\
&= \Phi\left(\frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right), \qquad (3.5)
\end{aligned}$$

wobei Φ die Verteilungsfunktion der Standardnormalverteilung bezeichnet. Analog berechnen wir

$$p_2 = \Phi\left(\frac{\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right). \tag{3.6}$$

Die relativen Effekte p_i sind bei normalverteilten Zufallsvariablen also äquivalent zur standardisierten Differenz der Mittelwerte

$$\frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}.$$

Integraldarstellung der relativen Effekte

Mit Hilfe des Satzes von Fubini können wir für die relativen Effekte eine Integraldarstellung mit den Verteilungsfunktionen gewinnen. Dabei ist in dieser Arbeit mit dem Begriff Verteilungsfunktion immer, wenn nicht ausdrücklich etwas anderes gesagt wird, die auf Lévy (1925) zurückgehende normalisierte Version der Verteilungsfunktion gemeint, also das Mittel aus rechts- und linksstetiger Version der Verteilungsfunktion. Für die in diesem Abschnitt verwendeten Verteilungsfunktionen F_i gilt demnach:

$$F_i(x) = \frac{1}{2}[F_i^-(x) + F_i^+(x)] = \mathrm{P}(X_{i1} < x) + \frac{1}{2}\,\mathrm{P}(X_{i1} = x), \quad i = 1, 2.$$

Durch Anwendung des Satzes von Fubini erhalten wir

$$\begin{aligned} p_1 &= \mathrm{P}(X_{21} < X_{11}) + \frac{1}{2}\,\mathrm{P}(X_{21} = X_{11}) \\ &= \int \mathrm{P}(X_{21} < x)\,dF_1(x) + \int \frac{1}{2}\,\mathrm{P}(X_{21} = x)\,dF_1(x) \\ &= \int \left[\mathrm{P}(X_{21} < x) + \frac{1}{2}\,\mathrm{P}(X_{21} = x)\right] dF_1(x) \\ &= \int F_2\,dF_1. \end{aligned} \tag{3.7}$$

Analog ergibt sich

$$p_2 = \int F_1\,dF_2.$$

Die eingeführte Integraldarstellung liefert zwar keine anschauliche Interpretation der relativen Effekte, wird aber unter anderem bei der Verkürzung von Schreibweisen und der Herleitung der Schätzer in Kapitel 4 hilfreich sein.

Relative Effekte als Wahrscheinlichkeiten

Grundlegend für den leicht verständlichen Umgang mit relativen Effekten sind ihre Definitionen als Wahrscheinlichkeiten (genauer: als Linearkombinationen von Wahrscheinlichkeiten). Der Begriff der Wahrscheinlichkeit ist elementarer und für den mathematischen Laien einfacher verständlich als die in der nichtparametrischen Statistik häufig verwendeten Begriffe des Medians und des Erwartungswerts. Insbesondere die allgemeine Definition des Erwartungswerts ist mathematisch aufwendig und dem Laien nur mit Hilfe der Gesetze der großen Zahlen als die Zahl zu erklären, der sich ein arithmetischer Mittelwert von unabhängigen Beobachtungen annähert. Die größere Verständlichkeit des Begriffs der Wahrscheinlichkeit im Vergleich mit den Begriffen des Medians und des Erwartungswerts zeigt sich auch daran, dass ersterer als einziger von den drei Begriffen Bestandteil der Alltagssprache ist. Die Bedeutung dieses Aspektes wird in der Literatur insbesondere von Wolfe und Hogg (1971), McGraw und Wong (1992) und Vargha und Delaney (2000) betont. Jedoch sollte ebenfalls bedacht werden, dass die relativen Effekte im Gegensatz zum Erwartungswert und Median nicht auf derselben Skala wie die Beobachtungen angegeben werden. Dies muss aber wegen der hohen Verständlichkeit der Skala der Wahrscheinlichkeiten kein Nachteil sein.

Das Einheitsintervall als Skala möglicher Wahrscheinlichkeiten bietet für die Werte der relativen Effekte feste Orientierungspunkte: Der Wert 0 für den relativen Effekt p_1 bedeutet, dass Beobachtungen mit der Verteilungsfunktion F_1 fast sicher kleiner sind als Beobachtungen mit der Verteilungsfunktion F_2. Umgekehrt bedeutet ein relativer Effekt $p_1 = 1$, dass Beobachtungen mit der Verteilungsfunktion F_1 fast sicher größer sind als Beobachtungen mit der Verteilungsfunktion F_2. Der Fall $p_1 = 1/2$ bedeutet, dass keine der beiden Verteilungsfunktionen zu größeren Werten als die andere tendiert. Aber auch anderen Werten als 0, 1/2 und 1 lassen sich leicht verständliche Bedeutungen zuordnen: Wenn p_1 kleiner als 1/2 ist, neigen Beobachtungen mit der Verteilungsfunktion F_1 eher zu kleineren Werten als Beobachtungen mit der Verteilungsfunktion F_2, und wenn p_1 größer als 1/2 ist, neigen Beobachtungen mit der Verteilungsfunktion F_1 eher zu größeren Werten. Falls relative Effekte in der Nähe der Ränder 0 und 1 liegen, kann man von sehr kleinen oder sehr großen relativen Effekten sprechen. Relative Effekte um die Mitte 1/2 bedeuten, dass keine der beiden Verteilungfunktionen in bedeutsamem Maß zu größeren Werten als die andere tendiert. Wir halten die soeben eingeführten Sprechweisen in einer Definition fest.

Definition 3.1. Die Verteilungsfunktion F_1 heißt tendenziell kleiner als die Verteilungsfunktion F_2, falls $p_1 < 1/2$ ist. Die Verteilungsfunktion F_1 heißt tendenziell größer als die Verteilungsfunktion F_2, falls $p_1 > 1/2$ ist. Die Verteilungsfunktionen F_1 und F_2 heißen tendenziell gleich, falls $p_1 = 1/2$ ist.

Für die Beziehung zwischen Zufallsvariablen $X_{1k} \sim F_1$ und $X_{2k} \sim F_2$ verwenden wir die gleichen Sprechweisen.

Aus Symmetriegründen sollte, wenn F_1 und F_2 tendenziell gleich sind, nicht nur $p_1 = 1/2$, sondern auch $p_2 = 1/2$ gelten. Der folgende Satz bestätigt die Gültigkeit dieser Forderung.

Satz 3.1. *Für die relativen Effekte p_1 und p_2 gilt*

$$p_1 + p_2 = 1.$$

Beweis. Die Ereignisse $\{X_{21} < X_{11}\}$, $\{X_{21} = X_{11}\}$ und $\{X_{11} < X_{21}\}$ bilden eine Zerlegung des Stichprobenraums. Damit erhalten wir

$$\begin{aligned} p_1 + p_2 &= \mathrm{P}(X_{21} < X_{11}) + \frac{1}{2}\,\mathrm{P}(X_{21} = X_{11}) + \mathrm{P}(X_{11} < X_{21}) + \frac{1}{2}\,\mathrm{P}(X_{11} = X_{21}) \\ &= \mathrm{P}(X_{21} < X_{11}) + \mathrm{P}(X_{21} = X_{11}) + \mathrm{P}(X_{11} < X_{21}) = 1. \end{aligned}$$

□

Als Folgerung aus obigem Satz erhalten wir die für das Rechnen mit Integralen nützliche Rechenregel der partiellen Integration.

Korollar 3.2 (Partielle Integration). *Für Verteilungsfunktionen F_1 und F_2 gilt*

$$\int F_2\,dF_1 = 1 - \int F_1\,dF_2.$$

Beweis. Aus den Integraldarstellungen der relativen Effekte und Satz 3.1 folgt

$$\int F_2\,dF_1 = p_1 = 1 - p_2 = 1 - \int F_1\,dF_2.$$

□

In bestimmten Situationen haben wir eine intuitive Vorstellung davon, ob die Verteilungsfunktion F_1 tendenziell kleiner als F_2, gleich F_2 oder größer als F_2 sein sollte. So gehen wir davon aus, dass zwei gleiche Verteilungen oder zwei symmetrische Verteilungen mit dem gleichen Symmetriezentrum auch tendenziell gleich sind. Wenn F_1 und F_2 symmetrisch sind und das Symmetriezentrum von F_1 kleiner (größer) als das von F_2 ist, scheint F_1 tendenziell kleiner (größer) als F_2 zu sein. Falls F_1 stochastisch kleiner bzw. größer als F_2 ist, erwarten wir, dass F_1 auch tendenziell kleiner bzw. größer als F_2 ist. Für die Spezialfälle der Bernoulli- und Normalverteilung können wir unsere intuitiven Vorstellungen anhand der Formeln (3.3) – (3.6) bestätigen. Der folgende Satz zeigt, dass die Begriffe aus Definition 3.1 den formulierten Erwartungen bis auf eine Ausnahme (siehe Beispiel 3.1) auch im Allgemeinen gerecht werden.

Satz 3.3. (a) *Falls $F_1 = F_2$ ist oder X_{11} und X_{21} symmetrisch um das gleiche Symmetriezentrum μ verteilt sind, so sind F_1 und F_2 tendenziell gleich.*

(b) *Falls F_1 und F_2 symmetrisch sind und das Symmetriezentrum μ_1 von F_1 kleiner (größer) ist als das Symmetriezentrum μ_2 von F_2, so ist F_1 tendenziell kleiner (größer) als F_2 oder tendenziell gleich F_2.*

(c) *Wenn F_1 stochastisch kleiner (größer) als F_2 ist, dann ist F_1 tendenziell kleiner (größer) als F_2.*

Beweis. Siehe Anhang A.1, Seite 95. □

Wenn wir die erste Aussage von Satz 3.3 (a) auf die Integraldarstellung der relativen Effekte anwenden, erhalten wir

Korollar 3.4. *Für jede Verteilungsfunktion F gilt*

$$\int F\,dF = \frac{1}{2}.$$

Wie folgendes Beispiel zeigt, kann Teil (b) von Satz 3.3 im Allgemeinen nicht verschärft werden.

Beispiel 3.1. Es gelte $\mathrm{P}(X_{11} = 1) = 1$ und $\mathrm{P}(X_{21} = -2) = \mathrm{P}(X_{21} = 2) = 1/2$. Obwohl F_1 und F_2 symmetrisch sind und das Symmetriezentrum von F_1 größer ist als das von F_2, sind F_1 und F_2 tendenziell gleich. □

Wenden wir die Teile (a) und (c) von Satz 3.3 in einem Verschiebungsmodell an, so ergibt sich

Korollar 3.5. *Für alle $x \in \mathbb{R}$ gelte*

$$F_2(x) = F_1(x + \mu).$$

Falls der Verschiebungsparameter μ negativ (positiv) ist, so ist F_1 tendenziell kleiner (größer) als F_2. Die Verteilungsfunktionen F_1 und F_2 sind genau dann tendenziell gleich, wenn $\mu = 0$ ist.

Relative Effekte für ordinale Daten

In der Definition der relativen Effekte wird für die Beobachtungen lediglich vorausgesetzt, dass sie miteinander verglichen werden können und dass somit festgestellt werden kann, ob zwei Beobachtungen gleich sind oder – falls das nicht der Fall ist – welche von zwei Beobachtungen größer ist. Es wird nicht vorausgesetzt, dass Summen von Beobachtungen erklärt sind, wie es beispielsweise zur sinnvollen Bildung von Erwartungswerten notwendig wäre. Relative Effekte können also immer verwendet werden, wenn das Skalenniveau der Beobachtungen mindestens ordinal ist. Die Beispiele aus Unterabschnitt 2.1.1 belegen die praktische Bedeutsamkeit dieser Tatsache, denn in allen Beispielen ist das Skalenniveau der Beobachtungen ordinal, aber nicht metrisch.

Die sehr allgemeine Verwendbarkeit von relativen Effekten für eine große Klasse von Skalenniveaus kann allerdings nur dadurch erreicht werden, dass die relativen Effekte ausschließlich Informationen über die Ordnungsstruktur der Beobachtungen enthalten. Jede darüber hinausgehende Information spiegelt sich in den relativen Effekten nicht mehr wieder, denn die relativen Effekte sind offensichtlich invariant unter streng wachsenden Transformationen der Beobachtungen. Wie Mantel (1979) in seiner Kritik an der Arbeit von Bross (1958) gezeigt hat, können dabei wichtige Informationen verloren gehen. Das folgende Beispiel stellt eine Abwandlung des von Mantel und Bross angegebenen Beispiels dar.

Beispiel 3.2. Für zwei verschiedene Fahrzeugtypen A und B werden die Verletzungen des Fahrers bei Crash-Tests mit Dummies in sechs Kategorien eingeteilt: keine, unwesentlich, leicht, mittel, schwer und tödlich. Den Kategorien werden die Zahlen von eins bis sechs zugeordnet. Wenn wir die Verletzungs-Kategorie für den Fahrzeugtyp A als Zufallsvariable X_{11} und die für den Fahrzeugtyp B als Zufallsvariable X_{21} ansehen, so sollten die relativen Effekte p_i Maße für den Unterschied der Sicherheit der beiden Fahrzeugtypen sein. Zwei mögliche Verteilungen für die Zufallsvariablen X_{i1} sind in Tabelle 3.1 angegeben.

Tabelle 3.1: Mögliche Verteilungen der Verletzungskategorien

	Kategorie i					
	1	2	3	4	5	6
$P(X_{11} = i)$	0	1/4	1/4	1/4	1/4	0
$P(X_{21} = i)$	1/4	0	1/4	1/4	0	1/4

Bei diesen beiden Verteilungen erscheint der Fahrzeugtyp A sicherer als der Fahrzeugtyp B zu sein, denn die meisten Personen werden bei der Beurteilung der Sicherheit besonderen Wert auf eine geringe Wahrscheinlichkeit einer tödlichen Verletzung legen. Unterschiedlichen Wahrscheinlichkeiten in den beiden untersten Kategorien wird dagegen normalerweise eine eher nebensächliche Bedeutung beigemessen. Mit anderen Worten: Der Unterschied zwischen einer schweren und einer tödlichen Verletzung wird als größer angesehen als der zwischen einer unwesentlichen und keiner Verletzung. In den relativen Effekten spiegelt sich dieser Unterschied aber nicht wieder. Wegen der Symmetrie der beiden Verteilungen um 7/2 gilt nach Satz 3.3 $p_1 = p_2 = 1/2$. Entsprechend der in diesem Unterabschnitt vorgeschlagenen Interpretation würde das bedeuten, dass keine der beiden Verteilungen zu größeren Werten neigt als die andere. Obwohl die Verletzungskategorien nicht metrisch, sondern nur ordinal sind, scheinen die relativen Effekte hier kein sinnvolles Maß zur Beurteilung des Unterschieds zwischen den Sicherheitseigenschaften der beiden Fahrzeugtypen zu sein. □

Der relative Charakter relativer Effekte

Die Definition der relativen Effekte für zwei unverbundene Stichproben werden wir in den folgenden Abschnitten auf umfangreichere Versuchspläne verallgemeinern. Für nur eine Stichprobe geben wir allerdings keine Definition an. Entsprechend dem Adjektiv 'relativ' in der Bezeichnung 'relativer Effekt' sind relative Effekte Größen, die Verhältnisse von Verteilungen zueinander beschreiben. Liegt nur eine Stichprobe und somit nur eine Verteilung vor, kann kein solches Verhältnis gebildet werden und demzufolge macht der Begriff des relativen Effektes keinen Sinn. Da aber der Vergleich ein zentrales Prinzip der Statistik ist (Freedman et al., 1998, Kapitel 1), bedeutet dies für die statistische Praxis nur eine geringe Einschränkung. Dennoch sollte beachtet werden, dass sich die relativen Effekte grundsätzlich von anderen Funktionalen wie dem Erwartungswert und dem Median unterscheiden, die für einzelne Verteilungen gebildet werden. Während zwei Verteilungen bereits anhand eines einzelnen relativen Effekts verglichen werden können, müssen bei Vergleichen anhand von Erwartungswerten oder Medianen diese Größen erst durch Differenz- oder Quotientenbildung in ein Verhältnis zueinander gesetzt werden.

Die Kritik von Senn (1997)

Wie wir an den Formeln (3.5) und (3.6) erkennen können, hängen relative Effekte bei normalverteilten Daten nicht nur von der Differenz der Erwartungswerte $\mu_1 - \mu_2$, sondern auch von den Varianzen σ_1^2 und σ_2^2 ab. Diese Eigenschaft schließt nach der

Auffassung von Senn (1997) eine sinnvolle Verwendung der relativen Effekte als statistische Funktionale aus: Er ist der Ansicht, dass die Variabilität von Beobachtungen beim Vergleich von zwei Verteilungen in keiner Weise berücksichtigt werden darf. Aus Sicht des Autors der vorliegenden Arbeit ist diese Auffassung in einigen bestimmten Situationen, aber nicht im Allgemeinen berechtigt. Ob die Variabilität berücksichtigt werden sollte, hängt nach Meinung des Autors von der Ursache für die Variabilität ab.

Für die Variabilität von zufälligen Beobachtungen gibt es in der Statistik zwei klassische Begründungen: den zufälligen Messfehler (Freedman et al., 1998, Kapitel 24) und das zufällige Ziehen von Stichproben aus Grundgesamtheiten (Freedman et al., 1998, Teil VI).

Bei zufälligen Messfehlern wird nicht die Versuchseinheit als zufällig angesehen, sondern der Vorgang des Messens an der Versuchseinheit. Die Zufallsvariablen setzen sich additiv aus einem wahren Wert μ_i und einem zufälligen Messfehler ϵ_i mit Erwartungswert 0 zusammen. In dieser Situation ist tatsächlich – solange es nicht um eine Prüfung der Genauigkeit des Messvorgangs geht – ausschließlich der Wert μ_i von Bedeutung. Die Varianz darf deshalb beim Vergleich der Verteilungen der Beobachtungen keine Rolle spielen. In der parametrischen Statistik tragen die Varianzen $\mathrm{Var}(\epsilon_i) = \sigma_i^2$ deshalb in diesem Zusammenhang auch den Namen Störparameter.

Falls aber die Variabilität der Beobachtungen durch das zufällige Ziehen von Stichproben aus verschiedenen Grundgesamtheiten zustande kommt, so spiegelt die Variabilität innerhalb der Stichproben charakteristische Eigenschaften dieser Grundgesamtheiten wieder. Deshalb gibt es in dieser Situation auch keinen Grund, die Berücksichtigung der Variabilität beim Vergleich der Verteilungen der Stichproben auszuschließen.

3.2 Zwei verbundene Stichproben

Wir betrachten nun den einfachen Versuchsplan mit zwei verbundenen Stichproben ($a = 1$ Gruppe, $t = 2$ Bedingungen, alle $m_{1ks} = 1$). Zur Vereinfachung der Schreibweise verwenden wir für diesen Versuchsplan in diesem Abschnitt nicht die allgemeine Notation aus Kapitel 2 sondern

$$\boldsymbol{X}_k = (X_{k1}, X_{k2})' \sim F, \quad k = 1, \ldots, n.$$

Die Randverteilungen der beiden Komponenten bezeichnen wir hier mit F_1 und F_2.

3.2.1 Definition

Als relative Effekte definieren wir für diesen Versuchsplan

$$p_1 = \mathrm{P}(X_{22} < X_{11}) + \frac{1}{2}\,\mathrm{P}(X_{22} = X_{11}), \tag{3.8}$$

$$p_2 = \mathrm{P}(X_{11} < X_{22}) + \frac{1}{2}\,\mathrm{P}(X_{11} = X_{22}). \tag{3.9}$$

Besonderes Augenmerk muss bei dieser Definition darauf gelegt werden, dass die betrachteten Zufallsvariablen X_{11} und X_{22} Komponenten von zwei unabhängigen

Zufallsvektoren sind. Der relative Effekt p_1 ist also die Summe aus der Wahrscheinlichkeit, dass eine Zufallsvariable mit der Verteilungsfunktion F_2 kleiner ist als eine unabhängige Zufallsvariable mit der Verteilungsfunktion F_1, und der Hälfte der Wahrscheinlichkeit, dass zwei solche Zufallsvariablen gleich sind. Eine analoge Aussage gilt für den relativen Effekt p_2. Die relativen Effekte hängen dementsprechend nur von den Randverteilungsfunktionen F_i ab und werden für diesen Versuchsplan deshalb auch als relative *Marginaleffekte* bezeichnet. Wie bei den relativen Effekten für zwei unverbundenen Stichproben können wir auch hier eine Integraldarstellung mit Hilfe der Randverteilungsfunktionen angeben:

$$p_1 = \int F_2 \, dF_1 \quad \text{und} \quad p_2 = \int F_1 \, dF_2.$$

3.2.2 Marginal- und Individualeffekte

Der bekannteste und einfachste Test für zwei verbundene Stichproben ist der Vorzeichentest (siehe z. B. Lehmann, 1975, Abschnitt 4.1). Dieser beruht aber nicht auf den relativen Marginaleffekten, sondern auf Funktionalen q_i, die man als relative *Individualeffekte* bezeichnen könnte:

$$\begin{aligned} q_1 &= \mathrm{P}(X_{12} < X_{11}) + \frac{1}{2}\,\mathrm{P}(X_{12} = X_{11}), \\ q_2 &= \mathrm{P}(X_{11} < X_{12}) + \frac{1}{2}\,\mathrm{P}(X_{11} = X_{12}). \end{aligned}$$

Der relative Individualeffekt q_1 ist also die Summe aus der Wahrscheinlichkeit, dass die zweite Komponente des Zufallsvektors $\boldsymbol{X}_k$ kleiner ist als die erste, und der Hälfte der Wahrscheinlichkeit, dass die beiden Komponenten identisch sind. Entsprechendes gilt für q_2.

Wegen der Ähnlichkeit der relativen Marginal- und Individualeffekte müssen diese Funktionale sorgfältig voneinander unterschieden werden. Während die relativen Marginaleffekte ausschließlich von den Randverteilungsfunktionen F_i abhängen, enthalten die relativen Individualeffekte auch Informationen über die gemeinsame Verteilungsfunktion F der Komponenten der Zufallsvektoren $\boldsymbol{X}_k$. Dies lässt sich mit einem einfachen Beispiel deutlich zeigen.

Beispiel 3.3. Für die Zufallsvektoren $\boldsymbol{X}_k$ gelte

$$\begin{aligned} \mathrm{P}(X_{11} = 100,\, X_{12} = 1) &= \frac{1}{100} \quad \text{und} \\ \mathrm{P}(X_{11} = i,\, X_{12} = i+1) &= \frac{1}{100} \quad \text{für } i = 1, \ldots, 99. \end{aligned}$$

Die Randverteilungsfunktionen F_1 und F_2 stimmen hier überein und deshalb gilt nach Satz 3.3 $p_1 = p_2 = 1/2$. Im Gegensatz dazu liegen die relativen Individualeffekte an den Rändern des Einheitsintervalls: $q_1 = 1/100$ und $q_2 = 99/100$. Obwohl also keine der beiden Randverteilungen zu größeren Werten als die andere tendiert, ist die zweite Komponente der Zufallsvektoren mit 99%iger Wahrscheinlichkeit größer als die erste.

Ohne die Randverteilungen zu verändern, erhalten wir durch Wahl einer anderen gemeinsamen Verteilung eine Umkehrung der relativen Individualeffekte. Im Fall

$$\begin{aligned} \mathrm{P}(X_{11} = 1,\, X_{12} = 100) &= \frac{1}{100} \quad \text{und} \\ \mathrm{P}(X_{11} = i+1,\, X_{12} = i) &= \frac{1}{100} \quad \text{für } i = 1, \ldots, 99 \end{aligned}$$

gilt nämlich $q_1 = 99/100$ und $q_2 = 1/100$. □

3.2.3 Geschichtlicher Hintergund

So wie die Betrachtungen relativer Effekte für zwei unverbundene Stichproben ihren Ursprung im Wilcoxon-Mann-Whitney-Test haben, beruht die Idee der Verwendung von Marginalverteilungen bei zwei verbundenen Stichproben auf den Tests der Hypothese $F_1 = F_2$ von Hollander et al. (1974) und Govindarajulu (1975). Diese stellen Übertragungen des Wilcoxon-Mann-Whitney-Tests auf verbundene Stichproben dar.

Das zu den relativen Marginaleffekten äquivalente Funktional

$$\mathrm{P}(X_{11} < X_{22}) - \mathrm{P}(X_{22} < X_{11}) = 2p_2 - 1 = 1 - 2p_1$$

wurde von Agresti (1983) vorgeschlagen und später unabhängig von Munzel (1999) behandelt.

3.3 Faktorielle Versuchspläne

Das Ziel dieses Abschnittes ist, eine sinnvolle Definition relativer Effekte für die allgemeine Klasse von Versuchsplänen aus Abschnitt 2.1 anzugeben und diese zu diskutieren. Die allgemeine Definition soll die Definitionen aus den beiden vorangehenden Abschnitten als Spezialfälle enthalten und die für die Spezialfälle diskutierten Eigenschaften und Interpretationen sollen dabei weitestmöglich erhalten bleiben.

3.3.1 Definition der relativen Effekte

Für zwei Stichproben sind die relativen Effekte die Summe aus der Wahrscheinlichkeit, dass eine Beobachtung mit einer der beiden Verteilungsfunktionen kleiner ist als eine davon unabhängige Beobachtung mit der anderen Verteilungsfunktion, und der Hälfte der Wahrscheinlichkeit, dass diese beiden Beobachtungen gleich sind. Wenn in faktoriellen Versuchsplänen Beobachtungen mit mehr als zwei verschiedenen Verteilungen vorliegen und wir beschreiben wollen, wie sich die Beobachtungen mit einer der Verteilungsfunktionen im Vergleich zu den Beobachtungen mit den anderen Verteilungsfunktionen verhalten, können wir das arithmetische Mittel aus allen paarweise gebildeten relativen Effekten verwenden. Wir treffen also die folgende

Definition 3.2. Der relative Effekt der i-ten Gruppe unter der s-ten Bedingung ist

$$p_{is} = \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} \left[\mathrm{P}(X_{u2v1} < X_{i1s1}) + \frac{1}{2}\, \mathrm{P}(X_{u2v1} = X_{i1s1}) \right].$$

Bei dieser Definition ist zu beachten, dass sich alle Wahrscheinlichkeiten in der Summe auf unabhängige Zufallsvariablen beziehen. Die Zufallsvariablen X_{u2v1} und X_{i1s1} sind auch für den Fall $u = i$ unabhängig, weil sie von verschiedenen Versuchseinheiten stammen. In der im vorangehenden Abschnitt eingeführten Sprechweise ausgedrückt handelt es sich bei den relativen Effekten p_{is} also um Mittelwerte von relativen Marginaleffekten.

Offensichtlich stimmen nach obiger Definition im Fall zweier unverbundener Stichproben ($a = 2$, $t = 1$) die relativen Effekte p_{i1} mit den in (3.1) und (3.2) definierten Größen p_i überein. Ebenso sind für zwei verbundene Stichproben ($a = 1$, $t = 2$) die relativen Effekte p_{1s} mit den Funktionalen p_s aus den Gleichungen (3.8) und (3.9) identisch.

Als abkürzende Schreibweise fassen wir die lexikographisch nach den Indizes geordneten relativen Effekte in Form eines Vektors

$$\boldsymbol{p} = (p_{11}, \ldots, p_{1t}, \ldots, p_{a1}, \ldots, p_{at})'$$

zusammen.

3.3.2 Eigenschaften und Interpretationen

Integral- und Wahrscheinlichkeitsdarstellungen

Wie bei zwei Stichproben können wir mit Hilfe des Satzes von Fubini Integraldarstellungen für die relativen Effekte p_{is} erreichen. Mit der Vereinbarung

$$H_{is} = \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} F_{uv}$$

erhalten wir

$$\begin{aligned}
p_{is} &= \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} \left[\mathrm{P}(X_{u2v1} < X_{i1s1}) + \frac{1}{2}\,\mathrm{P}(X_{u2v1} = X_{i1s1}) \right] \\
&= \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} \int F_{uv}\, dF_{is} \\
&= \int \left(\frac{1}{at-1} \sum_{(u,v)\neq(i,s)} F_{uv} \right) dF_{is} \\
&= \int H_{is}\, dF_{is}.
\end{aligned}$$

Da H_{is} als arithmetisches Mittel von Verteilungsfunktionen wieder eine Verteilungsfunktion ist, existiert eine Zufallsvariable $Y_{is} \sim H_{is}$. Wenn wir einen Wahrscheinlichkeitsraum konstruieren, auf dem Y_{is} und X_{i1s1} gemeinsam definiert und unabhängig sind, erhalten wir durch erneute Anwendung des Satzes von Fubini wieder eine Darstellung des relativen Effektes p_{is} als Linearkombination von Wahrscheinlichkeiten:

$$p_{is} = \int H_{is}\, dF_{is} = \mathrm{P}(Y_{is} < X_{i1s1}) + \frac{1}{2}\,\mathrm{P}(Y_{is} = X_{i1s1}). \tag{3.10}$$

Aufgrund dieser Darstellung übertragen sich die in den vorangehenden beiden Abschnitten diskutierten Eigenschaften der relativen Effekte für zwei Stichproben in wesentlichen Teilen auf die relativen Effekte für faktorielle Versuchspläne.

Wie die relativen Effekte fassen wir auch die lexikographisch nach den Indizes geordneten Verteilungsfunktionen H_{is} in Form eines Vektors

$$\boldsymbol{H} = (H_{11}, \ldots, H_{1t}, \ldots, H_{a1}, \ldots, H_{at})'$$

zusammen. Daraus ergibt sich als kurze Schreibweise für den Vektor der relativen Effekte

$$\boldsymbol{p} = \int \boldsymbol{H}\, d\boldsymbol{F} = \left(\int H_{11}\, dF_{11}, \ldots, \int H_{at}\, dF_{at} \right)'.$$

Relative Effekte als Wahrscheinlichkeiten

Die wichtigste Eigenschaft der relativen Effekte, die sich vom Zwei-Stichproben-Fall auf faktorielle Versuchspläne überträgt, ist die Interpretierbarkeit auf der Skala der Wahrscheinlichkeiten. Der Fall $p_{is} = 0$ bedeutet, dass die Beobachtungen aus der i-ten Gruppe unter der s-ten Bedingung fast sicher kleiner als alle anderen im Versuch vorkommenden Beobachtungen sind. Gilt $p_{is} = 1$, so sind Beobachtungen mit der Verteilungsfunktion F_{is} fast sicher größer als Beobachtungen mit einer der anderen Verteilungsfunktionen F_{uv}. Für die drei Fälle $p_{is} < 1/2$, $p_{is} = 1/2$ und $p_{is} > 1/2$ können wir die Sprechweisen aus Definition 3.1 übernehmen. Im ersten Fall ist F_{is} tendenziell kleiner als das Mittel der anderen Verteilungsfunktionen H_{is}, im zweiten Fall sind F_{is} und H_{is} tendenziell gleich und im letzten Fall ist F_{is} tendenziell größer als H_{is}. Eine Übersicht darüber, unter welchen Bedingungen welcher der drei Fälle eintritt, gibt der folgende

Satz 3.6. (a) *Falls $F_{is} = H_{is}$ ist oder $X_{iks\ell}$ und $Y_{is} \sim H_{is}$ symmetrisch um das gleiche Symmetriezentrum verteilt sind, so sind F_{is} und H_{is} tendenziell gleich.*

Insbesondere sind F_{is} und H_{is} dann tendenziell gleich, wenn alle Verteilungsfunktionen F_{uv}, $(u,v) \neq (i,s)$, mit F_{is} übereinstimmen oder wenn alle X_{uwvr}, $(u,v) \neq (i,s)$, symmetrisch um das gleiche Symmetriezentrum wie $X_{iks\ell}$ verteilt sind.

(b) *Falls F_{is} und H_{is} symmetrisch sind und das Symmetriezentrum von F_{is} kleiner (größer) ist als das Symmetriezentrum von H_{is}, so ist F_{is} tendenziell kleiner (größer) als H_{is} oder tendenziell gleich H_{is}.*

Insbesondere ist F_{is} dann tendenziell kleiner (größer) als H_{is} oder tendenziell gleich H_{is}, wenn das Symmetriezentrum von F_{is} kleiner (größer) als die Symmetriezentren aller F_{uv} mit $(u,v) \neq (i,s)$ ist.

(c) *Wenn F_{is} stochastisch kleiner (größer) als H_{is} ist, dann ist F_{is} tendenziell kleiner (größer) als H_{is}.*

Insbesondere ist F_{is} dann tendenziell kleiner (größer) als H_{is}, wenn F_{is} stochastisch kleiner (größer) als alle Verteilungsfunktionen F_{uv} mit $(u,v) \neq (i,s)$ ist.

Beweis. Alle Aussagen können mit Hilfe der Integraldarstellung (3.10) auf Satz 3.3 zurückgeführt werden. □

Vergleich von Verteilungen anhand von relativen Effekten

Die relativen Effekte p_{is} dienen nach Definition zunächst einmal zum Vergleich der Verteilungsfunktion F_{is} mit der gemittelten Verteilungsfunktion H_{is}. Es scheint aber naheliegend, anhand der relativen Effekte p_{is} und p_{uv} auch die beiden Verteilungsfunktionen F_{is} und F_{uv} zueinander in Beziehung zu setzen. Die Sinnhaftigkeit eines solchen Vorgehens beruht auf folgendem

Satz 3.7. *Für alle Paare (i, s) gilt*

$$\begin{aligned} p_{is} &= \frac{at}{at-1} \int \overline{F}_{\cdot\cdot}\, dF_{is} - \frac{1}{2(at-1)} \\ &= \frac{at}{at-1} \left[\mathrm{P}(Y < X_{i1s1}) + \frac{1}{2}\, \mathrm{P}(Y = X_{i1s1}) \right] - \frac{1}{2(at-1)}, \end{aligned}$$

wobei Y eine von X_{i1s1} unabhängige Zufallsvariable mit der gemittelten Verteilungsfunktion

$$\overline{F}_{\cdot\cdot} = \frac{1}{at} \sum_{i=1}^{a} \sum_{s=1}^{t} F_{is}$$

ist.

Beweis. Siehe Anhang A.2, Seite 97. □

Mit dem obigen Satz haben wir eine Darstellung der relativen Effekte p_{is} gefunden, die sich für alle Paare (i, s) nur noch auf F_{is} und die einheitliche Referenzverteilung $\overline{F}_{\cdot\cdot}$ an Stelle der verschiedenen H_{is} bezieht. Die Ordnungsbeziehungen zwischen den p_{is} entsprechen genau den Ordnungsbeziehungen zwischen den Integralen $\int \overline{F}_{\cdot\cdot}\, dF_{is}$. Für letztere führen wir in der folgenden Definition eine allgemeine Sprechweise ein.

Definition 3.3. G sei eine beliebige Verteilungsfunktion und $Y \sim G$ sei unabhängig von X_{i1s1} und X_{u1v1}.

(a) F_{is} heißt tendenziell kleiner als F_{uv} mit Bezug auf G, falls

$$\begin{aligned} \int G\, dF_{is} &= \mathrm{P}(Y < X_{i1s1}) + \frac{1}{2}\, \mathrm{P}(Y = X_{i1s1}) \\ < \int G\, dF_{uv} &= \mathrm{P}(Y < X_{u1v1}) + \frac{1}{2}\, \mathrm{P}(Y = X_{u1v1}). \end{aligned}$$

(b) F_{is} und F_{uv} heißen tendenziell gleich mit Bezug auf G, falls

$$\begin{aligned} \int G\, dF_{is} &= \mathrm{P}(Y < X_{i1s1}) + \frac{1}{2}\, \mathrm{P}(Y = X_{i1s1}) \\ = \int G\, dF_{uv} &= \mathrm{P}(Y < X_{u1v1}) + \frac{1}{2}\, \mathrm{P}(Y = X_{u1v1}). \end{aligned}$$

(c) F_{is} heißt tendenziell größer als F_{uv} mit Bezug auf G, falls

$$\begin{aligned} \int G\, dF_{is} &= \mathrm{P}(Y < X_{i1s1}) + \frac{1}{2}\, \mathrm{P}(Y = X_{i1s1}) \\ > \int G\, dF_{uv} &= \mathrm{P}(Y < X_{u1v1}) + \frac{1}{2}\, \mathrm{P}(Y = X_{u1v1}). \end{aligned}$$

Mit der speziellen Wahl $G = \overline{F}_{..}$ und Satz 3.7 erhalten wir

Korollar 3.8. (a) *F_{is} ist genau dann tendenziell kleiner als F_{uv} mit Bezug auf $\overline{F}_{..}$, wenn*

$$p_{is} < p_{uv} \quad \Leftrightarrow \quad \int \overline{F}_{..}\, dF_{is} < \int \overline{F}_{..}\, dF_{uv}.$$

(b) *F_{is} und F_{uv} sind genau dann tendenziell gleich mit Bezug auf $\overline{F}_{..}$, wenn*

$$p_{is} = p_{uv} \quad \Leftrightarrow \quad \int \overline{F}_{..}\, dF_{is} = \int \overline{F}_{..}\, dF_{uv}.$$

(c) *F_{is} ist genau dann tendenziell größer als F_{uv} mit Bezug auf $\overline{F}_{..}$, wenn*

$$p_{is} > p_{uv} \quad \Leftrightarrow \quad \int \overline{F}_{..}\, dF_{is} > \int \overline{F}_{..}\, dF_{uv}.$$

Die Beziehung zwischen den Definitionen 3.1 und 3.3 wird durch folgenden Satz geklärt.

Satz 3.9. (a) *F_{is} ist genau dann tendenziell kleiner als F_{uv}, wenn F_{is} mit Bezug auf $(F_{is} + F_{uv})/2$ tendenziell kleiner als F_{uv} ist.*

(b) *F_{is} und F_{us} sind genau dann tendenziell gleich, wenn F_{is} und F_{uv} mit Bezug auf $(F_{is} + F_{uv})/2$ tendenziell gleich sind.*

(c) *F_{is} ist genau dann tendenziell größer als F_{uv}, wenn F_{is} mit Bezug auf $(F_{is} + F_{uv})/2$ tendenziell größer als F_{uv} ist.*

Beweis. Siehe Anhang A.3, Seite 97. □

Es wäre wünschenswert, wenn die Relationen, die durch die relativen Effekte beschrieben werden (z. B. $p_{is} < p_{uv}$), und die Relationen, die durch die paarweise gebildeten relativen Effekte beschrieben werden (z. B. F_{is} tendenziell kleiner als F_{uv}) widerspruchsfrei wären. Mit anderen Worten: Die tendenziellen Beziehungen zwischen je zwei Verteilungsfunktionen F_{is} und F_{uv} sollten unabhängig davon sein, ob sie mit Bezug auf $\overline{F}_{..}$ oder $(F_{is} + F_{uv})/2$ gebildet werden. Dies trifft aber nicht immer zu. Wie wir sehen werden, sind schon die Relationen, die ausschließlich auf den paarweise gebildeten relativen Effekten beruhen, im Allgemeinen nicht widerspruchsfrei.

Zur Klärung dieses Problems beschäftigen wir uns allgemeiner mit der Frage nach Widersprüchlichkeiten, die zwischen tendenziellen Relationen mit Bezug auf die gemittelte Verteilungsfunktionen

$$\overline{F}_M = \frac{1}{|M|} \sum_{(i,k)\in M} F_{is}, \quad M \subset \{1,\ldots,a\} \times \{1,\ldots,t\},\ |M| \geq 2,$$

auftreten können. Wir führen dazu in Anlehnung an Saari (1989) zunächst die beiden Begriffe *Wort* und *Wörterbuch* ein.

Definition 3.4. $\mathcal{M}$ sei eine endliche Menge und $\mathcal{Q}(\mathcal{M})$ sei die Menge aller Teilmengen von $\mathcal{M}$ mit mindestens 2 Elementen. Ein Wort auf $\mathcal{M}$ ist eine Abbildung, die jedem Element M von $\mathcal{Q}(\mathcal{M})$ eine Rangfolge der Elemente von M zuordnet. Eine Menge von Worten auf $\mathcal{M}$ heißt Wörterbuch auf $\mathcal{M}$.

Beispiel 3.4. Sei $\mathcal{M} = \{1, 2, 3, 4\}$. Dann ist

$$\begin{aligned}\mathcal{Q}(\mathcal{M}) \quad = \quad & \{\{1,2\},\{1,3\},\{1,4\},\{2,3\},\{2,4\},\{3,4\},\\ & \{1,2,3\},\{1,2,4\},\{1,3,4\},\{2,3,4\},\\ & \{1,2,3,4\}\}.\end{aligned}$$

Ein Beispiel für ein Wort auf $\mathcal{M}$ ist

$$\begin{aligned}
\{1,2\} &\mapsto (1 \prec 2)\\
\{1,3\} &\mapsto (1 \prec 3)\\
\{1,4\} &\mapsto (1 \prec 4)\\
\{2,3\} &\mapsto (3 \prec 2)\\
\{2,4\} &\mapsto (4 \prec 2)\\
\{3,4\} &\mapsto (3 \prec 4)\\
\{1,2,3\} &\mapsto (3 \prec 2 = 1)\\
\{1,2,4\} &\mapsto (1 = 2 = 4)\\
\{1,3,4\} &\mapsto (3 = 4 \prec 1)\\
\{2,3,4\} &\mapsto (2 \prec 3 \prec 4)\\
\{1,2,3,4\} &\mapsto (1 \prec 4 = 2 \prec 3).
\end{aligned}$$

Dabei bedeutet zum Beispiel $(3 \prec 2 = 1)$, dass der Teilmenge $\{1, 2, 3\}$ durch das betrachtete Wort die Reihenfolge zugeordnet wird, bei der die 3 vor der 2 und die 2 mit der 1 gleichauf liegt. □

Für jedes (at)-Tupel von Verteilungsfunktionen $(F_{11}, F_{12}, \ldots, F_{at})$ erhalten wir ein Wort auf $\mathcal{M} = \{1, \ldots, a\} \times \{1, \ldots, t\}$, indem wir jeder Teilmenge M von $\mathcal{M}$ mit mindestens zwei Elementen die Rangfolge zuordnen, die sich aus den tendenziellen Relationen mit Bezug auf $\overline{F}_M$ ergeben. Die Gesamtheit aller dieser Wörter, die sich für alle (at)-Tupel von Verteilungsfunktionen $(F_{11}, F_{12}, \ldots, F_{at})$ ergibt, nennen wir das Wörterbuch der tendenziellen Relationen auf $\mathcal{M}$. Dieses Wörterbuch gibt Auskunft über alle Widersprüchlichkeiten, die bei tendenziellen Relationen zwischen Verteilungsfunktionen mit Bezug auf die $\overline{F}_M$ auftreten können.

Ein anderes Wörterbuch ist das sogenannte *Borda-Wörterbuch* auf einer Menge von Kandidaten $\{C_1, \ldots, C_d\}$, das in Saari (1990) charakterisiert ist. Wir können das Wörterbuch der tendenziellen Relationen auf $\{1, \ldots, a\} \times \{1, \ldots, t\}$ mit dem Borda-Wörterbuch auf $\{C_1, \ldots, C_{at}\}$ vergleichen, wenn wir die beiden Mengen $\{1, \ldots, a\} \times \{1, \ldots, t\}$ und $\{C_1, \ldots, C_{at}\}$ vermöge der Bijektion $(i, s) \mapsto C_{is}$ identifizieren. In diesem Sinne ist der folgende Satz zu verstehen.

Satz 3.10. *Für alle* $(a,t) \in \mathbb{N}^2$ *enthält das Wörterbuch der tendenziellen Relationen auf* $\{1,\ldots,a\} \times \{1,\ldots,t\}$ *das Borda-Wörterbuch auf* $\{C_1,\ldots,C_{at}\}$.

Beweis. Siehe Anhang A.4, Seite 98. □

Als Anwendung dieses Satzes geben wir die Übertragungen einiger Widersprüchlichkeiten aus dem Borda-Wörterbuch, die in Haunsperger (1992) genannt sind, in das Wörterbuch der tendenziellen Relationen an.

Beispiel 3.5. (a) Es existieren Verteilungsfunktionen $F_{11},\ldots,F_{at}$, so dass das zugehörige Wort im Wörterbuch der tendenziellen Relationen eingeschränkt auf die Menge

$$\mathcal{P} = \{\{(1,1),(1,2)\},\{(1,2),(1,3)\},\ldots,\{(a,t-1),(a,t)\},\{(a,t),(1,1)\}\}$$

die folgende Gestalt hat:

$$\begin{array}{rcl}
\{(1,1),(1,2)\} & \mapsto & ((1,1) \prec (1,2)), \\
\{(1,2),(1,3)\} & \mapsto & ((1,2) \prec (1,3)), \\
 & \ldots & \\
\{(a,t-1),(a,t)\} & \mapsto & ((a,t-1) \prec (a,t)), \\
\{(a,t),(1,1)\} & \mapsto & ((a,t) \prec (1,1)).
\end{array}$$

Die tendenziellen Relationen zwischen den Verteilungsfunktionen F_{is} bilden bei diesem Beispiel einen sogenannten *Condorcet-Kreis*: F_{11} ist tendenziell kleiner als F_{12}, F_{12} ist tendenziell kleiner als F_{13} und jede weitere Verteilungsfunktion ist tendenziell kleiner als ihr Nachfolger. Der Kreis wird dadurch geschlossen, dass F_{at} tendenziell kleiner als F_{11} ist.

(b) Es existieren Verteilungsfunktionen $F_{11},\ldots,F_{at}$, so dass das zugehörige Wort im Wörterbuch der tendenziellen Relationen eingeschränkt auf die Menge $\mathcal{N}$ der verschachtelten Teilmengen

$$\begin{array}{rcl}
N_2 & = & \{(1,1),(1,2)\}, \\
N_3 & = & \{(1,1),(1,2),(1,3)\}, \\
 & \ldots & \\
N_t & = & \{(1,1),\ldots,(1,t)\} \\
N_{t+1} & = & \{(1,1),\ldots,(1,t),(a,1)\} \\
 & \ldots, & \\
N_{at} & = & \{(1,1),\ldots,(a,t)\}
\end{array}$$

die folgende Form hat:

$$\begin{array}{rcl}
N_2 & \mapsto & ((1,1) \prec (1,2)) \\
N_3 & \mapsto & ((1,3) \prec (1,2) \prec (1,1)) \\
N_4 & \mapsto & ((1,1) \prec (1,2) \prec (1,3) \prec (1,4)) \\
 & \text{usw.} &
\end{array}$$

Für die Verteilungsfunktionen F_{is} heißt das:

- Mit Bezug auf $\overline{F}_{N_2}$ ist F_{11} tendenziell kleiner als F_{12}.
- Mit Bezug auf $\overline{F}_{N_3}$ ist F_{11} tendenziell größer als F_{12} und F_{12} tendenziell größer als F_{13}.
- Mit Bezug auf $\overline{F}_{N_4}$ ist F_{11} tendenziell kleiner als F_{12}, F_{12} tendenziell kleiner als F_{13} und F_{13} tendenziell kleiner als F_{14}.

Je nachdem ob N_u eine gerade oder ungerade Anzahl von Elementen hat, sind die F_{is} mit Bezug auf $\overline{F}_{N_u}$ also tendenziell lexikographisch oder umgekehrt lexikographisch nach den Indexpaaren (i,s) geordnet.

(c) Es existieren Verteilungsfunktionen $F_{11},\ldots,F_{at}$, so dass das zugehörige Wort im Wörterbuch der tendenziellen Relationen eingeschränkt auf die Menge aller Mengen mit zwei Elementen und $\{1,\ldots,a\}\times\{1,\ldots,t\}$ die folgende Form hat: Jede Menge mit zwei Elementen $\{(i,s),(u,v)\}$ wird genau dann auf $((i,s)\prec(u,v))$ abgebildet, wenn (i,s) lexikographisch kleiner als (u,v) ist. Das Bild von $\{1,\ldots,a\}\times\{1,\ldots,t\}$ dagegen hat die Gestalt:

$$((1,1)\prec(a,t)\prec(a,t-1)\prec(a,t-2)\prec\cdots\prec(1,2)).$$

Für diese Verteilungsfunktionen ist also F_{is} genau dann tendenziell kleiner als F_{uv}, wenn (i,s) lexikographisch kleiner als (u,v) ist. Mit Bezug auf $\overline{F}_{..}$ ergibt sich jedoch ein ganz anderes Bild:

$$p_{11}<p_{at}<p_{a(t-1)}<p_{a(t-2)}<\cdots<p_{12}.$$

Alle F_{is} mit Ausnahme von F_{11} sind also mit Bezug auf $\overline{F}_{..}$ tendenziell in der umgekehrten Reihenfolge geordnet, wie es nach Betrachtung der tendenziellen Ordnungen mit Bezug auf die paarweise gebildeten Mittel $\overline{F}_{\{(i,s),(u,v)\}}$ zu erwarten wäre. □

Wie wir gesehen haben, besteht zwischen den Aussagen über das tendenzielle Verhältnis von Verteilungen zueinander – mit und ohne Bezug auf $\overline{F}_{..}$ – keine allgemein gültige einfache Beziehung. Aber wenigstens für den Fall, dass die Verteilungsfunktionen stochastisch geordnet sind, dass heißt, wenn für je zwei Verteilungsfunktionen F_{is} und F_{uv} entweder $F_{is}\leq F_{uv}$ oder $F_{is}\geq F_{uv}$ gilt, können wir eine solche Beziehung nachweisen. Dieser Fall umfasst insbesondere Lokationsmodelle.

