

# Ranked Set Sampling: Praktische Probleme und deren Erkennung

**Diplomarbeit**

der Philosophisch-naturwissenschaftlichen Fakultät  
der Universität Bern

vorgelegt von  
Michael Vock

2001

Leiter der Arbeit:

Prof. Dr. Jürg Hüsler,  
Institut für mathematische Statistik und  
Versicherungslehre der Universität Bern

# Inhaltsverzeichnis

<b>Vorwort</b>	<b>3</b>
<b>1 Ordnungsstatistiken</b>	<b>4</b>
1.1 Verteilungs- und Dichtefunktion . . . . .	4
1.2 Spezialfall: Uniformverteilung . . . . .	5
<b>2 Mittelwert-Schätzung</b>	<b>7</b>
2.1 Methode nach McIntyre . . . . .	7
2.2 Bezeichnungen . . . . .	8
2.3 Varianzverkleinerung gegenüber einer zufälligen Stichprobe . . . . .	8
2.4 Obere und untere Schranke für die relative Präzision . . . . .	10
2.5 Werte der relativen Präzision für bekannte Verteilungen . . . . .	15
<b>3 Vorzeichentest</b>	<b>18</b>
3.1 RSS-Vorzeichen-Teststatistik . . . . .	18
3.2 Erwartungswert, Varianz . . . . .	18
3.3 Verhalten unter $H_0 : \theta = 0$ . . . . .	21
3.4 Berechnung von $\delta_{0,m}^2$ . . . . .	21
3.5 Ablehnungsbereich von $H_0 : \theta = 0$ . . . . .	24
3.6 Vergleich mit dem SRS-Vorzeichentest: Pitman-Effizienz . . . . .	24
<b>4 Praktische Probleme</b>	<b>30</b>
4.1 Rangierungsfehler . . . . .	30
4.2 Abhängigkeiten innerhalb der Sets . . . . .	34
<b>5 Kostenüberlegungen</b>	<b>37</b>
<b>6 Erkennungsmöglichkeiten für die praktischen Probleme</b>	<b>39</b>
6.1 Allgemeine Überlegungen . . . . .	39
6.2 Grafische Methode . . . . .	40

<b>7</b>	<b>Statistik <math>R_{m,n}(\mathbf{X})</math></b>	<b>43</b>
7.1	Verhalten unter der Nullhypothese . . . . .	44
7.2	Verhalten unter der speziellen Alternative . . . . .	48
7.3	Minimax-Test . . . . .	54
<b>8</b>	<b>Statistik <math>\tilde{R}_{m,n}(\mathbf{X})</math></b>	<b>57</b>
8.1	Verhalten unter der Nullhypothese . . . . .	57
8.2	Verhalten unter der speziellen Alternative . . . . .	58
8.3	Minimax-Test . . . . .	61
<b>9</b>	<b>Statistik <math>R_{m,n}^*(\mathbf{X})</math></b>	<b>63</b>
9.1	Verhalten unter der Nullhypothese . . . . .	63
9.2	Verhalten unter der speziellen Alternative . . . . .	66
9.3	Exkurs: Jonckheere-Test . . . . .	69
9.4	Minimax-Test . . . . .	71
<b>10</b>	<b>Ausblick</b>	<b>73</b>
	<b>Zusammenfassung</b>	<b>74</b>
	<b>Literaturverzeichnis</b>	<b>75</b>

# Vorwort

Es gibt Situationen, in denen eine ungefähre Rangierung einer kleinen Stichprobe nach der interessierenden Grösse leicht möglich ist (etwa von blossem Auge), die genaue Bestimmung eines Werts („Messung“) aber mit grossem Aufwand verbunden ist. In solchen Fällen kann die Anwendung von Ranked Set Sampling sinnvoll sein. Dabei handelt es sich im wesentlichen um ein Auswahlverfahren, das auf rangierten Stichproben (eben den „Ranked Sets“) beruht. Die Methode wurde durch McIntyre (1952) vorgestellt und seither von zahlreichen Autoren untersucht und weiterentwickelt. Dabei wurden auch Probleme aufgezeigt, die auftreten, wenn gewisse Annahmen nicht erfüllt sind, und die die erhofften Kosteneinsparungen gegenüber herkömmlichen Verfahren zunichte machen können.

In dieser Arbeit wird Ranked Set Sampling anhand zweier Anwendungen aus der Literatur vorgestellt, und die erwähnten Probleme werden erörtert. Anschliessend wird nach Möglichkeiten gesucht, wie die Bedeutung dieser Probleme nach Anwendung des Auswahlverfahrens alleine mit dem vorliegenden Zahlenmaterial beurteilt werden kann.

In den letzten Jahren wurden auch einige Abwandlungen des Auswahlverfahrens untersucht, die nur bestimmte Ordnungsstatistiken verwenden und dadurch genauere Schätzer bzw. effizientere Tests ermöglichen (vgl. z.B. Barabesi 2001). Solche Verfahren dürften aber Rückschlüsse von den ausgewählten Elementen auf die ursprünglichen Daten erschweren, wie sie in dieser Arbeit vorgesehen sind. Deshalb wird hier nur das Auswahlverfahren in seiner ursprünglichen Form betrachtet.

Ich möchte es an dieser Stelle nicht versäumen, Herrn Professor Hüsler sowie meinen Bürokollegen Marco Collenberg, Samuel Müller und Kaspar Rufibach für ihre fachliche und menschliche Unterstützung zu danken.

Bern, im November 2001

Michael Vock

# Kapitel 1

## Ordnungsstatistiken

Da die Ranked-Set-Sampling-Methode auf Ordnungsstatistiken beruht, sind zunächst einige Aussagen über deren Eigenschaften nötig.

