

DAS NICHTPARAMETRISCHE
BEHRENS-FISHER-PROBLEM MIT
APPROXIMATION NACH ATIQULLAH

Bachelor-Thesis an der Fakultät für
Mathematik und Informatik
Georg-August-Universität Göttingen

vorgelegt von:

Markus Harden

Erstgutachter:

Prof. Dr. Edgar Brunner

Zweitgutachter:

Prof. Dr. Tatyana Krivobokova

Danksagung

An dieser Stelle möchte ich mich bei den Menschen bedanken, ohne deren Unterstützung diese Arbeit nicht gelungen wäre.

Mein besonderer Dank gilt Herrn Prof. Dr. Edgar Brunner für die Vergabe des Themas und die hervorragende, überaus intensive fachliche Betreuung. Ich danke Frau Prof. Dr. Tatyana Krivobokova für die Übernahme des Co-Referats.

Dank schulde ich der Abteilung für Medizinische Statistik, da ich in einer freundlichen, vertrauensvollen Atmosphäre arbeiten durfte. Insbesondere danke ich Herrn Frank Konietschke für seine wertvolle Unterstützung.

Meinem guten Freund und Leidensgenossen Tobias Lohse danke ich an dieser Stelle ebenfalls. Die gemeinsame Arbeit an unserem Bachelorabschluss wird mir in Erinnerung bleiben. Seine besondere Disziplin und Zielstrebigkeit haben mich stets motiviert.

Zu guter Letzt bedanke ich mich bei meinen Eltern und meiner Schwester für ihre Anteilnahme und moralische Unterstützung, die mich während meines Studiums immer begleitet haben.

Göttingen, den 30.09.2009

Markus Harden

Inhaltsverzeichnis

| | |
|---|-----------|
| Einleitung | 1 |
| 1 Parametrische Modelle | 2 |
| 1.1 Normalverteilte Daten | 2 |
| 1.1.1 Gauss-Test | 2 |
| 1.1.2 t-Test | 3 |
| 1.1.3 Das klassische Behrens-Fisher-Problem | 4 |
| 1.2 Beispiel: Ferritin bei Kleinwuchs | 5 |
| 2 Nichtparametrische Modelle | 6 |
| 2.1 Grundlagen der Nichtparametrik | 6 |
| 2.1.1 Der relative Effekt | 6 |
| 2.1.2 Empirische Verteilungsfunktionen | 8 |
| 2.1.3 Ränge | 9 |
| 2.2 Das nichtparametrische Behrens-Fisher-Problem | 12 |
| 2.2.1 Hypothesen | 12 |
| 2.2.2 Verteilung des relativen Effektes | 13 |
| 2.2.3 t-Approximation | 16 |
| 3 Simulationen | 22 |
| 3.1 Anpassung der Schätzer | 22 |
| 3.1.1 Berechnung der Varianzen | 22 |
| 3.1.2 Schätzung des Freiheitsgrades | 24 |
| 3.2 Simulationen für Skalenalternativen | 25 |
| 3.3 Simulationen für beliebige Verteilungen | 26 |
| 3.3.1 Monte-Carlo-Simulationen | 26 |
| 3.3.2 Simulationen für schiefe Verteilungen | 26 |
| 3.4 Beispiel: Ferritin bei Kleinwuchs | 27 |
| 4 Zusammenfassung | 28 |
| A Notation und Mathematische Resultate | I |
| A.1 Matrizen und Vektoren | I |

| | |
|-----------------------------------|--------------|
| <i>INHALTSVERZEICHNIS</i> | v |
| A.2 Quadratische Formen | II |
| A.3 Ungleichungen | III |
| A.4 Konvergenzarten | III |
| B Tabellen und Abbildungen | V |
| C SAS-Code | XV |
| Literatur | XXIII |

Einleitung

Medizinische Studien haben zum Ziel, die Wirkungsweise neuer Substanzen nachzuweisen. Um zu validen Ergebnissen zu gelangen, muss bereits während der Planung des Versuches ein statistisches Modell formuliert werden. Ein mögliches Modell ist das *unverbundene Zwei-Stichproben-Problem*. Es werden zwei Behandlungsstufen betrachtet, wobei häufig ein Vergleich zwischen der neuen Substanz und einem Placebo durchgeführt wird. Unverbunden bedeutet in diesem Zusammenhang, dass sich die Probanden gegenseitig nicht beeinflussen können und in der Auswertung keine Korrelationen berücksichtigt werden muss.

Klassische Auswertungsverfahren wie der t-Test können sensibel auf Verletzungen ihrer Modellannahmen reagieren. Daher sind häufig allgemeine *nichtparametrische Modelle* gefordert, die mit wenigen Anforderungen an die Daten auskommen. Das Behrens-Fisher-Problem beschäftigt sich mit der Auswertung von Daten mit ungleichen Stichprobenumfängen und Varianzen.

In dieser Arbeit soll das nichtparametrische Behrens-Fisher-Problem untersucht werden. Als Ausgangspunkt dient hier der Brunner-Munzel-Test, der die Testgröße durch eine t-Verteilung approximiert (vgl. [2]). Insbesondere sollen Verteilungen mit betragsmäßig großer Kurtosis (Wölbung einer Verteilung) untersucht werden. Da der Brunner-Munzel-Test diese Größe völlig außer acht lässt, soll überprüft werden, ob mit Resultaten von Atiqullah eine Verbesserung des Tests durchgeführt werden kann. Dafür wird das entwickelte Testverfahren im Anschluss mit Hilfe von simulierten Zufallsdaten verschiedener Verteilungen getestet.

1 Parametrische Modelle

Parametrische Modelle basieren im Allgemeinen auf der Annahme, dass den gemessenen Daten eine Verteilung zugrunde liegt, die durch eine endliche Anzahl von Parametern eindeutig beschrieben werden kann. Diese Parameter lassen sich aus den Beobachtungen schätzen, wodurch die Möglichkeit entsteht, verschiedene Behandlungen miteinander zu vergleichen. Der Behandlungseffekt wird vor der Auswertung durch eine Hypothese (Nullhypothese) formuliert, deren Gültigkeit im Anschluss mit statistischen Methoden überprüft wird.

Das Zwei-Stichprobenproblem beschränkt sich dabei auf zwei Behandlungsstufen, wobei der Begriff *Behandlung* für eine Einteilung des Beobachtungskollektivs in verschiedene Gruppen steht. Um einen Unterschied zwischen den Behandlungsstufen aufzudecken, bildet man eine Teststatistik, die mit einer aus den Daten rekonstruierten Verteilungsfunktion verglichen wird. Liegt die Teststatistik außerhalb eines vorher festgelegten Annahmebereiches, wird die anfangs aufgestellte Hypothese abgelehnt. Dies geschieht unabhängig davon, ob die gemessenen Daten die Realität wiedergeben oder nicht.

1.1 Normalverteilte Daten

Seien $\mathbf{X}_1 = (X_{11}, \dots, X_{1n_1})'$ und $\mathbf{X}_2 = (X_{21}, \dots, X_{2n_2})'$ zwei Beobachtungsvektoren unabhängig identisch normalverteilter Daten $X_{ik} \sim N(\mu_i, \sigma_i^2)$; $i = 1, 2$; $k = 1, \dots, n_i$. Es ist zu untersuchen, ob die Erwartungswerte der beiden Verteilungen gleich sind oder nicht. Die Nullhypothese wird als $H_0^\mu : \mu_1 = \mu_2$ formuliert. Zu verschiedenen Konstellationen der Parameter μ_i und σ_i^2 gibt es unterschiedliche Testverfahren, die im Folgenden kurz vorgestellt werden.

1.1.1 Gauss-Test

Der Gauss-Test ist ein exaktes Testverfahren, das die Mittelwerte beider Stichproben, unter der *Annahme* gleicher Erwartungswerte, miteinander vergleicht. Da die Varianzen als bekannt vorausgesetzt werden, ist dieser Test in der Praxis nicht durchführbar. Er eignet sich aber, um zufällig erzeugte

Daten mit anderen Auswertungsmethoden zu vergleichen. Der Parameter μ_i kann erwartungstreu durch den Mittelwert

$$\hat{\mu}_i = \bar{X}_i = \frac{1}{n_i} \sum_{k=1}^{n_i} X_{ik}$$

geschätzt werden. Anschließend standardisiert man $\hat{\mu}_1 - \hat{\mu}_2$ und erhält als Teststatistik

$$U_n = \frac{(\hat{\mu}_1 - \hat{\mu}_2) - (\mu_1 - \mu_2)}{\sqrt{\text{Var}(\hat{\mu}_1 - \hat{\mu}_2)}} \stackrel{H_0^\mu}{=} \frac{\hat{\mu}_1 - \hat{\mu}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1).$$

Falls die Testgröße U_n nun außerhalb eines vorher festgelegten Intervalls liegt, wird die vorher getroffene Annahme gleicher Erwartungswerte verworfen.

1.1.2 t-Test

Wenn den beiden Stichproben die Verteilung $N(\mu_i, \sigma^2)$ zugrunde liegt, die Varianz σ^2 aber unbekannt ist, so verwendet man den t-Test, um die Hypothese $H_0^\mu : \mu_1 = \mu_2$ zu überprüfen. Eine erwartungstreue Schätzung von σ^2 kann durch die gepoolte empirische Varianz S_0^2 erfolgen:

$$S_0^2 = \frac{1}{n_1 + n_2 - 2} \sum_{i=1}^2 \sum_{k=1}^{n_i} (X_{ik} - \bar{X}_i)^2 \sim \frac{\sigma^2}{n_1 + n_2 - 2} \cdot \chi_{n_1+n_2-2}^2.$$

Die zu beurteilende Teststatistik setzt sich aus zwei Teilen zusammen. Der Zähler Z_n besteht aus der standardisierten Differenz der Mittelwerte, der Nenner N_n aus einer standardisierten Form des Varianzschätzers

$$Z_n = \frac{\hat{\mu}_1 - \hat{\mu}_2}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad N_n = \sqrt{\frac{1}{\sigma^2} S_0^2}.$$

Die Unabhängigkeit von Z_n und N_n kann mit Hilfe des Satzes von Craig-Sakamoto gezeigt werden (s. Anhang). Damit erfüllt die Teststatistik die

Definition einer t-Verteilung und hat unter Gültigkeit von H_0^μ die Gestalt

$$\frac{Z_n}{N_n} = \frac{1}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \frac{\hat{\mu}_1 - \hat{\mu}_2}{S_0} \quad \sim t_{n_1+n_2-2}.$$

1.1.3 Das klassische Behrens-Fisher-Problem

Unter dem klassischen Behrens-Fisher-Problem versteht man die Aufgabe, Erwartungswerte normalverteilter Daten mit ungleichen Varianzen zu vergleichen. Dazu betrachtet man unabhängig, identisch verteilte Zufallsvariablen $X_{ik} \sim N(\mu_i, \sigma_i^2)$, $i = 1, 2$; $k = 1, \dots, n_i$, und standardisiert die Statistik $\hat{\mu}_1 - \hat{\mu}_2$ wie in Abschnitt 1.1.1.

Analog zum Gauss-Test liefert die Standardisierung für bekannte Varianzen eine standardnormalverteilte Testgröße:

$$U_n \stackrel{H_0^\mu}{=} \frac{\hat{\mu}_1 - \hat{\mu}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad \sim N(0, 1).$$

Da die Varianzen unbekannt sind, müssen diese aus den Beobachtungen geschätzt werden. Einen Ansatz zur Lösung dieses Problems lieferte B. L. Welch (vgl. [7]). Der Ausdruck $S_1^2/n_1 + S_2^2/n_2$ wird durch eine gestreckte χ_f^2 -Verteilung approximiert, deren Freiheitsgrad durch eine Box-Approximation (vgl. [1]) berechnet wird:

$$f = \frac{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)^2}{\frac{\sigma_1^2/n_1}{n_1-1} + \frac{\sigma_2^2/n_2}{n_2-1}}.$$

Die Teststatistik hat die Gestalt

$$B_n \stackrel{H_0^\mu}{=} \frac{\hat{\mu}_1 - \hat{\mu}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} \quad \overset{\sim}{\simeq} t_f,$$

wobei

$$S_i^2 = \frac{1}{n_i - 1} \sum_{k=1}^{n_i} (X_{ik} - \bar{X}_{i\cdot})^2 \quad \sim \sigma_i^2 \cdot \chi_{n_i-1}^2$$

ein erwartungstreuer Schätzer für die Varianz der i-ten Behandlungsstufe ist. Ersetzt man nun f durch \hat{f} mit

$$\hat{f} = \frac{(S_1^2/n_1 + S_2^2/n_2)^2}{\frac{S_1^2/n_1}{n_1-1} + \frac{S_2^2/n_2}{n_2-1}},$$

lassen sich die Quantile der t-Verteilung berechnen.

Das folgende Beispiel soll dazu dienen, die Notwendigkeit nichtparametrischer Modelle zu zeigen.

1.2 Beispiel: Ferritin bei Kleinwuchs

In diesem Beispiel wurden die Ferritin-Werte bei Kindern mit hormonell bedingtem Kleinwuchs gemessen. Die zu untersuchende Fragestellung lautet, ob Kinder mit erniedrigtem IGF-1 Wert zu erhöhten Ferritin-Werten neigen. Die gemessenen Ferritin-Werte sind in Abbildung 1 im Anhang dargestellt.

Die Boxplots sprechen sowohl gegen die Annahme gleicher Varianzen, als auch die normalverteilter Daten. Das hat zur Folge, dass *keines* der gerade vorgestellten Verfahren zum Vergleich der Versuchsgruppen angewendet werden kann. Dies motiviert uns, ein allgemeineres Modell zu formulieren, dass mit minimalen Voraussetzungen arbeitet und den Vergleich zwischen verschiedenen Klassen von Verteilungsfunktionen ermöglicht.

2 Nichtparametrische Modelle

Nichtparametrische Modelle sind dann gefordert, wenn keine Informationen über die Verteilungen der Beobachtungen vorliegen. Als Ausgangspunkt dieser Arbeit dient der Brunner-Munzel-Test (vgl. [2]), der eine t-Approximation für kleine Stichproben vorgibt.

Die Grundlage für dieses Testverfahren stellt der relative Effekt dar, der aus den Beobachtungen geschätzt werden kann. Die Verteilung des Schätzers soll unter der Hypothese $H_0^p : p = \frac{1}{2}$ betrachtet werden. Wichtige Eigenschaften und Resultate werden in den folgenden Abschnitten dargestellt.

2.1 Grundlagen der Nichtparametrik

Um einen nichtparametrischen Test zu formulieren, benötigt man eine Hypothese, die für verschiedene Klassen von Verteilungen getestet werden kann. Beispielsweise kann die Hypothese $H_0 : \mu_1 = \mu_2$ für Cauchy-verteilte Zufallsvariablen nicht verifiziert werden, da die Erwartungswerte, unabhängig von der Wahl der Parameter, nicht existieren. Die folgende Darstellung des relativen Effektes und anderer Werkzeuge orientiert sich an dem Buch "Nichtparametrische Datenanalyse" (vgl. [6], Springer 2002). Da in dieser Arbeit das unverbundene Zwei-Stichprobenproblem behandelt wird, werden fortan ausschließlich einfaktorielle Modelle mit zwei Faktorstufen betrachtet.