Satz 3.11. *Die Verteilungsfunktionen F_{is}, $i=1,\ldots,a$, $s=1,\ldots,t$, seien stochastisch geordnet. Dann gelten folgende Äquivalenzen:*

(a) *F_{is} ist genau dann tendenziell kleiner als F_{uv}, wenn F_{is} mit Bezug auf $\overline{F}_{..}$ tendenziell kleiner als F_{uv} ist.*

(b) *F_{is} und F_{uv} sind genau dann tendenziell gleich, wenn F_{is} und F_{uv} mit Bezug auf $\overline{F}_{..}$ tendenziell gleich sind.*

(c) *F_{is} ist genau dann tendenziell größer als F_{uv}, wenn F_{is} mit Bezug auf $\overline{F}_{..}$ tendenziell größer als F_{uv} ist.*

Beweis. Siehe Anhang A.5, Seite 98. □

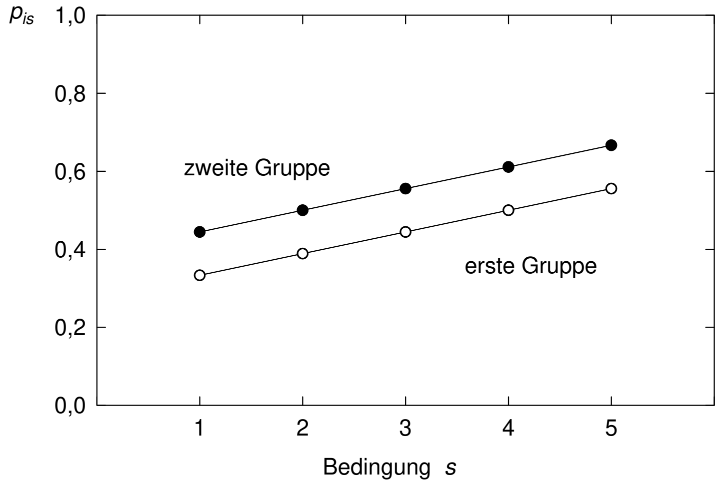

Abbildung 3.1: Relative Effekte p_{is} für die Bernoulli-Verteilungen in Beispiel 3.6.

Beispiele und graphische Darstellungen

Da für jede Verteilungsfunktion F_{is} ein relativer Effekt p_{is} definiert ist, eignen sich die relativen Effekte in ähnlicher Weise wie Erwartungswerte oder Mediane zur graphischen Darstellung. Anhand einer solchen Darstellung kann man sich – auch wenn man bei der Interpretation vorsichtig sein muss – schnell einen Überblick über die Verhältnisse der Verteilungen zueinander verschaffen. Dies soll hier für zwei einfache Beispiele exemplarisch gezeigt werden.

Beispiel 3.6. Die Beobachtungen $X_{iks\ell}$ seien Bernoulli-verteilt mit Erfolgswahrscheinlichkeit q_{is}. Nach Definition der relativen Effekte p_{is} und Formel (3.3) gilt

$$\begin{aligned} p_{is} &= \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} \left[\mathrm{P}(X_{u2v1} < X_{i1s1}) + \frac{1}{2}\, \mathrm{P}(X_{u2v1} = X_{i1s1}) \right] \\ &= \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} \left(\frac{q_{is} - q_{uv}}{2} + \frac{1}{2} \right). \end{aligned}$$

Die relativen Effekte sind also äquivalent zu den mittleren Differenzen zwischen den Erfolgswahrscheinlichkeiten q_{is} und q_{uv}, $(u,v) \neq (i,s)$.

Für einen Versuchsplan mit $a = 2$ Gruppen und $t = 5$ Bedingungen sind in Abbildung 3.1 die relativen Effekte dargestellt, die sich aus den Erfolgswahrscheinlichkeiten

$$q_{1s} = \frac{s+1}{10} \quad \text{und} \quad q_{2s} = \frac{s+3}{10}, \quad s = 1, \ldots, 5,$$

ergeben. □

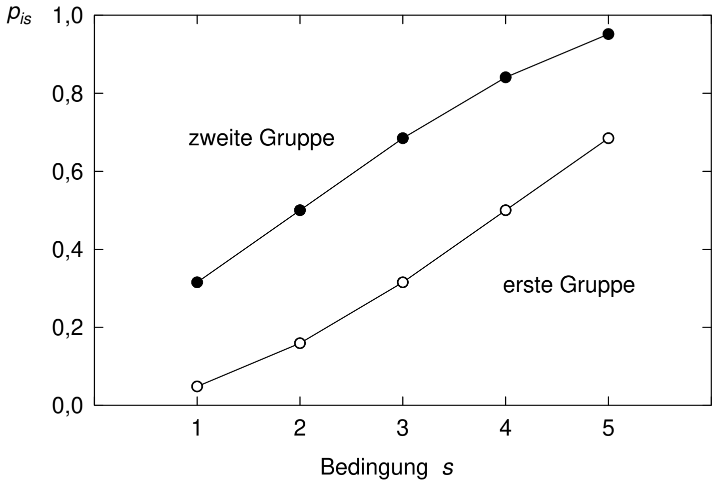

Abbildung 3.2: Relative Effekte p_{is} für die Normalverteilungen in Beispiel 3.7.

Beispiel 3.7. Die Beobachtungen $X_{iks\ell}$ seien normalverteilt mit Erwartungswert μ_{is} und Varianz σ^2_{is}. Nach Definition der relativen Effekte p_{is} und Formel (3.5) gilt

$$\begin{aligned} p_{is} &= \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} \left[\mathrm{P}(X_{u2v1} < X_{i1s1}) + \frac{1}{2}\, \mathrm{P}(X_{u2v1} = X_{i1s1}) \right] \\ &= \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} \Phi\left(\frac{\mu_{is} - \mu_{uv}}{\sqrt{\sigma^2_{is} + \sigma^2_{uv}}} \right). \end{aligned}$$

Die relativen Effekte sind also nicht äquivalent zu den mittleren standardisierten Differenzen zwischen den Erwartungswerten μ_{is} und μ_{uv}, $(u,v) \neq (i,s)$.

Für einen Versuchsplan mit $a = 2$ Gruppen und $t = 5$ Bedingungen sind in Abbildung 3.2 die relativen Effekte dargestellt, die sich aus den Erwartungswerten

$$\mu_{1s} = s \quad \text{und} \quad \mu_{2s} = s + 2, \quad s = 1, \ldots, 5,$$

und Varianzen $\sigma^2_{is} = 1$, $i = 1, 2$, $s = 1, \ldots, 5$, ergeben. □

3.3.3 Bisherige Verallgemeinerungsansätze

Die in diesem Abschnitt getroffene Definition der relativen Effekte in faktoriellen Versuchsplänen ist nicht die einzige Verallgemeinerung der relativen Effekte für Situationen mit mehr als zwei Stichproben. Wir stellen in diesem Unterabschnitt einige andere Verallgemeinerungsansätze aus der Literatur vor.

Einer dieser Ansätze beruht auf der Idee, eine der Verteilungen, zum Beispiel die Verteilung F_{is}, auszuwählen und als Referenz zu verwenden. Alle anderen Verteilungen können dann jeweils anhand der paarweise gebildeteten relativen Effekte mit der Referenzverteilung verglichen werden. Beim Vergleich der Referenzverteilung mit der Verteilung F_{uv} erhalten wir so die Effekte

$$\mathrm{P}(X_{i1s1} < X_{u2v1}) + \frac{1}{2}\, \mathrm{P}(X_{i1s1} = X_{u2v1}), \quad (u,v) \neq (i,s).$$

Diese Größen kann man zum Beispiel in einer graphischen Darstellung verwenden, um sich einen Überblick über das Verhältnis aller Verteilungen zur Referenzverteilung zu verschaffen. Der Vorschlag zur Bildung dieser Funktionale geht auf Bross (1958) zurück, der aber auch schon bemerkt hat, dass es wichtig ist, eine Referenzverteilung auszuwählen, die ein möglichst breites Spektrum an Werten abdeckt (Bross, 1958, Abschnitt 3). Falls das nicht möglich ist, sind die oben definierten Größen eventuell nur dazu geeignet, die Verhältnisse der Referenzverteilung zu den anderen Verteilungen zu beschreiben, nicht aber dazu, die Verhältnisse der anderen Verteilungen untereinander darzustellen. Das folgende überspitzte Beispiel verdeutlicht dieses Problem.

Beispiel 3.8. Es sei $a = 10$ und $t = 1$. Die Zufallsvariablen $X_{ik1\ell}$ seien gleichverteilt auf den offenen Intervallen $(i-1, i)$. Unabhängig von der Wahl der Referenzverteilungsfunktion F_{i1} gilt für alle $u \neq i$

$$\begin{aligned} \mathrm{P}(X_{i111} < X_{u111}) + \frac{1}{2}\,\mathrm{P}(X_{i111} = X_{u111}) &= 0 \quad \text{für } u < i, \\ \mathrm{P}(X_{i111} < X_{u111}) + \frac{1}{2}\,\mathrm{P}(X_{i111} = X_{u111}) &= 1 \quad \text{für } u > i. \end{aligned}$$

Obwohl die Träger aller Verteilungen disjunkt sind, gilt für alle Verteilungsfunktionen F_{u1} und $F_{u'1}$ mit $\min\{u, u'\} > i$

$$\begin{aligned} 1 &= \mathrm{P}(X_{i111} < X_{u111}) + \frac{1}{2}\,\mathrm{P}(X_{i111} = X_{u111}), \\ &= \mathrm{P}(X_{i111} < X_{u'111}) + \frac{1}{2}\,\mathrm{P}(X_{i111} = X_{u'111}). \end{aligned}$$

Die Verteilungen von X_{u111} und $X_{u'111}$ können trotz der disjunkten Träger durch Betrachtung der von Bross definierten Größen nicht auseinandergehalten werden.

Im Gegensatz dazu können wir bei Verwendung der relativen Effekte aus Definition 3.2 alle Verteilungen anhand der Werte

$$p_{i1} = \frac{1}{9}(i-1), \quad i = 1, \ldots, 10,$$

voneinander unterscheiden. Der Grund für die größere Trennschärfe der relativen Effekte p_{i1} liegt darin, dass jede der Verteilungsfunktionen H_{i1} ein größeres Spektrum an Werten abdeckt als jede nach Bross in Frage kommende Referenzverteilungsfunktion F_{i1}. □

Ein anderer Verallgemeinerungsvorschlag für faktorielle Versuchspläne stammt von Marden und Muyot (1995). Statt – wie von Bross vorgeschlagen – alle Verteilungen paarweise mit nur einer Referenzverteilung zu vergleichen, bilden Marden und Muyot alle möglichen Paare von Verteilungen. Sie betrachten also für alle (i, s) und alle (u, v) die paarweise gebildeten relativen Effekte

$$\mathrm{P}(X_{u2v1} < X_{i1s1}) + \frac{1}{2}\,\mathrm{P}(X_{u2v1} = X_{i1s1}).$$

Im Gegensatz zu der von Bross vorgeschlagenen Methode tritt hierbei das Problem auf, dass schon bei relativ wenigen Gruppen und Bedingungen eine so große Anzahl von Paaren gebildet wird, dass diese kaum noch dazu geeignet sind, sich einen Überblick über die Verhältnisse der Verteilungen zueinander zu verschaffen. Bei einem Versuchsplan mit $a = 4$ Gruppen und $t = 5$ Bedingungen gibt es 20 Verteilungsfunktionen F_{is} und deshalb 400 Paare von Verteilungsfunktionen. Eine solch große Anzahl von paarweise gebildeten relativen Effekten lässt sich auch mit Hilfe grafischer Darstellungen kaum noch überblicken. Aus diesem Grund widerspricht die von Marden und Muyot vorgeschlagene Strategie dem statistischen Grundprinzip der Datenreduktion und wird schon bei wenig umfangreichen Versuchsplänen unbrauchbar.

Neben den von Bross und Marden und Muyot vorgeschlagenen Verallgemeinerungsstrategien, die für die allgemeine Klasse von faktoriellen Versuchsplänen aus Abschnitt 2.1 durchführbar sind, gibt es auch Vorschläge für die Verwendung paarweise gebildeter relativer Effekte in speziellen Versuchsplänen. So beschäftigen sich Patel und Hoel (1973) anhand der Differenz zweier relativer Effekte mit der Wechselwirkung in einem 2×2-Versuchsplan mit unabhängigen Beobachtungen. Für den Versuchsplan mit $a = 2$ Gruppen und t Bedingungen vergleichen Thall und Lachin (1988) und Lachin (1992, 1996) die beiden Gruppen anhand der für die t Bedingungen paarweise gebildeten relativen Effekte.

3.3.4 Relative Gruppen- und Bedingungseffekte

Nach Definition dient ein relativer Effekt p_{is} zur Beschreibung des Verhältnisses der Verteilungsfunktion in der i-ten Gruppe unter der s-ten Bedingung zu den anderen Verteilungsfunktionen F_{uv} mit $(u, v) \neq (i, s)$. In bestimmten Situationen kann es aber auch von Interesse sein, mit Hilfe von relativen Effekten das Verhältnis der Verteilungsfunktionen in der i-ten Gruppe zu den Verteilungsfunktionen der anderen Gruppen oder das Verhältnis der Verteilungsfunktionen unter der s-ten Bedingung zu den Verteilungsfunktionen unter den anderen Bedingungen zu untersuchen. In der Panik-Skala-Studie II (siehe Unterabschnitt 2.1.1, Seite 9) könnte man zum Beispiel daran interessiert sein, die Entwicklung der Panikstörungen unter den unterschiedlichen Bedingungen (= Zeitpunkten) zu erfassen, ohne zu berücksichtigen, ob die Patienten aus der Agoraphobie-Gruppe stammen oder nicht.

Zu diesem Zweck definieren wir die sogenannten relativen *Gruppen- und Bedingungseffekte.*

Definition 3.5. (a) Sei $\overline{F}_{i\cdot}$ das Mittel der Verteilungsfunktionen aus der i-ten Gruppe und H_{i*} das Mittel aller Verteilungsfunktionen, die nicht aus der i-ten Gruppe stammen, das heißt,

$$\overline{F}_{i\cdot} = \frac{1}{t}\sum_{s=1}^{t} F_{is} \quad \text{und} \quad H_{i*} = \frac{1}{at-t}\sum_{u\neq i}\sum_{v=1}^{t} F_{uv}.$$

Die Zufallsvariablen $X_{i*} \sim \overline{F}_{i\cdot}$ und $Y_{i*} \sim H_{i*}$ seien unabhängig. Dann heißt

$$p_{i*} = \mathrm{P}(Y_{i*} < X_{i*}) + \frac{1}{2}\mathrm{P}(Y_{i*} = X_{i*})$$

relativer Effekt der i-ten Gruppe.

(b) Sei $\overline{F}_{\cdot s}$ das Mittel der Verteilungsfunktionen unter der s-ten Bedingung und H_{*s} das Mittel aller Verteilungsfunktionen, die nicht zu der s-ten Bedingung gehören, das heißt,

$$\overline{F}_{\cdot s} = \frac{1}{a}\sum_{i=1}^{a} F_{is} \quad \text{und} \quad H_{*s} = \frac{1}{at-a}\sum_{u=1}^{a}\sum_{v \neq s}^{t} F_{uv}.$$

Die Zufallsvariablen $X_{*s} \sim \overline{F}_{\cdot s}$ und $Y_{*s} \sim H_{*s}$ seien unabhängig. Dann heißt

$$p_{*s} = \mathrm{P}(Y_{*s} < X_{*s}) + \frac{1}{2}\,\mathrm{P}(Y_{*s} = X_{*s})$$

relativer Effekt der s-ten Bedingung.

Wie für die relativen Effekte p_{is} existieren auch für die Gruppen- und Bedingungseffekte Integraldarstellungen. Sie haben die Form

$$p_{i*} = \int H_{i*}\, d\overline{F}_{i\cdot} \quad \text{und} \quad p_{*s} = \int H_{*s}\, d\overline{F}_{\cdot s}.$$

Zu beachten ist bei der obigen Definition, dass die relativen Gruppen- und Bedingungseffekte nicht mit den entsprechenden arithmetischen Mitteln der relativen Effekte übereinstimmen:

$$p_{i*} \neq \overline{p}_{i\cdot} = \frac{1}{t}\sum_{s=1}^{t} p_{is} \quad \text{und} \quad p_{*s} \neq \overline{p}_{\cdot s} = \frac{1}{a}\sum_{i=1}^{a} p_{is}.$$

Der Grund dafür, dass wir in Definition 3.5 von den arithmetischen Mittelwerten der relativen Effekte abweichen, liegt in den unterschiedlichen Wertebereichen der p_{i*} und p_{*s} auf der einen und der $\overline{p}_{i\cdot}$ und $\overline{p}_{\cdot s}$ auf der anderen Seite. Für die relativen Gruppen- und Bedingungseffekte sind die Wertebereiche jeweils das gesamte Einheitsintervall. Die Grenzen dieses Intervalls bilden bei der Interpretation der Effekte hilfreiche Orientierungspunkte: $p_{i*} = 0$ gilt genau dann, wenn jede Beobachtung aus der i-ten Gruppe fast sicher kleiner ist als jede Beobachtung aus allen anderen Gruppen. Entsprechend gilt $p_{i*} = 1$ genau dann, wenn jede Beobachtung aus der i-ten Gruppe fast sicher größer ist als jede Beobachtung aus den anderen Gruppen. Analoge Aussagen gelten für die Bedingungseffekte.

Die Mittelwerte der relativen Effekte $\overline{p}_{i\cdot}$ und $\overline{p}_{\cdot s}$ liegen als Mittelwerte von Wahrscheinlichkeiten zwar auch immer im Einheitsintervall. Aber die Grenzen 0 und 1 sind mit Ausnahme der Sonderfälle $a = 1$ oder $t = 1$ als Werte ausgeschlossen.

Satz 3.12. *Für die gemittelten relativen Effekte gilt*

$$\overline{p}_{i\cdot} \in \left[\frac{t-1}{2(at-1)}, 1 - \frac{t-1}{2(at-1)}\right] \quad \textit{und} \quad \overline{p}_{\cdot s} \in \left[\frac{a-1}{2(at-1)}, 1 - \frac{a-1}{2(at-1)}\right].$$

Beweis. Siehe Anhang A.6, Seite 99. □

Die Wertebereiche der $\overline{p}_{i\cdot}$ und $\overline{p}_{\cdot s}$ hängen also im Gegensatz zu denen der relativen Gruppen- und Bedingungseffekte von der Anzahl der Gruppen a und der Anzahl der Bedingungen t ab. Um zu beurteilen, ob die gemittelten Effekte auf den Rändern oder in der Nähe der Ränder liegen, müssen diese erst ermittelt werden. Deshalb ist es bequemer, mit den relativen Gruppen- und Bedingungseffekten zu arbeiten.

Bei der Konstruktion von Schätzern für die relativen Gruppen- und Bedingungseffekte in den nächsten Kapiteln wird die folgende Darstellung mit Hilfe des Vektors $\boldsymbol{p} = (p_{11}, \ldots, p_{at})'$ von Nutzen sein.

Satz 3.13. *Für die (at)-dimensionalen Spaltenvektoren $\boldsymbol{\nu}_{i*}$ und $\boldsymbol{\nu}_{*s}$, deren Komponenten durch*

$$(\boldsymbol{\nu}_{i*})_{uv} = \begin{cases} \frac{at-1}{at^2-t^2} & : u = i, \\ 0 & : u \neq i, \end{cases} \quad \textit{und} \quad (\boldsymbol{\nu}_{*s})_{uv} = \begin{cases} \frac{at-1}{a^2t-a^2} & : v = s, \\ 0 & : v \neq s, \end{cases}$$

definiert sind, gilt

$$p_{i*} = \boldsymbol{\nu}'_{i*}\boldsymbol{p} - \frac{t-1}{2(at-t)} \quad \textit{und} \quad p_{*s} = \boldsymbol{\nu}'_{*s}\boldsymbol{p} - \frac{a-1}{2(at-a)}.$$

Beweis. Siehe Anhang A.7, Seite 100. □

Kapitel 4

Punktschätzer

Im vorangehenden Kapitel haben wir diskutiert, inwieweit relative Effekte als Funktionale dazu geeignet sind, die Verteilungen in einem faktoriellen Versuchsplan sinnvoll zusammenfassend zu beschreiben. Dies alleine ist aber noch nicht ausreichend dafür, dass relative Effekte in der Statistik nützlich sind. Es muss darüber hinaus auch möglich sein, auf Grundlage der aus dem Versuch vorliegenden Beobachtungen Rückschlüsse auf die relativen Effekte zu ziehen. Die einfachste Methode dazu ist die Berechnung von Punktschätzern, die der Gegenstand dieses Kapitels sind.

Bevor wir in diesem Kapitel mit der Konstruktion von Punktschätzern beginnen, geben wir eine kurze Übersicht über die Literatur, die sich mit Punktschätzern für relative Effekte im Fall zweier unabhängiger Stichproben beschäftigt. Die Idee zur Konstruktion allgemeinerer Punktschätzer wird im zweiten Abschnitt dieses Kapitels behandelt und beruht auf der Darstellung (3.10) der relativen Effekte mit Hilfe der Randverteilungsfunktionen. Aus diesem Grund beschäftigen wir uns im dritten Abschnitt mit Schätzern für die Randverteilungsfunktionen, bevor wir die Eigenschaften der sich ergebenden Schätzer für die relativen Effekte diskutieren. Abschließend geben wir noch Hinweise zur praktischen Berechnung der Punktschätzer und führen diese beispielhaft für die in Unterabschnitt 2.1.1 genannten Versuche durch.

4.1 Bisherige Arbeiten zu zwei unverbundenen Stichproben

Da die relativen Effekte für zwei unverbundene Stichproben, wie sie in Abschnitt 3.1 dargestellt sind, in der statistischen Fachliteratur seit der Arbeit von Birnbaum (1956) immer wieder diskutiert wurden, existiert auch eine beachtliche Anzahl von Veröffentlichungen, die sich mit Punktschätzern dazu beschäftigen.

Für das hier behandelte nichtparametrische Modell sind der bereits von Birnbaum verwendete Schätzer sowie dessen Modifikation für unstetige Verteilungsfunktionen von Bross nach Kenntnis des Autors jedoch die einzigen Vorschläge. Wir gehen hier auf diese Schätzer aber nicht näher ein, da sie Spezialfälle der in den noch folgenden Teilen dieses Kapitels behandelten Schätzer für die allgemeine Klasse von faktoriellen Versuchsplänen aus Kapitel 2 sind. Erwähnt sei aber noch die

Vorlesung von Lehmann (1950), in der – wenn auch mehr als Beispiel für einen mathematischen Sachverhalt – Optimalitätseigenschaften des Schätzers für den Fall stetiger Verteilungsfunktionen nachgewiesen werden.

Für verschiedene parametrische Situationen liegt aber eine Vielzahl von Arbeiten mit unterschiedlichen Vorschlägen zur Schätzung der relativen Effekte vor. Der interessierte Leser sei für normalverteilte Beobachtungen auf die Veröffentlichungen von Owen et al. (1964), Govindarajulu (1967), Church und Harris (1970), Mazumdar (1970), Downton (1973), Reiser und Guttman (1986) und Simonoff et al. (1986), für exponentialverteilte Beobachtungen auf Tong (1974), Johnson (1975), Tong (1975a), Kelley et al. (1976) und Sathe und Shah (1981) sowie für gammaverteilte Beobachtungen auf Tong (1975b) und Woodward und Gray (1975) verwiesen.

4.2 Idee für die Schätzer: Das Einsetzungsprinzip

In der Gleichung (3.10) auf Seite 31 haben wir eine Darstellung der relativen Effekte mit Hilfe der Verteilungsfunktionen F_{is} und H_{is} hergeleitet:

$$p_{is} = \int H_{is}\, dF_{is}.$$

Da sich Verteilungsfunktionen in unserer Situation gut schätzen lassen, worauf wir im folgenden Abschnitt genauer eingehen werden, ist es naheliegend, als Schätzer den Ausdruck zu verwenden, der sich ergibt, wenn wir in obiger Gleichung die Verteilungsfunktionen durch ihre Schätzer ersetzen. Dieses Vorgehen ist in der Literatur auch als Einsetzungsprinzip bekannt. Mit den im nächsten Abschnitt hergeleiteten Schätzern $\widehat{F}_{is,n}$ und $\widehat{H}_{is,n}$ für F_{is} und H_{is} erhalten wir als Schätzer für p_{is}

$$\widehat{p}_{is,n} = \int \widehat{H}_{is,n}\, d\widehat{F}_{is,n}.$$

Als bequeme Schreibweise für die Vektoren der Schätzer benutzen wir analog zu den Symbolen $\boldsymbol{p}, \boldsymbol{F}$ und $\boldsymbol{H}$

$$\begin{aligned}
\widehat{\boldsymbol{p}}_n &= (\widehat{p}_{11,n}, \ldots, \widehat{p}_{1t,n}, \ldots, \widehat{p}_{a1,n}, \ldots, \widehat{p}_{at,n})', \\
\widehat{\boldsymbol{F}}_n &= \left(\widehat{F}_{11,n}, \ldots, \widehat{F}_{1t,n}, \ldots, \widehat{F}_{a1,n}, \ldots, \widehat{F}_{at,n}\right)' \text{ und} \\
\widehat{\boldsymbol{H}}_n &= \left(\widehat{H}_{11,n}, \ldots, \widehat{H}_{1t,n}, \ldots, \widehat{H}_{a1,n}, \ldots, \widehat{H}_{at,n}\right)'.
\end{aligned}$$

Für den Vektor der Schätzer der relativen Effekte können wir damit auch

$$\widehat{\boldsymbol{p}}_n = \int \widehat{\boldsymbol{H}}_n\, d\widehat{\boldsymbol{F}}_n = \left(\int \widehat{H}_{11,n}\, d\widehat{F}_{11,n}, \ldots, \int \widehat{H}_{at,n}\, d\widehat{F}_{at,n}\right)'$$

schreiben.

Für die Gruppen- und Bedingungseffekte p_{i*} und p_{*s} verwenden wir die wegen der Darstellungen

$$p_{i*} = \boldsymbol{\nu}'_{i*}\boldsymbol{p} - \frac{t-1}{2(a-1)t} \quad \text{und} \quad p_{*s} = \boldsymbol{\nu}'_{*s}\boldsymbol{p} - \frac{a-1}{2a(t-1)}$$

aus Satz 3.13 (siehe Seite 43) naheliegenden Schätzer

$$\widehat{p}_{i*,n} = \boldsymbol{\nu}'_{i*}\widehat{\boldsymbol{p}}_n - \frac{t-1}{2(a-1)t} \quad \text{und} \quad \widehat{p}_{*s,n} = \boldsymbol{\nu}'_{*s}\widehat{\boldsymbol{p}}_n - \frac{a-1}{2a(t-1)}.$$

4.3 Empirische Verteilungsfunktionen

In diesem Abschnitt leiten wir die benötigten Schätzer für die Verteilungsfunktionen her. Diese Schätzer stimmen mit den Schätzern überein, die in Brunner et al. (1999) für die gleiche Klasse von Versuchsplänen angegeben sind.

Zunächst definieren wir punktweise einen erwartungstreuen Schätzer für F_{is}, der ausschließlich auf der Zufallsvariablen X_{iksr} beruht:

$$\widehat{F}_{iksr}(x) = \mathrm{c}(x - X_{iksr}).$$

Im nächsten Schritt geben wir einen Schätzer für F_{is} an, der die Informationen aus allen Komponenten von $\boldsymbol{X}_{iks}$ ausnutzt. Da die gemeinsame Verteilung dieser Komponenten durch unser Modell nicht näher bestimmt ist und die naheliegendste diesbezügliche Annahme die Austauschbarkeit ist, scheint es natürlich, den Schätzer $\widehat{F}_{iks}$ als ungewichtetes Mittel der $\widehat{F}_{iksr}$ zu wählen:

$$\widehat{F}_{iks} = \begin{cases} \frac{1}{m_{iks}} \sum_{r=1}^{m_{iks}} \widehat{F}_{iksr} & : m_{iks} > 0, \\ 0 & : m_{iks} = 0. \end{cases}$$

Dieser Schätzer ist wieder erwartungstreu, und wir können vermuten, dass er eine umso kleinere Varianz hat, je größer m_{iks} ist. Wenn wir also schließlich $F_{is}(x)$ durch ein gewichtetes Mittel der unabhängigen Zufallsvariablen $\widehat{F}_{iks}(x)$, $k = 1, \ldots, n_i$, erwartungstreu schätzen würden, so könnten wir die Varianz minimieren, indem wir die Gewichte umgekehrt proportional zu $\operatorname{Var}\left(\widehat{F}_{iks}(x)\right)$ wählen. Leider sind diese Varianzen unbekannt und darüber hinaus im Allgemeinen für unterschiedliche x verschieden, so dass wir hier trotz eines möglichen Effizienzverlustes aus Gründen der Einfachheit wieder ein ungewichtetes Mittel wählen:

$$\widehat{F}_{is,n} = \frac{1}{\lambda_{i\cdot s}} \sum_{k=1}^{n_i} \widehat{F}_{iks}.$$

Wie obige Diskussion gezeigt hat, sind unsere Definitionen der Schätzer für die Verteilungsfunktionen mit Ausnahme der einfachen Versuchspläne, in denen $m_{iks} = 1$ für alle Tripel (i, k, s) gilt, sicher nicht die einzigen sinnvollen Definitionen. Dennoch bezeichnen wir diese Schätzer schlichtweg als die empirischen Verteilungsfunktionen.

Als Schätzer für die gemittelten Verteilungsfunktionen H_{is} benutzen wir entsprechend der Definition

$$\widehat{H}_{is,n} = \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} \widehat{F}_{uv,n}.$$

Wir fassen nun einige Eigenschaften der empirischen Verteilungsfunktionen zusammen.

Satz 4.1. *Die empirischen Verteilungsfunktionen sind erwartungstreu und für ihre Varianzen gelten die Abschätzungen*

$$\begin{aligned}\operatorname{Var}\left(\widehat{F}_{is,n}(x)\right) &\leq \frac{1}{4\lambda_{i\cdot s}} \quad \textit{und} \\ \operatorname{Var}\left(\widehat{H}_{is,n}(x)\right) &\leq \frac{1}{4(at-1)^2} \sum_{(u,v)\neq(i,s)} \frac{1}{\lambda_{u\cdot v}}.\end{aligned}$$

Beweis. Siehe Anhang A.8, Seite 100. □

Aufgrund des obigen Satzes können wir vermuten, das die erste der drei asymptotischen Annahmen (A1) – (A3) an die Folge der Stichprobenumfänge (siehe Seite 13) bereits hinreichend ist, um die Konsistenz der empirischen Verteilungsfunktionen zu garantieren. Die Richtigkeit dieser Vermutung entnehmen wir der folgenden Verallgemeinerung des Satzes von Glivenko-Cantelli.

Satz 4.2. *Unter der Bedingung (A1) gilt für alle Paare* (i, s)

$$\left\|\widehat{F}_{is,n} - F_{is}\right\|_\infty \xrightarrow{f.s.} 0 \quad \textit{und} \quad \left\|\widehat{H}_{is,n} - H_{is}\right\|_\infty \xrightarrow{f.s.} 0.$$

Beweis. Siehe Anhang A.9, Seite 101. □

Wegen der Darstellungen $p_{is} = \int H_{is}\, dF_{is}$ und $\widehat{p}_{is,n} = \int \widehat{H}_{is,n}\, d\widehat{F}_{is,n}$ können wir nach diesem Satz erwarten, daß die Bedingung (A1) auch für die Konsistenz von $\widehat{p}_{is,n}$ ausreichend ist. Diese Vermutung wird im folgenden Abschnitt bestätigt.

4.4 Eigenschaften der Punktschätzer

In diesem Abschnitt werden wir zunächst untersuchen, inwiefern sich die im vorigen Abschnitt gefundenen Eigenschaften der empirischen Verteilungsfunktionen auf die Schätzer für die relativen Effekte übertragen. Darüber hinaus stellen wir Überlegungen zur Robustheit und Äquivarianz der Schätzer an.

Erwartungstreue

Zunächst befassen wir uns mit der Frage nach der Erwartungstreue. Dazu zerlegen wir den Schätzer in eine Summe, deren Summanden wir einzeln behandeln können:

$$\begin{aligned}\widehat{p}_{is,n} &= \int \widehat{H}_{is,n}\, d\widehat{F}_{is,n} = \int \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} \widehat{F}_{uv,n}\, d\widehat{F}_{is,n} \\ &= \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} \int \widehat{F}_{uv,n}\, d\widehat{F}_{is,n} \\ &= \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} \int \frac{1}{\lambda_{u\cdot v}} \sum_{w=1}^{n_u} \widehat{F}_{uwv}\, d\left(\frac{1}{\lambda_{i\cdot s}} \sum_{k=1}^{n_i} \widehat{F}_{iks}\right) \\ &= \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} \frac{1}{\lambda_{u\cdot v}\lambda_{i\cdot s}} \sum_{w=1}^{n_u} \sum_{k=1}^{n_i} \int \widehat{F}_{uwv}\, d\widehat{F}_{iks}.\end{aligned}$$

Für p_{is} existiert die analoge Zerlegung

$$p_{is} = \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} \frac{1}{\lambda_{u\cdot v}\lambda_{i\cdot s}} \sum_{w=1}^{n_u} \sum_{k=1}^{n_i} \int F_{uwv}\, dF_{iks},$$

wobei wir $F_{uwv} = \lambda_{uwv} F_{uv}$ und $F_{iks} = \lambda_{iks} F_{is}$ definieren.

Damit ist $\widehat{p}_{is,n}$ erwartungstreu, wenn für alle Tripel (u, w, v) und (i, k, s) mit $(u, v) \neq (i, s)$

$$\mathrm{E}\left(\int \widehat{F}_{uwv}\, d\widehat{F}_{iks}\right) = \int F_{uwv}\, dF_{iks}$$

gilt. Falls $(u, w) \neq (i, k)$ ist, so ergibt sich dies wegen der Unabhängigkeit von $\boldsymbol{X}_{uwv}$ und $\boldsymbol{X}_{iks}$ sowie wegen der Erwartungstreue von $\widehat{F}_{uwv}$ und $\widehat{F}_{iks}$ aus dem Satz von Fubini:

$$\begin{aligned}
\mathrm{E}\left(\int \widehat{F}_{uwv}\, d\widehat{F}_{iks}\right) &= \mathrm{E}\left[\mathrm{E}\left(\int \widehat{F}_{uwv}\, d\widehat{F}_{iks} \middle| \boldsymbol{X}_{uwv}\right)\right] \\
&= \mathrm{E}\left[\int \widehat{F}_{uwv}\, dF_{iks}\right] \\
&= \int F_{uwv}\, dF_{iks}.
\end{aligned}$$

Für $(u, w) = (i, k)$ sind $\boldsymbol{X}_{uwv}$ und $\boldsymbol{X}_{iks} = \boldsymbol{X}_{uws}$ aber möglicherweise abhängig, und dann kann auch

$$\mathrm{E}\left(\int \widehat{F}_{uwv}\, d\widehat{F}_{iks}\right) \neq \int F_{uwv}\, dF_{iks}$$

gelten. Der Schätzer $\widehat{p}_{is,n}$ ist also im Allgemeinen nicht erwartungstreu. Wir können aber zumindest die Verzerrung des Schätzers nach oben abschätzen:

$$\begin{aligned}
&|\mathrm{E}(\widehat{p}_{is,n}) - p_{is}| \\
&= \left| \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} \frac{1}{\lambda_{u\cdot v}\lambda_{i\cdot s}} \sum_{w=1}^{n_u} \sum_{k=1}^{n_i} \mathrm{E}\left(\int \widehat{F}_{uwv}\, d\widehat{F}_{iks}\right) - \int F_{uwv}\, dF_{iks} \right| \\
&= \left| \frac{1}{at-1} \sum_{v\neq s} \frac{1}{\lambda_{i\cdot v}\lambda_{i\cdot s}} \sum_{w=1}^{n_i} \sum_{k=1}^{n_i} \mathrm{E}\left(\int \widehat{F}_{iwv}\, d\widehat{F}_{iks}\right) - \int F_{iwv}\, dF_{iks} \right| \\
&= \left| \frac{1}{at-1} \sum_{v\neq s} \frac{1}{\lambda_{i\cdot v}\lambda_{i\cdot s}} \sum_{k=1}^{n_i} \mathrm{E}\left(\int \widehat{F}_{ikv}\, d\widehat{F}_{iks}\right) - \int F_{ikv}\, dF_{iks} \right| \\
&\leq \frac{1}{at-1} \sum_{v\neq s} \frac{1}{\lambda_{i\cdot v}\lambda_{i\cdot s}} \sum_{k=1}^{n_i} \left| \mathrm{E}\left(\int \widehat{F}_{ikv}\, d\widehat{F}_{iks}\right) - \int F_{ikv}\, dF_{iks} \right| \\
&\leq \frac{1}{at-1} \sum_{v\neq s} \frac{1}{\lambda_{i\cdot v}\lambda_{i\cdot s}} \sum_{k=1}^{n_i} 1 \\
&\leq \frac{1}{at-1} \sum_{v\neq s} \frac{1}{\lambda_{i\cdot s}} = \frac{t-1}{(at-1)\lambda_{i\cdot s}} \leq \frac{1}{a\lambda_{i\cdot s}}.
\end{aligned}$$

Die obige Abschätzung zeigt, dass die Verzerrungen bei hinreichend vielen Gruppen und hinreichend großen $\lambda_{i \cdot s}$ vernachlässigt werden können. Darüber hinaus können wir ablesen, dass der Schätzer für den Fall $t = 1$ unverzerrt ist.

Konsistenz

Wenden wir uns nun der Konsistenz von $\widehat{p}_{is,n}$ zu. Diese ergibt sich ohne Schwierigkeiten aus Satz 4.2.

Satz 4.3. *Unter der Bedingung (A1) auf Seite 13 ist $\widehat{p}_{is,n}$ stark konsistent.*

Beweis. Siehe Anhang A.10, Seite 102. □

Aus obigem Satz folgt unmittelbar

Korollar 4.4. *Unter der Bedingung (A1) sind die Schätzer $\widehat{p}_{i*,n}$ und $\widehat{p}_{*s,n}$ für die Gruppen- und Bedingungseffekte stark konsistent.*

Beweis. Die Behauptung ergibt sich mit dem voranstehenden Satz sofort aus den Darstellungen in Satz 3.13 (siehe Seite 43) und der Definition der Schätzer. □

Robustheit

Mit der Erwartungstreue und der Konsistenz haben wir zwei der wichtigsten mathematisch-statistischen Eigenschaften der Punktschätzer für die relativen Effekte bereits diskutiert. Eine Vielzahl weiterer Aspekte von Punktschätzern, die dann eine Rolle spielen, wenn die Beobachtungen eines Versuchs einem formulierten Modell nicht exakt, sondern nur näherungsweise genügen, sind Gegenstand der robusten Statistik. In diesem Unterabschnitt befassen wir uns mit einem dieser Aspekte, dem Grobe-Fehler Bruchpunkt des Schätzers $\widehat{p}_{is,n}$. Weitere Aspekte sind zum Beispiel in dem Lehrbuch von Hampel et al. (1986) oder im Zusammenhang mit nichtparametrischer Statistik in der Monographie von Hettmansperger und McKean (1998) dargestellt.

Der Grobe-Fehler Bruchpunkt eines Schätzers (genauer: einer Folge von Schätzern) gibt Auskunft darüber, welcher Anteil an Beobachtungen durch Ausreißer ersetzt werden kann, ohne dass der Schätzer alle in ihm enthaltene Information verliert. Um den Grobe-Fehler Bruchpunkt exakt definieren zu können, führen wir zunächst einige Schreibweisen ein. Das Wahrscheinlichkeitsmaß, das zu dem Wahrscheinlichkeitsraum gehört, auf dem die Zufallsvariablen $\boldsymbol{X}_{ik}$, $i = 1, \ldots, a$, $k = 1, \ldots, n_i$, definiert sind, bezeichnen wir mit $P_{\boldsymbol{F}}$. Der Vektor

$$\boldsymbol{G} = (G_{11}, \ldots, G_{1t}, \ldots, G_{a1}, \ldots, G_{at})'$$

sei wie $\boldsymbol{F}$ ein (at)-dimensionaler Vektor von Verteilungsfunktionen. Als Grobe-Fehler Bruchpunkt des Schätzers $\widehat{p}_{is,n}$ an der Stelle $\boldsymbol{F}$ definieren wir (siehe Hampel et al., 1986, Seite 97)

$$\epsilon_{is}^*(\boldsymbol{F}) = \sup \left\{ \epsilon \leq 1 : \exists K_\epsilon \text{ kompakt}, K_\epsilon \subsetneq [0,1] : \forall \boldsymbol{G} : P_{(1-\epsilon)\boldsymbol{F}+\epsilon\boldsymbol{G}}(\widehat{p}_{is,n} \in K_\epsilon) \to 1 \right\}.$$

Wie der folgende Satz zeigt, ist der Grobe-Fehler Bruchpunkt von $\widehat{p}_{is,n}$ für alle $\boldsymbol{F}$ maximal.

Satz 4.5. *Für alle $\boldsymbol{F}$ und alle Paare (i,s) gilt unter der Bedingung (A1)*

$$\epsilon^*_{is}(\boldsymbol{F}) = 1.$$

Beweis. Siehe Anhang A.11, Seite 102. □

Der obige Satz ist insbesondere deshalb bemerkenswert, weil sogar als besonders robust geltende Schätzer wie der empirische Median nur einen Grobe-Fehler Bruchpunkt von 1/2 haben.

Äquivarianz

Wenn wir die Beobachtungen X_{iksr} mit Hilfe einer streng monoton wachsenden Abbildung ϕ durch $\phi(X_{iksr})$ ersetzen und dann die relativen Effekte p^ϕ_{is} bilden, so erhalten wir – wie aus Definition 3.2 auf Seite 30 hervorgeht – wieder die ursprünglichen relativen Effekten, das heißt,

$$p^\phi_{is} = p_{is}.$$

Wenn wir die Beobachtungen $X_{iks\ell}$ dagegen mit Hilfe einer streng monoton fallenden Abbildung ϕ durch $\phi(X_{iksr})$ ersetzen und dann die relativen Effekte p^ϕ_{is} bilden, so gilt

$$\begin{aligned}
p^\phi_{is} &= \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} \left[\mathrm{P}(\phi(X_{u2v1}) < \phi(X_{i1s1})) + \frac{1}{2}\,\mathrm{P}(\phi(X_{u2v1}) = \phi(X_{i1s1}))\right] \\
&= \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} \left[\mathrm{P}(X_{u2v1} > X_{i1s1}) + \frac{1}{2}\,\mathrm{P}(X_{u2v1} = X_{i1s1})\right] \\
&= \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} \left[1 - \mathrm{P}(X_{u2v1} \leq X_{i1s1}) + \frac{1}{2}\,\mathrm{P}(X_{u2v1} = X_{i1s1})\right] \\
&= \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} \left[1 - \left(\mathrm{P}(X_{u2v1} < X_{i1s1}) + \frac{1}{2}\,\mathrm{P}(X_{u2v1} = X_{i1s1})\right)\right] \\
&= 1 - p_{is}.
\end{aligned}$$

Sowohl für streng monoton wachsende als auch für streng monoton fallende Abbildungen ϕ erscheint es unmittelbar einleuchtend, dass sich die Schätzer für die relativen Effekte in gleicher Weise wie die relativen Effekte selbst verhalten sollten. Im ersten Fall sollte also

$$\widehat{p}^\phi_{is,n} = \widehat{p}_{is,n}$$

und im zweiten Fall

$$\widehat{p}^\phi_{is,n} = 1 - \widehat{p}_{is,n}$$

gelten. Falls die Schätzer $\widehat{p}_{is,n}$ eine dieser beiden Eigenschaften haben, so heißen sie äquivariant bezüglich ϕ. Tatsächlich gilt

Satz 4.6. *Die Schätzer $\widehat{p}_{is,n}$ sind äquivariant bezüglich jeder streng monotonen Abbildung ϕ.*

Beweis. Wie wir im folgenden Abschnitt in den Gleichungen (4.1) und (4.2) sehen werden, lassen sich die Schätzer $\widehat{p}_{is,n}$ als Konvexkombinationen von Summanden der Form

$$c(X_{iksr} - X_{uwv\ell})$$

darstellen. Unter streng monoton wachsenden Abbildungen ϕ sind diese invariant, und für streng monoton fallende Abbildungen ϕ gilt

$$c(\phi(X_{iksr}) - \phi(X_{uwv\ell})) = 1 - c(X_{iksr} - X_{uwv\ell}).$$

Die Bildung von Konvexkombinationen überträgt dieses Verhalten auf $\widehat{p}_{is,n}$. □

Analoge Überlegungen für die relativen Gruppen- und Bedingungseffekte und die zugehörigen Schätzer zeigen, dass auch diese äquivariant unter streng monotonen Transformationen der Beobachtungen sind.

4.5 Praktische Berechnung der Punktschätzer

Die Schätzer $\widehat{p}_{is,n}$ sind exakt und ohne aufwendige Verfahren aus der numerischen Mathematik berechenbar. Die Umsetzung zum Beispiel mit der Programmiersprache IML, die in dem weit verbreiteten Programmpaket SAS enthalten ist, bereitet keinerlei prinzipielle Schwierigkeiten. Da Programme in dieser Programmiersprache jedoch nur interpretiert und nicht von einem Compiler übersetzt werden, ist je nach Größe und Struktur des Datensatzes die Laufzeit möglicherweise nicht zu vernachlässigen. Dies trifft insbesondere auf Simulationen zu, bei denen tausende von einem Zufallszahlengenerator erzeugte Datensätze ausgewertet werden. Deshalb werden wir in diesem Abschnitt Vorschläge machen, die bei der effizienten Berechnung der Punktschätzer hilfreich sein können.

Zunächst stellen wir $\widehat{p}_{is,n}$ als Linearkombination von Werten der empirischen Verteilungsfunktionen $\widehat{F}_{uv,n}$ mit $(u,v) \neq (i,s)$ an den Stellen X_{iksr}, $k = 1, \ldots, n_i$, $r = 1, \ldots, m_{iks}$, dar:

$$\begin{aligned}
\widehat{p}_{is,n} &= \int \widehat{H}_{is,n}\, d\widehat{F}_{is,n} \\
&= \int \widehat{H}_{is,n}\, d\left(\frac{1}{\lambda_{i\cdot s}} \sum_{k=1}^{n_i} \frac{1}{m_{iks}} \sum_{r=1}^{m_{iks}} \widehat{F}_{iksr}\right) \\
&= \frac{1}{\lambda_{i\cdot s}} \sum_{k=1}^{n_i} \frac{1}{m_{iks}} \sum_{r=1}^{m_{iks}} \int \widehat{H}_{is,n}\, d\widehat{F}_{iksr} \\
&= \frac{1}{\lambda_{i\cdot s}} \sum_{k=1}^{n_i} \frac{1}{m_{iks}} \sum_{r=1}^{m_{iks}} \widehat{H}_{is,n}(X_{iksr}) \\
&= \frac{1}{\lambda_{i\cdot s}} \sum_{k=1}^{n_i} \frac{1}{m_{iks}} \sum_{r=1}^{m_{iks}} \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} \widehat{F}_{uv,n}(X_{iksr}). \qquad (4.1)
\end{aligned}$$

Die Werte $\widehat{F}_{uv,n}(X_{iksr})$ sind Bausteine sowohl für die Schätzer der relativen Effekte als auch für den in Abschnitt 5.3 (siehe Seite 63) behandelten Schätzer $\widehat{\boldsymbol{V}}_n$ der

Kovarianzmatrix von $\sqrt{n}\,(\widehat{\boldsymbol{p}}_n - \boldsymbol{p})$. Wir gehen deshalb hier auf die Berechnung der $\widehat{F}_{uv,n}(X_{iksr})$ auch für den Fall $(u,v)=(i,s)$ ein, der zwar nicht an dieser Stelle, aber später für $\widehat{\boldsymbol{V}}_n$ benötigt wird.

Ein naiver Ansatz zur Berechnung der $\widehat{F}_{uv,n}(X_{iksr})$ führt auf die Zählfunktion c:

$$\begin{aligned} \widehat{F}_{uv,n}(X_{iksr}) &= \frac{1}{\lambda_{u\cdot v}} \sum_{w=1}^{n_u} \widehat{F}_{uwv}(X_{iksr}) \\ &= \frac{1}{\lambda_{u\cdot v}} \sum_{w=1}^{n_u} \frac{1}{m_{uwv}} \sum_{\ell=1}^{m_{uwv}} \widehat{F}_{uwv\ell}(X_{iksr}) \\ &= \frac{1}{\lambda_{u\cdot v}} \sum_{w=1}^{n_u} \frac{1}{m_{uwv}} \sum_{\ell=1}^{m_{uwv}} \mathrm{c}(X_{iksr} - X_{uwv\ell}). \end{aligned} \tag{4.2}$$

Die Auswertung von $\mathrm{c}(X_{iksr} - X_{uwv\ell})$ kann mit Hilfe einer einfachen Fallunterscheidung durchgeführt werden.

Falls für alle Versuchseinheiten aus der u-ten Gruppe unter der s-ten Bedingung die Anzahlen der Messwiederholungen m_{uwv} gleich einer Zahl m_{uv} oder 0 sind, das heisst, falls

$$m_{uwv} \in \{0, m_{uv}\} \quad w = 1, \ldots, n_u, \tag{4.3}$$

gilt, so existiert für $\widehat{F}_{uv,n}(X_{iksr})$ die folgende durch Abzählen nachvollziehbare Rangdarstellung:

$$\widehat{F}_{uv,n}(X_{iksr}) = \begin{cases} \frac{1}{\lambda_{u\cdot v} m_{uv}} \left(R^{(uv)}_{ukvr} - \frac{1}{2} \right) & : (u,v) = (i,s), \\ \frac{1}{\lambda_{u\cdot v} m_{uv}} \left(R_{iksr} - R^{(-uv)}_{iksr} \right) & : (u,v) \neq (i,s). \end{cases}$$

Hierbei bezeichnet $R^{(uv)}_{ukvr}$ den Mittelrang von X_{ukvr} unter allen Beobachtungen aus der Gruppe u unter der Bedingung v, R_{iksr} den Mittelrang von X_{iksr} unter allen Beobachtungen und $R^{(-uv)}_{iksr}$ den Mittelrang von X_{iksr} unter allen Beobachtungen mit Ausnahme derer aus Gruppe u unter Bedingung v. In IML und vielen anderen statistisch orientierten Programmiersprachen aus stehen Funktionen zur Berechnung von Rängen zur Verfügung. Da in interpretierten Programmen die Durchführung einzelner Programmschritte mit hohen Grundkosten bezüglich der Laufzeit behaftet ist und durch die Verwendung der Rangdarstellungen die Zahl an Operationen gegenüber der naiven Vorgehensweise jedenfalls für hinreichend große $\lambda_{u\cdot v} m_{uv}$ verringert werden kann, können wir auch insgesamt bei Verwendung der Rangdarstellungen auf eine geringere Laufzeit hoffen.

