Biometrische Methoden zur Planung und Auswertung von Sicherheitsstudien

Habilitationsschrift zur Erlangung der Lehrbefugnis für das Lehrgebiet Biometrie

Dem Fachbereich Statistik der Universität Dortmund vorgelegt von Dr. Dieter Hauschke aus Konstanz

Dortmund 1999
Meiner verstorbenen Mutter Josefina Hauschke in dankbarer Erinnerung gewidmet.

Mein Dank gilt Herrn Prof. Dr. Hothorn für die Anregung und Betreuung der vorliegenden Arbeit.
Ebenso danke ich Herrn Prof. Dr. Schach, Herrn Prof. Dr. Brunner und Herrn Prof. Dr. Kunert für ihr Interesse, das sie der Arbeit entgegengebracht haben. Für anregende Diskussionen danke ich meinen Kollegen, Herrn Burke, Herrn Diletti und Herrn Dr. Neuhäuser.
Biometrische Methoden zur Planung und Auswertung
von Sicherheitsstudien

Habilitationsschrift
zur Erlangung der Lehrbefugnis
für das Lehrgebiet
Biometrie

Dem Fachbereich Statistik der Universität Dortmund
vorgelegt von
Dr. Dieter Hauschke
aus Konstanz

Dortmund 1999
Inhaltsverzeichnis

1 Einleitung und Übersicht ........................................................................................................ 1

2 Einführung in die Problemstellung ..................................................................................... 4

3 Zweistichprobensituation ................................................................................................... 12
   3.1 Sicherheitsnachweis für die Differenz von Lokationsparametern .................................. 12
      3.1.1 Einseitiges Testproblem .................................................................................. 12
      3.1.2 Zweiseitiges Testproblem ........................................................................... 15
      3.1.3 Gütebestimmung und Fallzahlplanung ......................................................... 17
   3.2 Sicherheitsnachweis für den Quotienten von Lokationsparametern ............................. 25
      3.2.1 Einseitiges Testproblem ................................................................................ 25
      3.2.2 Zweiseitiges Testproblem ........................................................................... 27
      3.2.3 Gütebestimmung und Fallzahlplanung ......................................................... 28

4 Crossover-Situation ............................................................................................................. 38
   4.1 Sicherheitsnachweis für die Differenz von Lokationsparametern .................................. 39
      4.1.1 Einseitiges Testproblem ................................................................................ 40
      4.1.2 Zweiseitiges Testproblem ........................................................................... 44
      4.1.3 Gütebestimmung und Fallzahlplanung ......................................................... 45
   4.2 Sicherheitsnachweis für den Quotienten von Lokationsparametern ............................. 49
      4.2.1 Einseitiges Testproblem ................................................................................ 49
      4.2.2 Zweiseitiges Testproblem ........................................................................... 54
      4.2.3 Gütebestimmung und Fallzahlplanung ......................................................... 56

5 K-Stichproben many-to-one-Situation .............................................................................. 64
   5.1 Simple-tree Alternative ............................................................................................... 64
      5.1.1 Globaler Sicherheitsnachweis ........................................................................... 64
         5.1.1.1 Einseitiges Testproblem ......................................................................... 64
         5.1.1.2 Zweiseitiges Testproblem ......................................................................... 66
      5.1.2 Partieller Sicherheitsnachweis ............................................................................. 68
         5.1.2.1 Einseitiges Testproblem ........................................................................... 68
         5.1.2.2 Zweiseitiges Testproblem ......................................................................... 74
   5.2 Simple-order Alternative ............................................................................................... 76
      5.2.1 Bestimmung der maximal sicheren Dosis ........................................................... 77
      5.2.1 Alternative Konzepte ....................................................................................... 79
1 Einleitung und Übersicht

Absence of evidence is not evidence of absence (Altman und Bland, 1994)


Weitaus wichtiger ist jedoch die Kontrolle des Konsumentenrisikos, nämlich der irrtümlichen Annahme, daß die Substanz unbedenklich ist. Es stellt sich unmittelbar die Frage, ob ohne Berücksichtigung des Fehlers 2. Art $\beta$ dieses Entscheidungsverfahren korrekt sein kann, zumal in Sicherheitsstudien oftmals eine Begründung der gewählten Fallzahlen nicht auf einer vorgegebenen Güte (*Power*) von $1 - \beta$ basiert, sondern sich statt dessen an behördlichen Mindestanforderungen orientiert. Dabei wird übersehen, daß ein relevanter Unterschied bei zum Beispiel kleinen Fallzahlen und/oder großer Versuchsstreuung zu einem falsch-negativen Testergebnis führen kann; statistische Nichtsignifikanz wäre dann nicht kongruent mit vorliegender Unbedenklichkeit.

Vom ethischen, wissenschaftlichen und regulatorischen Standpunkt her gesehen, muß eine sicherheitsrelevante Studie eine ausreichende Güte aufweisen (Healey, 1987). Dieser Tatsache wurde erstmals in dem von der U.S. Food and Drug Administration (1993) erstellten Entwurf der *Guidance for Toxicity Tests* folgendermaßen Rechnung getragen: *A power calculation should be provided for tests that failed to reject the null hypothesis, particularly to justify the adequacy of the sample size.*

Eine vor Versuchsbeginn vorgenommene Gütebetrachtung dient generell dazu sicherzustellen, daß vorhandenene Effekte mit ausreichender Wahrscheinlichkeit entdeckt werden können. Wird jedoch die entsprechende Nullhypothese nicht verworfen, so kann eine *a-posteriori* vorgenommene Güteermittlung nur als eine indirekte Methode zum Nachweis
von nicht vorhandenen Unterschieden betrachtet werden (Hayes, 1987; Muller und Benignus, 1992).

Ein weiterer Nachteil des *proof of hazard* besteht ferner darin, daß ein irrelevanter Unterschied zu einem signifikanten Testergebnis führen kann; statistische Signifikanz wäre dann inkongruent mit Unbedenklichkeit der Substanz. Dieses Phänomen tritt zum Beispiel oftmals bei der Analyse von Mutagenitätsuntersuchungen auf, da in einzelnen Mutagenitätsassays die Zelle als Versuchseinheit vorliegt und daher große Fallzahlen bzw. wegen der diskreten Verteilung der Endpunkte eine Vielzahl von Bindungen und somit geringe Varianzschatzungen resultieren (Hauschke et al., 1997).

Insgesamt gesehen kann daher der *proof of hazard* sowohl behördlichen (Kontrolle des Konsumentenrisikos) als auch produzentenspezifischen (Ausschluß irrelevanter Unterschiede) Anforderungen nicht gerecht werden. Der wesentliche Grund besteht darin, daß auf der Neyman-Pearson Theorie basierende statistische Testverfahren das Vorliegen der Alternativhypothese anhand der empirischen Evidenz gegen die Nullhypothese demonstrieren. Klar beschrieben wurde dies bereits von Fisher (1935): *The null hypothesis is never proved or established, but is possibly disproved in the course of experimentation. Every experiment may be said to exist only in order to give the facts a chance of disproving the null hypothesis.* Es ist daher verwunderlich, daß erst Mitte der achtziger Jahre von Bross (1985) für den direkten Sicherheitsnachweis das Konzept des *proof of safety* vorgestellt wurde. Hierbei wird als Nullhypothese \( H_0 : \text{Substanz bedenklich} \) und als Alternative \( H_1 : \text{Substanz unbedenklich} \) gewählt und desweiteren wird ein Schwellenwert (nicht mehr akzeptables Risiko) mit in die Formulierung des Testproblems einbezogen. Bross (1985) und Millard (1987) benutzten den *proof of safety* für die Frage, ob gegen die Bewohnbarkeit einer ehemals belasteten Umgebung Sicherheitsbedenken bestehen. Beide Autoren kamen zum Schluß, daß für den direkten Sicherheitsnachweis umfangreichere Fallzahlen als bisher zu Grunde gelegt werden müssen. Basierend auf dieser Erkenntnis, empfahlen Holland und Ordoukhani (1990) den klassischen Nachweis des *proof of hazard* beizubehalten und dabei für den Fehler 1. Art höhere Niveaus zu wählen und zusätzlich gleich dem Fehler 2. Art zu setzen, d.h. \( \alpha = \beta = 0.10 \) oder sogar in bestimmten Fällen \( \alpha = \beta = 0.20 \).


Das Crossover-Design kann durch eine elementare Transformation der zugrundeliegenden Zufallsvariablen auf eine Zweistichprobensituation zurückgeführt werden. Daher werden die im dritten Kapitel hergeleiteten Verfahren im vierten Kapitel auf die Crossover-Situation übertragen.

Im fünften Kapitel werden die biometrischen Methoden für den proof of safety in der Mehrstichprobensituation für partiell bzw. vollständig geordnete Alternativen hergeleitet. Das zentrale Anliegen besteht dabei in der Ermittlung sicherer Substanzen bzw. in der Bestimmung der maximal sicheren Dosis.

Nach der Darstellung des Konzeptes der Bioäquivalenz im sechsten Kapitel, werden im siebten Kapitel die wesentlichen Ergebnisse der Arbeit zusammengefaßt und es wird im Ausblick auf noch offene Probleme hingewiesen.
2 Einführung in die Problematik

Zur Veranschaulichung der Fragestellung wird die Zweistichprobensituation mit $n_0$ bzw. $n_1$ Beobachtungen pro Gruppe betrachtet. Die zur Sicherheitsbeurteilung relevante Zielgröße sei beschrieben durch die Zufallsvariable $X_{ij}$, die unabhängig identisch verteilt ist (u.i.v. $\sim$) mit der stetigen Verteilungsfunktion $F_i(x) = F(x - \mu_i)$, $i = 0, 1$ und $j = 1, \ldots, n_i$. Dabei kennzeichnen $\mu_i$ die Lageparameter, der Index $i = 0$ die Kontrolle (Vehikel- oder Negativkontrolle) und der Index $i = 1$ die Behandlungsgruppe. Desweiteren wird angenommen, daß sich ein möglicher sicherheitsrelevanter Substanzeffekt durch höhere Ausprägungen der Zielgröße auszeichnet, d.h. $\mu_1 > \mu_0$. Das traditionelle Testproblem des proof of hazard läßt sich dann folgendermaßen formulieren:

\[ H_0 : \mu_1 - \mu_0 \leq 0 \quad \text{(Substanz in der vorliegenden Untersuchung unbedenklich)} \]

versus

\[ H_1 : \mu_1 - \mu_0 > 0 \quad \text{(Substanz in der vorliegenden Untersuchung bedenklich).} \]


Trifft man nun die parametrische Verteilungsannahme $F_i(x) = \Phi\left(\frac{x - \mu_i}{\sigma}\right)$, wobei $\Phi(\cdot)$ die Verteilungsfunktion der Standardnormalverteilung bezeichnet, und nimmt man ferner an, daß die Varianz $\sigma^2$ zwar unbekannt aber in beiden Gruppen gleich ist, so kann obiges Testproblem mittels des Zweistichproben-$t$-Tests überprüft werden. Die Nullhypothese $H_0$ wird genau dann verworfen, falls

\[ t_1 = \frac{\bar{X}_1 - \bar{X}_0}{S \left( \frac{1}{n_0} + \frac{1}{n_1} \right)} \geq t_{\alpha,n_0+n_1-2}, \]

wobei

\[ \bar{X}_0 = \frac{1}{n_0} \sum_{j=1}^{n_0} X_{0j}, \quad \bar{X}_1 = \frac{1}{n_1} \sum_{j=1}^{n_1} X_{1j} \]
die empirischen Mittelwerte der beiden Stichproben bezeichnen,

\[ S^2 = \frac{1}{n_0 + n_1 - 2} \sum_{i=0}^{n_0} \sum_{j=1}^{n_1} (X_{ij} - \bar{X}_i)^2 \]

die gepoolte empirische Varianz ist und der kritische Wert \( t_{\alpha,n} \) dem \((1-\alpha)\)-Quantil der zentralen \( t \)-Verteilung mit \( v \) Freiheitsgraden entspricht.

In Abbildung 1 wird die Abhängigkeit der Güte des Zweistichproben-\( t \)-Tests von dem relativ zur Kontrollgruppe dargestellten Variationskoeffizienten \((CV_0 = \frac{\sigma}{\mu_0})\) und der Alternativen dargestellt, wobei die Differenz \( \mu_1 - \mu_0 \) ebenfalls als Anteil von \( \mu_0 \) ausgedrückt wird. Mit \( \alpha = 0.05 \) und jeweils 10 Versuchseinheiten pro Gruppe wurden für subchronische Toxizitätsuntersuchungen übliche Größen gewählt. In Tabelle 1 werden exemplarisch die entsprechenden Gütwerte an der Stelle \( \mu_1 - \mu_0 = 0.1\mu_0 \) für ausgewählte haematologische Endpunkte an Ratten angegeben. Es ist offensichtlich, daß bei den in der Praxis vorkommenden Variabilitäten nur dramatische Effekte mit hinreichend hoher Wahrscheinlichkeit entdeckt werden können: Nimmt der Variationskoeffizient zu bzw. ist ein möglicher Unterschied zwischen den Lokationsparametern nicht sehr ausgeprägt, so verringert sich die Güte des Entscheidungsverfahrens dramatisch.

Abbildung 1: Die Gütefunktion des einseitigen Zweistichproben-\( t \)-Tests \( t_1 \) für \( \alpha = 0.05, n_0 = n_1 = 10 \) und verschiedene Variationskoeffizienten \( CV_0 \)
Tabelle 1: Gütabwerte für ausgewählte haematologische Endpunkte von Ratten (Chanter et al., 1987) an der Stelle \( m_1 - m_0 = 0.1\mu_0, \alpha = 0.05, n_0 = n_1 = 10 \)

<table>
<thead>
<tr>
<th>Endpunkt</th>
<th>( CV_0 )%</th>
<th>Güte</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haemoglobin</td>
<td>4</td>
<td>0.9999</td>
</tr>
<tr>
<td>Prothrombin Zeit</td>
<td>9</td>
<td>0.7717</td>
</tr>
<tr>
<td>Thrombozyten</td>
<td>15</td>
<td>0.4167</td>
</tr>
<tr>
<td>Leukozyten</td>
<td>28</td>
<td>0.1905</td>
</tr>
<tr>
<td>Neutrophile</td>
<td>66</td>
<td>0.0936</td>
</tr>
</tbody>
</table>

Zur praktischen Illustration diene die folgende chronische Toxizitätsprüfung an weiblichen Wistar-Ratten, in der unter anderem die Serum-ASAT-Werte nach einer Applikationsdauer von sechs Monaten erhoben wurden. In Tabelle 2 sind die Einzelwerte und in Abbildung 2 die entsprechenden Box-plots für die Kontrolle und für die mit der Substanz behandelten Gruppe dargestellt. Geplant waren 19 Tiere pro Gruppe; infolge mortalitätsbedingter Ausfälle innerhalb der Behandlungsgruppe ergaben sich jedoch unterschiedliche Fallzahlen.

Tabelle 2: Serum-ASAT-Werte(\( \mu mol/L \)) weiblicher Wistar-Ratten

<table>
<thead>
<tr>
<th></th>
<th>Kontrollgruppe</th>
<th>Behandlungsgruppe</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.33</td>
<td>1.53</td>
<td></td>
</tr>
<tr>
<td>1.78</td>
<td>1.75</td>
<td></td>
</tr>
<tr>
<td>1.53</td>
<td>2.12</td>
<td></td>
</tr>
<tr>
<td>1.95</td>
<td>2.83</td>
<td></td>
</tr>
<tr>
<td>1.83</td>
<td>2.58</td>
<td></td>
</tr>
<tr>
<td>1.47</td>
<td>2.37</td>
<td></td>
</tr>
<tr>
<td>1.87</td>
<td>2.92</td>
<td></td>
</tr>
<tr>
<td>1.55</td>
<td>2.08</td>
<td></td>
</tr>
<tr>
<td>2.58</td>
<td>3.01</td>
<td></td>
</tr>
<tr>
<td>2.17</td>
<td>2.67</td>
<td></td>
</tr>
<tr>
<td>1.97</td>
<td>2.06</td>
<td></td>
</tr>
<tr>
<td>1.62</td>
<td>2.25</td>
<td></td>
</tr>
<tr>
<td>2.25</td>
<td>3.08</td>
<td></td>
</tr>
<tr>
<td>3.53</td>
<td>2.01</td>
<td></td>
</tr>
<tr>
<td>2.92</td>
<td>1.58</td>
<td></td>
</tr>
<tr>
<td>1.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anzahl Tiere</td>
<td>19</td>
<td>15</td>
</tr>
<tr>
<td>Mittelwert</td>
<td>2.09</td>
<td>2.32</td>
</tr>
<tr>
<td>Standardabweichung</td>
<td>0.57</td>
<td>0.51</td>
</tr>
</tbody>
</table>
Abbildung 2: Box-Plots für die Serum-ASAT-Werte ($\mu$mol/L) weiblicher Wistar-Ratten

Die zugrundeliegende Fragestellung beinhaltete eine mögliche sicherheitsrelevante Erhöhung der ASAT-Werte, dabei wurde eine Erhöhung von $\delta = 0.45$ als minimal biologisch relevant angesehen (Hothorn, 1992). Führt man den klassischen Test auf $H_0: \mu_1 - \mu_0 \leq 0$ (keine Erhöhung in den ASAT-Werten) versus $H_1: \mu_1 - \mu_0 > 0$ (erhöhte ASAT-Werte), mittels des $t$-Tests durch, so erhält man einen $p$-Wert von 0.108. In der Praxis folgert man nun aus der Nichtverwerfung der Nullhypothese das fehlende Vorliegen der Erhöhung der ASAT-Werte und übersieht dabei die folgende von Altman und Bland (1994) treffend formulierte Tatsache: Absence of evidence is not evidence of absence. Vertauscht man Null- und Alternativhypothese und führt man zusätzlich mit $\delta = 0.45$ einen Schwellenwert (nicht mehr akzeptable Erhöhung) ein, so erhält man das Testproblem:

$H_0: \mu_1 - \mu_0 \geq 0.45$ (relevant erhöhte ASAT-Werte) versus $H_1: \mu_1 - \mu_0 < 0.45$ (keine relevante Erhöhung in den ASAT-Werten).
Die Auswertung dieser verschobenen Nullhypothese ergibt nun einen $p$-Wert von 0.132, somit kann nicht auf eine unbedenkliche Erhöhung der ASAT-Werte geschlossen werden.


**Tabelle 3:** Terminale Körpergewichte (g) männlicher Ratten in den Erholungsgruppen

<table>
<thead>
<tr>
<th></th>
<th>Kontrollgruppe</th>
<th>Hohe Dosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>519.8</td>
<td>465.0</td>
<td></td>
</tr>
<tr>
<td>535.7</td>
<td>492.2</td>
<td></td>
</tr>
<tr>
<td>491.7</td>
<td>466.3</td>
<td></td>
</tr>
<tr>
<td>508.7</td>
<td>555.5</td>
<td></td>
</tr>
<tr>
<td>565.5</td>
<td>528.1</td>
<td></td>
</tr>
<tr>
<td>485.8</td>
<td>483.8</td>
<td></td>
</tr>
<tr>
<td>520.7</td>
<td>502.8</td>
<td></td>
</tr>
<tr>
<td>507.5</td>
<td>468.1</td>
<td></td>
</tr>
<tr>
<td>428.9</td>
<td>519.2</td>
<td></td>
</tr>
<tr>
<td>466.0</td>
<td>500.2</td>
<td></td>
</tr>
<tr>
<td>Anzahl Tiere</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Mittelwert</td>
<td>503.0</td>
<td>498.1</td>
</tr>
<tr>
<td>Median</td>
<td>508.1</td>
<td>496.2</td>
</tr>
<tr>
<td>Standardabweichung</td>
<td>37.9</td>
<td>29.7</td>
</tr>
</tbody>
</table>
Die inhaltliche Fragestellung beinhaltet die Vergleichbarkeit der terminalen Körpergewichte beider Gruppen. Das klassische Testproblem lautet somit

\[ H_0: \mu_1 - \mu_0 = 0 \quad \text{versus} \quad H_1: \mu_1 - \mu_0 \neq 0 \]

und \( H_0 \) wird genau dann verworfen, falls \( |t_k| \geq t_{a, n_0 + n_1 - 2} \). Führt man den Test nun mit den obigen Daten durch, so erhält man einen \( p \)-Wert von 0.75. In der Praxis folgert man nun aus der Nichtverwerfung der Nullhypothese irrtümlicherweise die Vergleichbarkeit der Körpergewichte. Analog zur einseitigen Fragestellung führt die korrekte Formulierung des Testproblems auf das zweiseitige Äquivalenzproblem:

\[ H_0: \mu_1 - \mu_0 \leq \delta_1 \quad \text{oder} \quad \mu_1 - \mu_0 \geq \delta_2 \]

versus

\[ H_1: \delta_1 < \mu_1 - \mu_0 < \delta_2, \]

wobei das Intervall \((\delta_1, \delta_2)\), \(\delta_1 < 0 < \delta_2\), den Sicherheitsbereich definiert. Die entsprechende Auswertung dieses Testproblems wird im dritten Kapitel detailliert beschrieben.

Eine weitere Fragestellung in Sicherheitsuntersuchungen ist die Bestimmung einer maximal sicheren Dosis (Hauschke, 1995; Hothorn und Hauschke, 2000; Tamhane et al., 1999). Dazu wird ein Experiment mit einer Vehikelkontrolle (Versuchseinheiten erhalten nur...
das Trägermedium, zum Beispiel das Lösungsmittel) sowie meist drei Dosierungen der Substanz durchgeführt. Zur Veranschaulichung diene die folgende Untersuchung (Tabelle 4), in der Tieren die Substanz in unterschiedlichen Dosierungen verabreicht wird. Die Zielvariable ist der Anteil erkrankter Tiere; Experiment I und II unterscheiden sich nur durch die Fallzahlen (Gaylor, 1983) von 20 bzw. 60 pro Gruppe.

Eine in der Praxis noch übliche Auswertungsstrategie basiert auf dem Abschlußtestprinzip (Marcus et al., 1976) und besteht darin, daß im ersten Schritt der Cochran-Armitage-Test mit allen Behandlungsgruppen durchgeführt wird (Lüdin, 1985). Ist der \( p \)-Wert kleiner oder gleich 0.05, so wird der Test ohne die höchste Dosis wiederholt. Im Falle einer erneuten Signifikanz zum Niveau von 0.05 erfolgt der entsprechende Vergleich zwischen Kontrolle und kleiner Dosis. Liegt ein nichtsignifikantes Ergebnis (\( p > 0.05 \)) vor, so wird die in dem Vergleich vorliegende höchste Dosis als maximal sichere Dosis bezeichnet. Es seien darauf hingewiesen, daß diese Dosis in der Literatur auch mit \textit{NOED (no observed effect dose)} oder \textit{NOEL (no observed effect level)} bezeichnet wird (Williams, 1971, 1972; Shirley, 1977; Tukey et al., 1985; Rom et al., 1994).

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|}
\hline
\textbf{Behandlungsgruppe} & \textbf{Experiment 1} & \textbf{Experiment 2} \\
\hline
Kontrolle & 0/20 & 0/60 \\
Kleine Dosis \( D_1 \) & 1/20 & 3/60 \\
Mittlere Dosis \( D_2 \) & 2/20 & 6/60 \\
Hohe Dosis \( D_3 \) & 10/20 & 30/60 \\
\hline
\end{tabular}
\caption{Tabelle 4: Dosis-Wirkungsanalyse in zwei Experimenten (Gaylor, 1983)}
\end{table}

Die exakten, konditionalen \( p \)-Werte sind in der Tabelle 5 aufgelistet; zusätzlich wurde die asymptotische Güte nach Nam (1987) berechnet. Während im ersten Experiment mit jeweils 20 Tieren pro Gruppe die mittlere Dosis \( D_2 \) als maximal sichere Dosis resultiert, ist in dem zweiten Experiment mit Verdreifachung der Fallzahlen, aber ansonsten gleichbleibenden Anteilen, die kleinste Dosis \( D_1 \) die maximal sichere Dosis. Betrachtet man die jeweiligen Güten, so ist es offensichtlich, daß sich durch entsprechende Wahl der Stichprobengrößen unterschiedliche, maximal sichere Dosierungen ergeben. Je kleiner die Stichprobe, desto größer ist die maximal sichere Dosis; dieses stellt jedoch genau das Entgegengesetzte dar, was
experimentell beabsichtigt ist. Daher wird im fünften Kapitel die probate Auswertungsstrategie zur Bestimmung der maximal sicheren Dosis hergeleitet.

<table>
<thead>
<tr>
<th>Tabelle 5: Exakte ( p )-Werte des Cochran-Armitage-Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behandlungsgruppe</td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td>Kleine Dosis ( D_1 )</td>
</tr>
<tr>
<td>Mittlere Dosis ( D_2 )</td>
</tr>
<tr>
<td>Hohe Dosis ( D_3 )</td>
</tr>
</tbody>
</table>

\(^{\text{a}}\) Asymptotische Güte


3 Zweistichprobensituation


3.1 Sicherheitsnachweis für die Differenz von Lokationsparametern

3.1.1 Einseitiges Testproblem

Für den Vergleich einer Behandlungsgruppe mit der Kontrollgruppe stellt sich das einseitige Testproblem wie folgt dar:

\[ H_0^{\delta} : \mu_1 - \mu_0 \geq \delta \quad \text{(Substanz in der vorliegenden Untersuchung bedenklich)} \]

versus

\[ H_1^{\delta} : \mu_1 - \mu_0 < \delta \quad \text{(Substanz in der vorliegenden Untersuchung unbedenklich)}, \]

wobei $(-\infty, \delta), \delta > 0$, den Bereich des nicht relevanten Unterschiedes festlegt. Auf die folgenden zwei Aspekte sei nochmals hingewiesen: Zum einen wird ein irrelevanter Unterschied zugelassen, und zum anderen wird die zu zeigende Vermutung als Alternative $H_1^{\delta}$ formuliert. Letzteres hat zur Folge, daß die Wahrscheinlichkeit für die irrtümliche Annahme der Unbedenklichkeit direkt durch den Fehler I. Art $\alpha$, welcher $a$-priori vorgebar ist, begrenzt wird. Im Falle der Normalverteilung mit homogenen, aber unbekannten Varianzen, d.h. $F_i(x) = \Phi\left(\frac{x - \mu_i}{\sigma}\right), i = 0, 1$, wird die Nullhypothese $H_0^{\delta}$ abgelehnt, falls
\[ t_1^\delta = \frac{\bar{X}_1 - \bar{X}_0 - \delta}{S \sqrt{\frac{1}{n_0} + \frac{1}{n_1}}} \leq -t_{\alpha,n_0+n_1-2}. \]

Die Durchführung des Tests verlangt nun im Gegensatz zum klassischen proof of hazard die explizite Einbeziehung des Schwellenwertes \( \delta \) in den Zähler der Teststatistik. Sollte es jedoch nicht möglich sein, diesen Wert a-priori zu definieren, so kann die Berechnung von entsprechenden Konfidenzintervallen eine Alternative darstellen: Kann \( \delta \) somit nur a-posteriori definiert werden, so ist der Schluß auf Unbedenklichkeit zum Niveau \( \alpha \) möglich, falls die obere Grenze \( \delta^+ \) des einseitigen 100(1-\( \alpha \))%-Konfidenzintervales für die Differenz \( \mu_1 - \mu_0 \) unterhalb von \( \delta \) liegt:

\[ (-\infty, \delta^+] = \left(-\infty, \bar{X}_1 - \bar{X}_0 + t_{\alpha,n_0+n_1-2} S \sqrt{\frac{1}{n_0} + \frac{1}{n_1}} \right] \subseteq (-\infty, \delta). \]

Verzichtet man auf die Annahme der Normalverteilung und postuliert man nur das Lokationsmodell \( X_{ij} \) u.i.v. \( F_i(x) = F(x - \mu_i) \), stetig, \( i = 0, 1 \) und \( j = 1, ..., n_i \), so gestatten der Zweistichproben-Wilcoxon-Test und das entsprechende verteilungsfreie Konfidenzintervall nach Moses (Hollander und Wolfe, 1973; Lehmann, 1975) ein analoges, nichtparametrisches Entscheidungsverfahren zur Beurteilung der Sicherheit. Sei \( R_{ij}(\bar{\delta}) \) der Rang von \( X_{ij} - \delta \) in der kombinierten Stichprobe

\[ X_{01}, ..., X_{0n_0}, X_{11} - \delta, ..., X_{1n_1} - \delta, \]

und \( R_{1}(\bar{\delta}) = \sum_{j=1}^{n_1} R_{ij}(\bar{\delta}) \) die Rangsumme bezüglich der um \( \delta \) verschobenen Beobachtungen der Behandlungsgruppe. Die Nullhypothese \( H_0^\delta \) wird verworfen, falls

\[ R_{1}(\bar{\delta}) \leq n_1(n_0 + n_1 + 1) - r_{\alpha,n_0,n_1}^\delta, \]

wobei \( r_{\alpha,n_0,n_1}^\delta \) das (1-\( \alpha \))-Quantil der Wilcoxonstatistik \( R_{1}(\bar{\delta}) \) bezeichne. Analog zur parametrischen Auswertung wird \( H_0^\delta \) abgelehnt, falls das verteilungsfreie einseitige 100(1-\( \alpha \))%-Konfidenzintervall für die Differenz der Lokationsparameter \( \mu_1 - \mu_0 \) vollständig im Sicherheitsbereich \( (-\infty, \delta) \) enthalten ist. Die Dualität zwischen Teststatistik und

Die Konstruktion des Konfidenzintervales basiert auf den geordneten Werten \( D_1 \leq \ldots \leq D_{n_0 n_1} \) der insgesamt \( n_0 n_1 \) Differenzen \( X_{ij} - X_{0 j} \), \( j = 1, \ldots, n_1 \) und \( j^* = 1, \ldots, n_0 \).

Damit ergibt sich als einseitiges 100(1-\( \alpha \))%-Konfidenzintervall:

\[
(-\infty, U_1^\delta], \quad \text{wobei} \quad U_1^\delta = D_{n_0 n_1+1-C_\alpha^1}, \quad C_\alpha^1 = \frac{n_1 (2n_0 + n_1 + 1)}{2} + 1 - r_0^\delta.
\]

Zur Vereinfachung sind in Tabelle 6 die zur Konstruktion von einseitigen 95%-Konfidenzintervallen benötigten Werte \( u_1 \) für eine Gesamtstichprobe vom Umfang \( n_0 + n_1 = 8, \ldots, 30 \) angegeben. Für größere \( n_0 \) und \( n_1 \) läßt sich \( C_\alpha^1 \) mittels

\[
C_\alpha^1 = \frac{n_0 n_1}{2} - z_\alpha \sqrt{\frac{n_0 n_1 (n_0 + n_1 + 1)}{12}}
\]

approximieren (Hollander und Wolfe, 1973); dabei bezeichnet \( z_\alpha \) das (1-\( \alpha \))-Quantil der Standardnormalverteilung.