### 1.1 Verteilungs- und Dichtefunktion

Seien  $X_1, \dots, X_n$  i.i.d. mit Verteilung  $F$ . Dann ist die Verteilung  $F_{k:n}$  der  $k$ -ten Ordnungsstatistik  $X_{k:n}$

$$\begin{aligned} F_{k:n}(x) &= P(X_{k:n} \leq x) \\ &= P(\#\{j : X_j \leq x\} \geq k) \\ &= \sum_{i=k}^n \binom{n}{i} [F(x)]^i [1 - F(x)]^{n-i}. \end{aligned} \quad (1.1)$$

Man kann leicht folgende Eigenschaft zeigen:

$$\begin{aligned} \sum_{k=1}^n F_{k:n}(x) &= \sum_{k=1}^n \sum_{i=k}^n \binom{n}{i} [F(x)]^i [1 - F(x)]^{n-i} \\ &= \sum_{i=1}^n i \binom{n}{i} [F(x)]^i [1 - F(x)]^{n-i} \\ &= \sum_{j=0}^{n-1} \frac{n(n-1)!}{j!(n-1-j)!} [F(x)]^{j+1} [1 - F(x)]^{n-1-j} \\ &= nF(x) [F(x) + 1 - F(x)]^{n-1} \\ &= nF(x) \end{aligned} \quad (1.2)$$

Daraus folgt insbesondere

$$\sum_{k=1}^n E[X_{k:n}^r] = \sum_{k=1}^n \int_{-\infty}^{\infty} x^r dF_{k:n}(x)$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} x^r d \left[ \sum_{k=1}^n F_{k:n}(x) \right] \\
&= \int_{-\infty}^{\infty} x^r d [nF(x)] \\
&= n \int_{-\infty}^{\infty} x^r dF(x) \\
&= n E [X_1^r], \tag{1.3}
\end{aligned}$$

sofern das  $r$ -te Moment existiert.

Besitzt  $F$  eine Dichtefunktion  $f$ , so kann auch für  $X_{k:n}$  eine Dichtefunktion  $f_{k:n}$  berechnet werden:

$$\begin{aligned}
f_{k:n}(x) &= \frac{d}{dx} F_{k:n}(x) \\
&= \sum_{i=k}^n \binom{n}{i} i [F(x)]^{i-1} f(x) [1 - F(x)]^{n-i} \\
&\quad - \sum_{i=k}^{n-1} \binom{n}{i} [F(x)]^i (n-i) [1 - F(x)]^{n-i-1} f(x) \\
&= \frac{n!}{(k-1)!(n-k)!} [F(x)]^{k-1} f(x) [1 - F(x)]^{n-k} \\
&\quad + \sum_{i=k+1}^n \frac{n!}{(i-1)!(n-i)!} [F(x)]^{i-1} f(x) [1 - F(x)]^{n-i} \\
&\quad - \sum_{j=k+1}^n \frac{n!}{(j-1)!(n-j)!} [F(x)]^{j-1} f(x) [1 - F(x)]^{n-j} \\
&= \frac{n!}{(k-1)!(n-k)!} [F(x)]^{k-1} f(x) [1 - F(x)]^{n-k} \tag{1.4}
\end{aligned}$$

## 1.2 Spezialfall: Uniformverteilung

Im Fall der Uniformverteilung auf dem Intervall  $[0, 1]$  vereinfacht sich (1.4): Für  $0 \leq x \leq 1$  gilt

$$f_{k:n}(x) = \frac{n!}{(k-1)!(n-k)!} x^{k-1} (1-x)^{n-k}.$$

Es folgt

$$\begin{aligned}
E [X_{k:n}] &= \int_0^1 x f_{k:n}(x) dx \\
&= \int_0^1 \frac{n!}{(k-1)!(n-k)!} x^k (1-x)^{n-k} dx
\end{aligned}$$

$$\begin{aligned}
&= \frac{k}{n+1} \int_0^1 \underbrace{\frac{(n+1)!}{k!(n-k)!} x^k (1-x)^{n-k}}_{f_{k+1:n+1}(x) \text{ der Uniformverteilung}} dx \\
&= \frac{k}{n+1}, \\
\mathbb{E}[X_{k:n}^2] &= \int_0^1 x^2 f_{k:n}(x) dx \\
&= \int_0^1 \frac{n!}{(k-1)!(n-k)!} x^{k+1} (1-x)^{n-k} dx \\
&= \frac{k(k+1)}{(n+1)(n+2)} \int_0^1 \underbrace{\frac{(n+2)!}{(k+1)!(n-k)!} x^{k+1} (1-x)^{n-k}}_{f_{k+2:n+2}(x) \text{ der Uniformverteilung}} dx \\
&= \frac{k(k+1)}{(n+1)(n+2)},
\end{aligned}$$

und somit

$$\begin{aligned}
\text{Var}[X_{k:n}] &= \frac{k(k+1)}{(n+1)(n+2)} - \frac{k^2}{(n+1)^2} \\
&= \frac{k(n+1-k)}{(n+1)^2(n+2)}. \tag{1.5}
\end{aligned}$$

## Kapitel 2

# Mittelwert-Schätzung

### 2.1 Methode nach McIntyre

Ranked Set Sampling (RSS) geht auf McIntyre (1952) zurück. Dieser suchte eine möglichst genaue und dennoch nicht zu aufwendige Methode zur Bestimmung des Ertrags von Acker- und Weideland, wobei er folgende Überlegung anstellte: Ein Experte kann den Ertrag mehrerer gleich grosser Quadrate innerhalb eines Feldes von blossen Auge mit einiger Sicherheit vergleichen und diese Quadrate nach Ertrag rangieren. Wesentlich aufwendiger ist es, den genauen Ertrag eines Quadrates zu bestimmen. McIntyre versuchte deshalb, teure, genaue Messungen durch billigere Rangierungen zu „ersetzen“. Zu diesem Zweck schlug er sinngemäss folgendes Vorgehen vor:

Es werden  $m \geq 2$  Stichproben (Sets) mit je  $m$  Elementen gezogen. Jedes Set wird für sich rangiert. Das kleinste Element des ersten Sets wird ausgewählt und genau gemessen, ebenso das zweitkleinste Element des zweiten Sets usw., schliesslich das grösste Element des  $m$ -ten Sets. Damit sind  $m$  Elemente gemessen, und der erste Durchgang ist abgeschlossen. Es folgen  $n - 1$  weitere solche Durchgänge, so dass schliesslich  $mn$  Elemente gemessen sind (von insgesamt  $m^2n$  betrachteten Elementen).