Falls die Bedingung (4.3) nicht erfüllt ist, können wir immerhin noch eine allgemein gültige Rangdarstellung für die $\widehat{F}_{uwv}(X_{iksr})$ angeben:

$$\widehat{F}_{uwv}(X_{iksr}) = \begin{cases} \frac{1}{m_{uwv}} \left(R^{(uwv)}_{uwvr} - \frac{1}{2} \right) & : (u,w,v) = (i,k,s), \\ \frac{1}{m_{uwv}} \left(R_{iksr} - R^{(-uwv)}_{iksr} \right) & : (u,w,v) \neq (i,k,s). \end{cases}$$

Dabei bezeichnet $R^{(uwv)}_{uwvr}$ den Mittelrang von X_{uwvr} unter den Beobachtungen aus dem Vektor $\boldsymbol{X}_{uwv}$ und $R^{(-uwv)}_{iksr}$ den Mittelrang von X_{iksr} unter allen Beobachtungen mit Ausnahme derer aus dem Vektor $\boldsymbol{X}_{uwv}$. Die Darstellung kann wiederum durch Abzählen überprüft werden. Durch Verwendung dieser Rangdarstellung ist – wenn überhaupt – kein nennenswerter Gewinn bezüglich der Laufzeit zu erwarten, da die m_{iks} in der Regel relativ klein sein werden. Für die asymptotische Gültigkeit der in Kapitel 5 hergeleiteten Konfidenzintervalle muss jedenfalls die Bedingung (A3) auf Seite 13 erfüllt sein, die die gleichmäßige Beschränktheit der m_{iks} fordert.

Die in diesem Abschnitt geschilderten Überlegungen wurden in den SAS-Makros `RE_CI` und `SIMPLE_RE_CI` umgesetzt. Diese Makros sind in Anhang B beschrieben und wurden zur Berechnung der Punktschätzer für die Beispiele im folgenden Abschnitt verwendet.

4.6 Beispiele

Wir berechnen nun für die Beispiele aus Unterabschnitt 2.1.1 die in diesem Kapitel konstruierten Schätzer für die relativen Effekte. Diese Schätzer stellen wir jeweils in graphischer Form dar und geben kurze Hinweise zur Interpretation. Zu Bedenken ist aber bei jedem der Versuche, dass die Auswertung auf der alleinigen Grundlage von Punktschätzern zunächst insofern unvollständig bleibt, als noch keine Aussagen über die Genauigkeit der Schätzer gemacht werden. Aussagen über die Genauigkeit werden anhand von Konfidenzintervallen im anschließenden Kapitel behandelt.

Panik-Skala-Studie I

Das Ziel der Panik-Skala-Studie I war, die Entwicklung der CGI bei den mit Imipramin behandelten Patienten mit Panikstörung und Agoraphobie in einem Zeitraum von acht Wochen zu untersuchen. Die Schätzer für die relativen Effekte sind in Abbildung 4.1 dargestellt.

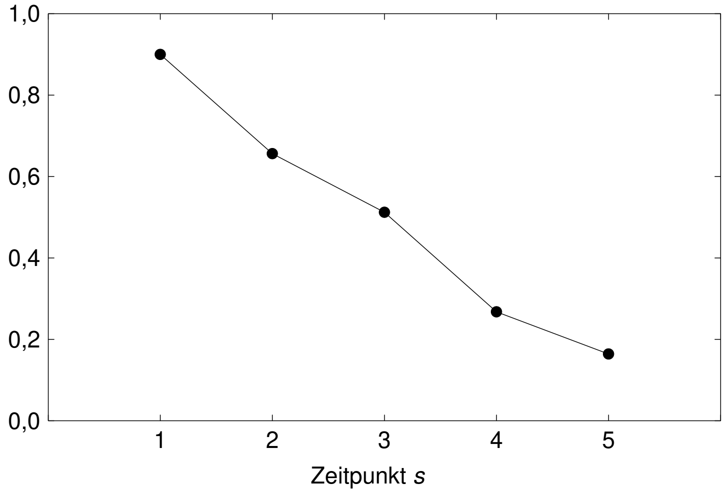

Abbildung 4.1: Schätzer $\widehat{p}_{1s}$ für die relativen Effekte in der Panik-Skala-Studie I.

Zweck der Behandlung mit Imipramin war eine Verbesserung der CGI, also eine Verringerung der CGI-Scores. Die Schätzer für die relativen Effekte deuten darauf hin, dass die Behandlung ihren Zweck erfüllt, weil sie von Zeitpunkt zu Zeitpunkt immer kleiner werden. Nach Definition der relativen Effekte haben wir also Grund zu der Annahme, dass die CGI-Scores der Patienten zu jedem der fünf Zeitpunkte mit umso größerer Wahrscheinlichkeit kleiner sind als die CGI-Scores zu den anderen Zeitpunkten, je später dieser Zeitpunkt liegt. Bemerkenswert ist auch der Wert 0,90 für den Schätzer des relativen Effektes zum Beginn der Behandlung. Wir können vermuten, dass Patienten zu diesem Zeitpunkt mit einer Wahrscheinlichkeit von ungefähr 90% einen größeren CGI-Score haben als die anderen Patienten im Mittel über die übrigen Zeitpunkte.

Panik-Skala-Studie II

Das Ziel der Panik Skala-Studie II war die Untersuchung der zeitlichen Entwicklung der P&A-Scores von Patienten mit Panikstörung während einer achtwöchigen Imipramin-Behandlung. Dabei sollten die Patienten mit und ohne Agoraphobie gegenübergestellt werden. In Abbildung 4.2 sind die Schätzer für die relativen Effekte angegeben.

Anhand der Grafik können wir erkennen, dass sowohl für die Patienten mit Agoraphobie als auch für die Patienten ohne Agoraphobie die geschätzten relativen Effekte im Laufe der Zeit immer kleiner werden. Wir haben also wie bei der Panik-Skala-Studie I Grund anzunehmen, dass die Panikstörung während der Behandlung mit Imipramin zurückgeht, und zwar für beide Gruppen von Patienten. Der zu Beginn der Behandlung sichtbare Unterschied zwischen den beiden Gruppen scheint durch die Behandlung innerhalb der acht Wochen nicht ausgeglichen werden zu können.

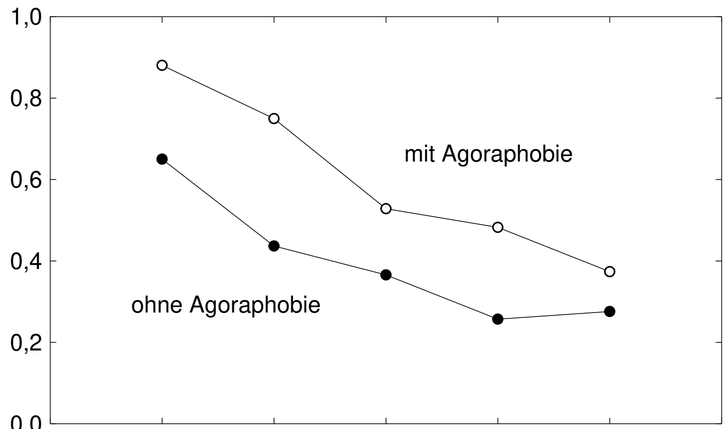

Abbildung 4.2: Schätzer $\widehat{p}_{is}$ für die relativen Effekte in der Panik-Skala-Studie II.

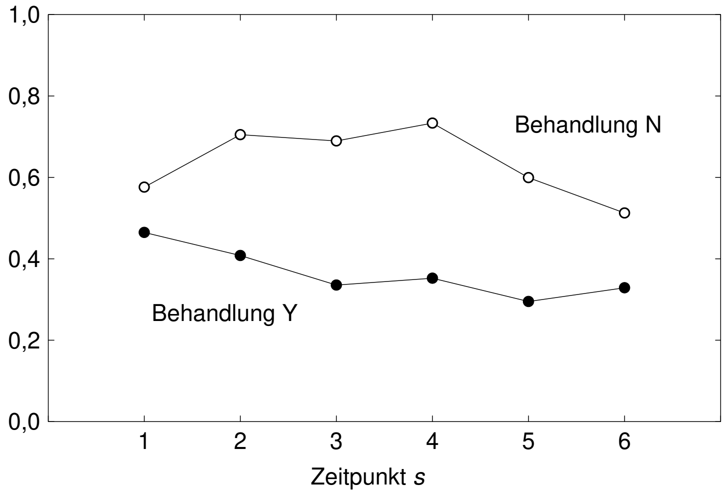

Abbildung 4.3: Schätzer $\widehat{p}_{is}$ für die relativen Effekte in der Schulterschmerz-Studie.

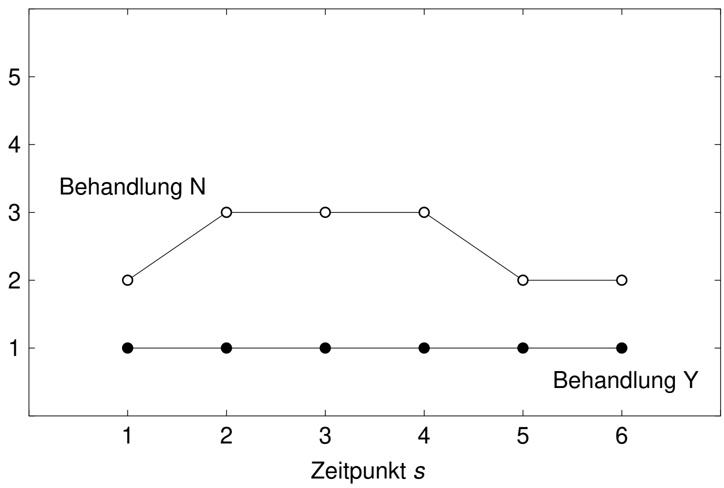

Abbildung 4.4: Empirische Mediane für die Daten aus der Schulterschmerz-Studie.

Schulterschmerz-Studie

Mit Hilfe der Schulterschmerz-Studie sollte untersucht werden, inwieweit ein spezielles Verfahren zur Absaugung der Luft nach einer laparoskopischen Operation in der Lage ist, die in der Folgezeit typischerweise auftretenden Schmerzen in der Schulterspitze zu reduzieren. Zum Vergleich wurden die Daten auch in einer Kontrollgruppe erhoben, in der keine Luft abgesaugt wurde. Die Werte der Schätzer für die relativen Effekte sind in Abbildung 4.3 dargestellt.

Die Schätzer für die relativen Effekte liefern einen Hinweis darauf, dass die Behandlung ihrem Zweck tatsächlich genügt: Wir können vermuten, dass jede Randverteilung der Schmerz-Scores in der Behandlungsgruppe (Behandlung Y) mit Bezug auf die mittlere Verteilung tendenziell kleiner ist als jede Randverteilung in der Kontrollgruppe (Behandlung N). Darüber hinaus scheinen die Randverteilungen in der

Kontrollgruppe mit Bezug auf die mittlere Verteilungsfunktion zunächst tendenziell größer und erst nach dem vierten Zeitpunkt tendenziell kleiner zu werden. In der Behandlungsgruppe dagegen sieht es so aus, als ob die Randverteilungen bereits vom ersten Zeitpunkt an mit Bezug auf die mittlere Verteilungsfunktion tendenziell kleiner würden. Die geschätzten relativen Effekte liefern hier also nicht nur Hinweise zur Wirksamkeit der Behandlung, sondern auch zum zeitlichen Verlauf der Schmerzintensität. Hier liefern die relativen Effekte genauere Informationen als die sonst in der nichtparametrischen Statistik häufig verwendeten empirischen Mediane (siehe Abbildung 4.4). Der Grund dafür liegt darin, dass die verwendete Skala der Schmerz-Scores mit nur fünf möglichen Werten sehr grob ist und deshalb auch für die Mediane nur diese fünf Werte möglich sind.

Wasser-Irrgarten-Test

Ziel des Wasser-Irrgarten-Tests war herauszufinden, inwiefern die verschiedenen Dosierungen der verabreichten Substanz sich auf das Geschick und die Lernfähigkeit der Ratten im Wasser-Irrgarten auswirkt. Die Schätzer $\widehat{p}_{is}$ für die relativen Effekte sind in Abbildung 4.5 dargestellt.

Die graphische Darstellung lässt vermuten, dass sich die Substanz insbesondere in der höchsten Dosierung negativ auswirkt. Darüber hinaus scheint sich bei allen Dosierungen das Verhalten der Tiere im Labyrinth von Tag 1 zu Tag 7 zu verbessern.

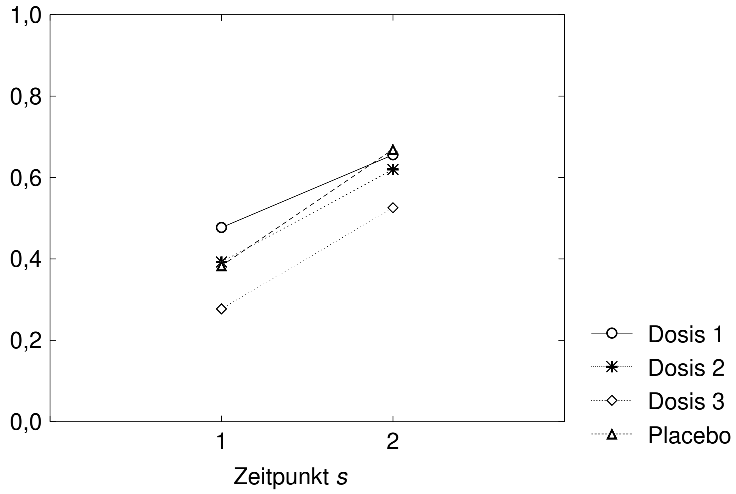

Abbildung 4.5: Schätzer $\widehat{p}_{is}$ für die relativen Effekte im Wasser-Irrgarten-Test.

Kapitel 5

Konfidenzintervalle

In diesem Kapitel leiten wir asymptotische Konfidenzintervalle für die relativen Effekte her. Diese Konfidenzintervalle beruhen auf der asymptotischen Normalverteilung von $\sqrt{n}\,(\widehat{\boldsymbol{p}}_n - \boldsymbol{p})$, sowie konsistenten Schätzern für die Kovarianzmatrix von $\sqrt{n}\,(\widehat{\boldsymbol{p}}_n - \boldsymbol{p})$. Zusätzlich zu den direkt auf der asymptotischen Normalverteilung beruhenden Konfidenzintervallen schlagen wir drei andere Konfidenzintervalle vor, die das vorgegebene Niveau insbesondere für kleine Stichprobenumfänge besser einhalten sollen. Durch Simulationen vergleichen wir schließlich die verschiedenen Verfahren und wenden sie auf die vier Beispiele aus Unterabschnitt 2.1.1 an.

5.1 Übersicht über Arbeiten zu zwei unverbundenen Stichproben

Bevor wir mit der Herleitung von Konfidenzintervallen für die allgemeine Klasse von Versuchsplänen aus Abschnitt 2.1 beginnen, geben wir noch eine kurze Übersicht über die Literatur zu Konfidenzintervallen für den einfachen Spezialfall zweier unverbundener Stichproben ($a = 2$, $t = 1$, alle $m_{ik1} = 1$).

Konfidenzintervalle für normalverteilte Beobachtungen werden von Owen et al. (1964), Govindarajulu (1967), Downton (1973) sowie Reiser und Guttman (1986) behandelt. Für stetige Verteilungen findet man Vorschläge in den Arbeiten von Birnbaum (1956), Birnbaum und McCarty (1958), Owen et al. (1964), Sen (1967), Govindarajulu (1968), Ury (1972), Cheng und Chao (1984), Halperin et al. (1987) und Mee (1990). Ury und Wiggins (1979) befassen sich mit Lokationsmodellen, Simonoff et al. (1986) mit kategorisierten Daten und Beder und Heim (1990, 1992) mit geordnet kategoriellen Daten. Chen und Kianifard (2000) schließlich haben Konfidenzintervalle für die hier behandelte allgemeine Klasse von Verteilungen angegeben.

5.2 Asymptotische Verteilung der Schätzer für die relativen Effekte

Grundlage für die Konstruktion asymptotischer Konfidenzintervalle für die relativen Effekte ist die asymptotische Normalverteilung von

$$\sqrt{n}\,(\widehat{p}_{is,n} - p_{is})\,, \quad \sqrt{n}\,(\widehat{p}_{i*,n} - p_{i*}) \quad \text{bzw.} \quad \sqrt{n}\,(\widehat{p}_{*s,n} - p_{*s})\,.$$

Statt diese getrennt nachzuweisen, behandeln wir allgemeiner die asymptotische multivariate Normalverteilung der Zufallsvektoren

$$\sqrt{n}\,(\widehat{\boldsymbol{p}}_n - \boldsymbol{p})\,.$$

Die Grundlage dafür und später auch für die konsistente Schätzung der Kovarianzmatrix bildet der sogenannte *asymptotische Äquivalenzsatz.*

An dieser Stelle wird zum ersten Mal die Bedingung (A2) an die Stichprobenumfänge benötigt, die bereits in Unterabschnitt 2.1.2 auf Seite 13 angegeben wurde.

Satz 5.1 (Asymptotischer Äquivalenzsatz). *Unter den Voraussetzungen (A1) und (A2) gilt*

$$\sqrt{n}\int \widehat{\boldsymbol{H}}_n\, d\left(\widehat{\boldsymbol{F}}_n - \boldsymbol{F}\right) - \sqrt{n}\int \boldsymbol{H}\, d\left(\widehat{\boldsymbol{F}}_n - \boldsymbol{F}\right) \xrightarrow{P} \mathbf{0}.$$

Beweis. Siehe Anhang A.12, Seite 104. □

Der asymptotische Äquivalenzsatz erlaubt eine Zerlegung von $\sqrt{n}\,(\widehat{\boldsymbol{p}}_n - \boldsymbol{p})$ in eine asymptotisch äquivalente zentrierte Summe von unabhängigen Zufallsvektoren:

Korollar 5.2. *Für die unabhängigen und unter Voraussetzung (A2) gleichmäßig beschränkten Zufallsvektoren $\boldsymbol{\Psi}_{ik,n} \in \mathbb{R}^{at}$, deren Komponenten $\Psi_{ikuv,n}$ durch*

$$\Psi_{ikuv,n} = \begin{cases} \frac{n_u}{\lambda_{u\cdot v}} \int H_{uv}\, d\widehat{F}_{ukv} - \frac{1}{at-1} \sum_{s\neq v} \frac{n_u}{\lambda_{u\cdot s}} \int F_{uv}\, d\widehat{F}_{uks} : & u = i, \\ \frac{-1}{at-1} \sum_{s=1}^{t} \frac{n_i}{\lambda_{i\cdot s}} \int F_{uv}\, d\widehat{F}_{iks} : & u \neq i, \end{cases}$$

definiert sind, gilt unter den Voraussetzungen (A1) und (A2)

$$\sqrt{n}\,(\widehat{\boldsymbol{p}}_n - \boldsymbol{p}) - \sqrt{n}\left(\sum_{i=1}^{a} \frac{1}{n_i} \sum_{k=1}^{n_i} [\boldsymbol{\Psi}_{ik,n} - \mathrm{E}(\boldsymbol{\Psi}_{ik,n})]\right) \xrightarrow{P} \mathbf{0}.$$

Beweis. Siehe Anhang A.13, Seite 107. □

Mit Hilfe des obigen Korollars können wir nun versuchen, die asymptotische Normalverteilung von $\sqrt{n}\,(\widehat{\boldsymbol{p}}_n - \boldsymbol{p})$ nachzuweisen, indem wir die asymptotische Normalverteilung einer Folge von zentrierten Summen gleichmäßig beschränkter und unabhängiger Zufallsvektoren beweisen. Wenn die Folge der Kovarianzmatrizen dieser Summen

$$\boldsymbol{V}_n = \mathrm{Cov}\left(\sqrt{n} \sum_{i=1}^{a} \frac{1}{n_i} \sum_{k=1}^{n_i} [\boldsymbol{\Psi}_{ik,n} - \mathrm{E}(\boldsymbol{\Psi}_{ik,n})]\right)$$

konvergieren würde, könnten wir dazu den wohlbekannten multivariaten Zentralen Grenzwertsatz (siehe z. B. Rao, 1973, Kapitel 2, Complements and Problems 4.7, oder van der Vaart, 1998, Proposition 2.27) verwenden. Da sich diese Konvergenz aus den Voraussetzungen (A1) – (A3) aber nicht ergibt und als Bedingung für die Versuchsplanung nur schwierig zu deuten ist, werden wir hier vermeiden, sie als weitere Annahme zu formulieren. Ohne diese Annahme ist aber zunächst einmal zu klären, was unter dem Begriff einer asymptotischen Normalverteilung verstanden werden soll.

Die einfachste Definition der asymptotischen Normalverteilung besagt, dass eine Folge von Zufallsvektoren $\boldsymbol{X}_n$ asymptotisch normalverteilt ist, wenn eine Kovarianzmatrix $\boldsymbol{\Sigma}$ existiert, so dass $\boldsymbol{X}_n \xrightarrow{\mathcal{D}} \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma})$ gilt. Dies ist, wenn die Folge von Kovarianzmatrizen $\boldsymbol{V}_n$ nicht konvergiert, aber nicht der Fall. Eine allgemeinere Definition von asymptotischer Normalverteilung, die auch für unsere Situation geeignet ist, können wir mit Hilfe der sogenannten *Prokhorov-Metrik* ρ (Prokhorov, 1956) angeben, die ein Abstandsmaß für Wahrscheinlichkeitsmaße darstellt. Eine Definition dieser Metrik findet sich zum Beispiel bei Dudley (1989) in Abschnitt 11.3. Für unsere Zwecke reichen die folgenden drei Lemmata zur Arbeit mit der Prokhorov-Metrik aus. Das erste Lemma erklärt, inwiefern Konvergenz bezüglich der Prokhorov-Metrik die schwache Konvergenz von Wahrscheinlichkeitsmaßen verallgemeinert. Die beiden anderen Lemmata zeigen, dass zwei der wichtigsten Hilfsmittel zur Arbeit mit der schwachen Konvergenz von Wahrscheinlichkeitsmaßen natürliche Verallgemeinerungen in Bezug auf die Prokhorov-Metrik besitzen.

Lemma 5.3. *Für ein Wahrscheinlichkeitsmaß P und eine Folge von Wahrscheinlichkeitsmaßen P_n auf dem $\mathbb{R}^k$ sind folgende Aussagen äquivalent:*

- $P_n \xrightarrow{w} P$
- $\rho(P_n, P) \to 0$

Beweis. Siehe Dudley (1989), Theorem 11.3.3. □

Das zweite Lemma ist eine Verallgemeinerung des Darstellungssatzes von Skorokhod und erklärt, was es bedeutet, wenn sich zwei Folgen von Wahrscheinlichkeitsmaßen bezüglich ρ beliebig nahe kommen.

Lemma 5.4. *Für zwei Folgen von Wahrscheinlichkeitsmaßen P_n und Q_n auf dem $\mathbb{R}^k$ sind folgende Aussagen äquivalent:*

- $\rho(P_n, Q_n) \to 0$
- *Es existieren Zufallsvektoren $\boldsymbol{X}_n \sim P_n$ und $\boldsymbol{Y}_n \sim Q_n$, die auf einem gemeinsamen Wahrscheinlichkeitsraum definiert sind und für die $\boldsymbol{X}_n - \boldsymbol{Y}_n \xrightarrow{P} \mathbf{0}$ gilt.*
- *Es existieren Zufallsvektoren $\boldsymbol{X}_n \sim P_n$ und $\boldsymbol{Y}_n \sim Q_n$, die auf einem gemeinsamen Wahrscheinlichkeitsraum definiert sind und für die $\boldsymbol{X}_n - \boldsymbol{Y}_n \xrightarrow{f.s.} \mathbf{0}$ gilt.*

Beweis. Siehe Dudley (1989), Theorem 11.7.1. □

Das dritte Lemma verallgemeinert den Satz von der stetigen Abbildung, der auch als Satz von Mann-Wald bekannt ist.

Lemma 5.5. *P_n und Q_n seien zwei Folgen von Wahrscheinlichkeitsmaßen auf dem $\mathbb{R}^k$ mit $\rho(P_n, Q_n) \to 0$. Für eine gleichmäßig stetige Abbildung $g : \mathbb{R}^k \to \mathbb{R}^d$ gilt dann auch $\rho(P_n \circ g^{-1}, Q_n \circ g^{-1}) \to 0$.*

Beweis. Die Behauptung ergibt sich aus dem vorangehenden Lemma. □

Wie man an den drei Lemmata 5.3 – 5.5 sieht, ermöglicht die Prokhorov-Metrik eine sinnvolle Verallgemeinerung des Begriffes der schwachen Konvergenz. Daher bietet es sich an, eine Folge von Zufallsvektoren $\boldsymbol{X}_n \sim P_n$ als asymptotisch normalverteilt anzusehen, wenn eine Folge von Kovarianzmatrizen $\boldsymbol{\Sigma}_n$ existiert, so dass $\rho[P_n, \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_n)] \to 0$ gilt. Zur Vereinfachung der Notation schreiben wir dafür auch $\rho[\boldsymbol{X}_n, \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_n)] \to 0$.

Mit dieser Definition und den Annahmen (A1) – (A3) aus Unterabschnitt 2.1.2 können wir nun die asymptotische Normalverteilung von $\sqrt{n}\,(\widehat{\boldsymbol{p}}_n - \boldsymbol{p})$ nachweisen.

Satz 5.6. *Unter den Voraussetzungen (A1) – (A3) gilt*

$$\rho\left[\sqrt{n}\,(\widehat{\boldsymbol{p}}_n - \boldsymbol{p})\,,\, \mathrm{N}(\boldsymbol{0}, \boldsymbol{V}_n)\right] \to 0.$$

Beweis. Siehe Anhang A.14, Seite 108. □

Korollar 5.7. *Unter den Voraussetzungen (A1) – (A3) gilt mit den Bezeichnungen $(\sigma^2_{11,n}, \ldots, \sigma^2_{at,n}) = \mathrm{diag}(\boldsymbol{V}_n)$, $\sigma^2_{i*,n} = \boldsymbol{\nu}'_{i*}\boldsymbol{V}_n\boldsymbol{\nu}_{i*}$ und $\sigma^2_{*s,n} = \boldsymbol{\nu}'_{*s}\boldsymbol{V}_n\boldsymbol{\nu}_{*s}$*

$$\begin{aligned}
\rho\left[\sqrt{n}(\widehat{p}_{is,n} - p_{is}),\, \mathrm{N}(0, \sigma^2_{is,n})\right] &\to 0, \\
\rho\left[\sqrt{n}(\widehat{p}_{i*,n} - p_{i*}),\, \mathrm{N}(0, \sigma^2_{i*,n})\right] &\to 0 \quad \textit{und} \\
\rho\left[\sqrt{n}(\widehat{p}_{*s,n} - p_{*s}),\, \mathrm{N}(0, \sigma^2_{*s,n})\right] &\to 0.
\end{aligned}$$

Beweis. Da die durch

$$\boldsymbol{x} = (x_{11}, \ldots, x_{at})' \mapsto x_{is}, \quad \boldsymbol{x} \mapsto \boldsymbol{\nu}'_{i*}\boldsymbol{x} \quad \text{und} \quad \boldsymbol{x} \mapsto \boldsymbol{\nu}'_{*s}\boldsymbol{x}$$

definierten Abbildungen mit Definitionsbereich $\mathbb{R}^{at}$ gleichmäßig stetig sind, ergibt sich die Behauptung durch Anwendung von Lemma 5.5 auf das Ergebnis aus Satz 5.6. □

Bemerkung. Satz 5.6 ist bewußt sehr allgemein gehalten. Er kann nicht nur zur Bestimmung der asymptotischen Verteilung von $\sqrt{n}(\widehat{p}_{is,n} - p_{is})$, $\sqrt{n}(\widehat{p}_{i*,n} - p_{i*})$ und $\sqrt{n}(\widehat{p}_{is,n} - p_{is})$ verwendet werden, sondern mit Hilfe von Lemma 5.5 kann für jede beliebige gleichmäßig stetige, vektorwertige Funktion die asymptotische Verteilung von $g[\sqrt{n}(\widehat{\boldsymbol{p}}_n - \boldsymbol{p})]$ ermittelt werden. Außer den in dieser Arbeit konstruierten Konfidenzintervallen für die relativen Effekte sind an weiteren Anwendungen zum Beispiel simultane Konfidenzintervalle für mehrere relative Effekte, Konfidenzintervalle für Minima oder Maxima von relativen Effekte sowie Tests für Hypothesen der Form $\boldsymbol{C}\boldsymbol{p} = \boldsymbol{0}$ mit $\boldsymbol{C} \in \mathbb{R}^{m \times at}$ auf der Basis von quadratischen Formen denkbar.

5.3 Varianzschätzer

In diesem Abschnitt befassen wir uns mit dem zweiten wichtigen Baustein für die Konstruktion der Konfidenzintervalle, dem konsistenten Schätzer für die Kovarianzmatrix $\boldsymbol{V}_n$.

Im vorigen Abschnitt haben wir die Bezeichnungen $\sigma^2_{uv,n}$ für die Diagonalelemente dieser Matrix verwendet. In diesem Abschnitt bezeichnen wir die Komponenten dieser Matrix einheitlich mit $\sigma_{uvu'v',n}$, um nicht zwischen verschiedenen Schreibweisen für die Varianzen und Kovarianzen unterscheiden zu müssen. Weil die Vektoren $\boldsymbol{\Psi}_{ik,n}$, $i = 1, \ldots, a$, $k = 1, \ldots, n_i$, unabhängig voneinander sind, gilt

$$\begin{aligned}\sigma_{uvu'v',n} &= \mathrm{Cov}\Bigg(\sqrt{n}\sum_{i=1}^{a}\frac{1}{n_i}\sum_{k=1}^{n_i}[\Psi_{ikuv,n} - \mathrm{E}(\Psi_{ikuv,n})],\\ &\qquad\qquad \sqrt{n}\sum_{i=1}^{a}\frac{1}{n_i}\sum_{k=1}^{n_i}[\Psi_{iku'v',n} - \mathrm{E}(\Psi_{iku'v',n})]\Bigg)\\ &= n\sum_{i=1}^{a}\frac{1}{n_i^2}\sum_{k=1}^{n_i}\mathrm{Cov}[\Psi_{ikuv,n} - \mathrm{E}(\Psi_{ikuv,n}), \Psi_{iku'v',n} - \mathrm{E}(\Psi_{iku'v',n})].\end{aligned}$$

Wegen

$$\begin{aligned}&\mathrm{Cov}(\Psi_{ikuv,n} - \mathrm{E}[\Psi_{ikuv,n}], \Psi_{iku'v',n} - \mathrm{E}[\Psi_{iku'v',n}])\\ &\qquad\qquad = \mathrm{E}[(\Psi_{ikuv,n} - \mathrm{E}[\Psi_{ikuv,n}])(\Psi_{iku'v',n} - \mathrm{E}[\Psi_{iku'v',n}])]\end{aligned}$$

ist

$$\widetilde{\sigma}_{uvu'v',n} = n\sum_{i=1}^{a}\frac{1}{n_i^2}\sum_{k=1}^{n_i}(\Psi_{ikuv,n} - \mathrm{E}[\Psi_{ikuv,n}])(\Psi_{iku'v',n} - \mathrm{E}[\Psi_{iku'v',n}]) \qquad (5.1)$$

ein erwartungstreuer und deshalb naheliegender Schätzer[1] für $\sigma_{uvu'v',n}$. Leider sind die Erwartungswerte $\mathrm{E}[\Psi_{ikuv,n}]$ und $\mathrm{E}[\Psi_{iku'v',n}]$ unbekannt. Wegen

$$\mathrm{E}[\Psi_{ikuv,n}] = \begin{cases} \frac{n_u}{\lambda_{u\cdot v}}\int H_{uv}\, d(\lambda_{ukv}F_{uv}) - \frac{1}{at-1}\sum_{s\neq v}\frac{n_u}{\lambda_{u\cdot s}}\int F_{uv}\, d(\lambda_{uks}F_{us}) & : u = i,\\ \frac{-1}{at-1}\sum_{s=1}^{t}\frac{n_i}{\lambda_{i\cdot s}}\int F_{uv}\, d(\lambda_{iks}F_{is}) & : u \neq i,\end{cases}$$

können wir durch

$$\check{\mu}_{ikuv,n} = \begin{cases} \frac{n_u}{\lambda_{u\cdot v}}\int H_{uv}\, d\left(\lambda_{ukv}\widehat{F}_{uv,n}\right) - \frac{1}{at-1}\sum_{s\neq v}\frac{n_u}{\lambda_{u\cdot s}}\int F_{uv}\, d\left(\lambda_{uks}\widehat{F}_{us,n}\right) & : u = i,\\ \frac{-1}{at-1}\sum_{s=1}^{t}\frac{n_i}{\lambda_{i\cdot s}}\int F_{uv}\, d\left(\lambda_{iks}\widehat{F}_{is,n}\right) & : u \neq i,\end{cases}$$

erwartungstreue Schätzer für diese Erwartungswerte definieren. Ersetzen wir in (5.1) die Erwartungswerte durch diese Schätzer und $1/n_i^2$ durch $1/(n_i(n_i-1))$, so erhalten wir mit

$$\check{\sigma}_{uvu'v',n} = n\sum_{i=1}^{a}\frac{1}{n_i(n_i-1)}\sum_{k=1}^{n_i}(\Psi_{ikuv,n} - \check{\mu}_{ikuv,n})(\Psi_{iku'v',n} - \check{\mu}_{iku'v',n})$$

[1] Der Begriff Schätzer ist in diesem Abschnitt nicht im Sinne seiner mathematischen Definition gemeint. Diese verlangt nämlich von einem Schätzer, dass er nicht von unbekannten Größen wie Verteilungfunktionen abhängt.

einen weiteren Schätzer für $\sigma_{uvu'v',n}$. Dieser ist zwar im Allgemeinen nicht mehr erwartungstreu, jedoch in dem Spezialfall, dass für jede Gruppe i alle Dimensionsvektoren $\boldsymbol{m}_{ik}$, $k = 1, \ldots, n_i$, identisch sind.

Aber auch diesen Schätzer können wir noch nicht verwenden, da er die unbekannten Verteilungsfunktionen F_{is} und H_{is} enthält. Ersetzen wir diese durch die entsprechenden empirischen Verteilungsfunktionen, so ergibt sich schließlich als praktisch verwendbarer Schätzer

$$\widehat{\sigma}_{uvu'v',n} = n \sum_{i=1}^{a} \frac{1}{n_i(n_i - 1)} \sum_{k=1}^{n_i} \left(\widehat{\Psi}_{ikuv,n} - \widehat{\mu}_{ikuv,n}\right) \left(\widehat{\Psi}_{iku'v',n} - \widehat{\mu}_{iku'v',n}\right), \tag{5.2}$$

wobei wir $\widehat{\Psi}_{ikuv,n}$ und $\widehat{\mu}_{ikuv,n}$ durch

$$\widehat{\Psi}_{ikuv,n} = \begin{cases} \frac{n_u}{\lambda_{u\cdot v}} \int \widehat{H}_{uv,n} \, d\widehat{F}_{ukv} - \frac{1}{at-1} \sum_{s \neq v} \frac{n_u}{\lambda_{u\cdot s}} \int \widehat{F}_{uv,n} \, d\widehat{F}_{uks} & : u = i, \\ \frac{-1}{at-1} \sum_{s=1}^{t} \frac{n_i}{\lambda_{i\cdot s}} \int \widehat{F}_{uv,n} \, d\widehat{F}_{iks} & : u \neq i, \end{cases}$$

und

$$\widehat{\mu}_{ikuv,n} = \begin{cases} \frac{n_u}{\lambda_{u\cdot v}} \int \widehat{H}_{uv,n} \, d\left(\lambda_{ukv}\widehat{F}_{uv,n}\right) - \frac{1}{at-1} \sum_{s \neq v} \frac{n_u}{\lambda_{u\cdot s}} \int \widehat{F}_{uv,n} \, d\left(\lambda_{uks}\widehat{F}_{us}\right) & : u = i, \\ \frac{-1}{at-1} \sum_{s=1}^{t} \frac{n_i}{\lambda_{i\cdot s}} \int \widehat{F}_{uv,n} \, d\left(\lambda_{iks}\widehat{F}_{is}\right) & : u \neq i, \end{cases}$$

definieren.

Zur praktischen Berechnung von $\widehat{\Psi}_{ikuv,n}$ und $\widehat{\mu}_{ikuv,n}$ kann man diese zunächst durch eine Linearkombination der Zufallsvariablen $\widehat{F}_{uv}(X_{iksr})$ darstellen, so wie das in Abschnitt 4.5 schon für $\widehat{\boldsymbol{p}}_n$ durchgeführt worden ist. Die $\widehat{F}_{uv}(X_{iksr})$ können wir dann – wie in Abschnitt 4.5 beschrieben – mittels einer Rangdarstellung berechnen. Die hier vorgeschlagene Vorgehensweise wurde bei den beiden in Anhang B beschriebenen SAS-Makros `RE_CI` und `SIMPLE_RE_CI` umgesetzt.

Obwohl zur Herleitung der in (5.2) angegebenen Kovarianz-Schätzer einige Schritte notwendig waren, sind diese Schätzer konsistent:

Satz 5.8. *Unter den Voraussetzungen (A1) und (A2) auf Seite 13 ist die Matrix* $\widehat{\boldsymbol{V}}_n$*, deren Komponenten* $\widehat{\sigma}_{uvu'v',n}$ *durch* (5.2) *definiert sind, ein konsistenter Schätzer für* $\boldsymbol{V}_n$ *in dem Sinne, dass*

$$\widehat{\boldsymbol{V}}_n - \boldsymbol{V}_n \xrightarrow{f.s.} \boldsymbol{0}$$

und

$$\sum_{u=1}^{a} \sum_{v=1}^{t} \sum_{u'=1}^{a} \sum_{v'=1}^{t} \mathrm{E}\left[(\widehat{\sigma}_{uvu'v',n} - \sigma_{uvu'v',n})^2\right] \to 0.$$

Beweis. Siehe Anhang A.15, Seite 111. □

Bemerkung. Im Zusammenhang mit fehlenden Werten kann es in anderen Situationen als der unseren vorkommen, dass Schätzer für Kovarianzmatrizen nicht positiv semidefinit sind oder die Cauchy-Schwarz-Ungleichung nicht erfüllen (siehe z. B. Rousseeuw und Molenberghs, 1993). Dieses Problem tritt hier nicht auf.

Korollar 5.9. *Unter den Voraussetzungen (A1) und (A2) sind die Schätzer*

$$\widehat{\sigma}^2_{is,n} = \widehat{\sigma}_{isis,n}, \quad \widehat{\sigma}^2_{i*,n} = \boldsymbol{\lambda}'_{i*}\widehat{\boldsymbol{V}}_n\boldsymbol{\lambda}_{i*} \quad \textit{und} \quad \widehat{\sigma}^2_{*s,n} = \boldsymbol{\lambda}'_{*s}\widehat{\boldsymbol{V}}_n\boldsymbol{\lambda}_{*s}$$

konsistent in dem Sinne, dass

$$\widehat{\sigma}^2_{is,n} - \sigma^2_{is,n} \xrightarrow{f.s.} 0, \quad \widehat{\sigma}^2_{i*,n} - \sigma^2_{i*,n} \xrightarrow{f.s.} 0 \quad \textit{und} \quad \widehat{\sigma}^2_{*s,n} - \sigma^2_{*s,n} \xrightarrow{f.s.} 0$$

sowie

$$\mathrm{E}\left[\left(\widehat{\sigma}^2_{is,n} - \sigma^2_{is,n}\right)^2\right] \to 0, \ \mathrm{E}\left[\left(\widehat{\sigma}^2_{i*,n} - \sigma^2_{i*,n}\right)^2\right] \to 0 \ \textit{und} \ \mathrm{E}\left[\left(\widehat{\sigma}^2_{*s,n} - \sigma^2_{*s,n}\right)^2\right] \to 0.$$

Beweis. Folgt aus dem voranstehenden Satz. □

5.4 Normal-Approximation

Die asymptotische Gültigkeit der Konfidenzintervalle für die relativen Effekte p_{is} in diesem und den folgenden zwei Abschnitten beruhen auf dem Resultat

$$\sqrt{n}\,\frac{\widehat{p}_{is,n} - p_{is}}{\widehat{\sigma}_{is,n}} \xrightarrow{\mathcal{D}} \mathrm{N}(0,\,1).$$

Dieses ist aber selbst unter allen bisher benutzten Voraussetzungen an die Stichprobenumfänge im Allgemeinen nicht richtig. Es wird eine weitere Voraussetzung benötigt, die aber keine Voraussetzung an die Folge der Stichprobenumfänge, sondern eine Annahme bezüglich der Varianzen $\sigma^2_{is,n}$ ist. Entsprechende Annahmen sind auch für die asymptotische Gültigkeit der Konfidenzintervalle für die Gruppen- und Bedingungseffekte p_{i*} und p_{*s} zu treffen.

(A4) $\liminf_{n\to\infty} \sigma^2_{is,n} > 0$

(A5) $\liminf_{n\to\infty} \sigma^2_{i*,n} > 0$

(A6) $\liminf_{n\to\infty} \sigma^2_{*s,n} > 0$

Satz 5.10. *Es seien die Voraussetzungen (A1) – (A3) auf Seite 13 erfüllt. Falls darüber hinaus eine der Voraussetzungen (A4), (A5) oder (A6) zutrifft, so gilt entsprechend*

$$\begin{aligned}
T_{is,n} &= \sqrt{n}\,\frac{\widehat{p}_{is,n} - p_{is}}{\widehat{\sigma}_{is,n}} &\xrightarrow{\mathcal{D}}&\ \mathrm{N}(0,\,1),\\
T_{i*,n} &= \sqrt{n}\,\frac{\widehat{p}_{i*,n} - p_{i*}}{\widehat{\sigma}_{i*,n}} &\xrightarrow{\mathcal{D}}&\ \mathrm{N}(0,\,1) \quad \textit{oder}\\
T_{*s,n} &= \sqrt{n}\,\frac{\widehat{p}_{*s,n} - p_{*s}}{\widehat{\sigma}_{*s,n}} &\xrightarrow{\mathcal{D}}&\ \mathrm{N}(0,\,1).
\end{aligned}$$

Beweis. Siehe Anhang A.16, Seite 116. □

Korollar 5.11. *Es seien die Voraussetzungen (A1) – (A3) erfüllt. Falls darüber hinaus eine der Voraussetzungen (A4), (A5) oder (A6) zutrifft, so ist entsprechend für alle* $\alpha \in (0, 1)$

$$\begin{aligned}
\left[u^N_{is,n},\, o^N_{is,n}\right] &= \left[\widehat{p}_{is,n} - \frac{\widehat{\sigma}_{is,n} z_{1-\alpha/2}}{\sqrt{n}},\, \widehat{p}_{is,n} + \frac{\widehat{\sigma}_{is,n} z_{1-\alpha/2}}{\sqrt{n}}\right], \\
\left[u^N_{i*,n},\, o^N_{i*,n}\right] &= \left[\widehat{p}_{i*,n} - \frac{\widehat{\sigma}_{i*,n} z_{1-\alpha/2}}{\sqrt{n}},\, \widehat{p}_{i*,n} + \frac{\widehat{\sigma}_{i*,n} z_{1-\alpha/2}}{\sqrt{n}}\right] \quad \text{oder} \\
\left[u^N_{*s,n},\, o^N_{is,n}\right] &= \left[\widehat{p}_{*s,n} - \frac{\widehat{\sigma}_{*s,n} z_{1-\alpha/2}}{\sqrt{n}},\, \widehat{p}_{*s,n} + \frac{\widehat{\sigma}_{*s,n} z_{1-\alpha/2}}{\sqrt{n}}\right]
\end{aligned}$$

ein asymptotisches $(1-\alpha)$*-Konfidenzintervall für* p_{is}*,* p_{i*} *oder* p_{*s}*.*

Beweis. Wegen Satz 5.10 gilt

$$\mathrm{P}\left(p_{is} \in \left[u^N_{is,n},\, o^N_{is,n}\right]\right) = \mathrm{P}\left(z_{\alpha/2} \leq \sqrt{n}\,\frac{\widehat{p}_{is,n} - p_{is}}{\widehat{\sigma}_{is,n}} \leq z_{1-\alpha/2}\right) \to 1 - \alpha.$$

Die anderen beiden Aussagen lassen sich analog nachweisen. □

5.5 *t*-Approximation

Im vorangehenden Abschnitt haben wir auf der Basis von Satz 5.10 eine Approximation der Verteilungen der Statistiken $T_{is,n}$, $T_{i*,n}$ und $T_{*s,n}$ durch eine Standard-Normalverteilung zur Konstruktion von Konfidenzintervallen für die relativen Effekte verwendet. Dass diese Konfidenzintervalle unter geeigneten Voraussetzungen asymptotisch das vorgegebene Niveau einhalten, entnehmen wir Korollar 5.11. Wir können aber vermuten, dass die Konfidenzintervalle bei kleinen Stichprobenumfängen das vorgegebene Niveau nicht einhalten. Denn selbst bei normalverteilten Daten

$$X_1, \ldots, X_n \text{ u. i. v. } \sim \mathrm{N}(\mu,\, \sigma^2)$$

hält das analog konstruierte Konfidenzintervall

$$\left[\widehat{\mu} - \frac{\widehat{\sigma} z_{1-\alpha/2}}{\sqrt{n}},\, \widehat{\mu} + \frac{\widehat{\sigma} z_{1-\alpha/2}}{\sqrt{n}}\right]$$

das vorgegebene Niveau nicht ein. Bei normalverteilten Daten kann man jedoch ein exaktes Konfidenzintervall erhalten, indem man an Stelle der Quantile der Normalverteilung $z_{1-\alpha/2}$ die Quantile $t_{n-1,\,1-\alpha/2}$ einer t-Verteilung mit $n-1$ Freiheitsgraden verwendet. Durch diesen Sachverhalt motiviert leiten wir auf heuristischer Basis Approximationen der Verteilungen der Statistiken $T_{is,n}$, $T_{i*,n}$ und $T_{*s,n}$ durch t-Verteilungen her und ersetzen die Quantile der Standard-Normalverteilung durch die der entsprechenden t-Verteilung. Da die Quantile $t_{f,1-\alpha/2}$ jeder t_f-Verteilung größer als $z_{1-\alpha/2}$ sind, können wir uns so aus einer konservativen Sichtweise nur verbessern. Die genannte t-Approximation scheint – wie sich noch herausstellen wird – allerdings nur in einem Spezialfall des allgemeinen Versuchsplans aus Abschnitt 2.1

sinnvoll zu sein: Wir können sie nur dann motivieren, wenn für alle Gruppen i die Dimensionsvektoren $\boldsymbol{m}_{ik}$ für alle k identisch sind.

Die zentrale Idee der t-Approximation ist die Approximation der Verteilung einer Linearkombination von unabhängigen χ^2-verteilten Zufallsvariablen durch eine gestreckte χ^2-Verteilung. Sie geht auf Arbeiten von Smith (1936), Welch (1936, 1938), Satterthwaite (1941), Patnaik (1949) und Box (1954) zurück und wurde in der nichtparametrischen Statistik zum Beispiel von Brunner et al. (1997) und Brunner und Munzel (2000) verwendet.

Wir leiten das Approximationsverfahren für $T_{is,n}$ ausführlich her. Für $T_{i*,n}$ und $T_{*s,n}$ verläuft die Herleitung analog, so dass wir uns für diese Statistiken auf die Nennung der Ergebnisse beschränken.

Die Statistik $T_{is,n}$ läßt sich darstellen als

$$T_{is,n} = \frac{\sqrt{n}\,\frac{\widehat{p}_{is,n}-p_{is}}{\sigma_{is,n}}}{\widehat{\sigma}_{is,n}/\sigma_{is,n}}.$$

Für den Zähler und den Nenner dieses Bruchs gelten unter den Voraussetzungen (A1) – (A4) (siehe Seite 13 und Seite 65)

$$\left(\sqrt{n}\,\frac{\widehat{p}_{is,n}-p_{is}}{\sigma_{is,n}},\ \frac{\widehat{\sigma}_{is,n}}{\sigma_{is,n}}\right) \xrightarrow{\mathcal{D}} N(0,1)\otimes\delta_1,$$

wobei wir mit $\otimes$ das Produkt zweier Maße und mit δ_1 das Einpunkt-Maß im Punkt 1 bezeichnen. Zähler und Nenner der Statistik sind also asymptotisch unabhängig, und der Zähler ist asymptotisch standardnormalverteilt. Um die Verteilung von $T_{is,n}$ durch eine $t_{f_{is,n}}$-Verteilung zu approximieren, müssen wir also die Verteilung des Nenners durch eine $\sqrt{\chi^2_{f_{is,n}}/f_{is,n}}$-Verteilung[2] annähern.

Wir approximieren zunächst die Verteilung von $\widehat{\sigma}^2_{is,n}$ durch eine $g_{is,n}\chi^2_{f_{is,n}}/f_{is,n}$-Verteilung. Um die Situation zu vereinfachen, ersetzen wir im folgenden den Schätzer $\widehat{\sigma}^2_{is,n}$ durch die – wie wir im Beweis von Satz 5.8 gesehen haben – asymptotisch äquivalente Zufallsvariable $\check{\sigma}^2_{is,n} = \check{\sigma}_{isis,n}$. Es gilt

$$\check{\sigma}^2_{is,n} = n\sum_{u=1}^{a}\frac{1}{n_u(n_u-1)}\sum_{w=1}^{n_u}(\Psi_{uwis,n}-\check{\mu}_{uwis,n})^2.$$

In dem hier betrachteten Spezialfall, dass für alle Gruppen u die Dimensionsvektoren $\boldsymbol{m}_{uw}$ für alle w übereinstimmen, sind die Zufallsvariablen $\Psi_{u1is,n},\ldots,\Psi_{un_uis,n}$ nicht nur unabhängig, sondern auch noch identisch verteilt. Darüber hinaus stimmen die Zufallsvariablen $\check{\mu}_{uwis,n}$ für $w\in\{1,\ldots,n_u\}$ mit dem arithmetischen Mittel der $\Psi_{uwis,n}$ überein, das heißt,

$$\check{\mu}_{uwis,n} = \overline{\Psi}_{u\cdot is,n} = \frac{1}{n_u}\sum_{w=1}^{n_u}\Psi_{uwis,n}.$$

[2]Wir schreiben hier $\sqrt{\chi^2_{f_{is,n}}/f_{is,n}}$ für die Verteilung einer Zufallsvariablen $\sqrt{Z/f_{is,n}}$ mit $Z\sim\chi^2_{f_{is,n}}$. Im Folgenden werden analoge abkürzende Schreibweisen verwendet.