### Tabelle 6: Zweistichprobensituation: Indizes \( u_1 \) zur Konstruktion des einseitigen nichtparametrischen 95%-Konfidenzintervales für \( \mu_1 - \mu_0 \)

<table>
<thead>
<tr>
<th>( n_0 ) =</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n_1 = 4 )</td>
<td>15</td>
<td>18</td>
<td>21</td>
<td>24</td>
<td>27</td>
<td>30</td>
<td>33</td>
<td>36</td>
<td>39</td>
<td>42</td>
<td>45</td>
<td>48</td>
</tr>
<tr>
<td>( n_1 = 5 )</td>
<td>18</td>
<td>21</td>
<td>25</td>
<td>29</td>
<td>32</td>
<td>36</td>
<td>39</td>
<td>43</td>
<td>47</td>
<td>50</td>
<td>54</td>
<td>57</td>
</tr>
<tr>
<td>( n_1 = 6 )</td>
<td>21</td>
<td>25</td>
<td>29</td>
<td>34</td>
<td>38</td>
<td>42</td>
<td>46</td>
<td>50</td>
<td>55</td>
<td>59</td>
<td>63</td>
<td>67</td>
</tr>
<tr>
<td>( n_1 = 7 )</td>
<td>24</td>
<td>29</td>
<td>34</td>
<td>38</td>
<td>43</td>
<td>48</td>
<td>53</td>
<td>58</td>
<td>63</td>
<td>67</td>
<td>72</td>
<td>77</td>
</tr>
<tr>
<td>( n_1 = 8 )</td>
<td>27</td>
<td>32</td>
<td>38</td>
<td>43</td>
<td>49</td>
<td>54</td>
<td>60</td>
<td>65</td>
<td>70</td>
<td>76</td>
<td>81</td>
<td>87</td>
</tr>
<tr>
<td>( n_1 = 9 )</td>
<td>30</td>
<td>36</td>
<td>42</td>
<td>48</td>
<td>54</td>
<td>60</td>
<td>66</td>
<td>72</td>
<td>78</td>
<td>84</td>
<td>90</td>
<td>96</td>
</tr>
<tr>
<td>( n_1 = 10 )</td>
<td>33</td>
<td>39</td>
<td>46</td>
<td>53</td>
<td>60</td>
<td>66</td>
<td>73</td>
<td>79</td>
<td>86</td>
<td>93</td>
<td>99</td>
<td>106</td>
</tr>
<tr>
<td>( n_1 = 11 )</td>
<td>36</td>
<td>43</td>
<td>50</td>
<td>58</td>
<td>65</td>
<td>72</td>
<td>79</td>
<td>87</td>
<td>94</td>
<td>101</td>
<td>108</td>
<td>115</td>
</tr>
<tr>
<td>( n_1 = 12 )</td>
<td>39</td>
<td>47</td>
<td>55</td>
<td>63</td>
<td>70</td>
<td>78</td>
<td>86</td>
<td>94</td>
<td>102</td>
<td>109</td>
<td>117</td>
<td>125</td>
</tr>
<tr>
<td>( n_1 = 13 )</td>
<td>42</td>
<td>50</td>
<td>59</td>
<td>67</td>
<td>76</td>
<td>84</td>
<td>93</td>
<td>101</td>
<td>109</td>
<td>118</td>
<td>126</td>
<td>134</td>
</tr>
<tr>
<td>( n_1 = 14 )</td>
<td>45</td>
<td>54</td>
<td>63</td>
<td>72</td>
<td>81</td>
<td>90</td>
<td>99</td>
<td>108</td>
<td>117</td>
<td>126</td>
<td>135</td>
<td>144</td>
</tr>
<tr>
<td>( n_1 = 15 )</td>
<td>48</td>
<td>57</td>
<td>67</td>
<td>77</td>
<td>87</td>
<td>96</td>
<td>106</td>
<td>115</td>
<td>125</td>
<td>134</td>
<td>144</td>
<td>153</td>
</tr>
</tbody>
</table>
3.1.2 Zweiseitiges Testproblem

Die einseitige Formulierung des Testproblems ist in der überwiegenden Anzahl von Sicherheitsstudien angezeigt. Es existieren jedoch auch Fragestellungen (Reversibilität von substanzbedingten Effekten), bei denen eher ein zweiseitiges Testproblem relevant ist. Das entsprechende Testproblem lautet

\[
H_0^\delta: \mu_1 - \mu_0 \leq \delta_1 \quad \text{oder} \quad \mu_1 - \mu_0 \geq \delta_2
\]

\[
H_1^\delta: \delta_1 < \mu_1 - \mu_0 < \delta_2,
\]

wobei das Intervall \((\delta_1, \delta_2)\), \(\delta_1 < 0 < \delta_2\), den Sicherheitsbereich oder, anders ausgedrückt, den Bereich des nicht relevanten Unterschiedes definiert. Eine Zerlegung in zwei einseitige Testprobleme ergibt:

\[
H_{01}^\delta: \mu_1 - \mu_0 \leq \delta_1 \quad \text{versus} \quad H_{11}^\delta: \mu_1 - \mu_0 > \delta_1
\]

und

\[
H_{02}^\delta: \mu_1 - \mu_0 \geq \delta_2 \quad \text{versus} \quad H_{12}^\delta: \mu_1 - \mu_0 < \delta_2.
\]

Die Nullhypothese \(H_0^\delta\) ist die Vereinigung von \(H_{01}^\delta\) und \(H_{02}^\delta\), und die Alternative \(H_1^\delta\) läßt sich als Durchschnitt von \(H_{11}^\delta\) und \(H_{12}^\delta\) darstellen:

\[
H_0^\delta = \bigcup_{r=1}^2 H_{0r}^\delta \quad \text{versus} \quad H_1^\delta = \bigcap_{r=1}^2 H_{1r}^\delta.
\]

Entsprechend des *intersection-union*-Prinzips (Berger, 1982) wird \(H_0^\delta\) zum Niveau \(\alpha\) genau dann abgelehnt, falls beide Hypothesen \(H_{01}^\delta\) und \(H_{02}^\delta\) jeweils zum Niveau \(\alpha\) verworfen werden können. Bei der parametrischen Annahme \(F_i(x) = \Phi\left(\frac{x - \mu_i}{\sigma}\right), i = 0, 1,\) mit unbekannter, aber homogener Varianz \(\sigma^2\) führt die Anwendung des Zweistichproben-\(t\)-Tests zur Ablehnung von \(H_{0i}^\delta\) bzw. von \(H_{02}^\delta\), falls

\[
t_{11}^\delta = \frac{\bar{X}_1 - \bar{X}_0 - \delta_1}{\sqrt{\frac{1}{n_0} + \frac{1}{n_1}}} \geq t_{\alpha,n_0+n_1-2} \quad \text{und} \quad t_{12}^\delta = \frac{\bar{X}_1 - \bar{X}_0 - \delta_2}{\sqrt{\frac{1}{n_0} + \frac{1}{n_1}}} \leq -t_{\alpha,n_0+n_1-2},
\]
Die Ablehnung von \( H_0^\delta \) mittels der zwei einseitigen Zweistichproben-\( t \)-Tests (Schuirmann, 1987) ist äquivalent zur Inklusion des zweiseitigen 100(1-2\( \alpha \))%-Konfidenzintervales \( [\delta_1^\delta, \delta_2^\delta] \) für die Differenz \( \mu_1 - \mu_0 \) im Sicherheitsbereich \( (\delta_1, \delta_2) \):

\[
[\delta_1^\delta, \delta_2^\delta] = \left[ \bar{X}_1 - \bar{X}_0 \pm t_{\alpha,n_0+n_1-2} S \sqrt{\frac{1}{n_0} + \frac{1}{n_1}} \right] \subset (\delta_1, \delta_2).
\]

Das analoge verteilungsfreie Verfahren wurde von Hauschke et al. (1990) publiziert, wobei anstelle des Zweistichproben-\( t \)-Tests der Zweistichproben-Wilcoxon-Test verwendet wird.

\( H_{01}^\delta \) wird abgelehnt, falls \( R_1(\delta_1) \geq r_{\alpha,n_0,n_1}^{\delta_1} \)

und \( H_{02}^\delta \) wird abgelehnt, falls \( R_1(\delta_2) \leq n_1(n_0 + n_1 + 1) - r_{\alpha,n_0,n_1}^{\delta_1} \).

Die Verwerfung beider Einzelhypothesen ist äquivalent zur Inklusion des verteilungsfreien zweiseitigen 100(1-2\( \alpha \))%-Konfidenzintervales \( [L_1^\delta, U_1^\delta] \) für \( \mu_1 - \mu_0 \) im Sicherheitsbereich \( (\delta_1, \delta_2) \) (Hauschke und Steinijans, 1991; Chow und Liu, 1992), wobei \( L_1^\delta = D_{n_1}^\delta = D_{n_1}^\delta \) und \( U_1^\delta = D_{n_0}^\delta = D_{n_0,n_1+1}^\delta \). In Tabelle 7 sind die zur Konstruktion von zweiseitigen 90%-Konfidenzintervallen benötigten Werte der Indizes \( l_1 \) und \( u_1 \) für eine Gesamtstichprobe vom Umfang \( n_0 + n_1 \) = 8, ..., 30 angegeben.

**Tabelle 7:** Zweiseitichprobensituation: Indizes \( l_1/u_1 \) zur Konstruktion des zweiseitigen nichtparametrischen 90%-Konfidenzintervales für \( \mu_1 - \mu_0 \)

<table>
<thead>
<tr>
<th>( n_0 )</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>( u_1 )</td>
<td>3/18</td>
<td>5/21</td>
<td>6/25</td>
<td>7/29</td>
<td>9/32</td>
<td>10/36</td>
<td>12/39</td>
<td>13/43</td>
<td>14/47</td>
<td>16/50</td>
<td>17/54</td>
<td>19/57</td>
</tr>
<tr>
<td>( l_1 )</td>
<td>4/21</td>
<td>6/25</td>
<td>8/29</td>
<td>9/34</td>
<td>11/38</td>
<td>13/42</td>
<td>15/46</td>
<td>17/50</td>
<td>18/55</td>
<td>20/59</td>
<td>22/63</td>
<td>24/67</td>
</tr>
<tr>
<td>( u_1 )</td>
<td>5/24</td>
<td>7/29</td>
<td>9/34</td>
<td>12/38</td>
<td>14/43</td>
<td>16/48</td>
<td>18/53</td>
<td>20/58</td>
<td>22/63</td>
<td>25/67</td>
<td>27/72</td>
<td>29/77</td>
</tr>
<tr>
<td>( l_1 )</td>
<td>6/27</td>
<td>9/32</td>
<td>11/38</td>
<td>14/43</td>
<td>16/49</td>
<td>19/54</td>
<td>22/60</td>
<td>24/65</td>
<td>27/70</td>
<td>29/76</td>
<td>32/81</td>
<td>34/87</td>
</tr>
<tr>
<td>( u_1 )</td>
<td>7/30</td>
<td>10/36</td>
<td>13/42</td>
<td>16/48</td>
<td>19/54</td>
<td>22/60</td>
<td>25/66</td>
<td>28/72</td>
<td>31/78</td>
<td>34/84</td>
<td>37/90</td>
<td>40/96</td>
</tr>
<tr>
<td>( l_1 )</td>
<td>8/33</td>
<td>12/39</td>
<td>15/46</td>
<td>18/53</td>
<td>21/60</td>
<td>25/66</td>
<td>28/73</td>
<td>32/79</td>
<td>35/86</td>
<td>38/93</td>
<td>42/99</td>
<td>45/106</td>
</tr>
<tr>
<td>( u_1 )</td>
<td>9/36</td>
<td>13/43</td>
<td>17/50</td>
<td>20/58</td>
<td>24/65</td>
<td>28/72</td>
<td>32/79</td>
<td>35/87</td>
<td>39/94</td>
<td>43/101</td>
<td>47/108</td>
<td>51/115</td>
</tr>
<tr>
<td>( l_1 )</td>
<td>10/39</td>
<td>14/47</td>
<td>18/55</td>
<td>22/63</td>
<td>27/70</td>
<td>31/78</td>
<td>35/86</td>
<td>39/94</td>
<td>43/102</td>
<td>48/109</td>
<td>52/117</td>
<td>56/125</td>
</tr>
<tr>
<td>( u_1 )</td>
<td>11/42</td>
<td>16/50</td>
<td>20/59</td>
<td>25/67</td>
<td>29/76</td>
<td>34/84</td>
<td>38/93</td>
<td>43/101</td>
<td>48/109</td>
<td>52/118</td>
<td>57/126</td>
<td>62/134</td>
</tr>
<tr>
<td>( l_1 )</td>
<td>12/45</td>
<td>17/54</td>
<td>22/63</td>
<td>27/72</td>
<td>32/81</td>
<td>37/90</td>
<td>42/99</td>
<td>47/108</td>
<td>52/117</td>
<td>57/126</td>
<td>62/135</td>
<td>67/144</td>
</tr>
<tr>
<td>( u_1 )</td>
<td>13/48</td>
<td>19/57</td>
<td>24/67</td>
<td>29/77</td>
<td>34/87</td>
<td>40/96</td>
<td>45/106</td>
<td>51/115</td>
<td>56/125</td>
<td>62/134</td>
<td>67/144</td>
<td>73/153</td>
</tr>
</tbody>
</table>
Zur Veranschaulichung werden für den Rattenversuch hinsichtlich der Reversibilität (Tabelle 3) die 90%-Konfidenzintervalle für $\mu_1 - \mu_0$ entsprechend des parametrischen bzw. nichtparametrischen Ansatzes in Tabelle 8 dargestellt. Bei einem Sicherheitsbereich $(\delta_1, \delta_2)$ von $\pm 40g$ führen beide Methoden zur Verwerfung der Nullhypothese und somit zur Annahme der Reversibilität der durch die Substanz bedingten Körpergewichtsveränderung.

<table>
<thead>
<tr>
<th>Tabelle 8: Zweiseitige 90%-Konfidenzintervalle für $\mu_1 - \mu_0$ für den Beispieldatensatz aus Tabelle 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%-Konfidenzintervall</td>
</tr>
<tr>
<td>Parametrisches Verfahren</td>
</tr>
<tr>
<td>Nichtparametrisches Verfahren</td>
</tr>
</tbody>
</table>

3.1.3 Gütebestimmung und Fallzahlplanung

Zum Zwecke der Güteberechnung oder der Fallzahlermittlung wird in dieser Arbeit generell ein balanciertes Design angenommen, d.h. $n_0 = n_1 = n$. Die Güte des einseitigen Zweistichproben-$t$-Tests $t_1^\delta$ ist die Wahrscheinlichkeit für die korrekte Annahme von $H_1^\delta$ und läßt sich über die nichtzentrale $t$-Verteilung ermitteln:

$$P\left[t_1^\delta \leq -t_{\alpha,2n-2}\left|\mu_1 - \mu_0 < \delta, \sigma\right.\right] = \frac{\sqrt{2\pi}}{\Gamma(n-1)2^{n-2}} \int_0^\infty \Phi\left(-\frac{t_{a,2n-2} x}{\sqrt{2n-2}} - \Delta\right) x^{2n-3} \Phi(x) \, dx,$$

wobei $\Phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ die Dichte der Standardnormalverteilung bezeichnet und der Nichtzentralitätsparameter sich wie folgt darstellt:

$$\Delta = \frac{\mu_1 - \mu_0 - \delta}{\sigma \sqrt{\frac{2}{n}}}.$$

Die nichtzentrale $t$-Verteilung ist in den gängigen Statistikprogrammen als Funktion enthalten, so daß die entsprechende Berechnung der Güte ohne Probleme erfolgen kann. In verschiedenen Publikationen, zum Beispiel Phillips (1990) und Liu und Chow (1992), wird zur Vereinfachung der Darstellung des Nichtzentralitätsparameters der Schwellenwert $\delta$, der Wert $\mu_1 - \mu_0$ aus der Alternative und die zugrundeliegende Standardabweichung $\sigma$ relativ zum Erwartungswert der Kontrollgruppe definiert:
\[ \delta = \bar{\theta} \cdot \mu_0, \quad \mu_1 - \mu_0 = \bar{\theta} \cdot \mu_0, \quad \sigma = CV_0 \mu_0 \]

und somit gilt für den Nichtzentralitätsparameter \( \Delta = \frac{\bar{\theta} - \bar{\theta}^*}{CV_0 \sqrt{\frac{2}{n}}} \).

In Abbildung 4 wird die Gütefunktion für verschiedene Variationskoeffizienten \( CV_0 \) dargestellt; dabei werden ein Sicherheitsbereich von \( (-\infty, \delta) = (-\infty, 0.2 \mu_0) \) und 10 Versuchs- einheiten pro Gruppe gewählt.

Abbildung 4: Zweistichprobensituation: Die Gütefunktion des einseitigen Tests \( t_1^\delta \) bei einem
Sicherheitsbereich von \( (-\infty, \delta) = (-\infty, 0.2 \mu_0) \) für \( \alpha = 0.05, n_0 = n_1 = 10 \) und verschiedene \( CV_0 \)

Den benötigten Stichprobenumfang \( n \) (pro Gruppe), um beim Testen von \( H_0^\delta \) versus \( H_1^\delta \)
einen Fehler 1. Art \( \alpha \) und einen Fehler 2. Art \( \beta \) an einer festgewählten Stelle \( \bar{\theta}, \bar{\theta} < \bar{\theta}^* \), nicht zu überschreiten, kann iterativ nach der folgenden bekannten Ungleichung (Mace, 1964) bestimmt werden:

\[ n \geq 2 \left( \frac{CV_0}{\bar{\theta}^* - \bar{\theta}} \right)^2 \left( t_{\alpha,2n-2} + t_{\beta,2n-2} \right)^2. \]

In Tabelle 9 sind die benötigten Fallzahlen angegeben, um eine Güte von mindestens 0.80 zu erreichen. Dabei werden die weiteren Einflußgrößen wie folgt vorgegeben: \( \alpha = 0.05, \bar{\theta}^* = 0.20, \bar{\theta} = 0.00, 0.05, 0.10, 0.15 \) und \( CV_0 = 5.0, 7.5, 10.0, ..., 30.0\% \). Die exakten Werte
stehen in der ersten Zeile, und die approximativen werden, falls sie von den exakten abweichen, in der zweiten Zeile angegeben; der Bindestrich bedeutet somit Übereinstimmung der Werte.

**Tabelle 9:** Zweistichprobensituation: Exakter und approximativer Stichprobenumfang pro Gruppe für eine Güte von mindestens 0.80 bei \( \alpha = 0.05 \), \( (-\infty, \delta) = (-\infty, 0.2 \mu_0) \) und verschiedene \( CV_0 \)

<table>
<thead>
<tr>
<th>( CV_0 (%) )</th>
<th>0.00</th>
<th>0.05</th>
<th>0.10</th>
<th>0.15</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.0</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>7.5</td>
<td>3</td>
<td>-</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td>10.0</td>
<td>4</td>
<td>7</td>
<td>14</td>
<td>51</td>
</tr>
<tr>
<td>12.5</td>
<td>6</td>
<td>10</td>
<td>21</td>
<td>78</td>
</tr>
<tr>
<td>15.0</td>
<td>8</td>
<td>14</td>
<td>29</td>
<td>112</td>
</tr>
<tr>
<td>17.5</td>
<td>11</td>
<td>18</td>
<td>39</td>
<td>153</td>
</tr>
<tr>
<td>20.0</td>
<td>14</td>
<td>23</td>
<td>51</td>
<td>199</td>
</tr>
<tr>
<td>22.5</td>
<td>17</td>
<td>29</td>
<td>64</td>
<td>252</td>
</tr>
<tr>
<td>25.0</td>
<td>21</td>
<td>36</td>
<td>78</td>
<td>310</td>
</tr>
<tr>
<td>27.5</td>
<td>25</td>
<td>43</td>
<td>95</td>
<td>375</td>
</tr>
<tr>
<td>30.0</td>
<td>29</td>
<td>51</td>
<td>112</td>
<td>446</td>
</tr>
</tbody>
</table>

Die statistische Methodik zur Bestimmung der Güte für das zweiseitige Testproblem wurde von Owen (1965) hergeleitet. Dabei wird die folgende Eigenschaft benötigt:

**Lemma 1:**

Die Korrelation zwischen den Teststatistiken \( t_{11}^\delta \) und \( t_{12}^\delta \) entspricht dem Wert 1.

**Beweis:**

Es gilt

\[
Var(\bar{X}_1 - \bar{X}_0 - \delta) = Var(\bar{X}_1 - \bar{X}_0 - \delta_2),
\]

\[
Kov(\bar{X}_1 - \bar{X}_0 - \delta_1, \bar{X}_1 - \bar{X}_0 - \delta_2) = Var(\bar{X}_1 - \bar{X}_0 - \delta_1)
\]

und daher

\[
Korr(t_{11}^\delta, t_{12}^\delta) = \rho(t_{11}^\delta, t_{12}^\delta) = \rho^\delta = 1.
\]
Unter Ausnutzung dieser Eigenschaft kann nach Owen (1965) die zugrundeliegende nichtzentrale bivariate $t$-Verteilung durch die entsprechende Differenz zweier Integrale bestimmt werden, und somit erhält man die Wahrscheinlichkeit für die korrekte Verwerfung von $H_0^\delta$:

$$P(t_{11}^\delta \geq t_{\alpha,2n-2}^\delta \land t_{12}^\delta \leq -t_{\alpha,2n-2}^\delta | \delta_1 < \mu_1 - \mu_0 < \delta_2, \sigma)$$

$$= Q(-t_{\alpha,2n-2}^\delta, \Delta_2, 0, \psi) - Q(t_{\alpha,2n-2}^\delta, \Delta_1, 0, \psi),$$

wobei

$$Q(t, \Delta, u, \psi) = \frac{\sqrt{2\pi}}{\Gamma((n-1)/2)^{n-2}} \int_u^{\infty} \Phi \left( \frac{tx + \Delta}{\sqrt{2n-2}} \right) x^{2n-3} \Phi(x) \, dx,$$

$$\psi = \frac{\Delta_1 - \Delta_2}{A_1 - A_2}, \quad A_1 = \frac{t_{\alpha,2n-2}^\delta}{\sqrt{2n-2}}, \quad A_2 = \frac{-t_{\alpha,2n-2}^\delta}{\sqrt{2n-2}}$$

und

$$\Delta_1 = \frac{\mu_1 - \mu_0 - \delta_1}{\sigma \sqrt{\frac{2}{n}}}, \quad \Delta_2 = \frac{\mu_1 - \mu_0 - \delta_2}{\sigma \sqrt{\frac{2}{n}}}.$$

Definiert man wieder entsprechend

$$\delta_1 = \vartheta_1 \mu_0, \quad \delta_2 = \vartheta_2 \mu_0, \quad \mu_1 - \mu_0 = \vartheta \mu_0, \quad \sigma = CV_0 \mu_0,$$

so folgt für die Nichtzentralitätsparameter:

$$\Delta_1 = \frac{\vartheta - \vartheta_1}{CV_0 \sqrt{\frac{2}{n}}} \quad \text{und} \quad \Delta_2 = \frac{\vartheta - \vartheta_2}{CV_0 \sqrt{\frac{2}{n}}}.$$

Die von Owen angegebenen Formeln zur Bestimmung von $Q(t, \Delta, u, \psi)$ wurden für die vorliegende Arbeit mittels SAS/IML (siehe Anhang) realisiert und dienen als Basis zur Berechnung der entsprechenden Gütefunktionen. In der Abbildung 5 werden die Gütefunktionen für einen symmetrischen Sicherheitsbereich von $[2.0, 2.0]$, für einen Variationskoeffizienten von 15.0% und für die Stichprobenumfänge von jeweils 10 bzw. 20 pro Gruppe dargestellt. Die Gütefunktionen sind bei dieser Wahl des Äquivalenzbereiches symmetrisch zu 0 und erreichen auch dort ihr Maximum. Man sieht, daß für den Stichprobenumfang von 10 pro Gruppe eine maximale Güte von etwa 0.80 erreicht wird.
Abbildung 5: Zweistichprobensituation: Die Gütefunktion des Entscheidungsverfahrens, basierend auf den zwei einseitigen Tests $t_{11}^δ$, $t_{12}^δ$ bei einem Sicherheitsbereich von $(δ_1, δ_2) = (-0.2μ_0, 0.2μ_0)$ für $α = 0.05, n_0 = n_1 = 10, 20$ und $CV_0 = 15.0\%$


**Satz 1:**

Die Approximationsformeln für die Fallzahlen (pro Gruppe) für eine Güte von mindestens $1−β$ bei vorgegebenem $α$, $CV_0$ und einem zu 0 symmetrisch gewählten Sicherheitsbereich, d.h. $θ_1 = −θ_2$, lauten:

i) falls $θ = 0$

$$n ≥ 2 \left( \frac{CV_0}{θ_2} \right)^2 \left( t_{α,2n-2} + t_{θ_2,2n-2} \right)^2,$$

ii) $0 < θ < θ_2$

$$n ≥ 2 \left( \frac{CV_0}{θ_2 − θ} \right)^2 \left( t_{α,2n-2} + t_{θ_2,2n-2} \right)^2$$

und

iii) $−θ_2 < θ < 0$

$$n ≥ 2 \left( \frac{CV_0}{−θ_2 − θ} \right)^2 \left( t_{α,2n-2} + t_{θ_2,2n-2} \right)^2.$$
Beweis:

i) Aus $\theta = 0$ folgt $\mu_1 = \mu_0$, und die Güte lautet

$$P \left[ \frac{X_1 - X_0 - \delta_1}{S \sqrt{\frac{2}{n}}} \geq t_{\alpha, 2n-2} \land \frac{X_1 - X_0 - \delta_2}{S \sqrt{\frac{2}{n}}} \leq -t_{\alpha, 2n-2} \right] =$$

$$P \left[ \delta_1 + t_{\alpha, 2n-2} S \sqrt{\frac{2}{n}} \leq X_1 - X_0 \leq \delta_2 - t_{\alpha, 2n-2} S \sqrt{\frac{2}{n}} \right] =$$

$$P \left[ \frac{\delta_1}{S \sqrt{\frac{2}{n}}} + t_{\alpha, 2n-2} \leq \frac{X_1 - X_0}{S \sqrt{\frac{2}{n}}} \leq \frac{\delta_2}{S \sqrt{\frac{2}{n}}} - t_{\alpha, 2n-2} \right].$$

Für $\mu_1 = \mu_0$ ist die Statistik $\frac{X_1 - X_0}{S \sqrt{\frac{2}{n}}}$ zentral $t$-verteilt. Zudem gilt mit $\delta_1 = -\delta_2$ die folgende Symmetrieeigenschaft

$$\frac{\delta_2}{S \sqrt{\frac{2}{n}}} - t_{\alpha, 2n-2} = \left\{ \begin{array}{ll} -\delta_2 & , t_{\alpha, 2n-2} \\ \delta_2 + t_{\alpha, 2n-2} & , S \sqrt{\frac{2}{n}} \end{array} \right\}.$$

Verlangt man nun, daß die Power mindestens $1 - \beta$ beträgt, und behandelt man die empirische Fehlervarianz $S^2$ wie eine deterministische Größe und ersetzt sie durch $\sigma^2$ (Chow und Liu, 1992), so folgt

$$\left| \frac{\delta_2}{\sigma \sqrt{\frac{2}{n}}} - t_{\alpha, 2n-2} \right| \geq t_{\frac{\alpha}{2}, 2n-2}$$

und dies ist äquivalent zu
\[ n \geq 2 \left( \frac{\sigma}{\delta_2} \right)^2 \left( t_{a,2n-2} + t_{\frac{\beta}{2},2n-2} \right)^2. \]

Mit \( \delta_2 = \theta \mu_0 \) und \( \sigma = CV_0 \mu_0 \) folgt

\[ n \geq 2 \left( \frac{CV_0}{\delta_2} \right)^2 \left( t_{a,2n-2} + t_{\frac{\beta}{2},2n-2} \right)^2. \]

ii) Für \( 0 < \mu_1 - \mu_0 < \delta_2 \) lautet die Güte

\[
P \left[ \frac{\delta_1 - (\mu_1 - \mu_0)}{S \sqrt{n}^2} + t_{a,2n-2} \leq \frac{\bar{X} - \bar{X}_0 - (\mu_1 - \mu_0)}{S \sqrt{n}^2} \leq \frac{\delta_2 - (\mu_1 - \mu_0)}{S \sqrt{n}^2} - t_{a,2n-2} \right] =
\]

\[
P \left[ -\delta_2 - (\mu_1 - \mu_0) + t_{a,2n-2} \leq \frac{\bar{X} - \bar{X}_0 - (\mu_1 - \mu_0)}{S \sqrt{n}^2} \leq \delta_2 - (\mu_1 - \mu_0) - t_{a,2n-2} \right],
\]

da \( \delta_1 = -\delta_2 \). Die Symmetrieigenschaft liegt nicht mehr vor:

\[
- \left( \frac{\delta_2 - (\mu_1 - \mu_0)}{S \sqrt{n}^2} - t_{a,2n-2} \right) = \frac{-\delta_2 + (\mu_1 - \mu_0)}{S \sqrt{n}^2} + t_{a,2n-2} \geq \frac{-\delta_2 - (\mu_1 - \mu_0)}{S \sqrt{n}^2} + t_{a,2n-2}.
\]

Verlangt man nun, daß die Power mindestens \( 1 - \beta \) beträgt und ersetzt man \( S^2 \) durch \( \sigma^2 \), so können bei Verwendung der Ungleichung

\[
\frac{\delta_2 - (\mu_1 - \mu_0)}{\sigma \sqrt{n}^2} - t_{a,2n-2} \geq t_{\frac{\beta}{2},2n-2}
\]

sehr große Stichprobengrößen resultieren. Daher wird die folgende Beziehung verwendet:
\[
\begin{align*}
P \left( t_{\alpha,2n-2} \leq \frac{\delta_1 - (\mu_1 - \mu_0)}{S \sqrt{\frac{2}{n}}} \leq \frac{\delta_2 - (\mu_1 - \mu_0)}{S \sqrt{\frac{2}{n}}} - t_{\alpha,2n-2} \right) \leq P \left( \frac{\delta_2 - (\mu_1 - \mu_0)}{\sigma \sqrt{\frac{2}{n}}} - t_{\alpha,2n-2} \geq t_{\beta,2n-2} \right). \\
\text{Dies führt zu}
\end{align*}
\]

\[
\frac{\delta_2 - (\mu_1 - \mu_0)}{\sigma \sqrt{\frac{2}{n}}} - t_{\alpha,2n-2} \geq t_{\beta,2n-2}
\]

oder äquivalent dazu

\[
n \geq 2 \left( \frac{\sigma}{\delta_2 - (\mu_1 - \mu_0)} \right)^2 \left( t_{\alpha,2n-2} + t_{\beta,2n-2} \right)^2.
\]

Mit \( \delta_2 = \hat{\theta}_2 \mu_0 \), \( \mu_1 - \mu_0 = \hat{\delta} \mu_0 \) und \( \sigma = CV_0 \mu_0 \) folgt

\[
n \geq 2 \left( \frac{CV_0}{\hat{\theta}_2 - \hat{\delta}} \right)^2 \left( t_{\alpha,2n-2} + t_{\beta,2n-2} \right)^2.
\]

iii) Erfolgt analog zu ii)

\[\triangleright\]

In der Tabelle 10 sind die exakten und approximativen Fallzahlen für \( \alpha = 0.05, \beta = 0.20, \hat{\theta}_1 = -\hat{\theta}_2 = -0.20, \) \( CV_0 = 5.0, 10.0, ..., 30.0\% \) und für \( |\hat{\delta}| = 0.00, 0.02, 0.05, 0.10, 0.15 \) angegeben. Es sei darauf hingewiesen, daß die Approximationsformeln für die Werte \( |\hat{\delta}| < 0.05 \) und \( \hat{\delta} \neq 0 \) zu unlogischen Resultaten führen können: Die ermittelten Werte sind an der Stelle 0.02 geringer als an der für die Fallzahlbestimmung optimalen Stelle 0. Dieses
Phänomen beruht auf dem diskreten Schritt von $t_{β,2n-2}$ für $\delta \neq 0$ nach $\frac{t_{β,2n-2}}{\sqrt{2n-2}}$ für $\delta = 0$.