Berechnet man nun das arithmetische Mittel der  $mn$  gemessenen Elemente, so erhält man einen Mittelwert-Schätzer, der eine kleinere Varianz hat als das arithmetische Mittel von  $mn$  zufällig ausgewählten Elementen (vgl. übernächsten Abschnitt). Umgekehrt ausgedrückt spart man gegenüber dem klassischen Mittelwert-Schätzer mit gleicher Varianz Messungen ein, muss aber dafür einen gewissen Aufwand für die Rangierungen in Kauf nehmen. Die Methode ist somit überall dort sinnvoll anwendbar, wo die Einsparungen bei den Messungen den Mehraufwand für die Rangierungen übersteigen.



## 2.2 Bezeichnungen

Zur Bezeichnung der betrachteten Elemente sind Indices für den Durchgang, für das Set innerhalb eines Durchgangs sowie für das Element innerhalb eines Sets nötig. Sei  $X_{ijk}$  das  $j$ -te Element des  $i$ -ten Sets im  $k$ -ten Durchgang. Weiter bezeichne  $X_{i(j)k}$  die  $j$ -te Ordnungsstatistik im  $i$ -ten Set des  $k$ -ten Durchgangs, und  $X_{i(i)k}$  werde kurz auch als  $X_{(i)k}$  bezeichnet. Dann lässt sich das Auswahlverfahren wie in Abbildung 2.1 darstellen.

## 2.3 Varianzverkleinerung gegenüber einer zufälligen Stichprobe

Die Verkleinerung der Varianz gegenüber dem Mittelwert einer zufälligen Stichprobe mit gleichvielen gemessenen Elementen soll nun genauer untersucht werden. Diese Varianzverkleinerung behandeln beispielsweise Dell und Clutter (1972) ausführlich.

Sei

$$\hat{\mu}_{m,n}^{\text{RSS}} = \frac{1}{mn} \sum_{k=1}^n \sum_{i=1}^m X_{(i)k}$$

der Mittelwertschätzer aus der Ranked-Set-Sampling-Prozedur wie oben beschrieben.  $\hat{\mu}_{m,n}^{\text{RSS}}$  ist erwartungstreu:

$$\begin{aligned} \text{E} [\hat{\mu}_{m,n}^{\text{RSS}}] &= \frac{1}{mn} \sum_{k=1}^n \sum_{i=1}^m \text{E} [X_{(i)k}] \\ &\stackrel{(1.3)}{=} \frac{1}{mn} \sum_{k=1}^n m \text{E} [X_{111}] \\ &= \text{E} [X_{111}] \end{aligned}$$

Die  $X_{(i)k}$  sind unabhängig, also gilt

$$\begin{aligned} \text{Var} [\hat{\mu}_{m,n}^{\text{RSS}}] &= \frac{1}{m^2 n^2} \sum_{k=1}^n \sum_{i=1}^m \text{Var} [X_{(i)k}] \\ &= \frac{1}{m^2 n} \sum_{i=1}^m \sigma_{i:m}^2, \end{aligned} \tag{2.1}$$

wobei  $\sigma_{i:m}^2$  die Varianz der  $i$ -ten Ordnungsstatistik aus einer Stichprobe des Umfangs  $m$  ist und natürlich von der zugrundeliegenden Verteilung abhängt.

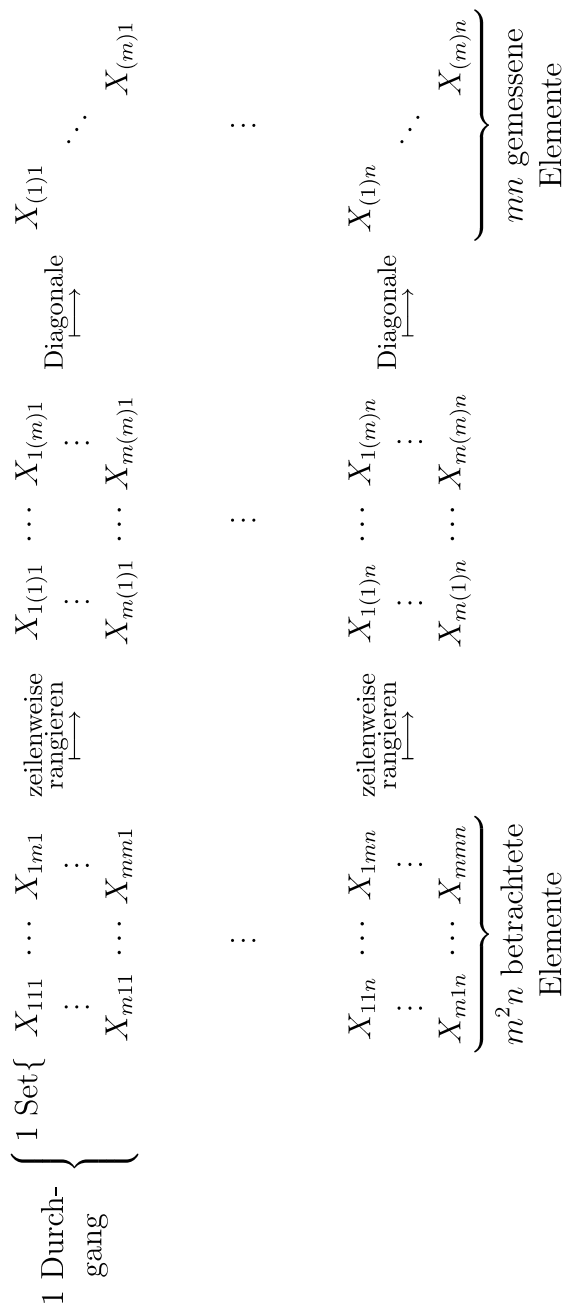


Abbildung 2.1: RSS-Auswahlverfahren/Bezeichnungen

Für den klassischen (Simple-Random-Sampling-)Mittelwertschätzer, der auf einer zufälligen Stichprobe  $(X_1, \dots, X_N)$  beruht, schreiben wir

$$\hat{\mu}_N^{\text{SRS}} = \frac{1}{N} \sum_{i=1}^N X_i.$$

Seine Varianz ist bekanntlich

$$\begin{aligned} \text{Var} [\hat{\mu}_N^{\text{SRS}}] &= \frac{1}{N^2} \sum_{i=1}^N \text{Var} [X_i] \\ &= \frac{\sigma^2}{N}, \end{aligned} \tag{2.2}$$

wobei  $\sigma^2$  die Varianz eines einzelnen Stichprobenelements bezeichnet.