Aus diesem Grund ist

$$\check{\sigma}^2_{uis,n} = \frac{1}{n_u - 1} \sum_{w=1}^{n_u} (\Psi_{uwis,n} - \check{\mu}_{uwis,n})^2 = \frac{1}{n_u - 1} \sum_{w=1}^{n_u} \left(\Psi_{uwis,n} - \overline{\Psi}_{u\cdot is,n}\right)^2$$

ein erwartungstreuer und konsistenter Schätzer für $\sigma^2_{uis,n} = \mathrm{Var}(\Psi_{uwis,n})$. Wären die Zufallsvariablen $\Psi_{u1is,n}, \ldots, \Psi_{un_uis,n}$ nicht nur unabhängig identisch verteilt, sondern auch noch normalverteilt, so würde

$$\check{\sigma}^2_{uis,n} \sim \frac{\sigma^2_{uis,n}\chi^2_{n_u-1}}{n_u - 1} \tag{5.3}$$

gelten. Da hier für die $\Psi_{uwis,n}$ lediglich die Voraussetzung der Normalverteilung nicht erfüllt ist, wollen wir bei der weiteren Approximation die Gültigkeit von (5.3) behelfsmäßig annehmen. Unter dieser Annahme können wir eine Verteilungsaussage für $\check{\sigma}^2_{is,n}$ machen:

$$\check{\sigma}^2_{is,n} = n \sum_{u=1}^{a} \frac{\check{\sigma}^2_{uis,n}}{n_u} \quad \sim \quad n \sum_{u=1}^{a} \frac{\sigma^2_{uis,n}\chi^2_{n_u-1}}{n_u(n_u - 1)}.$$

Um nun diese Verteilung durch eine $g_{is,n}\chi^2_{f_{is,n}}/f_{is,n}$-Verteilung zu approximieren, bestimmen wir $g_{is,n}$ und $f_{is,n}$ so, dass Erwartungswert und Varianz der beiden Verteilungen übereinstimmen. Wir erhalten das Gleichungssystem

$$\begin{aligned} \mathrm{E}\left(n \sum_{u=1}^{a} \frac{\sigma^2_{uis,n}\chi^2_{n_u-1}}{n_u(n_u-1)}\right) &= \sigma^2_{is,n} &&= g_{is,n} = \mathrm{E}\left(\frac{g_{is,n}\chi^2_{f_{is,n}}}{f_{is,n}}\right), \\ \mathrm{Var}\left(n \sum_{u=1}^{a} \frac{\sigma^2_{uis,n}\chi^2_{n_u-1}}{n_u(n_u-1)}\right) &= 2n^2 \sum_{u=1}^{a} \frac{\sigma^4_{uis,n}}{n_u^2(n_u-1)} &&= \frac{2g^2_{is,n}}{f_{is,n}} = \mathrm{Var}\left(\frac{g_{is,n}\chi^2_{f_{is,n}}}{f_{is,n}}\right) \end{aligned}$$

mit der Lösung

$$g_{is,n} = \sigma^2_{is,n} \quad \text{und} \quad f_{is,n} = \sigma^4_{is,n} \Big/ n^2 \sum_{u=1}^{a} \frac{\sigma^4_{uis,n}}{n_u^2(n_u-1)}.$$

Nach unseren Überlegungen sollte also näherungsweise $\widehat{\sigma}^2_{is,n}/\sigma^2_{is,n} \sim \chi^2_{f_{is,n}}/f_{is,n}$ und damit ungefähr $T_{is,n} \sim t_{f_{is,n}}$ gelten. Da nach Annahme (A4) $f_{is,n} \to \infty$ und deshalb $t_{f_{is,n}} \xrightarrow{\mathrm{w}} \mathrm{N}(0,1)$ gilt, befindet sich die hergeleitete Approximation im Einklang mit Satz 5.10.

Den Freiheitsgrad $f_{is,n}$ schätzen wir kanonisch durch

$$\widehat{f}_{is,n} = \widehat{\sigma}^4_{is,n} \Big/ n^2 \sum_{u=1}^{a} \frac{\widehat{\sigma}^4_{uis,n}}{n_u^2(n_u-1)}, \tag{5.4}$$

wobei die $\widehat{\sigma}^2_{uis,n}$ durch

$$\widehat{\sigma}^2_{uis,n} = \frac{1}{n_u - 1} \sum_{w=1}^{n_u} \left(\widehat{\Psi}_{uwis,n} - \widehat{\mu}_{uwis,n}\right)^2 = \frac{1}{n_u - 1} \sum_{w=1}^{n_u} \left(\widehat{\Psi}_{uwis,n} - \overline{\widehat{\Psi}}_{u\cdot is,n}\right)^2$$

definiert sind. Im Beweis von Satz 5.8 haben wir gezeigt, dass der Schätzer $\widehat{\sigma}^2_{uis,n}$ unter den Annahmen (A1) und (A2) konsistent für $\sigma^2_{uis,n}$ ist:

$$\widehat{\sigma}^2_{uis,n} - \sigma^2_{uis,n} \xrightarrow{\text{f.s.}} 0.$$

Unter der zusätzlichen Voraussetzung (A4) sind auch die geschätzten Freiheitsgrade konsistent für $f_{is,n}$ in dem Sinne, dass $\widehat{f}_{is,n} \xrightarrow{\text{f.s.}} \infty$.

Wir fassen unsere Ergebnisse zusammen und ergänzen sie bezüglich p_{i*} und p_{*s}.

Approximationsverfahren 5.12. *Falls für alle Gruppen $i = 1, \ldots, a$ die Dimensionsvektoren $\boldsymbol{m}_{ik}$, $k = 1, \ldots, n_i$, identisch sind, so erfüllen die Konfidenzintervalle*

$$\begin{aligned}
\left[u^t_{is,n},\, o^t_{is,n}\right] &= \left[\widehat{p}_{is,n} - \frac{\widehat{\sigma}_{is,n} t_{\widehat{f}_{is,n},1-\alpha/2}}{\sqrt{n}},\, \widehat{p}_{is,n} + \frac{\widehat{\sigma}_{is,n} t_{\widehat{f}_{is,n},1-\alpha/2}}{\sqrt{n}}\right], \\
\left[u^t_{i*,n},\, o^t_{i*,n}\right] &= \left[\widehat{p}_{i*,n} - \frac{\widehat{\sigma}_{i*,n} t_{\widehat{f}_{i*,n},1-\alpha/2}}{\sqrt{n}},\, \widehat{p}_{i*,n} + \frac{\widehat{\sigma}_{i*,n} t_{\widehat{f}_{i*,n},1-\alpha/2}}{\sqrt{n}}\right] \quad \textit{und} \\
\left[u^t_{*s,n},\, o^t_{is,n}\right] &= \left[\widehat{p}_{*s,n} - \frac{\widehat{\sigma}_{*s,n} t_{\widehat{f}_{*s,n},1-\alpha/2}}{\sqrt{n}},\, \widehat{p}_{*s,n} + \frac{\widehat{\sigma}_{*s,n} t_{\widehat{f}_{*s,n},1-\alpha/2}}{\sqrt{n}}\right]
\end{aligned}$$

für alle $\alpha \in (0, 1)$ bei hinreichend großen Stichprobenumfängen näherungsweise

$$\begin{aligned}
\mathrm{P}\left(p_{is} \in \left[u^t_{is,n},\, o^t_{is,n}\right]\right) &= 1 - \alpha, \\
\mathrm{P}\left(p_{i*} \in \left[u^t_{i*,n},\, o^t_{i*,n}\right]\right) &= 1 - \alpha \quad \textit{und,} \\
\mathrm{P}\left(p_{*s} \in \left[u^t_{*s,n},\, o^t_{*s,n}\right]\right) &= 1 - \alpha.
\end{aligned}$$

Die Freiheitsgrade sind dabei durch (5.4) *beziehungsweise durch*

$$\widehat{f}_{i*,n} = \widehat{\sigma}^4_{i*,n} \Big/ n^2 \sum_{u=1}^{a} \frac{\widehat{\sigma}^4_{ui*,n}}{n_u^2(n_u - 1)} \quad \textit{und} \quad \widehat{f}_{*s,n} = \widehat{\sigma}^4_{*s,n} \Big/ n^2 \sum_{u=1}^{a} \frac{\widehat{\sigma}^4_{u*s,n}}{n_u^2(n_u - 1)}$$

gegeben, wobei die $\widehat{\sigma}^2_{ui*,n}$ und die $\widehat{\sigma}^2_{u*s,n}$ durch

$$\begin{aligned}
\widehat{\sigma}^2_{ui*,n} &= \frac{(at-1)^2}{(at^2 - t^2)^2(n_u - 1)} \sum_{w=1}^{n_u} \left(\widehat{\Psi}_{uwi\cdot,n} - \widehat{\mu}_{uwi\cdot,n}\right)^2, \\
\widehat{\sigma}^2_{u*s,n} &= \frac{(at-1)^2}{(a^2t - a^2)^2(n_u - 1)} \sum_{w=1}^{n_u} \left(\widehat{\Psi}_{uw\cdot s,n} - \widehat{\mu}_{uw\cdot s,n}\right)^2, \\
\widehat{\Psi}_{uwi\cdot,n} &= \sum_{s=1}^{t} \widehat{\Psi}_{uwis,n}\,, \quad \widehat{\mu}_{uwi\cdot,n} = \sum_{s=1}^{t} \widehat{\mu}_{uwis,n}, \\
\widehat{\Psi}_{uw\cdot s,n} &= \sum_{i=1}^{a} \widehat{\Psi}_{uwis,n} \quad \textit{und} \quad \widehat{\mu}_{uw\cdot s,n} = \sum_{i=1}^{a} \widehat{\mu}_{uwis,n}
\end{aligned}$$

definiert sind.

Bemerkung. Falls alle Versuchseinheiten einer Gruppe angehören ($a = 1$), so ergibt sich erwartungsgemäß $\widehat{f}_{is,n} = f_{1s,n} = n - 1$. Für zwei unverbundene Stichproben ($a = 2$, $t = 1$, alle $m_{ik1} = 1$) stimmen die hergeleiteten Freiheitsgrade mit denen aus der Gleichung (5.9) in Brunner und Munzel (2000) überein.

5.6 Transformationsmethode

Die in den vorangehenden beiden Abschnitten vorgestellten Konfidenzintervalle haben einen wesentlichen Nachteil: Obwohl die relativen Effekte wie auch die dazugehörigen Punktschätzer Elemente des Einheitsintervalls $[0,1]$ sind, überdecken die $(1-\alpha)$-Konfidenzintervalle für hinreichend kleines α auch Punkte außerhalb des Einheitsintervalls. Sie werden also größer, ohne zusätzliche Information zu enthalten. Der gleichen Problematik steht man gegenüber, wenn man die Normalverteilung zur Approximation der Binomialverteilung verwendet und auf dieser Basis ein Konfidenzintervall für die Erfolgswahrscheinlichkeit $p \in [0,1]$ konstruiert.

In diesem und dem folgenden Abschnitt leiten wir Konfidenzintervalle für relative Effekte her, die diesen Nachteil nicht aufweisen, die also immer im Einheitsintervall enthalten sind. Konfidenzintervalle mit dieser Eigenschaft werden von Efron und Tibshirani (1993, Abschnitt 13.6) als *bereichserhaltend* bezeichnet.

Die zunächst vorgestellte sogenannte *Transformationsmethode* beruht auf folgendem einfachen Prinzip: Das offene Einheitsintervall kann durch umkehrbare Funktionen $g : (0,1) \to \mathbb{R}$ auf die Menge aller reellen Zahlen abgebildet werden. Die Bilder von $\widehat{p}_{is,n}$ und p_{is} unter g bezeichnen wir im Folgenden kurz mit $\widehat{p}^*_{is,n}$ und p^*_{is}. Unter bestimmten Voraussetzungen an die Bijektion g besitzt die Statistik

$$\sqrt{n}\left[\widehat{p}^*_{is,n} - p^*_{is,n}\right] = \sqrt{n}\left[g(\widehat{p}_{is,n}) - g\left(p_{is,n}\right)\right]$$

wieder eine asymptotische Normalverteilung, so dass wir ein asymptotisches Konfidenzintervall $\left[u^*_{is,n}, o^*_{is,n}\right]$ für die mit g transformierten relativen Effekte p^*_{is} angeben können. Wegen der Umkehrbarkeit von g ist dann $g^{-1}\left(\left[u^*_{is,n}, o^*_{is,n}\right]\right)$ ein asymptotischer Konfidenzbereich für p_{is}, der im Einheitsintervall enthalten ist. Damit $g^{-1}\left(\left[u^*_{is,n}, o^*_{is,n}\right]\right)$ wieder ein Intervall $\left[u^T_{is,n}, o^T_{is,n}\right]$ ist, sollte die Transformation streng monoton sein. Die soeben skizzierte Methode lässt sich in gleicher Weise auf die Effekte p_{i*} und p_{*s} anwenden. Sie ist anschaulich in Abbildung 5.1 dargestellt.

Der folgende Satz bildet die Grundlage zur Konstruktion von Konfidenzintervallen für die transformierten Effekte und schränkt unsere Wahl der Funktion g weiter ein.

$$\begin{array}{ccccccc} (0,1) & \ni & p_{is} & \xrightarrow{g} & p^*_{is} & \in & \mathbb{R} \\ &&&& \downarrow && \\ (0,1) & \supset & \left[u^T_{is,n}, o^T_{is,n}\right] & \xleftarrow{g^{-1}} & \left[u^*_{is,n}, o^*_{is,n}\right] & \subset & \mathbb{R} \end{array}$$

Abbildung 5.1: Schematische Darstellung der Transformationsmethode.

Satz 5.13. *Es seien die Voraussetzungen (A1) – (A3) auf Seite 13 erfüllt und die Abbildung* $g : (0,1) \to \mathbb{R}$ *sei stetig differenzierbar mit* $g'(x) \neq 0$ *für alle* $x \in (0,1)$*. Falls darüber hinaus eine der Annahmen (A4), (A5) oder (A6) zutrifft, so gilt entsprechend*

$$\begin{aligned}
\sqrt{n}\,\frac{\widehat{p}^*_{is,n} - p^*_{is}}{\widehat{\sigma}^*_{is,n}} &\xrightarrow{\mathcal{D}} \mathrm{N}(0,\,1), \\
\sqrt{n}\,\frac{\widehat{p}^*_{i*,n} - p^*_{i*}}{\widehat{\sigma}^*_{i*,n}} &\xrightarrow{\mathcal{D}} \mathrm{N}(0,\,1) \quad \textit{oder} \\
\sqrt{n}\,\frac{\widehat{p}^*_{*s,n} - p^*_{*s}}{\widehat{\sigma}^*_{*s,n}} &\xrightarrow{\mathcal{D}} \mathrm{N}(0,\,1),
\end{aligned}$$

wobei $\widehat{\sigma}^*_{is,n}$, $\widehat{\sigma}^*_{i*,n}$ *und* $\widehat{\sigma}^*_{*s,n}$ *durch*

$$\widehat{\sigma}^*_{is,n} = g'(\widehat{p}_{is,n})\,\widehat{\sigma}_{is,n}, \quad \widehat{\sigma}^*_{i*,n} = g'(\widehat{p}_{i*,n})\,\widehat{\sigma}_{i*,n} \quad \textit{und} \quad \widehat{\sigma}^*_{*s,n} = g'(\widehat{p}_{*s,n})\,\widehat{\sigma}_{*s,n}$$

definiert sind.

Beweis. Siehe Anhang A.17, Seite 116. □

Korollar 5.14. *Es seien die Voraussetzungen (A1) – (A3) erfüllt und die Abbildung* $g : (0,1) \to \mathbb{R}$ *sei streng monoton und stetig differenzierbar mit* $g'(x) \neq 0$ *für alle* $x \in (0,1)$*. Falls darüber hinaus eine der Annahmen (A4), (A5) oder (A6) zutrifft, so sind entsprechend für alle* $\alpha \in (0,\,1)$

$$\begin{aligned}
\left[u^T_{is,n},\, o^T_{is,n}\right] &= g^{-1}\left(\left[\widehat{p}^*_{is,n} - \frac{\widehat{\sigma}^*_{is,n} z_{1-\alpha/2}}{\sqrt{n}},\, \widehat{p}^*_{is,n} + \frac{\widehat{\sigma}^*_{is,n} z_{1-\alpha/2}}{\sqrt{n}}\right]\right) \subset (0,1), \\
\left[u^T_{i*,n},\, o^T_{i*,n}\right] &= g^{-1}\left(\left[\widehat{p}^*_{i*,n} - \frac{\widehat{\sigma}^*_{i*,n} z_{1-\alpha/2}}{\sqrt{n}},\, \widehat{p}^*_{i*,n} + \frac{\widehat{\sigma}^*_{i*,n} z_{1-\alpha/2}}{\sqrt{n}}\right]\right) \subset (0,1) \quad \textit{oder} \\
\left[u^T_{*s,n},\, o^T_{*s,n}\right] &= g^{-1}\left(\left[\widehat{p}^*_{*s,n} - \frac{\widehat{\sigma}^*_{*s,n} z_{1-\alpha/2}}{\sqrt{n}},\, \widehat{p}^*_{*s,n} + \frac{\widehat{\sigma}^*_{*s,n} z_{1-\alpha/2}}{\sqrt{n}}\right]\right) \subset (0,1)
\end{aligned}$$

asymptotische $(1-\alpha)$*-Konfidenzintervalle für* p_{is}, p_{i*} *oder* p_{*s}.

Beweis. Wegen der strengen Monotonie und Umkehrbarkeit von g sowie Satz 5.13 gilt

$$\begin{aligned}
\mathrm{P}\left(p_{is} \in \left[u^T_{is,n},\, o^T_{is,n}\right]\right) &= \mathrm{P}\left(p^*_{is} \in \left[\widehat{p}^*_{is,n} - \frac{\widehat{\sigma}^*_{is,n} z_{1-\alpha/2}}{\sqrt{n}},\, \widehat{p}^*_{is,n} + \frac{\widehat{\sigma}^*_{is,n} z_{1-\alpha/2}}{\sqrt{n}}\right]\right) \\
&= \mathrm{P}\left(z_{\alpha/2} \leq \sqrt{n}\,\frac{\widehat{p}^*_{is,n} - p^*_{is}}{\widehat{\sigma}^*_{is,n}} \leq z_{1-\alpha/2}\right) \to 1 - \alpha.
\end{aligned}$$

Die anderen beiden Aussagen lassen sich analog nachweisen. □

Für die Transformation g kommen alle streng monotonen und stetig differenzierbaren Bijektionen $g : (0,1) \to \mathbb{R}$ in Frage, deren erste Ableitungen auf dem offenen Einheitsintervall nullstellenfrei sind. Um daraus eine sinnvolle Auswahl zu treffen, können wir an die Konfidenzintervalle $\left[u^T_{is,n},\, o^T_{is,n}\right]$ eine Forderung stellen, die für die

Punktschätzer der relativen Effekte bereits in Satz 4.6 nachgewiesen wurde: Unter streng monoton fallenden Transformationen ϕ der Beobachtungen werden die relativen Effekte p_{is} wie die entsprechenden Schätzer $\widehat{p}_{is,n}$ am Wert 1/2 gespiegelt. Damit dies auch für die Konfidenzintervalle nach der Transformationsmethode der Fall ist, sollte der Graph von g punktsymmetrisch um $(1/2, g(1/2))$ sein. Darüber hinaus können wir uns ohne Verlust an Allgemeinheit auf Funktionen g mit $g(1/2) = 0$ beschränken.

Zwei in der Statistik bedeutsame Funktionen mit diesen Eigenschaften sind zum Beispiel die logit-Funktion und die probit-Funktion (siehe z. B. McCullagh und Nelder, 1989). Wegen ihrer guten Handhabbarkeit und großen Popularität verwenden wir in dieser Arbeit die logit-Funktion

$$x \mapsto \operatorname{logit}(x) = \log\left(\frac{x}{1-x}\right)$$

mit der Ableitung

$$\operatorname{logit}'(x) = \frac{1}{x(1-x)}$$

und der expit-Funktion als Umkehrfunktion:

$$\operatorname{logit}^{-1}(x) = \operatorname{expit}(x) = \frac{\exp(x)}{\exp(x)+1}.$$

5.7 Perzentilmethode

Eine zweite Methode zur Konstruktion bereichserhaltender Konfidenzintervalle ist die sogenannte *Perzentilmethode*, die auf Efron (1979) zurückgeht. Die Perzentilmethode gehört zu den sogenannten *Bootstrap-Methoden* und wird in den entsprechenden Lehrbüchern (z. B. Efron und Tibshirani, 1993; Davison und Hinkley, 1997) allgemein vorgestellt. Eine mathematisch anspruchsvollere Darstellung findet der interessierte Leser in van der Vaart (1998).

Grundlage der Perzentilmethode ist die Schätzung der multivariaten Verteilungsfunktionen $F_{i\boldsymbol{m}_{ik}}$ durch empirische Verteilungsfunktionen. Um in der Praxis bei diesen Schätzungen keine allzu kleinen Stichprobenumfänge zu haben und um die Dimensionen dieser Verteilungen nicht zu groß werden zu lassen, beschränken wir uns in diesem Abschnitt auf den Fall

$$\boldsymbol{m}_{ik} = \mathbf{1}_t, \; i = 1, \ldots, a, \; k = 1, \ldots, n_i.$$

Wir können wegen dieser Einschränkung für die Verteilungsfunktionen $F_{i\boldsymbol{m}_{ik}}$ kurz F_i schreiben und bei allen Symbolen auf den Index für die abhängigen Messwiederholungen verzichten. Als Schätzer für die Verteilungsfunktionen bieten sich die üblichen multivariaten empirischen Verteilungsfunktionen an:

$$\widehat{F}_{i,n}(\boldsymbol{x}) = \frac{1}{2n_i} \sum_{k=1}^{n_i} \left(1_{\{\boldsymbol{X}_{ik} < \boldsymbol{x}\}} + 1_{\{\boldsymbol{X}_{ik} \le \boldsymbol{x}\}}\right).$$

Die Relationszeichen in $\{\boldsymbol{X}_{ik} < \boldsymbol{x}\}$ und $\{\boldsymbol{X}_{ik} \leq \boldsymbol{x}\}$ sind dabei komponentenweise aufzufassen.

Der nächste Schritt der Perzentilmethode besteht in der Betrachtung einer sogenannten Bootstrap-Stichprobe

$$\boldsymbol{X}^*_{11,n}, \ldots, \boldsymbol{X}^*_{1n_1,n}, \ldots, \boldsymbol{X}^*_{a1,n}, \ldots, \boldsymbol{X}^*_{an_a,n}.$$

Die Zufallsvektoren $\boldsymbol{X}^*_{ik,n}$ seien dabei bedingt auf die empirischen Verteilungsfunktionen $\widehat{F}_{i,n}$ unabhängig, und die bedingten Verteilungsfunktionen der $\boldsymbol{X}^*_{ik,n}$ seien für alle Paare (i,k) gegeben durch

$$\boldsymbol{X}^*_{ik,n} \sim \widehat{F}_{i,n}.$$

In der gleichen Weise, wie wir aus der ursprünglichen Stichprobe $\boldsymbol{X}_{ik}$, $i = 1, \ldots, a$, $k = 1, \ldots, n_i$, die Schätzer $\widehat{p}_{is,n}$, $\widehat{p}_{i*,n}$ und $\widehat{p}_{*s,n}$ berechnen, können wir aus der Bootstrap-Stichprobe die Zufallsvariablen $\widehat{p}^*_{is,n}$, $\widehat{p}^*_{i*,n}$ und $\widehat{p}^*_{*s,n}$ ermitteln. Die $(1-\alpha)$-Konfidenzintervalle für die relativen Effekte nach der Perzentil-Methode sind durch die auf $\left(\widehat{F}_{1,n}, \ldots, \widehat{F}_{a,n}\right)$ bedingten Quantile dieser Zufallsvariablen definiert:

$$\begin{aligned}
\left[u^P_{is,n}, o^P_{is,n}\right] &= \left[\xi^{\alpha/2}_{is,n}, \xi^{1-\alpha/2}_{is,n}\right], \\
\left[u^P_{i*,n}, o^P_{i*,n}\right] &= \left[\xi^{\alpha/2}_{i*,n}, \xi^{1-\alpha/2}_{i*,n}\right], \\
\left[u^P_{*s,n}, o^P_{*s,n}\right] &= \left[\xi^{\alpha/2}_{*s,n}, \xi^{1-\alpha/2}_{*s,n}\right].
\end{aligned}$$

Dabei bezeichnet $\xi^{\beta}_{is,n}$ das β-Quantil der bedingten Verteilung von $\widehat{p}^*_{is,n}$, $\xi^{\beta}_{i*,n}$ das β-Quantil der bedingten Verteilung von $\widehat{p}^*_{i*,n}$ und $\xi^{\beta}_{*s,n}$ das β-Quantil der bedingten Verteilung von $\widehat{p}^*_{*s,n}$.

Dass die Konfidenzintervalle nach der Perzentilmethode bereichserhaltend sind, ergibt sich daraus, dass die Zufallsvariablen $\widehat{p}^*_{is,n}$, $\widehat{p}^*_{i*,n}$ und $\widehat{p}^*_{*s,n}$ Elemente des Einheitsintervalls sind. Darüber hinaus fällt auf, dass zur Berechnung der Intervalle im Gegensatz zu den bisher konstruierten Konfidenzintervallen kein Varianzschätzer benötigt wird.

Anders als bei den in den vorangehenden Abschnitten beschriebenen Konfidenzintervallen werden wir an dieser Stelle keinen Versuch unternehmen, die asymptotische Korrektheit der Konfidenzintervalle nach der Perzentilmethode zu beweisen. Eine nach Meinung des Autors gute heuristische Begründung für die asymptotische Gültigkeit findet man in Abschnitt 23.1 von van der Vaart (1998). Darüber hinaus findet man in den anschließenden Abschnitten des genannten Buches Konzepte für Beweise der asymptotischen Gültigkeit. Diese beinhalten jedoch tiefliegende und im Rahmen dieser Arbeit zu weit führende mathematische Methoden wie die Konvergenz empirischer Prozesse in der von Hoffmann-Jørgensen (1991) eingeführten Form und die Hadamard-Differenzierbarkeit statistischer Funktionale. Die Aufwendigkeit mathematischer Beweise für die Korrektheit von Bootstrap-Methoden kann man auch daran ablesen, dass sich die Dissertation von Steland (1996), die Statistiken behandelt, die den Statistiken der vorliegenden Arbeit ähnlich sind, zu großen Teilen mit solchen Beweisen beschäftigt.

Abschließend gehen wir noch kurz auf die praktische Berechnung der bedingten Quantile ein. Es scheint nicht einfach, Formeln zur Berechnung dieser Quantile auf Grundlage der empirischen Verteilungsfunktionen anzugeben. Dies ist typisch für Bootstrap-Verfahren, bereitet aber bei ausreichenden elektronischen Rechenkapazitäten keine ernsthaften Probleme. Die übliche Vorgehensweise besteht darin, durch Simulationen eine Anzahl von B auf $\left(\widehat{F}_{1,n}, \ldots, \widehat{F}_{a,n}\right)$ bedingt unabhängige Bootstrap-Stichproben

$$\begin{gathered}
\boldsymbol{X}^{*1}_{11,n}, \ldots, \boldsymbol{X}^{*1}_{1n_1,n}, \ldots, \boldsymbol{X}^{*1}_{a1,n}, \ldots, \boldsymbol{X}^{*1}_{an_a,n}, \\
\boldsymbol{X}^{*2}_{11,n}, \ldots, \boldsymbol{X}^{*2}_{1n_1,n}, \ldots, \boldsymbol{X}^{*2}_{a1,n}, \ldots, \boldsymbol{X}^{*2}_{an_a,n}, \\
\vdots \\
\boldsymbol{X}^{*B}_{11,n}, \ldots, \boldsymbol{X}^{*B}_{1n_1,n}, \ldots, \boldsymbol{X}^{*B}_{a1,n}, \ldots, \boldsymbol{X}^{*B}_{an_a,n}
\end{gathered}$$

zu erzeugen und daraus die Zufallsvariablen

$$\widehat{p}^{*1}_{is,n}, \widehat{p}^{*2}_{is,n}, \ldots, \widehat{p}^{*B}_{is,n}, \quad \widehat{p}^{*1}_{i*,n}, \widehat{p}^{*2}_{i*,n}, \ldots, \widehat{p}^{*B}_{i*,n} \quad \text{und} \quad \widehat{p}^{*1}_{*s,n}, \widehat{p}^{*2}_{*s,n}, \ldots, \widehat{p}^{*B}_{*s,n}$$

zu berechnen. Als Approximationen für die bedingten Quantile können die bedingten empirischen Quantile der $\widehat{p}^{*b}_{is,n}$, $\widehat{p}^{*b}_{i*,n}$ und $\widehat{p}^{*b}_{*s,n}$, $b = 1, \ldots, B$, dienen. Praktische Aspekte dieser Approximationsmethode werden mit ausführlichen Literaturhinweisen von Chernick (1999) diskutiert. Dort wird auch empfohlen, B nicht kleiner als 5000 zu wählen. Dementsprechend wurde dieser Wert bei den Beispielen und Simulationen im folgenden Abschnitt verwendet.

5.8 Beispiele und Simulationen

Bevor wir Konfidenzintervalle für die relativen Effekte in den Beispielen aus Unterabschnitt 2.1.1 angeben können, müssen wir uns zwischen den vier verschiedenen vorgestellten Methoden entscheiden. Hierbei sind im Wesentlichen zwei Kriterien zu berücksichtigen: Zunächst muss geprüft werden, wie gut die asymptotischen Konfidenzintervalle ihr Niveau bei endlichen Stichprobenumfängen einhalten. Darüber hinaus sollte auch beachtet werden, wie lang die Konfidenzintervalle sind. Je kürzer Konfidenzintervalle sind, umso genauere Aussagen können wir mit ihrer Hilfe treffen. Das zweite genannte Kriterium sollte allerdings nur dann eine Rolle spielen, wenn verschiedene Konfidenzintervalle in vergleichbarer Weise das Niveau einhalten oder wenn sie das Niveau sogar überschreiten.

Um die genannten Kriterien zu prüfen, stehen als wesentliches Hilfsmittel Simulationen zur Verfügung. Mit Hilfe von Simulationen, die sich mit der heutigen elektronischen Datenverarbeitung schnell und in hinreichendem Umfang durchführen lassen, können wir prüfen, wie sich die Konfidenzintervalle hinsichtlich der genannten Kriterien für bestimmte Versuchspläne und Wahrscheinlichkeitsverteilungen verhalten. Während die Versuchspläne in der Praxis bekannt sind, haben wir, wenn wir mit dem nichtparametrischen Modell aus Abschnitt 2.2 arbeiten, nur sehr wenige Informationen über die Wahrscheinlichkeitsverteilungen. Wir stehen hier einem grundsätzlichen Problem gegenüber, da das tatsächliche Niveau der Konfidenzintervalle und die Verteilung ihrer Längen sicherlich von den Wahrscheinlichkeitsverteilungen abhängt. Weil es immer unendlich viele in Frage kommende Verteilungen gibt, können wir nicht für alle Verteilungen Simulationen durchführen. Deshalb sind wir gezwungen, aus den in Frage kommenden Verteilungen eine Auswahl zu treffen und die Simulationen für diese Auswahl durchzuführen.

Hier haben wir es bei der Simulation des Niveaus von Konfidenzintervallen einfacher als bei der Simulation des Niveaus von statistischen Tests: Bei Konfidenzintervallen können wir die beobachteten (Marginal-)Verteilungen zur Grundlage der Simulationen machen. Dies ist bei statistischen Tests nicht möglich, weil die beobachteten Verteilungen entweder nicht der gestellten Hypothese genügen oder, wenn sie der Hypothese doch genügen, diese nicht verworfen werden kann und Niveausimulationen damit überflüssig werden.

Die Simulationen zu den vier Beispielen sind entsprechend den obigen Überlegungen analog aufgebaut: Die Versuchspläne der Simulationen stimmen (mit Ausnahme der fehlenden Werte bei der Panik-Skala-Studie II) mit denen der Beispiele überein. Die den $\boldsymbol{X}_{ik}$ entsprechenden Zufallsvektoren wurden in den Simulationen als unabhängige Zufallsvektoren

$$\boldsymbol{Y}_{ik} = (Y_{ik11}, \ldots, Y_{ik1m_{ik1}}, \ldots, Y_{ikt1}, \ldots, Y_{iktm_{ikt}})'$$

mit den Randverteilungsfunktionen $\widehat{F}_{is,n}$ erzeugt. Durch die Festlegung dieser Randverteilungen sind die relativen Effekte der simulierten Zufallsvektoren eindeutig auf die Werte $\widehat{p}_{is,n}$ festgelegt. Zur Erzeugung der Zufallsvektoren $\boldsymbol{Y}_{ik}$ wurden in einem ersten Schritt multivariat normalverteilte Zufallsvektoren

$$\widetilde{\boldsymbol{Y}}_{ik} = \left(\widetilde{Y}_{ik11}, \ldots, \widetilde{Y}_{ik1m_{ik1}}, \ldots, \widetilde{Y}_{ikt1}, \ldots, \widetilde{Y}_{iktm_{ikt}}\right)' \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{W})$$

generiert. Die Elemente der Kovarianzmatrix $\boldsymbol{W}$ bezeichnen wir mit

$$\tau_{srs'r'} = \mathrm{Cov}\left(\widetilde{Y}_{iksr}, \widetilde{Y}_{iks'r'}\right).$$

In einem zweiten Schritt wurden die Komponenten $Y_{iksr} \sim \widehat{F}_{is,n}$ der simulierten Zufallsvektoren durch eine geeignete Kategorisierung der stetigen Zufallsvariablen $\widetilde{Y}_{iksr}$ gebildet. Bei dieser Konstruktion wird die Abhängigkeitsstruktur innerhalb der Zufallsvektoren $\boldsymbol{Y}_{ik}$ durch die Kovarianzen $\tau_{srs'r'}$ bestimmt.

Die Anzahl an Simulationen wurde einheitlich auf die heute weithin üblichen 10000 festgelegt. Auch bei dem Niveau der Konfidenzintervalle wurde nicht von den gewöhnlichen 95% abgewichen. Die Ergebnisse befinden sich in Anhang C. Alle Simulationen wurden auf Grundlage der in Anhang B beschriebenen SAS-Makros `RE_CI` und `SIMPLE_RE_CI` durchgeführt.

Panik-Skala-Studie I

Für die Simulationen zur Panik-Skala-Studie I wurden die Elemente der Kovarianzmatrix $\boldsymbol{W}$ durch

$$\tau_{s1s'1} = \left(\frac{1}{2}\right)^{|s-s'|}$$

festgelegt. Somit hängen die simulierten Zufallsvariablen Y_{1ks1} und $Y_{1ks'1}$ umso stärker voneinander ab, je kleiner der Abstand zwischen den Zeitpunkten s und s' ist.

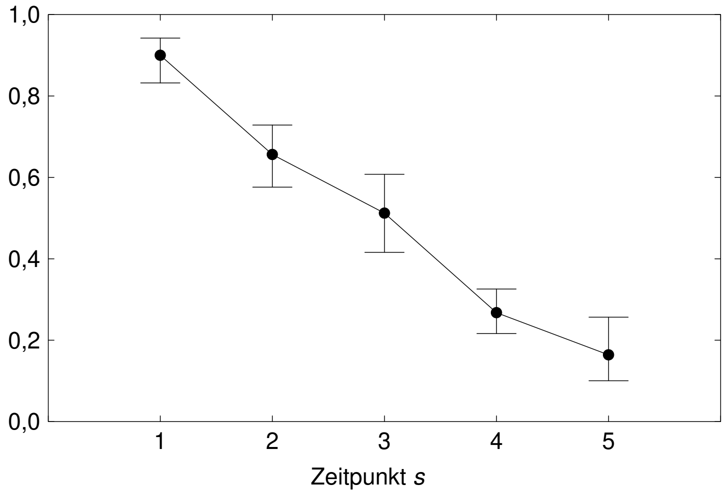

Abbildung 5.2: Schätzer $\widehat{p}_{is}$ mit 95%-Konfidenzintervallen auf Grundlage der Transformationsmethode für die relativen Effekte in der Panik-Skala-Studie I.

Durch diese Eigenschaft soll der longitudinalen Struktur der Zufallsvektoren Rechnung getragen werden.

Für den Versuchsplan der Panik-Skala-Studie I konnten die Simulationen auf Basis des SAS-Makros `SIMPLE_RE_CI` für alle vier verschiedenen Typen von Konfidenzintervallen durchgeführt werden. Die Ergebnisse befinden sich in Anhang C.1.

Bei den Niveausimulationen fällt zunächst auf, dass lediglich mit der t-Approximation zum dritten Zeitpunkt das Niveau eingehalten wird. Im schlechtesten Fall sinkt das Niveau auf 90,1% (Normal-Approximation und Perzentil-Methode zum ersten Zeitpunkt). Die beiden geeignetesten Methoden scheinen bei diesem Beispiel die t-Approximation und die Transformationsmethode zu sein. Die t-Approximation schneidet im Vergleich mit den anderen Methoden bei den mittleren drei Zeitpunkten am besten ab, die Transformationsmethode dagegen an den Rändern. Für die Transformationsmethode spricht zudem, dass die durchschnittliche Länge der simulierten Konfidenzintervalle im Vergleich zur t-Approximation geringer ist. Insgesamt scheint deshalb hier die Transformationsmethode am geeignetesten zu sein. Die entsprechenden mit dem SAS-Makro `SIMPLE_RE_CI` berechneten 95%-Konfidenzintervalle $\left[u_{is,n}^T,\, o_{is,n}^T\right]$ sind in Abbildung 5.2 dargestellt.

Im Vergleich zu dem von den relativen Effekten eingeschlossenen Bereich von 0,16 bis 0,90 sind die Konfidenzintervalle eher kurz. Die schon bei der Betrachtung der Punktschätzer $\widehat{p}_{1s}$ in Abschnitt 4.6 aufgestellte Vermutung, dass die relativen Effekte mit der Zeit immer kleiner werden, wird durch die Konfidenzintervalle weiter gestützt. Abgesehen von der inhaltlichen Interpretation fällt bei der Betrachtung der Konfidenzintervalle als technischer Aspekt die durch die Transformationsmethode verursachte Asymmetrie an den Rändern auf.

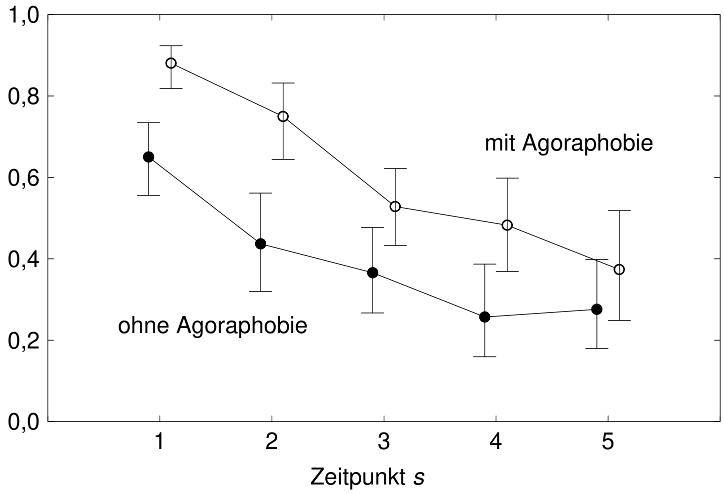

Abbildung 5.3: Schätzer $\widehat{p}_{is}$ mit 95%-Konfidenzintervallen auf Grundlage der Transformationsmethode für die relativen Effekte in der Panik-Skala-Studie II.

Panik-Skala-Studie II

Für die Simulationen zur Panik-Skala-Studie II wurden die Elemente der Kovarianzmatrix $\boldsymbol{W}$ wie bei der Panik-Skala-Studie I durch

$$\tau_{s1s'1} = \left(\frac{1}{2}\right)^{|s-s'|}$$

festgelegt. Somit hängen die simulierten Zufallsvariablen Y_{iks1} und $Y_{iks'1}$ wiederum umso stärker voneinander ab, je kleiner der Abstand zwischen den Zeitpunkten s und s' ist. Durch diese Eigenschaft soll die auch hier vorliegende longitudinale Struktur der Zufallsvektoren berücksichtigt werden.

Im Gegensatz zur Panik-Skala-Studie I fehlen bei der Panik-Skala-Studie II einige Beobachtungen. Wie in Abschnitt 2.2.2 erläutert, gehen wir davon aus, dass die Beobachtungen rein zufällig fehlen. Deshalb wurden auch bei den Simulationen Beobachtungen zufällig aus den Datensätzen entfernt. Als Wahrscheinlichkeiten für das Entfernen der simulierten Beobachtungen wurden die beobachteten relativen Häufigkeiten aus den Originaldaten verwendet. In der Agoraphobie-Gruppe wurden also simulierte Beobachtungen aus den Wochen 4 und 6 mit Wahrscheinlichkeit 3/24 und aus Woche 8 mit Wahrscheinlichkeit 4/24 gelöscht. In der Gruppe ohne Agoraphobie wurden Werte aus den Wochen 6 und 8 mit Wahrscheinlichkeit 3/13 entfernt. Wegen der fehlenden Werte können für dieses Beispiel Konfidenzintervalle nur mit der Normal-Approximation und der Transformationmethode berechnet und simuliert werden. Die Simulationsergebnisse sind in den Tabellen in Anhang C.2 angegeben.

Obwohl die simulierten Konfidenzintervalle auf Basis der Transformationsmethode im Mittel kürzer sind als die auf Basis der Normal-Approximation, halten sie das Niveau für beide Gruppen zu allen Zeitpunkten besser ein. Bei diesem Beispiel ist die Entscheidung für die Transformationsmethode also eindeutig. Die entsprechenden mit dem SAS-Makro `RE_CI` berechneten Konfidenzintervalle sind in Abbildung 5.3 dargestellt.

Wie bei der Panik-Skala-Studie I sind die Unterschiede zwischen den geschätzten relativen Effekten zu den verschiedenen Zeitpunkten im Vergleich zu den Längen der Konfidenzintervalle für beide Gruppen eher groß. Wir können somit die Vermutung aus Abschnitt 4.6 festigen, dass die Panikstörungen im Laufe der Zeit sowohl für Patienten mit als auch ohne Agoraphobie zurückgehen. Darüber hinaus haben wir, weil sich die Konfidenzintervalle der beiden Gruppen zu den ersten beiden Zeitpunkten nicht überschneiden, einen starken Hinweis darauf, dass Patienten mit Agoraphobie zumindest zu Beginn der Imipramin-Therapie zu stärkeren Panik-Störungen neigen als Patienten ohne Agoraphobie.

Schulterschmerz-Studie

Die Kovarianzmatrix $\boldsymbol{W}$ wurde aus den gleichen Gründen wie bei den Panik-Skala-Studien durch

$$\tau_{s1s'1} = \left(\frac{1}{2}\right)^{|s-s'|}$$

festgelegt.

Für den Versuchsplan der Schulterschmerz-Studie wurden die Simulationen wieder auf Basis des SAS-Makros `SIMPLE_RE_CI` und somit für alle vier Methoden durchgeführt. Die Ergebnisse befinden sich in Anhang C.3.

Von den vier Methoden hält hier die t-Approximation für beide Behandlungen zu allen Zeitpunkten das Niveau am besten ein. Obwohl diese Konfidenzintervalle im Durchschnitt nicht die kürzesten waren, scheinen sie am ehesten geeignet zu sein. Sie sind in Abbildung 5.4 dargestellt.

Die 95%-Konfidenzintervalle $[u_{is}^t, o_{is}^t]$ der beiden Gruppen überlappen sich zu den Zeitpunkten 2 bis 6 nicht und haben teilweise sogar große Abstände zueinan-

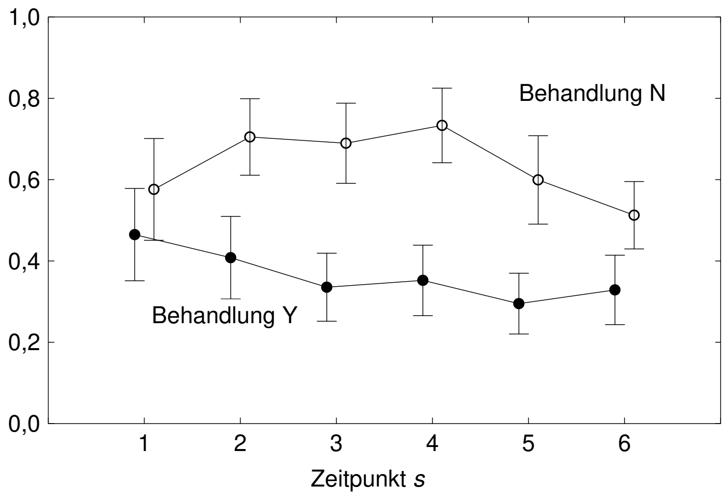

Abbildung 5.4: Schätzer $\widehat{p}_{is}$ mit 95%-Konfidenzintervallen auf Grundlage der t-Approximation für die relativen Effekte in der Schulterschmerz-Studie.

der. Damit gewinnen wir zusätzliche Sicherheit in Bezug auf unsere Vermutung aus Abschnitt 4.6, dass behandelte Patienten nach der Operation deutlich weniger unter Schulterschmerzen leiden als Patienten aus der Kontrollgruppe.

Wasser-Irrgarten-Test

Bei den Beobachtungen des Wasser-Irrgarten-Tests liegt eine etwas andere Abhängigkeitsstruktur vor als bei den drei anderen Beispielen: Die Zufallsvektoren

$$\boldsymbol{X}_{ik} = (X_{ik11}, X_{ik12}, X_{ik21}, X_{ik22})'$$

repräsentieren hier Beobachtungen an zwei jungen Ratten mit dem jeweils gleichen Muttertier an den Tagen 1 und 7. Die dritten Indizes an den Symbolen X_{iksr} stehen hier für die Zeitpunkte und die vierten für die Jungtiere. Es scheint naheliegend, dass Beobachtungen an demselben Jungtier stärker voneinander abhängen als Beobachtungen an verschiedenen Jungtieren. Um dies zu berücksichtigen, wurde bei den Simulationen als Kovarianzmatrix

$$\boldsymbol{W} = \begin{pmatrix} 1 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & 1 & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{4} & 1 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 1 \end{pmatrix}$$

verwendet. Für den Versuchsplan des Wasser-Irrgarten-Tests wurden auf Grundlage des SAS-Makros `RE_CI` Simulationen mit der Normal-Approximation und der Transformationsmethode durchgeführt. Die Ergebnisse befinden sich in Anhang C.4.

Wieder zeigt sich, dass mit der Transformationsmethode das Niveau eher eingehalten wird als bei der Normal-Approximation. Trotzdem sind die Konfidenzintervalle nach der Transformationsmethode im Durchschnitt kürzer. Dementsprechend sind diese Konfidenzintervalle in Abbildung 5.5 angegeben.

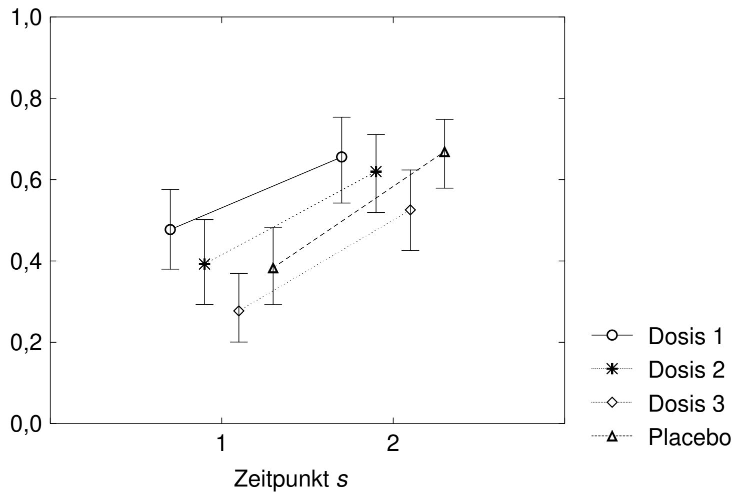

Abbildung 5.5: Schätzer $\widehat{p}_{is}$ mit 95%-Konfidenzintervallen auf Grundlage der Transformationsmethode für die relativen Effekte beim Wasser-Irrgarten-Test.

Die vielen Überschneidungen zwischen den 95%-Konfidenzintervallen bei diesem Beispiel warnen davor, den Unterschieden zwischen den Punktschätzern für die relativen Effekte bei den verschiedenen Dosierungen eine allzu große Bedeutung beizumessen. Dabei sollte auch berücksichtigt werden, dass der geschätzte relative Effekt für die Placebo-Gruppe zwischen den relativen Effekten für die Dosierungen 1 bis 3 liegt, was sich wegen der hohen Plausibilität einer monotonen Dosis-Wirkungsbeziehung nur schwer erklären lässt. Im Gegensatz zu den drei anderen Beispielen finden wir hier also kaum eine Bestätigung dafür, dass die beobachteten Unterschiede in den relativen Effekten auf systematische Unterschiede in den Verteilungen zurückzuführen sind.

Kapitel 6

Varianzreduktion in randomisierten Versuchsplänen

In diesem Kapitel geht es darum, wie in bestimmten Situationen zusätzliche Information zur Konstruktion von verbesserten Schätzern für die relativen Effekte verwendet werden kann. Mit bestimmten Situationen sind hier die in der Biometrie häufig behandelten randomisierten Studien gemeint. Unter zusätzlicher Information wollen wir entweder Ausgangswerte oder Kovariablen mit mindestens ordinalem Skalenniveau verstehen. Wir werden uns in den folgenden Abschnitten nicht darum bemühen, das Problem der Varianzreduktion in möglichst allgemeiner Form zu behandeln, wie wir es bei der Diskussion der relativen Effekte in den vorangehenden Kapiteln getan haben, sondern uns auf einen besonders einfachen Versuchsplan beschränken. Insgesamt sollen die folgenden Ausführungen im Wesentlichen als Anregung für die weitere Forschung dienen, weshalb hier auch auf die ausführliche Diskussion von Beispielen oder Simulationen verzichtet wird.