Daher sollte in diesen Situationen unbedingt die exakte Methode verwendet werden.

**Tabelle 10**: Zweiseitichprobensituation: Exakter und approximativer Stichprobenumfang pro Gruppe für eine Güte von mindestens 0.80 bei $\alpha = 0.05$, von $(δ_1, δ_2) = (-0.2µ_0, 0.2µ_0)$ und verschiedene $CV_0$

| $CV_0$ (%) | 0.00 | 0.02 | $|\delta|$ | 0.05 | 0.10 | 0.15 |
|-----------|------|------|----------|------|------|------|
| 5.0       | 3    | 3    | 3        | 4    | 14   |
|           | -    | -    | -        | 5    | -    |
| 7.5       | 4    | 4    | 5        | 8    | 29   |
|           | -    | -    | -        | -    | -    |
| 10.0      | 6    | 7    | 6        | 14   | 6    |
|           | -    | -    | -        | -    | -    |
| 12.5      | 8    | 7    | 10       | 21   | 78   |
|           | -    | -    | -        | -    | -    |
| 15.0      | 11   | 10   | 11       | 14   | 29   |
|           | -    | -    | -        | -    | -    |
| 17.5      | 14   | 13   | 14       | 18   | 112  |
|           | -    | -    | -        | -    | -    |
| 20.0      | 18   | 17   | 24       | 51   | 199  |
|           | -    | -    | -        | -    | -    |
| 22.5      | 23   | 21   | 23       | 64   | 252  |
|           | -    | -    | -        | -    | -    |
| 25.0      | 28   | 25   | 29       | 78   | 310  |
|           | -    | -    | -        | -    | -    |
| 27.5      | 34   | 35   | 36       | 79   | 375  |
|           | -    | -    | -        | -    | -    |
| 30.0      | 40   | 36   | 51       | 112  | 446  |
|           | -    | -    | -        | -    | -    |

3.2 Sicherheitsnachweis für den Quotienten von Lokationsparametern

Wie in den einleitenden Bemerkungen zu Kapitel 3 bereits erwähnt, ist die Festlegung des Sicherheitsbereiches in absoluten Einheiten in der Praxis oftmals sehr schwierig. Aus diesem Grunde werden im folgenden Methoden dargestellt, die es ermöglichen, diesen Wert relativ zum Erwartungswert der Kontrollgruppe zu definieren. Daraus resultiert die Notwendigkeit, entsprechende Testprobleme durch geeignete Verfahren bezüglich des Quotienten $\frac{\mu_1}{\mu_0}$, $\mu_0 \neq 0$, adäquat zu lösen.

3.2.1 Einseitiges Testproblem

Das einseitige Testproblem mit einem relativ zum Erwartungswert der Kontrolle definierten Schwellenwert, d.h. $δ = fµ_0$, $0 < f$, stellt sich wie folgt dar:
\[ H_0^\delta: \mu_1 - \mu_0 \geq \delta = f \mu_0 \iff H_0^0: \frac{\mu_1}{\mu_0} \geq 1 + f = \theta \]

versus

\[ H_1^\delta: \mu_1 - \mu_0 < \delta = f \mu_0 \iff H_1^0: \frac{\mu_1}{\mu_0} < 1 + f = \theta , \]

wobei \((-\infty, \theta), \theta > 1\), den auf den Quotienten der Erwartungswerte \(\frac{\mu_1}{\mu_0}\) bezogenen Bereich des nicht relevanten Unterschiedes definiert.

Sasabuchi (1988) zeigte, daß unter der parametrischen Annahme \(F_i(x) = \Phi \left( \frac{x - \mu_i}{\sigma} \right)\), \(i = 0, 1\), das folgende Testverfahren ein Likelihood-Ratio-Test zum Niveau \(\alpha\) ist: \(H_0^0\) wird genau dann ablehnt, falls

\[ t_1^0 = \frac{\bar{X}_1 - \theta \bar{X}_0}{S \sqrt{\frac{1}{n_0} + \frac{1}{n_1}}} \leq -t_{\alpha, n_0+n_1-2} . \]

Hauschke et al. (1999b) zeigten, daß die Bedingung \(t_1^0 \leq -t_{\alpha, n_0+n_1-2}\) äquivalent ist zu

\[ (-\infty, \theta_1^1] \subset (-\infty, \theta) \quad \text{und} \quad \bar{X}_0 > a_0 , \]

mit

\[ \theta_1^1 = \frac{\bar{X}_0 \bar{X}_1 + \sqrt{a_0 \bar{X}_0^2 + a_1 \bar{X}_1^2 - a_0 a_1}}{\bar{X}_0^2 - a_0} , \quad a_0 = \frac{S^2}{n_0} t_{\alpha, n_0+n_1-2}^2 \quad \text{und} \quad a_1 = \frac{S^2}{n_1} t_{\alpha, n_0+n_1-2}^2 . \]

Der Wert \(\theta_1^1\) ist die obere Grenze des einseitigen 100(1-\(\alpha\))-Konfidenzintervales nach Fieller (1954) für \(\frac{\mu_1}{\mu_0}\), daher gestattet die Verwendung dieses Konfidenzintervales ebenfalls eine \textit{a-posteriori} zu treffende Sicherheitsbeurteilung. Die Bedingung \(\bar{X}_0^2 > a_0\) bedeutet, daß der Lokationsparameter der Kontrolle \(\mu_0\) statistisch signifikant ungleich Null ist und stellt eine ausreichende Bedingung dafür dar, daß sich die Fiellersche Konfidenzmenge als ein reguläres Konfidenzintervall darstellt.

Für das Beispiel der chronischen Toxizitätsprüfung an weiblichen Wistar-Ratten (Tabelle 2) ergibt die Berechnung der oberen Grenze des einseitigen 95%-Konfidenzintervales für
\( \frac{\mu_1}{\mu_0} \) den Wert \( \theta_+^1 = 1.29 \). Postuliert man eine relevante Erhöhung von \( \delta = 0.25\mu_0 \), so folgt für die obere Schranke von \( \frac{\mu_1}{\mu_0} \) der Wert \( \theta = 1.25 \). Die obere Grenze des Konfidenzintervales liegt klar oberhalb von \( \theta \) und daher kann eine sicherheitsrelevante Erhöhung der ASAT-Werte nicht ausgeschlossen werden.

### 3.2.2 Zweiseitiges Testproblem

Das zur Formulierung
\[
H_0^\delta: \mu_1 - \mu_0 \leq \delta \quad \text{oder} \quad \mu_1 - \mu_0 \geq \delta
\]
versus
\[
H_1^\delta: \delta < \mu_1 - \mu_0 < \delta
\]
äquivalente Testproblem für den Quotienten der Erwartungswerte lautet
\[
H_0^\theta: \frac{\mu_1}{\mu_0} \leq \theta \quad \text{oder} \quad \frac{\mu_1}{\mu_0} \geq \theta
\]
versus
\[
H_1^\theta: \theta < \frac{\mu_1}{\mu_0} < \theta
\]
mit dem Sicherheitsbereich \((\theta_1, \theta_2)\), \(\theta_1 = 1 + f_1, \theta_2 = 1 + f_2\) und \(-1 < f_1 < 0 < f_2\).

Das Testen der zweiseitigen Fragestellung ist äquivalent zur simultanen Überprüfung der beiden einseitigen Testprobleme:

\[
H_{01}^\theta: \frac{\mu_1}{\mu_0} \leq \theta \quad \text{versus} \quad H_{11}^\theta: \frac{\mu_1}{\mu_0} > \theta
\]

und

\[
H_{02}^\theta: \frac{\mu_1}{\mu_0} \geq \theta \quad \text{versus} \quad H_{12}^\theta: \frac{\mu_1}{\mu_0} < \theta
\]

Entsprechend des intersection-union-Prinzips wird \( H_0^\theta = \bigcup_{r=1}^{2} H_{0r}^\theta \) zum Niveau \( \alpha \) zugunsten von \( H_1^\theta = \bigcap_{r=1}^{2} H_{1r}^\theta \) abgelehnt, falls beide Einzelhypothesen \( H_{01}^\theta \) und \( H_{02}^\theta \) jeweils zum Niveau
Die Bedingungen $t_{11}^0 \geq t_{\alpha,n_0+n_1-2}$ und $t_{12}^0 \leq -t_{\alpha,n_0+n_1-2}$ sind äquivalent zu

$$[\theta_{-1}^1, \theta_{+1}^1 ] \subset (\theta_1, \theta_2) \quad \text{und} \quad \bar{X}_0^2 > a_0,$$

wobei

$$\theta_{\pm}^1 = \frac{\bar{X}_0 \bar{X}_1 \pm a_0 \bar{X}_1^2 + a_1 \bar{X}_0^2 - a_0 a_1}{\bar{X}_0^2 - a_0}.$$

Der Schluß auf Sicherheit zum Niveau $\alpha$ ist somit möglich, falls das Fiellersche 100(1-2$\alpha$)%-Konfidenzintervall $[\theta_{-1}^1, \theta_{+1}^1 ]$ regulär und vollständig im Sicherheitsbereich enthalten $(\theta_1, \theta_2)$ ist (Hauschke et al., 1999c).

Berechnet man für den Reversibilitätsversuch (Tabelle 3) nach der obigen Methode das zweiseitige 90%-Konfidenzintervall für $\frac{\mu_1}{\mu_0}$, so erhält man [0.94, 1.04]. Bei einem Sicherheitsbereich von $(\delta_1, \delta_2) = (-0.1 \mu_0, +0.1 \mu_0)$ für die Differenz $\mu_1 - \mu_0$ folgt ein Bereich von $(\theta_1, \theta_2) = (0.9, 1.1)$ für $\frac{\mu_1}{\mu_0}$. Das Konfidenzintervall ist im Sicherheitsbereich enthalten, und somit ist die Annahme der Reversibilität der durch die Substanz bedingten Körperformgewichtsveränderung bei diesen Vorgaben gerechtfertigt.

### 3.2.3 Gütebestimmung und Fallzahlplanung

Die Güte des einseitigen Sasabuchi-Tests $t_i^0$ ist die Wahrscheinlichkeit für die korrekte Annahme von $H_{0i}$, und läßt sich über die nichtzentrale $t$-Verteilung ermitteln:

$$P \left[ t_i^0 \leq -t_{\alpha,2n-2} \left| \frac{\mu_i}{\mu_0} < \theta_1, \sigma \right. \right] = \frac{\sqrt{2\pi}}{\Gamma(n-1)2^{n-2}} \int_0^\infty \Phi \left( \frac{-t_{\alpha,2n-2} x}{\sqrt{2n-2} - \theta} \right) x^{2n-3} \Phi'(x) \, dx,$$
wobei der Nichtzentralitätsparameter lautet:

\[
\Theta = \frac{\mu_1 - \theta \mu_0}{\sigma \sqrt{\frac{1 + \theta^2}{n}}} = \frac{\mu_1 - \theta}{CV_0 \sqrt{\frac{1 + \theta^2}{n}}}
\]

In Abbildung 6 wird die Gütefunktion für verschiedene Variationskoeffizienten dargestellt; dabei werden ein Sicherheitsbereich von \((-\infty, \theta) = (-\infty, 1.2)\) und 10 Versuchseinheiten pro Gruppe gewählt.

\[\text{Abbildung 6: Zweistichprobensituation: Die Gütefunktion des einseitigen Tests } t_1^0 \text{ bei einem Sicherheitsbereich von } (-\infty, \theta) = (-\infty, 1.2) \text{ für } \alpha = 0.05, n_0 = n_1 = 10 \text{ und verschiedene } CV_0\] 

Den notwendigen Stichprobenumfang \(n\) (pro Gruppe), um beim Testen von \(H_0^0\) versus \(H_1^0\) die Fehler \(\alpha\) und \(\beta\) an der Stelle \(\frac{\mu_1}{\mu_0}, \frac{\mu_1}{\mu_0} < \theta\), nicht zu überschreiten, kann man iterativ nach der folgenden Ungleichung berechnen (Hauschke et al., 1999b):

**Satz 2:**

Die Approximationsformel für die Fallzahlen (pro Gruppe) für eine Güte von mindestens \(1 - \beta\) bei vorgegebenem \(\alpha, CV_0\) und einem Sicherheitsbereich \((-\infty, \theta)\) lautet:
\[ n \geq (1 + \theta^2) \left( \frac{CV_0}{\theta - \frac{\mu_1}{\mu_0}} \right)^2 \left( t_{\alpha,2n-2} + t_{\beta,2n-2} \right)^2. \]

**Beweis:**

Die Wahrscheinlichkeit \( H_0 \) zu verwerfen ist

\[
P\left[ t_i^0 \leq -t_{\alpha,2n-2} \left| \frac{\mu_1}{\mu_0} < \theta, \sigma \right. \right] = P \left[ \frac{\bar{X}_i - \theta \bar{X}_0}{S} \leq -t_{\alpha,2n-2} \left| \frac{\mu_1}{\mu_0} < \theta, \sigma \right. \right] =
\]

\[
P \left[ \frac{\bar{X}_i - \theta \bar{X}_0 - (\mu_1 - \theta \mu_0)}{S} \leq -t_{\alpha,2n-2} - \frac{\mu_1 - \theta \mu_0}{S} \left| \frac{\mu_1}{\mu_0} < \theta, \sigma \right. \right].
\]

Verlangt man nun, daß die Power mindestens \( 1 - \beta \) beträgt und ersetzt man \( S^2 \) durch \( \sigma^2 \), so folgt analog zu Satz 1 die Behauptung:

\[ n \geq (1 + \theta^2) \left( \frac{CV_0}{\theta - \frac{\mu_1}{\mu_0}} \right)^2 \left( t_{\alpha,2n-2} + t_{\beta,2n-2} \right)^2 \]

\[ \Delta \]

In Tabelle 11 sind die Fallzahlen angegeben, um eine Güte von mindestens 0.80 zu erreichen. Dabei werden die weiteren Einflußgrößen wie folgt vorgegeben: \( \alpha = 0.05, \theta = 1.20, \frac{\mu_1}{\mu_0} = 1.00, 1.05, 1.10, 1.15 \) und \( CV_0 = 5.0, 7.5, 10.0, ..., 30.0\% \).
Die exakte Gütebestimmung für das zweiseitige Testproblem erfolgt direkt über die bivariate nichtzentrale $t$-Verteilung, wobei das folgende Ergebnis benötigt wird:

**Lemma 2:**

Die Korrelation zwischen den beiden Teststatistiken $t_{11}^0$ und $t_{12}^0$ lautet

$$Korr(t_{11}^0, t_{12}^0) = p(t_{11}^0, t_{12}^0) = \rho^0 = \frac{1+\theta_{12}}{\sqrt{(1+\theta_1^2)(1+\theta_2^2)}}.$$

**Beweis:**

Mit $Var(\bar{X}_1) = Var(\bar{X}_0) = \frac{\sigma^2}{n}$ folgt

$$Var(\bar{X}_1 - \theta, \bar{X}_0) = \frac{\sigma^2(1+\theta_r^2)}{n}, r = 1, 2,$$

$$Kov(\bar{X}_1 - \theta, \bar{X}_0, \bar{X}_1 - \theta_2 \bar{X}_0) = Var(\bar{X}_1) + \theta_1 \theta_2 Var(\bar{X}_0) = \frac{\sigma^2(1+\theta_1 \theta_2)}{n}$$
und damit
\[
\rho^0 = \frac{\sigma^2(1+\theta_1\theta_2)}{n} = \frac{1+\theta_1\theta_2}{\sqrt{(1+\theta_1^2)(1+\theta_2^2)}}.
\]

Die Wahrscheinlichkeit der korrekten Annahme von $H_1^0$ ist
\[
P\left( t_{11}^0 \geq t_{a,2n-2} \land t_{12}^0 \leq -t_{a,2n-2} \theta_1 < \frac{\mu_1}{\mu_0} < \theta_2, \sigma \right) = Q(\infty, -t_{a,2n-2}, \Theta_1, \Theta_2, \rho^0) - Q(t_{a,2n-2}, -t_{a,2n-2}, \Theta_1, \Theta_2, \rho^0),
\]

wobei
\[
Q(t_1, t_2, \Theta_1, \Theta_2, \rho) = P\left( t_{11}^0 \leq t_1 \land t_{12}^0 \leq t_2 \theta_1 < \frac{\mu_1}{\mu_0} < \theta_2, \sigma \right) = \frac{\sqrt{2\pi}}{\Gamma(n-1)2^{n-2}} \int_0^\infty \Phi_2\left( \frac{t_1x}{\sqrt{2n-2}} - \Theta_1, \frac{t_2x}{\sqrt{2n-2}} - \Theta_2, \rho \right) x^{2n-3} \Phi(x) dx,
\]

\[
\Phi_2(x, y, \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^\infty \int_{-\infty}^\infty \exp\left( -\frac{u^2 - 2\rho uv + v^2}{2(1-\rho^2)} \right) du dv,
\]

und für die entsprechenden Nichtzentralitätsparameter gilt
\[
\Theta_1 = \frac{\mu_1 - \theta_1 \mu_0}{\sigma \sqrt{(1+\theta_1^2)n}} \quad \text{und} \quad \Theta_2 = \frac{\mu_1 - \theta_2 \mu_0}{\sigma \sqrt{(1+\theta_2^2)n}}.
\]

Die Berechnung der obigen Integrale erfolgt mittels entsprechender SAS-Algorithmen (Genz und Bretz, 1999, siehe Anhang). In der Abbildung 7 werden die Gütefunktionen für einen Sicherheitsbereich von $(\theta_1, \theta_2) = (0.8, 1.25)$, einen Variationskoeffizienten von 15.0% und für Stichprobenumfänge von jeweils 10 und 20 pro Gruppe dargestellt. Die
Gütefunktionen sind nicht symmetrisch zu 1 und erreichen ihr Maximum ebenfalls für kleine Fallzahlen und/oder große Variationskoeffizienten nicht an der Stelle 1.

Abbildung 7: Zweistichprobensituation: Die Gütefunktion des Entscheidungsverfahrens, basierend auf den zwei einseitigen Tests \( t_{11}^0, t_{12}^0 \) bei einem Sicherheitsbereich von \((\theta_1, \theta_2) = (0.8, 1.25)\) für \( \alpha = 0.05, n_0 = n_1 = 10, 20 \) und \( CV_0 = 15.0\% \)

In den Abbildungen 8 und 9 ist die Stelle der maximalen Güte \( X_{max} \) in Abhängigkeit von der Fallzahl, vom Variationskoeffizienten und für verschiedene Sicherheitsbereiche dargestellt. Man erkennt, daß für hinreichend große Fallzahlen das Maximum an der Stelle 1 genau dann angenommen wird, wenn ein Sicherheitsbereich der Gestalt \((\theta_1, \theta_2) = \left(\theta_1, \frac{1}{\theta_1}\right)\) gewählt wird (Kieser et al., 1997).
Abbildung 8: Zweistichprobensituation: Stelle der maximalen Güte $X_{\text{max}}$ in Abhängigkeit von der Fallzahl und vom Variationskoeffizienten, basierend auf den zwei einseitigen Tests $t_{11}^\theta$, $t_{12}^\theta$ für $\alpha = 0.05$ bei einem Sicherheitsbereich von (a) $(\theta_1, \theta_2) = (0.8, 1.25)$, (b) $(\theta_1, \theta_2) = (0.8, 1.20)$
Abbildung 9: Zweistichprobensituation: Stelle der maximalen Güte $X_{\text{max}}$ in Abhängigkeit von der Fallzahl und vom Variationskoeffizienten, basierend auf den zwei einseitigen Tests $t_1^\theta \cdot t_2^\theta$ für $\alpha = 0.05$ bei einem Sicherheitsbereich von (a) $(\theta_1, \theta_2) = (0.7, 1.43)$, (b) $(\theta_1, \theta_2) = (0.7, 1.30)$
Zur vereinfachten Berechnung der Fallzahlen haben Kieser und Hauschke (1999a) die folgenden Approximationsformeln hergeleitet.

**Satz 3:**

Die Approximationsformeln für die Fallzahlen (pro Gruppe) für eine Güte von mindestens \( 1 - \beta \) bei vorgegebenem \( \alpha \), \( CV_0 \) für einen Sicherheitsbereich der Gestalt \( (\theta_1, \theta_2) = \left( \frac{1}{\theta_2}, \theta_2 \right) \) lauten:

i) falls \( \frac{\mu_1}{\mu_0} = 1 \)

\[
n \geq \left( 1 + \theta_2^2 \right) \left( \frac{CV_0}{\theta_2 - 1} \right)^2 \left( t_{\alpha, 2n-2} + t_{\beta, 2n-2} \right)^2,
\]

ii) \( 1 < \frac{\mu_1}{\mu_0} < \theta_2 \)

\[
n \geq \left( 1 + \theta_2^2 \right) \left( \frac{CV_0}{\theta_2 - \frac{\mu_1}{\mu_0}} \right)^2 \left( t_{\alpha, 2n-2} + t_{\beta, 2n-2} \right)^2
\]

und

iii) \( \frac{1}{\theta_2} < \frac{\mu_1}{\mu_0} < 1 \)

\[
n \geq \left( 1 + \frac{1}{\theta_2^2} \right) \left( \frac{CV_0}{\frac{1}{\theta_2} - \frac{\mu_1}{\mu_0}} \right)^2 \left( t_{\alpha, 2n-2} + t_{\beta, 2n-2} \right)^2
\]

**Beweis:**

Die Herleitung erfolgt analog zu Satz 1.

\[ \triangleright \]

In der Tabelle 12 sind die exakten (Hauschke et al., 1999c) und die approximativen (Kieser und Hauschke, 1999a) Fallzahlen für \( \alpha = 0.05 \), \( \beta = 0.20 \), \( (\theta_1, \theta_2) = (0.8, 1.25) \), \( CV_0 = 5.0 \), 10.0, ..., 30.0% und für verschiedene Quotienten \( \frac{\mu_1}{\mu_0} \) angegeben.
In diesem dritten Kapitel wurden die Entscheidungsverfahren für den direkten SicherheitshexlNachweis dargestellt. Desweiteren erfolgte eine exakte Gütebestimmung, und neben der korrespondierenden exakten Fallzahlbestimmung wurden zusätzlich Approximationsformeln zur Verfügung gestellt. Vergleicht man in den jeweiligen Tabellen die approximativen mit den exakt ermittelten Werten, so läßt sich eine weitgehende Übereinstimmung feststellen.

Es sei darauf hingewiesen, daß für den Quotienten im Parallelgruppenvergleich bislang nur parametrische Methoden existieren. Eine finite Adaption auf nichtparametrische Verfahren ist nicht trivial, da schon unter der Normalverteilungsannahme \( F_i(x) = \Phi \left( \frac{x - \mu_i}{\sigma} \right) \), \( i = 0, 1 \), wegen \( Var(\bar{X}_i) \neq Var(\bar{\theta}X_0) \) die Voraussetzung des Lokationsmodells verletzt ist.

### Tabelle 12: Zweistichproben-Situation: Exakter und approximativer Stichprobenumfang pro Gruppe für eine Güte von mindestens 0.80 bei \( \alpha = 0.05 \), \( (\theta_1, \theta_2) = (0.8, 1.25) \) und verschiedene \( CV_0 \)

<table>
<thead>
<tr>
<th>( CV_0(%) )</th>
<th>0.85</th>
<th>0.90</th>
<th>0.95</th>
<th>( \mu_i/\mu_0 )</th>
<th>1.00</th>
<th>1.05</th>
<th>1.10</th>
<th>1.15</th>
<th>1.20</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.0</td>
<td>11</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>7.5</td>
<td>24</td>
<td>7</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>10</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>10.0</td>
<td>42</td>
<td>11</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>8</td>
<td>17</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>12.5</td>
<td>65</td>
<td>17</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>12</td>
<td>26</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>15.0</td>
<td>92</td>
<td>24</td>
<td>12</td>
<td>9</td>
<td>10</td>
<td>17</td>
<td>37</td>
<td>144</td>
<td></td>
</tr>
<tr>
<td>17.5</td>
<td>125</td>
<td>32</td>
<td>15</td>
<td>12</td>
<td>14</td>
<td>23</td>
<td>50</td>
<td>195</td>
<td></td>
</tr>
<tr>
<td>20.0</td>
<td>163</td>
<td>42</td>
<td>20</td>
<td>15</td>
<td>18</td>
<td>29</td>
<td>65</td>
<td>255</td>
<td></td>
</tr>
<tr>
<td>22.5</td>
<td>207</td>
<td>53</td>
<td>25</td>
<td>19</td>
<td>22</td>
<td>37</td>
<td>81</td>
<td>322</td>
<td></td>
</tr>
<tr>
<td>25.0</td>
<td>255</td>
<td>65</td>
<td>30</td>
<td>23</td>
<td>27</td>
<td>45</td>
<td>100</td>
<td>397</td>
<td></td>
</tr>
<tr>
<td>27.5</td>
<td>308</td>
<td>78</td>
<td>36</td>
<td>28</td>
<td>32</td>
<td>54</td>
<td>121</td>
<td>480</td>
<td></td>
</tr>
<tr>
<td>30.0</td>
<td>366</td>
<td>92</td>
<td>43</td>
<td>33</td>
<td>38</td>
<td>65</td>
<td>144</td>
<td>572</td>
<td></td>
</tr>
</tbody>
</table>

In diesem dritten Kapitel wurden die Entscheidungsverfahren für den direkten Sicherheits- nachweis dargestellt. Desweiteren erfolgte eine exakte Gütebestimmung, und neben der korrespondierenden exakten Fallzahlbestimmung wurden zusätzlich Approximationsformeln zur Verfügung gestellt. Vergleicht man in den jeweiligen Tabellen die approximativen mit den exakt ermittelten Werten, so läßt sich eine weitgehende Übereinstimmung feststellen.

Es sei darauf hingewiesen, daß für den Quotienten im Parallelgruppenvergleich bislang nur parametrische Methoden existieren. Eine finite Adaption auf nichtparametrische Verfahren ist nicht trivial, da schon unter der Normalverteilungsannahme \( F_i(x) = \Phi \left( \frac{x - \mu_i}{\sigma} \right) \), \( i = 0, 1 \), wegen \( Var(\bar{X}_i) \neq Var(\bar{\theta}X_0) \) die Voraussetzung des Lokationsmodells verletzt ist.
4 Crossover-Situation

Ein ursprünglich aus der Landwirtschaft stammender Versuchsplan (Jones und Kenward, 1989) ist das Crossover-Design mit zwei Perioden und zwei Behandlungen. Jede Versuchseinheit erhält beide Behandlungen, wobei in der ersten Sequenz die Behandlungen in der Reihenfolge (0, 1) und in der zweiten Sequenz in der umgekehrten Reihenfolge (1, 0) verabreicht werden. Die Zuordnung der Versuchseinheiten zu den beiden Sequenzen erfolgt zufällig. Ein wesentlicher Vorteil im Vergleich zum Parallelgruppenvergleich besteht darin, daß die intraindividuele Variabilität, d.h. Variabilität innerhalb der Versuchseinheit, meist wesentlich geringer ist als die interindividuelle Variabilität und somit geringere Stichprobenumfänge als im Parallelgruppenvergleich resultieren. Jedoch ist es auch offensichtlich, daß dieses Design nur dann zur Anwendung kommen kann, falls eine vollständige Reversibilität der zu untersuchenden Effekte vorliegt und daher nicht die Applikation der jeweiligen Behandlung in der zweiten Periode in Frage stellt. Zwischen den beiden Perioden ist somit eine hinreichend lange Washout-Phase zu wählen, so daß mögliche Carryover-Effekte auszuschließen sind. Die Zufallsvariable \( X_{ijk} \) bezeichnet die Beobachtung der Zielvariablen bei der \( j \)-ten Versuchseinheit, \( j = 1, \ldots, n_j \), in der \( i \)-ten Sequenz, \( i = 0, 1 \) aus der \( k \)-ten Periode, \( k = 1, 2 \).

Das allgemeine bivariate Modell, welches im folgenden als Modell I bezeichnet wird, lautet:

\[
X_{ijk} = \mu_h + \pi_k + \epsilon_{ijk},
\]

wobei \( \mu_h \) den Effekt der Behandlung \( h, h = 0 \) für \( i +1 = k \) und \( h = 1 \) für \( i +1 \neq k \) bezeichnet und \( \pi_k \) die Perioden-Effekte darstellen mit \( \pi_1 + \pi_2 = 0 \). Weiterhin wird angenommen, daß für die bivariaten Zufallsvektoren gilt: \( (\epsilon_{ij1}, \epsilon_{ij2})' \) u.i.v. \( \sim F_i(x_1, x_2, \text{stetig}, i = 0, 1 \). In der Tabelle 13 ist das entsprechende Schema der Beobachtungen angegeben.