Die Genauigkeit des RSS-Schätzers im Vergleich zum SRS-Schätzer mit gleichvielen ( $mn$ ) gemessenen Elementen lässt sich mit dem Verhältnis der Varianzen angeben:

$$\frac{\text{Var} [\hat{\mu}_{mn}^{\text{SRS}}]}{\text{Var} [\hat{\mu}_{m,n}^{\text{RSS}}]} = \frac{m\sigma^2}{\sum_{i=1}^m \sigma_{i:m}^2} \tag{2.3}$$

Dieses Verhältnis wird als relative Präzision (RP) bezeichnet.

## 2.4 Obere und untere Schranke für die relative Präzision

**Beispiel 2.1** Sind die  $X_{ijk}$  uniformverteilt auf  $[0, 1]$ , dann ist

$$\begin{aligned} \sum_{i=1}^m \sigma_{i:m}^2 &\stackrel{(1.5)}{=} \sum_{i=1}^m \frac{i(m+1-i)}{(m+1)^2(m+2)} \\ &= \frac{(m+1) \sum_{i=1}^m i - \sum_{i=1}^m i^2}{(m+1)^2(m+2)} \\ &= \frac{3m^2 + 3m - 2m^2 - m}{6(m+1)(m+2)} \\ &= \frac{m}{6(m+1)}, \end{aligned}$$

die relative Präzision ist also in diesem Fall

$$\frac{\text{Var} [\hat{\mu}_{mn}^{\text{SRS}}]}{\text{Var} [\hat{\mu}_{m,n}^{\text{RSS}}]} = \frac{\frac{m}{12}}{\frac{m}{6(m+1)}} = \frac{m+1}{2}.$$

Takahasi und Wakimoto (1968) zeigen, dass (2.3) für stetige Verteilungen mit endlicher Varianz kleiner oder gleich  $\frac{m+1}{2}$  ist. Der Wert für die Uniformverteilung ist also gerade eine obere Schranke.

Von grösserer praktischer Bedeutung ist eine untere Schranke für (2.3) – schliesslich möchte man eine gewisse Entschädigung für den Rangieraufwand garantiert haben. Da  $\sigma_{i:m}^2 \leq \sigma^2$ , ist 1 eine untere Schranke für die relative Präzision. Die RSS-Methode ist also mindestens so genau wie die SRS-Methode mit gleichvielen gemessenen Elementen. Eine bessere untere Schranke lässt sich leider nicht angeben, wie die folgenden Beispiele zeigen.

**Beispiel 2.2** *Wir konstruieren eine Folge von Verteilungen, für die die relative Präzision gegen 1 strebt. Es seien  $X_1^{(n)}, \dots, X_m^{(n)}$  i.i.d. Zufallsvariablen mit*

$$\begin{aligned} \mathrm{P}\left(X_i^{(n)} = 0\right) &= 1 - \frac{1}{n} \text{ und} \\ \mathrm{P}\left(X_i^{(n)} = 1\right) &= \frac{1}{n}. \end{aligned}$$

*Da diese Zufallsvariablen nur die Werte 0 und 1 annehmen, gilt*

$$\begin{aligned} \mathrm{E}\left[\left(X_i^{(n)}\right)^k\right] &= \frac{1}{n}, \quad k = 1, 2, 3, \dots, \text{ und somit} \\ \mathrm{Var}\left[X_i^{(n)}\right] &= \frac{1}{n} - \frac{1}{n^2} \\ &= \frac{n-1}{n^2}. \end{aligned}$$

*Für diese Stichprobe des Umfangs  $m$  betrachten wir nun die Ordnungsstatistiken  $X_{i:m}^{(n)}$  und insbesondere das Maximum  $X_{m:m}^{(n)}$  näher. (Die diskrete Verteilung verursacht hier keine Probleme: Im allgemeinen sind zwar die Ränge der einzelnen Beobachtungen nicht eindeutig definiert, wohl aber die Werte der Ordnungsstatistiken.)*

$$\begin{aligned} \mathrm{P}\left(X_{m:m}^{(n)} = 1\right) &= 1 - \mathrm{P}\left(X_{m:m}^{(n)} = 0\right) \\ &= 1 - \mathrm{P}\left(X_i^{(n)} = 0, i = 1, \dots, m\right) \\ &= 1 - \left(1 - \frac{1}{n}\right)^m \\ \mathrm{E}\left[\left(X_{m:m}^{(n)}\right)^k\right] &= 1 - \left(1 - \frac{1}{n}\right)^m, \quad k = 1, 2, 3, \dots \\ \mathrm{Var}\left[X_{m:m}^{(n)}\right] &= 1 - \left(1 - \frac{1}{n}\right)^m - \left[1 - \left(1 - \frac{1}{n}\right)^m\right]^2 \end{aligned}$$

$$\begin{aligned}
&= \left(1 - \frac{1}{n}\right)^m - \left(1 - \frac{1}{n}\right)^{2m} \\
&= \frac{n^m(n-1)^m - (n-1)^{2m}}{n^{2m}} \\
&= \frac{1}{n^{2m}} \left[ n^m \sum_{j=0}^m \binom{m}{j} n^j (-1)^{m-j} - \sum_{j=0}^{2m} \binom{2m}{j} n^j (-1)^{2m-j} \right] \\
&= \frac{1}{n^{2m}} \left[ n^{2m} - mn^{2m-1} - n^{2m} + 2mn^{2m-1} + O(n^{2m-2}) \right] \\
&= \frac{mn^{2m-1} + O(n^{2m-2})}{n^{2m}}
\end{aligned}$$