Im ersten Abschnitt dieses Kapitels beschreiben wir die betrachteten Versuchspläne und Modelle. Im Anschluss daran erläutern wir kurz die Idee der Varianzreduktion. Bevor diese umgesetzt wird, übertragen wir die Ergebnisse zur asymptotischen Verteilung der Schätzer und zur Schätzung der Kovarianzmatrix aus Kapitel 5 in den Kontext dieses Kapitels. Im letzten Abschnitt dieses Kapitels wird diskutiert, wie sich die Ideen der Varianzreduktion mit der Transformationsmethode aus Abschnitt 5.6 kombinieren lassen.

6.1 Versuchsplan und Modell

Wir befassen uns in diesem Kapitel mit dem folgenden einfachen Versuchsplan: Die Beobachtungspaare

$$(X_{ik}, Y_{ik})' \sim F_i, \quad i = 1, \dots, a,\ k = 1, \dots, n_i,$$

seien unabhängig. Für jede der insgesamt $n = \sum_{i=1}^{a} n_i$ Versuchseinheiten liegen also genau zwei Beobachtungen vor.

Die Beobachtungen X_{ik} seien die, für deren Randverteilungen F_i^X wir uns interessieren. Um die Verteilungen zu vergleichen, bilden wir die in Abschnitt 3.3 für eine allgemeinere Klasse von Versuchsplänen eingeführten relativen Effekte

$$p_i^X = \mathrm{P}\left(Z_i^X < X_{ik}\right) + \frac{1}{2}\,\mathrm{P}\left(Z_i^X = X_{ik}\right),$$

wobei Z_i^X eine von X_{ik} unabhängige Zufallsvariable mit der Verteilungsfunktion

$$H_i^X = \frac{1}{a-1}\sum_{u \neq i} F_i^X$$

bezeichnet. Diese relativen Effekte besitzen die Integraldarstellung

$$p_i^X = \int H_i^X \, dF_i^X.$$

Die Beobachtungen Y_{ik} seien Ausgangswerte oder Kovariablen, die vor der Zuteilung der Versuchseinheiten zu den Gruppen beobachtet wurden. Da wir uns hier nur mit randomisierten Studien befassen wollen, können wir annehmen, dass alle Y_{ik} die gleiche Verteilungsfunktion $F_i^Y = F^Y$ haben. Von den in Abschnitt 2.1.1 genannten Versuchen können wir hier die Schulterschmerz-Studie als Beispiel heranziehen, in der die Patienten zufällig den beiden Behandlungen zugeteilt wurden. Als Kovariable könnte hier zum Beispiel das Alter der Patienten dienen, dass in der Quelle für den Datensatz (Lumley, 1996) auch tatsächlich angegeben ist. Andere denkbare Kovariablen sind die Dauer der Operation oder die Schmerzempfindlichkeit der Patienten. Wichtig ist an dieser Stelle, dass es sich bei der Kovariablen um Beobachtungen mit mindestens ordinalem Skalenniveau handelt, die mit den eigentlichen Beobachtungen X_{ik} möglichst stark zusammenhängen.

Auch für die Kovariablen oder Ausgangswerte können wir die relativen Effekte

$$p_i^Y = \mathrm{P}\left(Z_i^Y < Y_{ik}\right) + \frac{1}{2}\,\mathrm{P}\left(Z_i^Y = Y_{ik}\right)$$

bilden, wobei Z_i^Y eine von Y_{ik} unabhängige Zufallsvariable mit der Verteilungsfunktion

$$H_i^Y = \frac{1}{a-1}\sum_{u \neq i} F_i^Y = F^Y$$

ist. Wegen der Integraldarstellung

$$p_i^Y = \int H_i^Y \, dF_i^Y = \int F^Y \, dF^Y$$

und Korollar 3.4 auf Seite 25 gilt aber $p_i^Y = 1/2$ für alle $i \in \{1, \ldots, a\}$.

Wie bisher müssen wir für die asymptotischen Ergebnisse in diesem Kapitel Annahmen an die Folgen der Stichprobenumfänge stellen. Die Annahmen (A1) und (A2) aus Abschnitt 2.1.2 nehmen für die Klasse von Versuchsplänen aus diesem Kapitel folgende Gestalt an:

(A1) $n_i \to \infty$ für alle $i \in \{1, \ldots, a\}$,

(A2) $n/n_i \leq n_0 < \infty$ für alle $i \in \{1, \ldots, a\}$.

Wenn wir uns in diesem Kapitel auf (A1) und (A2) beziehen, meinen wir immer diese übertragene Form. Auf die Bedingung (A3) können wir hier verzichten, da sie für alle in diesem Kapitel betrachteten Versuchspläne immer erfüllt ist.

6.2 Idee der Varianzreduktion

Wie die Punktschätzer für die relativen Effekte aus Kapitel 4 beruhen auch die Punktschätzer in diesem Kapitel auf den durch

$$\widehat{F}_{i,n}^X(x) = \frac{1}{n_i}\sum_{k=1}^{n_i} c(x - X_{ik}) \quad \text{und} \quad \widehat{H}_{i,n}^X = \frac{1}{a-1}\sum_{u\neq i}\widehat{F}_{u,n}^X$$

definierten empirischen Verteilungsfunktionen. Zusätzlich betrachten wir

$$\widehat{F}_{i,n}^Y(x) = \frac{1}{n_i}\sum_{k=1}^{n_i} c(x - Y_{ik}) \quad \text{und} \quad \widehat{H}_{i,n}^Y = \frac{1}{a-1}\sum_{u\neq i}\widehat{F}_{u,n}^Y.$$

Aus Abschnitt 4.4 wissen wir, dass

$$\widehat{p}_{i,n}^X = \int \widehat{H}_{i,n}^X \, d\widehat{F}_{i,n}^X \quad \text{und} \quad \widehat{p}_{i,n}^Y = \int \widehat{H}_{i,n}^Y \, d\widehat{F}_{i,n}^Y$$

erwartungstreue und unter Annahme (A1) stark konsistente Schätzer für p_i^X und $p_i^Y = 1/2$ sind. Auf den ersten Blick scheint es sinnlos zu sein, Schätzer für die bekannten p_i^Y anzugeben. Sinn und Zweck dieser Vorgehensweise deuten sich aber bereits im folgenden Satz an:

Satz 6.1. *Für alle beschränkten Folgen reeller Zahlen $\gamma_{i,n}$ und alle $i \in \{1,\ldots,a\}$ ist*

$$\widehat{p}_{i,n}^{\gamma_{i,n}} = \widehat{p}_{i,n}^X - \gamma_{i,n}\left(\widehat{p}_{i,n}^Y - \frac{1}{2}\right)$$

ein erwartungstreuer und unter Bedingung (A1) stark konsistenter Schätzer für p_i^X.

Beweis. Die Behauptungen ergeben sich aus den wohlbekannten Rechenregeln für Erwartungswerte und fast sichere Grenzwerte. □

Obiger Satz legt es nahe, unter allen Schätzern $\widehat{p}_{i,n}^{\gamma_{i,n}}$ denjenigen zu suchen, dessen asymptotische Varianz minimal ist. Diese Idee werden wir in den folgenden Abschnitten konkretisieren und umsetzen.

6.3 Asymptotische Verteilung und Schätzung der Kovarianzmatrix

Bevor wir mit der Umsetzung der Idee aus dem vorangehenden Abschnitt beginnen, übertragen wir die Ergebnisse zur asymptotischen Verteilung der Punktschätzer und zur Schätzung der Kovarianzmatrix aus Kapitel 5 auf die hier betrachteten Versuchpläne. Bei der Schätzung der Kovarianzmatrix können wir als zusätzliche Information die Annahme $F_i^Y = F^Y$, $i = 1,\ldots,a$, gewinnbringend verwenden.

Als abkürzende Schreibweise fassen wir die relativen Effekte und die Schätzer dafür in Form der Vektoren

$$\boldsymbol{p}^X = \left(p_1^X, \ldots, p_a^X\right)' \quad \text{und} \quad \boldsymbol{p}^Y = \frac{1}{2}\,\mathbf{1}_a$$

sowie
$$\widehat{\boldsymbol{p}}_n^X = \left(\widehat{p}_{1,n}^X, \ldots, \widehat{p}_{a,n}^X\right)' \quad \text{und} \quad \widehat{\boldsymbol{p}}_n^Y = \left(\widehat{p}_{1,n}^Y, \ldots, \widehat{p}_{a,n}^Y\right)'$$
zusammen.

Grundlage für die Resultate zur asymptotischen Verteilung und zur Konstruktion von Schätzern der Kovarianzmatrix ist wieder ein asymptotischer Äquivalenzsatz.

Satz 6.2. *Für die unabhängigen und unter Voraussetzung (A2) gleichmäßig beschränkten Zufallsvektoren*
$$\boldsymbol{\Psi}_{ik} = \left(\boldsymbol{\Psi}_{ik}^X, \boldsymbol{\Psi}_{ik}^Y\right)' = \left(\Psi_{ik1}^X, \ldots, \Psi_{ika}^X, \Psi_{ik1}^Y, \ldots, \Psi_{ika}^Y\right)',$$
deren Komponenten durch
$$\Psi_{iku}^X = \begin{cases} H_u^X(X_{uk}) & : u = i, \\ \frac{-1}{a-1} F_u^X(X_{ik}) & : u \neq i, \end{cases} \quad \textit{und} \quad \Psi_{iku}^Y = \begin{cases} H_u^Y(Y_{uk}) & : u = i, \\ \frac{-1}{a-1} F_u^Y(Y_{ik}) & : u \neq i, \end{cases}$$
definiert sind, gilt unter den Voraussetzungen (A1) und (A2)
$$\sqrt{n}\left[\left(\widehat{\boldsymbol{p}}_n^X - \boldsymbol{p}^X\right)', \left(\widehat{\boldsymbol{p}}_n^Y - \boldsymbol{p}^Y\right)'\right]' - \sqrt{n}\left(\sum_{i=1}^a \frac{1}{n_i}\sum_{k=1}^{n_i}[\boldsymbol{\Psi}_{ik} - \mathrm{E}(\boldsymbol{\Psi}_{ik})]\right) \xrightarrow{P} \mathbf{0}.$$

Beweis. Der Beweis der Behauptung kann vollständig analog zu dem Beweis von Korollar 5.2 (siehe Seite 60) geführt werden. □

In diesem Kapitel bezeichnen wir mit $\boldsymbol{V}_n$ die Kovarianzmatrix von
$$\sqrt{n}\left(\sum_{i=1}^a \frac{1}{n_i}\sum_{k=1}^{n_i}[\boldsymbol{\Psi}_{ik} - \mathrm{E}(\boldsymbol{\Psi}_{ik})]\right).$$
Als Symbole für die Komponenten von $\boldsymbol{V}_n$ vereinbaren wir
$$\begin{aligned}
\sigma_{uu',n}^X &= \mathrm{Cov}\left[\sqrt{n}\left(\sum_{i=1}^a \frac{1}{n_i}\sum_{k=1}^{n_i}\Psi_{iku}^X\right), \sqrt{n}\left(\sum_{i=1}^a \frac{1}{n_i}\sum_{k=1}^{n_i}\Psi_{iku'}^X\right)\right] \\
&= n\sum_{i=1}^a \frac{1}{n_i}\mathrm{Cov}\left(\Psi_{iku}^X, \Psi_{iku'}^X\right), \\
\sigma_{uu',n}^Y &= \mathrm{Cov}\left[\sqrt{n}\left(\sum_{i=1}^a \frac{1}{n_i}\sum_{k=1}^{n_i}\Psi_{iku}^Y\right), \sqrt{n}\left(\sum_{i=1}^a \frac{1}{n_i}\sum_{k=1}^{n_i}\Psi_{iku'}^Y\right)\right] \\
&= n\sum_{i=1}^a \frac{1}{n_i}\mathrm{Cov}\left(\Psi_{iku}^Y, \Psi_{iku'}^Y\right) \quad \text{sowie} \qquad (6.1) \\
\sigma_{uu',n}^{XY} &= \mathrm{Cov}\left[\sqrt{n}\left(\sum_{i=1}^a \frac{1}{n_i}\sum_{k=1}^{n_i}\Psi_{iku}^X\right), \sqrt{n}\left(\sum_{i=1}^a \frac{1}{n_i}\sum_{k=1}^{n_i}\Psi_{iku'}^Y\right)\right] \\
&= n\sum_{i=1}^a \frac{1}{n_i}\mathrm{Cov}\left(\Psi_{iku}^X, \Psi_{iku'}^Y\right).
\end{aligned}$$

Auf Grundlage von Satz 6.2 können wir die asymptotische Normalverteilung von

$$\sqrt{n}\left[\left(\widehat{\boldsymbol{p}}_n^X - \boldsymbol{p}^X\right)', \left(\widehat{\boldsymbol{p}}_n^Y - \boldsymbol{p}^Y\right)'\right]'$$

beweisen.

Satz 6.3. *Unter den Voraussetzungen (A1) und (A2) gilt*

$$\rho\left(\sqrt{n}\left[\left(\widehat{\boldsymbol{p}}_n^X - \boldsymbol{p}^X\right)', \left(\widehat{\boldsymbol{p}}_n^Y - \boldsymbol{p}^Y\right)'\right]', \mathrm{N}(\mathbf{0}, \boldsymbol{V}_n)\right) \quad \to \quad 0.$$

Beweis. Der Beweis kann mit Hilfe von Satz 6.2 genauso geführt werden wie der Beweis von Satz 5.6 (siehe Seite 62). □

Wenn wir die Komponenten von $\boldsymbol{V}_n$ nach den gleichen Prinzipien schätzen, wie wir das in Abschnitt 5.3 getan haben, so erhalten wir mit den Bezeichnungen

$$\widehat{\Psi}_{iku,n}^X = \begin{cases} \widehat{H}_{u,n}^X(X_{uk}) & : u = i, \\ \frac{-1}{a-1}\widehat{F}_{u,n}^X(X_{uk}) & : u \neq i, \end{cases} \quad \text{und} \quad \widetilde{\Psi}_{iku}^Y = \begin{cases} \widehat{H}_{u,n}^Y(Y_{ik}) & : u = i, \\ \frac{-1}{a-1}\widehat{F}_{u,n}^Y(Y_{ik}) & : u \neq i, \end{cases}$$

sowie

$$\overline{\widehat{\Psi}}_{i\cdot u,n}^X = \frac{1}{n_i}\sum_{k=1}^{n_i}\widehat{\Psi}_{iku,n}^X \quad \text{und} \quad \overline{\widetilde{\Psi}}_{i\cdot u,n}^Y = \frac{1}{n_i}\sum_{k=1}^{n_i}\widetilde{\Psi}_{iku,n}^Y$$

die Schätzer

$$\widehat{\sigma}_{uu',n}^X = n\sum_{i=1}^{a}\frac{1}{n_i(n_i-1)}\sum_{k=1}^{n_i}\left(\widehat{\Psi}_{iku,n}^X - \overline{\widehat{\Psi}}_{i\cdot u,n}^X\right)\left(\widehat{\Psi}_{iku',n}^X - \overline{\widehat{\Psi}}_{i\cdot u',n}^X\right),$$

$$\widetilde{\sigma}_{uu',n}^Y = n\sum_{i=1}^{a}\frac{1}{n_i(n_i-1)}\sum_{k=1}^{n_i}\left(\widetilde{\Psi}_{iku,n}^Y - \overline{\widetilde{\Psi}}_{i\cdot u,n}^Y\right)\left(\widetilde{\Psi}_{iku',n}^Y - \overline{\widetilde{\Psi}}_{i\cdot u',n}^Y\right)$$

sowie

$$\widetilde{\sigma}_{uu',n}^{XY} = n\sum_{i=1}^{a}\frac{1}{n_i(n_i-1)}\sum_{k=1}^{n_i}\left(\widehat{\Psi}_{iku,n}^X - \overline{\widehat{\Psi}}_{i\cdot u,n}^X\right)\left(\widetilde{\Psi}_{iku',n}^Y - \overline{\widetilde{\Psi}}_{i\cdot u',n}^Y\right).$$

Bei einer genaueren Betrachtung der Herleitung von $\widehat{\boldsymbol{V}}_n$ in Abschnitt 5.3 fällt auf, dass wir die hier angegebenen Schätzer noch verbessern können, wenn wir berücksichtigen, dass wegen der Randomisierung der Versuchseinheiten für alle $i \in \{1, \ldots, a\}$ die Gleichung $H_i^Y = F_i^Y = F^Y$ gilt.

Zunächst lesen wir aus Gleichung (6.1) unter Verwendung von

$$\mathrm{Cov}\left(\Psi_{iku}^Y, \Psi_{iku'}^Y\right) = \begin{cases} \mathrm{Var}\left(F^Y(Y_{ik})\right) & : u = i = u', \\ \frac{-1}{a-1}\,\mathrm{Var}\left(F^Y(Y_{ik})\right) & : u = i \neq u', \\ \frac{-1}{a-1}\,\mathrm{Var}\left(F^Y(Y_{ik})\right) & : u \neq i = u', \\ \frac{1}{(a-1)^2}\,\mathrm{Var}\left(F^Y(Y_{ik})\right) & : u \neq i \neq u', \end{cases}$$

die Identität

$$\sigma^Y_{uu',n} = \begin{cases} n\,\mathrm{Var}\left(F^Y(Y_{ik})\right)\left(\frac{1}{n_u} + \sum_{i\neq u} \frac{1}{n_i(a-1)^2}\right) & : u = u', \\ n\,\mathrm{Var}\left(F^Y(Y_{ik})\right)\left(\frac{-1}{n_u(a-1)} + \frac{-1}{n_{u'}(a-1)} + \sum_{i\notin\{u,u'\}} \frac{1}{n_i(a-1)^2}\right) & : u \neq u', \end{cases} \tag{6.2}$$

ab. Mit der Definition der empirischen Verteilungsfunktion $\widehat{F}^Y_n$ durch

$$\widehat{F}^Y_n(y) = \frac{1}{n}\sum_{i=1}^{a}\sum_{k=1}^{n_i} c(y - Y_{ik})$$

als Schätzer für F^Y und wegen $\mathrm{E}\left(F^Y(Y_{ik})\right) = \int F^Y\, dF^Y = 1/2$ bietet sich

$$\widehat{\tau}^2_{Y,n} = \frac{1}{n}\sum_{i=1}^{a}\sum_{k=1}^{n_i}\left(\widehat{F}^Y_n(Y_{ik}) - \frac{1}{2}\right)^2$$

als Schätzer für $\tau^2_Y = \mathrm{Var}\left(F^Y(Y_{ik})\right)$ an. Als verbesserten Schätzer für $\sigma^Y_{uu',n}$ erhalten wir damit

$$\widehat{\sigma}^Y_{uu',n} = \begin{cases} n\widehat{\tau}^2_{Y,n}\left(\frac{1}{n_u} + \sum_{i\neq u} \frac{1}{n_i(a-1)^2}\right) & : u = u', \\ n\widehat{\tau}^2_{Y,n}\left(\frac{-1}{n_u(a-1)} + \frac{-1}{n_{u'}(a-1)} + \sum_{i\notin\{u,u'\}} \frac{1}{n_i(a-1)^2}\right) & : u \neq u'. \end{cases} \tag{6.3}$$

Die zusätzliche Information über die Verteilung der Y_{ik} können wir auch in die Schätzer für $\sigma^{XY}_{uu',n}$ einfließen lassen, indem wir $\widetilde{\Psi}^Y_{iku,n}$ durch

$$\widehat{\Psi}^Y_{iku} = \begin{cases} \widehat{F}^Y_n(Y_{ik}) & : u = i, \\ \frac{-1}{a-1}\widehat{F}^Y_n(Y_{ik}) & : u \neq i, \end{cases}$$

und $\overline{\overline{\widetilde{\Psi}}}^Y_{i\cdot u,n}$ durch

$$\mathrm{E}\left(\Psi^Y_{iku,n}\right) = \begin{cases} \frac{1}{2} & : u = i, \\ \frac{-1}{2(a-1)} & : u \neq i, \end{cases}$$

ersetzen. Wir erhalten damit

$$\widehat{\sigma}^{XY}_{uu',n} = n\sum_{i=1}^{a}\frac{1}{n_i(n_i-1)}\sum_{k=1}^{n_i}\left(\widehat{\Psi}^X_{iku,n} - \overline{\overline{\widehat{\Psi}}}^X_{i\cdot u,n}\right)\left(\widehat{\Psi}^Y_{iku',n} - \mathrm{E}\left(\Psi^Y_{iku',n}\right)\right). \tag{6.4}$$

Als Schätzer $\widehat{\boldsymbol{V}}_n$ für die Kovarianzmatrix $\boldsymbol{V}_n$ verwenden wir die Matrix mit den Einträgen $\widehat{\sigma}^X_{uu',n}$, $\widehat{\sigma}^Y_{uu',n}$ und $\widehat{\sigma}^{XY}_{uu',n}$. Dieser Schätzer ist wie $\widehat{\boldsymbol{V}}_n$ aus Abschnitt 5.3 unter geeigneten Annahmen konsistent.

Satz 6.4. *Unter den Voraussetzungen (A1) und (A2) ist $\widehat{\boldsymbol{V}}_n$ konsistent für $\boldsymbol{V}_n$ in dem Sinne, dass*

$$\widehat{\boldsymbol{V}}_n - \boldsymbol{V}_n \xrightarrow{f.s.} \mathbf{0}$$

und

$$\sum_{u=1}^{a}\sum_{u'=1}^{a}\left[\mathrm{E}\left(\left(\widehat{\sigma}^X_{uu',n} - \sigma^X_{uu',n}\right)^2\right) + \mathrm{E}\left(\left(\widehat{\sigma}^Y_{uu',n} - \sigma^Y_{uu',n}\right)^2\right) + \mathrm{E}\left(\left(\widehat{\sigma}^{XY}_{uu',n} - \sigma^{XY}_{uu',n}\right)^2\right)\right] \to 0.$$

Beweis. Bis auf einige Vereinfachungen, die wegen des einfacheren Versuchsplans und der verwendeten zusätzlichen Information möglich sind, verläuft der Beweis analog zu dem von Satz 5.9 (siehe Seite 65). □

6.4 Punkt- und Intervallschätzer

Entsprechend der Idee aus Abschnitt 6.2 wollen wir uns mit Schätzern der Form

$$\widehat{p}^{\gamma_{i,n}}_{i,n} = \widehat{p}^X_{i,n} - \gamma_{i,n}\left(\widehat{p}^Y_{i,n} - \frac{1}{2}\right)$$

beschäftigen. Wir werden versuchen, die $\gamma_{i,n}$ so zu wählen, dass die Varianz der nach Satz 6.2 zu $\sqrt{n}\left(\widehat{p}^{\gamma_{i,n}}_{i,n} - p^X_i\right)$ asymptotisch äquivalenten Zufallsvariablen

$$\sqrt{n}\left(\sum_{u=1}^{a}\frac{1}{n_u}\sum_{w=1}^{n_u}\left[\Psi^X_{uwi} - \mathrm{E}\left(\Psi^X_{uwi}\right)\right]\right) - \gamma_{i,n}\sqrt{n}\left(\sum_{u=1}^{a}\frac{1}{n_u}\sum_{w=1}^{n_u}\left[\Psi^Y_{uwi} - \mathrm{E}\left(\Psi^Y_{uwi}\right)\right]\right)$$

möglichst klein wird. Unter Verwendung der Bezeichnungen aus dem letzten Abschnitt erhalten wir

$$\begin{aligned}\mathrm{Var}\Bigg[&\sqrt{n}\left(\sum_{u=1}^{a}\frac{1}{n_u}\sum_{w=1}^{n_u}\left[\Psi^X_{uwi} - \mathrm{E}\left(\Psi^X_{uwi}\right)\right]\right)\\ &- \gamma_{i,n}\sqrt{n}\left(\sum_{u=1}^{a}\frac{1}{n_u}\sum_{w=1}^{n_u}\left[\Psi^Y_{uwi} - \mathrm{E}\left(\Psi^Y_{uwi}\right)\right]\right)\Bigg] = \sigma^X_{ii,n} - 2\gamma_{i,n}\sigma^{XY}_{ii,n} + \gamma^2_{i,n}\sigma^Y_{ii,n}.\end{aligned}$$

Durch Bildung der ersten Ableitung bestimmen wir

$$\gamma^{\min}_{i,n} = \frac{\sigma^{XY}_{ii,n}}{\sigma^Y_{ii,n}} = \underset{\gamma_{i,n}\in\mathbb{R}}{\operatorname{argmin}}\left\{\sigma^X_{ii,n} - 2\gamma_{i,n}\sigma^{XY}_{ii,n} + \gamma^2_{i,n}\sigma^Y_{ii,n}\right\}.$$

Mit Hilfe der im vorangehenden Abschnitt definierten Schätzer $\widehat{\sigma}^{XY}_{ii,n}$ und $\widehat{\sigma}^Y_{ii,n}$ können wir $\gamma^{\min}_{i,n}$ durch

$$\widehat{\gamma}^{\min}_{i,n} = \frac{\widehat{\sigma}^{XY}_{ii,n}}{\widehat{\sigma}^Y_{ii,n}}$$

schätzen. Die Varianz

$$\left(\sigma_{i,n}^{\min}\right)^2 = \sigma_{ii,n}^X - 2\gamma_{i,n}^{min}\sigma_{ii,n}^{XY} + \left(\gamma_{i,n}^{\min}\right)^2 \sigma_{ii,n}^Y = \sigma_{ii,n}^X - \frac{\left(\sigma_{ii,n}^{XY}\right)^2}{\sigma_{ii,n}^Y} \tag{6.5}$$

schätzen wir kanonisch durch

$$\left(\widehat{\sigma}_{i,n}^{\min}\right)^2 = \widehat{\sigma}_{ii,n}^X - 2\widehat{\gamma}_{i,n}^{\min}\widehat{\sigma}_{ii,n}^{XY} + \left(\widehat{\gamma}_{i,n}^{\min}\right)^2 \widehat{\sigma}_{ii,n}^Y = \widehat{\sigma}_{ii,n}^X - \frac{\left(\widehat{\sigma}_{ii,n}^{XY}\right)^2}{\widehat{\sigma}_{ii,n}^Y}.$$

An der Gleichung (6.5) können wir ablesen, dass die Varianz $\left(\sigma_{i,n}^{\min}\right)^2$ umso kleiner ist, je größer der Betrag der Kovarianz $\sigma_{ii,n}^{XY}$ ist. Aus diesem Grund sollten, wie in Abschnitt 6.1 bereits erwähnt wurde, nach Möglichkeit Kovariablen gewählt werden, die mit den eigentlichen Beobachtungen möglichst stark zusammenhängen.

Der folgende Satz zeigt, dass wir mit dem verbesserten Schätzer

$$\widehat{p}_{i,n}^{\min} = \widehat{p}_{i,n}^{\widehat{\gamma}_{i,n}^{\min}}$$

und dem Varianzschätzer $\left(\widehat{\sigma}_{i,n}^{\min}\right)^2$ asymptotisch so arbeiten können, wie wir es in den vorangehenden Kapiteln mit $\widehat{p}_{is,n}$ und $\widehat{\sigma}_{is,n}^2$ getan haben.

Satz 6.5. *Die Verteilung der Y_{ik} sei keine Einpunktverteilung und es seien die Annahmen (A1) und (A2) erfüllt. Dann gilt:*

(a) $\widehat{p}_{i,n}^{\min}$ *ist für alle $i \in \{1, \ldots, a\}$ ein stark konsistenter Schätzer für p_i^X.*

(b) $\left(\widehat{\sigma}_{i,n}^{\min}\right)^2$ *ist für alle $i \in \{1, \ldots, a\}$ ein stark konsistenter Schätzer für $\left(\sigma_{i,n}^{\min}\right)^2$ in dem Sinne, dass*

$$\left(\widehat{\sigma}_{i,n}^{\min}\right)^2 - \left(\sigma_{i,n}^{\min}\right)^2 \xrightarrow{f.s.} 0.$$

(c) *Falls zusätzlich zu den am Anfang dieses Satzes gestellten Voraussetzungen noch* $\liminf_{n\to\infty} \sigma_{i,n}^{\min} > 0$ *gilt, so konvergiert*

$$\sqrt{n}\,\frac{\widehat{p}_{i,n}^{\min} - p_i^X}{\widehat{\sigma}_{i,n}^{\min}}$$

in Verteilung gegen eine Standardnormalverteilung.

Beweis. Siehe Anhang A.18, Seite 117. □

Bemerkung. Die in dem obigen Satz formulierte Annahme, die Verteilung der Y_{ik} sei keine Einpunktverteilung, stellt für die Praxis keine relevante Einschränkung dar. Wenn nämlich alle Y_{ik} identisch sind, besteht ohnehin kein Interesse, die Y_{ik} als Kovariablen oder Ausgangswerte zu berücksichtigen. Wenn aber nicht alle Y_{ik} identisch sind, wissen wir bereits, dass keine Einpunktverteilung vorliegt.

Auf Grundlage von Teil (c) des obigen Satzes können wir ohne Schwierigkeiten Konfidenzintervalle für die relativen Effekte p_i^X konstruieren:

Korollar 6.6. *Unter den Voraussetzungen von Teil (c) des obigen Satzes ist*

$$\left[u_{i,n}^{\min}, o_{i,n}^{\min}\right] = \left[\widehat{p}_{i,n}^{\min} - \frac{\widehat{\sigma}_{i,n}^{\min} z_{1-\alpha/2}}{\sqrt{n}}, \widehat{p}_{i,n}^{\min} + \frac{\widehat{\sigma}_{i,n}^{\min} z_{1-\alpha/2}}{\sqrt{n}}\right]$$

für jedes $\alpha \in (0,1)$ ein asymptotisches $(1-\alpha)$-Konfidenzintervall für p_i^X.

Beweis. Nach Teil (c) von Satz 6.5 gilt

$$\mathrm{P}\left(p_i^X \in \left[u_{i,n}^{\min}, o_{i,n}^{\min}\right]\right) = \mathrm{P}\left(z_{\alpha/2} \leq \sqrt{n}\,\frac{\widehat{p}_{i,n}^{\min} - p_i^X}{\widehat{\sigma}_{i,n}^{\min}} \leq z_{1-\alpha/2}\right) \to 1 - \alpha.$$

□

Mit dem Ansatz der Varianzreduktion erhalten wir wegen

$$\widehat{\sigma}_{ii,n}^X \geq \widehat{\sigma}_{ii,n}^X - \frac{\left(\widehat{\sigma}_{ii,n}^{XY}\right)^2}{\widehat{\sigma}_{ii,n}^Y} = \left(\widehat{\sigma}_{i,n}^{\min}\right)^2$$

immer Konfidenzintervalle, die kürzer sind als die ohne Varianzreduktion konstruierten Konfidenzintervalle

$$\left[\widehat{p}_{i,n}^X - \frac{\sqrt{\widehat{\sigma}_{ii,n}^X}\, z_{1-\alpha/2}}{\sqrt{n}}, \widehat{p}_{i,n}^X + \frac{\sqrt{\widehat{\sigma}_{ii,n}^X}\, z_{1-\alpha/2}}{\sqrt{n}}\right].$$

Erwähnenswert ist auch, dass wir zur Konstruktion der verbesserten Punkt- und Intervallschätzer keine neuen Modellannahmen stellen mussten. Aus diesem Grund erscheint es lohnenswert, die in diesem Kapitel entwickelten Techniken in den Kontext der Dissertation von Langer (1998) zu übertragen. Diese befasst sich mit Modellen und Statistiken, die den in der vorliegenden Arbeit behandelten Modellen und Statistiken ähnlich sind, kommt aber bei der Varianzreduktion nicht ohne eine zusätzliche Modellannahme aus (siehe Langer, 1998, Annahme (R2), Seite 31).

Andererseits entsteht durch die Varianzreduktion auch eine neue Schwierigkeit: Während die Punktschätzer $\widehat{p}_{i,n}^X$ bereichserhaltend, also immer im Einheitsintervall $[0,1]$ enthalten sind, trifft diese Eigenschaft nicht mehr auf die Punktschätzer

$$\widehat{p}_{i,n}^{\min} = \widehat{p}_{i,n}^X - \widehat{\gamma}_{i,n}^{\min}\left(\widehat{p}_{i,n}^Y - \frac{1}{2}\right)$$

zu. Bei hinreichend großen $\widehat{\gamma}_{i,n}^{\min}$ liegt im Fall $\widehat{p}_{i,n}^Y \neq 1/2$ der Schätzer $\widehat{p}_{i,n}^{\min}$ außerhalb von $[0,1]$. Deshalb bietet sich hier schon bei der Konstruktion der Punktschätzer an, die Bereichserhaltung mit Hilfe der Transformationsmethode zu erzwingen. Dies beschreiben wir im abschließenden Abschnitt dieses Kapitels.

6.5 Transformationsmethode

Im Folgenden betrachten wir streng monotone, bijektive Transformationen $g : (0,1) \to \mathbb{R}$ mit den Eigenschaften, die sich in Abschnitt 5.6 als sinnvoll erwiesen haben. Das heißt, g sei stetig differenzierbar mit nullstellenfreier erster Ableitung und punktsymmetrisch um $(1/2, 0)$. Mit einer solchen Funktion g transformieren wir die relativen Effekte p_i^X auf $p_i^* = g\left(p_i^X\right)$ und schätzen diese durch

$$\widehat{p}_{i,n}^{*\gamma_i^*} = g\left(\widehat{p}_{i,n}^X\right) - \gamma_i^* \left[g\left(\widehat{p}_{i,n}^Y\right) - g\left(p_i^Y\right)\right].$$

Entsprechend der Grundidee dieses Kapitels werden wir versuchen, γ_i^* so zu wählen, dass die asymptotische Varianz von

$$\sqrt{n}\left(\widehat{p}_{i,n}^{*\gamma_i^*} - p_i^*\right)$$

möglichst klein wird. Mit diesem Ansatz erhalten wir Punktschätzer für die transformierten relativen Effekte, können aber auch Konfidenzintervalle konstruieren. Durch Anwendung der Umkehrfunktion g^{-1} auf die Schätzer ergeben sich wie in Abschnitt 5.6 Schätzer für die (nicht transformierten) relativen Effekte p_i^X.

Mathematische Grundlage für die Transformationsmethode ist der multivariate Delta-Satz. Um diesen bequem anwenden zu können, ersetzen wir in diesem Abschnitt die Bedingung (A2) durch die stärkere Annahme

(A2)' $n/n_i \to c_i \in \mathbb{R}$ für alle $i \in \{1, \dots, a\}$.

Aus dieser Annahme ergeben sich die Vereinfachungen

$$\begin{aligned}
\sigma_{uu',n}^X &= n \sum_{i=1}^a \frac{1}{n_i} \operatorname{Cov}\left(\Psi_{iku}^X, \Psi_{iku'}^X\right) \to \sum_{i=1}^a c_i \operatorname{Cov}\left(\Psi_{iku}^X, \Psi_{iku'}^X\right) = \sigma_{uu'}^X, \\
\sigma_{uu',n}^Y &= n \sum_{i=1}^a \frac{1}{n_i} \operatorname{Cov}\left(\Psi_{iku}^Y, \Psi_{iku'}^Y\right) \to \sum_{i=1}^a c_i \operatorname{Cov}\left(\Psi_{iku}^Y, \Psi_{iku'}^Y\right) = \sigma_{uu'}^Y, \\
\sigma_{uu',n}^{XY} &= n \sum_{i=1}^a \frac{1}{n_i} \operatorname{Cov}\left(\Psi_{iku}^X, \Psi_{iku'}^Y\right) \to \sum_{i=1}^a c_i \operatorname{Cov}\left(\Psi_{iku}^X, \Psi_{iku'}^Y\right) = \sigma_{uu'}^{XY}
\end{aligned}$$

sowie $\boldsymbol{V}_n \to \boldsymbol{V}$.

An Stelle von Satz 6.3 und Satz 6.4 erhalten wir die folgenden beiden stärkeren Resultate.

Satz 6.7. *Unter den Voraussetzungen (A1) und (A2)' gilt*

$$\sqrt{n}\left[\left(\widehat{\boldsymbol{p}}_n^X - \boldsymbol{p}^X\right)', \left(\widehat{\boldsymbol{p}}_n^Y - \boldsymbol{p}^Y\right)'\right]' \xrightarrow{\mathcal{D}} \mathrm{N}(\boldsymbol{0}, \boldsymbol{V}).$$

Beweis. Siehe Anhang A.19, Seite 120. □

Satz 6.8. *Unter den Voraussetzungen (A1) und (A2)' ist $\widehat{\boldsymbol{V}}_n$ ein stark konsistenter Schätzer für $\boldsymbol{V}$ und es gilt*

$$\sum_{u=1}^{a}\sum_{u'=1}^{a}\left[\mathrm{E}\left(\left(\widehat{\sigma}^X_{uu',n}-\sigma^X_{uu'}\right)^2\right)+\mathrm{E}\left(\left(\widehat{\sigma}^Y_{uu',n}-\sigma^Y_{uu'}\right)^2\right)+\mathrm{E}\left(\left(\widehat{\sigma}^{XY}_{uu',n}-\sigma^{XY}_{uu'}\right)^2\right)\right]\to 0.$$

Beweis. Der Beweis ergibt sich sofort aus Satz 6.4 und $\boldsymbol{V}_n \to \boldsymbol{V}$. □

Der multivariate Delta-Satz liefert

Satz 6.9. *Unter den Annahmen (A1) und (A2)' gilt für jede stetig differenzierbare Funktion $g:(0,1)\to\mathbb{R}$ und alle $i\in\{1,\ldots,a\}$ mit $0<p_i^X<1$*

$$\sqrt{n}\begin{pmatrix} g\left(\widehat{p}^X_{i,n}\right)-g\left(p^X_i\right)\\ g\left(\widehat{p}^Y_{i,n}\right)-g\left(p^Y_i\right)\end{pmatrix}$$
$$\xrightarrow{\mathcal{D}} \mathrm{N}\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}\left[g'\left(p^X_i\right)\right]^2\sigma^X_{ii} & g'\left(p^X_i\right)g'\left(p^Y_i\right)\sigma^{XY}_{ii}\\ g'\left(p^X_i\right)g'\left(p^Y_i\right)\sigma^{XY}_{ii} & \left[g'\left(p^Y_i\right)\right]^2\sigma^Y_{ii}\end{pmatrix}\right).$$

Beweis. Siehe Anhang A.20, Seite 120. □

Korollar 6.10. *Unter den Annahmen (A1) und (A2)' konvergiert für jede stetig differenzierbare Funktion $g:(0,1)\to\mathbb{R}$ und alle $i\in\{1,\ldots,a\}$ mit $0<p_i^X<1$ und*

$$0<\left(\sigma_i^{*\gamma_i^*}\right)^2=\left[g'\left(p^X_i\right)\right]^2\sigma^X_{ii}-2\gamma_i^* g'\left(p^X_i\right)g'\left(p^Y_i\right)\sigma^{XY}_{ii}+\left(\gamma_i^*\right)^2\left[g'\left(p^Y_i\right)\right]^2\sigma^Y_{ii}$$

die Verteilung von

$$\sqrt{n}\,\frac{\widehat{p}^{*\gamma_i^*}_{i,n}-p_i^*}{\sigma_i^{*\gamma_i^*}}$$

schwach gegen eine Standardnormalverteilung.

Beweis. Der Beweis ergibt sich sofort aus dem obigen Satz und dem Satz von der stetigen Abbildung. □

Die asymptotische Varianz $\left(\sigma_i^{*\gamma_i^*}\right)^2$ wird minimal, wenn wir als Wert für γ_i^*

$$\gamma_i^{*\,\min}=\frac{g'\left(p^X_i\right)g'\left(p^Y_i\right)\sigma^{XY}_{ii}}{\left[g'\left(p^Y_i\right)\right]^2\sigma^Y_{ii}}=\frac{g'\left(p^X_i\right)\sigma^{XY}_{ii}}{g'\left(p^Y_i\right)\sigma^Y_{ii}}$$

wählen.

Mit Hilfe der Schätzer $\widehat{p}^X_{i,n}$, $\widehat{\sigma}^{XY}_{ii,n}$ und $\widehat{\sigma}^Y_{ii,n}$ können wir $\gamma_i^{*\,\min}$ durch

$$\widehat{\gamma}^{*\,\min}_{i,n}=\frac{g'\left(\widehat{p}^X_{i,n}\right)\widehat{\sigma}^{XY}_{ii,n}}{g'\left(\frac{1}{2}\right)\widehat{\sigma}^Y_{ii,n}}$$

schätzen. Die Varianz

$$\begin{aligned}\left(\sigma_i^{*\min}\right)^2 &= \left[g'\left(p_i^X\right)\right]^2 \sigma_{ii}^X - 2\gamma_i^{*\min} g'\left(p_i^X\right) g'\left(p_i^Y\right) \sigma_{ii}^{XY} + \left(\gamma_i^{*\min}\right)^2 \left[g'\left(p_i^Y\right)\right]^2 \sigma_{ii}^Y \\ &= \left[g'\left(p_i^X\right)\right]^2 \left(\sigma_{ii}^X - \frac{\left(\sigma_{ii}^{XY}\right)^2}{\sigma_{ii}^Y}\right)\end{aligned}$$

schätzen wir kanonisch durch

$$\begin{aligned}\left(\widehat{\sigma}_{i,n}^{*\min}\right)^2 &= \left[g'\left(\widehat{p}_{i,n}^X\right)\right]^2 \widehat{\sigma}_{ii,n}^X - 2\widehat{\gamma}_{i,n}^{*\min} g'\left(\widehat{p}_{i,n}^X\right) g'\left(\frac{1}{2}\right) \widehat{\sigma}_{ii,n}^{XY} \\ &\quad + \left(\widehat{\gamma}_{i,n}^{*\min}\right)^2 \left[g'\left(\frac{1}{2}\right)\right]^2 \widehat{\sigma}_{ii,n}^Y \\ &= \left[g'\left(\widehat{p}_{i,n}^X\right)\right]^2 \left(\widehat{\sigma}_{ii,n}^X - \frac{\left(\widehat{\sigma}_{ii,n}^{XY}\right)^2}{\widehat{\sigma}_{ii,n}^Y}\right).\end{aligned}$$

Der folgende Satz zeigt, dass wir mit dem verbesserten Schätzer

$$\widehat{p}_{i,n}^{*\min} = \widehat{p}_{i,n}^{*\widehat{\gamma}_{i,n}^{*\min}}$$

und dem Varianzschätzer $\left(\widehat{\sigma}_{i,n}^{*\min}\right)^2$ asymptotisch so arbeiten können, wie wir es in Abschnitt 5.6 mit $\widehat{p}_{is,n}^*$ und $\left(\widehat{\sigma}_{is,n}^*\right)^2$ getan haben.

Satz 6.11. *Die Verteilung der Y_{ik} sei keine Einpunktverteilung, $g : (0,1) \to \mathbb{R}$ sei eine streng monotone, stetig differenzierbare Bijektion mit nullstellenfreier erster Ableitung und, es seien die Annahmen (A1) und (A2)' erfüllt. Dann gilt für alle $i \in \{1, \ldots, a\}$ mit $0 < p_i^X < 1$:*

(a) *$\widehat{p}_{i,n}^{*\min}$ ist ein stark konsistenter Schätzer für p_i^*.*

(b) *$\left(\widehat{\sigma}_{i,n}^{*\min}\right)^2$ ist ein stark konsistenter Schätzer für $\left(\sigma_{i,n}^{*\min}\right)^2$.*

(c) *Falls zusätzlich zu den am Anfang dieses Satzes gestellten Voraussetzungen noch $\sigma_i^{*\min} > 0$ gilt, so konvergiert*

$$\sqrt{n}\,\frac{\widehat{p}_{i,n}^{*\min} - p_i^*}{\widehat{\sigma}_{i,n}^{*\min}}$$

in Verteilung gegen eine Standardnormalverteilung.

Beweis. Siehe Anhang A.21, Seite 121. □

Durch Anwendung der inversen Transformation g^{-1} erhalten wir bereichserhaltende Punkt- und Intervallschätzer:

Korollar 6.12. *Unter den am Anfang des obigen Satzes formulierten Bedingungen gilt für alle $i \in \{1, \ldots, a\}$ mit $0 < p_i^X < 1$:*

(a) *$g^{-1}\left(\widehat{p}_{i,n}^{*\min}\right) \in (0,1)$ ist ein stark konsistenter Schätzer für p_i^X.*

(b) *Unter der zusätzlichen Bedingung $\sigma_i^{*\min} > 0$ ist*

$$\left[u_{i,n}^{T\min}, o_{i,n}^{T\min}\right] = g^{-1}\left(\left[\widehat{p}_{i,n}^{*\min} - \frac{\widehat{\sigma}_{i,n}^{*\min} z_{1-\alpha/2}}{\sqrt{n}}, \widehat{p}_{i,n}^{*\min} + \frac{\widehat{\sigma}_{i,n}^{*\min} z_{1-\alpha/2}}{\sqrt{n}}\right]\right) \subset (0,1)$$

für alle $\alpha \in (0,1)$ ein asymptotisches $(1-\alpha)$-Konfidenzintervall für p_i^X.

Beweis. Behauptung (a) ergibt sich aus obigem Satz und der Stetigkeit von g^{-1}. Teil (b) folgt sofort aus

$$\mathrm{P}\left(p_i^X \in \left[u_{i,n}^{T\min}, o_{i,n}^{T\min}\right]\right) = \mathrm{P}\left(z_{\alpha/2} \leq \sqrt{n}\,\frac{\widehat{p}_{i,n}^{*\min} - p_i^*}{\widehat{\sigma}_{i,n}^{*\min}} \leq z_{1-\alpha/2}\right) \to 1 - \alpha.$$

□

Anhang A

Beweise

In diesem Anhang befinden sich die Beweise zu den Ergebnissen des Hauptteils dieser Arbeit, deren Umfang über wenige Zeilen hinausgeht. Dementsprechend sollte dieses Kapitel nicht für sich allein gelesen werden, sondern es sollte während der Lektüre des Hauptteils bei Bedarf zum Nachschlagen verwendet werden. Um das Suchen der Beweise zu erleichtern, ist für jeden Beweis ein eigener Abschnitt angelegt worden.

Bei allen Beweisen in dieser Arbeit werden die grundlegenden Regeln der Maß- und Wahrscheinlichkeitstheorie, wie zum Beispiel die Monotonie und die Linearität des Erwartungswerts, stillschweigend verwendet. Eine kurze und präzise Übersicht zu den wichtigsten Ausgangspunkten der asymptotischen Statistik findet man in den Kapiteln 2 und 3 des Lehrbuchs von van der Vaart (1998).

A.1 Beweis von Satz 3.3

Beweis von Satz 3.3. (a) Falls $F_1 = F_2$ ist, X_{11} also die gleiche Verteilung wie X_{21} hat, so ergibt sich aus Symmetriegründen

$$\begin{aligned} p_1 &= \mathrm{P}(X_{21} < X_{11}) + \frac{1}{2}\,\mathrm{P}(X_{21} = X_{11}) \\ &= \mathrm{P}(X_{11} < X_{21}) + \frac{1}{2}\,\mathrm{P}(X_{11} = X_{21}) = p_2. \end{aligned}$$

Mit $p_1 + p_2 = 1$ folgt sofort $p_1 = p_2 = \frac{1}{2}$.

Falls X_{11} und X_{21} symmetrisch um das gleiche Symmetriezentrum μ verteilt sind, so gilt

$$X_{i1} - \mu \sim \mu - X_{i1}, \quad i = 1, 2,$$

und deshalb

$$\begin{aligned} p_1 &= \mathrm{P}(X_{21} < X_{11}) + \frac{1}{2}\,\mathrm{P}(X_{21} = X_{11}) \\ &= \mathrm{P}(X_{21} - \mu < X_{11} - \mu) + \frac{1}{2}\,\mathrm{P}(X_{21} = X_{11}) \\ &= \mathrm{P}(\mu - X_{21} < \mu - X_{11}) + \frac{1}{2}\,\mathrm{P}(X_{21} = X_{11}) \\ &= \mathrm{P}(X_{11} < X_{21}) + \frac{1}{2}\,\mathrm{P}(X_{11} = X_{21}) = p_2. \end{aligned}$$

Der zweite Teil der Behauptung aus Teil (a) des Satzes ergibt sich also wieder aus $p_1 + p_2 = 1$.

(b) Sei o. E. d. A. $\mu_1 < \mu_2$. Da die Verteilungsfunktion $F_2(\cdot + \mu_2 - \mu_1)$ das Symmetriezentrum μ_1 hat, folgt aus Teil (a)

$$\int F_2(x + \mu_2 - \mu_1)\, dF_1(x) = \frac{1}{2}.$$

Wegen $F_2 \leq F_2(\cdot + \mu_2 - \mu_1)$ erhalten wir

$$p_1 = \int F_2(x)\, dF_1(x) \leq \int F_2(x + \mu_2 - \mu_1)\, dF_1(x) = \frac{1}{2}.$$

(c) Sei o. E. d. A. F_1 stochastisch größer als F_2, das heißt, es gelte $F_1(x) \leq F_2(x)$ für alle $x \in \mathbb{R}$ und es existiere ein x_0 mit $F_1(x_0) < F_2(x_0)$. Aus Teil (a) dieses Satzes und der Integraldarstellung relativer Effekte ergibt sich

$$\int F_1\, dF_1 = \frac{1}{2}.$$

Deshalb genügt es

$$\int (F_2 - F_1)\, dF_1 = p_1 - \int F_1\, dF_1 > 0$$

zu zeigen.

Nach Voraussetzung gilt

$$F_2(x_0) - F_1(x_0) = \epsilon > 0.$$

Wegen $F_1(x) \to 1$ für $x \to \infty$ erhalten wir

$$y_0 = \sup\left\{ y : y \geq x_0,\ F_2(x_0) - F_1(y) \geq \frac{\epsilon}{2} \right\} < \infty.$$

Wir unterscheiden nun zwei Fälle in Abhängigkeit davon, ob F_1 in y_0 stetig ist.