<table>
<thead>
<tr>
<th>Tabelle 13: Schema der Zufallsvariablen im Zwei-Perioden-Crossover-Design mit zwei Behandlungen</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>(0, 1)</td>
</tr>
<tr>
<td>(1.0)</td>
</tr>
</tbody>
</table>
Das klassische univariate Modell (Grizzle, 1965):

\[ X_{ijk} = \mu + \pi_k + b_{ij} + e_{ijk}, \]

bezeichnet im folgenden als Modell II, ist mit \( e_{ijk} = b_{ij} + e_{ijk} \) als Spezialfall im allgemeinen Modell I enthalten. Dabei bezeichnet \( b_{ij} \) den zufälligen Effekt der \( j \)-ten Versuchseinheit mit \( b_{ij} \), u.i.v. \( \sim F_{b}(x) = \Phi \left( \frac{x}{\sigma_b} \right) \) und \( e_{ijk} \) den zufälligen Fehlerterm mit \( e_{ijk} \), u.i.v. \( \sim F_{e}(x) = \Phi \left( \frac{x}{\sigma_e} \right) \); die Zufallsvariablen \( b_{ij} \) und \( e_{ijk} \) werden als unabhängig vorausgesetzt. Neben der Annahme identischer Varianzen setzt Modell II auch stets positive Kovarianzen für die Beobachtungen der ersten und zweiten Periode voraus:

\[ \text{Kov}(X_{ijk}, X_{ijl}) = \begin{cases} \sigma_b^2 & k \neq l \\ \sigma_b^2 + \sigma_e^2 & k = l \end{cases}, \]

Da diese Voraussetzungen sicherlich sehr restriktiv sein können (Lehmacher 1987, 1991; Hauschke et al., 1996), werden statistische Entscheidungsverfahren sowohl für Modell I als auch für Modell II hergeleitet.

4.1 Sicherheitsnachweis für die Differenz von Lokationsparametern

Seien \( Y_{ij}^{\delta} \), die \( j \)-ten intraindividuellen Periodendifferenzen der Beobachtungen in der 1. und 2. Periode innerhalb der \( i \)-ten Sequenz: \( Y_{ij}^{\delta} = X_{ij1} - X_{ij2} \), und ferner seien \( H_i(y) \) die stetigen Verteilungsfunktionen der Differenzen der zugehörigen Fehlerterme:

\[ H_i(y) = P(\varepsilon_{ij1} - \varepsilon_{ij2} \leq y), \quad i = 0, 1, j = 1, ..., n_i. \]

Verlangt man \( H_0(y) = H_1(y) = H(y) \), so folgt unmittelbar:

\[ Y_{ij}^{\delta} \text{ u.i.v. } \sim H(y + \mu_0 - \mu_1 + \pi_2 - \pi_1), \quad j = 1, ..., n_0, \]

\[ Y_{1j}^{\delta} \text{ u.i.v. } \sim H(y + \mu_0 - \mu_1 + \pi_2 - \pi_1), \quad j = 1, ..., n_1. \]
d.h. die Verteilungen der Differenzen der beiden Sequenzen unterscheiden sich nur durch den Lokationsunterschied von \(2(\mu_1 - \mu_0)\). Daher können Hypothesen über \(\mu_1 - \mu_0\) durch Verfahren für zwei unverbundene Stichproben untersucht werden.

### 4.1.1 Einseitiges Testproblem

Für die parametrische Analyse im Modell I wird eine bivariate Normalverteilung mit Erwartungswertvektor \((0, 0)\) und Varianz-Kovarianzmatrizen \(\Sigma_i\) angenommen, d.h. \((\epsilon_{y1}, \epsilon_{y2})'\) u.i.v. \(\sim F_i(x_1, x_2) = N_2(0, \Sigma_i)\), wobei

\[
\Sigma_0 = \begin{pmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{pmatrix} \quad \text{und} \quad \Sigma_1 = \begin{pmatrix} \sigma_1^2 & \sigma_{01} \\ \sigma_{01} & \sigma_0^2 \end{pmatrix}.
\]

Zur Konstruktion eines Testverfahren benötigt man das folgende Lemma.

**Lemma 3:**

Der Schätzer \(\frac{\bar{X}_1 - \bar{X}_0}{2}\) ist normalverteilt mit Erwartungswert \(\mu_1 - \mu_0\) und Varianz

\[
\sigma_\delta^2 = \frac{1}{4} \left( \frac{1}{n_0} + \frac{1}{n_1} \right),
\]

wobei

\[
\sigma_\delta^2 = \sigma_0^2 + \sigma_1^2 - 2\sigma_{01}
\]

und

\[
\bar{X}_0 = \left( \frac{1}{n_0} \sum_{j=1}^{n_0} X_{0j1} + \frac{1}{n_1} \sum_{j=1}^{n_1} X_{1j2} \right) \quad \text{und} \quad \bar{X}_1 = \left( \frac{1}{n_0} \sum_{j=1}^{n_0} X_{0j2} + \frac{1}{n_1} \sum_{j=1}^{n_1} X_{1j1} \right).
\]

**Beweis:**

Die Verteilungsfunktionen der Periodendifferenzen sind:

\[
Y_{y_i}^\delta \text{ u.i.v. } H(y - E(Y_{y_i}^\delta)) = \Phi \left( \frac{y - E(Y_{y_i}^\delta)}{\text{Var}(Y_{y_i}^\delta)} \right),
\]

mit

\[
E(Y_{y_i}^\delta) = \mu_0 - \mu_1 + \pi_1 - \pi_2, \quad E(Y_{i_j}^\delta) = \mu_1 - \mu_0 + \pi_1 - \pi_2,
\]

\[
\text{Var}(Y_{y_i}^\delta) = \sigma_\delta^2 = \sigma_0^2 + \sigma_1^2 - 2\sigma_{01}, \quad i = 0, 1 \quad \text{und} \quad j = 1, \ldots, n_i.
\]
Der Erwartungswert für die Differenz der transformierten Zufallsvariablen aus den beiden Stichproben beträgt \( E(Y_{1j}^\delta - Y_{0j}^\delta) = 2(\mu_1 - \mu_0) \); ein erwartungstreuer Schätzer für die Differenz \( \mu_1 - \mu_0 \) ist somit:

\[
\frac{\bar{Y}_1^\delta - \bar{Y}_0^\delta}{2} = \frac{1}{2} \left( \frac{1}{n_1} \sum_{j=1}^{n_1} (X_{1j1} - X_{1j2}) - \frac{1}{n_0} \sum_{j=1}^{n_0} (X_{0j1} - X_{0j2}) \right) = \frac{\bar{X}_1 - \bar{X}_0}{2}.
\]

Die Varianz des Schätzers ist

\[
Var\left( \frac{\bar{X}_1 - \bar{X}_0}{2} \right) = Var\left( \frac{1}{2} \left( \frac{1}{n_1} \sum_{i=1}^{n_1} Y_{1ij} + \frac{1}{n_0} \sum_{i=1}^{n_0} Y_{0ij} \right) \right) = \sigma_\delta^2 \frac{1}{4} \left( \frac{1}{n_0} + \frac{1}{n_1} \right),
\]
da

\[
Var(Y_{ij}^\delta) = \sigma_\delta^2, \quad i = 0, 1 \text{ und } j = 1, \ldots, n_i.
\]

Mittels der Periodendifferenzen wurde das \emph{Crossover}-Design auf die Zweistichproben-situation reduziert

\[
Y_{ij}^\delta \text{ u.i.v. } \sim \Phi \left( \frac{y - E(Y_{ij}^\delta)}{Var(Y_{ij}^\delta)} \right),
\]

\[
E(Y_{ij}^\delta) = (-1)^{i+1} (\mu_1 - \mu_0) + \pi_1 - \pi_2, \quad Var(Y_{ij}^\delta) = \sigma_\delta^2.
\]

\[
i = 0, 1 \text{ und } j = 1, \ldots, n_i.
\]

Die Varianz \( \sigma_\delta^2 \) kann somit erwartungstreue durch \( S_\delta^2 \) geschätzt werden, mit

\[
S_\delta^2 = \frac{1}{n_0 + n_1 - 2} \sum_{j=0}^{1} \sum_{i=1}^{n_i} (Y_{ij}^\delta - \bar{Y}_i^\delta)^2.
\]

Das einseitig formulierte Testproblem

\[
H_0^\delta: \quad \mu_1 - \mu_0 \geq \delta \quad \text{versus} \quad H_1^\delta: \quad \mu_1 - \mu_0 < \delta
\]
wird somit abgelehnt, falls

\[ t_1^\delta = \frac{\bar{X}_1 - \bar{X}_0 - \delta}{S_\delta} \leq -t_{\alpha, n_0 + n_1 - 2}. \]

Bei einer *a-posteriori* vorgenommenen Festlegung von \( \delta \) kann mit einer Irrtumswahrscheinlichkeit von \( \alpha \) auf Unbedenklichkeit geschlossen werden, falls die obere Grenze \( \delta^1 \) des einseitigen 100(1-\( \alpha \))-\%-Konfidenzintervales für die Differenz \( \mu_1 - \mu_0 \) unterhalb von \( \delta \) liegt:

\[ (-\infty, \delta^1] = \left[ -\infty, \frac{\bar{X}_1 - \bar{X}_0}{2} + t_{\alpha, n_0 + n_1 - 2} S_\delta \sqrt{\frac{1}{4} \left( \frac{1}{n_0} + \frac{1}{n_1} \right)} \right] \subset (-\infty, \delta). \]

Im Modell II ergibt sich als Unterschied zu Modell I nur ein anderer Varianzschätzer, da

\[ \text{Var}(Y^\delta_0) = \sigma^2_\delta = \sigma^2_0 + \sigma^2_1 - 2\sigma_{01} = 2(\sigma^2_\delta + \sigma^2_\epsilon) - 2\sigma^2_\delta = 2\sigma^2_\epsilon, \]

\[ i = 0, 1 \text{ und } j = 1, ..., n_i. \]

Die Varianz des Schätzers \( \frac{\bar{X}_1 - \bar{X}_0}{2} \) beträgt nun \( \sigma^2_\epsilon \frac{1}{2} \left( \frac{1}{n_0} + \frac{1}{n_1} \right) \), und ein erwartungstreuer Schätzer (Jones und Kenward, 1989) für \( \sigma^2_\epsilon \) ist

\[ S^2_\epsilon = \frac{1}{n_0 + n_1 - 2} \left( \sum_{i=0}^{n_0} \sum_{j=1}^{n_1} X^2_{ijk} - \sum_{i=0}^{n_0} \sum_{j=1}^{n_1} \frac{X^2_{ij}}{2} - \sum_{i=0}^{n_0} \frac{X^2_{i1}}{n_j} + \sum_{i=0}^{n_0} \frac{X^2_{i2}}{2n_j} \right), \]

mit

\[ X^2_{ij} = (X^2_{ij1} + X^2_{ij2}), \quad X^2_{ijk} = \left( \sum_{j=1}^{n_1} X_{ijk} \right)^2 \text{ und } X^2_{k} = \left( \sum_{j=1}^{n_1} \sum_{k=1}^{n} X_{ijk} \right)^2. \]

Aufgrund der Reduktion auf die Zweistichprobensituation, gilt zwischen den beiden Varianzschätzern die folgende Beziehung (Hauschke und Steinijans, 1991):

\[ S^2_\epsilon = \frac{S^2_{\delta}}{2}. \]
$H_0^\delta$ wird im Modell II abgelehnt, falls

$$
t^\delta_1 = \frac{\bar{X}_i - \bar{X}_0}{\frac{2}{S_\varepsilon}} \cdot \left( \frac{1}{\sqrt{n_0 + 1}} \right) \cdot S \cdot \left( \frac{1}{\sqrt{2(n_0 + n_1)}} \right) \leq -t_{\alpha,n_0+n_1-2}
$$

$$
\Leftrightarrow
$$

$$
(-\infty, \delta^\ast_1] = \left[ -\infty, \frac{\bar{X}_i - \bar{X}_0}{2} + t_{\alpha,n_0+n_1-2} \cdot \frac{1}{\sqrt{2(n_0 + n_1)}} \right] \subset (-\infty, \delta).
$$

Verzichtet man auf die Normalverteilungsannahme, so kann wegen des vorliegenden Lokationsmodells

$$
Y_{0j}^\delta \text{ u.i.v. } \sim H(y + \mu_1 - \mu_0 + \pi_2 - \pi_1), \quad j = 1, ..., n_0,
$$

$$
Y_{1j}^\delta \text{ u.i.v. } \sim H(y + \mu_0 - \mu_1 + \pi_2 - \pi_1), \quad j = 1, ..., n_1,
$$

für beide Modelle der Wilcoxon-Test und das verteilungsfreie Konfidenzintervall nach Moses als Entscheidungsverfahren herangezogen werden. Dabei muß jedoch beachtet werden, daß die Rangsumme der um $2\delta$ verschobenen $j$-ten intraindividuellen Periodendifferenzen verwendet wird. Sei $R_{ij}(2\delta)$ der Rang von $Y_{ij}^\delta - 2\delta$ in der kombinierten Stichprobe

$$
Y_{0j}^\delta, ..., Y_{0n_0}^\delta, Y_{1j}^\delta - 2\delta, ..., Y_{in_1}^\delta - 2\delta,
$$

und $R_i(2\delta) = \sum_{j=1}^{n} R_{ij}(2\delta)$. $H_0^\delta$ wird verworfen, falls $R_i(2\delta) \leq n_1(n_0 + n_1 + 1) - r_{\alpha,n_0,n_1}^{2\delta}$, wobei $r_{\alpha,n_0,n_1}^{2\delta}$ das (1-$\alpha$)-Quantil der Wilcoxonstatistik $R_i(2\delta)$ ist. Die Konstruktion des Konfidenzintervales basiert wiederum auf den geordneten Werten $D_{1}^{\delta} \leq ... \leq D_{n_{0}n_{1}}^{\delta}$ der insgesamt $n_{0}n_{1}$ Differenzen $Y_{ij}^\delta - Y_{0j}^\delta$, $j = 1, ..., n_1$ und $j^* = 1, ..., n_0$. Damit ergibt sich als einseitiges 100(1-$\alpha$)%-Konfidenzintervall für $\mu_1 - \mu_0 :$

$$
\left[ -\infty, \frac{U_{1}^{\delta}}{2} \right], \quad \text{wobei} \quad U_{1}^{\delta} = D_{n}^{\delta} = D_{n_{0}n_{1}+1-c_{2}}^{\delta}, \quad C_{\alpha}^{1} = \frac{n_{1}(2n_{0} + n_{1} + 1)}{2} + 1 - r_{\alpha,n_0,n_1}^{2\delta}.$$
4.1.2 Zweiseitiges Testproblem

Die zweiseitige Formulierung des Entscheidungsproblems

\[ H_0^\delta : \; \mu_1 - \mu_0 \leq \delta \quad \text{oder} \quad \mu_1 - \mu_0 \geq \delta \]

versus

\[ H_1^\delta : \; \delta < \mu_1 - \mu_0 < \delta \]

wird wieder in zwei einseitige Testprobleme zerlegt

\[ H_{01}^\delta : \; \mu_1 - \mu_0 \leq \delta_1 \quad \text{versus} \quad H_{11}^\delta : \; \mu_1 - \mu_0 > \delta_1 \]

und

\[ H_{02}^\delta : \; \mu_1 - \mu_0 \geq \delta_2 \quad \text{versus} \quad H_{12}^\delta : \; \mu_1 - \mu_0 < \delta_2 \]

Die Nullhypothese \( H_0^\delta = \bigcup_{r=1}^{2} H_{0r}^\delta \) wird im Modell I zum Niveau \( \alpha \) genau dann abgelehnt, falls

\[
 t_{11}^\delta = \frac{\bar{X}_1 - \bar{X}_0}{S_{\delta}} \geq t_{\alpha,n_0+n_1-2} \quad \text{und} \quad t_{12}^\delta = \frac{\bar{X}_1 - \bar{X}_0}{S_{\delta}} \leq -t_{\alpha,n_0+n_1-2},
\]

was wiederum äquivalent ist zur Inklusion des 100(1-2\( \alpha \))%-Konfidenzintervales \([\delta_1^-, \delta_1^+]\) für \( \mu_1 - \mu_0 \) im Sicherheitsbereich:

\[
 [\delta_1^-, \delta_1^+] = \left[ \frac{\bar{X}_1 - \bar{X}_0}{2} \pm t_{\alpha,n_0+n_1-2} S_{\delta} \sqrt{\frac{1}{n_0} + \frac{1}{n_1}} \right] \subset (\delta_1, \delta_2).
\]

Analog gilt im Modell II, daß \( H_0^\delta \) zum Niveau \( \alpha \) genau dann abgelehnt wird, falls

\[
 t_{11}^\delta = \frac{\bar{X}_1 - \bar{X}_0}{S_{\delta}} \geq t_{\alpha,n_0+n_1-2} \quad \text{und} \quad t_{12}^\delta = \frac{\bar{X}_1 - \bar{X}_0}{S_{\delta}} \leq -t_{\alpha,n_0+n_1-2},
\]

\( \iff \)
\[ [\delta_1^-, \delta_1^+] = \left[ \frac{\bar{X}_1 - \bar{X}_0}{2} \right. \pm \frac{t_{a,n_0+n_1-2} S}{\sqrt{2 \left( \frac{1}{n_0} + \frac{1}{n_1} \right)}} \left. \right] \subset (\delta_1, \delta_2). \]

Für das entsprechende verteilungsfreie Verfahren in beiden Modellen mit Zweistichproben-Wilcoxon-Tests gilt

\[ H_{01}^\delta \text{ wird abgelehnt, falls } R_1(2\delta_1) \geq r_{a,n_0,n_1}^{2\delta_1}. \]

\[ \text{und } H_{02}^\delta \text{ wird abgelehnt, falls } R_1(2\delta_2) \leq n_1(n_0 + n_1 + 1) - r_{a,n_0,n_1}^{2\delta_2}. \]

Die Verwerfung beider Einzelhypothesen ist äquivalent zur Inklusion des verteilungsfreien zweiseitigen 100(1-2\alpha)%-Konfidenzintervales \[ \left[ \frac{L_1^\delta}{2}, \frac{U_1^\delta}{2} \right] \] für \( \mu_1 - \mu_0 \) im Sicherheitsbereich \( (\delta_1, \delta_2) \) wobei \( L_1^\delta = D_1^\delta = D_{\epsilon_a}^\delta \) und \( U_1^\delta = D_1^\delta = D_{n_1+n_1-1-C_1^\delta}^\delta \).

### 4.1.3 Gütebestimmung und Fallzahlbestimmung

Unter Annahme des Modells II kann die Güte des einseitigen Zweistichproben-\( t \)-Tests \( t_1^\delta \) mittels der nichtzentralen \( t \)-Verteilung analog zur Zweistichprobensituation bestimmt werden:

\[ P[t_1^\delta \leq -t_{a,2n-2} | \mu_1 - \mu_0 < \delta, \sigma_e] = \frac{\sqrt{2\pi}}{\Gamma(n-1)2^{n-2}} \int_0^\infty \Phi \left( \frac{-t_{a,2n-2} x + \Delta}{\sqrt{2n-2}} \right) x^{2n-3} \Phi(x) \, dx, \]

mit Nichtzentralitätsparameter

\[ \Delta = \frac{\hat{\delta} - \hat{\delta}^*}{CV_e \sqrt{\frac{1}{n}}}, \text{ wobei } \delta = \hat{\delta}^* \mu_0, \mu_1 - \mu_0 = \hat{\delta} \mu_0, CV_e = \frac{\sigma_e}{\mu_0}. \]

In Abbildung 10 wird die Gütefunktion für verschiedene Variationskoeffizienten bei einem Sicherheitsbereich von \( (-\infty, \delta) = (-\infty, 0.2\mu_0) \) und jeweils 10 Versuchseinheiten pro Sequenz dargestellt.
Abbildung 10: Crossover-Situation: Die Gütefunktion des einseitigen Tests $t_1^\delta$ bei einem Sicherheitsbereich von $(-\infty, \delta) = (-\infty, 0.2\mu_0)$ für $\alpha = 0.05$, $n_0 = n_1 = 10$ und verschiedene $CV_e$.

Den benötigten Stichprobenumfang $n$ (pro Sequenz) kann man iterativ nach der folgenden Ungleichung bestimmen:

$$n \geq \left( \frac{CV_e}{\delta^* - \delta} \right)^2 \left( t_{1,2n-2} + t_{p,2n-2} \right)^2,$$

wobei die Herleitung analog zur entsprechenden Ungleichung in der Zweistichproben-situation erfolgt.

In Tabelle 14 sind die benötigten Fallzahlen angegeben, um eine Güte von mindestens 0.80 zu erreichen. Dabei werden die weiteren Einflussgrößen wie folgt vorgegeben: $\alpha = 0.05$, $\delta^* = 0.20$, $\delta = 0.00$, 0.05, 0.10, 0.15 und $CV_e = 5.0$, 7.5, 10.0, ..., 30.0%.

Da die Korrelation $\rho^\delta$ zwischen den Teststatistiken $t_{11}^\delta$ und $t_{12}^\delta$ genau dem Wert 1 entspricht, wird die Wahrscheinlichkeit für die korrekte Verwerfung für das zweiseitige Testproblem unter Ausnutzung der Owen-Technik berechnet:

$$P[t_{11}^\delta \geq t_{a,2n-2} \wedge t_{12}^\delta \leq -t_{a,2n-2} | \delta_1 < \mu_1 - \mu_0 < \delta_2, \sigma_e] =$$

$$Q(-t_{a,2n-2}, \Delta_z, 0, \psi) - Q(t_{a,2n-2}, \Delta_z, 0, \psi),$$

wobei
\[ \Delta_i = \frac{\hat{\theta} - \hat{\theta}_1}{CV \sqrt{\frac{1}{n}}} \quad \text{und} \quad \Delta_2 = \frac{\hat{\theta} - \hat{\theta}_2}{CV \sqrt{\frac{1}{n}}} . \]

Tabelle 14: Crossover-Situation: Exakter und approximativer Stichprobenumfang pro Sequenz für eine Güte von mindestens 0.80 bei \( \alpha = 0.05 \), \((-\infty, \delta) = (-\infty, 0.2\mu_0)\) und verschiedene \( CV_e \)

<table>
<thead>
<tr>
<th>( CV_e (%) )</th>
<th>0.00</th>
<th>0.05</th>
<th>0.10</th>
<th>0.15</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.0</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>7.5</td>
<td>3</td>
<td>3</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>10.0</td>
<td>3</td>
<td>4</td>
<td>7</td>
<td>26</td>
</tr>
<tr>
<td>12.5</td>
<td>4</td>
<td>6</td>
<td>11</td>
<td>40</td>
</tr>
<tr>
<td>15.0</td>
<td>5</td>
<td>7</td>
<td>15</td>
<td>57</td>
</tr>
<tr>
<td>17.5</td>
<td>6</td>
<td>10</td>
<td>20</td>
<td>77</td>
</tr>
<tr>
<td>20.0</td>
<td>7</td>
<td>12</td>
<td>26</td>
<td>100</td>
</tr>
<tr>
<td>22.5</td>
<td>9</td>
<td>15</td>
<td>32</td>
<td>126</td>
</tr>
<tr>
<td>25.0</td>
<td>11</td>
<td>18</td>
<td>40</td>
<td>156</td>
</tr>
<tr>
<td>27.5</td>
<td>13</td>
<td>22</td>
<td>48</td>
<td>188</td>
</tr>
<tr>
<td>30.0</td>
<td>15</td>
<td>26</td>
<td>57</td>
<td>224</td>
</tr>
</tbody>
</table>

In der Abbildung 11 werden die Gütefunktionen für einen symmetrischen Sicherheitsbereich von \((-0.2\mu_0, 0.2\mu_0)\), einen Variationskoeffizienten von 15.0\% und für die Stichprobenumfänge von jeweils 10 bzw. 20 pro Sequenz dargestellt.

Die vereinfachte Berechnung der Fallzahlen erfolgt analog zur Zweistichprobensituation nach den folgenden Approximationsformeln, wobei ebenfalls ein zu 0 symmetrischer Sicherheitsbereich gewählt wurde, d.h. \( \hat{\theta}_1 = -\hat{\theta}_2 \):

falls \( \hat{\theta} = 0 \):

\[ n \geq \left( \frac{CV_e}{\hat{\theta}_2} \right)^2 \left( t_{\alpha,2n-2} + t_{\beta,2n-2} \right)^2 , \]

\( 0 < \hat{\theta} < \hat{\theta}_2 \):

\[ n \geq \left( \frac{CV_e}{\hat{\theta}_2 - \hat{\theta}} \right)^2 \left( t_{\alpha,2n-2} + t_{\beta,2n-2} \right)^2 \]

und \(-\hat{\theta}_2 < \hat{\theta} < 0 \):

\[ n \geq \left( \frac{CV_e}{-\hat{\theta}_2 - \hat{\theta}} \right)^2 \left( t_{\alpha,2n-2} + t_{\beta,2n-2} \right)^2 \].
Abbildung 11: Crossover-Situation: Die Gütefunktion des Entscheidungsverfahrens, basierend auf den zwei einseitigen Tests $t_{11}^{\delta_1}, t_{12}^{\delta_2}$ bei einem Sicherheitsbereich von $(\delta_1, \delta_2) = (-0.2\mu_0, 0.2\mu_0)$ für $\alpha = 0.05, n_0 = n_1 = 10, 20$ und $CV_e = 15.0\%$

In der Tabelle 15 sind die exakten und approximativen Fallzahlen für $\alpha = 0.05, \beta = 0.20, \delta_1 = -\delta_2 = -0.20, CV_e = 5.0, 10.0, \ldots, 30.0\%$ und für $|\delta| = 0.00, 0.02, 0.05, 0.10, 0.15$ angegeben.

**Tabelle 15: Crossover-Situation: Exakter und approximativer Stichprobenumfang pro Sequenz für eine Güte von mindestens 0.80 bei $\alpha = 0.05, (\delta_1, \delta_2) = (-0.2\mu_0, 0.2\mu_0)$ und verschiedene $CV_e$**

<table>
<thead>
<tr>
<th>$CV_e$ (%)</th>
<th>0.00</th>
<th>0.05</th>
<th>0.10</th>
<th>0.15</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.0</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>7.5</td>
<td>3</td>
<td>3</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>10.0</td>
<td>4</td>
<td>4</td>
<td>7</td>
<td>26</td>
</tr>
<tr>
<td>12.5</td>
<td>5</td>
<td>6</td>
<td>11</td>
<td>40</td>
</tr>
<tr>
<td>15.0</td>
<td>6</td>
<td>7</td>
<td>15</td>
<td>57</td>
</tr>
<tr>
<td>17.5</td>
<td>8</td>
<td>10</td>
<td>20</td>
<td>77</td>
</tr>
<tr>
<td>20.0</td>
<td>10</td>
<td>12</td>
<td>26</td>
<td>100</td>
</tr>
<tr>
<td>22.5</td>
<td>12</td>
<td>15</td>
<td>32</td>
<td>126</td>
</tr>
<tr>
<td>25.0</td>
<td>15</td>
<td>19</td>
<td>40</td>
<td>156</td>
</tr>
<tr>
<td>27.5</td>
<td>17</td>
<td>22</td>
<td>48</td>
<td>188</td>
</tr>
<tr>
<td>30.0</td>
<td>20</td>
<td>26</td>
<td>57</td>
<td>224</td>
</tr>
</tbody>
</table>
4.2 Sicherheitsnachweis für den Quotienten von Lokationsparametern

Seien
\[ y_{i}^{0} = \begin{cases} \theta X_{i1} - X_{i2} & i = 0 \\ X_{i1} - \theta X_{i2} & i = 1 \end{cases} \]
die um \( \theta \) korrigierten \( j \)-ten intraindividuellen Periodendifferenzen, \( j = 1, \ldots, n_{j} \), und seien ferner \( G_{i}(y) \) die Verteilungsfunktionen der Differenzen der zugehörigen Fehlerterme:

\[ G_{i}(y) = \begin{cases} P(\theta \varepsilon_{g1} - \varepsilon_{g2} \leq y) & i = 0 \\ P(\varepsilon_{g1} - \theta \varepsilon_{g2} \leq y) & i = 1 \end{cases} . \]

Mit der Voraussetzung \( G_{0}(y) = G_{1}(y) = G(y) \) folgt unmittelbar:

\[ Y_{0j}^{0} \text{ u.i.v. } \sim G(y + \mu_{1} - \theta \mu_{0} + \pi_{2} - \theta \pi_{1}) , \quad j = 1, \ldots, n_{0} , \]
\[ Y_{1j}^{0} \text{ u.i.v. } \sim G(y + \theta \mu_{0} - \mu_{1} + \theta \pi_{2} - \pi_{1}) , \quad j = 1, \ldots, n_{1} , \]
d.h. die Verteilungen der Differenzen der beiden Sequenzen unterscheiden sich nur durch den Lokationsunterschied von

\[ 2(\mu_{1} - \theta \mu_{0}) + \pi_{1} - \theta \pi_{2} + \pi_{2} - \theta \pi_{1} = 2(\mu_{1} - \theta \mu_{0}) + (\pi_{1} + \pi_{2})(1 - \theta) = 2(\mu_{1} - \theta \mu_{0}) , \]
da \( \pi_{1} + \pi_{2} = 0 \). Somit können Hypothesen über \( \mu_{1} - \theta \mu_{0} \leq 0 \Leftrightarrow \frac{\mu_{1}}{\mu_{0}} \leq \theta \) durch Verfahren für zwei unverbundene Stichproben untersucht werden.