Im Folgenden seien $X_{ik} \sim F_i$ unabhängig identisch verteilte Zufallsvariablen, die in den Vektoren $\mathbf{X}_i = (X_{i1}, \dots, X_{in_i}), i = 1, 2; k = 1, \dots, n_i$, angeordnet werden. Erläuterungen zur Notation können dem Anhang entnommen werden.

2.1.1 Der relative Effekt

Der relative Effekt p ist definiert als

$$p := P(X_{11} < X_{21}) + \frac{1}{2}P(X_{11} = X_{21}).$$

Er gibt Auskunft darüber, ob die Zufallsvariable X_{11} in Relation zu X_{21} zu kleineren, gleichen oder größeren Werten tendiert. Dazu definieren Brunner und Munzel ([6], Seite 17) den Begriff der *stochastischen Tendenz*:

Definition 2.1.1. *Seien $X_i \sim F_i$, $i = 1, 2$, unabhängige Zufallsvariablen. Dann tendiert X_1*

- *zu größeren Werten als X_2 falls $p < \frac{1}{2}$,*
- *zu kleineren Werten als X_2 falls $p > \frac{1}{2}$.*

Die Zufallsvariablen X_1 und X_2 heißen tendentiell gleich, falls $p = \frac{1}{2}$.

Für die Betrachtung unabhängiger Zufallsvariablen lässt sich der relative Effekt auch äquivalent in der Integralform darstellen:

$$p := \int F_1 dF_2.$$

Dabei ist F_i die normalisierte Version der Verteilungsfunktion von X_i :

$$\begin{aligned} F_i(x) &= \frac{1}{2} [F_i^+(x) + F_i^-(x)] \\ &= \frac{1}{2} [P(X_i < x) + P(X_i \leq x)]. \end{aligned}$$

Falls die Verteilungsfunktion stetig ist, gilt $F_i(x) = F_i^+(x) = F_i^-(x)$, da in diesem Fall $P(X_i = x) = 0$, $\forall x$.

Für unabhängig, identisch verteilte Zufallsvariablen kann eine Beziehung zwischen $H_0^F : F_1 = F_2$ und $H_0^p : p = \frac{1}{2}$ gezeigt werden:

Proposition 2.1.2 (Eigenschaften des relativen Effektes). *Der Relative Effekt p ist gleich $1/2$, falls X_1 und X_2 unabhängige, identisch verteilte Zufallsvariablen sind.*

Beweis. s. [6], Seite 19. □

Im Allgemeinen gibt es keine Beziehung zwischen dem relativen Effekt p und der Differenz von Erwartungswerten. Das hat zur Folge, dass die Hypothesen $H_0^\mu : \mu_1 = \mu_2$ und $H_0^p : p = \frac{1}{2}$ unterschiedliche Resultate liefern

können. Für Skalenalternativen lässt sich zeigen, dass aus $p = \frac{1}{2}$ die Gleichheit der Erwartungswerte $\mu_1 = \mu_2$ folgt (vgl. [6] S.18).

2.1.2 Empirische Verteilungsfunktionen

Da die Verteilungsfunktionen und damit auch der relative Effekt im Allgemeinen unbekannt sind, müssen diese geschätzt werden. Dafür definiert man zunächst die Zählfunktionen mit:

$$\begin{aligned} c_i^-(x) &= \begin{cases} 0 & , x \leq 0 \\ 1 & , x > 0 \end{cases} \quad \textit{links-stetige Version} \\ c_i^+(x) &= \begin{cases} 0 & , x < 0 \\ 1 & , x \geq 0 \end{cases} \quad \textit{rechts-stetige Version} \\ c_i(x) &= \frac{1}{2} [c_i^-(x) + c_i^+(x)] \quad \textit{normalisierte Version} \end{aligned}$$

Zur Herleitung eines Schätzers \hat{p} werden einige Eigenschaften der Zählfunktionen benötigt.

Lemma 2.1.3 (Erwartungswert der Zählfunktion). *Seien $X_{ik} \sim F_i$ unabhängig identisch verteilte Zufallsvariablen, $i = 1, 2$; $k = 1, \dots, n_i$. Dann gilt für alle $i, r \in \{1, 2\}$ und $k, s \in \{1, \dots, n_i\}$:*

$$\begin{aligned} E[c(x - X_{ik})] &= F_i(x) \\ E[c(X_{ik} - X_{rs})] &= \int F_r dF_i. \end{aligned}$$

Beweis. s. [6], Seite 191. □

Mithilfe der normalisierten Zählfunktion lässt sich die *normalisierte empirische Verteilungsfunktion* als gemittelte Summe über die gezählten Beob-

achtungen X_{ik} definieren:

$$\widehat{F}_i(x) := \frac{1}{n_i} \sum_{k=1}^{n_i} c(x - X_{ik}). \quad (1)$$

Das folgende Lemma listet einige wichtige Eigenschaften und Abschätzungen der empirischen Verteilungsfunktion auf:

Lemma 2.1.4 (Eigenschaften von \widehat{F}). *Seien $X_{ik} \sim F_i$, unabhängig identisch verteilte Zufallsvariablen, $i = 1, 2$; $k = 1, \dots, n_i$ und \widehat{F}_i die in (1) definierte empirische Verteilungsfunktion. Dann gilt:*

$$\begin{aligned} E \left[\widehat{F}_i(x) \right] &= F_i(x) \quad \forall x \\ E \left[\widehat{F}_i(X_{rs}) \right] &= \int F_i dF_r \\ E \left[\widehat{F}_i(x) - F_i(x) \right]^2 &\leq \frac{1}{n_i} \\ E \left[\widehat{F}_i(X_{rs}) - F_i(X_{rs}) \right]^2 &\leq \frac{1}{n_i}. \end{aligned}$$

Beweis. s. [6], Seite 192 ff. □

Mit Hilfe der empirischen Verteilungsfunktionen lässt sich nun ein intuitiver Schätzer für den relativen Effekt p angeben:

$$\widehat{p} := \int \widehat{F}_2 d\widehat{F}_1 = \frac{1}{n_1} \sum_{k=1}^{n_1} \widehat{F}_2(X_{1k}).$$

Dabei heißt $\widehat{F}_2(X_{1k})$ *normierte Platzierung* von X_{1k} unter den Zufallsvariablen X_{21}, \dots, X_{2n_2} . Erwartungstreue und Konsistenz des Schätzers folgen direkt aus dem vorherigen Lemma.

2.1.3 Ränge

Da sich die Berechnungen der normierten Platzierungen über die Zählfunktion sehr unübersichtlich gestalten, verwendet man stattdessen Ränge. Dabei stellt der Rang r_i die Platznummer einer Zahl $a_i \in \{a_1, \dots, a_n\}$ in der geordneten

Reihe $a_{(1)} \leq a_{(2)} \leq \dots \leq a_{(n-1)} \leq a_{(n)}$ dar. Im Falle von *Bindungen* (gleiche Werte) werden Mittelränge vergeben.

Definition 2.1.5 (Ränge von Zufallsvariablen). *Seien $c^-(x)$, $c^+(x)$, $c(x)$ die zuvor definierten Versionen der Zählfunktion und X_1, \dots, X_n Zufallsvariablen, die auf einer metrischen oder ordinalen Skala beobachtet werden. Dann heißt*

$$\begin{aligned} R_i^- &:= 1 + \sum_{k=1}^n c^-(X_i - X_k) && \text{Minimum-Rang} \\ R_i^+ &:= 1 + \sum_{k=1}^n c^+(X_i - X_k) && \text{Maximum-Rang} \\ R_i &:= 1 + \sum_{k=1}^n c(X_i - X_k) && \text{Mittel-Rang} \end{aligned}$$

von X_i unter allen Zufallsvariablen X_1, \dots, X_n .

Die drei Rangversionen unterscheiden sich, falls Bindungen vorliegen. Treten keine gleichen Werte auf, sind Minimum- Mittel- und Maximum-Rang identisch. In der weiteren Herleitung werden die Mittelränge verwendet, da sich die Summe

$$\sum_{k=1}^n R_k = \frac{n(n+1)}{2}$$

durch Bindungen nicht verändert.

Für die Berechnung der normierten Platzierungen muss der Begriff des Ranges verallgemeinert werden. Deshalb führt man die Begriffe des Gesamt-Ranges und des Intern-Ranges ein.

Definition 2.1.6 (Gesamt-Ränge und Intern-Ränge). *Sei $c(x)$ die normalisierte Zählfunktion und $X_{ik} \sim F_i$, identisch verteilte Zufallsvariablen, $i = 1, 2$; $k = 1, \dots, n_i$, die auf einer metrischen oder ordinalen Skala gemessen werden, $N = n_1 + n_2$. Dann heißt*

1. $R_{ik} = \frac{1}{2} + \sum_{j=1}^2 \sum_{\ell=1}^{n_j} c(X_{ik} - X_{j\ell})$ Gesamt-Rang von X_{ik} unter allen N Zufallsvariablen.

2. $R_{ik}^{(i)} = \frac{1}{2} + \sum_{\ell=1}^{n_i} c(X_{ik} - X_{i\ell})$ Intern-Rang von X_{ik} in der Stichprobe i .

Mit Hilfe der Intern- und Gesamt-Ränge können die normierten Platzierungen nun folgendermaßen dargestellt werden:

Lemma 2.1.7 (Berechnung der normierten Platzierungen). *Seien $X_{ik} \sim F_i$ identisch verteilte Zufallsvariablen, $i = 1, 2; k = 1, \dots, n_i$, und $\hat{F}_i(x)$ die empirische Verteilungsfunktion zu $F_i(x)$. Des Weiteren sei R_{ik} der Gesamt-Rang von X_{ik} unter allen $N = n_1 + n_2$ Beobachtungen und $R_{ik}^{(i)}$ der Intern-Rang in der Gruppe i . Dann lassen sich die normierten Platzierungen für $r \neq i$ schreiben als:*

$$\begin{aligned}\hat{F}_i(X_{ik}) &= \frac{1}{n_i} \left(R_{ik}^{(i)} - \frac{1}{2} \right) \\ \hat{F}_r(X_{ik}) &= \frac{1}{n_r} \left(R_{ik} - R_{ik}^{(i)} \right).\end{aligned}$$

Beweis.

$$\begin{aligned}\hat{F}_i(X_{ik}) &= \frac{1}{n_i} \sum_{\ell=1}^{n_i} c(X_{ik} - X_{i\ell}) \\ &= \frac{1}{n_i} \left(\left[\frac{1}{2} + \sum_{\ell=1}^{n_i} c(X_{ik} - X_{i\ell}) \right] - \frac{1}{2} \right) \\ &= \frac{1}{n_i} \left(R_{ik}^{(i)} - \frac{1}{2} \right) \\ \hat{F}_r(X_{ik}) &= \frac{1}{n_r} \left(\sum_{\ell=1}^{n_r} c(X_{ik} - X_{r\ell}) \right) \\ &\stackrel{i \neq r}{=} \frac{1}{n_r} \left(\left[\frac{1}{2} + \sum_{j=1}^2 \sum_{\ell=1}^{n_j} c(X_{ik} - X_{j\ell}) \right] - \left[\frac{1}{2} + \sum_{h=1}^{n_i} c(X_{ik} - X_{ih}) \right] \right) \\ &= \frac{1}{n_r} \left(R_{ik} - R_{ik}^{(i)} \right)\end{aligned}$$

□

Ein weiterer Vorteil der Vergabe von Rängen besteht darin, dass deren direkte Berechnung in vielen Statistik-Programmen enthalten ist und damit bei umfangreichen Simulationen wertvolle Rechenzeit eingespart werden kann.

Für den Schätzer des relativen Effektes erhält man die folgende Rangdarstellung

$$\hat{p} = \frac{1}{n_1} \sum_{k=1}^{n_1} \hat{F}_2(X_{1k}) = \frac{1}{n_1} \left(\bar{R}_2 - \frac{n_2 + 1}{2} \right) \text{ mit } \bar{R}_2 := \frac{1}{n_2} \sum_{k=1}^{n_2} R_{2k}.$$

2.2 Das nichtparametrische Behrens-Fisher-Problem

Wie bereits im vorherigen Kapitel angesprochen, liegt dem Behrens-Fisher-Problem die Annahme ungleicher Varianzen zugrunde. Sind die Beobachtungen nicht normalverteilt, kann es bei den parametrischen Testverfahren zu starken Niveau-Überschreitungen kommen.

2.2.1 Hypothesen

Der relative Effekt ist ein Vergleichsinstrument für Zufallsvariablen, das für beliebige Verteilungsklassen angewendet werden kann. Sind die zugrunde liegenden Verteilungen identisch, so sind die Aussagen

- die Erwartungswerte sind gleich, und
- der relative Effekt zwischen X_1 und X_2 ist gleich $\frac{1}{2}$

äquivalent. Da der relative Effekt keine Eigenschaften der Verteilungsfunktionen voraussetzt, kann er für beliebige Stichprobenvergleiche verwendet werden. Daher formuliert man, angelehnt an die Definition der stochastischen Tendenz, die Hypothese

$$H_0^p : p = \frac{1}{2} \quad \text{gegen die Alternative} \quad H_1^p : p \neq \frac{1}{2}.$$

Um die Hypothese zu überprüfen, muss der relative Effekt aus den Daten geschätzt werden. Das folgende Kapitel untersucht die asymptotische Verteilung des relativen Effektschätzers und führt zu einer Approximation des nichtparametrischen Behrens-Fisher-Problems.

2.2.2 Verteilung des relativen Effektes

Man betrachte zwei Datenvektoren $\mathbf{X}_i = (X_{i1}, \dots, X_{in_i})'$, bestehend aus unabhängig identisch verteilten Zufallsvariablen $X_{ik} \sim F_i$, $i = 1, 2$; $k = 1, \dots, n_i$. Dabei seien die Verteilungsfunktionen F_1 und F_2 (abgesehen von 1-Punkt-Verteilungen) beliebig. Es ist zu untersuchen, ob eine der beiden Beobachtungsgruppen zu größeren Werten als die andere tendiert. Für den Vergleich der Beobachtungen arbeitet man unter der Hypothese: $H_0^p : p = \frac{1}{2}$ und betrachtet im Folgenden die asymptotische Verteilung des Schätzers \hat{p} :

$$\begin{aligned} \sqrt{N}(\hat{p} - p) &\stackrel{H_0}{=} \sqrt{N} \left[\frac{1}{n_1} \left(\bar{R}_{2\cdot} - \frac{n_2 + 1}{2} \right) - \frac{1}{2} \right] \\ &= \sqrt{N} \left[\frac{1}{N} (\bar{R}_{2\cdot} - \bar{R}_{1\cdot}) \right] \\ &= \frac{1}{\sqrt{N}} (\bar{R}_{2\cdot} - \bar{R}_{1\cdot}). \end{aligned}$$

Da keine Verteilungsannahmen getroffen werden, muss auf asymptotische Resultate der Wahrscheinlichkeitstheorie, wie den zentralen Grenzwertsatz, zurückgegriffen werden. Dafür benötigt man eine standardisierte Teststatistik unabhängiger Zufallsvariablen. Da die Ränge R_{ik} nicht unabhängig sind, gestaltet sich die Berechnung der Varianz von $\sqrt{N}(\hat{p} - p)$ kompliziert. Man könnte daher eine Teststatistik unabhängiger Zufallsvariablen, deren Verteilung asymptotisch äquivalent zu der von $\sqrt{N}(\hat{p} - p)$ ist, zur Bestimmung der Varianz benutzen. Der asymptotische Äquivalenzsatz liefert genau diese Größe, die durch die asymptotisch normierten Platzierungen dargestellt wird:

Satz 2.2.1 (Asymptotischer Äquivalenzsatz). *Die Zufallsvariablen $X_{ik} \sim F_i$ seien unabhängig und identisch verteilt. Bezeichne $\sigma_i^2 = \text{Var}(F_r(X_{i1}))$ die Varianz der asymptotisch normierten Platzierung und \hat{p} den Schätzer für den relativen Effekt p . Falls $N/n_i \leq N_0 < \infty$ und $\sigma_i^2 > 0$ gilt:*

$$\sqrt{N}(\hat{p} - p) \doteq \sqrt{N} \left(\frac{1}{n_2} \sum_{k=1}^{n_2} F_1(X_{2k}) - \frac{1}{n_1} \sum_{k=1}^{n_1} F_2(X_{1k}) + 1 - 2p \right). \quad (2)$$

Beweis. s. [6], Seite 207. □

Bemerkung.