Sei zunächst F_1 stetig in y_0. Daraus folgt $F_2(x_0) - F_1(y_0) = \epsilon/2$ und deshalb $F_1(y_0) - F_1(x_0) = \epsilon/2$. Für alle $x \in [x_0, y_0]$ gilt $F_2(x) - F_1(x) \geq \epsilon/2$. Insgesamt erhalten wir

$$\int (F_2 - F_1)\, dF_1 \geq \int_{x_0}^{y_0} (F_2 - F_1)\, dF_1 \geq \int_{x_0}^{y_0} \frac{\epsilon}{2}\, dF_1 \geq \frac{\epsilon}{2} \cdot \frac{\epsilon}{2} = \frac{\epsilon^2}{4} > 0.$$

Sei nun F_1 nicht stetig in y_0, das heißt, F_1 habe eine Sprungstelle mit einer Höhe $\delta > 0$ in y_0. Wegen

$$F_2^-(y_0) - F_1^-(y_0) = \lim_{x \uparrow y_0} (F_2(x) - F_1(x)) \geq \frac{\epsilon}{2}$$

und

$$F_2^+(y_0) - F_1^+(y_0) = \lim_{x \downarrow y_0} (F_2(x) - F_1(x)) \geq 0$$

gilt

$$F_2(y_0) - F_1(y_0) = \frac{1}{2}[F_2^-(y_0) - F_1^-(y_0)] + \frac{1}{2}[F_2^+(y_0) - F_1^+(y_0)] \geq \frac{\epsilon}{4}.$$

Wir erhalten

$$\int (F_2 - F_1)\, dF_1 \geq \int_{\{y_0\}} (F_2 - F_1)\, dF_1 \geq \int_{\{y_0\}} \frac{\epsilon}{4}\, dF_1 \geq \frac{\epsilon}{4} \cdot \delta > 0.$$

□

A.2 Beweis von Satz 3.7

Beweis von Satz 3.7. Wegen Korollar 3.4 auf Seite 25 gilt

$$\begin{aligned}
p_{is} &= \int H_{is}\, dF_{is} = \int \frac{1}{at-1} \left(\sum_{(u,v)\neq(i,s)} F_{uv} \right) dF_{is} \\
&= \int \frac{1}{at-1} \left(\sum_{(u,v)\neq(i,s)} F_{uv} + F_{is} - F_{is} \right) dF_{is} \\
&= \frac{1}{at-1} \int \left(\sum_{u=1}^{a} \sum_{v=1}^{t} F_{uv} \right) dF_{is} - \frac{1}{2(at-1)} \\
&= \frac{at}{at-1} \int \overline{F}_{..}\, dF_{is} - \frac{1}{2(at-1)}.
\end{aligned}$$

Die zweite Gleichung ergibt sich wie in (3.7) auf Seite 23 aus dem Satz von Fubini.

□

A.3 Beweis von Satz 3.9

Beweis von Satz 3.9. Sei zunächst F_{is} mit Bezug auf $(F_{is}+F_{uv})/2$ tendenziell kleiner als F_{uv}. Unter Verwendung der Korollare 3.2 und 3.4 auf Seite 25 erhalten wir durch Äquivalenzumformungen

$$\begin{aligned}
& \int \frac{F_{is}+F_{uv}}{2}\, dF_{is} < \int \frac{F_{is}+F_{uv}}{2}\, dF_{uv} \\
\Leftrightarrow\ & \int F_{is}\, dF_{is} + \int F_{uv}\, dF_{is} < \int F_{is}\, dF_{uv} + \int F_{uv}\, dF_{uv} \\
\Leftrightarrow\ & \int F_{uv}\, dF_{is} < \int F_{is}\, dF_{uv} \\
\Leftrightarrow\ & \int F_{uv}\, dF_{is} < 1 - \int F_{uv}\, dF_{is} \\
\Leftrightarrow\ & \int F_{uv}\, dF_{is} < \frac{1}{2}\,.
\end{aligned}$$

Damit ist Teil (a) bewiesen. Die anderen beiden Teile können analog bewiesen werden.

□

A.4 Beweis von Satz 3.10

Beweis von Satz 3.10. Für eine natürliche Zahl m und reelle Zahlen x_{iks}, $i = 1, \ldots, a$, $k = 1, \ldots, m$, $s = 1, \ldots, t$, bezeichne r_{iks}^M für jede Teilmenge $M \subset \{1, \ldots, a\} \times \{1, \ldots, t\}$ den Mittelrang von x_{iks} unter allen Zahlen

$$\{x_{iks} : (i, s) \in M,\ k = 1, \ldots, m\}.$$

Die Rangsummen

$$r_{i \cdot s}^M = \sum_{k=1}^{m} r_{iks}^M$$

definieren für gegebene x_{iks} in natürlicher Weise ein Wort auf $\{1, \ldots, a\} \times \{1, \ldots, t\}$.

Identifizieren wir $\{C_1, \ldots, C_{at}\}$ durch die Bijektion $C_{is} \mapsto (i, s)$ mit $\{1, \ldots, a\} \times \{1, \ldots, t\}$, so können wir dieses Wort als ein Wort auf $\{C_1, \ldots, C_{at}\}$ auffassen. Die Menge aller auf diese Art erzeugter Wörter heißt nach Haunsperger (1992) *Kruskal-Wallis-Wörterbuch* auf $\{C_1, \ldots, C_{at}\}$. In Haunsperger (1992) wird die Übereinstimmung des Kruskal-Wallis-Wörterbuches auf $\{C_1, \ldots, C_{at}\}$ mit dem Borda-Wörterbuch auf $\{C_1, \ldots, C_{at}\}$ gezeigt. Wir müssen also nur noch beweisen, dass das Kruskal-Wallis-Wörterbuch in dem Wörterbuch der tendenziellen Relationen enthalten ist.

Dazu betrachten wir das Kruskal-Wallis-Wörterbuch wieder als ein Wörterbuch auf $\{1, \ldots, a\} \times \{1, \ldots, t\}$. Für gegebene x_{iks}, $i = 1, \ldots, a$, $k = 1, \ldots, m$, $s = 1, \ldots, t$ sei F_{is} die durch

$$F_{is}(x) = \frac{1}{m} \sum_{k=1}^{m} c(x - x_{iks})$$

definierte Verteilungsfunktion. Durch Abzählen erhalten wir für alle Teilmengen $M \subset \{1, \ldots, a\} \times \{1, \ldots, t\}$ mit den Bezeichnungen $\overline{F}_M = \sum_{(i,s) \in M} F_{is} / |M|$ und $\overline{r}_{i \cdot s}^M = r_{i \cdot s}^M / m$ die Identität

$$\int \overline{F}_M \, dF_{is} = \frac{1}{m\,|M|} \left(\overline{r}_{i \cdot s}^M - \frac{1}{2} \right).$$

Also stimmt das zu den x_{ik} gehörige Wort im Kruskal-Wallis-Wörterbuch mit dem zu den Verteilungsfunktionen F_{is} gehörenden Wort im Wörterbuch der tendenziellen Relationen überein. □

A.5 Beweis von Satz 3.11

Beweis von Satz 3.11. Wir zeigen die folgenden Implikationen:

(a) Wenn F_{is} tendenziell kleiner als F_{uv} ist, dann ist F_{is} mit Bezug auf $\overline{F}_{\cdot\cdot}$ tendenziell kleiner als F_{uv}.

(b) Wenn F_{is} und F_{uv} tendenziell gleich sind, dann sind F_{is} und F_{uv} mit Bezug auf $\overline{F}_{\cdot\cdot}$ tendenziell gleich.

(c) Wenn F_{is} tendenziell größer als F_{uv} ist, dann ist F_{is} mit Bezug auf $\overline{F}_{\cdot\cdot}$ tendenziell größer als F_{uv}.

Zusammen ergeben diese Implikationen die behaupteten Äquivalenzen.

(a) F_{is} sei tendenziell kleiner als F_{uv}, das heißt,

$$\int F_{uv}\,dF_{is} = \mathrm{P}(X_{u2v1} < X_{i1s1}) + \frac{1}{2}\,\mathrm{P}(X_{u2v1} = X_{i1s1}) < \frac{1}{2}.$$

Da die Verteilungsfunktionen stochastisch geordnet sind, muss nach Satz 3.3 (siehe Seite 25) F_{is} auch stochastisch kleiner als F_{uv} sein, also $F_{is} \geq F_{uv}$ gelten. Wir zeigen unter Verwendung der Darstellung aus Satz 3.7 (siehe Seite 33), partieller Integration und Korollar 3.4 (siehe Seite 25) die zur Behauptung, F_{is} sei mit Bezug auf $\overline{F}_{\cdot\cdot}$ tendenziell kleiner als F_{uv}, äquivalente Aussage $p_{uv} - p_{is} > 0$.

$$\begin{aligned}
p_{uv} - p_{is} &= \frac{at}{at-1}\int \overline{F}_{\cdot\cdot}\,dF_{uv} - \frac{1}{2(at-1)} \\
&\quad - \left(\frac{at}{at-1}\int \overline{F}_{\cdot\cdot}\,dF_{is} - \frac{1}{2(at-1)}\right) \\
&= \frac{at}{at-1}\int \overline{F}_{\cdot\cdot}\,d(F_{uv} - F_{is}) \\
&= \frac{at}{at-1}\int (F_{is} - F_{uv})\,d\overline{F}_{\cdot\cdot} \\
&= \frac{at}{at-1}\int (F_{is} - F_{uv})\,d\left(\frac{1}{at}\sum_{i'=1}^{a}\sum_{s'=1}^{t} F_{i's'}\right) \\
&\geq \frac{1}{at-1}\int (F_{is} - F_{uv})\,dF_{is} \\
&= \frac{1}{at-1}\left(\frac{1}{2} - \int F_{uv}\,dF_{is}\right) > 0
\end{aligned}$$

(b) Nach Satz 3.3 können die stochastisch geordneten Verteilungsfunktionen F_{is} und F_{uv} nur dann tendenziell gleich sein, wenn sie identisch sind. Mit der Darstellung aus Satz 3.7 folgt sofort $p_{is} = p_{uv}$.

(c) Ergibt sich bei Vertauschung der Verteilungsfunktionen aus Teil (a).

□

A.6 Beweis von Satz 3.12

Das folgende Lemma dient als Vorbereitung für die Beweise der Sätze 3.12 und 3.13.

Lemma A.1. *Für die Mittelwerte der relativen Effekte $\overline{p}_{i\cdot}$ und $\overline{p}_{\cdot s}$ gilt*

$$\overline{p}_{i\cdot} = \frac{at-t}{at-1}\,p_{i*} + \frac{t-1}{2(at-1)} \quad \textit{und} \quad \frac{at-a}{at-1}\,p_{*s} + \frac{a-1}{2(at-1)}.$$

Beweis. Nach Satz 3.7 auf Seite 33 und Korollar 3.4 auf Seite 25 gilt

$$\begin{aligned}
\overline{p}_{i\cdot} &= \frac{1}{t}\sum_{s=1}^{t}\int H_{is}\,dF_{is} \\
&= \frac{1}{t}\sum_{s=1}^{t}\left(\frac{at}{at-1}\int \overline{F}_{\cdot\cdot}\,dF_{is} - \frac{1}{2(at-1)}\right) \\
&= \frac{at}{at-1}\int \overline{F}_{\cdot\cdot}\,d\overline{F}_{i\cdot} - \frac{1}{2(at-1)} \\
&= \frac{at}{at-1}\int \left(\frac{a-1}{a}H_{i*} + \frac{1}{a}\overline{F}_{i\cdot}\right)d\overline{F}_{i\cdot} - \frac{1}{2(at-1)} \\
&= \frac{at}{at-1}\left(\frac{a-1}{a}\int H_{i*}\,d\overline{F}_{i\cdot} + \frac{1}{a}\int \overline{F}_{i\cdot}\,d\overline{F}_{i\cdot}\right) - \frac{1}{2(at-1)} \\
&= \frac{at-t}{at-1}p_{i*} + \frac{t}{2(at-1)} - \frac{1}{2(at-1)} \\
&= \frac{at-t}{at-1}p_{i*} + \frac{t-1}{2(at-1)}.
\end{aligned}$$

Die zweite Behauptung ergibt sich analog. □

Beweis von Satz 3.12. Die Behauptungen folgen mit Lemma A.1 aus $p_{i*} \in [0,1]$ und $p_{*s} \in [0,1]$. □

A.7 Beweis von Satz 3.13

Beweis von Satz 3.13. Nach Definition von $\boldsymbol{\nu}_{i*}$ und $\boldsymbol{\nu}_{*s}$ genügt es,

$$p_{i*} = \frac{at-1}{at-t}\overline{p}_{i\cdot} - \frac{t-1}{2(at-t)} \quad \text{und} \quad p_{*s} = \frac{at-1}{at-a}\overline{p}_{\cdot s} - \frac{a-1}{2(at-a)}$$

zu zeigen. Dies ergibt sich nach elementaren Umformungen aus Lemma A.1. □

A.8 Beweis von Satz 4.1

Lemma A.2. *Für alle $x \in \mathbb{R}$ gilt*

$$\operatorname{Var}\left[\widehat{F}_{iks}(x)\right] \leq \frac{1}{4}.$$

Beweis. Mit Hilfe der Jensenschen Ungleichung ergibt sich

$$\begin{aligned}
\operatorname{Var}\left[\widehat{F}_{iks}(x)\right] &= \operatorname{Var}\left[\frac{1}{m_{iks}}\sum_{r=1}^{m_{iks}}\widehat{F}_{iksr}(x)\right] \\
&= \operatorname{E}\left[\frac{1}{m_{iks}}\sum_{r=1}^{m_{iks}}\widehat{F}_{iksr}(x) - F_{is}(x)\right]^2
\end{aligned}$$

$$
\begin{aligned}
&\leq \mathrm{E}\left[\frac{1}{m_{iks}}\sum_{r=1}^{m_{iks}}\left[\widehat{F}_{iksr}(x) - F_{is}(x)\right]^2\right] \\
&= \mathrm{E}\left[\widehat{F}_{iksr}(x) - F_{is}(x)\right]^2 \\
&= \mathrm{Var}\left[\widehat{F}_{iksr}(x)\right] \\
&= F_{is}(x)[1 - F_{is}(x)] \leq \frac{1}{4}.
\end{aligned}
$$

□

Beweis von Satz 4.1. Die Erwartungstreue haben wir bereits durch die Konstruktion sichergestellt.

Die erste Abschätzung für die Varianzen ergibt sich mit Hilfe von Lemma A.2 folgendermaßen:

$$
\begin{aligned}
\mathrm{Var}\left(\widehat{F}_{is,n}(x)\right) &= \mathrm{Var}\left(\frac{1}{\lambda_{i\cdot s}}\sum_{k=1}^{n_i}\frac{1}{m_{iks}}\sum_{r=1}^{m_{iks}}\widehat{F}_{iksr}(x)\right) \\
&= \frac{1}{\lambda_{i\cdot s}^2}\sum_{k=1}^{n_i}\mathrm{Var}\left(\frac{1}{m_{iks}}\sum_{r=1}^{m_{iks}}\widehat{F}_{iksr}(x)\right) \\
&\leq \frac{1}{\lambda_{i\cdot s}^2}\sum_{k=1}^{n_i}\mathrm{Var}\left(\widehat{F}_{iks}(x)\right) \leq \frac{1}{4\lambda_{i\cdot s}}.
\end{aligned}
$$

Die zweite Abschätzung folgt aus dieser Ungleichung mit Hilfe der üblichen Rechenregeln für Varianzen. □

A.9 Beweis von Satz 4.2

Beweis von Satz 4.2. Wegen der gleichmäßigen Beschränktheit der Zufallsvariablen $\widehat{F}_{iks}(x)$, $k = 1, \ldots, n_i$, gilt nach dem Starken Gesetz der großen Zahl in der Darstellung von Krengel (1991, Satz 11.4) für alle $x \in \mathbb{R}$

$$
\widehat{F}_{is,n}(x) = \frac{1}{\lambda_{i\cdot s}}\sum_{k=1}^{n_i}\widehat{F}_{iks}(x) \xrightarrow{\text{f.s.}} F_{is}(x).
$$

Das gleiche Resultat kann man völlig analog für die links- und rechtsstetigen Versionen der Verteilungsfunktionen und empirischen Verteilungsfunktionen beweisen.

Für jedes $\epsilon > 0$ existiert nun ein Menge $\boldsymbol{x} = \{x_0, x_1, \ldots, x_m\}$ mit $-\infty = x_0 < x_1 < \cdots < x_m = \infty$, so dass

$$
F_{is}^-(x_i) - F_{is}^+(x_{i-1}) < \epsilon
$$

für alle $i \in \{1, \ldots, m\}$ gilt. Für alle $x \in \mathbb{R} \setminus \boldsymbol{x}$ gibt es ein $i \in \{1, \ldots, m\}$ mit $x_{i-1} < x < x_i$ und deshalb

$$
\widehat{F}_{is,n}^+(x_{i-1}) - F_{is}^+(x_{i-1}) - \epsilon \leq \widehat{F}_{is,n}(x) - F_{is}(x) \leq \widehat{F}_{is,n}^-(x_i) - F_{is}^-(x_i) + \epsilon.
$$

Da die fast sichere Konvergenz von $\widehat{F}_{is,n}$, $\widehat{F}_{is}^{-}$ und $\widehat{F}_{is}^{+}$ sicherlich gleichmäßig auf der endlichen Menge $\boldsymbol{x}$ ist, erhalten wir

$$\limsup_{n\to\infty} \left\| \widehat{F}_{is,n} - F_{is} \right\|_{\infty} \overset{\text{f. s.}}{\leq} \epsilon.$$

Da dies für jedes $\epsilon > 0$ zutrifft, folgt die erste Behauptung des Satzes. Die zweite ergibt sich jetzt aus der Dreiecksungleichung. □

A.10 Beweis von Satz 4.3

Beweis von Satz 4.3. Es gilt

$$\begin{aligned} |\widehat{p}_{is,n} - p_{is}| &= \left| \int \widehat{H}_{is,n}\, d\widehat{F}_{is,n} - \int H_{is}\, dF_{is} \right| \\ &= \left| \int \left(\widehat{H}_{is,n} - H_{is}\right)\, d\widehat{F}_{is,n} + \int H_{is}\, d\left(\widehat{F}_{is,n} - F_{is}\right) \right| \\ &\leq \left| \int \left(\widehat{H}_{is,n} - H_{is}\right)\, d\widehat{F}_{is,n} \right| + \left| \int H_{is}\, d\left(\widehat{F}_{is,n} - F_{is}\right) \right|. \end{aligned}$$

Auf den ersten Summanden können wir Satz 4.2 (siehe Seite 48) sofort anwenden.

$$\left| \int \left(\widehat{H}_{is,n} - H_{is}\right)\, d\widehat{F}_{is,n} \right| \leq \left\| \widehat{H}_{is,n} - H_{is} \right\|_{\infty} \xrightarrow{\text{f.s.}} 0$$

Den zweiten Summanden integrieren wir zunächst partiell,

$$\begin{aligned} \int H_{is}\, d\left(\widehat{F}_{is,n} - F_{is}\right) &= \int H_{is}\, d\widehat{F}_{is,n} - \int H_{is}\, dF_{is} \\ &= 1 - \int \widehat{F}_{is,n}\, dH_{is} - \left(1 - \int F_{is}\, dH_{is}\right) \\ &= \int \left(F_{is} - \widehat{F}_{is,n}\right)\, dH_{is}, \end{aligned}$$

um dann auch Satz 4.2 anwenden zu können:

$$\left| \int \left(F_{is} - \widehat{F}_{is,n}\right)\, dH_{is} \right| \leq \left\| \widehat{F}_{is,n} - F_{is} \right\|_{\infty} \xrightarrow{\text{f.s.}} 0.$$

□

A.11 Beweis von Satz 4.5

Beweis von Satz 4.5. Es sei $0 < \epsilon < 1$ und die Verteilung der Zufallsvariablen $\boldsymbol{X}_{ik}$, $i = 1, \ldots, a$, $k = 1, \ldots, n_i$, sei durch $P_{(1-\epsilon)\boldsymbol{F}+\epsilon\boldsymbol{G}}$ gegeben. Unter der Bedingung (A1) auf Seite 13 folgt aus Satz 4.3 (siehe Seite 50)

$$\begin{aligned}
\widehat{p}_{is,n} &\xrightarrow{\text{f.s.}} \int \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} \left[(1-\epsilon)F_{uv}+\epsilon G_{uv}\right] \, d\left[(1-\epsilon)F_{is}+\epsilon G_{is}\right] \\
&= (1-\epsilon)^2 p_{is} + (1-\epsilon)\epsilon \int H_{is}\, dG_{is} \\
&\quad +\epsilon(1-\epsilon) \int \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} G_{uv}\, dF_{is} \\
&\quad +\epsilon^2 \int \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} G_{uv}\, dG_{is}
\end{aligned} \tag{A.1}$$

Wegen

$$\begin{aligned}
\int H_{is}\, dG_{is} &\in [0,1], \\
\int \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} G_{uv}\, dF_{is} &\in [0,1] \quad \text{und} \\
\int \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} G_{uv}\, dG_{is} &\in [0,1]
\end{aligned}$$

gilt

$$\begin{aligned}
(1-\epsilon)\epsilon \int H_{is}\, dG_{is} + \epsilon(1-\epsilon) \int \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} G_{uv}\, dF_{is} \\
+ \epsilon^2 \int \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} G_{uv}\, dG_{is} \in \left[0, 1-(1-\epsilon)^2\right].
\end{aligned}$$

Daraus ergibt sich

$$\begin{aligned}
(1-\epsilon)^2 p_{is} + (1-\epsilon)\epsilon \int H_{is}\, dG_{is} + \epsilon(1-\epsilon) \int \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} G_{uv}\, dF_{is} \\
+ \epsilon^2 \int \frac{1}{at-1} \sum_{(u,v)\neq(i,s)} G_{uv}\, dG_{is} \in \left[(1-\epsilon)^2 p_{is}, 1-(1-\epsilon)^2(1-p_{is})\right].
\end{aligned} \tag{A.2}$$

Falls $p_{is} < 1$ ist, sei

$$K_\epsilon = \left[0, 1-\frac{1}{2}(1-\epsilon)^2(1-p_{is})\right],$$

und für $p_{is} = 1$ sei

$$K_\epsilon = \left[\frac{1}{2}(1-\epsilon)^2, 1\right].$$

In jedem der beiden Fälle ist K_ϵ kompakt und eine echte Teilmenge des Einheitsintervalls $[0,1]$. Wegen (A.1) und (A.2) gilt wie behauptet

$$P_{(1-\epsilon)\boldsymbol{F}+\epsilon\boldsymbol{G}}(\widehat{p}_{is,n} \in K_\epsilon) \to 1.$$

$\square$

A.12 Beweis von Satz 5.1

Die Technik des Beweises von Satz 5.1 ist der des Beweises von Theorem 3.2 aus Brunner et al. (1999) nachempfunden. Wir bereiten den Beweis durch ein Lemma vor.

Lemma A.3. *Mit der Bezeichnung $F_{iks} = \lambda_{iks} F_{is}$ gilt:*

(a) *Falls eins der vier Indexpaare $(i,k),(i,k'),(u,w)$ und (u',w') von den drei anderen verschieden ist, so ist*

$$\mathrm{E}\left[\int \left(\widehat{F}_{uwv} - F_{uwv}\right) d\left(\widehat{F}_{iks} - F_{iks}\right) \int \left(\widehat{F}_{u'w'v'} - F_{u'w'v'}\right) d\left(\widehat{F}_{ik's} - F_{ik's}\right)\right] = 0.$$

(b) *Falls nicht, so ist*

$$\left|\mathrm{E}\left[\int \left(\widehat{F}_{uwv} - F_{uwv}\right) d\left(\widehat{F}_{iks} - F_{iks}\right) \int \left(\widehat{F}_{u'w'v'} - F_{u'w'v'}\right) d\left(\widehat{F}_{ik's} - F_{ik's}\right)\right]\right| \leq 4.$$

Beweis. (a) Einer der vier Zufallsvektoren $\boldsymbol{X}_{iks}, \boldsymbol{X}_{ik's}, \boldsymbol{X}_{uwv}$ und $\boldsymbol{X}_{u'w'v'}$ ist unabhängig von den drei anderen, je nachdem welches Indexpaar von den drei anderen verschieden ist. Falls (i,k) dieses Paar ist, folgt aus dem Satz von Fubini

$$\begin{aligned}
&\mathrm{E}\left[\int \left(\widehat{F}_{uwv} - F_{uwv}\right) d\left(\widehat{F}_{iks} - F_{iks}\right) \int \left(\widehat{F}_{u'w'v'} - F_{u'w'v'}\right) d\left(\widehat{F}_{ik's} - F_{ik's}\right)\right] \\
&= \mathrm{E}\left[\mathrm{E}\left(\int \left(\widehat{F}_{uwv} - F_{uwv}\right) d\left(\widehat{F}_{iks} - F_{iks}\right)\right.\right. \\
&\qquad \left.\left.\int \left(\widehat{F}_{u'w'v'} - F_{u'w'v'}\right) d\left(\widehat{F}_{ik's} - F_{ik's}\right) \middle| \boldsymbol{X}_{ik's}, \boldsymbol{X}_{uwv}, \boldsymbol{X}_{u'w'v'}\right)\right] \\
&= \mathrm{E}\left[\int \left(\widehat{F}_{uwv} - F_{uwv}\right) d(F_{iks} - F_{iks}) \int \left(\widehat{F}_{u'w'v'} - F_{u'w'v'}\right) d\left(\widehat{F}_{ik's} - F_{ik's}\right)\right] \\
&= 0.
\end{aligned}$$

In den drei anderen Fällen verläuft der Beweis analog.

(b) Zunächst gelten

$$\begin{aligned}
&\left|\int \left(\widehat{F}_{uwv} - F_{uwv}\right) d\left(\widehat{F}_{iks} - F_{iks}\right)\right| \\
&\leq \left|\int \left(\widehat{F}_{uwv} - F_{uwv}\right) d\widehat{F}_{iks}\right| + \left|\int \left(\widehat{F}_{uwv} - F_{uwv}\right) dF_{iks}\right| \leq 2
\end{aligned}$$

und aus dem gleichen Grund

$$\left|\int \left(\widehat{F}_{u'w'v'} - F_{u'w'v'}\right) d\left(\widehat{F}_{ik's} - F_{ik's}\right)\right| \leq 2.$$

Daraus folgt

$$\left|\left[\int \left(\widehat{F}_{uwv} - F_{uwv}\right) d\left(\widehat{F}_{iks} - F_{iks}\right) \int \left(\widehat{F}_{u'w'v'} - F_{u'w'v'}\right) d\left(\widehat{F}_{ik's} - F_{ik's}\right)\right]\right| \leq 4$$

und somit die Behauptung. ☐

Beweis von Satz 5.1. Wir zeigen in diesem Beweis sogar für alle Paare (i,s)

$$\mathrm{E}\left[\left(\sqrt{n}\int \widehat{H}_{is,n}\, d\left(\widehat{F}_{is,n}-F_{is}\right)-\sqrt{n}\int H_{is}\, d\left(\widehat{F}_{is,n}-F_{is}\right)\right)^2\right]\to 0, \tag{A.3}$$

was sicherlich die Behauptung beinhaltet.

Wegen

$$\begin{aligned}
&\sqrt{n}\int \widehat{H}_{is,n}\, d\left(\widehat{F}_{is,n}-F_{is}\right)-\sqrt{n}\int H_{is}\, d\left(\widehat{F}_{is,n}-F_{is}\right)\\
&= \sqrt{n}\int \left(\widehat{H}_{is,n}-H_{is}\right)\, d\left(\widehat{F}_{is,n}-F_{is}\right)\\
&= \sqrt{n}\int \frac{1}{at-1}\sum_{(u,v)\neq(i,s)}\left(\widehat{F}_{uv,n}-F_{uv}\right)\, d\left(\widehat{F}_{is,n}-F_{is}\right)\\
&= \frac{\sqrt{n}}{at-1}\sum_{(u,v)\neq(i,s)}\int \left(\widehat{F}_{uv,n}-F_{uv}\right)\, d\left(\widehat{F}_{is,n}-F_{is}\right)\\
&= \frac{\sqrt{n}}{at-1}\sum_{(u,v)\neq(i,s)}\int \left[\frac{1}{\lambda_{u\cdot v}}\sum_{w=1}^{n_u}\left(\widehat{F}_{uv,n}-F_{uv}\right)\right]\, d\left[\frac{1}{\lambda_{i\cdot s}}\sum_{k=1}^{n_i}\left(\widehat{F}_{is,n}-F_{is}\right)\right]\\
&= \frac{\sqrt{n}}{at-1}\sum_{(u,v)\neq(i,s)}\frac{1}{\lambda_{u\cdot v}\lambda_{i\cdot s}}\sum_{w=1}^{n_u}\sum_{k=1}^{n_i}\int \left(\widehat{F}_{uwv}-F_{uwv}\right)\, d\left(\widehat{F}_{iks}-F_{iks}\right)
\end{aligned}$$

erhalten wir durch Ausmultiplizieren

$$\begin{aligned}
&\mathrm{E}\left[\left(\sqrt{n}\int \widehat{H}_{is,n}\, d\left(\widehat{F}_{is,n}-F_{is}\right)-\sqrt{n}\int H_{is}\, d\left(\widehat{F}_{is,n}-F_{is}\right)\right)^2\right]\\
&= \mathrm{E}\Bigg[\frac{n}{(at-1)^2}\sum_{(u,v)\neq(i,s)}\sum_{(u',v')\neq(i,s)}\frac{1}{\lambda_{u\cdot v}\lambda_{u'\cdot v'}\lambda_{i\cdot s}^2}\sum_{w=1}^{n_u}\sum_{w'=1}^{n_{u'}}\sum_{k=1}^{n_i}\sum_{k'=1}^{n_i}\\
&\qquad \int \left(\widehat{F}_{uwv}-F_{uwv}\right)\, d\left(\widehat{F}_{iks}-F_{iks}\right)\int \left(\widehat{F}_{u'w'v'}-F_{u'w'v'}\right)\, d\left(\widehat{F}_{ik's}-F_{ik's}\right)\Bigg]\\
&= \frac{n}{(at-1)^2}\sum_{(u,v)\neq(i,s)}\sum_{(u',v')\neq(i,s)}\frac{1}{\lambda_{u\cdot v}\lambda_{u'\cdot v'}\lambda_{i\cdot s}^2}\sum_{w=1}^{n_u}\sum_{w'=1}^{n_{u'}}\sum_{k=1}^{n_i}\sum_{k'=1}^{n_i}\\
&\qquad \underbrace{\mathrm{E}\left[\int \left(\widehat{F}_{uwv}-F_{uwv}\right)\, d\left(\widehat{F}_{iks}-F_{iks}\right)\int \left(\widehat{F}_{u'w'v'}-F_{u'w'v'}\right)\, d\left(\widehat{F}_{ik's}-F_{ik's}\right)\right]}_{\phi_{uwviksu'w'v'k'}}\\
&= \frac{n}{(at-1)^2}\sum_{(u,v)\neq(i,s)}\sum_{(u',v')\neq(i,s)}\frac{1}{\lambda_{u\cdot v}\lambda_{u'\cdot v'}\lambda_{i\cdot s}^2}\sum_{w=1}^{n_u}\sum_{w'=1}^{n_{u'}}\sum_{k=1}^{n_i}\sum_{k'=1}^{n_i}\phi_{uwviksu'w'v'k'}
\end{aligned}$$

$$= \underbrace{\frac{n}{(at-1)^2} \sum_{(u,v)\neq(i,s)} \sum_{(u',v')\neq(i,s)} \frac{1}{\lambda_{u\cdot v}\lambda_{u'\cdot v'}\lambda_{i\cdot s}^2} \sum_{w=1}^{n_u} \sum_{w'=1}^{n_{u'}} \sum_{k=1}^{n_i} \phi_{uwviksu'w'v'k}}_{S_{is}}$$

$$+ \underbrace{\frac{n}{(at-1)^2} \sum_{(u,v)\neq(i,s)} \sum_{(u',v')\neq(i,s)} \frac{1}{\lambda_{u\cdot v}\lambda_{u'\cdot v'}\lambda_{i\cdot s}^2} \sum_{w=1}^{n_u} \sum_{w'=1}^{n_{u'}} \sum_{k=1}^{n_i} \sum_{k'\neq k} \phi_{uwviksu'w'v'k'}}_{T_{is}} .$$

Nun schätzen wir die Summanden S_{is} und T_{is} getrennt ab. Dabei benutzen wir Lemma A.3. Die Aussagen dieses Lemmas lauten in der hier eingeführten Schreibweise:

$$|\phi_{uwviksu'w'v'iksk'}| \leq \begin{cases} 0: & \text{eins der Indexpaare } (u,w),(u',w'),(i,k),(i,k') \text{ ist} \\ & \text{von den drei anderen verschieden,} \\ 4: & \text{sonst.} \end{cases}$$

Betrachten wir zunächst S_{is}. Da für $k = k'$ auch $(i,k) = (i,k')$ gilt, folgt aus $(u,w) \neq (u',w')$, dass wenigstens eins der beiden Paare (u,w) und (u',w') von (i,k) und damit von allen drei anderen Paaren verschieden ist, und deshalb also $\phi_{uwviksu'w'v'k} = 0$. Wir erhalten

$$\begin{aligned} |S_{is}| &= \left| \frac{n}{(at-1)^2} \sum_{(u,v)\neq(i,s)} \sum_{v'\neq s} \frac{1}{\lambda_{u\cdot v}\lambda_{u\cdot v'}\lambda_{i\cdot s}^2} \sum_{w=1}^{n_u} \sum_{k=1}^{n_i} \phi_{uwviksuwv'k} \right| \\ &\leq \frac{n}{(at-1)^2} \sum_{(u,v)\neq(i,s)} \sum_{v'\neq s} \frac{4}{\lambda_{u\cdot v}\lambda_{i\cdot s}}. \end{aligned}$$

Aus den Bedingungen (A1) und (A2) auf Seite 13 ergibt sich

$$\begin{aligned} \frac{n}{(at-1)^2} \sum_{(u,v)\neq(i,s)} \sum_{v'\neq s} \frac{4}{\lambda_{u\cdot v}\lambda_{i\cdot s}} &= \sum_{(u,v)\neq(i,s)} \sum_{v'\neq s} \frac{4n}{(at-1)^2\lambda_{u\cdot v}\lambda_{i\cdot s}} \\ &\leq \sum_{(u,v)\neq(i,s)} \sum_{v'\neq s} \frac{4n_0}{(at-1)^2\lambda_{i\cdot s}} \\ &\leq \frac{4tn_0}{(at-1)\lambda_{i\cdot s}} \to 0. \end{aligned}$$

Jetzt schätzen wir $|T_{is}|$ ab. Falls $u \neq u'$ ist, gilt $u \neq i$ oder $u' \neq i$, und deshalb unterscheidet sich dann (u,w) oder (u',w') von allen drei anderen Indexpaaren. Wir erhalten

$$|T_{is}| = \frac{n}{(at-1)^2} \sum_{(u,v)\neq(i,s)} \sum_{v'\neq s} \frac{1}{\lambda_{u\cdot v}\lambda_{u\cdot v'}\lambda_{i\cdot s}^2} \sum_{w=1}^{n_u} \sum_{w'=1}^{n_u} \sum_{k=1}^{n_i} \sum_{k'\neq k} \phi_{uwviksuw'v'k'} .$$

Für $u \neq i$ ist (i,k) von allen anderen Indexpaaren verschieden. Es bleibt also nur noch

$$|T_{is}| = \frac{n}{(at-1)^2} \sum_{v\neq s} \sum_{v'\neq s} \frac{1}{\lambda_{i\cdot v}\lambda_{i\cdot v'}\lambda_{i\cdot s}^2} \sum_{w=1}^{n_i} \sum_{w'=1}^{n_i} \sum_{k=1}^{n_i} \sum_{k'\neq k} \phi_{iwviksiw'v'k'}$$

übrig. Wegen $u = u' = i$ ist nur noch zu prüfen, wann einer der Indizes w, w', k und k' von den drei anderen verschieden ist. Dies gilt wegen $k \neq k'$ sicher für $w = w'$. Falls aber $w \neq w'$ ist, sind sogar alle vier Indizes w, w', k und k' paarweise verschieden, wenn nicht entweder $w = k$ und $w' = k'$ oder $w = k'$ und $w' = k$ gilt. Es ergibt sich

$$\begin{aligned}
|T_{is}| &= \frac{n}{(at-1)^2} \sum_{v \neq s} \sum_{v' \neq s} \frac{1}{\lambda_{i \cdot v} \lambda_{i \cdot v'} \lambda_{i \cdot s}^2} \sum_{k=1}^{n_i} \sum_{k' \neq k} \left(\phi_{ikviksik'v'k'} + \phi_{ik'viksikv'k'} \right) \\
&\leq \frac{n}{(at-1)^2} \sum_{v \neq s} \sum_{v' \neq s} \frac{8}{\lambda_{i \cdot s}^2}.
\end{aligned}$$

Nun können wir wieder auf die Annahmen (A1) und (A2) zurückgreifen und erhalten

$$\begin{aligned}
\frac{n}{(at-1)^2} \sum_{v \neq s} \sum_{v' \neq s} \frac{8}{\lambda_{i \cdot s}^2} &= \sum_{v \neq s} \sum_{v' \neq s} \frac{8n}{(at-1)^2 \lambda_{i \cdot s}^2} \\
&\leq \sum_{v \neq s} \sum_{v' \neq s} \frac{8n_0}{(at-1)^2 \lambda_{i \cdot s}} \leq \frac{8t^2 n_0}{(at-1)^2 \lambda_{i \cdot s}} \to 0.
\end{aligned}$$

Damit ist (A.3) schließlich bewiesen. □

A.13 Beweis von Korollar 5.2

Beweis von Korollar 5.2. Aufbauend auf den asymptotischen Äquivalenzsatz nehmen wir – unter anderem mittels komponentenweiser partieller Integration – einige Umformungen an $\sqrt{n} \int \widehat{\boldsymbol{H}}_n \, d\left(\widehat{\boldsymbol{F}}_n - \boldsymbol{F}\right) - \sqrt{n} \int \boldsymbol{H} \, d\left(\widehat{\boldsymbol{F}}_n - \boldsymbol{F}\right)$ vor.

$$\begin{aligned}
&\sqrt{n} \int \widehat{\boldsymbol{H}}_n \, d\left(\widehat{\boldsymbol{F}}_n - \boldsymbol{F}\right) - \sqrt{n} \int \boldsymbol{H} \, d\left(\widehat{\boldsymbol{F}}_n - \boldsymbol{F}\right) \\
&= \sqrt{n} \left(\int \widehat{\boldsymbol{H}}_n \, d\widehat{\boldsymbol{F}}_n - \int \boldsymbol{H} \, d\boldsymbol{F} \right) - \sqrt{n} \left(\int \boldsymbol{H} \, d\widehat{\boldsymbol{F}}_n + \int \widehat{\boldsymbol{H}}_n \, d\boldsymbol{F} - 2 \int \boldsymbol{H} \, d\boldsymbol{F} \right) \\
&= \sqrt{n}(\widehat{\boldsymbol{p}}_n - \boldsymbol{p}) - \sqrt{n} \left(\int \boldsymbol{H} \, d\widehat{\boldsymbol{F}}_n + \boldsymbol{1}_{at} - \int \boldsymbol{F} \, d\widehat{\boldsymbol{H}}_n - 2\boldsymbol{p} \right)
\end{aligned}$$

Es genügt also, die Identität

$$\begin{aligned}
&\sqrt{n} \left(\int \boldsymbol{H} \, d\widehat{\boldsymbol{F}}_n + \boldsymbol{1}_{at} - \int \boldsymbol{F} \, d\widehat{\boldsymbol{H}}_n - 2\boldsymbol{p} \right) \\
&= \sqrt{n} \left(\sum_{i=1}^{a} \frac{1}{n_i} \sum_{k=1}^{n_i} [\boldsymbol{\Psi}_{ik,n} - \mathrm{E}(\boldsymbol{\Psi}_{ik,n})] \right)
\end{aligned}$$

zu zeigen. Wegen

$$\begin{aligned}
\mathrm{E}\left(\int \boldsymbol{H} \, d\widehat{\boldsymbol{F}}_n + \boldsymbol{1}_{at} - \int \boldsymbol{F} \, d\widehat{\boldsymbol{H}}_n - 2\boldsymbol{p} \right) &= \int \boldsymbol{H} \, d\boldsymbol{F} + \boldsymbol{1}_{at} - \int \boldsymbol{F} \, d\boldsymbol{H} - 2\boldsymbol{p} \\
&= \boldsymbol{p} + \boldsymbol{1}_{at} - \left(\boldsymbol{1}_{at} - \int \boldsymbol{H} \, d\boldsymbol{F} \right) - 2\boldsymbol{p} = \boldsymbol{0}
\end{aligned}$$

reicht es,

$$\int \boldsymbol{H}\, d\widehat{\boldsymbol{F}}_n - \int \boldsymbol{F}\, d\widehat{\boldsymbol{H}}_n = \sum_{i=1}^{a} \frac{1}{n_i} \sum_{k=1}^{n_i} \boldsymbol{\Psi}_{ik,n}$$

zu überprüfen. Wir führen dies komponentenweise durch.

$$\begin{aligned}
&\int H_{uv}\, d\widehat{F}_{uv,n} - \int F_{uv}\, d\widehat{H}_{uv,n} \\
= \;& \frac{1}{\lambda_{u\cdot v}} \sum_{k=1}^{n_u} \int H_{uv}\, d\widehat{F}_{ukv} - \frac{1}{at-1} \sum_{(i,s)\neq(u,v)} \int F_{uv}\, d\widehat{F}_{is,n} \\
= \;& \frac{1}{\lambda_{u\cdot v}} \sum_{k=1}^{n_u} \int H_{uv}\, d\widehat{F}_{ukv} - \frac{1}{at-1} \sum_{(i,s)\neq(u,v)} \frac{1}{\lambda_{i\cdot s}} \sum_{k=1}^{n_i} \int F_{uv}\, d\widehat{F}_{iks} \\
= \;& \frac{1}{\lambda_{u\cdot v}} \sum_{k=1}^{n_u} \int H_{uv}\, d\widehat{F}_{ukv} - \frac{1}{at-1} \sum_{s\neq v} \frac{1}{\lambda_{u\cdot s}} \sum_{k=1}^{n_u} \int F_{uv}\, d\widehat{F}_{uks} \\
& - \frac{1}{at-1} \sum_{i\neq u} \sum_{s=1}^{t} \frac{1}{\lambda_{i\cdot s}} \sum_{k=1}^{n_i} \int F_{uv}\, d\widehat{F}_{iks} \\
= \;& \sum_{k=1}^{n_u} \left(\frac{1}{\lambda_{u\cdot v}} \int H_{uv}\, d\widehat{F}_{ukv} - \frac{1}{at-1} \sum_{s\neq v} \frac{1}{\lambda_{u\cdot s}} \int F_{uv}\, d\widehat{F}_{uks} \right) \\
& + \sum_{i\neq u} \sum_{k=1}^{n_i} \left(\frac{-1}{at-1} \sum_{s=1}^{t} \frac{1}{\lambda_{i\cdot s}} \int F_{uv}\, d\widehat{F}_{iks} \right) \\
= \;& \frac{1}{n_u} \sum_{k=1}^{n_u} \left(\frac{n_u}{\lambda_{u\cdot v}} \int H_{uv}\, d\widehat{F}_{ukv} - \frac{1}{at-1} \sum_{s\neq v} \frac{n_u}{\lambda_{u\cdot s}} \int F_{uv}\, d\widehat{F}_{uks} \right) \\
& + \sum_{i\neq u} \frac{1}{n_i} \sum_{k=1}^{n_i} \left(\frac{-1}{at-1} \sum_{s=1}^{t} \frac{n_i}{\lambda_{i\cdot s}} \int F_{uv}\, d\widehat{F}_{iks} \right) \\
= \;& \frac{1}{n_u} \sum_{k=1}^{n_u} \Psi_{ukuv,n} + \sum_{i\neq u} \frac{1}{n_i} \sum_{k=1}^{n_i} \Psi_{ikuv,n} \\
= \;& \sum_{i=1}^{a} \frac{1}{n_i} \sum_{k=1}^{n_i} \Psi_{ikuv,n}
\end{aligned}$$

Die gleichmäßige Beschränktheit der Zufallsvektoren $\boldsymbol{\Psi}_{ik,n}$ ergibt sich wegen $n_i \leq n$ sofort aus Annahme (A2) auf Seite 13. □

A.14 Beweis von Satz 5.6

Wir benötigen zunächst folgendes

Lemma A.4. *Für jede Konstante $M \in \mathbb{R}$ und jede natürliche Zahl d gibt es eine Nullfolge positiver Zahlen a_n mit folgender Eigenschaft: Für alle Folgen unabhängig* *identisch verteilter und zentrierter d-dimensionaler Zufallsvektoren $\boldsymbol{Y}_k$ mit Kovarianzmatrix $\boldsymbol{\Sigma}$, die durch M fast sicher beschränkt sind, gilt*

$$\rho\left[\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\boldsymbol{Y}_k, \mathrm{N}(\boldsymbol{0},\boldsymbol{\Sigma})\right] \leq a_n.$$

Beweis. Nach Proposition A.5.2 auf Seite 457 in van der Vaart und Wellner (1996) gilt für eine Konstante $C \in \mathbb{R}$ und jedes $\epsilon > 0$

$$\rho\left[\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\boldsymbol{Y}_k, \mathrm{N}(\boldsymbol{0},\boldsymbol{\Sigma})\right] \leq f(\epsilon, n),$$

wobei

$$\begin{aligned} f(\epsilon, n) = \max\left\{2\epsilon^{-2}M^2\,\mathrm{P}\left(\|\boldsymbol{Y}_1\| \geq \epsilon n^{1/2}\right), \epsilon\right\} + 4^{1/3}M^{2/3}\left[\mathrm{P}\left(\|\boldsymbol{Y}_1\| \geq \epsilon n^{1/2}\right)\right]^{1/3} \\ + Cd^{1/4}\epsilon^{1/4}M^{1/2} + Cd^{1/4}\epsilon^{1/4}M^{1/2}\left|\log\frac{\epsilon M^2}{d}\right|^{1/2}. \end{aligned}$$

Mit der Wahl $a_n = f\left(n^{-1/4}, n\right)$ erhalten wir wegen der Beschränktheit der $\boldsymbol{Y}_k$ und wegen des langsamen Wachstums der Logarithmusfunktion (siehe Forster (1983), Seite 77)

$$\rho\left[\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\boldsymbol{Y}_k, \mathrm{N}(\boldsymbol{0},\boldsymbol{\Sigma})\right] \leq a_n \to 0.$$

□

Nach dieser Vorbereitung kommen wir zum

Beweis von Satz 5.6. Wegen Korollar 5.2 auf Seite 60 und Lemma 5.4 auf Seite 61 genügt es,

$$\rho\left[\sqrt{n}\left(\sum_{i=1}^{a}\frac{1}{n_i}\sum_{k=1}^{n_i}\{\boldsymbol{\Psi}_{ik,n} - \mathrm{E}(\boldsymbol{\Psi}_{ik,n})\}\right), \mathrm{N}(\boldsymbol{0},\boldsymbol{V}_n)\right] \to 0. \tag{A.4}$$

zu zeigen.

Wir vereinbaren zunächst einige Bezeichnungen. Die Menge $\boldsymbol{M}_0$ sei das t-fache kartesische Produkt der Menge $\{0, 1, \ldots, m_0\}$. Für jedes $\boldsymbol{m} \in \boldsymbol{M}_0$ sei $n_{i\boldsymbol{m}}$ die Anzahl der Versuchseinheiten in der i-ten Gruppe mit Dimensionsvektor $\boldsymbol{m}$, das heißt,

$$n_{i\boldsymbol{m}} = \#\{(i,k) : \boldsymbol{m}_{ik} = \boldsymbol{m}\}.$$

Darüber hinaus gelte

$$\begin{aligned} \boldsymbol{S}_{i\boldsymbol{m},n} &= \frac{1}{\sqrt{n_{i\boldsymbol{m}}}}\sum_{\boldsymbol{m}_{ik}=\boldsymbol{m}}\frac{\sqrt{n n_{i\boldsymbol{m}}}}{n_i}\left[\boldsymbol{\Psi}_{ik,n} - \mathrm{E}(\boldsymbol{\Psi}_{ik,n})\right], \\ \boldsymbol{V}_{i\boldsymbol{m},n} &= \mathrm{Cov}(\boldsymbol{S}_{i\boldsymbol{m},n}). \end{aligned}$$

Da die Verteilung der Summanden $\boldsymbol{\Psi}_{ik,n} - \mathrm{E}(\boldsymbol{\Psi}_{ik,n})$ nur von i und $\boldsymbol{m}_{ik}$ abhängt, ist $\boldsymbol{S}_{\boldsymbol{im},n}$ eine Summe unabhängig identisch verteilter Zufallsvektoren.

Wir unterscheiden nun zwei Fälle bezüglich der monoton wachsenden Folgen $(n_{\boldsymbol{im}})$: Diese sind entweder beschränkt oder konvergieren gegen ∞.

Im ersten Fall gilt

$$\begin{aligned}\boldsymbol{S}_{\boldsymbol{im},n} &= \frac{1}{\sqrt{n_{\boldsymbol{im}}}} \sum_{\boldsymbol{m}_{ik}=\boldsymbol{m}} \frac{\sqrt{n n_{\boldsymbol{im}}}}{n_i} \left[\boldsymbol{\Psi}_{ik,n} - \mathrm{E}(\boldsymbol{\Psi}_{ik,n})\right] \\ &= \frac{\sqrt{n}}{n_i} \sum_{\boldsymbol{m}_{ik}=\boldsymbol{m}} \left[\boldsymbol{\Psi}_{ik,n} - \mathrm{E}(\boldsymbol{\Psi}_{ik,n})\right] \to \mathbf{0},\end{aligned}$$

da nach den Annahmen (A1) und (A2) $\sqrt{n}/n_i \to 0$ konvergiert und die Anzahl der gleichmäßig beschränkten Summanden in der Summe in der letzten Zeile beschränkt bleibt. Folglich gelten auch die Aussagen $\boldsymbol{V}_{\boldsymbol{im},n} \to \mathbf{0}$ und

$$\rho\left[\boldsymbol{S}_{\boldsymbol{im},n}, \mathrm{N}(\mathbf{0}, \boldsymbol{V}_{\boldsymbol{im},n})\right] \to 0. \tag{A.5}$$

Im zweiten Fall können wir Lemma A.4 anwenden und erhalten eine Nullfolge (a_n) mit

$$\rho\left[\boldsymbol{S}_{\boldsymbol{im},n}, \mathrm{N}(\mathbf{0}, \boldsymbol{V}_{\boldsymbol{im},n})\right] \le a_{n_{\boldsymbol{im}}}.$$

Wegen $n_{\boldsymbol{im}} \to \infty$ gilt $a_{n_{\boldsymbol{im}}} \to 0$ und also auch in diesem Fall (A.5).