4.2.1 Einseitiges Testproblem

Im Model I mit der bivariaten Normalverteilungsannahme \( (\varepsilon_{g1}, \varepsilon_{g2})^{\prime} \text{ u.i.v. } \sim N_{2}(0, \Sigma) \),

\[ \Sigma_{0} = \begin{pmatrix} \sigma_{0}^{2} & \sigma_{01} \\ \sigma_{01} & \sigma_{1}^{2} \end{pmatrix} \text{ und } \Sigma_{1} = \begin{pmatrix} \sigma_{1}^{2} & \sigma_{01} \\ \sigma_{01} & \sigma_{0}^{2} \end{pmatrix} , \]

benötigt man zur Konstruktion von Entscheidungsverfahren das folgende Lemma.
Lemma 4:

Der Schätzer \( \frac{X_1 - \theta X_0}{2} \) ist normalverteilt mit Erwartungswert \( \mu_1 - \theta \mu_0 \) und Varianz

\[
\sigma_0^2 \cdot \frac{1}{4} \left( \frac{1}{n_0} + \frac{1}{n_1} \right),
\]

wobei

\[
\sigma_0^2 = \theta^2 \sigma_0^2 + \sigma_1^2 - 2 \theta \sigma_{01}.
\]

Beweis:

Die Verteilungsfunktionen der um \( \theta \) korrigierten Periodendifferenzen sind:

\[
Y_{ij}^\theta \text{ u.i.v. } \sim G(y - E(Y_{ij}^\theta)) = \Phi \left( \frac{y - E(Y_{ij}^\theta)}{\text{Var}(Y_{ij}^\theta)} \right),
\]

mit

\[
E(Y_{0i}^\theta) = \theta \mu_0 - \mu_1 + \theta \pi_1 - \pi_2, \quad E(Y_{1i}^\theta) = \mu_1 - \theta \mu_0 + \pi_1 - \theta \pi_2,
\]

\[
\text{Var}(Y_{ij}^\theta) = \sigma_0^2 = \theta^2 \sigma_0^2 + \sigma_1^2 - 2 \theta \sigma_{01}, \quad i = 0, 1 \text{ und } j = 1, \ldots, n_i.
\]

Der Erwartungswert für die Differenz der transformierten Zufallsvariablen aus den beiden Stichproben beträgt

\[
E(Y_{1ij}^\theta - Y_{0ij}^\theta) = 2(\mu_1 - \theta \mu_0) + (\pi_1 + \pi_2)(1 - \theta) = 2(\mu_1 - \theta \mu_0),
\]

und ein erwartungstreuer Schätzer für die Differenz \( \mu_1 - \theta \mu_0 \) ist

\[
\frac{\bar{Y}_{iij}^\theta - \bar{Y}_{0ij}^\theta}{2} = \frac{1}{2} \left( \frac{1}{n_1} \sum_{j=1}^{n_1} (X_{1ijj} - \theta X_{1ijj}) - \frac{1}{n_0} \sum_{j=1}^{n_0} (\theta X_{0ijj} - X_{0ijj}) \right) = \frac{X_1 - \theta X_0}{2}.
\]

Die Varianz des Schätzers ist

\[
\text{Var} \left( \frac{X_1 - \theta X_0}{2} \right) = \text{Var} \left( \frac{\bar{Y}_{iij}^\theta - \bar{Y}_{0ij}^\theta}{2} \right) = \sigma_0^2 \cdot \frac{1}{4} \left( \frac{1}{n_0} + \frac{1}{n_1} \right)
\]

da

\[
\text{Var}(Y_{ij}^\theta) = \sigma_0^2, \quad i = 0, 1 \text{ und } j = 1, \ldots, n_i.
\]
Die Varianz $\sigma^2_0$ kann erwartungstreu durch $S_0^2 = \theta^2 S_1^2 + S_1^2 - 2\theta S_{01}$ geschätzt werden, wobei

$$S_0^2 = \frac{1}{n_0 + n_1 - 2} \left( \sum_{j=1}^{n_0} (X_{0j} - \bar{X}_{0.1})^2 + \sum_{j=1}^{n_0} (X_{1j} - \bar{X}_{1.2})^2 \right),$$

$$S_1^2 = \frac{1}{n_0 + n_1 - 2} \left( \sum_{j=1}^{n_0} (X_{0j} - \bar{X}_{0.2})^2 + \sum_{j=1}^{n_0} (X_{1j} - \bar{X}_{1.1})^2 \right),$$

$$S_{01} = \frac{1}{n_0 + n_1 - 2} \left( \sum_{j=1}^{n_0} (X_{0j} - \bar{X}_{1.1})(X_{0j} - \bar{X}_{0.2}) + \sum_{j=1}^{n_0} (X_{1j} - \bar{X}_{1.1})(X_{1j} - \bar{X}_{1.2}) \right)$$

und $\bar{X}_{ik} = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ijk}$.

Die Verteilung von $\frac{S_0^2}{\sigma^2_0}(n_0 + n_1 - 2)$ ist $\chi^2$-verteilt mit $n_0 + n_1 - 2$ Freiheitsgraden (Locke, 1984). Das auf den Quotienten der Erwartungswerte bezogene einseitige Testproblem $H^0_\theta$ mit einem irrelevanten Bereich von $(0, \infty)$, $\theta > 1$,

$$H^0_\theta: \frac{\mu_1}{\mu_0} \geq \theta \quad \text{versus} \quad H^1_\theta: \frac{\mu_1}{\mu_0} < \theta,$$

wird genau dann abgelehnt, falls

$$t_1^\theta = \frac{\bar{X}_1 - \theta \bar{X}_0}{S_0} \leq -t_{\alpha, n_0 + n_1 - 2}.$$

Die Bedingung $t_1^\theta \leq -t_{\alpha, n_0 + n_1 - 2}$ ist äquivalent zu

$$(-\infty, \theta^*_{-1}) \subset (-\infty, \theta) \quad \text{und} \quad \bar{X}_0^2 > a_0,$$

wobei

$$\theta^*_{-1} = \frac{(\bar{X}_1 \bar{X}_0 - a_{01}) + \sqrt{(\bar{X}_1 \bar{X}_0 - a_{01})^2 - (\bar{X}_1^2 - a_1)(\bar{X}_0^2 - a_0)}}{(\bar{X}_0^2 - a_0)}.$$
\[ a_1 = \frac{1}{4} \left( \frac{1}{n_0} + \frac{1}{n_1} \right) S^2_{\alpha,n_0+n_1-2}, \quad a_0 = \frac{1}{4} \left( \frac{1}{n_0} + \frac{1}{n_1} \right) S^2_{0,n_0+n_1-2}, \quad a_{01} = \frac{1}{4} \left( \frac{1}{n_0} + \frac{1}{n_1} \right) S_{01} S^2_{\alpha,n_0+n_1-2}. \]

Ist die obere Grenze \( \theta^1_+ \) des einseitigen 100(1-\( \alpha \))%-Konfidenzintervales für \( \mu_1 \) unterhalb von \( \theta \), so kann zum Niveau \( \alpha \) auf vorliegende Sicherheit geschlossen werden.

Im Modell II wird \( H_0^\theta \) genau dann abgelehnt, falls

\[ t^0_i = \frac{\bar{X}_i - \theta \bar{X}_0}{\sqrt{S^2_b (1 + \theta^2) + S^2_e (1 - \theta)^2}} \leq -t_{\alpha,n_0+n_1-2}, \]

mit

\[ S^2_b = \frac{1}{2} \left( \frac{1}{2} \sum_{i=0}^{n_0} X^2_i + \sum_{j=0}^{n_1} X^2_j - S^2_e \right). \]

Kieser und Hauschke (1999b) haben gezeigt, daß die Bedingung \( t^0_i \leq -t_{\alpha,n_0+n_1-2} \) äquivalent ist zu

\[ (-\infty, \theta^0_+) \subset (-\infty, \theta) \quad \text{und} \quad \bar{X}_0^2 > a_{e+b}, \]

wobei

\[ \theta^0_+ = \frac{(\bar{X}_1 \bar{X}_0 - a_b) + \sqrt{(\bar{X}_1 \bar{X}_0 - a_b)^2 - (\bar{X}_1^2 - a_{e+b})(\bar{X}_0^2 - a_{e+b})}}{(\bar{X}_0^2 - a_{e+b})}, \]

\[ a_b = \frac{1}{4} \left( \frac{1}{n_0} + \frac{1}{n_1} \right) S^2_{\alpha,n_0+n_1-2}, \quad a_{e+b} = \frac{1}{4} \left( \frac{1}{n_0} + \frac{1}{n_1} \right) (S^2_e + S^2_b) t^2_{\alpha,n_0+n_1-2}. \]

Man beachte, daß der Schätzer \( S^2_b \) aufgrund der folgenden Beziehung nach der Satterthwaites Methode berechnet wurde:

\[ E \left( \frac{1}{n_0 + n_1 - 2} \left( \sum_{i=0}^{n_0} \sum_{j=1}^{n_1} \frac{X^2_{ij}}{2} - \sum_{j=0}^{n_1} X^2_j \right) \right) = 2\sigma^2_b + \sigma^2_e. \]
Es sei daher darauf hingewiesen, daß im Vorliegen der Varianzhomogenität (Modell II) die Teststatistik $t_1^0$ nur approximativ $t$-verteilt ist (Vuorinen und Tuominen, 1994), wobei für die nach der Satterthwaites Approximation berechneten Freiheitsgrade $\nu$ gilt

$$n_0 + n_1 - 2 \leq \nu \leq 2(n_0 + n_1 - 2).$$


Verzichtet man auf die Normalverteilungsannahme, so kann wegen des vorliegenden Lokationsmodells

$$Y^0_{0j} \text{ u.i.v. } \sim G(y + \mu_j - \theta \mu_0 + \pi_2 - \theta \pi_1), \ j = 1, ..., n_0,$$

$$Y^0_{1j} \text{ u.i.v. } \sim G(y + \theta \mu_0 - \mu_j + \theta \pi_2 - \pi_1), \ j = 1, ..., n_1,$$

sowohl für Model I als auch für Modell II der Wilcoxon-Test (Vuorinen und Turunen, 1997) und das verteilungsfreie Konfidenzintervall nach Moses als Entscheidungsverfahren herangezogen werden. Unter $H^0_0$ gilt

$$Y^0_{0j} \text{ u.i.v. } \sim G(y + \pi_2 - \theta \pi_1) = G(z), \ j = 1, ..., n_0$$

$$Y^0_{1j} \text{ u.i.v. } \sim G(y + \theta \pi_2 - \pi_1) = G(z), \ j = 1, ..., n_1,$$

da aus $\pi_1 + \pi_2 = 0$ folgt, daß $\pi_1 - \theta \pi_2 = \theta \pi_1 - \pi_2$. Dabei muß jedoch nun beachtet werden, daß die Rangsumme der um $\theta$ korrigierten $j$-ten intraindividuellen Periodendifferenzen $Y^0_{ij}$ verwendet wird. Sei $R_{ij}(\theta)$ der Rang von $Y^0_{ij} = X_{ij1} - \theta X_{ij2}$ in der kombinierten Stichprobe

$$Y^0_{01}, ..., Y^0_{0n_0}, Y^0_{11}, ..., Y^0_{1n_1} =$$

$$\theta X_{011} - X_{012}, ..., \theta X_{0n_01} - X_{0n_02}, X_{111} - \theta X_{112}, ..., X_{1n_11} - \theta X_{1n_12}$$

und $R_i(\theta) = \sum_{j=1}^{n_i} R_{ij}(\theta)$. Die Nullhypothese $H^0_0$ wird verworfen, falls

$$R_i(\theta) \leq n_i(n_0 + n_1 + 1) - t^\theta_{\alpha, n_0, n_1},$$
wobei \( r^0_{\alpha,n_0,n_1} \) das \((1-\alpha)\)-Quantil der Wilcoxonstatistik \( R_i(\theta) \) ist. Die Konstruktion des Konfidenzintervallges basiert wiederum auf den geordneten Werten \( D^0_1 \leq \ldots \leq D^0_{n_0n_1} \) der insgesamt \( n_0n_1 \) Differenzen \( Y^\theta_{ij} - Y^\theta_{0j} \), \( j = 1, \ldots, n_1 \) und \( j^* = 1, \ldots, n_0 \). Damit ergibt sich als einseitiges \( 100(1-\alpha)\% \)-Konfidenzintervall für \( \mu_1 - \theta \mu_0 \):

\[
\left( -\infty, \frac{U^\theta_1}{2} \right], \text{ wobei } U^\theta_1 = D^0_{n_1} = D^0_{n_0n_1+1-C^1_\alpha}, \quad C^1_\alpha = \frac{n_1(2n_0 + n_1 + 1)}{2} + 1 - r^\theta_{\alpha,n_0,n_1}.
\]

Liegt die obere Grenze des einseitigen Konfidenzintervallges unterhalb von 0, so kann die Nullhypothese verworfen und auf Sicherheit geschlossen werden.

### 4.2.2 Zweiseitiges Testproblem

Das Testen der zweiseitigen Fragestellung

\[
H^\theta_0: \frac{\mu_1}{\mu_0} \leq \theta \quad \text{oder} \quad \frac{\mu_1}{\mu_0} \geq \theta_2
\]

versus

\[
H^\theta_1: \theta_1 < \frac{\mu_1}{\mu_0} < \theta_2,
\]

mit \( \theta_1 = 1 + f_1, \theta_2 = 1 + f_2 \) und \(-1 < f_1 < 0 < f_2 \) ist äquivalent zur simultanen Überprüfung der beiden einseitigen Testprobleme:

\[
H^\theta_{01}: \frac{\mu_1}{\mu_0} \leq \theta_1 \quad \text{versus} \quad H^\theta_{11}: \frac{\mu_1}{\mu_0} > \theta_1
\]

und

\[
H^\theta_{02}: \frac{\mu_1}{\mu_0} \geq \theta_2 \quad \text{versus} \quad H^\theta_{12}: \frac{\mu_1}{\mu_0} < \theta_2.
\]

\( H^\theta_0 = \bigcup_{r=1}^2 H^\theta_{0r} \) wird bei der bivariaten Normalverteilungsannahme (Modell I) verworfen, falls
\[ t_{11}^0 = \frac{\bar{X}_1 - \theta_1 \bar{X}_0}{\frac{2}{S_{\theta_1} \left( \frac{1}{n_0} + \frac{1}{n_1} \right)}} \geq t_{\alpha, n_0+n_1-2} \quad \text{und} \quad t_{12}^0 = \frac{\bar{X}_1 - \theta_2 \bar{X}_0}{\frac{2}{S_{\theta_2} \left( \frac{1}{n_0} + \frac{1}{n_1} \right)}} \leq -t_{\alpha, n_0+n_1-2}. \]

Die Verwerfung beider einseitigen Hypothesen ist äquivalent zu

\[
[\theta_1^+, \theta_2^+] \subset (\theta_1, \theta_2) \quad \text{und} \quad \bar{X}_0^2 > a_0, \quad \text{wobei}
\]

\[
\theta_2^+ = \frac{\bar{X}_1 \bar{X}_0 - a_{\theta_1}}{\sqrt{\bar{X}_1^2 (1 + \theta_1^2) + \bar{X}_0^2 (1 - \theta_1^2)}} \leq -t_{\alpha, n_0+n_1-2}.
\]

Der Schluß auf Sicherheit zum Niveau \( \alpha \) ist somit möglich, falls das 100(1-2\( \alpha \))%-Konfidenzintervall für \( \frac{\mu_1}{\mu_0} \) regulär und vollständig im Sicherheitsbereich enthalten \((\theta_1, \theta_2)\) ist (Hauschke et al., 1999c).

Im Modell II (Kieser und Hauschke, 1999b) wird \( H_0^0 = \bigcup_{r=1}^2 H_{0r}^0 \) genau dann abgelehnt, falls

\[
t_{11}^0 = \frac{\bar{X}_1 - \theta_1 \bar{X}_0}{\sqrt{S_{\bar{X}_1}^2 (1 + \theta_1^2) + S_{\bar{X}_0}^2 (1 - \theta_1^2)}} \leq -t_{\alpha, n_0+n_1-2}
\]

\[
\text{und}
\]

\[
t_{12}^0 = \frac{\bar{X}_1 - \theta_2 \bar{X}_0}{\sqrt{S_{\bar{X}_1}^2 (1 + \theta_2^2) + S_{\bar{X}_0}^2 (1 - \theta_2^2)}} \geq t_{\alpha, n_0+n_1-2}.
\]

Die Verwerfung beider Hypothesen ist äquivalent zu

\[
[\theta_1^+, \theta_2^+] \subset (\theta_1, \theta_2) \quad \text{und} \quad \bar{X}_0^2 > a_{\theta_1 \theta_2}, \quad \text{wobei}
\]

\[
\theta_2^+ = \frac{\bar{X}_1 \bar{X}_0 - a_{\theta_1}}{\sqrt{\bar{X}_1^2 (1 + \theta_1^2) + \bar{X}_0^2 (1 - \theta_1^2)}} \leq -t_{\alpha, n_0+n_1-2}.
\]
Für das entsprechende verteilungsfreie Verfahren mit Zweistichproben-Wilcoxon-Tests gilt

\[ H_{01}^\theta \text{ wird abgelehnt, falls } R_1(\theta_1) \geq r_{\alpha,n_0,n_1}^\theta \]

und \( H_{02}^\theta \text{ wird abgelehnt, falls } R_2(\theta_2) \leq n_1(n_0 + n_1 + 1) - r_{\alpha,n_0,n_1}^\theta \).

Die Verwerfung beider Einzelhypotesen ist äquivalent zur Bedingung, daß die untere Grenze des einseitigen 100(1-\(\alpha\))%-Konfidenzintervales für \( \mu_1 - \theta_1 \mu_0 \) oberhalb von 0 liegt, und ferner, daß die obere Grenze des einseitigen 100(1-\(\alpha\))%-Konfidenzintervales für \( \mu_1 - \theta_2 \mu_0 \) unterhalb von 0 liegt:

\[
\left[ \frac{\bar{L}_1^\theta}{2}, \infty \right] \subset (0, \infty) \quad \text{und} \quad \left( -\infty, \frac{\bar{U}_1^\theta}{2} \right] \subset (-\infty, 0),
\]

wobei \( \bar{L}_1^\theta = D_{n_0}^\theta = D_{c_0}^\theta \) und \( \bar{U}_1^\theta = D_{n_1}^\theta = D_{c_1}^\theta \).

### 4.2.3 Gütebestimmung und Fallzahlplanung

Die Güte des einseitigen Sasabuchi-Tests \( t_1^\theta \) im Crossover-Design unter Annahme des Modells II läßt sich über die nichtzentrale \( t \)-Verteilung ermitteln:

\[
P \left[ t_1^\theta \leq -t_{\alpha,2n-2} \left| \frac{\mu_1}{\mu_0} < \theta, \sigma_\theta \right. \right] = \frac{\sqrt{2\pi}}{\Gamma(n-1)2^{n-2}} \int_0^\frac{\pi}{2} \Phi \left( \frac{-t_{\alpha,2n-2} x}{\sqrt{2n-2}} - \Theta \right) x^{2n-3} \Phi(x) \, dx,
\]

mit dem Nichtzentralitätsparameter:

\[
\Theta = \frac{\mu_1}{\sigma_\theta} - \theta \mu_0 = \frac{\mu_1 - \theta \mu_0}{\sigma_\theta} = \frac{\mu_1 - \theta}{\mu_0} = \sqrt{2n} \frac{CV_\epsilon (1+\theta^2) + CV_b (1-\theta)^2}{2n},
\]

mit \( \sigma_\epsilon = CV_\epsilon \mu_0 \) und \( \sigma_b = CV_b \mu_0 \).

In Abbildung 12 wird die Gütefunktion für verschiedene Variationskoeffizienten \( CV_\epsilon \) bei einem festen \( CV_b = 30\% \), einem Sicherheitsbereich von \( (-\infty, \theta) = (-\infty, 1.2) \) und jeweils 10 Versuchseinheiten pro Sequenz dargestellt.
Abbildung 12: Crossover-Situation: Die Gütefunktion des einseitigen Tests $t_1^0$ bei einem Sicherheitsbereich von $(-\infty, \theta) = (-\infty, 1.2)$ für $\alpha = 0.05$, $n_0 = n_1 = 10$, $CV_b = 0.30$ und verschiedene $CV_c$

Den notwendigen Stichprobenumfang $n$ (pro Sequenz), um beim Testen von $H_0^\theta$ versus $H_1^\theta$ die Fehler $\alpha$ und $\beta$ an der Stelle $\frac{\mu_1}{\mu_0} < \theta$, nicht zu überschreiten, kann man iterativ nach der folgenden Ungleichung berechnen:

$$n \geq \frac{1}{2} \frac{CV_c^2(1+\theta^2) + CV_b^2(1-\theta)^2}{(\theta - \frac{\mu_1}{\mu_0})^2} \left(t_{\alpha,2n-2} + t_{\beta,2n-2}\right)^2.$$

Die Herleitung erfolgt analog zur Zweistichprobensituation.

In den Tabellen 16 und 17 sind die exakten und approximativen Fallzahlen angegeben, um eine Güte von mindestens 0.80 zu erreichen. Dabei sind die weiteren Einflussgrößen wie folgt vorgegeben: $\alpha = 0.05$, $\theta = 1.20$, $\frac{\mu_1}{\mu_0} = 1.00, 1.05, 1.10, 1.15$ und $CV_0 = 5.0, 7.5, 10.0, \ldots, 30.0\%$. 
Tabelle 16: Crossover-Situation: Exakter Stichprobenumfang pro Sequenz für eine Güte von mindestens 0.80 bei $\alpha = 0.05$, $(\infty,0) = (\infty,12)$ und verschiedene $CV_e$, $CV_b$.

<table>
<thead>
<tr>
<th>$CV_e$ (%)</th>
<th>$CV_b$ (%)</th>
<th>1.00</th>
<th>1.05</th>
<th>1.10</th>
<th>1.15</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.0</td>
<td>20.0</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>40.0</td>
<td></td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>17</td>
</tr>
<tr>
<td>60.0</td>
<td></td>
<td>3</td>
<td>4</td>
<td>8</td>
<td>27</td>
</tr>
<tr>
<td>80.0</td>
<td></td>
<td>4</td>
<td>6</td>
<td>11</td>
<td>40</td>
</tr>
<tr>
<td>100.0</td>
<td></td>
<td>5</td>
<td>8</td>
<td>15</td>
<td>58</td>
</tr>
<tr>
<td>10.0</td>
<td>20.0</td>
<td>4</td>
<td>5</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>40.0</td>
<td></td>
<td>4</td>
<td>6</td>
<td>11</td>
<td>39</td>
</tr>
<tr>
<td>60.0</td>
<td></td>
<td>4</td>
<td>7</td>
<td>13</td>
<td>49</td>
</tr>
<tr>
<td>80.0</td>
<td></td>
<td>5</td>
<td>8</td>
<td>17</td>
<td>63</td>
</tr>
<tr>
<td>100.0</td>
<td></td>
<td>6</td>
<td>10</td>
<td>21</td>
<td>81</td>
</tr>
<tr>
<td>15.0</td>
<td>20.0</td>
<td>6</td>
<td>9</td>
<td>19</td>
<td>71</td>
</tr>
<tr>
<td>40.0</td>
<td></td>
<td>6</td>
<td>10</td>
<td>20</td>
<td>77</td>
</tr>
<tr>
<td>60.0</td>
<td></td>
<td>7</td>
<td>11</td>
<td>23</td>
<td>87</td>
</tr>
<tr>
<td>80.0</td>
<td></td>
<td>8</td>
<td>12</td>
<td>26</td>
<td>101</td>
</tr>
<tr>
<td>100.0</td>
<td></td>
<td>9</td>
<td>14</td>
<td>31</td>
<td>119</td>
</tr>
<tr>
<td>20.0</td>
<td>20.0</td>
<td>9</td>
<td>15</td>
<td>32</td>
<td>124</td>
</tr>
<tr>
<td>40.0</td>
<td></td>
<td>9</td>
<td>16</td>
<td>33</td>
<td>130</td>
</tr>
<tr>
<td>60.0</td>
<td></td>
<td>10</td>
<td>17</td>
<td>36</td>
<td>140</td>
</tr>
<tr>
<td>80.0</td>
<td></td>
<td>11</td>
<td>18</td>
<td>39</td>
<td>154</td>
</tr>
<tr>
<td>100.0</td>
<td></td>
<td>12</td>
<td>20</td>
<td>44</td>
<td>171</td>
</tr>
<tr>
<td>25.0</td>
<td>20.0</td>
<td>13</td>
<td>22</td>
<td>49</td>
<td>192</td>
</tr>
<tr>
<td>40.0</td>
<td></td>
<td>14</td>
<td>23</td>
<td>50</td>
<td>198</td>
</tr>
<tr>
<td>60.0</td>
<td></td>
<td>14</td>
<td>24</td>
<td>53</td>
<td>208</td>
</tr>
<tr>
<td>80.0</td>
<td></td>
<td>15</td>
<td>26</td>
<td>56</td>
<td>221</td>
</tr>
<tr>
<td>100.0</td>
<td></td>
<td>16</td>
<td>28</td>
<td>61</td>
<td>239</td>
</tr>
<tr>
<td>30.0</td>
<td>20.0</td>
<td>18</td>
<td>32</td>
<td>70</td>
<td>275</td>
</tr>
<tr>
<td>40.0</td>
<td></td>
<td>19</td>
<td>32</td>
<td>71</td>
<td>281</td>
</tr>
<tr>
<td>60.0</td>
<td></td>
<td>19</td>
<td>33</td>
<td>74</td>
<td>291</td>
</tr>
<tr>
<td>80.0</td>
<td></td>
<td>20</td>
<td>35</td>
<td>77</td>
<td>304</td>
</tr>
<tr>
<td>100.0</td>
<td></td>
<td>21</td>
<td>37</td>
<td>81</td>
<td>322</td>
</tr>
</tbody>
</table>

Für das zweiseitige Testproblem erfolgt die Gütebestimmung nach

\[
P\left( t_{11}^0 \geq t_{\alpha,2n-2} \land t_{12}^0 \leq t_{\alpha,2n-2} \mid \theta_1 < \frac{\mu_1}{\mu_0} < \theta_2, \sigma_{\theta_1}, \sigma_{\theta_2} \right) =
\]

\[
Q(\infty,-t_{\alpha,2n-2},\Theta_1,\Theta_2,\rho) - Q(t_{\alpha,2n-2},-t_{\alpha,2n-2},\Theta_1,\Theta_2,\rho)
\]

wobei

\[
Q(t_1, t_2, \Theta_1, \Theta_2, \rho) = P\left( t_{11}^0 \leq t_1 \land t_{12}^0 \leq t_2 \mid \theta_1 < \frac{\mu_1}{\mu_0} < \theta_2, \sigma \right) =
\]
\[
\frac{\sqrt{2\pi}}{\Gamma(n-1)2^{n-2}} \int_0^\infty \Phi_2 \left( \frac{t_1x}{\sqrt{2n-2}} - \Theta_1, \frac{t_2x}{\sqrt{2n-2}} - \Theta_2, \rho \right) x^{2n-3} \, \Phi(x) \, dx
\]

und

\[
\Theta_1 = \frac{\mu_1 - \theta_1 \mu_0}{\sigma_0 \sqrt{\frac{1}{2n}}} \quad \text{und} \quad \Theta_2 = \frac{\mu_1 - \theta_2 \mu_0}{\sigma_0 \sqrt{\frac{1}{2n}}}
\]

**Tabelle 17:** Crossover-Situation: Approximater Stichprobenumfang pro Sequenz für eine Güte von mindestens 0.80 bei \( \alpha = 0.05 \), \((\rightarrow, 0) = (\infty, 1.2)\) und verschiedene \( CV_a \), \( CV_b \).
Im Modell II gilt

\[ \Theta_1 = \frac{\mu_1 - \theta_1}{\sqrt{\frac{CV_e^2(1 + \theta_1^2) + CV_b^2(1 - \theta_1)^2}{2n}}} \quad \text{und} \quad \Theta_2 = \frac{\mu_1 - \theta_2}{\sqrt{\frac{CV_e^2(1 + \theta_2^2) + CV_b^2(1 - \theta_2)^2}{2n}}}. \]

Der benötigte Korrelationskoeffizient \( \rho^0 \) folgt mit Lemma 5.

**Lemma 5:**

Es gilt

\[ \rho^0 = \frac{CV_e^2(1 + \theta_1, \theta_2) + CV_b^2(1 + \theta_1, \theta_2 - \theta_1, -\theta_2)}{\sqrt{(CV_e^2(1 + \theta_1^2) + CV_b^2(1 - \theta_1)^2)(CV_e^2(1 + \theta_2^2) + CV_b^2(1 - \theta_2)^2)}}. \]

**Beweis:**

\[ Var(\bar{X}_1 - \theta_1, \bar{X}_0) = \frac{1}{2n} \left( \sigma_e^2(1 + \theta_1^2) + \sigma_b^2(1 - \theta_1)^2 \right) \]
\[ = \frac{\mu_0^2}{2n} \left( CV_e^2(1 + \theta_1^2) + CV_b^2(1 - \theta_1)^2 \right), \quad r = 1, 2. \]

\[ Kov(\bar{X}_1 - \theta_1, \bar{X}_0, \bar{X}_1 - \theta_2, \bar{X}_0) = Var(\bar{X}_1) + Var(\bar{X}_0)\theta_1\theta_2 - Kov(\bar{X}_1, \bar{X}_0)(\theta_1 + \theta_2) = \]
\[ \frac{1}{2n} \left( \sigma_e^2 + \sigma_b^2 + (\sigma_e^2 + \sigma_b^2)\theta_1\theta_2 - \sigma_b^2(\theta_1 + \theta_2) \right) = \frac{1}{2n} \left( \sigma_e^2(1 + \theta_1\theta_2) + \sigma_b^2(1 + \theta_1\theta_2 - \theta_1 - \theta_2) \right) = \]
\[ \frac{\mu_0^2}{2n} \left( CV_e^2(1 + \theta_1\theta_2) + CV_b^2(1 + \theta_1\theta_2 - \theta_1 - \theta_2) \right) \]
und daher
\[ \rho^0 = \frac{CV_e^2(1 + \theta_1, \theta_2) + CV_b^2(1 + \theta_1, \theta_2 - \theta_1, -\theta_2)}{\sqrt{(CV_e^2(1 + \theta_1^2) + CV_b^2(1 - \theta_1)^2)(CV_e^2(1 + \theta_2^2) + CV_b^2(1 - \theta_2)^2)}}. \]

In Abbildung 13 werden die Gütefunktionen für einen Sicherheitsbereich von \( (\theta_1, \theta_2) = (0.8, 1.25) \), für die Variationskoeffizienten \( CV_e = 32.5\% \), \( CV_b = 30.0\% \) und für verschiedene Fallzahlen dargestellt. Anhand der Grafik erkennt man, daß bei kleinen Stichproben das aktuelle Niveau an der Stelle 0.8 kleiner ist als das nominelle Niveau. Bei größer werdenden Stichprobenanzahlen konvergiert es jedoch gegen das nominelle Niveau (Schuirmann, 1989).