1. Das Zeichen \doteq bedeutet asymptotisch äquivalent. Die Definition des Begriffs ist dem Anhang zu entnehmen.
2. Die Bedingung $N/n_i \leq N_0 < \infty$ ist so zu verstehen, dass für $N \rightarrow \infty$ der Quotient $\frac{N}{n_i} \leq N_0 < \infty$ für beide Stichproben durch eine obere Grenze $N_0 \in \mathbb{N}$ beschränkt wird.

Die asymptotisch normierten Platzierungen $Y_{1k} = F_2(X_{1k})$, $k = 1, \dots, n_1$ und $Y_{2k} = F_1(X_{2k})$, $k = 1, \dots, n_2$, sind aufgrund der Unabhängigkeit der X_{ik} unabhängige, gleichmäßig beschränkte Zufallsvariablen zwischen 0 und 1. Wegen der asymptotischen Äquivalenz aus (2) kann der Schätzer des relativen Effektes nun durch die Varianzen σ_1^2 und σ_2^2 standardisiert werden:

$$\begin{aligned} T_N &= \frac{\sqrt{N}(\hat{p} - p)}{\sqrt{\text{Var}\left(\sqrt{N}(\bar{Y}_2 - \bar{Y}_1 + 1 - 2p)\right)}} \\ &\stackrel{H_0}{=} \frac{\frac{1}{\sqrt{N}}(\bar{R}_2 - \bar{R}_1)}{\sqrt{N\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)}} \\ &= \frac{\bar{R}_2 - \bar{R}_1}{\sqrt{N(n_2\sigma_1^2 + n_1\sigma_2^2)}} \sqrt{\frac{n_1n_2}{N}}. \end{aligned}$$

Die Voraussetzungen für den zentralen Grenzwertsatz von Lindeberg-Levy sind erfüllt. Deshalb gilt für die Teststatistik:

$$T_N = \frac{\bar{R}_2 - \bar{R}_1}{\sqrt{N(n_2\sigma_1^2 + n_1\sigma_2^2)}} \sqrt{\frac{n_1n_2}{N}} \dot{\sim} N(0, 1). \quad (3)$$

Im Folgenden wird die Bezeichnung

$$\sigma_N^2 = N(n_2\sigma_1^2 + n_1\sigma_2^2) \quad (4)$$

verwendet.

Für die erwartungstreue Schätzung der σ_i^2 könnten die empirischen Varianzschätzer

$$\tilde{\sigma}_i^2 = \frac{1}{n_i - 1} \sum_{k=1}^{n_i} (Y_{ik} - \bar{Y}_{i\cdot})^2 = \frac{1}{n_i - 1} \mathbf{Y}'_i \mathbf{P}_{n_i} \mathbf{Y}_i, \quad (5)$$

mit $\mathbf{Y}'_i = (Y_{i1}, \dots, Y_{in_i})$ und $\mathbf{P}_{n_i} = \mathbf{J}_{n_i} - \frac{1}{n_i} \mathbf{I}_{n_i}$, verwendet werden und man würde für (4) den Schätzer

$$\tilde{\sigma}_N^2 = N \left(n_2 \tilde{\sigma}_1^2 + n_1 \tilde{\sigma}_2^2 \right) \quad (6)$$

erhalten - wenn die Zufallsvariablen Y_{ik} beobachtbar wären.

Da eine Berechnung von (5) aufgrund unbekannter Verteilungsfunktionen nicht möglich ist, ersetzt man die nichtbeobachtbaren Y_{ik} durch die beobachtbaren normierten Platzierungen $\hat{Y}_{ik} = \hat{F}_r(X_{ik})$. Es muss dabei gezeigt werden, dass die \hat{Y}_{ik} für große Stichprobenumfänge nahe genug an den Y_{ik} liegen, was aber aus Lemma 2.1.4 folgt.

Die normierten Platzierungen lassen sich folgendermaßen umformen:

$$\begin{aligned} \hat{Y}_{ik} &= \frac{1}{n_r} \left[R_{ik} - R_{ik}^{(i)} \right] \\ \hat{Y}_{ik} - \hat{\bar{Y}}_{i\cdot} &= \frac{1}{n_r} \left[R_{ik} - R_{ik}^{(i)} - \bar{R}_{i\cdot} + \frac{n_i + 1}{2} \right]. \end{aligned}$$

Der folgende Satz beschreibt, dass die quadratischen Formen der normierten Platzierungen konsistent für die Varianzen σ_i^2 sind.

Satz 2.2.2. *Seien $X_{ik} \sim F_i$ unabhängig identisch verteilte Zufallsvariablen und \hat{Y}_{ik} die normierten Platzierungen der X_{ik} , $i = 1, 2$; $k = 1, \dots, n_i$. Dann gilt mit $\hat{\mathbf{Y}}_i = \frac{1}{N - n_i} (R_{i1} - R_{i1}^{(i)}, \dots, R_{in_i} - R_{in_i}^{(i)})'$ und $\mathbf{V} = \text{Cov}(\mathbf{Y}_i) = \sigma_i^2 \cdot \mathbf{I}_{n_i}$:*

$$\hat{\sigma}_i^2 = \frac{1}{n_i - 1} \hat{\mathbf{Y}}'_i \mathbf{P}_{n_i} \hat{\mathbf{Y}}_i \quad (7)$$

ist konsistenter Schätzer für σ_i^2 in dem Sinne, dass $E[\hat{\sigma}_i^2 / \sigma_i^2 - 1]^2 \rightarrow 0$.

Beweis. s. [6], Seite 229. □

Mit diesen Ergebnissen lässt sich die Varianz (4) durch die empirischen Varianzen aus (7) schätzen:

$$\hat{\sigma}_N^2 = N (n_2 \hat{\sigma}_1^2 + n_1 \hat{\sigma}_2^2). \quad (8)$$

Unter Anwendung des asymptotischen Äquivalenzsatzes und des zentralen Grenzwertsatzes konnte gezeigt werden, dass die Teststatistik T_n asymptotisch standardnormalverteilt ist. Für die Varianz (4) wurde in (8) ein Schätzer angegeben. Da in der Anwendung mit endlichen Stichprobengrößen gearbeitet werden muss, wird eine t-Approximation für kleine Stichproben vorgestellt.

2.2.3 t-Approximation

Für kleine Stichproben kann ein approximatives Verfahren zur Lösung des nichtparametrischen Behrens-Fisher-Problems hergeleitet werden, indem man die Teststatistik aus (3) durch eine t-Verteilung approximiert. Dieser Ansatz wird dadurch motiviert, dass die t-Verteilung für große Freiheitsgrade gegen die Normalverteilung konvergiert. Dafür approximiert man die Varianz (6) durch eine gestreckte χ_f^2 -Verteilung mit dem Streckungsfaktor g . Mit dem Ansatz der Box-Approximation (vgl. [1]) lässt sich das folgende Gleichungssystem aufstellen:

$$E \left[\tilde{\sigma}_N^2 \right] = N(n_2 \sigma_1^2 + n_1 \sigma_2^2) \stackrel{!}{=} g \cdot f \quad (9)$$

$$Var \left[\tilde{\sigma}_N^2 \right] = V_{\text{Atq}} \stackrel{!}{=} 2g^2 \cdot f \quad (10)$$

Die Varianz V_{Atq} kann mit Hilfe des Satzes von Atiqullah berechnet werden.

Satz 2.2.3 (Atiqullah, 1962). *Sei $\mathbf{Y} = (Y_1, \dots, Y_n)'$ ein Zufallsvektor bestehend aus $Y_i \sim F$ unabhängig identisch verteilten Zufallsvariablen, $i = 1, \dots, n$ mit $E(Y_1) = \mu$, $\text{Var}(Y_1) = \sigma^2$, $\mu_3 := E(Y_1 - \mu)^3$ und $\mu_4 := E(Y_1 - \mu)^4$. Für*

jede symmetrische Matrix \mathbf{A} und den Vektor $\mathbf{a} := \text{diag}\{\mathbf{A}\}$ gilt:

$$\text{Var}(\mathbf{Y}'\mathbf{A}\mathbf{Y}) = (\mu_4 - 3 \cdot \sigma^4) \mathbf{a}'\mathbf{a} + 2 \cdot \sigma^4 \text{Sp}[\mathbf{A}^2] + 4 \cdot \sigma^4 \boldsymbol{\mu}'\mathbf{A}^2\boldsymbol{\mu} + 4 \cdot \mu_3 \boldsymbol{\mu}'\mathbf{A}\mathbf{a}.$$

Beweis. Der Satz ist ein Spezialfall des Theorems 1.8 aus [5], Seite 14. \square

Für $\mathbf{A} = \mathbf{P}_{n_i}$ vereinfacht sich das Resultat zu

$$\text{Var}(\mathbf{Y}'\mathbf{P}_n\mathbf{Y}) = (\mu_4 - 3 \cdot \sigma^4) \frac{(n-1)^2}{n} + 2 \cdot \sigma^4 \cdot (n-1). \quad (11)$$

Der Brunner-Munzel-Test approximiert an dieser Stelle mit unabhängigen normalverteilten Zufallsvariablen $Z_i \sim N(\mu, \sigma^2)$, $i = 1, \dots, n$, und erhält für die Varianz der quadratischen Form:

$$\text{Var}(\mathbf{Y}'\mathbf{P}_n\mathbf{Y}) = \text{Var}(\mathbf{Z}'\mathbf{P}_n\mathbf{Z}) = 2 \cdot \sigma^4 \cdot (n-1).$$

Daraus resultiert für die Varianz aus (10):

$$V_{\text{BM}} = N^2 \sum_{i=1}^2 (N - n_i)^2 \left[\frac{2 \cdot \sigma_i^4}{n_i - 1} \right]. \quad (12)$$

Für endliche Stichproben ist die Annahme der Kurtosis $\beta_2 = 3$ eine sehr grobe Approximation. Für Verteilungen, deren Wölbung stark von diesem Wert abweicht, könnte dies zu schlechten Auswertungsergebnissen führen. Daher wird mit (11) die folgende Varianz gebildet:

$$V_{\text{Atq}} = N^2 \sum_{i=1}^2 (N - n_i)^2 \left[\frac{\mu_{i,4} - 3 \cdot \sigma_i^4}{n_i} + \frac{2 \cdot \sigma_i^4}{n_i - 1} \right], \quad (13)$$

An dieser Stelle bezeichnet $\mu_{i,4}$ das 4. zentrale Moment und σ_i^4 die quadrierte Varianz der Stichprobe i . Mit diesen Resultaten lässt sich das Gleichungssystem der Box-Approximation umformen und man erhält:

$$f = \frac{2 \cdot N^2 (n_2 \sigma_1^2 + n_1 \sigma_2^2)^2}{V_{\text{Atq}}} \quad g = \frac{V_{\text{Atq}}}{2 \cdot N (n_2 \sigma_1^2 + n_1 \sigma_2^2)}$$

Da $f \cdot g = N(n_2\sigma_1^2 + n_1\sigma_2^2) = \sigma_N^2$ lässt sich (3) umformen zu:

$$T_N = \frac{\bar{R}_2 - \bar{R}_1}{\sqrt{f \cdot g}} \sqrt{\frac{n_1 n_2}{N}} \quad \overset{\cdot}{\sim} N(0, 1). \quad (14)$$

Die Momente μ_4 und σ^4 können erwartungstreu aus den Daten geschätzt werden. Aus Gründen der Übersichtlichkeit wird der Satz für eine Stichprobe mit n Beobachtungen angegeben. Anschließend wird der Bezug zu den Platzierungen hergestellt und die Konsistenz des benötigten Schätzers gezeigt.

Satz 2.2.4. *Seien $Y_1, \dots, Y_n \sim F$ unabhängig identisch verteilte Zufallsvariablen und bezeichne μ_4 das 4. zentrale Moment und σ^4 die quadrierte Varianz von Y_1 . Dann sind*

$$\begin{aligned} \tilde{\mu}_4 &= \frac{n(n^2 - 2n + 3)}{(n-1)(n-2)(n-3)} \cdot \left(\frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^4 \right) \\ &\quad - \frac{3n(2n-3)}{(n-1)(n-2)(n-3)} \cdot \left(\frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2 \right)^2 \\ \tilde{\sigma}^4 &= \frac{n(n^2 - 3n + 3)}{(n-1)(n-2)(n-3)} \cdot \left(\frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2 \right)^2 \\ &\quad - \frac{n}{(n-2)(n-3)} \cdot \left(\frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^4 \right) \end{aligned}$$

erwartungstreu für μ_4 und σ^4 .

Beweis. s. [4], Seite 161, Korollar 2.1.4. □

Der Satz lässt sich direkt auf die asymptotisch normierten Platzierungen anwenden und man erhält, wie im Fall der Varianzschätzer nichtbeobachtbare Größen, die wiederum durch die normierten Platzierungen approximiert werden müssen. An dieser Stelle muss erneut die Konsistenz der gebildeten Schätzer gezeigt werden.

Satz 2.2.5 (Konsistenz von $\hat{\mu}_4$ und $\hat{\sigma}_i^4$). *Seien \hat{Y}_{ik} die normierten Platzierungen der unabhängig, identisch verteilten Zufallsvariablen $X_{ik} \sim F_i$, $i = 1, 2$;*

$k = 1, \dots, n_i$. Dann sind

$$\begin{aligned}\widehat{\mu}_{i,4} &= \frac{n_i(n_i^2 - 2n_i + 3)}{(n_i - 1)(n_i - 2)(n_i - 3)} \cdot \left(\frac{1}{n_i} \sum_{k=1}^{n_i} (\widehat{Y}_{ik} - \widehat{Y}_{i\cdot})^4 \right) \\ &\quad - \frac{3n_i(2n_i - 3)}{(n_i - 1)(n_i - 2)(n_i - 3)} \cdot \left(\frac{1}{n_i} \sum_{k=1}^{n_i} (\widehat{Y}_{ik} - \widehat{Y}_{i\cdot})^2 \right)^2 \\ \widehat{\sigma}_i^4 &= \frac{n_i(n_i^2 - 3n_i + 3)}{(n_i - 1)(n_i - 2)(n_i - 3)} \cdot \left(\frac{1}{n_i} \sum_{k=1}^{n_i} (\widehat{Y}_{ik} - \widehat{Y}_{i\cdot})^2 \right)^2 \\ &\quad - \frac{n_i}{(n_i - 2)(n_i - 3)} \cdot \left(\frac{1}{n_i} \sum_{k=1}^{n_i} (\widehat{Y}_{ik} - \widehat{Y}_{i\cdot})^4 \right)\end{aligned}$$

konsistente Schätzer für $\widetilde{\mu}_{i,4}$ und $\widetilde{\sigma}_i^4$, $i = 1, 2$.