Wir erhalten als Abschätzung für (2.3):

$$\begin{aligned}
\frac{m \operatorname{Var} [X_i^{(n)}]}{\sum_{i=1}^m \operatorname{Var} [X_{i:m}^{(n)}]} &\leq \frac{m \operatorname{Var} [X_i^{(n)}]}{\operatorname{Var} [X_{m:m}^{(n)}]} \\
&= \frac{m \frac{n-1}{n^2}}{\frac{mn^{2m-1} + O(n^{2m-2})}{n^{2m}}} \\
&= \frac{mn^{2m-2}(n-1)}{mn^{2m-1} + O(n^{2m-2})} \\
&= \frac{n^{2m-1} - n^{2m-2}}{n^{2m-1} + O(n^{2m-2})} \\
&\xrightarrow{n \rightarrow \infty} 1
\end{aligned}$$

Da die relative Präzision aber mindestens 1 beträgt, gilt in diesem Beispiel tatsächlich

$$\frac{m \operatorname{Var} [X_i^{(n)}]}{\sum_{i=1}^m \operatorname{Var} [X_{i:m}^{(n)}]} \xrightarrow{n \rightarrow \infty} 1.$$

Damit ist 1 die grösste untere Schranke für die relative Präzision.

Auch durch eine Beschränkung auf Verteilungen mit Dichte erhält man keine grössere untere Schranke. Die Konstruktion einer entsprechenden Folge und die Berechnung von (2.3) werden allerdings aufwendiger. Auch hierfür ein Beispiel:

**Beispiel 2.3** Wir konstruieren nun eine Folge von stetigen Verteilungen, für die die relative Präzision ebenfalls gegen 1 strebt. Es seien  $X_1^{(n)}, \dots, X_m^{(n)}$

*i.i.d. Zufallsvariablen mit Verteilung*

$$F^{(n)}(x) = \begin{cases} 0 & x < -\frac{1}{n} \\ \left(x + \frac{1}{n}\right) (n-1) & -\frac{1}{n} \leq x < 0 \\ \frac{n-1+x}{n} & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

*und dazugehöriger Dichte*

$$f^{(n)}(x) = \begin{cases} n-1 & -\frac{1}{n} \leq x < 0 \\ \frac{1}{n} & 0 \leq x < 1 \\ 0 & \text{sonst.} \end{cases}$$

Für  $n = 1$  entspricht dies der Uniformverteilung.

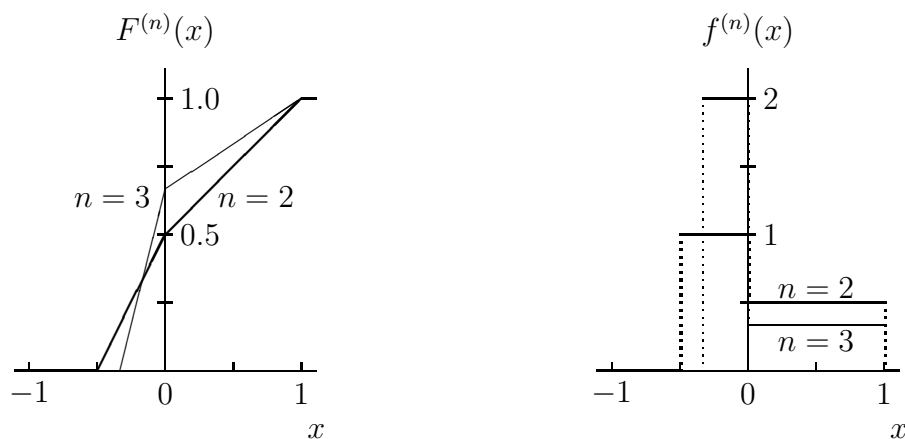


Abbildung 2.2: Verteilungs- und Dichtefunktionen zu Beispiel 2.3 (für  $n = 2, 3$ )

*Allgemein gilt*

$$\begin{aligned} \mathbb{E} [X_i^{(n)}] &= \int_{-\frac{1}{n}}^0 x(n-1) dx + \int_0^1 x \frac{1}{n} dx \\ &= \frac{(n-1)x^2}{2} \Big|_{-\frac{1}{n}}^0 + \frac{x^2}{2n} \Big|_0^1 \\ &= -\frac{n-1}{2n^2} + \frac{1}{2n} \\ &= \frac{1}{2n^2}, \end{aligned}$$

analog auch

$$\mathbb{E} \left[ \left( X_i^{(n)} \right)^2 \right] = \frac{n^2 + n - 1}{3n^3}$$

und somit

$$\begin{aligned} \text{Var} \left[ X_i^{(n)} \right] &= \frac{n^2 + n - 1}{3n^3} - \frac{1}{4n^4} \\ &= \frac{1}{3n} + O\left(\frac{1}{n^2}\right). \end{aligned}$$

Für diese Stichprobe des Umfangs  $m$  betrachten wir wiederum das Maximum  $X_{m:m}^{(n)}$  näher.

$$\begin{aligned} f_{m:m}^{(n)}(x) &\stackrel{(1.4)}{=} \begin{cases} \frac{m!}{(m-1)!0!} \left[ \left( x + \frac{1}{n} \right) (n-1) \right]^{m-1} (n-1) & -\frac{1}{n} \leq x < 0 \\ \frac{m!}{(m-1)!0!} \left[ \frac{n-1+x}{n} \right]^{m-1} \frac{1}{n} & 0 \leq x < 1 \\ 0 & \text{sonst} \end{cases} \\ &= \begin{cases} m(n-1)^m \left( x + \frac{1}{n} \right)^{m-1} & -\frac{1}{n} \leq x < 0 \\ \frac{m}{n^m} (n-1+x)^{m-1} & 0 \leq x < 1 \\ 0 & \text{sonst} \end{cases} \\ \mathbb{E} \left[ X_{m:m}^{(n)} \right] &= \int_{-\frac{1}{n}}^0 x m (n-1)^m \left( x + \frac{1}{n} \right)^{m-1} dx \\ &\quad + \int_0^1 x \frac{m}{n^m} (n-1+x)^{m-1} dx \\ &= m(n-1)^m \sum_{j=0}^{m-1} \binom{m-1}{j} \frac{x^{j+2}}{(j+2)n^{m-1-j}} \Big|_{-\frac{1}{n}}^0 \\ &\quad + \frac{m}{n^m} \sum_{j=0}^{m-1} \binom{m-1}{j} (n-1)^j \frac{x^{m-j+1}}{m-j+1} \Big|_0^1 \\ &= \frac{m(n-1)^m}{n^{m+1}} \sum_{j=0}^{m-1} \binom{m-1}{j} \frac{(-1)^{j+1}}{j+2} \\ &\quad + \frac{m}{n^m} \sum_{j=0}^{m-1} \binom{m-1}{j} \frac{(n-1)^j}{m-j+1} \\ &= O\left(\frac{1}{n}\right) \\ \mathbb{E} \left[ \left( X_{m:m}^{(n)} \right)^2 \right] &= \frac{m(n-1)^m}{n^{m+2}} \sum_{j=0}^{m-1} \binom{m-1}{j} \frac{(-1)^j}{j+3} \\ &\quad + \frac{m}{n^m} \sum_{j=0}^{m-1} \binom{m-1}{j} \frac{(n-1)^j}{m-j+2} \end{aligned}$$