Abbildung 13: Crossover-Situation: Die Gütefunktion des Entscheidungsverfahrens, basierend auf den zwei einseitigen Tests $t_{11}^θ, t_{12}^θ$ bei einem Sicherheitsbereich von $(\theta_1, \theta_2) = (0.8, 1.25)$ für $\alpha = 0.05$, $n_0 = n_1 = 10, 20, 30$, $CV_a = 32.5$ und $CV_b = 30.0\%$

Entsprechende Fallzahlen, um eine Güte von mindestens 0.80 zu erreichen, werden in Tabelle 18 angegeben. Die vereinfachte Berechnung der Fallzahlen erfolgt nach den folgenden Approximationsformeln. Dabei wurde folgender Sicherheitsbereich gewählt, d.h. $\theta_1 = \frac{1}{\theta_2}$:

falls $\frac{\mu_1}{\mu_0} = 1$

\[
 n \geq \frac{1}{2} \left( \frac{CV_0^2 (1+\theta_2^2) + CV_a^2 (1-\theta_2^2)}{(\theta_2 - 1)^2} \left( t_{\alpha,2n-2} + t_{\beta,2n-2} \right)^2 \right),
\]

1 < $\frac{\mu_1}{\mu_0} < \theta_2$

\[
 n \geq \frac{1}{2} \left( \frac{CV_0^2 (1+\theta_2^2) + CV_a^2 (1-\theta_2^2)}{(\theta_2 - \mu_1/\mu_0)^2} \left( t_{\alpha,2n-2} + t_{\beta,2n-2} \right)^2 \right).
\]
\[
\frac{1}{\theta_2} < \frac{\mu_1}{\mu_0} < 1 \quad \Rightarrow \quad n \geq \frac{1}{2} \left( \frac{1}{1 - \frac{1}{\theta_2}} \right)^2 \left( 1 + \frac{1}{\theta_2^2} \right) + CV_b^2 \left( 1 - \frac{1}{\theta_2} \right)^2 \left( t_{0,2n-2} + t_{1,2n-2}^2 \right) \]

Die Herleitung dieser Formeln erfolgt analog zur Zweistichprobensituation.

Die der Tabelle 18 entsprechenden approximativen Fallzahlen sind in Tabelle 19 aufgelistet.

**Tabelle 18: Crossover-Situation: Exakter Stichprobenumfang pro Sequenz für eine Güte von mindestens 0.80 bei \( \alpha = 0.05 \) \((\theta_1, \theta_2) = (0.8, 1.25)\), und verschiedene \( CV_e \), \( CV_b \).**

<table>
<thead>
<tr>
<th>( CV_e ) (%)</th>
<th>( CV_b ) (%)</th>
<th>0.85</th>
<th>0.90</th>
<th>0.95</th>
<th>( \frac{\mu_1}{\mu_0} )</th>
<th>1.00</th>
<th>1.05</th>
<th>1.10</th>
<th>1.15</th>
<th>1.20</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.0</td>
<td>20.0</td>
<td>8</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>40.0</td>
<td>24</td>
<td>7</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>10</td>
<td>37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60.0</td>
<td>38</td>
<td>10</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>8</td>
<td>16</td>
<td>59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>80.0</td>
<td>56</td>
<td>15</td>
<td>7</td>
<td>5</td>
<td>7</td>
<td>11</td>
<td>23</td>
<td>86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100.0</td>
<td>23</td>
<td>7</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>10</td>
<td>36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40.0</td>
<td>29</td>
<td>8</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>12</td>
<td>45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60.0</td>
<td>39</td>
<td>11</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>8</td>
<td>16</td>
<td>61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>80.0</td>
<td>53</td>
<td>14</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>10</td>
<td>21</td>
<td>82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100.0</td>
<td>71</td>
<td>19</td>
<td>9</td>
<td>7</td>
<td>8</td>
<td>13</td>
<td>28</td>
<td>110</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15.0</td>
<td>49</td>
<td>13</td>
<td>7</td>
<td>5</td>
<td>6</td>
<td>10</td>
<td>20</td>
<td>76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40.0</td>
<td>55</td>
<td>15</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>11</td>
<td>22</td>
<td>85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60.0</td>
<td>65</td>
<td>17</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>12</td>
<td>26</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>80.0</td>
<td>78</td>
<td>21</td>
<td>10</td>
<td>8</td>
<td>9</td>
<td>15</td>
<td>31</td>
<td>122</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100.0</td>
<td>96</td>
<td>25</td>
<td>12</td>
<td>9</td>
<td>11</td>
<td>18</td>
<td>38</td>
<td>150</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20.0</td>
<td>84</td>
<td>22</td>
<td>11</td>
<td>8</td>
<td>10</td>
<td>16</td>
<td>34</td>
<td>131</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40.0</td>
<td>90</td>
<td>23</td>
<td>11</td>
<td>9</td>
<td>10</td>
<td>17</td>
<td>36</td>
<td>140</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60.0</td>
<td>100</td>
<td>26</td>
<td>13</td>
<td>10</td>
<td>11</td>
<td>18</td>
<td>40</td>
<td>156</td>
<td></td>
<td></td>
</tr>
<tr>
<td>80.0</td>
<td>114</td>
<td>29</td>
<td>14</td>
<td>11</td>
<td>13</td>
<td>21</td>
<td>45</td>
<td>177</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100.0</td>
<td>132</td>
<td>34</td>
<td>16</td>
<td>12</td>
<td>14</td>
<td>24</td>
<td>52</td>
<td>205</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25.0</td>
<td>130</td>
<td>33</td>
<td>16</td>
<td>12</td>
<td>14</td>
<td>24</td>
<td>51</td>
<td>202</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40.0</td>
<td>136</td>
<td>35</td>
<td>17</td>
<td>13</td>
<td>15</td>
<td>25</td>
<td>54</td>
<td>212</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60.0</td>
<td>146</td>
<td>37</td>
<td>18</td>
<td>14</td>
<td>16</td>
<td>26</td>
<td>58</td>
<td>227</td>
<td></td>
<td></td>
</tr>
<tr>
<td>80.0</td>
<td>160</td>
<td>41</td>
<td>19</td>
<td>15</td>
<td>17</td>
<td>29</td>
<td>63</td>
<td>249</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100.0</td>
<td>177</td>
<td>45</td>
<td>21</td>
<td>16</td>
<td>19</td>
<td>32</td>
<td>70</td>
<td>276</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30.0</td>
<td>186</td>
<td>47</td>
<td>22</td>
<td>17</td>
<td>20</td>
<td>33</td>
<td>73</td>
<td>289</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40.0</td>
<td>192</td>
<td>49</td>
<td>23</td>
<td>18</td>
<td>20</td>
<td>34</td>
<td>76</td>
<td>299</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60.0</td>
<td>201</td>
<td>51</td>
<td>24</td>
<td>19</td>
<td>21</td>
<td>36</td>
<td>79</td>
<td>314</td>
<td></td>
<td></td>
</tr>
<tr>
<td>80.0</td>
<td>215</td>
<td>55</td>
<td>26</td>
<td>20</td>
<td>23</td>
<td>38</td>
<td>85</td>
<td>336</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100.0</td>
<td>233</td>
<td>59</td>
<td>28</td>
<td>21</td>
<td>25</td>
<td>41</td>
<td>92</td>
<td>364</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Mittels der Periodendifferenzen bzw. der um $\theta$ korrigierten Periodendifferenzen kann das Crossover-Design auf die Zweistichprobensituation reduziert werden. Daher konnten die im dritten Kapitel hergeleiteten biometrischen Methoden auf die Crossover-Situation übertragen werden. Bemerkenswert ist wiederum die gute Übereinstimmung der mittels Approximationsformeln ermittelten Fallzahlen mit den exakt ermittelten Werten.

Im Kapitel 6 wird die entsprechende Auswertung des Crossover-Designs an einem realen Datensatz veranschaulicht.
5 K-Stichproben many-to-one-Situation

5.1 Simple-tree Alternative

Ein für Sicherheitsuntersuchungen typisches Design ist die folgende Versuchsanordnung:

\[ \{ \text{Kontrolle, Behandlung}_1, \text{Behandlung}_2, \ldots, \text{Behandlung}_k \}, \]


Es bezeichnen \( X_{ij} \) die Zielgrößen, die unabhängig identisch verteilt sind nach den stetigen Verteilungsfunktionen \( F_i(x) = F(x - \mu_i) \), \( i = 0, 1, \ldots, k \) und \( j = 1, \ldots, n_i \), wobei \( \mu_i \) die Lageparameter bezeichnen, der Index \( i \geq 1 \) die Behandlungsgruppen und der Index \( i = 0 \) die Kontrollgruppe. Die zur Sicherheitsuntersuchung relevanten Parameter sind entweder die insgesamt \( k \) Differenzen \( \mu_i - \mu_0 \) oder die Quotienten \( \frac{\mu_i}{\mu_0} \), \( i = 1, \ldots, k \), der Lokationsparameter, je nachdem, ob der Schwellenwert absolut oder nur relativ zur Kontrolle angegeben werden kann.

5.1.1 Globaler Sicherheitsnachweis

5.1.1.1 Einseitiges Testproblem

Die einseitig formulierten Einzelhypothesen für die Differenzen und die Quotienten lauten:

\[ H_{0i}^\delta : \mu_i - \mu_0 \geq \delta \quad \text{versus} \quad H_{1i}^\delta : \mu_i - \mu_0 < \delta \]

bzw.

\[ H_{00}^\theta : \frac{\mu_i}{\mu_0} \geq \theta \quad \text{versus} \quad H_{10}^\theta : \frac{\mu_i}{\mu_0} < \theta . \]

In der Situation des globalen Sicherheitsnachweises muß die Unbedenklichkeit aller \( k \) Behandlungen nachgewiesen werden. Die einseitige Formulierung des Testproblems ist:
Unter der Normalverteilungsannahme $F_i(x) = \Phi\left(\frac{x - \mu_i}{\sigma}\right)$, $i = 0, \ldots, k$, wird $H_0^\delta$ unter Ausnutzung des intersection-union-Prinzips genau dann zum Niveau $\alpha$ abgelehnt, falls alle Einzelhypothesen $H^\delta_{0i}$ jeweils zum Niveau $\alpha$ abgelehnt werden können:

$$t_i^\delta = \frac{\bar{X}_i - \bar{X}_0 - \delta}{S \sqrt{\frac{1}{n_0} + \frac{1}{n_i}}} \leq -t_{\alpha,\nu} \quad \forall \ i \in (1, \ldots, k),$$

mit

$$\bar{X}_0 = \frac{1}{n_0} \sum_{j=1}^{n_0} X_{0j}, \quad \bar{X}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}, \quad S^2 = \frac{1}{\nu} \sum_{i=0}^{k} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2 \quad \text{und} \quad \nu = \sum_{i=0}^{k} n_i - (k + 1).$$

Äquivalent dazu ist die Inklusion aller einseitigen $100(1-\alpha)$-%-Konfidenzintervalle für die Differenzen $\mu_i - \mu_0$, $i = 1, \ldots, k$, im Sicherheitsbereich:

$$(-\infty, \bar{X}_i - \bar{X}_0 + t_{\alpha,\nu} S \sqrt{\frac{1}{n_0} + \frac{1}{n_i}}] \subset (-\infty, \delta) \quad \forall \ i \in (1, \ldots, k).$$

Das nichtparametrische Verfahren basiert auf der paarweise vorgenommenen Rangvergabe (separate ranking), d.h. $R_{y_i}(\delta)$ bezeichnet den Rang von $X_{y_i} - \delta$ in der kombinierten Stichprobe vom Umfang $n_0 + n_i$:

$$X_{01}, \ldots, X_{0n_0}, X_{i1} - \delta, \ldots, X_{in_i} - \delta.$$

$H_0^\delta$ wird abgelehnt, falls $R_i(\delta) = \sum_{j=1}^{n_i} R_{y_j}(\delta) \leq n_i(n_0 + n_i + 1) - r_{\alpha,n_0,n_i}^\delta$ für alle $i \in (1, \ldots, k)$, dabei bezeichnet $r_{\alpha,n_0,n_i}^\delta$ das $(1-\alpha)$-Quantil der Wilcoxonstatistik $R_i(\delta)$. Die einseitigen $100(1-\alpha)$-%-Konfidenzintervalle für $\mu_i - \mu_0$ ergeben sich analog zur Zweistichproben-situation:
\((-\infty, \lambda_i^\delta]\), wobei \(U_i^\delta = D_{\alpha i}^\delta = D_{\alpha, n_i + 1 - C_{\alpha i}}^\delta\), \(C_{\alpha i}^i = \frac{n_i(2n_0 + n_i + 1)}{2} + 1 - t_{\alpha, n_i, n_i}\)

und globale Sicherheit liegt zum Niveau \(\alpha\) vor, falls \((-\infty, \lambda_i^\delta]\subset(-\infty, \delta)\ \forall \ i \in (1, \ldots, k)\).

Die auf den Quotienten basierende Globalhypothese \(H_0^\delta\) wird verworfen, falls

\[
t_i^\delta = \frac{\bar{X}_i - \theta_0}{S\sqrt{\frac{\theta_0^2 + 1}{n_0 + n_i}}} \leq -t_{\alpha, v} \quad \forall \ i \in (1, \ldots, k),
\]

was wiederum äquivalent ist zu

\[
(-\infty, \theta_i^\delta]\subset(-\infty, \theta) \quad \forall \ i \in (1, \ldots, k) \text{ und } \bar{X}_0^\delta > a_0,
\]

mit

\[
\theta_i^\delta = \frac{\bar{X}_0 \bar{X}_i + a_0 \bar{X}_0^2 + a_i \bar{X}_i^2 - a_0 a_i}{\bar{X}_0^2 - a_0}, \quad a_0 = \frac{S^2}{n_0} t_{\alpha, v}^2 \quad \text{und} \quad a_i = \frac{S^2}{n_i} t_{\alpha, v}^2.
\]

Liegen alle oberen Grenzen \(\theta_i^\delta\) der einseitigen 100(1-\(\alpha\))-Konfidenzintervalle für die Quotienten \(\frac{\mu_i}{\mu_0}\), \(i = 1, \ldots, k\) unterhalb von \(\theta\), kann zum Niveau \(\alpha\) auf globale Sicherheit geschlossen werden.

5.1.1.2 Zweiseitiges Testproblem

Die \(k\) zweiseitig formulierten Einzelhypothesen

\[
H_{0i}^\delta: \ \mu_i - \mu_0 \leq \delta_1 \text{ oder } \mu_i - \mu_0 \geq \delta_2 \quad \text{versus} \quad H_{1i}^\delta: \ \delta_1 < \mu_i - \mu_0 < \delta_2
\]

bzw.

\[
H_{0i}^\theta: \ \frac{\mu_i}{\mu_0} \leq \theta_1 \text{ oder } \frac{\mu_i}{\mu_0} \geq \theta_2 \quad \text{versus} \quad H_{1i}^\theta: \ \theta_1 < \frac{\mu_i}{\mu_0} < \theta_2
\]

werden in jeweils zwei einseitige Testprobleme zerlegt:

\[
H_{0i1}^\delta: \ \mu_i - \mu_0 \leq \delta_1 \quad \text{versus} \quad H_{1i1}^\delta: \ \mu_i - \mu_0 > \delta_1
\]

und

\[
H_{0i2}^\delta: \ \mu_i - \mu_0 \geq \delta_2 \quad \text{versus} \quad H_{1i2}^\delta: \ \mu_i - \mu_0 < \delta_2
\]
bzw.

\[ H_{01}^\theta : \frac{\mu_i}{\mu_0} \leq \theta_1 \quad \text{versus} \quad H_{11}^\theta : \frac{\mu_i}{\mu_0} > \theta_1 \]

und

\[ H_{02}^\theta : \frac{\mu_i}{\mu_0} \geq \theta_2 \quad \text{versus} \quad H_{12}^\theta : \frac{\mu_i}{\mu_0} < \theta_2 . \]

Die Globalhypothesen lauten dann:

\[ H_0^\delta : \bigcup_{i=1}^k \bigcup_{r=1}^2 H_{0ir}^\delta \quad \text{versus} \quad H_1^\delta : \bigcap_{i=1}^k \bigcap_{r=1}^2 H_{1ir}^\delta \]

bzw.

\[ H_0^\theta : \bigcup_{i=1}^k \bigcup_{r=1}^2 H_{0ir}^\theta \quad \text{versus} \quad H_1^\theta : \bigcap_{i=1}^k \bigcap_{r=1}^2 H_{1ir}^\theta . \]

Bei Normalverteilungsannahme wird \( H_0^\delta \) verworfen, falls jeweils alle Einzelhypothesen mittels zweier einseitiger \( t \)-Tests abgelehnt werden, oder äquivalent dazu, falls alle korrespondierenden zweiseitigen 100(1-\( \alpha \))%-Konfidenzintervalle im Sicherheitsbereich enthalten sind:

\[
t_{i1}^\delta = \frac{\bar{X}_i - \bar{X}_0 - \delta_1}{S \sqrt{\frac{1}{n_0} + \frac{1}{n_i}}} \geq t_{\alpha,0} \quad \text{und} \quad t_{i2}^\delta = \frac{\bar{X}_i - \bar{X}_0 - \delta_2}{S \sqrt{\frac{1}{n_0} + \frac{1}{n_i}}} \leq -t_{\alpha,0} \quad \forall i \in (1, ..., k)
\]

\[
[\delta_-, \delta_+] = \left[ \bar{X}_i - \bar{X}_0 \pm t_{\alpha,0} S \sqrt{\frac{1}{n_0} + \frac{1}{n_i}} \right] \subset (\delta_1, \delta_2) \quad \forall i \in (1, ..., k).
\]

Es sei darauf hingewiesen, daß Bofinger und Bofinger (1993) für den globalen Sicherheitsnachweis den maximalen Abstand \( \mu^* = \max_{i=1}^k |\mu_i - \mu_0| \) betrachteten. Dabei definierten die Autoren als Nullhypothese \( H_0^\delta : \mu^* \leq \delta \), welche verworfen werden kann, falls die obere Grenze des 100(1-\( \alpha \))%-Konfidenzintervales für \( \mu^* \) unterhalb von \( \delta \) liegt. Giani und Straßburger (1994) hingegen definierten die globale Sicherheit in der standardisierten Form \( \frac{\mu^*}{\sigma} \) und leiteten ein entsprechendes Testverfahren her. Der Nachteil beider Methoden besteht in der komplexen Bestimmung der kritischen Werte, welche nur mit
aufwendigen Algorithmen zu lösen sind. Unter Ausnutzung des intersection-union-Prinzips ergibt sich jedoch, wie oben dargestellt, ein wesentlich einfacheres Verfahren zum direkten Nachweis der globalen Sicherheit.

Für das verteilungsfreie Verfahren gilt, daß \( H_0^\delta \) abgelehnt wird, falls

\[
R_i(\delta_1) \geq \frac{r_{a_{n_0,n_i}}}{} \quad \text{und} \quad R_i(\delta_2) \leq n_i(n_0 + n_i + 1) - \frac{r_{a_{n_0,n_i}}}{n_i} \quad \forall i \in (1, ..., k)
\]

\[\Leftrightarrow\]

\[
(L^\delta_i, U^\delta_i) \subset (\delta_1, \delta_2) \quad \forall i \in (1, ..., k), \quad L^\delta_i = D^\delta_i = D^\delta_{n_i} \quad \text{und} \quad U^\delta_i = D^\delta_{n_i} = D^\delta_{n_i+1 - c_n}.
\]

Die Entscheidungsregel hinsichtlich der auf den Quotienten basierenden Nullhypothese \( H_0^\theta \) erfolgt ebenfalls mittels der entsprechenden Testverfahren bzw. mit der dazu äquivalenten Inklusionsregel:

\[
i^{(0)}_1 = \frac{X_i - \theta_1 X_0}{\sqrt{\frac{\theta_1^2}{n_0} + \frac{1}{n_i}}} \geq t_{\alpha,0} \quad \text{und} \quad i^{(0)}_2 = \frac{X_i - \theta_2 X_0}{\sqrt{\frac{\theta_2^2}{n_0} + \frac{1}{n_i}}} \leq -t_{\alpha,0} \quad \forall i \in (1, ..., k)
\]

\[\Leftrightarrow\]

\[
[\theta^{(i)}_1, \theta^{(i)}_2] \subset (\theta_1, \theta_2) \quad \forall i \in (1, ..., k) \quad \text{und} \quad X_0^2 > a_0,
\]

mit

\[
\theta^{(i)}_1 = \frac{X_0 X_i + \sqrt{a_0 X_i^2 + a_i X_0^2 - a_0 a_i}}{X_0^2 - a_0}, \quad a_0 = \frac{S^2}{n_0} t_{\alpha,0}^2 \quad \text{und} \quad a_i = \frac{S^2}{n_i} t_{\alpha,0}^2.
\]

5.1.2 Partieller Sicherheitsnachweis

5.1.2.1 Einseitiges Testproblem

Der Nachweis der globalen Sicherheit bedingt, daß für alle einzelnen Behandlungen der Nachweis der Sicherheit geführt werden muß. Diese Anforderung ist sicherlich ein sehr restriktives Kriterium und wird daher im folgenden bei der Untersuchung der partiellen Sicherheit abgeschwächt, welche im Nachweis \textit{mindestens einer} sicheren Behandlung besteht. Die entsprechenden Einzelhypothesen lauten:
Diese Formulierung entspricht dem union-intersection-Problem (Roy, 1953). Im Gegensatz zum intersection-union-Problem, bei dem alle Einzelhypothesen abgelehnt werden müssen, wird nun die Globalhypothese verworfen, falls mindestens eine der $k$ Einzelhypothesen zum entsprechend adjustierten Niveau abgelehnt werden kann.

Basierend auf der Normalverteilungsannahme $F_i(x) = \Phi\left(\frac{x - \mu_i}{\sigma}\right)$, $i = 0, 1, ..., k$, kann mittels des Dunnett-Verfahrens (Dunnett und Tamhane, 1991) der Nachweis der partiellen Sicherheit durchgeführt werden. Wesentlich ist dabei, daß dieses Verfahren den multiplen Fehler 1. Art kontrolliert, d.h. die Wahrscheinlichkeit, mindestens eine wahre Einzelhypothese irrtümlicherweise abzulehnen. Die $i$-te Einzelhypothese $H_i^\delta$ für die Differenz der Erwartungswerte oder $H_i^\theta$ für den Quotienten der Erwartungswerte wird verworfen, falls

$$
t_i^\delta = \frac{\bar{X}_i - \bar{X}_0 - \delta}{S \sqrt{\frac{1}{n_0} + \frac{1}{n_i}}} \leq -t_{a,v,k}^\delta
$$

bzw.

$$
t_i^\theta = \frac{\bar{X}_i - \theta \bar{X}_0}{S \sqrt{\frac{\theta^2}{n_0} + \frac{1}{n_i}}} \leq -t_{a,v,k}^\theta
$$

wobei die kritischen Werte $t_{a,v,k}^\delta$ und $t_{a,v,k}^\theta$ folgendermaßen hergeleitet werden.

**Lemma 6:**

Unter der jeweiligen Globalhypothese $H_0^\delta$ bzw. $H_0^\theta$ sind die Teststatistiken $(t_1^\delta, t_2^\delta, ..., t_k^\delta)$' und $(t_1^\theta, t_2^\theta, ..., t_k^\theta)$' $k$-dimensional zentral $t$-verteilt mit $v$ Freiheitsgraden und Korrelationsmatrix $R_k^\delta = (\rho_{ij}^\delta)_{i,j=1,...,k}$ bzw. $R_k^\theta = (\rho_{ij}^\theta)_{i,j=1,...,k}$, wobei

$$
H_0^\delta: \bigcap_{i=1}^k H_{0i}^\delta \text{ versus } \bigcup_{i=1}^k H_{1i}^\delta
$$

bzw.

$$
H_0^\theta: \bigcap_{i=1}^k H_{0i}^\theta \text{ versus } \bigcup_{i=1}^k H_{1i}^\theta.
$$
i) \( \rho_{ij}^\delta = \lambda_i^\delta \lambda_j^\delta \), \( 1 \leq i \neq j \leq k \), wobei

\[
\lambda_i^\delta = \frac{1}{\sqrt{1 + \frac{n_0}{n_i}}} , \quad i = 1, \ldots, k.
\]

ii) \( \rho_{ij}^0 = \lambda_i^0 \lambda_j^0 \), \( 1 \leq i \neq j \leq k \), wobei

\[
\lambda_i^0 = \frac{1}{\sqrt{\theta^2 + \frac{n_0}{n_i}}} , \quad i = 1, \ldots, k.
\]

Beweis:

i)

\[
\rho_{ij}^\delta = \frac{\text{Kov}(\bar{X}_i - \bar{X}_0 - \delta, \bar{X}_j - \bar{X}_0 - \delta)}{\sqrt{\text{Var}(\bar{X}_i - \bar{X}_0 - \delta) \text{Var}(\bar{X}_j - \bar{X}_0 - \delta)}} = \frac{\sigma^2}{n_0} = \frac{\sigma^2}{n_0} \left( \frac{1}{n_0 + \frac{n_0}{n_i}} \right) \left( \frac{1}{n_0 + \frac{n_0}{n_j}} \right)
\]

\[
\frac{1}{\sqrt{1 + \frac{n_0}{n_i}}} \cdot \frac{1}{\sqrt{1 + \frac{n_0}{n_j}}} = \lambda_i^\delta \lambda_j^\delta.
\]

ii)

\[
\rho_{ij}^0 = \frac{\text{Kov}(\bar{X}_i - \theta \bar{X}_0, \bar{X}_j - \theta \bar{X}_0)}{\sqrt{\text{Var}(\bar{X}_i - \theta \bar{X}_0) \text{Var}(\bar{X}_j - \theta \bar{X}_0)}} = \frac{\sigma^2}{n_0} \theta^2 = \frac{\sigma^2}{n_0} \theta^2 \left( \frac{\theta^2}{n_0 + \frac{n_0}{n_i}} \right) \left( \frac{\theta^2}{n_0 + \frac{n_0}{n_j}} \right)
\]

\[
\frac{\theta}{\sqrt{\theta^2 + \frac{n_0}{n_i}}} \cdot \frac{\theta}{\sqrt{\theta^2 + \frac{n_0}{n_j}}} = \lambda_i^0 \lambda_j^0.
\]

\( t_{\alpha,v,n_0^i} \) bezeichnet das \((1-\alpha)\)-Quantil der Statistik \( \max_{i=1}^k t_i^\delta \) und \( t_{\alpha,v,n_0^0} \) das \((1-\alpha)\)-Quantil der Statistik \( \max_{i=1}^k t_i^0 \). Diese Quantile können wegen der Produkteigenschaft der Korrelations-
koeffizienten mittels der SAS-Funktion PROBMC berechnet werden (Hauschke et al., 1999d).

Die äquivalente Beurteilung ergibt sich unter Verwendung der simultanen 100(1-α)%-Konfidenzintervalle. Die i-te Einzelhypothese $H_{0i}$ oder $H_{a0i}$, $i = 1, \ldots, k$, wird abgelehnt, falls

$$(-\infty, \delta_{i}^{-}) = \left( -\infty, \bar{X}_{i} - \bar{X}_{0} + t_{\alpha,v,i} \sqrt{\frac{1}{n_{0}} + \frac{1}{n_{i}}} \right) \subset (-\infty, \delta)$$

bzw.

$$(-\infty, \theta_{i}^{-}) \subset (-\infty, \theta) \quad \text{und} \quad \bar{X}_{0}^{2} > a_{0},$$

mit

$$\theta_{i}^{-} = \frac{\bar{X}_{0} \bar{X}_{i} + a_{0} \bar{X}_{i}^{2} + a_{i} \bar{X}_{0}^{2} - a_{i} a_{0}}{\bar{X}_{0}^{2} - a_{0}}, \quad a_{0} = \frac{S^{2}}{n_{0}} t_{\alpha,v,i}^{2} \quad \text{und} \quad a_{i} = \frac{S^{2}}{n_{i}} t_{\alpha,v,i}^{2}.$$ 

Es sei darauf hingewiesen, daß die obigen Konfidenzintervalle $(-\infty, \theta_{i}^{-})$ für die jeweiligen Quotienten $\frac{\mu_{i}}{\mu_{0}}, i = 1, \ldots, k$, nicht mehr für eine a-posteriori vorgenommene Sicherheitsbeurteilung verwendet werden können. Dies liegt an der Tatsache, daß zur Konstruktion dieser Konfidenzintervalle der Wert $\theta$ zur Berechnung der Korrelation und somit zur Bestimmung der kritischen Werte $t_{\alpha,v,i}^{2}$ benötigt wird. Auf Kosten eines Güteverlustes ist dies jedoch unter Ausnutzung der $\alpha$-Korrektur nach Bonferroni leicht zu beheben, d.h. man ersetzt die kritische Grenze $t_{\alpha,v,i}^{2}$ durch den Wert der univariaten $t$-Verteilung $t_{\alpha,v}^{2}$.

Zur Veranschaulichung diene die von Bedotto et al. (1989) veröffentlichte Studie Cardiac hypertrophy induced by thyroid hormone is independent of loading conditions and beta-adrenoceptor blockade. In diesem pharmakologischen Experiment an Ratten untersuchten die Autoren in einem Teilexperiment, ob die durch die Applikation des Schilddrüsenhormons T₄ induzierte links-ventrikuläre Hypertrophie durch gleichzeitige Behandlung mit Captopril, Hydralazin, Propranolol bzw. einer Kombination von Captopril und Propranolol gehemmt werden kann. Die getroffenen Schlußfolgerungen basierten auf dem indirekten Ansatz des fehlenden Unterschiedes zwischen jeweiliger Behandlung und der Kontrolle (T₄). Postuliert man jedoch die Hypothese, daß bei gleichzeitiger Verabreichung dieser Substanzen die durch T₄ induzierte Hypertrophie nur irrelevant erhöht wird, so folgt der direkte Sicherheitsnachweis als probate Methode (Hauschke und Hothorn, 1998). In der folgenden
Tabelle 20 sind die empirischen Daten und die oberen Grenzen $\theta_i$ der simultanen 95%-Konfidenzintervalle für die Quotienten $\frac{\mu_i}{\mu_0}$, $i = 1, ..., 4$, dargestellt. Dabei wurden die simultanen 95%-Konfidenzintervalle bei a-priori Vorgabe von $\theta = 1.07$ (fiktive Vorgabe) und zusätzlich unter Verwendung der Bonferroni-Korrektur ermittelt; dazu wurde die für das Gesamtexperiment vorliegende Varianzschätzung $S = 0.1538$ mit 93 Freiheitsgraden benutzt (Dunnett und Tamhane, 1991). Mit a-priori vorgegebenem $\theta = 1.07$ folgt eine irrelevante Erhöhung bei gleichzeitiger Gabe von Captopril bzw. Hydralazin. Verwendet man die konservativere Bonferroni-Adjustierung, so kann die Unbedenklichkeit nur bei gleichzeitiger Gabe von Captopril gefolgt werden.