Beweis.

Bezeichne $A = \frac{n_i(n_i^2 - 2n_i + 3)}{(n_i - 1)(n_i - 2)(n_i - 3)}$ und $B = \frac{3n_i(2n_i - 3)}{(n_i - 1)(n_i - 2)(n_i - 3)}$.

Dann gilt:

$$\begin{aligned}E \left[\widehat{\mu}_{i,4} - \widetilde{\mu}_{i,4} \right]^2 &= E \left(A \cdot \frac{1}{n_i} \sum_{k=1}^{n_i} \left[(\widehat{Y}_{ik} - \widehat{Y}_{i\cdot})^4 - (Y_{ik} - \bar{Y}_{i\cdot})^4 \right] \right. \\ &\quad \left. + B \cdot \left(\left[\frac{1}{n_i} \sum_{k=1}^{n_i} (Y_{ik} - \bar{Y}_{i\cdot})^2 \right]^2 - \left[\frac{1}{n_i} \sum_{k=1}^{n_i} (\widehat{Y}_{ik} - \widehat{Y}_{i\cdot})^2 \right]^2 \right) \right)^2\end{aligned}$$

Da die Platzierungen zwischen 0 und 1 liegen lässt sich der zweite Teil mit B abschätzen. Durch Abschätzung mit der Jensen- und c_r -Ungleichung folgt:

$$\leq 2 \cdot E \left(A \cdot \frac{1}{n_i} \sum_{k=1}^{n_i} \left[(\widehat{Y}_{ik} - \widehat{Y}_{i\cdot})^4 - (Y_{ik} - \bar{Y}_{i\cdot})^4 \right] \right)^2 + O \left(\frac{1}{n_i} \right)$$

Mit der 3. binomischen Formel, einer erneuten Abschätzung der Platzierungen und dem Maximum über alle k, ergibt sich:

$$\leq \max_{k=1, \dots, n_i} 8 \cdot A^2 E \left(\left((\widehat{Y}_{ik} - \widehat{Y}_{i\cdot})^2 - (Y_{ik} - \bar{Y}_{i\cdot})^2 \right)^2 \right) + O \left(\frac{1}{n_i} \right)$$

Erneutes Anwenden der 3. binomischen Formel und Abschätzen der Platzierungen führt in Verbindung mit der Jensen- und der c_r -Ungleichung zu:

$$\leq \max_{k=1, \dots, n_i} 64 \cdot A^2 \left[E \left(\widehat{Y}_{ik} - \widehat{Y}_{i \cdot} \right)^2 + E \left(\overline{Y}_{i \cdot} - \widehat{Y}_{i \cdot} \right)^2 \right] + O \left(\frac{1}{n_i} \right)$$

Nach Lemma 2.1.4 sind die normierten Platzierungen konsistent für Y_{ik} . Daher folgt die Konsistenz von $\widehat{\mu}_{i,4}$.

$$\text{Bezeichne } C = \frac{n_i(n_i^2 - 3n_i + 3)}{(n_i - 1)(n_i - 2)(n_i - 3)} \text{ und } D = \frac{n_i}{(n_i - 2)(n_i - 3)}.$$

Dann gilt:

$$\begin{aligned} E \left[\widehat{\sigma}_{i,4} - \widetilde{\sigma}_{i,4} \right]^2 &= E \left(C \left(\left[\frac{1}{n_i} \sum_{k=1}^{n_i} \left(\widehat{Y}_{ik} - \widehat{Y}_{i \cdot} \right)^2 \right]^2 - \left[\frac{1}{n_i} \sum_{k=1}^{n_i} \left(Y_{ik} - \overline{Y}_{i \cdot} \right)^2 \right]^2 \right) \right. \\ &\quad \left. + D \frac{1}{n_i} \sum_{k=1}^{n_i} \left[\underbrace{\left(Y_{ik} - \overline{Y}_{i \cdot} \right)^4}_{\leq 1} - \underbrace{\left(\widehat{Y}_{ik} - \widehat{Y}_{i \cdot} \right)^4}_{\geq 0} \right] \right)^2 \end{aligned}$$

Mit der 3. binomischen Formel folgt:

$$\begin{aligned} &\leq E \left(C \cdot \left(\frac{1}{n_i} \sum_{k=1}^{n_i} \left[\left(\widehat{Y}_{ik} - \widehat{Y}_{i \cdot} \right)^2 - \left(Y_{ik} - \overline{Y}_{i \cdot} \right)^2 \right] \right) \right. \\ &\quad \left. \left(\frac{1}{n_i} \sum_{k=1}^{n_i} \left[\left(\widehat{Y}_{ik} - \widehat{Y}_{i \cdot} \right)^2 + \left(Y_{ik} - \overline{Y}_{i \cdot} \right)^2 \right] \right) + D \right)^2 \end{aligned}$$

Benutzt man die Beschränkung der Platzierungen, die 3. binomische Formel und die Jensen- und c_r -Ungleichung gelangt man schließlich zu

$$\leq \max_{k=1, \dots, n_i} 64 \cdot C^2 \cdot \left[E \left(\widehat{Y}_{ik} - Y_{ik} \right)^2 + E \left(\overline{Y}_{i \cdot} - \widehat{Y}_{i \cdot} \right)^2 \right] + O \left(\frac{1}{n_i} \right)$$

Die Konsistenz folgt wieder aus Lemma 2.1.4. \square

Mit diesen Ergebnissen approximiert man (14), motiviert durch die Arbei-

ten von Greenwood und Nikulin (vgl. [3]), durch die t-verteilte Teststatistik

$$t_N = \frac{\bar{R}_2 - \bar{R}_1}{\sqrt{N(n_2\hat{\sigma}_1^2 + n_1\hat{\sigma}_2^2)}} \sqrt{\frac{n_1 n_2}{N}}, \quad (15)$$

mit f Freiheitsgraden. Da die t-Verteilung mit wachsendem Freiheitsgrad gegen die Standardnormalverteilung konvergiert, konvergiert der Fehler, der durch die t-Approximation entsteht, für große Stichprobenumfänge gegen 0. Setzt man in f die empirischen Varianzschätzer ein, erhält man für den geschätzten Freiheitsgrad der Verteilung

$$\hat{f} = \frac{2 \cdot N^2 (n_2\hat{\sigma}_1^2 + n_1\hat{\sigma}_2^2)^2}{\hat{V}_{\text{Atq}}}, \quad (16)$$

mit

$$\hat{V}_{\text{Atq}} = N^2 \sum_{i=1}^2 (N - n_i)^2 \left[\frac{\hat{\mu}_{i,4} - 3 \cdot \hat{\sigma}_i^4}{n_i} + \frac{2 \cdot \hat{\sigma}_i^4}{n_i - 1} \right].$$

Im folgenden Kapitel soll der Test durch zufällig erzeugte Datensätze geprüft werden. Sollte er das getestete Niveau für Verteilungen mit extremer Wölbung besser einhalten als der Brunner-Munzel-Test, wäre die Verwendung des neuen Schätzers aus (16) zu empfehlen.

3 Simulationen

Um die Qualität des entwickelten nichtparametrischen Tests zu untersuchen, wurden Datensätze zu verschiedenen Verteilungen in dem Statistikprogramm SAS erzeugt und unter der Fragestellung untersucht, ob eine der beiden Gruppen zu größeren Werten tendiert als die andere. Da die Stärke der nichtparametrischen Tests in der universellen Anwendbarkeit liegt, soll diese Eigenschaft für unterschiedliche parametrische Datensätze belegt werden.

Zur Generierung und Auswertung der Daten wird das Statistikprogramm SAS verwendet. Für Stichprobenumfänge zwischen 10 und 30 werden Zufallsvektoren mit der RANDOM-Prozedur erzeugt und anschließend mit der in Abschnitt 2.2.3 beschriebenen t-Approximation ausgewertet. Falls die Auswertung in einzelnen Fällen durch schlecht geschätzte Parameter verhindert wird, muss eine Parameteranpassung durch das Programm erfolgen.

3.1 Anpassung der Schätzer

Unter bestimmten Voraussetzungen kann es bei kleinen Stichprobenumfängen zu fehlerhaften Auswertungen der Daten kommen. Die nächsten beiden Abschnitte erläutern, wie das Programm in diesen Sonderfällen verfährt.

3.1.1 Berechnung der Varianzen

Bei der Simulation von Zufallsvariablen mit gleichem Erwartungswert kann es vorkommen, dass die Datensätze getrennt, das bedeutet ohne Überschneidung, vorliegen. Da die Verteilung des relativen Effektes über die Gesamt- und Intern-Ränge standardisiert wird, bricht die SAS-Prozedur die Auswertung an dieser Stelle ab, da die Aufforderung übergeben wird durch Null zu teilen. Um die Datensätze dennoch auszuwerten, wird die Varianz abgeschätzt.

Lemma 3.1.1 ($\hat{\sigma}_N^2$ für zwei Stichproben mit einer Bindung). *Seien $\mathbf{X}_i = (X_{i1}, \dots, X_{in_i})'$, $i = 1, 2$, zwei Zufallsvektoren und es gelte*

$$X_{11} \leq X_{12} \leq \dots \leq X_{1(n_1-1)} < X_{1n_1} = X_{21} < X_{22} \leq \dots \leq X_{2n_2}.$$

Bezeichne R_{ik} den Gesamt-Rang und $R_{ik}^{(i)}$ den Intern-Rang von X_{ik} , $i = 1, 2$; $k = 1, \dots, n_i$. Dann gilt für den in (7) angegebenen Varianzschätzer:

$$\hat{\sigma}_N^2 = N (n_2 \hat{\sigma}_1^2 + n_1 \hat{\sigma}_2^2) = \frac{N}{2n_1 n_2}. \quad (17)$$

Beweis. Zuerst berechnet man $A = \sum_{k=1}^{n_1} \left(R_{1k} - R_{1k}^{(1)} - \bar{R}_{1\cdot} + \bar{R}_{1\cdot}^{(1)} \right)^2$:

$$\begin{aligned} A &= \sum_{k=1}^{n_1} \left(R_{1k} - R_{1k}^{(1)} - \bar{R}_{1\cdot} + \bar{R}_{1\cdot}^{(1)} \right)^2 \\ &= \sum_{k=1}^{n_1-1} \left(\bar{R}_{1\cdot}^{(1)} - \bar{R}_{1\cdot} \right)^2 + \left(\frac{1}{2} - \bar{R}_{1\cdot} + \bar{R}_{1\cdot}^{(1)} \right)^2 \\ &= \sum_{k=1}^{n_1-1} \left(-\frac{1}{2n_1} \right)^2 + \left(\frac{1}{2} - \frac{1}{2n_1} \right)^2 \\ &= (n_1 - 1) \left[\frac{1}{4n_1^2} \right] + \left(\frac{n_1 - 1}{2n_1} \right)^2 \\ &= \frac{n_1 - 1}{4n_1} \end{aligned}$$

Analog erhält man für $B = \sum_{k=1}^{n_2} \left(R_{2k} - R_{2k}^{(2)} - \bar{R}_{2\cdot} + \bar{R}_{2\cdot}^{(2)} \right)^2$:

$$\begin{aligned} B &= \sum_{k=1}^{n_2} \left(R_{2k} - R_{2k}^{(2)} - \bar{R}_{2\cdot} + \bar{R}_{2\cdot}^{(2)} \right)^2 \\ &= \left(n_1 - \frac{1}{2} - \frac{1}{n_2} \left(\sum_{\ell=1}^{n_2} n_1 - \frac{1}{2} \right) \right)^2 + \sum_{k=2}^{n_2} \left(n_1 - \frac{1}{n_2} \left(\sum_{\ell=1}^{n_2} n_1 - \frac{1}{2} \right) \right)^2 \\ &= \sum_{k=1}^{n_2} \left(n_1 - \frac{1}{n_2} \left(\sum_{\ell=1}^{n_2} n_1 - \frac{1}{2} \right) \right)^2 + \frac{1}{4} - \left(n_1 - \frac{1}{n_2} \left(\sum_{\ell=1}^{n_2} n_1 - \frac{1}{2} \right) \right) \\ &= \frac{1}{4n_2} + \frac{1}{4} - \frac{1}{2n_2} \\ &= \frac{n_2 - 1}{4n_2} \end{aligned}$$

Daraus ergibt sich schließlich:

$$\hat{\sigma}_N^2 = N \left(\frac{A}{(n_1 - 1)n_2} + \frac{B}{(n_2 - 1)n_1} \right) = \frac{N}{2n_1n_2}$$

□

3.1.2 Schätzung des Freiheitsgrades

Der Freiheitsgrad der t-Verteilung ist ein positiver Parameter, der die Verteilungsfunktion vollständig beschreibt. Wird dieser Parameter aus den Daten geschätzt, muss überprüft werden, ob durch diesen Wert überhaupt eine t-Verteilung gebildet werden kann. Deshalb soll in diesem Abschnitt der Schätzer des Freiheitsgrades aus (16) untersucht werden. In ausgeschriebener Form hat (16) die Gestalt

$$\hat{f} = \frac{2 \cdot (n_2 \hat{\sigma}_1^2 + n_1 \hat{\sigma}_2^2)^2}{\sum_{i=1}^2 (N - n_i)^2 \left[\frac{\hat{\mu}_{i,4} - 3 \cdot \hat{\sigma}_i^4}{n_i} + \frac{2 \cdot \hat{\sigma}_i^4}{n_i - 1} \right]}.$$

Da die Varianz des Schätzers (6) durch den Satz von Atiqullah berechnet wird, kann es, im Gegensatz zu V_{BM} , in zwei Situationen zu einer fehlerhaften Auswertung kommen:

1. Sollten die beobachteten Daten eine sehr große Wölbung besitzen, d.h. $\hat{\mu}_{i,4} \gg 3 \cdot \hat{\sigma}_i^4$, könnte daraus ein negativer Schätzer für den Freiheitsgrad resultieren. Dann wäre keine Approximation durch eine t-Verteilung möglich. In diesem Fall wird der Freiheitsgrad auf den Wert 1 gesetzt.
2. Für eine sehr kleine geschätzte Kurtosis kann der Nenner negativ werden. Um den Datensatz dennoch auszuwerten, wird auch in diesem Fall der Wert 1 übergeben, da die t-Verteilung flacher als die Normalverteilung, d.h. *platykurtisch* verläuft.

Da die t-Verteilung für $f \rightarrow \infty$ gegen die Normalverteilung konvergiert, ist für sehr große Werte des Schätzers keine Anpassung notwendig.