$$\begin{aligned}
&= \frac{m}{3n} + O\left(\frac{1}{n^2}\right) \\
\text{Var} [X_{m:m}^{(n)}] &= \frac{m}{3n} + O\left(\frac{1}{n^2}\right) - \left[O\left(\frac{1}{n}\right)\right]^2 \\
&= \frac{m}{3n} + O\left(\frac{1}{n^2}\right)
\end{aligned}$$

Wir erhalten als Abschätzung für (2.3):

$$\begin{aligned}
\frac{m \text{Var} [X_i^{(n)}]}{\sum_{i=1}^m \text{Var} [X_{i:m}^{(n)}]} &\leq \frac{m \text{Var} [X_i^{(n)}]}{\text{Var} [X_{m:m}^{(n)}]} \\
&= \frac{m \left[ \frac{1}{3n} + O\left(\frac{1}{n^2}\right) \right]}{\frac{m}{3n} + O\left(\frac{1}{n^2}\right)} \\
&\xrightarrow{n \rightarrow \infty} 1
\end{aligned}$$

In diesem Beispiel gilt also ebenfalls

$$\frac{m \text{Var} [X_i^{(n)}]}{\sum_{i=1}^m \text{Var} [X_{i:m}^{(n)}]} \xrightarrow{n \rightarrow \infty} 1.$$

Damit ist 1 auch für stetige Verteilungen die grösste untere Schranke für die relative Präzision.

## 2.5 Werte der relativen Präzision für bekannte Verteilungen

Dell und Clutter (1972) geben Werte der relativen Präzision für zahlreiche Verteilungen an. Eine Auswahl davon ist in Tabelle 2.1 wiedergegeben.

Dell und Clutter weisen auf einen recht starken Zusammenhang zwischen

$$\alpha_4 = \frac{\text{E}[(X - \text{E}(X))^4]}{\sigma^4}$$

und der relativen Präzision für ein gegebenes  $m$  hin. Dieser Zusammenhang ist in der grafischen Darstellung (Abbildung 2.3) gut sichtbar – für die Uniformverteilung ( $\alpha_4 = 1.8$ ) ergibt sich die maximale relative Präzision, für kleinere bzw. grössere Werte von  $\alpha_4$  nimmt sie ab.

Allerdings ist auch in den Beispielen von Dell und Clutter bereits erkennbar, dass die relative Präzision noch von weiteren Faktoren abhängt –



die Gammaverteilung mit Parameter  $\alpha = 2$  sowie die doppelt exponentielle Verteilung haben den gleichen Wert  $\alpha_4$ , aber verschiedene relative Präzision. Noch deutlicher wird dieses Phänomen, wenn man etwa die Werte für die oben konstruierte Folge von stetigen Verteilungen (Beispiel 2.3) mit  $m = 2$  betrachtet: Der Fall  $n = 1$  entspricht der Uniformverteilung, die übrigen Werte weichen aber erheblich von denjenigen der „klassischen“ Verteilungen ab.

Tabelle 2.1: Relative Präzision von  $\hat{\mu}_{m,n}^{\text{RSS}}$  gegenüber  $\hat{\mu}_{mn}^{\text{SRS}}$ ; aus Dell und Clutter (1972).  $\mu$ : Mittelwert,  $\sigma^2$ : Varianz,  $\alpha_4$ :  $E[(X - E(X))^4]/\sigma^4$ . Ein zufällig entdeckter Fehler beim Wert  $\alpha_4$  der U-förmigen Verteilung (Dichte  $f(x) = \frac{3x^2}{2}$  für  $-1 < x < 1$ ) wurde korrigiert.

Verteilung	$\mu$	$\sigma^2$	$\alpha_4$	RP für $m =$			
				2	3	4	5
Uniform [0,1]	0.500	0.0833	1.80	1.500	2.000	2.500	3.000
Exponential	1.000	1.0000	9.00	1.333	1.636	1.920	2.190
Gamma (1.5, 1)	1.500	1.5000	7.00	1.370	1.710	2.030	2.336
Gamma (2.0, 1)	2.000	2.0000	6.00	1.391	1.753	2.096	2.424
Gamma (2.5, 1)	2.500	2.5000	5.40	1.405	1.781	2.139	2.487
Gamma (3.0, 1)	3.000	3.0000	5.00	1.414	1.801	2.169	2.524
Gamma (3.5, 1)	3.500	3.5000	4.71	1.421	1.816	2.192	2.556
Gamma (4.0, 1)	4.000	4.0000	4.50	1.427	1.827	2.210	2.575
Gamma (5.0, 1)	5.000	5.0000	4.20	1.434	1.843	2.236	2.615
Normal	0.000	1.0000	3.00	1.467	1.914	2.347	2.770
Beta (4, 4)	0.500	0.0278	2.45	1.484	1.958	2.425	2.890
Beta (7, 4)	0.636	0.0193	2.70	1.475	1.936	2.389	2.834
Beta (13, 4)	0.765	0.0100	3.14	1.460	1.903	2.333	2.765
U-förmig	0.000	0.6000	1.19	1.441	1.849	2.234	2.603
Chi-Quadrat (1)	0.798	0.3634	3.87	1.430	1.841	2.239	
Doppelt exponentiell	0.000	2.0000	6.00	1.381	1.730	2.038	2.327