<table>
<thead>
<tr>
<th>Behandlungsgruppe</th>
<th>Mittelwert</th>
<th>Fallzahl</th>
<th>Obere Konfidenzgrenze für $\theta = 1.07$</th>
<th>Obere Konfidenzgrenze Bonferroni-adjustiert</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_4$</td>
<td>2.52</td>
<td>10</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$T_4 +$ Captopril</td>
<td>2.49</td>
<td>9</td>
<td>1.051</td>
<td>1.054</td>
</tr>
<tr>
<td>$T_4 +$ Propranolol</td>
<td>2.60</td>
<td>12</td>
<td>1.092</td>
<td>1.094</td>
</tr>
<tr>
<td>$T_4 +$ Hydralazin</td>
<td>2.54</td>
<td>10</td>
<td>1.069</td>
<td>1.072</td>
</tr>
<tr>
<td>$T_4 +$ Captopril + Propranolol</td>
<td>2.56</td>
<td>10</td>
<td>1.078</td>
<td>1.081</td>
</tr>
</tbody>
</table>

Basierend auf der Abschlußtestprozedur (Marcus et al., 1976), haben Dunnett und Tamhane (1991) ein step-down Verfahren hergeleitet, welches im Vergleich zur obigen single-step Methode eine größere Güte aufweist. Dabei werden die Teststatistiken der Größe nach geordnet:

$$ t_{(i)}^\delta \leq t_{(2)}^\delta \leq \ldots \leq t_{(k)}^\delta \quad \text{bzw.} \quad t_{(i)}^\theta \leq t_{(2)}^\theta \leq \ldots \leq t_{(k)}^\theta .$$

Die jeweils korrespondierenden Hypothesen seien mit $H_{0(1)}^\delta, H_{0(2)}^\delta, \ldots, H_{0(k)}^\delta$ bzw. $H_{0(1)}^\theta, H_{0(2)}^\theta, \ldots, H_{0(k)}^\theta$ bezeichnet. Die Hypothesen $H_{0(i)}^\delta$ bzw. $H_{0(i)}^\theta$, $i = 1, ..., k$, werden abgelehnt, falls

$$ t_{(i)}^\delta \leq -t_{\alpha, i, \bar{R}_{1-n}} \quad \text{und alle Hypothesen} \ H_{0(1)}^\delta, H_{0(2)}^\delta, \ldots, H_{0(i-1)}^\delta \ \text{verworfen wurden}$$
Das *step-down* Verfahren wird für den Beispielldatensatz von Bedotto *et al.* (1989) verwendet. In der Tabelle 21 sind die geordneten Teststatistiken $t^0_{(i)}$ und die kritischen Werte $-t^0_{0.05, R_{k+1}}$, $i = 1, ..., 4$, dargestellt. Mit $\theta = 1.07$ folgt neben der irrelevanten Erhöhung bei gleichzeitiger Gabe von Captopril ebenfalls eine nicht relevante Erhöhung für die gleichzeitige Verabreichung von Hydralazin.

**Tabelle 21:** Geordnete Teststatistiken und kritische Werte für die Vergleiche mit der Kontrolle $T_4$

<table>
<thead>
<tr>
<th>Behandlungsgruppe</th>
<th>$t^0_{(i)}$</th>
<th>$-t^0_{0.05, R_{k+1}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_4 + \text{Captopril}$</td>
<td>-2.825</td>
<td>-2.178</td>
</tr>
<tr>
<td>$T_4 + \text{Hydralazin}$</td>
<td>-2.196</td>
<td>-2.077</td>
</tr>
<tr>
<td>$T_4 + \text{Captopril + Propranolol}$</td>
<td>-1.915</td>
<td>-1.930</td>
</tr>
<tr>
<td>$T_4 + \text{Propranolol}$</td>
<td>-1.409</td>
<td>-1.661</td>
</tr>
</tbody>
</table>

Eine analoge *step-up* Prozedur wurde von Dunnett und Tamhane (1992, 1995) vorgeschlagen, wobei jedoch der Nachweis der Kontrolle des multiplen Fehlers 1. Art für ungleiche Korrelationen bislang noch nicht erbracht wurde. Wesentlich einfacher gestaltet sich die Verwendung der $\alpha$-Korrektur nach Bonferroni; dabei werden die kritischen Werte $t^0_{\alpha, R_k}$ und $t^0_{\alpha, R_k}$ der $k$-dimensionalen $t$-Verteilung durch den Wert der univariaten $t$-Verteilung $t^{\alpha}_{k}$ ersetzt. Insbesondere für großes $k$ ist die Vorgehensweise konservativ, daher wurde diese durch die *step-down* Methode nach Holm (1979) verbessert. Dazu werden nun die $p$-Werte der Teststatistiken der Größe nach geordnet:

$$p^\alpha_{(1)} \leq p^\alpha_{(2)} \leq ... \leq p^\alpha_{(k)} \quad \text{bzw.} \quad p^\theta_{(1)} \leq p^\theta_{(2)} \leq ... \leq p^\theta_{(k)},$$

die korrespondierenden Hypothesen $H^\delta_{0(i)}$ bzw. $H^\theta_{0(i)}$, $i = 1, ..., k$, werden abgelehnt, falls

$$p^\alpha_{(i)} \leq \frac{\alpha}{k - i + 1} \quad \text{und alle Hypothesen} \quad H^\delta_{0(1)}, H^\delta_{0(2)}, ..., H^\delta_{0(k)},$$

bzw.
\[ p_{(i)}^0 \leq \frac{\alpha}{k - i + 1} \] und alle Hypothesen \( H_{0(i)}, H_{0(2)}, \ldots, H_{0(k-i)} \) verworfen wurden.


Eine verteilungsfreie Prozedur zum multiplen Niveau \( \alpha \) hat Steel (1959) für den \textit{many-to-one} Vergleich vorgeschlagen: Für identische Stichprobenumfänge der Behandlungsgruppen \( n_1 = n_2 = \ldots = n_k = m \) und \( m \) nicht notwendigerweise identisch mit \( n_0 \), wird die Hypothese \( H_{0(i)} \) abgelehnt, falls

\[
R_{(i)}(\delta) = \sum_{j=1}^{n_i} R_{y_j}(\delta) \leq m(n_0 + m + 1) - \delta_{\alpha,i,n_0,m}, \text{ dabei bezeichnet } \delta_{\alpha,i,n_0,m} \text{ das (1-}\alpha\text{-Quantil der Statistik } \max_{i=1}^{k} R_{k}(\delta). \text{ Exakte kritische Werte hat Steel für } k = 2, 3 \text{ und } n_0 = m = 3, 4, 5 \text{ tabelliert; einfache Approximationsformeln wurden von Hochberg und Tamhane (1987) und Hsu (1996) angegeben. Die einseitigen 100(1-}\alpha\text{-\%-Konfidenzintervalle für } \mu_i - \mu_0 \text{ sind}

\[ (-\infty, U_i^\delta] \], \ wobei \( U_i^\delta = D_{a_i}^\delta = D_{a,m+1-c_i}^\delta \), \( C_i = \frac{m(2n_0 + m + 1)}{2} + 1 - r_{\alpha,k,n_0,m}^\delta \).

5.1.2.2 Zweiseitiges Testproblem

 Analog zum globalen Sicherheitsnachweis lautet die zweiseitige Hypothese für den partiellen Nachweis der Unbedenklichkeit:

\[ H_{0}^\delta : \bigcap_{i=1}^{k} \bigcup_{r=1}^{2} H_{0ir}^\delta \text{ versus } H_{1}^\delta : \bigcup_{i=1}^{k} \bigcap_{r=1}^{2} H_{1ir}^\delta \]

bzw.

\[ H_{0}^\theta : \bigcap_{i=1}^{k} \bigcup_{r=1}^{2} H_{0ir}^\theta \text{ versus } H_{1}^\theta : \bigcup_{i=1}^{k} \bigcap_{r=1}^{2} H_{1ir}^\theta . \]

Bei Normalverteilungsannahme kann \( H_{0}^\delta \) unter Verwendung der Bonferroni-Korrektur überprüft werden, wobei die \( k \) Einzelhypothesen mittels der zwei einseitigen \( t \)-Tests getestet werden oder äquivalent dazu, mittels der Inklusionsregel:
\[ t_{1_i}^\delta = \frac{X_i - \overline{X}_0 - \delta_{i}}{S \sqrt{\frac{1}{n_0} + \frac{1}{n_i}}} \geq t_{\alpha/2}^{\overline{X}^0} \quad \text{und} \quad t_{2_i}^\delta = \frac{X_i - \overline{X}_0 - \delta_{i}}{S \sqrt{\frac{1}{n_0} + \frac{1}{n_i}}} \leq -t_{\alpha/2}^{\overline{X}^0} \]
\[ \iff \]
\[ [\delta_{\overline{i}}, \delta_{\overline{j}}] = \left[ \overline{X}_i - \overline{X}_0 \pm t_{\alpha/2}^{\overline{X}^0} S \sqrt{\frac{1}{n_0} + \frac{1}{n_i}} \right] \subset (\delta_{\overline{i}}, \delta_{\overline{j}}). \]

Die Bonferroni-Adjustierung kann wiederum durch die step-down Prozedur nach Holm verbessert werden: es seien \( p_{i_1}^\delta = \max(p_{i_1}^\delta, p_{i_2}^\delta) \) und \( p_{i_1}^\theta = \max(p_{i_1}^\theta, p_{i_2}^\theta) \), wobei \( p_{i_1}^\delta \), \( p_{i_1}^\theta \) die \( p \)-Werte der Teststatistiken \( t_{i_1}^\delta \) und \( t_{i_1}^\theta \) bezeichnen, \( r = 1, 2 \). Die \( p \)-Werte werden der Größe nach geordnet:
\[ p_{(1)}^\delta \leq p_{(2)}^\delta \leq \ldots \leq p_{(k)}^\delta \quad \text{bzw.} \quad p_{(1)}^\theta \leq p_{(2)}^\theta \leq \ldots \leq p_{(k)}^\theta, \]

und die korrespondierenden Hypothesen \( H_{0(i)}^\delta \) bzw. \( H_{0(i)}^\theta \), \( i = 1, \ldots, k \), werden abgelehnt, falls
\[ p_{(i)}^\delta \leq \frac{\alpha}{k - i + 1} \quad \text{und alle Hypothesen} \quad H_{0(1)}^\delta, H_{0(2)}^\delta, \ldots, H_{0(k-i)}^\delta \quad \text{verworfen wurden} \]

bzw.
\[ p_{(i)}^\theta \leq \frac{\alpha}{k - i + 1} \quad \text{und alle Hypothesen} \quad H_{0(1)}^\theta, H_{0(2)}^\theta, \ldots, H_{0(k-i)}^\theta \quad \text{verworfen wurden}. \]

Unter der Annahme identischer Stichprobenumfänge der Behandlungsgruppen, d.h. \( n_1 = n_2 = \ldots = n_k = m \), haben Bofinger und Bofinger (1995) die folgenden simultanen 100(1-\( \alpha \))-\%-Konfidenzintervalle für \( \mu_i - \mu_0 \) hergeleitet:
\[
\left[ \overline{X}_i - \overline{X}_0 - c_k S \sqrt{\frac{1}{n_0} + \frac{1}{m}} \right] \leq \left[ \overline{X}_i - \overline{X}_0 + c_k S \sqrt{\frac{1}{n_0} + \frac{1}{m}} \right],
\]

wobei \( x^- = \min(0, x) \), \( x^+ = \max(0, x) \) und \( c_k \) spezielle, aus der multivariaten \( t \)-Verteilung abgeleitete kritische Werte darstellen. Die Einzelhypothesen lassen sich dann mit der Inklusionsregel testen:
\[
\left( \bar{X}_i - \bar{X}_0 - c_k S \sqrt{\frac{1}{n_0} + \frac{1}{m}} \right)^- \left( \bar{X}_i - \bar{X}_0 + c_k S \sqrt{\frac{1}{n_0} + \frac{1}{m}} \right)^+ \subset (\delta_1, \delta_2).
\]

Für das analoge verteilungsfreie Verfahren gilt, daß \( H_0^\delta \) abgelehnt wird, falls

\[
R_i(\delta_1) \geq r_{\alpha/n_0,n_i}^{\delta_i} \quad \text{und} \quad R_i(\delta_2) \leq n_i(n_0 + n_i + 1) - r_{\alpha/n_0,n_i}^{\delta_2},
\]

\[
\iff [L_i^\delta, U_i^\delta] \subset (\delta_1, \delta_2), \quad L_i^\delta = D_i^\delta = D_{c^\alpha}^{\delta_i} \quad \text{und} \quad U_i^\delta = D_i^\delta = D_{n_0n_i+1-c^\alpha}^{\delta_i}.
\]

Die Überprüfung von \( H_0^\theta \) und der Einzelhypothesen erfolgt analog:

\[
t_i^{0 \Downarrow 0} = \frac{\bar{X}_i - \theta_0 \bar{X}_0}{S \sqrt{\frac{1}{n_0} + \frac{1}{n_i}}} \geq t_{\alpha/n_0,n_i}^{0 \Downarrow 0} \quad \text{und} \quad t_i^{0 \Downarrow 2} = \frac{\bar{X}_i - \theta_2 \bar{X}_0}{S \sqrt{\frac{1}{n_0} + \frac{1}{n_i}}} \leq -t_{\alpha/n_0,n_i}^{0 \Downarrow 2},
\]

\[
\iff [\theta_0^\downarrow, \theta_2^\downarrow] \subset (\theta_1, \theta_2) \quad \text{und} \quad \bar{X}_0^2 > a_0,
\]

mit

\[
\theta_i^\downarrow = \frac{\bar{X}_0 \bar{X}_i + \sqrt{a_0 \bar{X}_i^2 + a_i \bar{X}_0^2 - a_0 a_i}}{\bar{X}_0^2 - a_0}, \quad a_0 = \frac{S^2}{n_0} \frac{t_{\alpha/n_0,n_i}^{0 \Downarrow 0}^2}{\bar{X}_0^2 - a_0}, \quad \text{und} \quad a_i = \frac{S^2}{n_i} \frac{t_{\alpha/n_0,n_i}^{0 \Downarrow 0}^2}{\bar{X}_0^2 - a_0}.
\]

5.2 Simple-order Alternative

Statt der verschiedenen Substanzen oder entsprechender Kombinationen liegen nun \( k \) aufsteigende Dosierungen vor:

\[
\{\text{Kontrolle, Dosis}_1, \text{Dosis}_2, \ldots, \text{Dosis}_k\}.
\]

Die Zielgrößen \( X_{ij} \) sind unabhängig identisch verteilt nach den stetigen Verteilungs- funktionen \( F_i(x) = F(x - \mu_i), \quad i = 0, 1, \ldots, k \) und \( j = 1, \ldots, n_i \); zusätzlich wird für die Lokationsparameter meist eine monotone Ordnung postuliert: \( \mu_0 \leq \mu_1 \leq \ldots \leq \mu_k \). Das Ziel besteht in der Bestimmung der maximalen Dosierung, die im Vergleich zur Kontrolle nur irrelevant erhöht ist.
5.2.1 Bestimmung der maximal sicheren Dosis

Kann der Schwellenwert $\delta$ absolut angegeben werden, so ist maximal sicherere Dosis wie folgt definiert:

$$ MAXSD^\delta = D_i \quad \text{wobei} \quad i^* = \max(i : \mu_i - \mu_0 < \delta) $$

und zusätzlich

$$ \forall j, j = 1, ..., i^* - 1 : \mu_j - \mu_0 < \delta. $$

Unter der Monotoneannahme $\mu_0 \leq \mu_1 \leq ... \leq \mu_i$ ist die zusätzliche Bedingung automatisch erfüllt, daß alle Dosierungen, die kleiner als die maximal sichere Dosis sind, auch sicher sind (Hothorn und Hauschke, 1999). Ausgehend von dem Konzept der a-priori geordneten Hypothesen (Hothorn und Lehmacher, 1991; Maurer et al., 1995; Hsu und Berger, 1998) kann eine Prozedur abgeleitet werden, die einerseits den multiplen Fehler 1. Art auf $\alpha$ begrenzt und andererseits die Annahme der monotonen Ordnung nicht benötigt (Neuhäuser und Hothorn, 1995, 1997; Hauschke, 1995; Hauschke und Hothorn, 1998). Beginnend mit dem Testproblem:

$$ H_{01}^\delta : \mu_1 - \mu_0 \geq \delta $$

versus

$$ H_{11}^\delta : \mu_1 - \mu_0 < \delta $$

wird ein Test zum Niveau $\alpha$ durchgeführt. Falls $H_{01}^\delta$ abgelehnt werden kann, wird

$$ H_{02}^\delta : \mu_2 - \mu_0 \geq \delta $$

versus

$$ H_{12}^\delta : \mu_2 - \mu_0 < \delta $$

getestet. Wird $H_{01}^\delta$ nicht abgelehnt, so ist die Aussage zu treffen, daß im vorliegenden Design die $MAXSD^\delta$ nicht bestimmt werden kann; inhaltlich hat dies zur Folge, daß ein neues Experiment mit kleiner gewählten Dosierungen erfolgen muß. Allgemein erfolgt der Übergang von

$$ H_{0i}^\delta : \mu_i - \mu_0 \geq \delta $$

versus

$$ H_{1i}^\delta : \mu_i - \mu_0 < \delta $$
zu

\[ H^\delta_{0i+1} : \mu_{ri1} - \mu_0 \geq \delta \]

versus

\[ H^\delta_{1i+1} : \mu_{ri1} - \mu_0 < \delta \]

genaus dann, falls \( H^\delta_{0i} \) zum Niveau \( \alpha \) verworfen werden konnte. Als Schätzer für die \( MAXSD^\delta \) dient die maximale Dosis, die im Vergleich zur Kontrolle keine relevante Erhöhung aufweist. Es sei darauf hingewiesen, daß bei der Bestimmung der \( MAXSD^\delta \) die Auswahl der Dosierungen von eminenter Bedeutung ist. Ohne Angabe des funktionalen Zusammenhanges zwischen Dosis und Effekt kann - falls überhaupt möglich - nur eine der gewählten Dosierungen als \( MAXSD^\delta \) abgeleitet werden.

Kann der Schwellenwert \( \delta \) nur relativ zum Erwartungswert der Kontrolle angegeben werden, so lautet die entsprechende Definition der maximal sichereren Dosis:

\[
MAXSD^\theta = D_{i^*} \text{ wobei } i^* = \max(i : \frac{\mu_i}{\mu_0} < \theta) 
\]

und zusätzlich

\[
\forall j, j = 1, ..., i^* - 1 : \frac{\mu_j}{\mu_0} < \delta .
\]

Die Bestimmung der maximal sicheren Dosis \( MAXSD^\theta \) erfolgt analog zur Ermittlung von \( MAXSD^\delta \) (Tamhane et al., 1999) und wird im folgenden an einem Beispiel mit drei Dosierungen mit der normalverteilten Zielvariablen Serum-ASAT-Werte erläutert. Hierbei wird die Varianzschätzung \( S = 0.5694 \) mit 59 Freiheitsgraden benutzt und als irrelevante Erhöhung sei \( \theta = 1.25 \) gewählt. Die Teststatistiken lauten:

\[
t^\theta_i = \frac{\overline{X}_i - 1.25 \overline{X}_0}{S \sqrt{\frac{1.25^2}{n_0} + \frac{1}{n_i}}}, \text{ } i = 1, 2, 3.
\]

In der Tabelle 22 sind die Ergebnisse für den Vergleich der Dosisgruppen mit der Kontrollgruppe dargestellt. Bei einem Niveau von \( \alpha = 0.05 \) bricht das Verfahren der \textit{a-priori} geordneten Hypothesen beim Vergleich der Dosisgruppe 200mg/kg, da der zugehörige \textit{p}-Wert größer ist als 0.05. Als maximal sichere Dosis folgt somit die Dosis 20mg/kg.
Tabelle 22: Mittelwerte, Fallzahlen für die Zielvariable ASAT-Werte (µmol/L),
Werte der Teststatistiken und entsprechende p-Werte

<table>
<thead>
<tr>
<th>Dosisgröße (mg/kg)</th>
<th>Mittelwert</th>
<th>Fallzahl</th>
<th>$t_i^0$</th>
<th>p-Wert</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2.087</td>
<td>19</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>20</td>
<td>2.064</td>
<td>16</td>
<td>-2.5147</td>
<td>0.0073</td>
</tr>
<tr>
<td>200</td>
<td>2.323</td>
<td>15</td>
<td>-1.3005</td>
<td>0.0992</td>
</tr>
<tr>
<td>1000</td>
<td>2.753</td>
<td>13</td>
<td>0.6350</td>
<td>-</td>
</tr>
</tbody>
</table>

5.2.2 Alternative Konzepte

Eine übliche Auswertungsstrategie basiert auf dem Abschlußtestprinzip (Marcus et al., 1976) und besteht darin, daß im ersten Schritt ein Trendtest (Jonckheere, 1954; Bartholomew, 1961) mit allen Behandlungsgruppen durchgeführt wird:

\[ H_0 : \mu_0 = \mu_1 = \ldots = \mu_k \]

versus

\[ H_1 : \mu_0 \leq \mu_1 \leq \ldots \leq \mu_k, \quad \mu_0 < \mu_k. \]

Ist der p-Wert kleiner oder gleich α, so wird der Test ohne die höchste Dosis $D_k$ wiederholt (Hothorn et al., 1997). Im Falle einer erneuten Signifikanz zum Niveau α erfolgt der entsprechende Vergleich ohne die Dosierungen $D_k$ und $D_{k-1}$. Diese Verfahrensweise des Ausschlusses der jeweils höchsten Dosis wird solange fortgesetzt, bis erstmalig ein nichtsignifikantes Ergebnis ($p > \alpha$) vorliegt. Die in diesem letzten Vergleich vorliegende höchste Dosis wird dann als maximal sichere Dosis MAXSD (MAXIMUM NO EFFECT DOSE) oder NOED (NO observed effect dose) oder NOEL (no observed effect level) bezeichnet (Williams, 1971, 1972; Shirley, 1977; Tukey et al., 1985; Rom et al., 1994). Bei Verletzung der Monotonieannahme, zum Beispiel bei Vorliegen einer Umbrella-Alternativen sind entsprechend modifizierte Verfahren zu benutzen (Mack und Wolfe, 1981; Hettmansperger und Norton, 1987).

schätzen, die im Vergleich zur Kontrolle eine nur irrelevante Erhöhung aufweist. Um der vorliegenden Variabilität Rechnung zu tragen, wird auch die untere Grenze des entsprechenden Konfidenzintervales als Schätzer verwendet.

6 Konzept der Bioäquivalenz


6.1 Modell, Testproblem und Entscheidungsverfahren

Die Wirksamkeit eines Arzneimittels ist in der Regel an eine bestimmte minimale Plasmakonzentration gebunden, damit der erwünschte Effekt eintritt. Zu hohe Plasmaspiegel sind dagegen oft mit unerwünschten Nebenwirkungen verbunden. Deshalb besteht die Notwendigkeit, zum Nachweis der Bioäquivalenz die Gleichwertigkeit sowohl für das Ausmaß als auch für die Geschwindigkeit der Resorption zu bestimmen (CPMP, 1991; FDA, 1992). Statt der Plasmaspiegelkurve selbst werden daraus abgeleitete pharmakokinetische Kenngrößen berechnet. Als charakterisierende Größen dienen für das Ausmaß die Fläche unter der Plasmaspiegelkurve \( (\text{AUC} = \text{area under the curve}) \) und für die Geschwindigkeit bei schnell freisetzenden Formulierungen Zeitpunkt \( (t_{\text{max}}) \) und Höhe des maximalen Serumspiegels \( (C_{\text{max}}) \). Bei kontrolliert freisetzenden Arzneimittelformen hingegen sind zur Charakterisierung der Resorptionsgeschwindigkeit andere Zielgrößen aussagefähiger (Hauschke und Steinijans, 1993).

Für viele Substanzen ist eine große interindividuelle Variabilität bekannt. Während zum Beispiel für die \( \text{AUC} \) zwischen den Individuen eines Kollektivs Variationskoeffizienten von 30% üblich sind, beträgt in einem Crossover-Design der intraindividuelle Variationskoeffizient oft nur noch 15% (Hauschke et al., 1994). Aus diesem Grunde werden Bioäquivalenzstudien als Crossover angelegt, wobei Carryover-Effekte durch eine hinreichend lang gewählte Washout-Phase vernachlässigt werden können. Beim einfachen
Zwei-Perioden-Crossover-Design werden Test (T) und Referenz (R) somit intraindividuell verglichen, wobei die Probanden den Behandlungssequenzen RT bzw. TR zufällig zugeordnet werden.

Für die meisten Kenngrößen wird angenommen, daß sie logarithmisch normalverteilt sind. Diese Annahme ist für die AUC durch die pharmacokinetische Grundgleichung

\[ AUC = \frac{f \cdot Dosis}{Clearance} \]

gerechtfertigt, wobei \( f \), \( 0 < f \leq 1 \), den bioverfügbaren Dosisanteil und damit den Formulierungseffekt darstellt, während die Clearance (Maß für die Eliminationsgeschwindigkeit) dem Individuumseffekt entspricht (Steinijans und Hauschke, 1990). Eine analoge Argumentation gilt für alle konzentrationsabhängigen Kenngrößen, insbesondere für \( C_{\text{max}} \). Das entsprechende multiplikative Modell lautet:

\[ X_{jk} = \mu_h \cdot \pi_k \cdot b_{ij} \cdot e_{jk}, \]

wobei \( X_{jk} \) die Beobachtung der pharmakokinetischen Kenngröße bei dem \( j \)-ten Probanden, \( j = 1, \ldots, n_j \), in der \( i \)-ten Sequenz, \( i = 0 \) (RT), \( 1 \) (TR) aus der \( k \)-ten Periode, \( k = 1, 2 \). Der Effekt der Formulierung \( h \) wird durch \( \mu_h \) gekennzeichnet, \( h = R \) für \( i + 1 = k \) und \( h = T \) für \( i + 1 \neq k \), für die Periodeneffekte gilt \( \pi_1 \cdot \pi_2 = 0 \). Für den zufälligen Probandeneffekt \( b_{ij} \) gelte \( \ln b_{ij} \) und für den zufälligen Fehlerterm \( e_{jk} \) gelte

\[ \ln e_{jk} \text{ u.i.v } \sim \Phi \left( \frac{x}{\sigma_e} \right). \]

Es sei darauf hingewiesen, daß für einige Kenngrößen der Resorptionsgeschwindigkeit, zum Beispiel \( t_{\text{max}} \), ein additives Modell angenommen wird (Hauschke und Steinijans, 1993). Die Auswertung erfolgt üblicherweise mittels des verteilungsfreien Verfahrens nach Hauschke et al. (1990).

Aus Behördensicht ist die irrtümliche Ablehnung der Bioäquivalenz (Produzentenrisiko) von sekundärer Bedeutung; wesentlich ist die Kontrolle der irrtümlichen Annahme der Bioäquivalenz (Konsumentenrisiko). Aus diesem Grund wird das Entscheidungsproblem mit einer Intervallalternative formuliert:

\[ H_0^0: \frac{\mu_T}{\mu_R} \leq \theta_1 \text{ oder } \frac{\mu_T}{\mu_R} \geq \theta_2. \]
versus

\[ H_0^\delta: \theta_1 < \frac{\mu_T}{\mu_R} < \theta_2, \]

wobei \((\theta_1, \theta_2), 0 < \theta_1 < 1 < \theta_2, \theta_1 = \frac{1}{\theta_2}\) den Äquivalenzbereich bezeichne. Im Falle der \(AUC\) und \(C_{max}\) werden üblicherweise \(\theta_1 = 0.80\) und \(\theta_2 = 1.25\) gewählt (FDA, 1992), wobei in Europa für \(C_{max}\) auch die Grenzen \(\theta_1 = 0.70\) und \(\theta_2 = 1.43\) zulässig sind (CPMP, 1991).

Durch logarithmische Transformation der Kenngröße wird das multiplikative Modell in das additive Modell übergeführt und die äquivalente Formulierung des Testproblems lautet

\[ H_0^\delta: \ln \mu_T - \ln \mu_R \leq \ln \theta_1 \quad \text{oder} \quad \ln \mu_T - \ln \mu_R \geq \ln \theta_2 \]

versus

\[ H_1^\delta: \ln \theta_1 < \ln \mu_T - \ln \mu_R < \ln \theta_2. \]

Daher können die im Kapitel 4.1 dargestellten Methoden für die Differenz von Lokationsparametern direkt verwendet werden. Man berechnet zum Beispiel zunächst das 100(1-2\(\alpha\))-\%-Konfidenzintervall für \(\ln \mu_T - \ln \mu_R\) und erhält durch anschließende inverse Transformation (\(\exp(\cdot)\)) der Intervallgrenzen ein 100(1-2\(\alpha\))-\%-Konfidenzintervall für \(\frac{\mu_T}{\mu_R}\). Ist die Annahme der Log-Normalverteilung für die entsprechenden Kenngrößen zweifelhaft, so kann das nichtparametrische Verfahren nach Hauschke et al. (1990) verwendet werden.

**6.2 Anwendungsbeispiel**

Das Beispiel bezieht sich auf eine Dosisäquivalenzstudie, bei der im randomisierten Zwei-Perioden-Crossover Versuch mit 7tägiger Washout-Phase an 18 gesunden Probanden 1000mg Theophyllin als Pellets in unterschiedlicher Kapselgrößen verglichen wurden (Steinijans et al., 1989). Je zwei Kapseln mit 200 und 300mg Theophyllin dienten als Referenz, zwei Kapseln mit 500mg dienten als Test. Im Akutteil der Studie, d.h. nach Gabe der ersten Dosis, erfolgte die Blutentnahmen zur Theophyllinbestimmung im Plasma vor und 1, 2, 3, 4, 5, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 28, 36, 40, 44 und 48 Stunden nach der abendlichen Applikation um 20 Uhr. Die Berechnung der \(AUC\) (Tabelle 23) erfolgte bis 48 Stunden mit der linearen Trapezregel und wurde aufgrund des terminalen monoexponentiellen
Konzentrationsverlaufes bis Unendlich extrapoliert (Sauter et al., 1992). Der Anteil der bis zum letzten Messzeitpunkt bestimmten AUC zu der Gesamt-AUC betrug 94 (91-99)% in der Testsituation und 95 (84-99)% in der Referenzsituation; Angaben als Median (Minimum – Maximum).