3.2 Simulationen für Skalenalternativen

In Kapitel 2.1.1 wurde bereits erwähnt, dass der relative Effekt für Skalenalternativen F_1, F_2 gleich $1/2$ ist, falls die Zentren der Verteilungen identisch sind. Daher kann unter diesen Voraussetzungen die Hypothese $H_0^\mu : \mu_1 = \mu_2$ mit dem Atiqullah-Test überprüft werden.

Als Beispiele für Skalenalternativen wurden normal- und gleichverteilte zentrierte Zufallsvariablen erzeugt. Die Anzahl der Beobachtungen variiert zwischen 10 und 30 und die Varianzen zwischen 1 und 10. Die Daten wurden mit dem Gauss-Test, t-Test und den t-Approximationen nach Atiqullah und Brunner-Munzel zu den 1%-, 5%- und 10%-Niveaus ausgewertet. Die Simulationsergebnisse befinden sich in den Tabellen 2 bis 5.

Bei der Simulation wird der Gauss-Test verwendet, um eine sehr genaue Auswertung für die gewählten Parameter zu demonstrieren. Der t-Test dient als zweites Vergleichsinstrument, da er einen der am häufigsten verwendeten Tests darstellt. Die Niveau-Überschreitungen sind darauf zurückzuführen, dass der gepoolte Varianzschätzer verwendet wird, der die ungleichen Varianzen mittelt.

Betrachtet man die verschiedenen Parameterkonstellationen, so fällt auf, dass die nichtparametrischen Tests die Niveaus sehr gut einhalten. Beide Testverfahren werten die Daten etwas *liberal* aus, was bedeutet, dass die Tests das vorgegebene Niveau leicht *überschreiten*. Dies gilt für alle Parameterkonstellationen, bis auf eine Ausnahme: Wird für die kleinere Stichprobe eine größere Varianz als in der Vergleichsgruppe simuliert, verhält sich der Test sehr *konservativ*, was bedeutet, dass das Niveau *unterschritten* wird und die Hypothese stochastisch gleicher Tendenzen häufiger als gewünscht verworfen wird. Ab 15 Beobachtungen in der kleineren Versuchsgruppe werden auch diese Verteilungen gut approximiert.

Der Brunner-Munzel-Test liefert etwas bessere Resultate als die Approximation nach Atiqullah. Das könnte damit zusammenhängen, dass die Schätzer

der Momente bei kleinen Stichproben zu sehr von der wahren Werten abweichen. Zusammenfassend stellt man fest, dass Skalenalternativen mit den nichtparametrischen Tests sehr genau ausgewertet werden können, wobei der Brunner-Munzel-Test die geringfügig besseren Ergebnisse liefert.

3.3 Simulationen für beliebige Verteilungen

3.3.1 Monte-Carlo-Simulationen

Für die Simulation von Zufallsvariablen müssen die Datensätze so erzeugt werden, dass die Bedingung $p = \frac{1}{2}$ erfüllt ist. Da sich die Berechnung der Faltung von verschiedenen Verteilungsfunktionen schwierig gestalten kann, wird vor der Datenerzeugung eine Monte-Carlo-Simulation durchgeführt. Der Vorteil dieses Verfahrens liegt darin, dass man den relativen Effekt mit großen Stichproben, aufgrund der Konsistenz von \hat{p} , gut approximieren kann. Die Prozedur der Parameteranpassung führt die folgenden Schritte aus.

1. Für 100.000 Simulationen werden jeweils 100.000 Beobachtungen in beiden Stichproben generiert und der Schätzer \hat{p} berechnet.
2. Ist der berechnete Wert größer als $\frac{1}{2}$ wird der Lokationsparameter verkleinert, für den umgekehrten Fall vergrößert.
3. Dieses Verfahren wird solange ausgeführt, bis die Differenz zwischen \hat{p} und $\frac{1}{2}$ kleiner als 0.0001 ist.

Ist der Parameter für die Hypothese $H_0 : p = \frac{1}{2}$ kalibriert, kann die Simulation für kleine Stichproben erfolgen.

3.3.2 Simulationen für schiefe Verteilungen

Mit dem relativen Effekt ist es möglich, verschiedene Verteilungsfunktionen miteinander zu vergleichen. Für den Vergleich zwischen schiefen und normalverteilten Daten bieten sich sowohl die Gamma-, als auch die nicht-zentrale χ^2 -Verteilung mit 3 Freiheitsgraden an. Die Stichprobenumfänge variieren auch hier zwischen 10 und 30 Beobachtungen. Da keine symmetrischen Verteilungen betrachtet werden, muss die beschriebene Monte-Carlo-Simulation

für jede Parameterkonstellation durchgeführt werden. Die Simulationsergebnisse stehen im Anhang in den Tabellen 6 bis 9.

Auch für diese Verteilungsfunktionen liefern beide t-Approximationen gute Resultate mit einer liberalen Tendenz, solange die Gruppe mit weniger Beobachtungen keine große Varianz aufweist. Obwohl beide Testverfahren fast identische Ergebnisse liefern, ist der Brunner-Munzel-Test auch für diese Verteilungen die bessere Wahl, da das Niveau für große Varianzunterschiede besser eingehalten wird. Damit wird auch für die Hypothese $H_0^p : p = \frac{1}{2}$ gezeigt, dass die einfache Schätzung der Varianz zu erstaunlich guten Auswertungen, unabhängig von der Kurtosis der Verteilungsfunktion, führt.

3.4 Beispiel: Ferritin bei Kleinwuchs

Die in Kapitel 2.2.3 vorgestellte t-Approximation soll nun zur Auswertung des Ferritin-Beispiels aus Kapitel 1.2 verwendet werden. Die gemessenen Ferritinwerte und daraus resultierenden Intern- und Gesamt-Ränge sind im Anhang in Tabelle 1 zusammengestellt.

Für die Mittelwerte der Gesamt-Ränge errechnet man $\bar{R}_1 = 5,71$ und $\bar{R}_2 = 12,5$. Damit lassen sich die Varianzschätzer $\hat{\sigma}_1^2 = 0,048$ und $\hat{\sigma}_2^2 = 0,026$ bestimmen. Die Teststatistik hat den Wert $t_N = 3,76$ und wird durch eine t-Verteilung mit $\hat{f} = 18,35$ Freiheitsgraden approximiert. Daraus resultiert für die Hypothese $H_0^p : p = \frac{1}{2}$ ein p-Wert von 0,0014. Damit kann die Hypothese zu einem 5%-Niveau valide verworfen werden.

4 Zusammenfassung

In dieser Bachelorarbeit sollte der Brunner-Munzel-Test für Verteilungen untersucht werden, deren Kurtosis stark von der normalverteilter Daten abweicht. Dafür wurde unter Berücksichtigung der allgemeinen Varianz einer quadratischen Form der Atiqullah-Test hergeleitet. Durch die Simulationen konnte für die betrachteten Verteilungen gezeigt werden, dass beide Testverfahren gute Resultate für das nichtparametrische Behrens-Fisher-Problem liefern können, solange die Varianz der kleineren Stichprobe nicht allzu groß wird, da die Tests ansonsten zu sehr konservativen Ergebnissen führen.

Erstaunlicherweise lieferte der Brunner-Munzel-Test bessere Resultate als die Approximation nach Atiqullah. Dies könnte zum einen darauf zurückzuführen sein, dass die Schätzer für das 4. Moment und die quadrierte Varianz bei kleinen Stichproben zu sehr variieren können, zum anderen mussten die geschätzten Freiheitsgrade für negative Werte brutal mit einem festen Wert abgeschätzt werden. Zusammenfassend kann man festhalten, dass die simple Abschätzung der Varianz in (12) keinen negativen Einfluss auf die Qualität der resultierenden t-Approximation hat.

A Notation und Mathematische Resultate

A.1 Matrizen und Vektoren

Matrizen und Vektoren werden in der gesamten Arbeit durch **Fettdruck** hervorgehoben. Dies erhöht die Lesbarkeit deutlich, da eine einfache Unterscheidung zwischen Skalaren und Matrizen möglich ist.

$$\mathbf{A} := \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

bezeichnet eine $m \times n$ -Matrix mit m Zeilen und n Spalten. Die transponierte Matrix \mathbf{A}' ist gegeben durch

$$\mathbf{A}' := \begin{pmatrix} a_{11} & \dots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \dots & a_{mn} \end{pmatrix}.$$

Ist $m = n$ so nennt man \mathbf{A} quadratisch. Falls eine quadratische Matrix die Bedingung $\mathbf{A} = \mathbf{A}'$ erfüllt, so nennt man diese *symmetrisch*. Ein Vektor ist eine Matrix mit nur einer Spalte oder Zeile. Setzt man $n = 1$ so bezeichnet

$$\mathbf{a} := \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}$$

einen Spaltenvektor mit m Einträgen. Transponiert man den Vektor \mathbf{a} , so erhält man einen Zeilenvektor

$$\mathbf{a}' := \left(a_1 \quad \dots \quad a_m \right).$$

Einige spezielle Matrizen sollen jetzt vorgestellt werden:

1. Die Identität ist eine symmetrische Matrix mit Einsen auf der Diago-

nalen:

$$\mathbf{I}_n := \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

2. Die Matrix \mathbf{J} ist eine beliebig dimensionierte Matrix die mit lauter Einsen besetzt ist. Der zugehörige Vektor wird mit $\mathbf{1}$ bezeichnet und es gilt:

$$\mathbf{J}_{(m \times n)} = \mathbf{1}_m \cdot \mathbf{1}'_n = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix}.$$

3. Die zentrierende Matrix \mathbf{P}_n zentriert einen Vektor um den Mittelwert seiner Einträge:

$$\mathbf{P}_n := \mathbf{I}_n - \frac{1}{n} \mathbf{J}_n = \begin{pmatrix} 1 - \frac{1}{n} & & -\frac{1}{n} \\ & \ddots & \\ -\frac{1}{n} & & 1 - \frac{1}{n} \end{pmatrix}$$

A.2 Quadratische Formen

Definition A.2.1 (Quadratische Form). Sei $\mathbf{X} = (X_1, \dots, X_n)'$ ein Zufallsvektor und \mathbf{A} eine symmetrische Matrix. Dann heißt

$$Q = \mathbf{X}' \mathbf{A} \mathbf{X}$$

quadratische Form.

Satz A.2.2 (Lancaster). Sei $\mathbf{X} = (X_1, \dots, X_n)'$ ein Zufallsvektor mit Erwartungswert $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)'$ und $\mathbf{V} = \text{Var}(\mathbf{X})$. Sei \mathbf{A} eine symmetrische Matrix und bezeichne $\text{Sp}(\cdot)$ die Spur einer Matrix. Dann erhält man den

Erwartungswert der quadratischen Form $Q = \mathbf{X}' \mathbf{A} \mathbf{X}$ durch:

$$E[Q] = \text{Sp}(\mathbf{A}\mathbf{V}) + \boldsymbol{\mu}' \mathbf{A} \boldsymbol{\mu}$$

Satz A.2.3 (Craig-Sakamoto). Sei $\mathbf{X} = (X_1, \dots, X_n)' \sim N(\boldsymbol{\mu}, \mathbf{V})$, \mathbf{A} und \mathbf{B} symmetrische, positiv semidefinite Matrizen und \mathbf{b} ein konstanter Vektor. Dann gilt:

$$\mathbf{B}\mathbf{V}\mathbf{A} = 0 \implies \mathbf{X}' \mathbf{A} \mathbf{X} \text{ und } \mathbf{X}' \mathbf{B} \mathbf{X} \text{ sind stochastisch unabhängig}$$

$$\mathbf{b}' \mathbf{V} \mathbf{A} = 0 \implies \mathbf{A} \mathbf{X} \text{ und } \mathbf{b}' \mathbf{X} \text{ sind stochastisch unabhängig}$$

A.3 Ungleichungen

Satz A.3.1 (c_r -Ungleichung). Für beliebige Zufallsvariablen X, Y gilt:

$$E(|X + Y|^r) \leq c_r \cdot [E(|X|^r) + E(|Y|^r)]$$

mit

$$c_r = \begin{cases} 1, & \text{für } 0 < r \leq 1 \\ 2^{r-1}, & \text{für } r > 1 \end{cases}$$

Satz A.3.2 (Jensen-Ungleichung). Sei X eine Zufallsvariable mit endlichem Erwartungswert $E[X] < \infty$ und $g(\cdot)$ eine konvexe Funktion. Dann gilt:

$$g(E[X]) \leq E[g(X)]$$

A.4 Konvergenzarten

Definition A.4.1 (p-Konvergenz). Eine Zufallsvariable X_n konvergiert in Wahrscheinlichkeit $X_n \xrightarrow{p} X$, falls für alle $\epsilon > 0$ gilt:

$$P(|X_n - X| \geq \epsilon) \xrightarrow{n \rightarrow \infty} 0$$

Definition A.4.2 (\mathcal{L}_p -Konvergenz). Eine Zufallsvariable X_n konvergiert im p -ten Mittel, falls für $p > 0$ gilt:

$$E[X_n - X]^p \xrightarrow{n \rightarrow \infty} 0$$

Satz A.4.3. Es gilt die Implikation:

$$X_n \xrightarrow{\mathcal{L}_p} X \implies X_n \xrightarrow{p} X$$

Satz A.4.4 (Slutzky). Seien X_n und Y_n zwei Folgen von Zufallsvariablen mit $X_n \xrightarrow{p} X$ und $Y_n \xrightarrow{p} b$ mit $0 \neq b < \infty$. Dann folgt:

$$\begin{aligned} X_n + Y_n &\xrightarrow{p} X + b \\ X_n \cdot Y_n &\xrightarrow{p} X \cdot b \\ \frac{X_n}{Y_n} &\xrightarrow{p} \frac{X}{b} \end{aligned}$$

Zentraler Grenzwertsatz (Lindeberg-Levy). Seien $X_n \stackrel{uiv}{\sim} F(x)$ Zufallsvariablen mit $E[X] = \mu$ und $\text{Var}[X] = \sigma^2 < \infty$. Dann gilt:

$$\lim_{n \rightarrow \infty} P\left(\frac{\bar{X}_n - \mu}{\sigma} \sqrt{n} \leq z\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{1}{2}x^2} dx$$

Definition A.4.5 (Asymptotische Äquivalenz). Die Folgen von Zufallsvariablen X_n und Y_n heißen asymptotisch äquivalent, $X_n \doteq Y_n$, falls

$$X_n - Y_n \xrightarrow{p} 0.$$

B Tabellen und Abbildungen

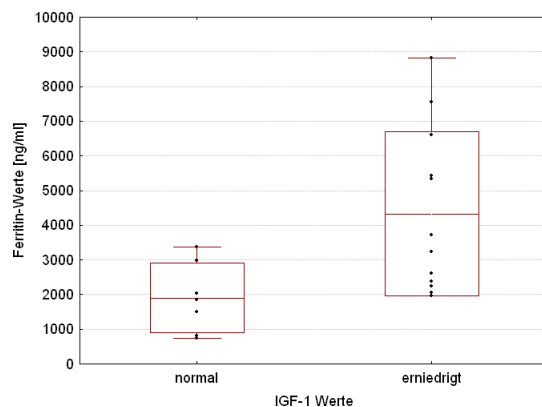


Abbildung 1: Boxplots zu den gemessenen Ferritin-Werten für $N = 19$ Patienten, eingeteilt in die Gruppen *IGF-1 erniedrigt* und *IGF-1 normal*