Relative Präzision

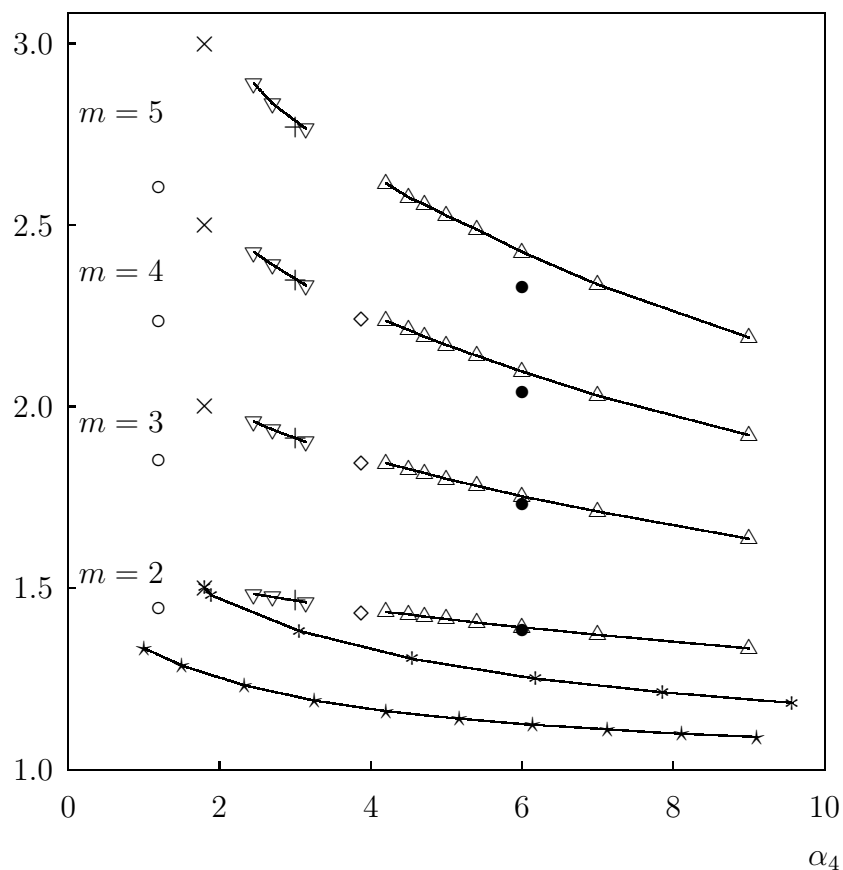


Abbildung 2.3: Relative Präzision für  $m = 2, 3, 4, 5$  in Abhängigkeit von  $\alpha_4$  gemäss Tabelle 2.1, wobei die Verteilungen bzw. Verteilungsfamilien folgendermassen gekennzeichnet sind:  $\circ$  U-förmig,  $\times$  Uniformverteilung,  $\nabla$  Beta-Verteilungen,  $+$  Normalverteilung,  $\triangle$  Gammaverteilungen (inkl. Exponentialverteilung),  $\bullet$  doppelt exponentielle Verteilung,  $\diamond$  Chi-Quadrat-Verteilung (nur für  $m = 2, 3, 4$ ). Ausserdem sind für  $m = 2$  die entsprechenden Werte für die zuvor konstruierten Beispiele eingezeichnet:  $\star$  diskretes Beispiel 2.2 für  $n = 2, \dots, 11$ ,  $\ast$  stetiges Beispiel 2.3 für  $n = 1, \dots, 7$ .

# Kapitel 3

## Vorzeichentest

In den ersten Beschreibungen des RSS-Verfahrens steht eindeutig die Schätzung des Mittelwerts im Vordergrund. Zwar untersucht bereits McIntyre (1952) Anwendungsmöglichkeiten zur Schätzung von höheren Momenten, kommt aber zum Schluss, dass die Varianzverkleinerung dort bescheiden ist.

Selbstverständlich lassen sich aber auf eine mittels Ranked Set Sampling gewonnene Auswahl von Stichprobenelementen verschiedenste statistische Verfahren anwenden. Als weiteres Beispiel wird im folgenden der RSS-Vorzeichentest nach Hettmansperger (1995) vorgestellt.

### 3.1 RSS-Vorzeichen-Teststatistik

Wir nehmen an, dass die Stichprobenelemente  $X_{ijk}$  die Verteilung  $G(x) = F(x - \theta)$  haben, wobei  $\theta$  der (unbekannte) Median ist: Es gelte  $F(0) = \frac{1}{2}$ . Die Statistik für den RSS-Vorzeichentest ist

$$S_{m,n}^{+ \text{RSS}} = \sum_{k=1}^n \sum_{i=1}^m \mathbb{I}(X_{(i)k} > 0),$$

also die Anzahl positiver Messungen ( $\mathbb{I}(\cdot)$  sei die Indikatorfunktion).

### 3.2 Erwartungswert, Varianz

**Satz 3.1** *Für die oben definierte RSS-Vorzeichen-Teststatistik gilt*

$$\mathbb{E} [S_{m,n}^{+ \text{RSS}}] = mn [1 - G(0)], \quad (3.1)$$

$$\text{Var} [S_{m,n}^{+ \text{RSS}}] = mn G(0) [1 - G(0)] \left[ 1 - \frac{\sum_{i=1}^m [G_{i:m}(0) - G(0)]^2}{m G(0) [1 - G(0)]} \right]. \quad (3.2)$$