<table>
<thead>
<tr>
<th>Sequenz</th>
<th>Periode 1</th>
<th>Periode 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>TR</td>
<td>228.04</td>
<td>288.79</td>
</tr>
<tr>
<td>RT</td>
<td>339.03</td>
<td>329.76</td>
</tr>
<tr>
<td>TR</td>
<td>288.21</td>
<td>343.37</td>
</tr>
<tr>
<td>RT</td>
<td>242.64</td>
<td>258.19</td>
</tr>
<tr>
<td>TR</td>
<td>249.94</td>
<td>201.56</td>
</tr>
<tr>
<td>TR</td>
<td>217.97</td>
<td>225.77</td>
</tr>
<tr>
<td>TR</td>
<td>133.13</td>
<td>235.89</td>
</tr>
<tr>
<td>RT</td>
<td>184.32</td>
<td>249.64</td>
</tr>
<tr>
<td>TR</td>
<td>213.78</td>
<td>215.14</td>
</tr>
<tr>
<td>TR</td>
<td>248.98</td>
<td>245.48</td>
</tr>
<tr>
<td>TR</td>
<td>163.93</td>
<td>134.89</td>
</tr>
<tr>
<td>RT</td>
<td>209.30</td>
<td>231.98</td>
</tr>
<tr>
<td>RT</td>
<td>207.40</td>
<td>234.19</td>
</tr>
<tr>
<td>TR</td>
<td>245.92</td>
<td>223.39</td>
</tr>
<tr>
<td>RT</td>
<td>239.84</td>
<td>241.25</td>
</tr>
<tr>
<td>RT</td>
<td>211.24</td>
<td>255.60</td>
</tr>
<tr>
<td>TR</td>
<td>188.05</td>
<td>169.70</td>
</tr>
<tr>
<td>RT</td>
<td>230.36</td>
<td>256.55</td>
</tr>
</tbody>
</table>

In Tabelle 24 ist das parametrische und das verteilungsfreie 90%-Konfidenzintervall für \( \frac{\mu_T}{\mu_R} \) aufgeführt. Beide Konfidenzintervalle sind vollständig im Äquivalenzbereich (0.80, 1.25) enthalten, so daß mit einer maximalen Irrtumswahrscheinlichkeit (Konsumentenrisiko) von 0.05 auf Äquivalenz hinsichtlich des Ausmaßes der Resorption geschlossen werden kann.

<table>
<thead>
<tr>
<th>Tabelle 24: Zweiseitige 90%-Konfidenzintervalle für ( \frac{\mu_T}{\mu_R} ) für den Beispieldatensatz aus Tabelle 23</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parametrisches Verfahren</td>
</tr>
<tr>
<td>Nichtparametrisches Verfahren</td>
</tr>
</tbody>
</table>

Die klassische Beurteilung der Bioäquivalenz (average bioequivalence) basiert bislang nur auf den Lokationsparametern. Zur Beurteilung der Ähnlichkeit hinsichtlich der Variabilität wird routinemäßig ein Groups-by-periods-Plot erstellt; für die Dosisäquivalenzstudie erkennt
man anhand der Abbildung 14, daß die Referenzformulierung eine größere Variabilität aufweist.

Abbildung 14: Groups-by-Periods-Plot für den Beispieldatensatz aus Tabelle 23 ( ● : Test, ○ : Referenz):

Geometrische Mittel und 68%-Bereiche

Neben der Forderung der Vergleichbarkeit hinsichtlich der Variabilität ist das Nichtvorhandensein einer ausgeprägten subject-by-formulation Interaktion ein wesentliches Maß für die Austauschbarkeit von Referenz- und Testformulierung; in der Abbildung 15 wird dies grafisch dargestellt. Man erkennt, daß - bezogen auf die individuellen Bioverfügbarkeits-

\[
\frac{X_{0j2}}{X_{0j1}} \text{ bzw. } \frac{X_{1j1}}{X_{1j2}}, \ j = 1, ..., n_j, \ i = 0, 1 - \text{ die Mehrzahl der Werte zwar sehr nahe bei 1 ist, aber auch extreme Werten oberhalb und unterhalb von 1 vorliegen.}
\]

Diese beiden zusätzlichen Anforderungen für die Austauschbarkeit von Formulierungen wurden in dem von der Food and Drug Administration (1997) erstellten Entwurf In vivo bioequivalence studies based on population and individual bioequivalence approaches gestellt.
Es sei jedoch darauf hingewiesen, daß die in dem Entwurf geforderte statistische Methodik noch Unzulänglichkeiten aufweist (Hauschke et al., 1999e), zum Beispiel die fehlende Interpretierbarkeit des gewählten Parameters (aggregate criterion) oder die nicht adäquate Ermittlung der Stichprobengröße.
7 Zusammenfassung und Ausblick

Bislang ist es in der Praxis üblich, bei der Beurteilung von Substanzen hinsichtlich ihrer Sicherheit den indirekten Ansatz des proof of hazard anzuwenden. Dabei wird die Nullhypothese $H_0$: Substanz unbedenklich gegen die Alternative $H_1$: Substanz bedenklich getestet, und man folgert die Unbedenklichkeit der Substanz, falls die Nullhypothese zum Niveau $\alpha$ nicht verworfen werden kann. Der entscheidende Nachteil dieses Ansatzes ist, daß nur die Wahrscheinlichkeit der irrtümlichen Verwerfung der Nullhypothese (Produzentenrisiko) direkt durch den vorgegebenen Fehler 1. Art $\alpha$ begrenzt ist. Weitaus wichtiger ist jedoch die Kontrolle des Konsumentenrisikos, nämlich der irrtümlichen Annahme, daß die Substanz unbedenklich ist.


Häufig wird der Schwellenwert auch relativ zum Erwartungswert einer Kontrollgruppe definiert, so daß sich das daraus resultierende Testproblem auf den Quotienten der Lokationsparameter bezieht. Daher wurden für diese Fragestellung die adäquaten biometrischen Methoden vorgestellt. Es sei darauf hingewiesen, daß die Methoden sich ausschließlich auf Lokationsvergleiche beziehen. Für einen simultanen Lokations- und Dispersionsvergleich kann das von Bauer und Bauer (1994) vorgestellte Konzept übertragen werden. Bauer und Bauer publizierten ein Verfahren zum simultanen Äquivalenznachweis für die Differenz von Lokationsparametern und für die Skalenparameter. Eine analoge Vorgehensweise für die Situation des simultanen Äquivalenznachweises für den Quotienten von Lokationsparametern und für die Skalenparameter erfordert jedoch zunächst, daß die Teststatistik für den Quotienten der Lokationsparameter auf den Fall ungleicher Varianzen verallgemeinert wird.
In der Zweistichprobensituation existieren bislang für die Beurteilung des Quotienten nur parametrische Methoden, da schon unter der Normalverteilungsannahme die Voraussetzung des Lokationsmodells verletzt ist. Es sollte daher untersucht werden, ob in dieser Situation sich alternative Methoden, zum Beispiel Bootstrap-Verfahren, als geeignet erweisen.

Das Crossover-Design kann durch die Betrachtung der Periodendifferenzen bzw. der korrigierten Periodendifferenzen auf eine Zweistichprobensituation zurückgeführt werden. Daher wurden die für die Zweistichprobensituation hergeleiteten Verfahren auf die Crossover-Situation übertragen. Bemerkenswert ist die Tatsache, daß die Transformation zu einer Varianzhomogenität innerhalb der beiden resultierenden Stichproben führt. Dies ist unabhängig davon, ob die den jeweiligen Behandlungen zugrundeliegenden Varianzen homogen sind oder nicht.

In der Mehrstichprobensituation bestand das zentrale Anliegen in der Ermittlung sicherer Substanzen bzw. in der Bestimmung der maximal sicheren Dosis. In der many-to-one-Situation wurden zum Nachweis der globalen Sicherheit das intersection-union-Prinzip verwendet, so daß wiederum die für die Zweistichprobensituation hergeleiteten Verfahren zur Anwendung kamen. Für den partiellen Sicherheitsnachweis wurde in der einseitigen Fragestellung die Dunnett-Prozedur verwendet bzw. zur Beurteilung der Quotienten der Lokationsparameter entsprechend verallgemeinert. Bei der zweiseitigen Fragestellung wurde zur Konstruktion der simultanen Konfidenzintervalle die Bonferroni-Korrektur verwendet. Daher sollte untersucht werden, ob diese konservative Vorgehensweise nicht durch ein Verfahren ersetzt werden kann, das das Niveau vollständiger ausschöpfen kann.

Ausgehend von dem Konzept der a-priori geordneten Hypothesen wurde zur Bestimmung der maximal sicheren Dosis eine wiederum auf Zweistichprobenverfahren basierende Prozedur zum multiplen Niveau α hergeleitet.

Abschließend wurde das Konzept der Bioäquivalenz dargestellt und anhand eines realen Datensatzes die zusätzliche Anforderung der individuellen Bioäquivalenz illustriert.

In der vorliegenden Arbeit wurden generell stetige Verteilungsfunktionen vorausgesetzt. Häufig liegen jedoch Ordinalskalen vor, anhand derer eine Beurteilung hinsichtlich der Sicherheit erfolgt. Die Herleitung exakter Testverfahren für den Äquivalenznachweis im Falle dichotomer Zufallsvariablen wurde erst in jüngster Zeit erbracht und die Verallgemeinerung dieser Verfahren auf ordinal skalierte Daten wird ein Forschungsschwerpunkt der kommenden Jahre sein.
8 Literaturverzeichnis


Barnes, D. G., Daston, G. P., Evans, J. S., Jarabek, A. M., Kavlock, R. J., Kimmel, C. A., Park, C.,


induced by thyroid hormone is independent of loading conditions and beta-adrenoeceptor blockade.

Berger, R. L. (1982). Multiparameter hypothesis testing and acceptance sampling. Technometrics, 24:
295-300.


Fundamental and Applied Toxicology, 13: 235-244.

Bross, I. D. (1985). Why proof of safety is much more difficult than proof of hazard. Biometrics, 41:
785-793.

Fundamental and Applied Toxicology, 13: 235-244.

Chan, I. S. F. (1998). Exact tests of equivalence and efficacy with a non-zero lower bound for
comparative studies. Statistics in Medicine, 17: 1403-1413.

conventional toxicology studies with rats. Toxicology, 43: 65-74.

Chow, S.-C., Liu, J.-P. Design and Analysis of Bioavailability and Bioequivalence Studies, Marcel

bioequivalence, III/54/89-EN.

Toxicology, 4: 854-871.


Williams, D. A. (1971). A test for differences between treatment means when several dose levels are compared with a zero dose control. *Biometrics, 27*: 103-117.


Anhang: SAS-Programme zur Gütebestimmung

Kursiv geschriebene Zeilen beinhalten z.B. Plausibilitätsprüfungen zu den jeweiligen Funktionsparametern und haben mit den Algorithmen nichts zu tun. Deswegen werden auch Funktionen, die in diesen Zeilen vorzufinden sind, nicht im Listing aufgeführt.

Zeilen wie:

```sas
onerr='if fa00002=0 then do;ok=0;resume;end;';
call push(onerr);
```

werden als Vorbereitung zur Laufzeitfehler-Erkennung benutzt.

Modul PSA_ERR gibt Fehlermeldungen aus, hier die Fehlerliste:

- PSA (-1) Parameter(s) not of appropriate type
- PSA (-2) BVAL not 0 or 1
- PSA (-3) Check parameter(s): No of rows,columns
- PSA (-4) Integer parameter(s) are not integer
- PSA (-5) NF < 2
- PSA (-6) Call to SAS-function PROBT failed
- PSA (-7) R < 0.001
- PSA (-8) Error number -8
- PSA (-9) PS_POW2R: Problems with integration
- SSC (-401) Errorreturn from a NCT-module

```sas
/* SSC.INC
Start    : 28.10.94
Last edit: <13-NOV-1998 08:35:31.88>
Version   : V1.1
Power calculation
*/

start ssc_pow(design     /* Design Crossover, 2-unpaired groups */
   ,model      /* Additive Multiplicative Sasabuchi */
   ,sideval    /* 1-2-Sided */
   ,sidedir    /* Direction if 1-Sided */
   ,alpha      /* Alpha */
   ,n          /* Samplesize */
   ,x          /* X-value (Vector) */
   ,cv         /* Coefficient of variation */
   ,cv2        /* 2nd CV (Sasabuchi Crossover) */
   ,th1        /* Lower limit */
   ,th2        /* Upper limit */
   ,powerval   /* Returned Powervalue */
) global(ps_dp1,ps_dp2,ps_cnt,ps_tech);

run i_psa001(dummy);
```
modnam='SSC_POW';

ps_dp1=0;
ps_dp2=0;

tt=type(design)+type(model)+type(sideval)+type(sidedir)+type(alpha) +type(n)+type(x)+type(cv)+type(th1)+type(th2)+type(powerval);

if tt='NNNNNNNNNNN' then do;
ierr=-1;
run psa_err(modnam,ierr);
return(ierr);
end;

if t1+t2 ^= 20 then do;
ierr=-3;
run psa_err(modnam,ierr);
return(ierr);
end;

powerval = -1;

pp=0;
run i_ssc001(1,design,model,sideval,sidedir,alpha,n,pp,x, cv,cv2,th1,th2,ierr);
if ierr ^= 0 then do;
run psa_err(modnam,ierr);
return(ierr);
end;

nx=nrow(x);
powerval=j(nx,1,-1);
do m=1 to nx;
  if x[m]=. then do;
powerval[m]=.;
end;
if model=3 then do;
  nf=2#n-2;              /* Sasabuchi Crossover and Unpaired 2 groups */
end; else do;
  if design=1 then do;
    nf = n-2;
c=2;                 /* Crossover */
  end; else do;
    nf = n-2;
c=4;                 /* Unpaired two groups */
  end;
end;

if model = 1 then do;
  u = cv # sqrt(c/n);/* Add. */
d1= (x[m]-th1)/u;   /* Left or twosided */
d2= (x[m]-th2)/u;   /* Right or twosided */
end; else do;
  if model=2 then do;
    u = sqrt( log(1+cv#cv)*(c/n) ); /* Mult. */
v = log(x[m]);
d1=(v-log(th1))/u;    /* Left or twosided */
d2 = (v - log(th2))/u; /* Right or twosided */
end; else do;

/* MODEL = 3  SASABUCHI */
if design=1 then do; /* Crossover */
ut1 = (1-th1)#2;
ut2 = (1-th2)#2;
wt1 = cv#cv#(1+th1#th1)+cv2#cv2#ut1;
wt2 = cv#cv#(1+th2#th2)+cv2#cv2#ut2;
u1 = sqrt(wt1/2/n);
u2 = sqrt(wt2/2/n);
rx1 = cv#cv#(1+th1#th2)+cv2#cv2#(1+th1#th2-th1-th2);
corr = rx1/sqrt(wt1#wt2);
end; else do; /* Unpaired two groups */
ut1 = 1+th1#th1;
ut2 = 1+th2#th2;
u1 = cv # sqrt(ut1/n);
u2 = cv # sqrt(ut2/n);
corr = (1+th1#th2)/sqrt(ut1#ut2);
end;
d1 = (x[m]-th1)/u1;
d2 = (x[m]-th2)/u2;
end;
end;
ierr = -401;
kerr = 0;
temp=0;
if sideval = 1 then do;
if sidedir < 0 then do; /* Onesided */
kerr = ps_nct(cv,nf,d1,temp);
if kerr ^=0 then return(ierr);
power = 1 - temp;
end; else do;
kerr = ps_nct(-tv,nf,d2,temp);
if kerr ^=0 then return(ierr);
power = temp;
end;
end; else do;
if model=3 then do; /* SASABUCHI */
part1=0;part2=0;
kerr=ps_pow2r(1e6,-tv,nf,d1,d2,corr,part1);
if kerr ^=0 then return(ierr);
ps_dp1=ps_cnt;
ps_dp2=ps_tech#10;
kerr=ps_pow2r( tv,-tv,nf,d1,d2,corr,part2);
if kerr ^=0 then return(ierr);
ps_dp1=ps_dp1+ps_cnt;
ps_dp2=ps_dp2+ps_tech;
power = part1 - part2;
end; else do;
fakt = sqrt(nf)/(2#tv);
r = (d1-d2) # fakt; /* Twosided */
q1=0;q2=0;
kerr=ps_nctfr(-tv,nf,d2,r,q1);
if kerr ^=0 then return(ierr);
kerr=ps_nctfr( tv,nf,d1,r,q2);
if kerr ^=0 then return(ierr);
power = q1 - q2;
end;
end;
powerval[m]=power;
end;
end;
return(0);
finish;
Noncentral t-distribution

start ps_nct (tv, /* T-value */
    nf, /* Degrees of freedom */
    d, /* Noncentral parameter */
    nct /* nct (tv, nf, d) */)
    global (fa00002);
run i_psa001 (dummy);

modnam='PS_NCT';

if type (tv) + type (nf) + type (d) + type (nct) ^= 'NNNN' then do;
    ierr = -1;
    run psa_err (modnam, ierr);
    return (ierr);
end;

if nrow (tv) + nrow (nf) + nrow (d) + nrow (nct) +
    ncol (tv) + ncol (nf) + ncol (d) + ncol (nct) ^= 8 then do;
    ierr = -3;
    run psa_err (modnam, ierr);
    return (ierr);
end;

if (ceil (nf) - floor (nf))^=0 then do;
    ierr = -4;
    run psa_err (modnam, ierr);
    return (ierr);
end;

nct = -1;

if nf < 2 then do;
    ierr = -5;
    run psa_err (modnam, ierr);
    return (ierr);
end;

ok=1;
onerr='if fa00002=0 then do;ok=0;resume;end;';
call push (onerr);

if d < 0 then do;
    fa00002=0;
    nct = 1 - probt (-tv, nf, -d);
    fa00002=1;
end; else do;
    fa00002=0;
    nct = probt (tv, nf, d);
    fa00002=1;
end;

if ok=0 then do;
    ierr = -6;
    run psa_err (modnam, ierr);
    return (ierr);
end;

return (0);
finish;

start ps_nctfr (tv, /* T-value */
    nf, /* Degrees of freedom */
    d, /* Noncentral parameter */
    r /* Integration-limit */
    ...
Fraction of the UNIVARIATE NON-CENTRAL-t-DISTRIBUTION
(-> Bivariate )
Fraction of Cumulative distribution function $Q(t,d;0,R)$

The integral in the noncentral-t-distribution
\[
\int_0^R \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \, dx
\]
is now reduced to
\[
\int_0^{NCTFR} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \, dx
\]
Formulas from D.B.Owen p.439

```
run i_psa001(dummy);
modnam='PS_NCTFR';
if (ceil(nf)-floor(nf))^=0 then do;
ierr= -4;
run psa_err(modnam,ierr);
return(ierr);
end;
nctfr = -1;
if nf < 2 then do;
ierr= -5;
run psa_err(modnam,ierr);
return(ierr);
end;
if r < 1e-3 then do;
ierr= -7;
run psa_err(modnam,ierr);
return(ierr);
end;
sq2phi= 2.506628274631; /* SQRT(2*PHI) */
a=tv/sqrt(nf);
b=nf/(nf+tv#tv);
sqb=sqrt(b);
par1=d#a#sqb;
par2=(d#a#b-r)/sqb;
fak= 0.3989422804014; /* 1/sqrt(2*PHI) */
x=d#sqb;  dens =fak#exp(-x#x/2); /* Density of normal distribution */
x=par1;  dens1 =fak#exp(-x#x/2);
x=par2;  dens2 =fak#exp(-x#x/2);
x=r;    densr =fak#exp(-x#x/2);
```
x=a#r-d;  densa=fak#exp(-x#x/2);

mk0=a#sqb#dens#(probnorm(par1)-probnorm(par2)); /* M0 */
hk0=-densr#probnorm(a#r-d); /* H0 */
meven=mk0 + hk0; /* Even sum */

if nf>=3 then do;
  mk=b#(d#a#mk0+a#dens#(dens1-dens2)); /* M1 */
hk=r#hk0; /* A3 = 1 */ /* H1 */
  modd =mk + hk; /* Odd sum */
end;

if nf>=4 then do;
  /* Usage of scalars, because they are faster as matrices and we must
   use (sure?) a loop to calculate our result!
   At K=2 (Loopbegin) ak2 : A4   (Loopend) ak : A3   ak1 : A4
   At K=3 (Loopbegin) ak2 : A5 = 1/3/A4 = 1/3/ak1
   General : ak2 = 1/k/ak1   with ak2==A[k+2] , ak1==A[k+1]
   
   At K=2 L1 = L[k-1]
   At K=3 L2 = L[k-1] = ak2#r#L[k-2]
   */
  lv=a#b#r#densr#densa/2.;
  ak =1;
  ak1=1;
  iflag=1;
  do k=2 to nf-2;
    ak2=1/k/ak1;
    if k>2 then lv = ak2#r#lv;
    mk1=(k-1)/k # b # (ak#d#a#mk + mk0) - lv;
    hk1=ak2#r#hk;
    if iflag=1 then do;
      meven=meven+mk1+hk1;
      end; else do;
        modd =modd +mk1+hk1;
        end;
    iflag = -iflag;
    ak =ak1;  ak1=ak2;
    mk0=mk;
    mk =mk1;
    hk =hk1;
    end;
  end;
  if int(mod(nf,2)) = 0 then do;
    nctfr=probnorm(-d)+sq2phi#meven;
    end; else do;
    tfunc1=0;tfunc2=0;tfunc3=0;
    idummy=ps_t_fct(d#sqb,a ,tfunc1);
    idummy=ps_t_fct(r,(a#r-d)/r,tfunc2);
    if abs(d) < 1e-10 then do;
      dval = 1e-10;
      if d<0 then dval = -dval;
      end; else dval=d;
    aval= a - r/b/dval;
    idummy=ps_t_fct(d#sqb,aval,tfunc3);
    nctfr=probnorm(r) + 2#(tfunc1-tfunc2-tfunc3) + 2#modd;
    if d >= 0 then nctfr= nctfr-1;
    end;
  return(0);
end;

start ps_t_fct(h  /* H-value */
/* A-value */
/* T-function returnvalue */

-- T-function, see OWEN "Tables for computing bivariate normal probabilities"
Annals of Mathematical Statistics, 27 (1956)
Page 1075..1090

run i_psa001(dummy);

modnam='PS_T_FCT';

if type(h)+type(a)+type(tfunc) ^= 'NNN' then do;
  ierr= -1;
  run psa_err(modnam,ierr);
  return(ierr);
end;
if nrow(h)+nrow(a)+nrow(tfunc)+
   ncol(h)+ncol(a)+ncol(tfunc) ^= 6 then do;
  ierr= -3;
  run psa_err(modnam,ierr);
  return(ierr);
end;

phi=3.14159265359;

-- Integral for A= 0..1

if a < 0 then do;
aa=-a;
ineg=1;
end; else do;
aa=a;
ineg=0;
end;
if aa > 1 then do;
  mode=2;
  gl=probnorm(abs(h));  /* If A<0 , use T(h,-a) = -T(h,a) */
  g2=probnorm(abs(aa#h));
  hh=aa#h;
aa=1/aa;
end; else do;
  mode=1;
  hh=h;
end;
if abs(hh) > 6 then do;
  tf=0;  /* Convergence to zero */
end; else do;
  hh=hh##2;
  hf=exp(-hh/2);
  sf=sf#hh/2/jx;    /* i=j : H^(2*i)/(2^i*i!) */
  sumf=sumf+sf;    /* SUM i=0 */
  vfak=1;          /* (-1)^0 */
  sum=(1-hf)#aa;   /* c0*A^(2*0+1) */
  jj=1;
c=1;
do while(abs(c)>1e-10);
  sf=sf#hh/2/jx;    /* i=j : H^(2*i)/2^i/i! */
  sumf=sumf+sf;    /* Sum i=0..j */
  vfak=-vfak;      /* (-1)^j */
  ind = 2*jx+1;
c=vfak##(1-hf#sumf)/ind # aa##ind;  /* Cj * a^(2j+1) */
  sum=sum+c;
  jj=jj+1;
end;
tf=(atan(aa)-sum)/2/phi;
end;

if mode = 1 then tfunc=tf; else tfunc=func-tf;
if ineg = 1 then tfunc = -tfunc;

return(0);
finish;

start ps_pow2r(t1, t2, nf, d1, d2, r, pow2r)
/ * Part of 1st integration limit to PROBBNRM */
, t1 /* Part of 2nd integration limit to PROBBNRM */
, nf /* Degrees of freedom */
, d1 /* 1st noncentrality value */
, d2 /* 2nd noncentrality value */
, r /* Correlation */
, pow2r) /* Result: POWER */
global(ps_cnt, ps_tech, ps_zt1, ps_zt2, ps_znf, ps_zd1, ps_zd2, ps_zr);

run i_psa001(dummy);

modnam='PS_POW2R';
pow2r = -1; /* Returnvalue if error */
ps_cnt=0;
ps_tech=0;

if type(t1)+type(t2)+type(nf)+type(d1)+type(d2)+type(r) ^= 'NNNNNN' then do;
ierr= -1;
run psa_err(modnam,ierr);
return(ierr);
end;
if nrow(t1)+nrow(t2)+nrow(nf)+nrow(d1)+nrow(d2)+nrow(r) ^= 6 then do;
ierr= -3;
run psa_err(modnam,ierr);
return(ierr);
end;
if (ceil(nf)-floor(nf))^=0 then do;
ierr= -4;
run psa_err(modnam,ierr);
return(ierr);
end;
if nf < 2  then do;
ierr= -5;
run psa_err(modnam,ierr);
return(ierr);
end;
ierr=0; /* Errorreturn of function */
xmax = sqrt(nf-1); /* Approximative maximum */

/* 1st try with QUAD */
ps_zt1=t1;
ps_zt2=t2;
pz_znf=nf;
pz_zd1=d1;
pz_zd2=d2;
pz_zr =r;

call quad(pret,"PS_FPW2Q",{1e-6 .p}) eps=1e-9 peak=xmax scale=0.1 msg="NO";

if pret^=. then do;
pow2r=pret;
ps_tech=1; /* Technique used */
return(ierr);
end;
/* QUAD has problems, calculate integral with own technique */

ps_tech=2;               /* Technique used now */

delta=0.1;
if 2#xmax/delta<200 then delta=xmax/100;

flag=0;
count=0;
do until(flag=1);
  count=count+1;
  xvec= do(1e-5,2#xmax,delta);
  yvec= ps_fpw2r(xvec,t1,t2,nf,d1,d2,r);
  nxy=ncol(xvec);
  ymaxind=loc(yvec=max(yvec));
  ymaxind=ymaxind[1];             /* If more than 1 maximum */
  if ymaxind<0.1#nxy | ymaxind>0.9#nxy then do;
    if count>=2 then do;
      if sum(yvec)#delta < 1e-15 then do;    /* Practically no area */
        pow2r=sum(yvec)#delta;               /* 'pseudo' trapezoidal */
        ierr=-9;
        run psa_err(modnam,ierr);
      end;
    end;
    flag=1;
    end;
  end;
  ymax=yvec[ymaxind];
  ycrit = ymax/1e8;
  xrange=loc(yvec>ycrit);    /* Get range of Y > YCRIT */
  goon=1;
  if type(xrange)="U" then do;
    goon=0;
  end; else do;
    xindlow=xrange[1];
    xindup=xrange[ncol(xrange)];
    if (xindup-xindlow)<5 then goon=0;
  end;
  if goon=0 then do;
    ierr=-9;                             /* Could not calculate */
    run psa_err(modnam,ierr);
    return(ierr);
  end;
  xbeg=xvec[xindlow];
  xend=xvec[xindup];
  if xindup>=(nxy-1) then do;
    flag=0;
    count=0;
    do until(flag=1);
      count=count+1;
      if count>=10 then do;
        ierr=-9;
        run psa_err(modnam,ierr);
        return(ierr);
      end;
      xend=xend#1.2;
      ylast= ps_fpw2r(xend,t1,t2,nf,d1,d2,r);
      if ylast<ycrit then flag=1;
    end;
  end;
ninter=1000; /* Must be EVEN */
delta=(xend-xbeg)/ninter;
xvec=do(xbeg,xend,delta);
yvec= ps_fpw2r(xvec,t1,t2,nf,d1,d2,r);
ny=ncol(xvec);

/* Simpsons formula

\[
\int_0^N \left[ y_0 + y_n + 4 \left( y_1 + y_3 + \ldots + y_{n-1} \right) + 2 \left( y_2 + y_4 + \ldots + y_{n-2} \right) \right] \, dx
\]

N even integer: NINTER == N NXY = NINTER + 1

*/

iodd=t(do(1,ninter-1,2)); /* odd index Start 1 == y1 == YVEC[2] */
ieven=t(do(2,ninter-2,2)); /* even index Start 2 == y2 == YVEC[3] */

psum =yvec[1]+yvec[ninter+1]+4#sum(yvec[iodd+1])+2#sum(yvec[ieven+1]);
pow2r=psum#delta/3;
return(ierr);
finish;

start ps_fpw2r(xvec,t1,t2,nf,d1,d2,r);
g1=(xvec#t1/sqrt(nf)-d1);
g2=(xvec#t2/sqrt(nf)-d2);
qwer=probbrnm(g1,g2,r); /* 1 row NX columns */

lfak=(nf-1)#log(xvec) - (nf/2-1)#log(2) - lgamma(nf/2) - xvec#xvec/2;
yvec=qwer#exp(lfak);
return(yvec);
finish;

/* Interface module to QUAD */

start ps_fpw2q(xvec) global(ps_cnt,ps_zt1,ps_zt2,ps_znf,ps_zd1,ps_zd2,ps_zr);
ps_cnt=ps_cnt+1;
yvec=ps_fpw2r(xvec,ps_zt1,ps_zt2,ps_znf,ps_zd1,ps_zd2,ps_zr);
return(yvec);
finish;

/*
About PS_FPW2R

Basic procedure received from Frank Bretz (Uni Hannover) (Feb 1998)

\[
\int_0^\infty \frac{f-1}{\Gamma(f/2) \sqrt{2\pi}} \exp\left(\frac{-x^2}{2}ight) \, dx
\]

** Interface module to QUAD **

\[
BI(a,b) = \frac{1}{2^*\pi^*\sqrt{1-r^*r}} \frac{a^b}{2^*(1-r^*r)} \int_{-\infty}^{-\infty} dv du
\]

Due to numeric problems we take the logarithm :
LFAK = (f-1)*LOG(x) - (f/2-1)*LOG(2) - LGAMMA(f/2) - x*x/2

(LGAMMA: Log of Gammafunction which is usable with great f)

----------------------------------------------------------------------

To know the maximum is a great help. Let us determine an approximate maximum.

Consider only following part:

\[
y = x^a \cdot e^{-x^2/2}
\]

\[
y' = x^{a-1} \cdot e^{-x^2/2} + a \cdot x \cdot e^{-x^2/2}
\]

\[
y' = 0 \rightarrow x = a^x \rightarrow x = a \rightarrow x = \sqrt{a} = \sqrt{f-1}
\]