Tabelle 1: Messwerte, Intern- und Gesamt-Ränge des Ferritin-Beispiels.

| IGF-1 gering | | | IGF-1 normal | | |
|--------------|----------|----------------|--------------|----------|----------------|
| Ferritin | R_{ik} | $R_{ik}^{(i)}$ | Ferritin | R_{ik} | $R_{ik}^{(i)}$ |
| 820 | 2 | 2 | 1956 | 5 | 1 |
| 3364 | 13 | 7 | 8828 | 19 | 12 |
| 1497 | 3 | 3 | 2051 | 7 | 2 |
| 1851 | 4 | 4 | 3721 | 14 | 7 |
| 2984 | 11 | 6 | 3233 | 12 | 6 |
| 744 | 1 | 1 | 6606 | 17 | 10 |
| 2044 | 6 | 5 | 2244 | 8 | 3 |
| | | | 5332 | 15 | 8 |
| | | | 5428 | 16 | 9 |
| | | | 2603 | 10 | 5 |
| | | | 2370 | 9 | 4 |
| | | | 7564 | 18 | 11 |

Tabelle 2: Vergleich zweier Normalverteilungen mit identischen Erwartungswerten und verschiedenen Varianzen.

| Parameter | | Schätzer | | Niveau | Tests | | | |
|-----------|-----------|-----------------|----------------|--------|--------|--------|--------|--------|
| n_1/n_2 | v_1/v_2 | \hat{f}_{Atq} | \hat{f}_{BM} | | Gauss | t-Test | ATQ | BM |
| 15/10 | 1/1 | 56 | 18 | 0.10 | 0.0988 | 0.0994 | 0.1039 | 0.1007 |
| | | | | 0.05 | 0.0490 | 0.0496 | 0.0555 | 0.0556 |
| | | | | 0.01 | 0.0099 | 0.0097 | 0.0139 | 0.0158 |
| 15/10 | 5/1 | 107 | 19 | 0.10 | 0.1000 | 0.0644 | 0.1055 | 0.0986 |
| | | | | 0.05 | 0.0491 | 0.0285 | 0.0561 | 0.0532 |
| | | | | 0.01 | 0.0101 | 0.0048 | 0.0120 | 0.0144 |
| 15/10 | 1/5 | 93 | 11 | 0.10 | 0.1010 | 0.1572 | 0.1053 | 0.1017 |
| | | | | 0.05 | 0.0507 | 0.0932 | 0.0517 | 0.0577 |
| | | | | 0.01 | 0.0106 | 0.0283 | 0.0119 | 0.0184 |
| 15/10 | 10/1 | 107 | 19 | 0.10 | 0.0999 | 0.0586 | 0.1053 | 0.0969 |
| | | | | 0.05 | 0.0498 | 0.0257 | 0.0543 | 0.0506 |
| | | | | 0.01 | 0.0103 | 0.0043 | 0.0107 | 0.0141 |
| 15/10 | 1/10 | 107 | 10 | 0.10 | 0.0983 | 0.1743 | 0.1041 | 0.1002 |
| | | | | 0.05 | 0.0492 | 0.1060 | 0.0466 | 0.0536 |
| | | | | 0.01 | 0.0099 | 0.0355 | 0.0084 | 0.0182 |
| 20/10 | 1/1 | 57 | 19 | 0.10 | 0.0991 | 0.0995 | 0.1035 | 0.1003 |
| | | | | 0.05 | 0.0495 | 0.0504 | 0.0562 | 0.0550 |
| | | | | 0.01 | 0.0098 | 0.0096 | 0.0147 | 0.0156 |
| 20/10 | 5/1 | 57 | 19 | 0.10 | 0.0976 | 0.0417 | 0.1035 | 0.0994 |
| | | | | 0.05 | 0.0483 | 0.0158 | 0.0549 | 0.0513 |
| | | | | 0.01 | 0.0097 | 0.0019 | 0.0128 | 0.0127 |
| 20/10 | 1/5 | 92 | 11 | 0.10 | 0.1010 | 0.1958 | 0.1048 | 0.1010 |
| | | | | 0.05 | 0.0499 | 0.1231 | 0.0506 | 0.0563 |
| | | | | 0.01 | 0.0095 | 0.0431 | 0.0116 | 0.0175 |
| 20/10 | 10/1 | 185 | 24 | 0.10 | 0.1001 | 0.0346 | 0.1056 | 0.0996 |
| | | | | 0.05 | 0.0503 | 0.0129 | 0.0570 | 0.0522 |
| | | | | 0.01 | 0.0101 | 0.0013 | 0.0137 | 0.0132 |
| 20/10 | 1/10 | 107 | 10 | 0.10 | 0.1007 | 0.2262 | 0.1062 | 0.1023 |
| | | | | 0.05 | 0.0502 | 0.1494 | 0.0477 | 0.0551 |
| | | | | 0.01 | 0.0101 | 0.0589 | 0.0086 | 0.0188 |

Tabelle 3: Vergleich zweier Normalverteilungen mit identischen Erwartungswerten und verschiedenen Varianzen.

| Parameter | | Schätzer | | Niveau | Tests | | | |
|-----------|-----------|-----------------|----------------|--------|--------|--------|--------|--------|
| n_1/n_2 | v_1/v_2 | \hat{f}_{Atq} | \hat{f}_{BM} | | Gauss | t-Test | ATQ | BM |
| 30/10 | 1/1 | 54 | 17 | 0.10 | 0.1001 | 0.1015 | 0.1061 | 0.1026 |
| | | | | 0.05 | 0.0507 | 0.0506 | 0.0565 | 0.0561 |
| | | | | 0.01 | 0.0104 | 0.0102 | 0.0169 | 0.0168 |
| 30/10 | 5/1 | 121 | 33 | 0.10 | 0.1000 | 0.0224 | 0.1024 | 0.0987 |
| | | | | 0.05 | 0.0495 | 0.0069 | 0.0555 | 0.0524 |
| | | | | 0.01 | 0.0100 | 0.0004 | 0.0140 | 0.0129 |
| 30/10 | 1/5 | 90 | 10 | 0.10 | 0.1000 | 0.2498 | 0.1053 | 0.1003 |
| | | | | 0.05 | 0.0492 | 0.1694 | 0.0493 | 0.0555 |
| | | | | 0.01 | 0.0094 | 0.0711 | 0.0121 | 0.0173 |
| 30/10 | 10/1 | 214 | 35 | 0.10 | 0.1009 | 0.0141 | 0.1048 | 0.1001 |
| | | | | 0.05 | 0.0505 | 0.0036 | 0.0564 | 0.0526 |
| | | | | 0.01 | 0.0105 | 0.0002 | 0.0138 | 0.0126 |
| 30/10 | 1/10 | 107 | 9 | 0.10 | 0.0988 | 0.2985 | 0.1066 | 0.1020 |
| | | | | 0.05 | 0.0492 | 0.2151 | 0.0467 | 0.0548 |
| | | | | 0.01 | 0.0097 | 0.1028 | 0.0085 | 0.0184 |
| 30/15 | 1/1 | 74 | 29 | 0.10 | 0.0987 | 0.0984 | 0.1032 | 0.0992 |
| | | | | 0.05 | 0.0492 | 0.0491 | 0.0549 | 0.0527 |
| | | | | 0.01 | 0.0099 | 0.0100 | 0.0134 | 0.0133 |
| 30/15 | 5/1 | 160 | 40 | 0.10 | 0.0982 | 0.0413 | 0.1029 | 0.0989 |
| | | | | 0.05 | 0.0494 | 0.0159 | 0.0549 | 0.0517 |
| | | | | 0.01 | 0.0099 | 0.0019 | 0.0133 | 0.0122 |
| 30/15 | 1/5 | 120 | 17 | 0.10 | 0.0998 | 0.1950 | 0.1090 | 0.1009 |
| | | | | 0.05 | 0.0497 | 0.1225 | 0.0583 | 0.0535 |
| | | | | 0.01 | 0.0099 | 0.0423 | 0.0120 | 0.0143 |
| 30/15 | 10/1 | 264 | 37 | 0.10 | 0.0989 | 0.0325 | 0.1025 | 0.0972 |
| | | | | 0.05 | 0.0489 | 0.0118 | 0.0540 | 0.0497 |
| | | | | 0.01 | 0.0099 | 0.0010 | 0.0132 | 0.0118 |
| 30/15 | 1/10 | 160 | 15 | 0.10 | 0.0997 | 0.2201 | 0.1073 | 0.0979 |
| | | | | 0.05 | 0.0495 | 0.1446 | 0.0578 | 0.0527 |
| | | | | 0.01 | 0.0099 | 0.0563 | 0.0107 | 0.0144 |

Tabelle 4: Vergleich zwischen zwei stetigen Gleichverteilungen mit identischem Erwartungswert $\mu = 1/2$ und unterschiedlichen Varianzen.

| Parameter | | | Schätzer | | Niveau | Tests | | | |
|-----------|-----------|-----------|-----------------|----------------|--------|--------|--------|--------|--------|
| n_1/n_2 | a_1/b_1 | a_2/b_2 | \hat{f}_{Atq} | \hat{f}_{BM} | | Gauss | t-Test | ATQ | BM |
| 10/15 | 0/1 | 0/1 | 56 | 18 | 0.10 | 0.1003 | 0.0994 | 0.1024 | 0.0991 |
| | | | | | 0.05 | 0.0484 | 0.0502 | 0.0544 | 0.0548 |
| | | | | | 0.01 | 0.0095 | 0.0114 | 0.0146 | 0.0165 |
| 10/15 | -1/2 | 0/1 | 102 | 10 | 0.10 | 0.0998 | 0.1716 | 0.1060 | 0.1018 |
| | | | | | 0.05 | 0.0484 | 0.1071 | 0.0450 | 0.0515 |
| | | | | | 0.01 | 0.0081 | 0.0390 | 0.0077 | 0.0179 |
| 10/15 | 0/1 | -1/2 | 173 | 16 | 0.10 | 0.1012 | 0.0614 | 0.1092 | 0.0999 |
| | | | | | 0.05 | 0.0500 | 0.0289 | 0.0572 | 0.0530 |
| | | | | | 0.01 | 0.0089 | 0.0057 | 0.0104 | 0.0142 |
| 10/15 | -10/11 | 0/1 | 26 | 9 | 0.10 | 0.0992 | 0.1918 | 0.1098 | 0.1066 |
| | | | | | 0.05 | 0.0476 | 0.1260 | 0.0123 | 0.0271 |
| | | | | | 0.01 | 0.0085 | 0.0515 | 0.0019 | 0.0200 |
| 10/15 | 0/1 | -10/11 | 211 | 14 | 0.10 | 0.1000 | 0.0548 | 0.1212 | 0.1153 |
| | | | | | 0.05 | 0.0493 | 0.0259 | 0.0470 | 0.0447 |
| | | | | | 0.01 | 0.0093 | 0.0058 | 0.0083 | 0.0099 |
| 10/20 | 0/1 | 0/1 | 58 | 19 | 0.10 | 0.0986 | 0.0978 | 0.1035 | 0.1008 |
| | | | | | 0.05 | 0.0487 | 0.0497 | 0.0552 | 0.0549 |
| | | | | | 0.01 | 0.0088 | 0.0104 | 0.0154 | 0.0162 |
| 10/20 | -1/2 | 0/1 | 99 | 9 | 0.10 | 0.1021 | 0.2235 | 0.1087 | 0.1042 |
| | | | | | 0.05 | 0.0496 | 0.1497 | 0.0461 | 0.0536 |
| | | | | | 0.01 | 0.0089 | 0.0630 | 0.0083 | 0.0191 |
| 10/20 | 0/1 | -1/2 | 230 | 22 | 0.10 | 0.0998 | 0.0354 | 0.1060 | 0.0983 |
| | | | | | 0.05 | 0.0486 | 0.0137 | 0.0554 | 0.0503 |
| | | | | | 0.01 | 0.0090 | 0.0019 | 0.0129 | 0.0125 |
| 10/20 | -10/11 | 0/1 | 24 | 9 | 0.10 | 0.0998 | 0.2599 | 0.1123 | 0.1098 |
| | | | | | 0.05 | 0.0490 | 0.1831 | 0.0127 | 0.0288 |
| | | | | | 0.01 | 0.0086 | 0.0884 | 0.0023 | 0.0216 |
| 10/20 | 0/1 | -10/11 | 177 | 19 | 0.10 | 0.0995 | 0.0275 | 0.1156 | 0.1055 |
| | | | | | 0.05 | 0.0499 | 0.0101 | 0.0505 | 0.0475 |
| | | | | | 0.01 | 0.0093 | 0.0013 | 0.0128 | 0.0126 |

Tabelle 5: Vergleich zwischen zwei stetigen Gleichverteilungen mit identischem Erwartungswert $\mu = 1/2$ und unterschiedlichen Varianzen.

| Parameter | | | Schätzer | | Niveau | Tests | | | |
|-----------|-----------|-----------|-----------------|----------------|--------|--------|--------|--------|--------|
| n_1/n_2 | a_1/b_1 | a_2/b_2 | \hat{f}_{Atq} | \hat{f}_{BM} | | Gauss | t-Test | ATQ | BM |
| 10/30 | 0/1 | 0/1 | 54 | 17 | 0.10 | 0.1008 | 0.1008 | 0.1063 | 0.1028 |
| | | | | | 0.05 | 0.0488 | 0.0502 | 0.0573 | 0.0565 |
| | | | | | 0.01 | 0.0091 | 0.0104 | 0.0173 | 0.0171 |
| 10/30 | -1/2 | 0/1 | 98 | 9 | 0.10 | 0.1009 | 0.2935 | 0.1075 | 0.1024 |
| | | | | | 0.05 | 0.0491 | 0.2114 | 0.0454 | 0.0533 |
| | | | | | 0.01 | 0.0082 | 0.1039 | 0.0084 | 0.0185 |
| 10/30 | 0/1 | -1/2 | 313 | 34 | 0.10 | 0.1002 | 0.0143 | 0.1039 | 0.0987 |
| | | | | | 0.05 | 0.0507 | 0.0043 | 0.0559 | 0.0516 |
| | | | | | 0.01 | 0.0097 | 0.0003 | 0.0133 | 0.0115 |
| 10/30 | -10/11 | 0/1 | 98 | 9 | 0.10 | 0.1012 | 0.3557 | 0.1119 | 0.1089 |
| | | | | | 0.05 | 0.0497 | 0.2735 | 0.0129 | 0.0286 |
| | | | | | 0.01 | 0.0081 | 0.1581 | 0.0021 | 0.0212 |
| 10/30 | 0/1 | -10/11 | 382 | 29 | 0.10 | 0.0991 | 0.0070 | 0.1052 | 0.0992 |
| | | | | | 0.05 | 0.0496 | 0.0018 | 0.0525 | 0.0493 |
| | | | | | 0.01 | 0.0090 | 0.0001 | 0.0147 | 0.0126 |
| 15/30 | 0/1 | 0/1 | 74 | 29 | 0.10 | 0.0986 | 0.0994 | 0.1031 | 0.0996 |
| | | | | | 0.05 | 0.0495 | 0.0497 | 0.0556 | 0.0534 |
| | | | | | 0.01 | 0.0089 | 0.0103 | 0.0136 | 0.0135 |
| 15/30 | -1/2 | 0/1 | 168 | 15 | 0.10 | 0.1004 | 0.2173 | 0.1097 | 0.1088 |
| | | | | | 0.05 | 0.0501 | 0.1443 | 0.0595 | 0.0578 |
| | | | | | 0.01 | 0.0095 | 0.0582 | 0.0102 | 0.0148 |
| 15/30 | 0/1 | -1/2 | 335 | 34 | 0.10 | 0.0997 | 0.0341 | 0.1050 | 0.0990 |
| | | | | | 0.05 | 0.0496 | 0.0129 | 0.0561 | 0.0512 |
| | | | | | 0.01 | 0.0097 | 0.0016 | 0.0140 | 0.0122 |
| 15/30 | -10/11 | 0/1 | 213 | 14 | 0.10 | 0.1028 | 0.2554 | 0.1198 | 0.1134 |
| | | | | | 0.05 | 0.0508 | 0.1799 | 0.0466 | 0.0444 |
| | | | | | 0.01 | 0.0095 | 0.0833 | 0.0088 | 0.0106 |
| 15/30 | 0/1 | -10/11 | 381 | 29 | 0.10 | 0.1003 | 0.0250 | 0.1057 | 0.0999 |
| | | | | | 0.05 | 0.0494 | 0.0089 | 0.0523 | 0.0495 |
| | | | | | 0.01 | 0.0097 | 0.0009 | 0.0148 | 0.0127 |