Wegen Lemma 5.4 und nach (A.5) gibt es für alle $(i, \boldsymbol{m})$ Wahrscheinlichkeitsräume $(\Omega_{\boldsymbol{im}}, \mathcal{A}_{\boldsymbol{im}}, P_{\boldsymbol{im}})$ und darauf definierte Zufallsvektoren $\boldsymbol{Y}_{\boldsymbol{im},n}$ und $\boldsymbol{Z}_{\boldsymbol{im},n}$ mit

$$\boldsymbol{Y}_{\boldsymbol{im},n} \sim \boldsymbol{S}_{\boldsymbol{im},n}, \quad \boldsymbol{Z}_{\boldsymbol{im},n} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{V}_{\boldsymbol{im},n}) \quad \text{und} \quad \boldsymbol{Y}_{\boldsymbol{im},n} - \boldsymbol{Z}_{\boldsymbol{im},n} \xrightarrow{\text{f.s.}} \mathbf{0}.$$

Wir bilden das Produkt dieser Wahrscheinlichkeitsräume

$$(\Omega, \mathcal{A}, P) = \bigotimes_{i=1}^{a} \bigotimes_{\boldsymbol{m} \in \boldsymbol{M}_0} (\Omega_{\boldsymbol{im}}, \mathcal{A}_{\boldsymbol{im}}, P_{\boldsymbol{im}})$$

mit den Elementen

$$\boldsymbol{\omega} = (\omega_{\boldsymbol{im}})_{i \in \{1,\dots,a\}, \boldsymbol{m} \in \boldsymbol{M}_0}.$$

Die darauf definierten Zufallsvektoren $\widetilde{\boldsymbol{Y}}_{\boldsymbol{im},n}$ und $\widetilde{\boldsymbol{Z}}_{\boldsymbol{im},n}$ seien durch

$$\widetilde{\boldsymbol{Y}}_{\boldsymbol{im},n} : \boldsymbol{\omega} \mapsto \boldsymbol{Y}_{\boldsymbol{im},n}(\omega_{\boldsymbol{im}}) \quad \text{und} \quad \widetilde{\boldsymbol{Z}}_{\boldsymbol{im},n} : \boldsymbol{\omega} \mapsto \boldsymbol{Z}_{\boldsymbol{im},n}(\omega_{\boldsymbol{im}})$$

erklärt. Daraus ergibt sich

$$\widetilde{\boldsymbol{Y}}_{\boldsymbol{im},n} \sim \boldsymbol{S}_{\boldsymbol{im},n}, \quad \widetilde{\boldsymbol{Z}}_{\boldsymbol{im},n} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{V}_{\boldsymbol{im},n}) \quad \text{sowie} \quad \widetilde{\boldsymbol{Y}}_{\boldsymbol{im},n} - \widetilde{\boldsymbol{Z}}_{\boldsymbol{im},n} \xrightarrow{\text{f.s.}} \mathbf{0},$$

und damit auch

$$\begin{aligned}\sum_{i=1}^{a} \sum_{\boldsymbol{m} \in \boldsymbol{M}_0} \widetilde{\boldsymbol{Y}}_{\boldsymbol{im},n} &\sim \sum_{i=1}^{a} \sum_{\boldsymbol{m} \in \boldsymbol{M}_0} \boldsymbol{S}_{\boldsymbol{im},n} \\ &= \sqrt{n} \left(\sum_{i=1}^{a} \frac{1}{n_i} \sum_{k=1}^{n_i} \{\boldsymbol{\Psi}_{ik,n} - \mathrm{E}(\boldsymbol{\Psi}_{ik,n})\} \right),\end{aligned}$$

$$\sum_{i=1}^{a} \sum_{\boldsymbol{m} \in \boldsymbol{M}_0} \widetilde{\boldsymbol{Z}}_{\boldsymbol{im},n} \sim \mathrm{N}\left(\boldsymbol{0}, \sum_{i=1}^{a} \sum_{\boldsymbol{m} \in \boldsymbol{M}_0} \boldsymbol{V}_{\boldsymbol{im},n}\right) = \mathrm{N}(\boldsymbol{0}, \boldsymbol{V}_n)$$

sowie

$$\sum_{i=1}^{a} \sum_{\boldsymbol{m} \in \boldsymbol{M}_0} \widetilde{\boldsymbol{Y}}_{\boldsymbol{im},n} - \sum_{i=1}^{a} \sum_{\boldsymbol{m} \in \boldsymbol{M}_0} \widetilde{\boldsymbol{Z}}_{\boldsymbol{im},n} \xrightarrow{\text{f.s.}} \boldsymbol{0}.$$

Die erneute Anwendung von Lemma 5.4 liefert schließlich wie verlangt (A.4). □

A.15 Beweis von Satz 5.8

Lemma A.5. *$X_{1n}, \ldots, X_{nn}$, $n \in \mathbb{N}$, seien unabhängige, gleichmäßig beschränkte Zufallsvariablen mit $\mathrm{E}(X_{kn}) = 0$. Dann konvergiert der Mittelwert*

$$\frac{1}{n} \sum_{k=1}^{n} X_{kn}$$

fast sicher gegen 0.

Beweis. Der Beweis kann völlig analog zu dem Beweis von Theorem 7.2 in Williams (1991) geführt werden, wenn wir die dort verwendeten Zufallsvariablen $X_1, \ldots, X_n$ jeweils durch $X_{1n}, \ldots, X_{nn}$ ersetzen. □

Beweis von Satz 5.8. Wir führen den Beweis komponentenweise und gehen dabei in den gleichen Schritten wie bei der Herleitung der Schätzer vor. Wir zeigen also nacheinander

$$\sigma_{uvu'v',n} - \widetilde{\sigma}_{uvu'v',n} \xrightarrow{\text{f.s.}} 0, \tag{A.6}$$

$$\widetilde{\sigma}_{uvu'v',n} - \check{\sigma}_{uvu'v',n} \xrightarrow{\text{f.s.}} 0 \tag{A.7}$$

und

$$\check{\sigma}_{uvu'v',n} - \widehat{\sigma}_{uvu'v',n} \xrightarrow{\text{f.s.}} 0. \tag{A.8}$$

Aus Lemma A.5 ergibt sich wegen der gleichmäßigen Beschränktheit der Zufallsvariablen $(\Psi_{ikuv,n} - \mathrm{E}[\Psi_{ikuv,n}])(\Psi_{iku'v',n} - \mathrm{E}[\Psi_{iku'v',n}])$ für alle $i \in \{1, \ldots, a\}$

$$\frac{1}{n_i} \sum_{k=1}^{n_i} \big[(\Psi_{ikuv,n} - \mathrm{E}[\Psi_{ikuv,n}])(\Psi_{iku'v',n} - \mathrm{E}[\Psi_{iku'v',n}]) \\ - \mathrm{E}[(\Psi_{ikuv,n} - \mathrm{E}[\Psi_{ikuv,n}])(\Psi_{iku'v',n} - \mathrm{E}[\Psi_{iku'v',n}])]\big] \xrightarrow{\text{f.s.}} 0.$$

Damit folgt (A.6) aus Annahme (A2) auf Seite 13.

Für den Beweis von (A.7) zerlegen wir $\widetilde{\sigma}_{uvu'v',n} - \check{\sigma}_{uvu'v',n}$ in zwei Summanden und behandeln diese getrennt.

$$\begin{aligned}
&\widetilde{\sigma}_{uvu'v',n} - \check{\sigma}_{uvu'v',n} \\
&= n\sum_{i=1}^{a}\frac{1}{n_i^2}\sum_{k=1}^{n_i}(\Psi_{ikuv,n} - \mathrm{E}[\Psi_{ikuv,n}])(\Psi_{iku'v',n} - \mathrm{E}[\Psi_{iku'v',n}]) \\
&\quad -n\sum_{i=1}^{a}\frac{1}{n_i(n_i-1)}\sum_{k=1}^{n_i}(\Psi_{ikuv,n} - \check{\mu}_{ikuv,n})(\Psi_{iku'v',n} - \check{\mu}_{iku'v',n}) \\
&= n\sum_{i=1}^{a}\frac{1}{n_i^2}\sum_{k=1}^{n_i}(\Psi_{ikuv,n} - \mathrm{E}[\Psi_{ikuv,n}])(\Psi_{iku'v',n} - \mathrm{E}[\Psi_{iku'v',n}]) \\
&\quad -n\sum_{i=1}^{a}\frac{1}{n_i^2}\sum_{k=1}^{n_i}(\Psi_{ikuv,n} - \check{\mu}_{ikuv,n})(\Psi_{iku'v',n} - \check{\mu}_{iku'v',n}) \\
&\quad -n\sum_{i=1}^{a}\frac{1}{n_i^2(n_i-1)}\sum_{k=1}^{n_i}(\Psi_{ikuv,n} - \check{\mu}_{ikuv,n})(\Psi_{iku'v',n} - \check{\mu}_{iku'v',n}) \\
&= n\sum_{i=1}^{a}\frac{1}{n_i^2}\sum_{k=1}^{n_i}[(\Psi_{ikuv,n} - \mathrm{E}[\Psi_{ikuv,n}])(\Psi_{iku'v',n} - \mathrm{E}[\Psi_{iku'v',n}]) \\
&\qquad\qquad -(\Psi_{ikuv,n} - \check{\mu}_{ikuv,n})(\Psi_{iku'v',n} - \check{\mu}_{iku'v',n})] && \text{(A.9)} \\
&\quad -n\sum_{i=1}^{a}\frac{1}{n_i^2(n_i-1)}\sum_{k=1}^{n_i}(\Psi_{ikuv,n} - \check{\mu}_{ikuv,n})(\Psi_{iku'v',n} - \check{\mu}_{iku'v',n}) && \text{(A.10)}
\end{aligned}$$

Der Subtrahend aus (A.10) konvergiert wegen der Annahmen (A1) und (A2) sowie wegen der gleichmäßigen Beschränktheit der Zufallsvariablen $(\Psi_{ikuv,n} - \check{\mu}_{ikuv,n})$ $(\Psi_{iku'v',n} - \check{\mu}_{iku'v',n})$ gegen 0.

Zur Behandlung des Summanden aus (A.9) zeigen wir zunächst

$$\max_{k=1,\ldots,n_i} |\mathrm{E}(\Psi_{ikuv,n}) - \check{\mu}_{ikuv,n}| \xrightarrow{\text{f.s.}} 0. \tag{A.11}$$

Im Fall $u = i$ erhalten wir durch partielle Integration, die Dreiecksungleichung und Anwendung von Satz 4.2 (siehe Seite 48)

$$\begin{aligned}
&\max_{k=1,\ldots,n_i} |\mathrm{E}(\Psi_{ikuv,n}) - \check{\mu}_{ikuv,n}| \\
&= \max_{k=1,\ldots,n_i}\left|\frac{n_u}{\lambda_{u\cdot v}}\int H_{uv}\, d(\lambda_{ukv}F_{uv}) - \frac{1}{at-1}\sum_{s\neq v}\frac{n_u}{\lambda_{u\cdot s}}\int F_{uv}\, d(\lambda_{uks}F_{us})\right. \\
&\qquad\left. -\frac{n_u}{\lambda_{u\cdot v}}\int H_{uv}\, d\left(\lambda_{ukv}\widehat{F}_{uv,n}\right) + \frac{1}{at-1}\sum_{s\neq v}\frac{n_u}{\lambda_{u\cdot s}}\int F_{uv}\, d\left(\lambda_{uks}\widehat{F}_{us,n}\right)\right| \\
&= \max_{k=1,\ldots,n_i}\left|\frac{n_u\lambda_{ukv}}{\lambda_{u\cdot v}}\int H_{uv}\, d\left(F_{uv} - \widehat{F}_{uv,n}\right)\right. \\
&\qquad\left. -\frac{1}{at-1}\sum_{s\neq v}\frac{n_u\lambda_{uks}}{\lambda_{u\cdot s}}\int F_{uv}\, d\left(F_{us} - \widehat{F}_{us,n}\right)\right|
\end{aligned}$$

$$
\begin{aligned}
&= \max_{k=1,\ldots,n_i} \left| \frac{n_u \lambda_{ukv}}{\lambda_{u\cdot v}} \int \left(\widehat{F}_{uv,n} - F_{uv} \right) dH_{uv} \right. \\
&\qquad \left. - \frac{1}{at-1} \sum_{s \neq v} \frac{n_u \lambda_{uks}}{\lambda_{u\cdot s}} \int \left(\widehat{F}_{us,n} - F_{us} \right) dF_{uv} \right| \\
&\leq \max_{k=1,\ldots,n_i} \frac{n_u \lambda_{ukv}}{\lambda_{u\cdot v}} \int \left| \widehat{F}_{uv,n} - F_{uv} \right| dH_{uv} \\
&\quad + \max_{k=1,\ldots,n_i} \frac{1}{at-1} \sum_{s \neq v} \frac{n_u \lambda_{uks}}{\lambda_{u\cdot s}} \int \left| \widehat{F}_{us,n} - F_{us} \right| dF_{uv} \\
&\leq \frac{n_u}{\lambda_{u\cdot v}} \left\| \widehat{F}_{uv,n} - F_{uv} \right\|_\infty + \frac{1}{at-1} \sum_{s \neq v} \frac{n_u}{\lambda_{u\cdot s}} \left\| \widehat{F}_{us,n} - F_{us} \right\|_\infty \xrightarrow{\text{f.s.}} 0.
\end{aligned}
$$

Im Fall $u \neq i$ kann man (A.11) analog beweisen.

Wir formen den Summanden aus (A.9) weiter um, damit wir (A.11) anwenden können:

$$
\begin{aligned}
&\left| n \sum_{i=1}^{a} \frac{1}{n_i^2} \sum_{k=1}^{n_i} \left[(\Psi_{ikuv,n} - \mathrm{E}[\Psi_{ikuv,n}])(\Psi_{iku'v',n} - \mathrm{E}[\Psi_{iku'v',n}]) \right. \right. \\
&\qquad \left. \left. - (\Psi_{ikuv,n} - \check{\mu}_{ikuv,n})(\Psi_{iku'v',n} - \check{\mu}_{iku'v',n}) \right] \right| \\
&= \left| n \sum_{i=1}^{a} \frac{1}{n_i^2} \sum_{k=1}^{n_i} \left[(\mathrm{E}[\Psi_{ikuv,n}] - \Psi_{ikuv,n})(\mathrm{E}[\Psi_{iku'v',n}] - \check{\mu}_{iku'v',n}) \right. \right. \\
&\qquad \left. \left. + (\check{\mu}_{iku'v',n} - \Psi_{iku'v',n})(\mathrm{E}[\Psi_{ikuv,n}] - \check{\mu}_{ikuv,n}) \right] \right| \\
&\leq n \sum_{i=1}^{a} \frac{1}{n_i^2} \sum_{k=1}^{n_i} \Big[\left| (\mathrm{E}[\Psi_{ikuv,n}] - \Psi_{ikuv,n})(\mathrm{E}[\Psi_{iku'v',n}] - \check{\mu}_{iku'v',n}) \right| \\
&\qquad + \left| (\check{\mu}_{iku'v',n} - \Psi_{iku'v',n})(\mathrm{E}[\Psi_{ikuv,n}] - \check{\mu}_{ikuv,n}) \right| \Big] \\
&\leq n \sum_{i=1}^{a} \frac{1}{n_i} \max_{k=1,\ldots,n_i} \Big[\left| (\mathrm{E}[\Psi_{ikuv,n}] - \Psi_{ikuv,n})(\mathrm{E}[\Psi_{iku'v',n}] - \check{\mu}_{iku'v',n}) \right| \\
&\qquad + \left| (\check{\mu}_{iku'v',n} - \Psi_{iku'v',n})(\mathrm{E}[\Psi_{ikuv,n}] - \check{\mu}_{ikuv,n}) \right| \Big] \\
&\leq n \sum_{i=1}^{a} \frac{1}{n_i} \Big(\max_{k=1,\ldots,n_i} \left| (\mathrm{E}[\Psi_{ikuv,n}] - \Psi_{ikuv,n})(\mathrm{E}[\Psi_{iku'v',n}] - \check{\mu}_{iku'v',n}) \right| \\
&\qquad + \max_{k=1,\ldots,n_i} \left| (\check{\mu}_{iku'v',n} - \Psi_{iku'v',n})(\mathrm{E}[\Psi_{ikuv,n}] - \check{\mu}_{ikuv,n}) \right| \Big).
\end{aligned}
$$

Aus der Annahme (A2), der gleichmäßigen Beschränktheit der Zufallsvariablen

$\mathrm{E}[\Psi_{ikuv,n}] - \Psi_{ikuv,n}$ und $\check{\mu}_{iku'v',n} - \Psi_{iku'v',n}$ sowie (A.11) ergibt sich nun

$$n \sum_{i=1}^{a} \frac{1}{n_i} \Big(\max_{k=1,\dots,n_i} \left| (\mathrm{E}[\Psi_{ikuv,n}] - \Psi_{ikuv,n})(\mathrm{E}[\Psi_{iku'v',n}] - \check{\mu}_{iku'v',n}) \right|$$
$$+ \max_{k=1,\dots,n_i} \left| (\check{\mu}_{iku'v',n} - \Psi_{iku'v',n})(\mathrm{E}[\Psi_{ikuv,n}] - \check{\mu}_{ikuv,n}) \right| \Big) \xrightarrow{\text{f.s.}} 0,$$

so dass (A.7) schließlich bewiesen ist.

Zur Vorbereitung des Beweises von (A.8) zeigen wir

$$\max_{k=1,\dots,n_i} \left| \Psi_{ikuv,n} - \widehat{\Psi}_{ikuv,n} \right| \xrightarrow{\text{f.s.}} 0 \tag{A.12}$$

und

$$\max_{k=1,\dots,n_i} \left| \check{\mu}_{ikuv,n} - \widehat{\mu}_{ikuv,n} \right| \xrightarrow{\text{f.s.}} 0. \tag{A.13}$$

Für $u = i$ erhalten wir durch die Dreiecksungleichung und Satz 4.2

$$\begin{aligned}
&\max_{k=1,\dots,n_u} \left| \Psi_{ikuv,n} - \widehat{\Psi}_{ikuv,n} \right| \\
&= \max_{k=1,\dots,n_u} \Bigg| \frac{n_u}{\lambda_{u\cdot v}} \int H_{uv}\, d\widehat{F}_{ukv} - \frac{1}{at-1} \sum_{s \neq v} \frac{n_u}{\lambda_{u\cdot s}} \int F_{uv}\, d\widehat{F}_{uks} \\
&\qquad - \frac{n_u}{\lambda_{u\cdot v}} \int \widehat{H}_{uv,n}\, d\widehat{F}_{ukv} + \frac{1}{at-1} \sum_{s \neq v} \frac{n_u}{\lambda_{u\cdot s}} \int \widehat{F}_{uv,n}\, d\widehat{F}_{uks} \Bigg| \\
&= \max_{k=1,\dots,n_u} \Bigg| \frac{n_u}{\lambda_{u\cdot v}} \int \left(H_{uv} - \widehat{H}_{uv,n} \right) d\widehat{F}_{ukv} \\
&\qquad - \frac{1}{at-1} \sum_{s \neq v} \frac{n_u}{\lambda_{u\cdot s}} \int \left(F_{uv} - \widehat{F}_{uv,n} \right) d\widehat{F}_{uks} \Bigg| \\
&\leq \max_{k=1,\dots,n_u} \frac{n_u}{\lambda_{u\cdot v}} \int \left| \widehat{H}_{uv,n} - H_{uv} \right| d\widehat{F}_{ukv} \\
&\quad + \max_{k=1,\dots,n_u} \frac{1}{at-1} \sum_{s \neq v} \frac{n_u}{\lambda_{u\cdot s}} \int \left| \widehat{F}_{uv,n} - F_{uv} \right| d\widehat{F}_{uks} \\
&\leq \frac{n_u}{\lambda_{u\cdot v}} \left\| \widehat{H}_{uv,n} - H_{uv} \right\|_\infty + \frac{1}{at-1} \sum_{s \neq v} \frac{n_u}{\lambda_{u\cdot s}} \left\| \widehat{F}_{uv,n} - F_{uv} \right\|_\infty \xrightarrow{\text{f.s.}} 0.
\end{aligned}$$

Für $u \neq i$ ergibt sich (A.12) analog, so wie sich auch (A.13) analog nachweisen lässt.

Nun formen wir zum Nachweis von (A.8) den Ausdruck $|\check{\sigma}_{uvu'v',n} - \widehat{\sigma}_{uvu'v',n}|$ so um, dass wir (A.12) und (A.13) anwenden können:

$$\begin{aligned}
&\left| \check{\sigma}_{uvu'v',n} - \widehat{\sigma}_{uvu'v',n} \right| \\
&= \Bigg| n \sum_{i=1}^{a} \frac{1}{n_i^2} \sum_{k=1}^{n_i} \Big[\left(\Psi_{ikuv,n} - \check{\mu}_{ikuv,n} \right) \left(\Psi_{iku'v',n} - \check{\mu}_{iku'v',n} \right) \\
&\qquad - \left(\widehat{\Psi}_{ikuv,n} - \widehat{\mu}_{ikuv,n} \right) \left(\widehat{\Psi}_{iku'v',n} - \widehat{\mu}_{iku'v',n} \right) \Big] \Bigg|
\end{aligned}$$

$$
\begin{aligned}
= \; & \Bigg| n \sum_{i=1}^{a} \frac{1}{n_i^2} \sum_{k=1}^{n_i} \Bigg[\Big(\Psi_{ikuv,n} - \check{\mu}_{ikuv,n}\Big)\Big(\Psi_{iku'v',n} - \widehat{\Psi}_{iku'v',n}\Big) \\
& \qquad - \Big(\Psi_{ikuv,n} - \check{\mu}_{ikuv,n}\Big)\Big(\check{\mu}_{iku'v',n} - \widehat{\mu}_{iku'v',n}\Big) \\
& \qquad + \Big(\widehat{\Psi}_{iku'v',n} - \widehat{\mu}_{iku'v',n}\Big)\Big(\Psi_{ikuv,n} - \widehat{\Psi}_{ikuv,n}\Big) \\
& \qquad - \Big(\widehat{\Psi}_{iku'v',n} - \widehat{\mu}_{iku'v',n}\Big)\Big(\check{\mu}_{ikuv,n} - \widehat{\mu}_{ikuv,n}\Big)\Bigg]\Bigg| \\
\leq \; & n \sum_{i=1}^{a} \frac{1}{n_i^2} \sum_{k=1}^{n_i} \Bigg[\Big|\Big(\Psi_{ikuv,n} - \check{\mu}_{ikuv,n}\Big)\Big(\Psi_{iku'v',n} - \widehat{\Psi}_{iku'v',n}\Big)\Big| \\
& \qquad + \Big|\Big(\Psi_{ikuv,n} - \check{\mu}_{ikuv,n}\Big)\Big(\check{\mu}_{iku'v',n} - \widehat{\mu}_{iku'v',n}\Big)\Big| \\
& \qquad + \Big|\Big(\widehat{\Psi}_{iku'v',n} - \widehat{\mu}_{iku'v',n}\Big)\Big(\Psi_{ikuv,n} - \widehat{\Psi}_{ikuv,n}\Big)\Big| \\
& \qquad + \Big|\Big(\widehat{\Psi}_{iku'v',n} - \widehat{\mu}_{iku'v',n}\Big)\Big(\check{\mu}_{ikuv,n} - \widehat{\mu}_{ikuv,n}\Big)\Big|\Bigg] \\
\leq \; & n \sum_{i=1}^{a} \frac{1}{n_i} \max_{k=1,\ldots,n_i} \Bigg[\Big|\Big(\Psi_{ikuv,n} - \check{\mu}_{ikuv,n}\Big)\Big(\Psi_{iku'v',n} - \widehat{\Psi}_{iku'v',n}\Big)\Big| \\
& \qquad + \Big|\Big(\Psi_{ikuv,n} - \check{\mu}_{ikuv,n}\Big)\Big(\check{\mu}_{iku'v',n} - \widehat{\mu}_{iku'v',n}\Big)\Big| \\
& \qquad + \Big|\Big(\widehat{\Psi}_{iku'v',n} - \widehat{\mu}_{iku'v',n}\Big)\Big(\Psi_{ikuv,n} - \widehat{\Psi}_{ikuv,n}\Big)\Big| \\
& \qquad + \Big|\Big(\widehat{\Psi}_{iku'v',n} - \widehat{\mu}_{iku'v',n}\Big)\Big(\check{\mu}_{ikuv,n} - \widehat{\mu}_{ikuv,n}\Big)\Big|\Bigg] \\
\leq \; & n \sum_{i=1}^{a} \frac{1}{n_i} \Bigg(\max_{k=1,\ldots,n_i} \Big|\Big(\Psi_{ikuv,n} - \check{\mu}_{ikuv,n}\Big)\Big(\Psi_{iku'v',n} - \widehat{\Psi}_{iku'v',n}\Big)\Big| \\
& \qquad + \max_{k=1,\ldots,n_i} \Big|\Big(\Psi_{ikuv,n} - \check{\mu}_{ikuv,n}\Big)\Big(\check{\mu}_{iku'v',n} - \widehat{\mu}_{iku'v',n}\Big)\Big| \\
& \qquad + \max_{k=1,\ldots,n_i} \Big|\Big(\widehat{\Psi}_{iku'v',n} - \widehat{\mu}_{iku'v',n}\Big)\Big(\Psi_{ikuv,n} - \widehat{\Psi}_{ikuv,n}\Big)\Big| \\
& \qquad + \max_{k=1,\ldots,n_i} \Big|\Big(\widehat{\Psi}_{iku'v',n} - \widehat{\mu}_{iku'v',n}\Big)\Big(\check{\mu}_{ikuv,n} - \widehat{\mu}_{ikuv,n}\Big)\Big|\Bigg).
\end{aligned}
$$

Aus der Annahme (A2), der gleichmäßigen Beschränktheit der Zufallsvariablen $\Psi_{ikuv,n} - \check{\mu}_{ikuv,n}$ und $\widehat{\Psi}_{iku'v',n} - \widehat{\mu}_{iku'v',n}$ sowie (A.12) und (A.13) ergibt sich

$$
\begin{aligned}
& n \sum_{i=1}^{a} \frac{1}{n_i} \Bigg(\max_{k=1,\ldots,n_i} \Big|\Big(\Psi_{ikuv,n} - \check{\mu}_{ikuv,n}\Big)\Big(\Psi_{iku'v',n} - \widehat{\Psi}_{iku'v',n}\Big)\Big| \\
& \qquad + \max_{k=1,\ldots,n_i} \Big|\Big(\Psi_{ikuv,n} - \check{\mu}_{ikuv,n}\Big)\Big(\check{\mu}_{iku'v',n} - \widehat{\mu}_{iku'v',n}\Big)\Big| \\
& \qquad + \max_{k=1,\ldots,n_i} \Big|\Big(\widehat{\Psi}_{iku'v',n} - \widehat{\mu}_{iku'v',n}\Big)\Big(\Psi_{ikuv,n} - \widehat{\Psi}_{ikuv,n}\Big)\Big| \\
& \qquad + \max_{k=1,\ldots,n_i} \Big|\Big(\widehat{\Psi}_{iku'v',n} - \widehat{\mu}_{iku'v',n}\Big)\Big(\check{\mu}_{ikuv,n} - \widehat{\mu}_{ikuv,n}\Big)\Big|\Bigg) \stackrel{\text{f.s.}}{\longrightarrow} 0.
\end{aligned}
$$

Somit ist schließlich auch (A.8) bewiesen.

Die zweite Aussage des Satzes ergibt sich aus der Tatsache, dass aus der fast sicheren Konvergenz gleichmäßig beschränkter Zufallsvariablen auch die Konvergenz aller Momente folgt (siehe z. B. van der Vaart, 1998, Abschnitt 2.5). □

A.16 Beweis von Satz 5.10

Beweis von Satz 5.10. Nach Satz 5.6 auf Seite 62 und Lemma 5.4 auf Seite 61 existieren Zufallsvariablen

$$Y_n \sim \sqrt{n}(\widehat{p}_{is,n} - p_{is}) \quad \text{und} \quad Z_n \sim N\left(0,\, \sigma^2_{is,n}\right) \quad \text{mit} \quad Y_n - Z_n \xrightarrow{\text{f.s.}} 0.$$

Wegen Annahme (A4) folgt $Y_n/\sigma_{is,n} - Z_n/\sigma_{is,n} \xrightarrow{\text{f.s.}} 0$, also wiederum nach Lemma 5.4

$$\rho\left[\sqrt{n}\,\frac{\widehat{p}_{is,n} - p_{is}}{\sigma_{is,n}},\, \mathrm{N}(0,\,1)\right] \to 0$$

und damit aufgrund von Lemma 5.3 $\sqrt{n}(\widehat{p}_{is,n} - p_{is})/\sigma_{is,n} \xrightarrow{\mathcal{D}} \mathrm{N}(0,\,1)$. Weiterhin gilt wegen Annahme (A4) und Korollar 5.9 auf Seite 65

$$\frac{\sigma_{is,n}}{\widehat{\sigma}_{is,n}} = \frac{\sigma_{is,n} - \widehat{\sigma}_{is,n}}{\widehat{\sigma}_{is,n}} + 1 \xrightarrow{\text{f.s.}} 1.$$

Mit dem Satz von Slutsky erhalten wir schließlich

$$\sqrt{n}\,\frac{\widehat{p}_{is,n} - p_{is}}{\widehat{\sigma}_{is,n}} = \frac{\sigma_{is,n}}{\widehat{\sigma}_{is,n}}\sqrt{n}\,\frac{\widehat{p}_{is,n} - p_{is}}{\sigma_{is,n}} \xrightarrow{\mathcal{D}} \mathrm{N}(0,\,1).$$

Die anderen beiden Aussagen des Satzes können analog bewiesen werden. □

A.17 Beweis von Satz 5.13

Beweis von Satz 5.13. Wir führen den Beweis für p_{is}. Für p_{i*} und p_{*s} sind analoge Beweise möglich.

Zunächst schließen wir die Fälle $p_{is} = 0$ und $p_{is} = 1$ aus. Im Fall $p_{is} = 0$ gilt nach Definition 3.2 auf Seite 30 für alle Paare $(u, v) \neq (i, s)$

$$\mathrm{P}(X_{u2v1} < X_{i1s1}) + \frac{1}{2}\,\mathrm{P}(X_{u2v1} = X_{i1s1}) = 0.$$

Deshalb gilt für alle Indizes (k, ℓ) und (u, w, v, r)

$$\mathrm{P}(X_{uwvr} < X_{iks\ell}) + \frac{1}{2}\,\mathrm{P}(X_{uwvr} = X_{iks\ell}) = 0.$$

Also ist $\Psi_{uwis,n}$ für alle Paare (u, w) fast sicher konstant:

$$\Psi_{uwis,n} = \begin{cases} \frac{n_i}{\lambda_{i\cdot s}} \int H_{is}\, d\widehat{F}_{iws} - \frac{1}{at-1} \sum_{v\neq s} \frac{n_i}{\lambda_{i\cdot s}} \int F_{is}\, d\widehat{F}_{iwv} \overset{\text{f.s.}}{=} -\frac{1}{at-1} \sum_{v\neq s} \frac{n_i}{\lambda_{i\cdot s}} & : i = u, \\ \frac{-1}{at-1} \sum_{v=1}^{t} \frac{n_u}{\lambda_{u\cdot v}} \int F_{is}\, d\widehat{F}_{uwv} \overset{\text{f.s.}}{=} \frac{-1}{at-1} \sum_{v=1}^{t} \frac{n_u}{\lambda_{u\cdot v}} & : i \neq u. \end{cases}$$

Damit folgt im Widerspruch zu Annahme (A4) auf Seite 65 für alle $n \in \mathbb{N}$

$$\sigma_{is,n}^2 = \mathrm{Var}\left(\sqrt{n}\sum_{u=1}^{a}\frac{1}{n_u}\sum_{w=1}^{n_u}\Psi_{uwis,n}\right) = n\sum_{u=1}^{a}\frac{1}{n_u^2}\sum_{w=1}^{n_u}\mathrm{Var}(\Psi_{uwis,n}) = 0.$$

Ebenso lässt sich zeigen, dass $p_{is} = 1$ im Widerspruch zu Annahme (A4) steht.

Wenden wir nun die Delta-Methode in Form von Theorem 3.1 aus van der Vaart (1998) mit den Bezeichnungen $\phi = g$, $\theta = p_{is}$, $T_n = \widehat{p}_{is,n}$ und $r_n = \sqrt{n}/\sigma_{is,n}$ an, so erhalten wir

$$\sqrt{n}\frac{g(\widehat{p}_{is,n}) - g(p_{is})}{g'(p_{is})\sigma_{is,n}} \xrightarrow{\mathcal{D}} \mathrm{N}(0,1).$$

Wegen der Konsistenz von $\widehat{p}_{is}$, der Stetigkeit von g' und $g'(p_{is}) \neq 0$ gilt

$$\frac{g'(p_{is})}{g'(\widehat{p}_{is,n})} \xrightarrow{\text{f.s.}} 1.$$

Aus dem Beweis von Satz 5.10 (siehe Seite 65) wissen wir bereits $\sigma_{is,n}/\widehat{\sigma}_{is,n} \xrightarrow{\text{f.s.}} 1$ konvergiert. Nun folgt die Behauptung aus dem Satz von Slutsky:

$$\sqrt{n}\frac{g(\widehat{p}_{is,n}) - g(p_{is})}{g'(\widehat{p}_{is,n})\widehat{\sigma}_{is,n}} = \frac{g'(p_{is})}{g'(\widehat{p}_{is,n})}\,\frac{\sigma_{is,n}}{\widehat{\sigma}_{is,n}}\,\sqrt{n}\,\frac{g(\widehat{p}_{is,n}) - g(p_{is})}{g'(p_{is})\sigma_{is,n}} \xrightarrow{\mathcal{D}} \mathrm{N}(0,1).$$

□

A.18 Beweis von Satz 6.5

Beweis von Satz 6.5. Wir zeigen zunächst

$$\widehat{\gamma}_{i,n}^{\min} - \gamma_{i,n}^{\min} \xrightarrow{\text{f.s.}} 0. \tag{A.14}$$

Es gilt

$$\widehat{\gamma}_{i,n}^{\min} - \gamma_{i,n}^{\min} = \frac{\widehat{\sigma}_{ii,n}^{XY}}{\widehat{\sigma}_{ii,n}^{Y}} - \frac{\sigma_{ii,n}^{XY}}{\sigma_{ii,n}^{Y}} = \frac{\widehat{\sigma}_{ii,n}^{XY}\left(\sigma_{ii,n}^{Y} - \widehat{\sigma}_{ii,n}^{Y}\right) + \widehat{\sigma}_{ii,n}^{Y}\left(\widehat{\sigma}_{ii,n}^{XY} - \sigma_{ii,n}^{XY}\right)}{\widehat{\sigma}_{ii,n}^{Y}\sigma_{ii,n}^{Y}}.$$

Wenden wir Annahme (A2) auf Seite 13 auf die Gleichungen (6.3) und (6.4) (siehe Seite 86) an, so erkennen wir, dass die Schätzer $\widehat{\sigma}_{ii,n}^{Y}$ und $\widehat{\sigma}_{ii,n}^{XY}$ gleichmäßig beschränkt sind. Da die Y_{ik} keine Einpunktverteilung haben, gilt $\tau_Y^2 > 0$. Mit der Identität (6.2) auf Seite 86 folgt $\sigma_{ii,n}^{Y} \geq \tau_Y^2 > 0$, also $1/\sigma_{ii,n}^{Y} \leq 1/\tau_Y^2$. Wegen Satz 6.4 (siehe Seite 87) ergibt sich

$$\mathrm{P}\left(\liminf_{n\to\infty}\widehat{\sigma}_{ii,n}^{Y} \geq \tau_Y^2\right) = 1$$

und damit

$$\mathrm{P}\left(\limsup_{n\to\infty}\frac{1}{\widehat{\sigma}_{ii,n}^{Y}} \leq \frac{1}{\tau_Y^2}\right) = 1. \tag{A.15}$$

Durch Anwendung von Satz 6.4 auf $\sigma_{ii,n}^{Y} - \widehat{\sigma}_{ii,n}^{Y}$ und $\widehat{\sigma}_{ii,n}^{XY} - \sigma_{ii,n}^{XY}$ erhalten wir schließlich (A.14).

(a) Um die starke Konsistenz von $\widehat{p}_{i,n}^{\min}$ nachzuweisen, müssen wir

$$\widehat{p}_{i,n}^{\min} - p_i^X = \widehat{p}_{i,n}^X - p_i^X - \widehat{\gamma}_{i,n}^{\min}\left(\widehat{p}_{i,n}^Y - p_i^Y\right) \xrightarrow{\text{f.s.}} 0$$

zeigen. Aus Satz 4.3 auf Seite 50 wissen wir bereits

$$\widehat{p}_{i,n}^X - p_i^X \xrightarrow{\text{f.s.}} 0 \quad \text{und} \quad \widehat{p}_{i,n}^Y - p_i^Y \xrightarrow{\text{f.s.}} 0.$$

Die starke Konsistenz ist deshalb bewiesen, wenn für eine Konstante $M \in \mathbb{R}$

$$\mathrm{P}\left(\limsup_{n\to\infty}\left|\widehat{\gamma}_{i,n}^{\min}\right| \le M\right) = 1 \tag{A.16}$$

gilt. Wegen $\widehat{\gamma}_{i,n}^{\min} = \widehat{\sigma}_{ii,n}^{XY}/\widehat{\sigma}_{ii,n}^{Y}$ ergibt sich dies aus der gleichmäßigen Beschränktheit der $\widehat{\sigma}_{ii,n}^{XY}$ und (A.15).

(b) Es ist

$$\begin{aligned}&\left(\widehat{\sigma}_{i,n}^{\min}\right)^2 - \left(\sigma_{i,n}^{\min}\right)^2 \\ &= \underbrace{\widehat{\sigma}_{ii,n}^X - \sigma_{ii,n}^X}_{A_n} - \underbrace{2\left(\widehat{\gamma}_{i,n}^{\min}\widehat{\sigma}_{ii,n}^{XY} - \gamma_{i,n}^{\min}\sigma_{ii,n}^{XY}\right)}_{B_n} + \underbrace{\left[\widehat{\gamma}_{i,n}^{\min}\right]^2\widehat{\sigma}_{ii,n}^Y - \left[\gamma_{i,n}^{\min}\right]^2\sigma_{ii,n}^Y}_{C_n} \xrightarrow{\text{f.s.}} 0\end{aligned}$$

zu zeigen. Dass A_n fast sicher gegen 0 konvergiert, wissen wir bereits aus Satz 6.4. Um nachzuweisen, dass B_n und C_n fast sicher gegen 0 konvergieren, zerlegen wir diese weiter:

$$\begin{aligned}B_n &= 2\widehat{\gamma}_{i,n}^{\min}\left(\widehat{\sigma}_{ii,n}^{XY} - \sigma_{ii,n}^{XY}\right) + 2\sigma_{ii,n}^{XY}\left(\widehat{\gamma}_{i,n}^{\min} - \gamma_{i,n}^{\min}\right)\\ C_n &= \left[\widehat{\gamma}_{i,n}^{\min}\right]^2\left(\widehat{\sigma}_{ii,n}^Y - \sigma_{ii,n}^Y\right) + \sigma_{ii,n}^Y\left(\left[\widehat{\gamma}_{i,n}^{\min}\right]^2 - \left[\gamma_{i,n}^{\min}\right]^2\right)\\ &= \left[\widehat{\gamma}_{i,n}^{\min}\right]^2\left(\widehat{\sigma}_{ii,n}^Y - \sigma_{ii,n}^Y\right) + \sigma_{ii,n}^Y\left(\widehat{\gamma}_{i,n}^{\min} + \gamma_{i,n}^{\min}\right)\left(\widehat{\gamma}_{i,n}^{\min} - \gamma_{i,n}^{\min}\right).\end{aligned}$$

Nun ergeben sich $B_n \xrightarrow{\text{f.s.}} 0$ und $C_n \xrightarrow{\text{f.s.}} 0$ aus der gleichmäßigen Beschränktheit der $\sigma_{ii,n}^{XY}$, der $\sigma_{ii,n}^Y$ und der $\gamma_{i,n}^{\min}$, sowie aus (A.14), (A.16) und Satz 6.4.

(c) Nach Satz 6.3 auf Seite 85 und Lemma 5.4 auf Seite 61 existieren Zufallsvektoren

$$\boldsymbol{\Gamma}_n = \begin{pmatrix}\Gamma_n^X\\ \Gamma_n^Y\end{pmatrix} \sim \sqrt{n}\begin{pmatrix}\widehat{p}_{i,n}^X - p_i^X\\ \widehat{p}_{i,n}^Y - p_i^Y\end{pmatrix}$$

und

$$\boldsymbol{\Delta}_n = \begin{pmatrix}\Delta_n^X\\ \Delta_n^Y\end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}\sigma_{ii,n}^X & \sigma_{ii,n}^{XY}\\ \sigma_{ii,n}^{XY} & \sigma_{ii,Y}^Y\end{pmatrix}\right)$$

mit

$$\boldsymbol{\Gamma}_n - \boldsymbol{\Delta}_n \xrightarrow{\text{f.s.}} \mathbf{0}.$$

Wegen der gleichmäßigen Beschränktheit der $\gamma_{i,n}^{\min}$ und wegen der Annahme $\liminf_{n\to\infty}\sigma_{i,n}^{\min} > 0$ folgt daraus

$$\frac{\Gamma_n^X - \gamma_{i,n}^{\min}\Gamma_n^Y}{\sigma_{i,n}^{\min}} - \frac{\Delta_n^X - \gamma_{i,n}^{\min}\Delta_n^Y}{\sigma_{i,n}^{\min}} = \frac{\Gamma_n^X - \Delta_n^X}{\sigma_{i,n}^{\min}} - \gamma_{i,n}^{\min}\frac{\Gamma_n^Y - \Delta_n^Y}{\sigma_{i,n}^{\min}} \xrightarrow{\text{f.s.}} 0.$$

Unter Beachtung von

$$\frac{\Gamma_n^X - \gamma_{i,n}^{\min}\Gamma_n^Y}{\sigma_{i,n}^{\min}} \sim \sqrt{n}\,\frac{\widehat{p}_{i,n}^X - p_i^X - \gamma_{i,n}^{\min}\left(\widehat{p}_{i,n}^Y - p_i^Y\right)}{\sigma_{i,n}^{\min}} \quad \text{und} \quad \frac{\Delta_n^X - \gamma_{i,n}^{\min}\Delta_n^Y}{\sigma_{i,n}^{\min}} \sim \mathrm{N}(0,1)$$

ergibt sich somit durch erneute Anwendung von Lemma 5.4

$$\rho\left(\sqrt{n}\,\frac{\widehat{p}_{i,n}^X - p_i^X - \gamma_{i,n}^{\min}\left(\widehat{p}_{i,n}^Y - p_i^Y\right)}{\sigma_{i,n}^{\min}}, \mathrm{N}(0,1)\right) \to 0$$

und nach Lemma 5.3 auf Seite 61

$$\sqrt{n}\,\frac{\widehat{p}_{i,n}^X - p_i^X - \gamma_{i,n}^{\min}\left(\widehat{p}_{i,n}^Y - p_i^Y\right)}{\sigma_{i,n}^{\min}} \xrightarrow{\mathcal{D}} \mathrm{N}(0,1).$$

Zum Beweis der Behauptung fehlt nun nur noch der Nachweis von

$$\sqrt{n}\,\frac{\widehat{p}_{i,n}^X - p_i^X - \gamma_{i,n}^{\min}\left(\widehat{p}_{i,n}^Y - p_i^Y\right)}{\sigma_{i,n}^{\min}} - \sqrt{n}\,\frac{\widehat{p}_{i,n}^X - p_i^X - \widehat{\gamma}_{i,n}^{\min}\left(\widehat{p}_{i,n}^Y - p_i^Y\right)}{\sigma_{i,n}^{\min}} \xrightarrow{\mathrm{P}} 0 \qquad \text{(A.17)}$$

sowie

$$\sqrt{n}\,\frac{\widehat{p}_{i,n}^X - p_i^X - \widehat{\gamma}_{i,n}^{\min}\left(\widehat{p}_{i,n}^Y - p_i^Y\right)}{\sigma_{i,n}^{\min}} - \sqrt{n}\,\frac{\widehat{p}_{i,n}^X - p_i^X - \widehat{\gamma}_{i,n}^{\min}\left(\widehat{p}_{i,n}^Y - p_i^Y\right)}{\widehat{\sigma}_{i,n}^{\min}} \xrightarrow{\mathrm{P}} 0. \qquad \text{(A.18)}$$

Zum Nachweis von (A.17) bemerken wir zunächst

$$\begin{aligned}\sqrt{n}\,\frac{\widehat{p}_{i,n}^X - p_i^X - \gamma_{i,n}^{\min}\left(\widehat{p}_{i,n}^Y - p_i^Y\right)}{\sigma_{i,n}^{\min}} - \sqrt{n}\,\frac{\widehat{p}_{i,n}^X - p_i^X - \widehat{\gamma}_{i,n}^{\min}\left(\widehat{p}_{i,n}^Y - p_i^Y\right)}{\sigma_{i,n}^{\min}}& \\ = \frac{\widehat{\gamma}_{i,n}^{\min} - \gamma_{i,n}^{\min}}{\sigma_{i,n}^{\min}}\sqrt{n}\left(\widehat{p}_{i,n}^Y - p_i^Y\right).&\end{aligned}$$

Wegen $\widehat{\gamma}_{i,n}^{\min} - \gamma_{i,n}^{\min} \xrightarrow{\text{f.s.}} 0$ und $\liminf_{n\to\infty} \sigma_{i,n}^{\min} > 0$ gilt

$$\frac{\widehat{\gamma}_{i,n}^{\min} - \gamma_{i,n}^{\min}}{\sigma_{i,n}^{\min}} \xrightarrow{\text{f.s.}} 0.$$

Weil zudem auf Grund von Satz 6.3 auch

$$\rho\left[\sqrt{n}\left(\widehat{p}_{i,n}^Y - p_i^Y\right), \mathrm{N}\left(0, \sigma_{ii,n}^Y\right)\right] \to 0$$

gilt, ergibt sich (A.17) und damit

$$\sqrt{n}\,\frac{\widehat{p}_{i,n}^X - p_i^X - \widehat{\gamma}_{i,n}^{\min}\left(\widehat{p}_{i,n}^Y - p_i^Y\right)}{\sigma_{i,n}^{\min}} \xrightarrow{\mathcal{D}} \mathrm{N}(0,1) \qquad \text{(A.19)}$$

aus Lemma D.15 in Domhof (1999).