Tabelle 6: Vergleich zwischen einer Gammaverteilung mit Shape-Parameter k und einer Normalverteilung, deren Erwartungswert mittels Simulation auf den Wert μ_2 verschoben wird, sodass $p \approx 0.50$.

| Parameter | | | Schätzer | | Niveau | Tests | |
|-----------|-----|-------------|-----------------|----------------|--------|--------|--------|
| n_1/n_2 | k | μ_2/v_2 | \hat{f}_{Atq} | \hat{f}_{BM} | | ATQ | BM |
| 10/15 | 0.1 | 0.085/1 | > 500 | > 500 | 0.10 | 0.1149 | 0.1070 |
| | | | | | 0.05 | 0.0543 | 0.0500 |
| | | | | | 0.01 | 0.0100 | 0.0123 |
| 10/15 | 1 | 0.087/1 | > 500 | > 500 | 0.10 | 0.1026 | 0.0990 |
| | | | | | 0.05 | 0.0549 | 0.0544 |
| | | | | | 0.01 | 0.0132 | 0.0155 |
| 10/15 | 10 | 9.69/1 | > 500 | > 500 | 0.10 | 0.1049 | 0.1010 |
| | | | | | 0.05 | 0.0473 | 0.0534 |
| | | | | | 0.01 | 0.0090 | 0.0176 |
| 15/10 | 0.1 | 0.086/1 | > 500 | > 500 | 0.10 | 0.1084 | 0.1074 |
| | | | | | 0.05 | 0.0308 | 0.0447 |
| | | | | | 0.01 | 0.0024 | 0.0218 |
| 15/10 | 1 | 0.087/1 | > 500 | > 500 | 0.10 | 0.1038 | 0.1021 |
| | | | | | 0.05 | 0.0544 | 0.0570 |
| | | | | | 0.01 | 0.0156 | 0.0170 |
| 15/10 | 10 | 9.69/1 | > 500 | > 500 | 0.10 | 0.1066 | 0.0982 |
| | | | | | 0.05 | 0.0555 | 0.0518 |
| | | | | | 0.01 | 0.0110 | 0.0141 |
| 10/20 | 0.1 | 0.86/1 | > 500 | > 500 | 0.10 | 0.1138 | 0.1071 |
| | | | | | 0.05 | 0.0560 | 0.0519 |
| | | | | | 0.01 | 0.0144 | 0.0137 |
| 10/20 | 1 | 0.87/1 | > 500 | > 500 | 0.10 | 0.1030 | 0.1004 |
| | | | | | 0.05 | 0.0551 | 0.0538 |
| | | | | | 0.01 | 0.0144 | 0.0145 |
| 10/20 | 10 | 9.68/1 | > 500 | > 500 | 0.10 | 0.1049 | 0.1004 |
| | | | | | 0.05 | 0.0466 | 0.0537 |
| | | | | | 0.01 | 0.0084 | 0.0183 |
| 20/10 | 0.1 | 0.86/1 | > 500 | > 500 | 0.10 | 0.1073 | 0.1051 |
| | | | | | 0.05 | 0.0299 | 0.0445 |
| | | | | | 0.01 | 0.0026 | 0.0211 |
| 20/10 | 1 | 0.87/1 | > 500 | > 500 | 0.10 | 0.1048 | 0.1033 |
| | | | | | 0.05 | 0.0544 | 0.0570 |
| | | | | | 0.01 | 0.0168 | 0.0172 |
| 20/10 | 10 | 9.68/1 | > 500 | > 500 | 0.10 | 0.1041 | 0.0982 |
| | | | | | 0.05 | 0.0562 | 0.0517 |
| | | | | | 0.01 | 0.0133 | 0.0128 |

Tabelle 7: Vergleich zwischen einer Gammaverteilung mit Shape-Parameter k und einer Normalverteilung, deren Erwartungswert mittels Simulation auf den Wert μ_2 verschoben wird, sodass $p \approx 0.50$.

| Parameter | | | Schätzer | | Niveau | Tests | |
|-----------|-----|-------------|-----------------|----------------|--------|--------|--------|
| n_1/n_2 | k | μ_2/v_2 | \hat{f}_{Atq} | \hat{f}_{BM} | | ATQ | BM |
| 10/30 | 0.1 | 0.081/1 | > 500 | > 500 | 0.10 | 0.1052 | 0.0998 |
| | | | | | 0.05 | 0.0547 | 0.0514 |
| | | | | | 0.01 | 0.0154 | 0.0130 |
| 10/30 | 1 | 0.087/1 | > 500 | > 500 | 0.10 | 0.1048 | 0.1024 |
| | | | | | 0.05 | 0.0572 | 0.0550 |
| | | | | | 0.01 | 0.0168 | 0.0162 |
| 10/30 | 10 | 9.7/1 | > 500 | > 500 | 0.10 | 0.1074 | 0.1028 |
| | | | | | 0.05 | 0.0468 | 0.0551 |
| | | | | | 0.01 | 0.0088 | 0.0185 |
| 30/10 | 0.1 | 0.081/1 | > 500 | > 500 | 0.10 | 0.1082 | 0.1076 |
| | | | | | 0.05 | 0.0240 | 0.0478 |
| | | | | | 0.01 | 0.0025 | 0.0217 |
| 30/10 | 1 | 0.87/1 | > 500 | > 500 | 0.10 | 0.1046 | 0.1031 |
| | | | | | 0.05 | 0.0547 | 0.0573 |
| | | | | | 0.01 | 0.0180 | 0.0181 |
| 30/10 | 10 | 9.7/1 | > 500 | > 500 | 0.10 | 0.1029 | 0.0983 |
| | | | | | 0.05 | 0.0541 | 0.0505 |
| | | | | | 0.01 | 0.0127 | 0.0116 |
| 15/30 | 0.1 | 0.086/1 | > 500 | > 500 | 0.10 | 0.1075 | 0.1010 |
| | | | | | 0.05 | 0.0553 | 0.0517 |
| | | | | | 0.01 | 0.0140 | 0.0117 |
| 15/30 | 1 | 0.87/1 | > 500 | > 500 | 0.10 | 0.1025 | 0.1000 |
| | | | | | 0.05 | 0.0541 | 0.0525 |
| | | | | | 0.01 | 0.0139 | 0.0136 |
| 15/30 | 10 | 9.69/1 | > 500 | > 500 | 0.10 | 0.1085 | 0.0990 |
| | | | | | 0.05 | 0.0577 | 0.0527 |
| | | | | | 0.01 | 0.0109 | 0.0142 |
| 30/15 | 0.1 | 0.086/1 | > 500 | > 500 | 0.10 | 0.1123 | 0.0984 |
| | | | | | 0.05 | 0.0598 | 0.0550 |
| | | | | | 0.01 | 0.0070 | 0.0131 |
| 30/15 | 1 | 0.87/1 | > 500 | > 500 | 0.10 | 0.1052 | 0.1006 |
| | | | | | 0.05 | 0.0553 | 0.0530 |
| | | | | | 0.01 | 0.0137 | 0.0146 |
| 30/15 | 10 | 9.69/1 | > 500 | > 500 | 0.10 | 0.1050 | 0.0995 |
| | | | | | 0.05 | 0.0566 | 0.0524 |
| | | | | | 0.01 | 0.0130 | 0.0117 |

Tabelle 8: Vergleich von nicht-zentraler Chi_3^2 -Verteilung mit Normalverteilung für verschiedene Nichtzentralitätsparameter und Varianzen.

| Parameter | | | Schätzer | | Niveau | Tests | |
|-----------|---------------|-------|-----------------|----------------|--------|--------|--------|
| n_1/n_2 | μ_1/μ_2 | v_2 | \hat{f}_{Atq} | \hat{f}_{BM} | | ATQ | BM |
| 10/15 | 0/2.49 | 1 | > 500 | > 500 | 0.10 | 0.1047 | 0.1010 |
| | | | | | 0.05 | 0.0505 | 0.0555 |
| | | | | | 0.01 | 0.0112 | 0.0172 |
| 10/15 | 5/76 | 1 | > 500 | > 500 | 0.10 | 0.1111 | 0.1079 |
| | | | | | 0.05 | 0.0169 | 0.0328 |
| | | | | | 0.01 | 0.0022 | 0.0212 |
| 10/15 | 25/1876 | 1 | > 500 | > 500 | 0.10 | 0.1101 | 0.1091 |
| | | | | | 0.05 | 0.0055 | 0.0238 |
| | | | | | 0.01 | 0.0019 | 0.0210 |
| 15/10 | 0/2.49 | 1 | > 500 | > 500 | 0.10 | 0.1071 | 0.1009 |
| | | | | | 0.05 | 0.0572 | 0.0548 |
| | | | | | 0.01 | 0.0124 | 0.0147 |
| 15/10 | 5/76 | 1 | > 500 | > 500 | 0.10 | 0.1201 | 0.1133 |
| | | | | | 0.05 | 0.0507 | 0.0459 |
| | | | | | 0.01 | 0.0082 | 0.0105 |
| 15/10 | 25/1876 | 1 | > 500 | > 500 | 0.10 | 0.1183 | 0.1163 |
| | | | | | 0.05 | 0.0397 | 0.0380 |
| | | | | | 0.01 | 0.0071 | 0.0080 |
| 10/20 | 0/2,48 | 1 | > 500 | > 500 | 0.10 | 0.1062 | 0.1023 |
| | | | | | 0.05 | 0.0512 | 0.0564 |
| | | | | | 0.01 | 0.0126 | 0.0181 |
| 10/20 | 5/76 | 1 | > 500 | > 500 | 0.10 | 0.1119 | 0.1090 |
| | | | | | 0.05 | 0.0175 | 0.0343 |
| | | | | | 0.01 | 0.0024 | 0.0213 |
| 10/20 | 25/1876 | 1 | > 500 | > 500 | 0.10 | 0.1097 | 0.1089 |
| | | | | | 0.05 | 0.0050 | 0.0239 |
| | | | | | 0.01 | 0.0019 | 0.0213 |
| 20/10 | 0/2,48 | 1 | > 500 | > 500 | 0.10 | 0.1036 | 0.0996 |
| | | | | | 0.05 | 0.0564 | 0.0530 |
| | | | | | 0.01 | 0.0143 | 0.0140 |
| 20/10 | 5/76 | 1 | > 500 | > 500 | 0.10 | 0.1164 | 0.1067 |
| | | | | | 0.05 | 0.0549 | 0.0501 |
| | | | | | 0.01 | 0.0134 | 0.0131 |
| 20/10 | 25/1876 | 1 | > 500 | > 500 | 0.10 | 0.1150 | 0.1111 |
| | | | | | 0.05 | 0.0439 | 0.0424 |
| | | | | | 0.01 | 0.0117 | 0.0118 |

Tabelle 9: Vergleich von nicht-zentraler Chi_3^2 -Verteilung mit Normalverteilung für verschiedene Nichtzentralitätsparameter und Varianzen.

| Parameter | | | Schätzer | | Niveau | Tests | |
|-----------|---------------|-------|-----------------|----------------|--------|--------|--------|
| n_1/n_2 | μ_1/μ_2 | v_2 | \hat{f}_{Atq} | \hat{f}_{BM} | | ATQ | BM |
| 10/30 | 0/2.49 | 1 | > 500 | > 500 | 0.10 | 0.1059 | 0.1012 |
| | | | | | 0.05 | 0.0508 | 0.0563 |
| | | | | | 0.01 | 0.0124 | 0.0179 |
| 10/30 | 5/76 | 1 | > 500 | > 500 | 0.10 | 0.1120 | 0.1090 |
| | | | | | 0.05 | 0.0175 | 0.0348 |
| | | | | | 0.01 | 0.0024 | 0.0216 |
| 10/30 | 25/1876 | 1 | > 500 | > 500 | 0.10 | 0.1106 | 0.1097 |
| | | | | | 0.05 | 0.0056 | 0.0250 |
| | | | | | 0.01 | 0.0021 | 0.0218 |
| 30/10 | 0/2.49 | 1 | > 500 | > 500 | 0.10 | 0.1026 | 0.0993 |
| | | | | | 0.05 | 0.0559 | 0.0530 |
| | | | | | 0.01 | 0.0145 | 0.0134 |
| 30/10 | 5/76 | 1 | > 500 | > 500 | 0.10 | 0.1047 | 0.0975 |
| | | | | | 0.05 | 0.0541 | 0.0499 |
| | | | | | 0.01 | 0.0136 | 0.0115 |
| 30/10 | 25/1876 | 1 | > 500 | > 500 | 0.10 | 0.1024 | 0.0998 |
| | | | | | 0.05 | 0.0479 | 0.0465 |
| | | | | | 0.01 | 0.0163 | 0.0151 |
| 15/30 | 0/2,48 | 1 | > 500 | > 500 | 0.10 | 0.1087 | 0.1011 |
| | | | | | 0.05 | 0.0585 | 0.0546 |
| | | | | | 0.01 | 0.0121 | 0.0144 |
| 15/30 | 5/77 | 1 | > 500 | > 500 | 0.10 | 0.1173 | 0.1099 |
| | | | | | 0.05 | 0.0494 | 0.0461 |
| | | | | | 0.01 | 0.0081 | 0.0109 |
| 15/30 | 25/1876 | 1 | > 500 | > 500 | 0.10 | 0.1190 | 0.1172 |
| | | | | | 0.05 | 0.0400 | 0.0388 |
| | | | | | 0.01 | 0.0065 | 0.0076 |
| 30/15 | 0/2,48 | 1 | > 500 | > 500 | 0.10 | 0.1038 | 0.1004 |
| | | | | | 0.05 | 0.0545 | 0.0516 |
| | | | | | 0.01 | 0.0135 | 0.0127 |
| 30/15 | 5/77 | 1 | > 500 | > 500 | 0.10 | 0.1077 | 0.1009 |
| | | | | | 0.05 | 0.0547 | 0.0505 |
| | | | | | 0.01 | 0.0151 | 0.0126 |
| 30/15 | 25/1876 | 1 | > 500 | > 500 | 0.10 | 0.1027 | 0.1008 |
| | | | | | 0.05 | 0.0471 | 0.0460 |
| | | | | | 0.01 | 0.0160 | 0.0149 |