Um (A.18) zu beweisen, schreiben wir

$$\sqrt{n}\,\frac{\widehat{p}^{X}_{i,n}-p^{X}_{i}-\widehat{\gamma}^{\min}_{i,n}\left(\widehat{p}^{Y}_{i,n}-p^{Y}_{i}\right)}{\sigma^{\min}_{i,n}}-\sqrt{n}\,\frac{\widehat{p}^{X}_{i,n}-p^{X}_{i}-\widehat{\gamma}^{\min}_{i,n}\left(\widehat{p}^{Y}_{i,n}-p^{Y}_{i}\right)}{\widehat{\sigma}^{\min}_{i,n}}$$
$$=\frac{\widehat{\sigma}^{\min}_{i,n}-\sigma^{\min}_{i,n}}{\widehat{\sigma}^{\min}_{i,n}}\sqrt{n}\,\frac{\widehat{p}^{X}_{i,n}-p^{X}_{i}-\widehat{\gamma}^{\min}_{i,n}\left(\widehat{p}^{Y}_{i,n}-p^{Y}_{i}\right)}{\sigma^{\min}_{i,n}}.$$

Wegen Teil (b) dieses Satzes gilt

$$\mathrm{P}\left(\liminf_{n\to\infty}\widehat{\sigma}^{\min}_{i,n}\geq\liminf_{n\to\infty}\sigma^{\min}_{i,n}\right)=1,$$

so dass wir auf Grund der Annahme $\liminf_{n\to\infty}\sigma^{\min}_{i,n}>0$ bei erneuter Anwendung von Teil (b)

$$\frac{\widehat{\sigma}^{\min}_{i,n}-\sigma^{\min}_{i,n}}{\widehat{\sigma}^{\min}_{i,n}}\xrightarrow{\text{f.s.}}0$$

erhalten. Nun ergibt sich (A.18) unter Berücksichtigung von (A.19) aus dem Satz von Slutsky. □

A.19 Beweis von Satz 6.7

Beweis von Satz 6.7. Nach Satz 6.3 auf Seite 85 gilt jedenfalls

$$\rho\left[\sqrt{n}\left[\left(\widehat{\boldsymbol{p}}^{X}_{n}-\boldsymbol{p}^{X}\right)',\left(\widehat{\boldsymbol{p}}^{Y}_{n}-\boldsymbol{p}^{Y}\right)'\right]',\mathrm{N}(\mathbf{0},\boldsymbol{V}_{n})\right]\to 0.$$

Da ρ eine Metrik ist und somit der Dreiecksungleichung genügt, müssen wir wegen Lemma 5.3 (siehe Seite 61) nur noch

$$\rho[\mathrm{N}(\mathbf{0},\boldsymbol{V}_{n}),\mathrm{N}(\mathbf{0},\boldsymbol{V})]\to 0$$

zeigen. Dies ergibt sich auf Grund der gleichmäßigen Beschränktheit der $\boldsymbol{V}_n$ aus den Sätzen D.12 und D.14 in Domhof (1999). □

A.20 Beweis von Satz 6.9

Beweis von Satz 6.9. Die Abbildung

$$\phi:(0,1)^2\to\mathbb{R}^2,\ \begin{pmatrix}x\\y\end{pmatrix}\mapsto\begin{pmatrix}g(x)\\g(y)\end{pmatrix}$$

ist differenzierbar und hat an der Stelle (x,y) die Jacobische Funktionalmatrix

$$\boldsymbol{\phi}'_{(x,y)}=\begin{pmatrix}g'(x)&0\\0&g'(y)\end{pmatrix}.$$

Wegen
$$\sqrt{n}\begin{pmatrix} g\left(\widehat{p}_{i,n}^X\right)-g\left(p_i^X\right) \\ g\left(\widehat{p}_{i,n}^Y\right)-g\left(p_i^Y\right)\end{pmatrix}=\sqrt{n}\left(\phi\begin{pmatrix}\widehat{p}_{i,n}^X \\ \widehat{p}_{i,n}^Y\end{pmatrix}-\phi\begin{pmatrix}p_i^X \\ p_i^Y\end{pmatrix}\right)$$
und Satz 6.7 auf Seite 90 liefert die multivariate Delta-Methode (siehe z. B. van der Vaart, 1998, Theorem 3.1)
$$\begin{aligned}\sqrt{n}\begin{pmatrix} g\left(\widehat{p}_{i,n}^X\right)-g\left(p_i^X\right) \\ g\left(\widehat{p}_{i,n}^Y\right)-g\left(p_i^Y\right)\end{pmatrix} &\xrightarrow{\mathcal{D}} \mathrm{N}\left(\mathbf{0},\boldsymbol{\phi}'_{\left(p_i^X,p_i^Y\right)}\boldsymbol{V}\boldsymbol{\phi}'_{\left(p_i^X,p_i^Y\right)}\right)\\
&= \mathrm{N}\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}\left[g'\left(p_i^X\right)\right]^2\sigma_{ii}^X & g'\left(p_i^X\right)g'\left(p_i^Y\right)\sigma_{ii}^{XY} \\ g'\left(p_i^X\right)g'\left(p_i^Y\right)\sigma_{ii}^{XY} & \left[g'\left(p_i^Y\right)\right]^2\sigma_{ii}^Y\end{pmatrix}\right).\end{aligned}$$
□

A.21 Beweis von Satz 6.11

Beweis von Satz 6.11. (a) Aus Satz 4.3 auf Seite 50 wissen wir bereits $\widehat{p}_{i,n}^X \xrightarrow{\text{f.s.}} p_i^X$ und $\widehat{p}_{i,n}^Y \xrightarrow{\text{f.s.}} p_i^Y$. Wegen der Stetigkeit von g und g' erhalten wir
$$g\left(\widehat{p}_{i,n}^X\right)\xrightarrow{\text{f.s.}} g\left(p_i^X\right),\quad g'\left(\widehat{p}_{i,n}^X\right)\xrightarrow{\text{f.s.}} g'\left(p_i^X\right)\quad\text{und}\quad g\left(\widehat{p}_{i,n}^Y\right)\xrightarrow{\text{f.s.}} g\left(p_i^Y\right). \tag{A.20}$$
Wegen Satz 6.8 auf Seite 91 gilt außerdem
$$\widehat{\sigma}_{ii,n}^X\xrightarrow{\text{f.s.}}\sigma_{ii}^X,\quad \widehat{\sigma}_{ii,n}^Y\xrightarrow{\text{f.s.}}\sigma_{ii}^Y\quad\text{und}\quad \widehat{\sigma}_{ii,n}^{XY}\xrightarrow{\text{f.s.}}\sigma_{ii}^{XY}. \tag{A.21}$$
Da die Verteilung der Y_{ik} keine Einpunktverteilung ist, gilt $\sigma_{ii}^Y>0$ und deshalb auch
$$\frac{1}{\widehat{\sigma}_{ii,n}^Y}\xrightarrow{\text{f.s.}}\frac{1}{\sigma_{ii}^Y}. \tag{A.22}$$
Anhand der Gleichungen (A.20), (A.21) und (A.22) ergibt sich aus den wohlbekannten Rechenregeln für die fast sichere Konvergenz schrittweise zunächst
$$\widehat{\gamma}_{i,n}^{*\min}=\frac{g'\left(\widehat{p}_{i,n}^X\right)\widehat{\sigma}_{ii,n}^{XY}}{g'\left(\frac{1}{2}\right)\widehat{\sigma}_{ii,n}^Y}\xrightarrow{\text{f.s.}}\frac{g'\left(p_i^X\right)\sigma_{ii}^{XY}}{g'\left(p_i^Y\right)\sigma_{ii}^Y}=\gamma_i^{*\min} \tag{A.23}$$
und damit auch
$$\begin{aligned}\widehat{p}_{i,n}^{*\min} &= g\left(\widehat{p}_{i,n}^X\right)-\widehat{\gamma}_{i,n}^{*\min}\left[g\left(\widehat{p}_{i,n}^Y\right)-g\left(p_i^Y\right)\right]\\ &\xrightarrow{\text{f.s.}} g\left(p_i^X\right)-\gamma_i^{*\min}\left[g\left(p_i^Y\right)-g\left(p_i^Y\right)\right]=p_i^*.\end{aligned}$$

(b) Mit den Gleichungen (A.20), (A.21) und (A.22) und den wohlbekannten Rechenregeln für die fast sichere Konvergenz können wir auch Behauptung (b) beweisen:
$$\begin{aligned}\left(\widehat{\sigma}_{i,n}^{*\min}\right)^2 &= \left[g'\left(\widehat{p}_{i,n}^X\right)\right]^2\left(\widehat{\sigma}_{ii,n}^X-\frac{\left(\widehat{\sigma}_{ii,n}^{XY}\right)^2}{\widehat{\sigma}_{ii,n}^Y}\right)\\ &\xrightarrow{\text{f.s.}}\left[g'\left(p_i^X\right)\right]^2\left(\sigma_{ii}^X-\frac{\left(\sigma_{ii}^{XY}\right)^2}{\sigma_{ii}^Y}\right)=\left(\sigma_i^{*\min}\right)^2.\end{aligned}$$

(c) Wegen (A.23) und wegen des Lemmas von Slutsky gilt

$$\sqrt{n}\,\frac{\widehat{p}_{i,n}^{*\gamma_i^{*\,\min}} - p_i^*}{\sigma_i^{*\,\min}} - \sqrt{n}\,\frac{\widehat{p}_{i,n}^{*\,\min} - p_i^*}{\sigma_i^{*\,\min}} = \left(\gamma_i^{*\,\min} - \widehat{\gamma}_{i,n}^{*\,\min}\right)\sqrt{n}\,\frac{g\left(\widehat{p}_{i,n}^{Y}\right) - g\left(p_i^{Y}\right)}{\sigma_i^{*\,\min}} \xrightarrow{\mathrm{P}} 0$$

und damit nach Korollar 6.10

$$\sqrt{n}\,\frac{\widehat{p}_{i,n}^{*\,\min} - p_i^*}{\sigma_i^{*\,\min}} \xrightarrow{\mathcal{D}} \mathrm{N}(0,1).$$

Aus Teil (b) dieses Satzes und wiederum dem Lemma von Slutsky folgt nun

$$\sqrt{n}\,\frac{\widehat{p}_{i,n}^{*\,\min} - p_i^*}{\sigma_i^{*\,\min}} - \sqrt{n}\,\frac{\widehat{p}_{i,n}^{*\,\min} - p_i^*}{\widehat{\sigma}_i^{*\,\min}} = \frac{\widehat{\sigma}_i^{*\,\min} - \sigma_i^{*\,\min}}{\widehat{\sigma}_i^{*\,\min}}\sqrt{n}\,\frac{\widehat{p}_{i,n}^{*\,\min} - p_i^*}{\sigma_i^{*\,\min}} \xrightarrow{\mathrm{P}} 0.$$

und deshalb wie behauptet

$$\sqrt{n}\,\frac{\widehat{p}_{i,n}^{*\,\min} - p_i^*}{\widehat{\sigma}_i^{*\,\min}} \xrightarrow{\mathcal{D}} \mathrm{N}(0,1).$$

□

Anhang B

Software

Zur praktischen Umsetzung von Teilen der in dieser Arbeit vorgestellten Verfahren existieren zwei vom Autor der vorliegenden Arbeit erstellte SAS-Makros: `RE_CI` und `SIMPLE_RE_CI`. Die beiden Makros sind über das Internet unter der URL

`http://www.ams.med.uni-goettingen.de/Projekte/LD/Makros_LD.html`

erhältlich. In den folgenden beiden Abschnitten wird kurz die Handhabung der beiden Makros beschrieben.

B.1 Das Makro RE_CI

Ist das Makro `RE_CI` in der Datei `RE_CI.SAS` im Pfad `PATH` enthalten, so kann es in SAS mit dem Befehl

```
%INCLUDE 'PATH/RE_CI.SAS';
```

eingebunden werden.

Die SAS-Datensätze müssen, um mit dem Makro ausgewertet werden zu können, für jede Beobachtung $X_{iks\ell}$ eine Zeile enthalten. Jede Zeile muss die Informationen darüber enthalten, zu welcher Gruppe und welcher Versuchseinheit die Beobachtung gehört, sowie unter welcher Bedingung die Beobachtung gemacht wurde. Somit besteht ein Datensatz für das Makro `RE_CI` in der Regel aus vier und in den Spezialfällen $a = 1$ oder $t = 1$ aus drei Variablen mit der folgenden Bedeutung:

- Gruppenvariable (falls $a \neq 1$): Bezeichnungen für die Gruppen.
- Versuchseinheitsvariable: Bezeichnungen für die Versuchseinheiten. Alle Versuchseinheiten müssen auch dann, wenn sie zu unterschiedlichen Gruppen gehören, verschiedene Bezeichnungen tragen.
- Bedingungsvariable (falls $t \neq 1$): Bezeichnungen für die Bedingungen.
- Beobachtungsvariable: Werte der Beobachtungen. Fehlende Werte können durch einen Punkt dargestellt werden.

Der Datensatz für den Wasser-Irrgarten-Test kann zum Beispiel auf folgende Weise erzeugt werden:

```
DATA wasser;
  INPUT behandlung $ mutter tag score;
  CARDS;
  Placebo  1 1 13
  Placebo  1 2 21
  Placebo  1 1 11
  Placebo  1 2 21
  Placebo  2 1 14
  Placebo  2 2 13
  Placebo  2 1 20
  Placebo  2 2 21
  ...
  Dosis3  19 2 21
RUN;
```

Das Makro `RE_CI` verfügt über sechs Schlüsselwort-Parameter:

- `DATA`: Name des auszuwertenden Datensatzes.
- `VAR`: Name der Beobachtungsvariablen.
- `GROUP` (falls $a \neq 1$): Name der Gruppenvariablen.
- `TIME` (falls $t \neq 1$): Name der Bedingungsvariablen.
- `SUBJECT`: Name der Versuchseinheitsvariablen.
- `ALPHA`: Legt das Niveau der Konfidenzintervalle auf 1 - `ALPHA` fest. Die Voreinstellung ist `0.05`.

Der Aufruf des Makros zur Berechnung von 90%-Konfidenzintervallen für das oben genannte Beispiel lautet dementsprechend:

```
%RE_CI(DATA=wasser, VAR=score, GROUP=behandlung, TIME=tag,
       SUBJECT=mutter, ALPHA=0.1)
```

Auf der ersten Seite der Ausgabe des Makros befinden sich allgemeine Informationen über den Datensatz, wie Namen, Anzahlen und Stichprobenumfänge. Für das in diesem Abschnitt gewählte Beispiel hat die erste Seite die folgende Gestalt.

```
                                RE_CI

         Point Estimators and Confidence-Intervals for Relative Effects

                          SAS-Data-Filename: wasser

     Response-Variable: score        Observations:        288
     Group-Variable:    gruppe       Groups:                4
     Time-Variable:     tag          Timepoints:            2
     Subject-Variable:  mutter       Subjects:             72
                                     Missing Values:        0
```

Group-Nr.	Group-Name	Time Point-Nr.	Time Point-Name
1	Dosis1	1	1
2	Dosis2	2	7
3	Dosis3		
4	Placebo		

Group	Time	Subjects	Observations	Missing Values
1	1	17	34	0
1	2	17	34	0
2	1	19	38	0
2	2	19	38	0
3	1	19	38	0
3	2	19	38	0
4	1	17	34	0
4	2	17	34	0

Die zweite Seite der Ausgabe enthält die Punktschätzer (Überschrift `RE`), Schätzer für die Varianzen (Überschrift `Variance`) und die Konfidenzintervalle auf Grundlage der Normal-Approximation (Überschriften `Normal Low` und `Normal Up`) und der Transformationsmethode (Überschriften `Trans Low` und `Trans Up`). Zu beachten ist, dass es sich bei den Varianzschätzern nicht um die $\widehat{\sigma}^2_{is,n}$, sondern um die Varianzschätzer für die $\widehat{p}_{is,n}$, also um $\widehat{\sigma}^2_{is,n}/n$ handelt. Für das hier betrachtete Beispiel sieht die zweite Seite der Ausgabe folgendermaßen aus.

Relative Effects, Variances and Confidence-Limits (alpha=0.1)

Group	Time	RE	Variance	Normal Low	Normal Up	Trans Low	Trans Up
1	1	.47709	.00257	.39370	.56049	.39509	.56035
1	2	.65563	.00298	.56590	.74535	.56131	.73910
2	1	.39234	.00292	.30344	.48124	.30781	.48385
2	2	.61980	.00246	.53829	.70131	.53564	.69733
3	1	.27720	.00188	.20594	.34846	.21181	.35371
3	2	.52551	.00263	.44121	.60981	.44128	.60831
4	1	.38329	.00241	.30251	.46407	.30633	.46659
4	2	.66914	.00189	.59758	.74070	.59413	.73643

B.2 Das Makro SIMPLE_RE_CI

Das SAS-Makro `SIMPLE_RE_CI` unterscheidet sich von dem Makro `RE_CI` insofern, als es nur Datensätze aus Versuchen mit einer einfacheren Strukur auswerten kann. Für alle Tripel (i, k, s) muss bei diesem Makro $m_{iks} = 1$ gelten. Es dürfen also keine Werte fehlen und keine abhängigen Messwiederholungen vorliegen. Zusätzlich zu den vom Makro `RE_CI` berechneten Konfidenzintervallen auf Grundlage der Normal-Approximation und der Transformationsmethode können mit dieser Einschränkung mit dem Makro `SIMPLE_RE_CI` aber auch Konfidenzintervalle auf Grundlage der t-Approximation und der Perzentilmethode ermittelt werden.

Die beiden in diesem Anhang vorgestellten Makros stimmen in Bezug auf ihre

Einbindung, die Struktur der zugehörigen SAS-Datensätze und die Syntax im Wesentlichen überein. Um also zum Beispiel den Datensatz aus der Schulterschmerz-Studie zu bearbeiten, könnte der SAS-Datensatz zunächst durch folgendes Programm erzeugt werden:

```
DATA schulter;
  INPUT patient behand $ tag score;
  CARDS;
   1 Y 1 1
   1 Y 2 1
   1 Y 3 1
   1 Y 4 1
   1 Y 5 1
   1 Y 6 1
   2 Y 1 3
  ...
  19 N 6 1
RUN;
```

Beim Aufruf des Makros `SIMPLE_RE_CI` steht im Vergleich zum Makro `RE_CI` der zusätzliche Schlüsselwort-Parameter `BOOT` zur Verfügung, durch den die Anzahl der erzeugten Bootstrap-Stichproben festgelegt wird. Die Voreinstellung beträgt 5000. Um zum Beispiel 95%-Konfidenzintervalle mit 10000 Bootstrap-Stichproben zu berechnen, kann das Makro für obiges Beispiel wie folgt aufgerufen werden:

```
%SIMPLE_RE_CI(DATA=schulter, VAR=score, GROUP=behand, TIME=tag,
              SUBJECT=patient, BOOT=10000)
```

Die erste Seite der Ausgabe besteht wie bei dem anderen Makro aus allgemeinen Informationen, also Namen, Anzahlen und Stichprobenumfängen:

```
                              SIMPLE_RE_CI

        Point Estimators and  Confidence-Intervals for Relative Effects

                         SAS-Data-Filename: schulter

            Response-Variable: score        Observations:      246
            Group-Variable:    behand       Groups:              2
            Time-Variable:     tag          Timepoints:          6
            Subject-Variable:  patient      Subjects:           41

      Group-Nr. Group-Name  Subjects      Time Point-Nr. Time Point-Name
              1 N                 19                   1               1
              2 Y                 22                   2               2
                                                       3               3
                                                       4               4
                                                       5               5
                                                       6               6
```

Die zweite Seite informiert über Punktschätzer, Varianzschätzer und Konfidenzintervalle. Wir geben zunächst eine Liste der Überschriften in der Ausgabe und den darunter zu findenden Größen an:

- `Group`: Gruppennummer.
- `Time`: Bedingungsnummer.
- `RE`: Punktschätzer $\widehat{p}_{is,n}$.
- `Var`: Varianzschätze $\widehat{\sigma}^2_{is,n}/n$.
- `N Low` / `N Up`: Untere / obere Grenzen der Konfidenzintervalle $\left[u^N_{is,n}, o^N_{is,n}\right]$.
- `t Low` / `t Up`: Untere / obere Grenzen der Konfidenzintervalle $\left[u^t_{is,n}, o^t_{is,n}\right]$.
- `Tr Low` / `Tr Up`: Untere / obere Grenzen der Konfidenzintervalle $\left[u^T_{is,n}, o^T_{is,n}\right]$.
- `Bt Low` / `Bt Up`: Untere / obere Grenzen der Konfidenzintervalle $\left[u^P_{is,n}, o^P_{is,n}\right]$.

Für unser Beispiel sieht die zweite Seite der Ausgabe folgendermaßen aus:

```
          Relative Effects, Variances and Confidence-Limits (alpha=0.05)

Group Time   RE     Var    N Low  N Up   t Low  t Up   Tr Low Tr Up  Bt Low Bt Up
    1    1  .5760  .0036  .4583  .6936  .4508  .7011  .4562  .6874  .4610  .6879
    1    2  .7050  .0020  .6162  .7937  .6109  .7991  .6093  .7854  .6164  .7911
    1    3  .6895  .0022  .5967  .7822  .5909  .7880  .5901  .7740  .5960  .7787
    1    4  .7333  .0019  .6470  .8196  .6416  .8249  .6388  .8104  .6459  .8155
    1    5  .5994  .0027  .4970  .7018  .4907  .7081  .4942  .6962  .4962  .6979
    1    6  .5125  .0016  .4334  .5915  .4296  .5953  .4338  .5906  .4283  .5845
    2    1  .4647  .0031  .3552  .5743  .3512  .5783  .3585  .5743  .3592  .5736
    2    2  .4080  .0025  .3099  .5061  .3065  .5095  .3146  .5085  .3157  .5103
    2    3  .3355  .0017  .2544  .4165  .2518  .4191  .2597  .4207  .2611  .4209
    2    4  .3523  .0018  .2683  .4363  .2656  .4389  .2735  .4400  .2749  .4413
    2    5  .2951  .0014  .2229  .3674  .2205  .3698  .2283  .3721  .2300  .3705
    2    6  .3288  .0018  .2460  .4116  .2434  .4142  .2519  .4162  .2544  .4149
```

Anhang C

Simulationsergebnisse

C.1 Panik-Skala-Studie I

Tabelle C.1: Simulierte Niveaus für die Konfidenzintervalle zur Panik-Skala-Studie I bei einem vorgegebenen Niveau von 95%.

		Methode			
Woche	$\widehat{p}_{1s,n}$	Normal	t	Transf.	Perzentil
0	0,90	90,1%	91,9%	93,6%	90,1%
2	0,66	92,6%	94,4%	93,1%	92,1%
4	0,51	93,2%	95,0%	93,6%	92,5%
6	0,27	93,3%	94,9%	93,9%	92,6%
8	0,16	91,1%	92,7%	93,2%	91,3%

Tabelle C.2: Simulierte durchschnittliche Längen der Konfidenzintervalle zur Panik-Skala-Studie I bei einem vorgegebenen Niveau von 95%.

		Methode			
Woche	$\widehat{p}_{1s,n}$	Normal	t	Transf.	Perzentil
0	0,90	0,12	0,13	0,13	0,12
2	0,66	0,15	0,17	0,15	0,15
4	0,51	0,17	0,19	0,17	0,17
6	0,27	0,15	0,16	0,14	0,14
8	0,16	0,19	0,20	0,19	0,18

C.2 Panik-Skala-Studie II

Tabelle C.3: Simulierte Niveaus für die Konfidenzintervalle zur Panik-Skala-Studie II bei einem vorgegebenen Niveau von 95%.

			Methode	
Agoraphobie	Woche	$\widehat{p}_{is,n}$	Normal	Transformation
mit	0	0,88	93,8%	94,9%
	2	0,75	92,5%	93,7%
	4	0,53	94,1%	94,6%
	6	0,48	94,1%	94,9%
	8	0,37	93,2%	94,8%
ohne	0	0,65	93,9%	94,4%
	2	0,44	92,3%	93,3%
	4	0,37	92,7%	93,6%
	6	0,26	91,7%	93,4%
	8	0,28	92,9%	94,1%

Tabelle C.4: Simulierte durchschnittliche Längen der Konfidenzintervalle zur Panik-Skala-Studie II bei einem vorgegebenen Niveau von 95%.

			Methode	
Agoraphobie	Woche	$\widehat{p}_{is,n}$	Normal	Transformation
mit	0	0,88	0,10	0,10
	2	0,75	0,18	0,18
	4	0,53	0,20	0,19
	6	0,48	0,25	0,24
	8	0,37	0,29	0,28
ohne	0	0,65	0,16	0,16
	2	0,44	0,27	0,26
	4	0,37	0,22	0,22
	6	0,26	0,25	0,25
	8	0,28	0,24	0,24

C.3 Schulterschmerz-Studie

Tabelle C.5: Simulierte Niveaus für die Konfidenzintervalle zur Schulterschmerz-Studie bei einem vorgegebenen Niveau von 95%.

			Methode			
Behandlung	Zeitpunkt	$\widehat{p}_{is,n}$	Normal	t	Transf.	Perzentil
Y	1	0,46	93,6%	94,7%	94,3%	93,6%
	2	0,41	93,7%	94,5%	94,3%	93,8%
	3	0,34	91,6%	92,4%	92,2%	91,7%
	4	0,35	93,2%	93,9%	93,6%	93,3%
	5	0,30	92,2%	93,1%	92,8%	91,9%
	6	0,33	93,0%	93,8%	93,4%	93,0%
N	1	0,58	93,0%	94,4%	94,0%	93,0%
	2	0,70	93,0%	94,3%	94,1%	93,1%
	3	0,69	93,0%	94,3%	94,0%	93,2%
	4	0,73	92,2%	93,7%	93,7%	92,7%
	5	0,60	93,2%	94,4%	93,9%	93,2%
	6	0,51	93,8%	95,0%	94,4%	93,7%

Tabelle C.6: Simulierte durchschnittliche Längen der Konfidenzintervalle zur Schulterschmerz-Studie bei einem vorgegebenen Niveau von 95%.

			Methode			
Behandlung	Zeitpunkt	$\widehat{p}_{is,n}$	Normal	t	Transf.	Perzentil
Y	1	0,46	0,21	0,22	0,20	0,20
	2	0,41	0,19	0,19	0,18	0,18
	3	0,34	0,15	0,16	0,15	0,15
	4	0,35	0,16	0,16	0,16	0,15
	5	0,30	0,13	0,13	0,13	0,12
	6	0,33	0,15	0,15	0,15	0,14
N	1	0,58	0,25	0,27	0,25	0,25
	2	0,70	0,21	0,22	0,21	0,20
	3	0,69	0,22	0,23	0,21	0,21
	4	0,73	0,21	0,22	0,21	0,20
	5	0,60	0,24	0,25	0,23	0,23
	6	0,51	0,20	0,22	0,20	0,20

C.4 Wasser-Irrgarten-Test

Tabelle C.7: Simulierte Niveaus für die Konfidenzintervalle zum Wasser-Irrgarten-Test bei einem vorgegebenen Niveau von 95%.

			Methode	
Dosis	Tag	$\widehat{p}_{is,n}$	Normal	Transformation
Placebo	1	0,38	93,6%	94,3%
	7	0,67	93,4%	94,1%
1	1	0,48	93,7%	94,4%
	7	0,66	93,3%	93,9%
2	1	0,39	93,8%	94,5%
	7	0,62	93,3%	94,2%
3	1	0,28	93,6%	94,4%
	7	0,53	93,3%	94,0%

Tabelle C.8: Simulierte durchschnittliche Längen der Konfidenzintervalle zum Wasser-Irrgarten-Test bei einem vorgegebenen Niveau von 95%.

			Methode	
Dosis	Tag	$\widehat{p}_{is,n}$	Normal	Transformation
Placebo	1	0,38	0,19	0,19
	7	0,67	0,17	0,17
1	1	0,48	0,21	0,21
	7	0,66	0,18	0,18
2	1	0,39	0,19	0,19
	7	0,62	0,20	0,19
3	1	0,28	0,16	0,16
	7	0,53	0,22	0,21

Notation

Vektoren und Matrizen

Während wir für eindimensionale Größen Schriften normaler Stärke benutzen, verwenden wir für Vektoren und Matrizen Fettschriften. Für den zu einem Vektor $\boldsymbol{x}$ transponierten Vektor schreiben wir $\boldsymbol{x}'$, die Transponierte einer Matrix $\boldsymbol{A}$ bezeichnen wir mit $\boldsymbol{A}'$. Als Symbol für den aus n Einsen bestehenden Spaltenvektor benutzen wir $\mathbf{1}_n$.

Konventionen für das Rechnen mit 0 und ∞

Für jede Zahl $x \in \mathbb{R} \setminus \{0\}$ ordnen wir dem Bruch $x/0$ in Abhängigkeit von dem Vorzeichen von x den Wert $+\infty$ oder $-\infty$ zu. Für den Fall $x = 0$ gelte $x/0 = 0$. Für die Ausdrücke $0 \cdot \infty$ und $0 \cdot (-\infty)$ vereinbaren wir als Wert die Zahl 0.

Summen und Mittelwerte

Um Summen über die möglichen Werte von Indizes abzukürzen, ersetzen wir die entsprechenden Indizes durch das Symbol '$\cdot$'. Zur Darstellung von Mittelwerten verwenden wir zusätzlich Überstreichungen. Sei zum Beispiel

$$(x_{11}, \ldots, x_{1n}, \ldots, x_{m1}, \ldots, x_{mn})'$$

ein (mn)-dimensionaler reeller Vektor. Für die Summen und die Mittelwerte

$$\sum_{j=1}^{n} x_{ij} \quad \text{und} \quad \frac{1}{n}\sum_{j=1}^{n} x_{ij}, \quad i = 1, \ldots, n,$$

schreiben wir dann $x_{i\cdot}$ und $\overline{x}_{i\cdot}$.

Spezielle Funktionen

Die Indikatorfunktion einer Menge A bezeichnen wir mit 1_A. Die Zählfunktion $c : \mathbb{R} \to \mathbb{R}$ definieren wir durch

$$c(x) = \begin{cases} 0 & : x < 0 \\ \frac{1}{2} & : x = 0 \\ 1 & : x > 0 \end{cases} .$$

Wahrscheinlichkeitsmaße und Verteilungsfunktionen

Wahrscheinlichkeitsmaße bezeichnen wir im Allgemeinen mit dem Buchstaben P. Wenn wir von einer Verteilungsfunktion sprechen, so meinen wir, wenn wir nicht ausdrücklich etwas anderes sagen, stets die normalisierte Version dieser Verteilungsfunktion. Für die linksstetige Version einer Verteilungsfunktion F schreiben wir F^- und für die rechtsstetige Version F^+.

Spezielle Wahrscheinlichkeitsverteilungen

Für die univariate Normalverteilung mit Erwartungswert μ und Varianz σ^2 schreiben wir kurz $\mathrm{N}(\mu, \sigma^2)$. Die Verteilungsfunktion der Standardnormalverteilung bezeichnen wir mit Φ und das α-Quantil dieser Verteilung bezeichnen wir mit z_α. Als Symbol für die multivariate Normalverteilung mit Erwartungswert $\boldsymbol{\mu}$ und Kovarianzmatrix $\boldsymbol{\Sigma}$ verwenden wir $\mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Die χ^2-Verteilung mit f Freiheitsgraden bezeichnen wir mit χ_f^2 und die t-Verteilung mit der gleichen Anzahl an Freiheitsgraden mit t_f. Als Schreibweise für das α-Quantil dieser t-Verteilung benutzen wir $t_{f,\alpha}$.

Zufallsvariablen

Als Symbole für Zufallsvariablen wählen wir Großbuchstaben. Hat eine Zufallsvariable X die Verteilungsfunktion F oder die Verteilung P, so schreiben wir dafür kurz

$$X \sim F \quad \text{oder} \quad X \sim P\,.$$

Haben zwei Zufallsvariablen X und Y die gleiche Verteilung, so verwenden als analoge Schreibweise $X \sim Y$.

Den Erwartungswert einer Zufallsvariablen X bezeichnen wir mit $\mathrm{E}(X)$, die Varianz mit $\mathrm{Var}(X)$. Als Symbol für die Kovarianzmatrix eines Zufallsvektors $\boldsymbol{X}$ schreiben wir $\mathrm{Cov}(\boldsymbol{X})$. Die bedingte Erwartung eines Zufallsvektors $\boldsymbol{X}$ bei gegebenem $\boldsymbol{Y}$ stellen wir durch das Symbol $\mathrm{E}(\boldsymbol{X}|\boldsymbol{Y})$ dar.

Zufallsvariablen, die als Schätzer für ein statistisches Funktional dienen, sind häufig durch den mathematischen Akzent '$\widehat{\ }$', von Zeit zu Zeit aber auch durch '$\check{\ }$' oder '$\widetilde{\ }$' gekennzeichnet.

Konvergenzbegriffe der Wahrscheinlichkeitstheorie

Konvergiert eine Folge von Zufallsvariablen X_n, $n \in \mathbb{N}$, in Verteilung gegen eine Zufallsvariable $X \sim P$, so stellen wir dies symbolisch durch

$$X_n \xrightarrow{\mathcal{D}} X \quad \text{oder} \quad X_n \xrightarrow{\mathcal{D}} P$$

dar. Konvergiert X_n stochastisch oder fast sicher gegen X, so schreiben wir dafür kurz

$$X_n \xrightarrow{\mathrm{P}} X \quad \text{oder} \quad X_n \xrightarrow{\text{f.s.}} X\,.$$

Für eine Folge von Wahrscheinlichkeitsmaßen P_n und ein Wahrscheinlichkeitsmaß P bedeutet der Ausdruck $P_n \xrightarrow{\mathrm{w}} P$, dass P_n schwach gegen P konvergiert.

Literaturverzeichnis

Agresti, A. (1983). Testing marginal homogeneity for ordinal categorical variables. *Biometrics*, 39:505–510.

Akritas, M. G. und Arnold, S. F. (1994). Fully nonparametric hypotheses for factorial designs I: Multivariate repeated measures designs. *Journal of the American Statistical Association*, 89:336–343.

Akritas, M. G., Arnold, S. F. und Brunner, E. (1997). Nonparametric hypotheses and rank statistics for unbalanced factorial designs. *Journal of the American Statistical Association*, 92:258–265.

Akritas, M. G. und Brunner, E. (1997). A unified approach to rank tests for mixed models. *Journal of Statistical Planning and Inference*, 61:249–279.

Bandelow, B. (1997). *Panik- und Agoraphobieskala (PAS).* Hogrefe, Göttingen.

Bandelow, B., Brunner, E., Broocks, A., Beinroth, D., Hajak, G., Pralle, L. und Rüther, E. (1998). The use of the panic and agoraphobia scale in a clinical trial. *Psychiatry Research*, 77:43–49.

Beder, J. H. und Heim, R. C. (1990). On the use of ridit analysis. *Psychometrika*, 55:603–616.

Beder, J. H. und Heim, R. C. (1992). Errata: On the use of ridit analysis. *Psychometrika*, 57:160.

Berkson, J. (1939). Some difficulties of interpretation encountered in the application of the chi-square test. *Journal of the American Statistical Association*, 33:526–536.

Berkson, J. (1942). Tests of significance as evidence. *Journal of the American Statistical Association*, 37:325–335.

Berkson, J. (1943). Experience with tests of significance: A reply to Professor R. A. Fisher. *Journal of the American Statistical Association*, 38:242–246.

Berkson, J. und Geary, R. C. (1941). Comments on Dr. Madow's "Note on Tests of Departure from Normality" with some remarks concerning tests of significance. *Journal of the American Statistical Association*, 36:539–543.

Bhattacharyya, G. K. und Johnson, R. A. (1974). Estimation of reliability in a multicomponent stress-strength model. *Journal of the American Statistical Association*, 69:966–970.

Billingsley, P. (1995). *Probability and Measure.* John Wiley & Sons, New York, Third Edition.

Birnbaum, Z. W. (1956). On a use of the Mann-Whitney statistic. In Neyman, J., editor, *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, Volume I, 13–17. University of California Press, Berkeley.

Birnbaum, Z. W. und McCarty, R. C. (1958). A distribution-free upper confidence bound for $\Pr\{Y < X\}$, based on independent samples of X and Y. *The Annals of Mathematical Statistics*, 29:558–562.

Box, G. E. P. (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Effect of inequality of variance in the one-way classification. *The Annals of Mathematical Statistics*, 25:290–302.

Bross, I. D. J. (1958). How to use ridit analysis. *Biometrics*, 14:18–38.

Bross, I. D. J. (1979). Reply by Dr. Bross: On biases in judging statistical methods. *American Journal of Epidemiology*, 109:30–32.

Brunner, E., Dette, H. und Munk, A. (1997). Box-type approximations in nonparametric factorial designs. *Journal of the American Statistical Association*, 92:1494–1502.

Brunner, E. und Munzel, U. (2000). The nonparametric Behrens-Fisher problem: Asymptotic theory and a small-sample approximation. *Biometrical Journal*, 42:1–9.

Brunner, E., Munzel, U. und Puri, M. L. (1999). Rank-score tests in factorial designs with repeated measures. *Journal of Multivariate Analysis*, 70:286–317.

Chandra, S. und Owen, D. B. (1975). On estimating the reliability of a component subject to several different stresses (strengths). *Naval Research Logistics Quarterly*, 22:31–39.

Chen, M. und Kianifard, F. (2000). A nonparametric procedure associated with a clinically meaningful efficacy measure. *Biostatistics*, 1:293–298.

Cheng, K. F. und Chao, A. (1984). Confidence intervals for reliability from stress-strength relationships. *IEEE Transactions on Reliabilty*, 33:246–249.

Chernick, M. R. (1999). *Bootstrap Methods.* John Wiley & Sons, New York.

Church, J. D. und Harris, B. (1970). The estimation of reliability from stress-strength relationships. *Technometrics*, 12:49–54.

Cramér, H. (1928). On the composition of elementary errors. *Skandinavisk Aktuarietidskrift*, 11:13–74 and 141–180.

Davison, A. C. und Hinkley, D. V. (1997). *Bootstrap Methods and their Applications.* Cambridge University Press, Cambridge.

Deuchler, G. (1914). Über die Methoden der Korrelationsrechnung in der Pädagogik und Psychologie. *Zeitschrift für Pädagogische Psychologie und Experimentelle Pädagogik*, 15:114–131, 145–159 und 229–242.

Dodge, Y. (1985). *Analysis of Experiments with Missing Data.* John Wiley & Sons, New York.

Domhof, S. (1999). Rangverfahren mit unbeschränkten Scorefunktionen in faktoriellen Versuchsplänen. Diplomarbeit, Institut für Mathematische Stochastik, Universität Göttingen, Göttingen.

Downton, F. (1973). The estimation of $\Pr(Y < X)$ in the normal case. *Technometrics*, 15:551–558.

Dudley, R. M. (1989). *Real Analysis and Probability.* Wadsworth & Brooks/Cole, Pacific Grove.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7:1–26.

Efron, B. und Tibshirani, R. J. (1993). *An Introduction to the Bootstrap.* Chapman & Hall, New York.

Enis, P. und Geisser, S. (1971). Estimation of the probability that $Y < X$. *Journal of the American Statistical Association*, 66:162–168.

Forster, O. (1983). *Analysis 1.* Friedrich Vieweg & Sohn, Braunschweig, Vierte Auflage.

Freedman, D., Pisani, R. und Purves, P. (1998). *Statistics.* W. W. Norton & Company, New York, Third Edition.

Govindarajulu, Z. (1967). Two-sided confidence limits for $P(X < Y)$ based on normal samples of X and Y. *Sankhyā: Series B*, 29:35–40.

Govindarajulu, Z. (1968). Distribution-free confidence bounds for $P(X < Y)$. *The Annals of the Institute of Statistical Mathematics*, 20:229–238.

Govindarajulu, Z. (1975). Robustness of Mann-Whitney-Wilcoxon test to dependence in the variables. *Studia Scientiarium Mathematicarum Hungarica*, 10:39–45.

Halperin, M., Gilbert, P. R. und Lachin, J. M. (1987). Distribution-free confidence intervals for $\Pr(X_1 < X_2)$. *Biometrics*, 43:71–90.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. und Stahel, W. A. (1986). *Robust Statistics.* John Wiley & Sons, New York.

Harris, B. und Soms, A. P. (1983). A note on a difficulty in estimating reliability from stress-strength relationships. *Naval Research Logistics Quarterly*, 30:659–663.

Hauck, W. W., Hyslop, T. und Anderson, S. (2000). Generalized treatment effects for clinical trials. *Statistics in Medicine*, 19:887–899.

Haunsperger, D. B. (1992). Dictionaries of paradoxes for statistical tests on k samples. *Journal of the American Statistical Association*, 87:149–155.

Hettmansperger, T. P. und McKean, J. W. (1998). *Robust Nonparametric Statistical Methods*. Arnold, London.

Hoffmann-Jørgensen, J. (1991). *Stochastic Processes on Polish Spaces*. Aarhus Universitet, Aarhus.

Hollander, M., Pledger, G. und Lin, P.-E. (1974). Robustness of the Wilcoxon test to a certain dependency between samples. *The Annals of Statistics*, 2:177–181.

Johnson, N. L. (1975). Letter to the editor. *Technometrics*, 17:393.

Kelley, G. D., Kelley, J. A. und Schucany, W. R. (1976). Efficient estimation of $P(Y < X)$ in the exponential case. *Technometrics*, 18:359–360.

Kiefer, J. (1959). K-sample analogues of the Kolmogorov-Smirnov and Cramer-V. Mises tests. *The Annals of Mathematical Statistics*, 30:420–447.

Kolmogorov, A. (1933). Sulla determinatione empirica di una legge di distributione. *Giornale dell'Istituto Italiano degli Attuari*, 4:83–91.

Krengel, U. (1991). *Einführung in die Wahrscheinlichkeitstheorie und Statistik*. Friedrich Vieweg und Sohn, Braunschweig, Dritte Auflage.

Kruskal, W. H. (1952). A nonparametric test for the several sample problem. *The Annals of Mathematical Statistics*, 23:525–540.

Kruskal, W. H. (1957). Historical notes on the Wilcoxon unpaired two-sample test. *Journal of the American Statistical Association*, 52:356–360.

Kruskal, W. H. und Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47:583–621.

Kruskal, W. H. und Wallis, W. A. (1953). Errata: Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 48:907–911.

Kulle, B. (1999). Nichtparametrisches Behrens-Fisher Problem im Mehr-Stichprobenfall. Diplomarbeit, Institut für Mathematische Stochastik, Universität Göttingen, Göttingen.

Lachin, J. M. (1992). Some large-sample distribution-free estimators and tests for multivariate partially incomplete data from two populations. *Statistics in Medicine*, 11:1151–1170.

Lachin, J. M. (1996). Distribution-free marginal analysis of repeated measures. *Drug Information Journal*, 30:1017–1028.

Langer, F. (1998). *Berücksichtigung von Kovariablen im nichtparametrischen gemischten Modell.* Dissertation, Universität Göttingen, Göttingen.

Lehmann, E. L. (1950). *Notes on the Theory of Estimation.* University of California, Berkeley. Recorded by Colin Blyth.

Lehmann, E. L. (1951). Consistency and unbiasedness of certain nonparametric tests. *The Annals of Mathematical Statistics*, 22:165–179.

Lehmann, E. L. (1975). *Nonparametrics.* Holden-Day, San Francisco.

Lévy, P. (1925). *Calcul des Probabilités.* Gauthier-Villars, Paris.

Little, R. J. A. und Rubin, D. B. (1987). *Statistical Analysis with Missing Data.* John Wiley & Sons, New York.

Lumley, T. (1996). Generalized estimation equation equations for ordinal data: A note on working correlation structures. *Biometrics*, 52:354–361.

Mann, H. B. und Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 17:50–60.

Mantel, N. (1963). Chi-square tests with one degree of freedom: Extensions of the Mantel-Haenszel procedure. *Journal of the American Statistical Association*, 58:690–670.

Mantel, N. (1979). Ridits analysis and related ranking procedures – use at your own risk. *American Journal of Epidemiology*, 109:25–29.

Marden, J. I. und Muyot, M. E. T. (1995). Rank tests for main and interaction effects in analysis of variance. *Journal of the American Statistical Association*, 90:1388–1398.

Mazumdar, M. (1970). Some estimates of reliability using interference theory. *Naval Research Logistics Quarterly*, 17:159–165.

McCullagh, P. und Nelder, J. A. (1989). *Generalized Linear Models.* Chapman & Hall, London, Second Edition.

McGraw, K. O. und Wong, S. P. (1992). A common language effect size statistic. *Psychological Bulletin*, 1992:361–365.

Mee, R. W. (1990). Confidence intervals for probabilities and tolerance regions based on a generalization of the Mann-Whitney statistics. *Journal of the American Statistical Association*, 85:793–800.

Munzel, U. (1999). Nonparametric methods for paired samples. *Statistica Neerlandica*, 53:277–286.

Ottaviani, G. (1939). Sulla probabilità che una prova su due variabili casuali X e Y verifichi la disuguaglianza $X < Y$ e sul corrispondente scarto quadratico medio. *Giornale dell'Istituto Italiano degli Attuari*, 10:186–192.

Owen, D. B., Craswell, K. J. und Hanson, D. L. (1964). Nonparametric upper confidence bounds for $\Pr\{Y < X\}$ and confidence limits for $\Pr\{Y < X\}$ when X and Y are normal. *Journal of the American Statistical Association*, 59:906–924.

Patel, K. M. und Hoel, D. G. (1973). A nonparametric test for interaction in factorial experiments. *Journal of the American Statistical Association*, 68:615–620.

Patnaik, P. B. (1949). The non-central χ^2 and F-distributions and their applications. *Biometrika*, 36:202–232.

Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science, Series 5*, 50:157–172.

Prokhorov, Y. V. (1956). Convergence of random processes and limit theorems in probability theory. *Theory of Probability and its Applications*, I:157–214.

Putter, J. (1955). The treatment of ties in some nonparametric tests. *The Annals of Mathematical Statistics*, 26:368–386.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications.* John Wiley & Sons, New York, Second Edition.

Reiser, B. und Guttman, I. (1986). Statistical inference for $\Pr(Y < X)$: the normal case. *Technometrics*, 28:253–257.

Rothman, K. J. und Greenland, S. (1998). *Modern Epidemiology.* Lippincott-Raven, Philadelphia, Second Edition.

Rousseeuw, P. J. und Molenberghs, G. (1993). Transformation of non positive semidefinite correlation matrices. *Communications in Statistics – Theory and Methods*, A22:965–984.

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63:581–590.

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys.* John Wiley & Sons, New York.

Saari, D. G. (1989). A dictionary for voting paradoxes. *Journal of Economic Theory*, 48:443–475.

Saari, D. G. (1990). The Borda dictionary. *Social Choice and Welfare*, 7:279–317.

Sathe, Y. S. und Shah, S. P. (1981). On estimating $P(X > Y)$ for the exponential distribution. *Communications in Statistics – Theory and Methods*, A10:39–47.

Satterthwaite, F. E. (1941). Synthesis of variance. *Psychometrika*, 6:309–316.

Savage, I. R. (1957). Nonparametric statistics. *Journal of The American Statistical Association*, 52:331–344.

Schafer, J. L. (1997). *Analysis of Incomplete Multivariate Data.* Chapman & Hall, Boca Raton.

Schwartz, L. und Wearden, S. (1959). A distribution-free asymptotic mathod of estimating, testing, and setting confidence limits for hertability. *Biometrics*, 15:227–235.

Selvin, S. (1979). Reply by Dr. Selvin. *American Journal of Epidemiology*, 109:29–30.

Sen, P. K. (1967). A note on asymptotically distribution-free confidence bounds for $P(X < Y)$, based on two independent samples. *Sankhyā: Series A*, 29:95–102.

Senn, S. (1997). Letter to the editor. *Statitics in Medicine*, 16:1301–1305.

Shao, J. (1999). *Mathematical Statistics.* Springer, New York.

Siemer, A. (1999). Die Berücksichtigung von heterogen verteilten Kovariablen in einem nichtparametrischen Modell. Diplomarbeit, Institut für Mathematische Stochastik, Universität Göttingen, Göttingen.

Simonoff, J. S., Hochberg, Y. und Reiser, B. (1986). Alternative estimation procedures for $\Pr(X < Y)$ in categorized data. *Biometrics*, 42:895–907.

Smirnov, N. (1939a). On the estimation of the discrepancy between empirical curves of distribution for two independent samples. *Bulletin de l'Université de Moscou, Série Internationale (Mathématiques)*, 2:3–16.

Smirnov, N. (1939b). Sur les éscarts de la courbe de distribution empirique. *Recueil Mathématique de la Societe Mathématique de Moscou*, 6:3–26.

Smith, H. F. (1936). The problem of comparing the results of two experiments with unequal errors. *Journal of the Council for Scientific and Industrial Research (Australia)*, 9:211–212.

Steland, A. (1996). *Bootstrapping linearer Rangstatistiken mit Anwendungen in nichtparametrischen Modellen.* Dissertation, Universität Göttingen, Göttingen.

Thall, P. F. und Lachin, J. M. (1988). Analysis of recurrent events: Nonparametric methods for random interval count data. *Journal of the American Statitical Association*, 83:339–347.

Tong, H. (1974). A note on the estimation of $\Pr(Y < X)$ in the exponential case. *Technometrics*, 16:625.

Tong, H. (1975a). Errata: A note on the estimation of $\Pr(Y < X)$ in the exponential case. *Technometrics*, 17:395.

Tong, H. (1975b). Letter to the editor. *Technometrics*, 17:393.

Ury, H. K. (1972). On distribution-free confidence bounds for $\Pr(Y < X)$. *Technometrics*, 14:577–581.

Ury, H. K. und Wiggins, A. D. (1979). Distribution-free confidence bounds for $\Pr(Y < X)$ when $F(x)$ and $G(y) = F(x - \theta)$ are continuous and symmetric. *Communications in Statistics – Theory and Methods*, A8:1247–1253.

van Dantzig, D. (1951). On the consistency and the power of Wilcoxon's two sample test. *Indigationes Mathematicae*, 13:1–8.

van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press, Cambridge.

van der Vaart, A. W. und Wellner, J. A. (1996). *Weak Convergence and Empirical Processes*. Springer, New York.

Vargha, A. und Delaney, H. D. (2000). A critique and improvement of the CL common language effect size statistic of McGraw and Wong. *Journal of Educational and Behavioral Statistics*, 25:101–132.

von Mises, R. (1931). *Vorlesungen aus dem Gebiete der angewandten Mathematik*, Band 1. Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik. Franz Deuticke, Leipzig.

Wald, A. und Wolfowitz, J. (1940). On a test whether two samples are from the same distribution. *The Annals of Mathematical Statistics*, 11:147–162.

Welch, B. L. (1936). The specification of rules for rejecting too variable a product, with particular reference to an electric lamp problem. *Supplement to the Journal of the Royal Statistical Society*, 3:29–48.

Welch, B. L. (1938). The significance of the difference between two means when the population variances are unequal. *Biometrika*, 29:350–362.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, 1:80–83.

Williams, D. (1991). *Probability with Martingales*. Cambridge University Press, Cambridge.

Wolfe, D. A. und Hogg, R. V. (1971). On constructing statistics and reporting data. *The American Statistician*, 25:27–30.

Woodward, W. A. und Gray, H. L. (1975). Minimum variance unbiased estimation in the gamma distribution. *Communications in Statistics*, 4:907–922.

Lebenslauf

Am 8. Juni 1974 wurde ich in Schwerte an der Ruhr als Sohn von Ursula Domhof, geb. Buchmeier, und Winfried Domhof geboren.

Von 1980 bis 1984 besuchte ich die Fröbelschule (Grundschule) in Peine, im Anschluss daran bis 1986 die Orientierungsstufe am Schlossberg. Meine schulische Ausbildung beendete ich 1993 nach dem siebenjährigen Besuch des Peiner Ratsgymnasiums mit dem Erwerb der Allgemeinen Hochschulreife. Vom Juli 1993 bis Juni 1994 leistete ich den Grundwehrdienst bei der Bundeswehr in Lüneburg.

Zum Wintersemester 1994/95 nahm ich das Studium der Mathematik mit dem Nebenfach Informatik an der Georg-August-Universität Göttingen auf. Die Diplom-Vorprüfung legte ich im Oktober 1996 ab. Während des Sommersemesters 1997 war ich als studentische Hilfskraft am Institut für Numerische und Angewandte Mathematik beschäftigt. Im anschließenden Wintersemester nahm ich meine Tätigkeit als studentische Hilfskraft in der Abteilung Medizinische Statistik auf, die ich bis zum Abschluss meines Studiums im Oktober 1999 fortführte. Während dieser Zeit fertigte ich unter Anleitung von Prof. Dr. Edgar Brunner die Diplomarbeit mit dem Titel "Rangverfahren mit unbeschränkten Scorefunktionen in faktoriellen Versuchsplänen" an.

Seit November 1999 bin ich als wissenschaftlicher Mitarbeiter im Rahmen des DFG-Forschungsprojekts "Nichtparametrische Modelle und Verfahren für longitudinale Daten" an der Abteilung Medizinische Statistik angestellt. Im Juli 2000 hatte ich Gelegenheit zu einem Forschungsaufenthalt an der Indiana University, Bloomington. Die vorliegende Dissertation habe ich während dieses Forschungsaufenthalts und der Zeit als wissenschaftlicher Mitarbeiter in der Abteilung Medizinische Statistik erstellt.