C SAS-Code

```

proc iml;
start mh(n1, n2, v1, mu2, v2, nsim, schalter);
gn = n1 + n2;
s1 = sqrt(1);
s2 = sqrt(v2);
pd = 0;
quantil = {0.90 0.95 0.99};
quant_calc = {0.95 0.975 0.995};
qnv = probit(quant_calc);
qt = tinv(quant_calc, gn-2);
tests = {"Gauss", "t-Test", "Rang-/ART", "t-Approx"};
nquant = ncol(quantil);
simerg = j(6,3,0);
fd_vz_count = 0;
fehler_vz = 0;
fd_count = 0;
fehler = 0;
fd_alt_count = 0;
fehler_alt = 0;
fd_c = 0;
fd_vz_c = 0;
fd_alt_c = 0;
**Simulationsschleife;
DO isim=1 to nsim;
**Datensaetze erzeugen;
x1 = j(n1,1,0);
DO ij1 =1 to n1 ;
x1[ ij1 ] = RAND( 'Gamma',v1);
end ;
x2 =j(n2,1,mu2) + rannor(j(n2, 1, 1)) # j(n2, 1, s2);
**Berechnungen;
**Mittelwerte;
x1pq = x1[+]/n1;

```

```

x2pq = x2[+]/n2;
**Varianzen;
*Normalverteilungsmodell;
var_nv = 1/n1+v2/n2;
std_nv = sqrt(var_nv);
*t-Test;
s0sq = 1/(gn-2)*(ssq(x1-j(n1,1,x1pq))+ssq(x2-j(n2,1,x2pq)))
;
std_ttest = sqrt(s0sq);
**Formeln aus Makro von Prof. Brunner;
sdqa1 = ssq(x1 - j(n1,1,x1pq));
sdqa2 = ssq(x2 - j(n2,1,x2pq));
***** Unbekannte Varianzen - SSW;
vssw = sdqa1/(n1*(n1-1)) + sdqa2/(n2*(n2-1));
ssw = (x1pq - x2pq) / sqrt(vssw);
s1q=sdqa1/(n1-1);
s2q=sdqa2/(n2-1);
df = vssw*vssw / ( (S1q/n1)**2/(n1-1)+ (S2q/n2)**2/(n2
-1));
qssw = tinv(quant_calc,df);
**Formeln aus Makro Ende;
**Raenge;
**Intern-Raenge berechnen;
ir1 = ranktie(x1);
ir2 = ranktie(x2);
**Gesamtraenge berechnen;
vek = j(gn, 1, 0);
vek = x1/x2;
gr = ranktie(vek);
r1 = gr[1:n1];
r2 = gr[n1+1:gn];
r1pq = r1[+]/n1;
r2pq = r2[+]/n2;
**Berechnung von pd;
pd = pd + 1/n2*(r1pq-(n2+1)/2);

```

```

**Asymptotische Rangtransformationen;
**Berechnung der Varianzschaezter der ART;
s1sq = 1/(n1-1)/(n2**2)*ssq(r1-ir1-j(n1,1,(r1pq-(n1+1)/2)))
;
s2sq = 1/(n2-1)/(n1**2)*ssq(r2-ir2-j(n2,1,(r2pq-(n2+1)/2)))
;
**Anpassung für getrennte Datensätze;
s1sq = (s1sq=0)/(n1-1)/(n2**2)*(n1-1)/(4*n2)+(s1sq>0)*s1sq;
s2sq = (s2sq=0)/(n2-1)/(n1**2)*(n2-1)/(4*n1)+(s2sq>0)*s2sq;
**Weitere Berechnungen;
tart_var = gn*(n2*s1sq+n1*s2sq);
tart_std = sqrt(tart_var);
**Varianzschaezter der ART;
std_art = sqrt(tart_var);
**Approximative t-Verteilung;
t11 = (n1*(n1**2-2*n1+3))/((n1-1)*(n1-2)*(n1-3))*1/(n1)/(n2
**4)*((r1-ir1-j(n1,1,(r1pq-(n1+1)/2)))###4)[+];
t12 = -(3*n1*(2*n1-3))/((n1-1)*(n1-2)*(n1-3))*(1/(n1)/(n2**
2)*ssq(r1-ir1-j(n1,1,(r1pq-(n1+1)/2))))**2;
t21 = (n2*(n2**2-2*n2+3))/((n2-1)*(n2-2)*(n2-3))*1/(n2)/(n1
**4)*((r2-ir2-j(n2,1,(r2pq-(n2+1)/2)))###4)[+];
t22 = -(3*n2*(2*n2-3))/((n2-1)*(n2-2)*(n2-3))*(1/(n2)/(n1**
2)*ssq(r2-ir2-j(n2,1,(r2pq-(n2+1)/2))))**2;
mu14d = t11 + t12;
mu24d = t21 + t22;
a11 = (n1*(n1**2-3*n1+3))/((n1-1)*(n1-2)*(n1-3))*(1/(n1)/(
n2**2)*ssq(r1-ir1-j(n1,1,(r1pq-(n1+1)/2))))**2;
a12 = -(n1)/((n1-2)*(n1-3))*1/(n1)/(n2**4)*((r1-ir1-j(n1
,1,(r1pq-(n1+1)/2)))###4)[+];
sig1d_4 = a11 + a12 ;
a21 = (n2*(n2**2-3*n2+3))/((n2-1)*(n2-2)*(n2-3))*(1/(n2)/(
n1**2)*ssq(r2-ir2-j(n2,1,(r2pq-(n2+1)/2))))**2;
a22 = -(n2)/((n2-2)*(n2-3))*1/(n2)/(n1**4)*((r2-ir2-j(n2
,1,(r2pq-(n2+1)/2)))###4)[+];
sig2d_4 = a21 + a22;

```

```

v1d_alt = 2*sig1d_4*(n1-1);
vq1d = (mu14d-3*sig1d_4)*(n1-1)**2/(n1)+v1d_alt ;
v2d_alt = 2*sig2d_4*(n2-1);
vq2d = (mu24d-3*sig2d_4)*(n2-1)**2/(n2)+v2d_alt ;
gd = 1;
**Berechnung der Teststatistiken;
**Gaußtest;
z_nv = (x1pq-x2pq);
gauss = z_nv/std_nv;
**t-Test;
z_ttest = sqrt(n1*n2/gn)*z_nv;
ttest = z_ttest/std_ttest;
**Rang-Teststatistik;
z_rang = sqrt((n1*n2)/gn)*(r2pq-r1pq);
wn = z_rang/std_art ;
**t-Approximation ;
t_app = z_rang/tart_std;
**Ergebnisse ;
simerg[1,1] = simerg[1,1] + ((gauss<=qnv[1])&(gauss>=-qnv
[1]));
simerg[1,2] = simerg[1,2] + ((gauss<=qnv[2])&(gauss>=-qnv
[2]));
simerg[1,3] = simerg[1,3] + ((gauss<=qnv[3])&(gauss>=-qnv
[3]));
simerg[2,1] = simerg[2,1] + ((ttest<=qt[1])&(ttest>=-qt[1])
);
simerg[2,2] = simerg[2,2] + ((ttest<=qt[2])&(ttest>=-qt[2])
);
simerg[2,3] = simerg[2,3] + ((ttest<=qt[3])&(ttest>=-qt[3])
);
fd_vz_z = (2*(n2*s1sq+n1*s2sq)**2);
fd_vz_n = ((n2**2)/((n1-1)**2)*vq1d+(n1**2)/((n2-1)**2)*
vq2d);
IF (fd_vz_n>0) THEN fd_vz = (fd_vz_z>fd_vz_n)*fd_vz_z/fd_vz
_n + (fd_vz_z<fd_vz_n);

```

```

ELSE fd_vz = 1          ;
fehler_vz = fehler_vz + 1*(fd_vz=2);
fd_vz_count = fd_vz_count + 1*(fd_vz=1);
IF (fd_vz>1000) THEN fd_vz = 1000;
ELSE fd_vz = fd_vz      ;
qtapp_vz = tinv(quant_calc , fd_vz);
simerg[3,1] = simerg[3,1] + ((wn<=qtapp_vz[1])&(wn>=qtapp_vz[1]));
simerg[3,2] = simerg[3,2] + ((wn<=qtapp_vz[2])&(wn>=qtapp_vz[2]));
simerg[3,3] = simerg[3,3] + ((wn<=qtapp_vz[3])&(wn>=qtapp_vz[3]));
**fd_alt Ueberpruefung der Grenzen fuer df;
fd_alt_z = ((n2*s1sq+n1*s2sq)**2);
fd_alt_n = ((n2*s1sq)**2/(n1-1)+(n1*s2sq)**2/(n2-1));
IF (fd_alt_n>0) THEN fd_alt = (fd_alt_z>fd_alt_n)*fd_alt_z/
    fd_alt_n + (fd_alt_z<fd_alt_n);
ELSE fd_alt = 1          ;
fehler_alt = fehler_alt + (fd_alt=2)*1;
fd_alt_count = fd_alt_count + (fd_alt=1)*1;
IF (fd_alt>1000) THEN fd_alt = 1000;
ELSE fd_alt = fd_alt;
qtapp_alt = tinv(quant_calc , fd_alt);
simerg[4,1] = simerg[4,1] + ((t_app<=qtapp_alt[1])&(t_app
    >=qtapp_alt[1]));
simerg[4,2] = simerg[4,2] + ((t_app<=qtapp_alt[2])&(t_app
    >=qtapp_alt[2]));
simerg[4,3] = simerg[4,3] + ((t_app<=qtapp_alt[3])&(t_app
    >=qtapp_alt[3]));
**fd Berechnungen und so weiter;
fd_z = (2*(n2**2*sig1d_4+n1**2*sig2d_4+2*n1*n2*s1sq*s2sq));
fd_n = ((n2**2)/((n1-1)**2)*vq1d+(n1**2)/((n2-1)**2)*vq2d);
IF (fd_n>0) THEN fd = (fd_z>fd_n)*fd_z/fd_n + (fd_z<fd_n);
ELSE fd = 1          ;
fehler = fehler + 1*(fd=2);

```

```

fd_count = fd_count + 1*(fd=1);
IF (fd>1000) THEN fd = 1000;
ELSE fd = fd          ;
qtapp = tinv(quant_calc,fd);
simerg[5,1] = simerg[5,1] + ((t_app<=qtapp[1])&(t_app>=
qtapp[1]));
simerg[5,2] = simerg[5,2] + ((t_app<=qtapp[2])&(t_app>=
qtapp[2]));
simerg[5,3] = simerg[5,3] + ((t_app<=qtapp[3])&(t_app>=
qtapp[3]));
**Berechnungen für SSW-Test;
simerg[6,1] = simerg[6,1] + ((ssw<=qssw[1])&(ssw>=qssw[1]
));
simerg[6,2] = simerg[6,2] + ((ssw<=qssw[2])&(ssw>=qssw[2]
));
simerg[6,3] = simerg[6,3] + ((ssw<=qssw[3])&(ssw>=qssw[3]
));
fd_c = fd_c + fd          ;
fd_vz_c = fd_vz_c + fd_vz;
fd_alt_c = fd_alt_c + fd_alt;
end          ;
fd_vz_count = fd_vz_count / nsim;
fd_count = fd_count / nsim;
fd_alt_count = fd_alt_count / nsim;
simerg = simerg/nsim;
simerg = round(simerg,0.0001);
simerg = j(6,1,1)*j(1,3,1)-simerg;
pd = pd / nsim          ;
IF (schalter=0) THEN return (pd);
IF (schalter>0) THEN DO;
**Ausgabe der Daten          ;
print "-----";
print "Auswertung_EIN";
print n1 n2 v1 mu2 v2 nsim          pd;
print fd_c fd_alt_c fd_vz_c;

```

```

print  simerg[colname = {"p_10" "p_5" "p_1"} rowname = {"
      Gauss" "T-Test" "verzerrter_t-Test" "alte_t-Approx" "t-
      Approx_neu" "SSW-Test"}];
print "Auswertung_AUS";
print "_____";
return (5)          ;
end                ;
finish              ;
print "#####";
print "Simulationsanfang_für_Gamma_vs_Normal";
n11=10000;
stp2=10000;
var1=10;
var2=1;
*n11 = {10, 15, 20, 30};
*stp2 = {10, 15, 20, 30};
*var1 = {0.001, 0.1, 1, 10};
*var2 = {1, 5, 10}      ;
DO j1=1 to nrow(stp2);
  DO jk1=1 to nrow(n11);
    DO j2=1 to nrow(var1);
      DO j3=1 to nrow(var2);
        **Anpassung von mu2 für Hypothese;
          mu2 = 0          ;
          nsim = 1        ;
          pd = 0          ;
          i = 1           ;
          print "++++++++++++++++++++";
          print "Anfangswerte_von_pd_und_mu2";
          print pd mu2    ;
          DO while (sqrt((pd-1/2)**2) > 0.0001);
            mu2 = mu2 -(pd<1/2)*(1/2-pd) +(pd>1/2)*(pd-1/2);
            pd = mh(10000, 10000, var1[j2], mu2, var2[j3], 1,
              0);
            i = i+1      ;

```

```

    *start mh(n1, n2, v1, mu2, v2, nsim, schalter);
    end
    print "Justierung von m2 und pd ENDE";
    print pd mu2 ;
    print (n11[jk1]) (stp2[j1]) (var1[j2]) mu2 (var2[j3
    ]) nsim i;
    ohnefkt = mh(n11[jk1], stp2[j1], var1[j2], mu2,
    var2[j3], nsim, 1);
    ohnefkt = mh(stp2[j1], n11[jk1], var1[j2], mu2,
    var2[j3], nsim, 1);
    end
end
end
end
    print "Simulationsende";
    print "#####";
quit

```


Literatur

- [1] G. E. P. Box. 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(2):290–302, 1954.
- [2] E. Brunner and U. Munzel. The nonparametric behrens-fisher problem: Asymptotic theory and a small-sample approximation. *Biometrical Journal*, 42:17–25, 2000.
- [3] P. E. Greenwood and M. S. Nikulin. *A Guide to Chi-Squared Testing*. John Wiley & Sons, 1996.
- [4] K. J. Schroeter. *Verfahren zur Approximation der Gesamtschadenverteilung: Systematisierung, Techniken und Vergleiche*. Verlag Versicherungswirtschaft, 1995.
- [5] G. A. F. Seber. *Linear Regression Analysis*. John Wiley & Sons, 1977.
- [6] E. Brunner und U. Munzel. *Nichtparametrische Datenanalyse*. Springer, 2002.
- [7] B. L. Welch. The significance of the difference between two means when the population variances are unequal. *Biometrika*, 29:350–263, 1938.