**Beweis:** Da die  $X_{(i)k}$  für ein festes  $i$  i.i.d. sind (mit Verteilung  $G_{i:m}$ ), ist folgende Zerlegung von  $S_{m,n}^{+RSS}$  zweckmässig:

$$S_{m,n}^{+RSS} = \sum_{i=1}^m S_{i:m,n}^{+RSS}$$

mit

$$S_{i:m,n}^{+RSS} = \sum_{k=1}^n \mathbf{I}(X_{(i)k} > 0).$$

$S_{i:m,n}^{+RSS}$  ist binomialverteilt mit Parametern  $n$  und  $1 - G_{i:m}(0)$ , und für  $i = 1, \dots, m$  sind die  $S_{i:m,n}^{+RSS}$  unabhängig. Folglich gilt

$$\begin{aligned} \mathbb{E} [S_{m,n}^{+RSS}] &= \sum_{i=1}^m \mathbb{E} [S_{i:m,n}^{+RSS}] \\ &= n \sum_{i=1}^m [1 - G_{i:m}(0)] \\ &\stackrel{(1.2)}{=} mn [1 - G(0)] \end{aligned}$$

und

$$\begin{aligned} \text{Var} [S_{m,n}^{+RSS}] &= \sum_{i=1}^m \text{Var} [S_{i:m,n}^{+RSS}] \\ &= n \sum_{i=1}^m [1 - G_{i:m}(0)] G_{i:m}(0) \\ &\stackrel{(1.2)}{=} mn G(0) - n \sum_{i=1}^m [G_{i:m}(0)]^2. \end{aligned}$$

Mit

$$\begin{aligned} n \sum_{i=1}^m [G_{i:m}(0)]^2 &\stackrel{(1.2)}{=} n \sum_{i=1}^m [(G_{i:m}(0))^2 - 2G_{i:m}(0)G(0) + (G(0))^2] \\ &\quad + mn(G(0))^2 \\ &= n \sum_{i=1}^m [G_{i:m}(0) - G(0)]^2 + mn(G(0))^2 \end{aligned}$$

erhält man

$$\begin{aligned} \text{Var} [S_{m,n}^{+RSS}] &= mn G(0) - n \sum_{i=1}^m [G_{i:m}(0) - G(0)]^2 - mn(G(0))^2 \\ &= mn G(0) [1 - G(0)] \left[ 1 - \frac{\sum_{i=1}^m [G_{i:m}(0) - G(0)]^2}{m G(0) [1 - G(0)]} \right], \end{aligned}$$

was den Beweis vervollständigt.

Der letzte Faktor im Ausdruck für die Varianz ist kleiner oder gleich 1. In dieser Form lässt sich leicht eine Verkleinerung der Varianz gegenüber der gewöhnlichen (SRS-)Vorzeichen-Teststatistik

$$S_{mn}^{+\text{SRS}} = \sum_{i=1}^{mn} \mathbf{I}(X_i > 0)$$

(basierend auf einer Stichprobe  $(X_1, \dots, X_{mn})$ ) zeigen:

$$\begin{aligned} \text{Var} [S_{mn}^{+\text{SRS}}] &= \sum_{i=1}^{mn} \text{Var} [\mathbf{I}(X_i > 0)] \\ &= mn G(0) [1 - G(0)] \\ &\geq \text{Var} [S_{m,n}^{+\text{RSS}}] \end{aligned}$$

Wir setzen

$$\delta_m^2 = \frac{\text{Var} [S_{m,n}^{+\text{RSS}}]}{\text{Var} [S_{mn}^{+\text{SRS}}]} = 1 - \frac{\sum_{i=1}^m [G_{i:m}(0) - G(0)]^2}{m G(0) [1 - G(0)]}$$

(Das Verhältnis zweier Varianzen ist nichtnegativ, was die Bezeichnung  $\delta_m^2$  rechtfertigt.)

**Satz 3.2** Mit  $\delta_m^2$  wie oben gilt

$$\frac{S_{m,n}^{+\text{RSS}} - mn [1 - G(0)]}{\sqrt{mn G(0) [1 - G(0)] \delta_m^2}} \xrightarrow{d} Z \quad (n \rightarrow \infty), \quad (3.3)$$

wobei  $Z$  eine standardnormalverteilte Zufallsvariable sei.

**Beweis:**  $S_{m,n}^{+\text{RSS}}$  ist die Summe von  $n$  i.i.d.-Summanden  $\sum_{i=1}^m \mathbf{I}(X_{(i)k} > 0)$ . Somit kann der zentrale Grenzwertsatz angewendet werden, und die Konvergenz von

$$\frac{S_{m,n}^{+\text{RSS}} - mn [1 - G(0)]}{\sqrt{mn G(0) [1 - G(0)] \delta_m^2}} = \frac{S_{m,n}^{+\text{RSS}} - \mathbf{E} [S_{m,n}^{+\text{RSS}}]}{\sqrt{\text{Var} [S_{m,n}^{+\text{RSS}}]}}$$

folgt unmittelbar.

### 3.3 Verhalten unter $H_0 : \theta = 0$

Wir gehen nun von der Nullhypothese aus, dass der Median  $\theta = 0$  sei. Dann ist  $G(0) = F(0) = \frac{1}{2}$ . Weiter gilt

$$\begin{aligned} G_{i:m}(0) &\stackrel{(1.1)}{=} \sum_{j=i}^m \binom{m}{j} [F(0)]^j [1 - F(0)]^{m-j} \\ &= \frac{1}{2^m} \sum_{j=i}^m \binom{m}{j}, \end{aligned} \quad (3.4)$$

$$E_0 [S_{m,n}^{+RSS}] \stackrel{(3.1)}{=} \frac{mn}{2}, \quad (3.5)$$

$$\text{Var}_0 [S_{m,n}^{+RSS}] \stackrel{(3.2)}{=} \frac{mn}{4} \underbrace{\left[ 1 - \frac{4}{m} \sum_{i=1}^m \left( G_{i:m}(0) - \frac{1}{2} \right)^2 \right]}_{=: \delta_{0,m}^2}. \quad (3.6)$$

Somit vereinfacht sich (3.3) zu

$$\frac{S_{m,n}^{+RSS} - \frac{mn}{2}}{\sqrt{\frac{mn}{4} \delta_{0,m}^2}} \xrightarrow{d} Z \quad (n \rightarrow \infty). \quad (3.7)$$

### 3.4 Berechnung von $\delta_{0,m}^2$

Die Werte von  $\delta_{0,m}^2$  lassen sich für kleine  $m$  (und nur solche werden realistisch benötigt) unter Verwendung von (3.4) relativ leicht berechnen:

$$\delta_{0,m}^2 = 1 - \frac{4}{m} \sum_{i=1}^m \left( \frac{1}{2^m} \sum_{j=i}^m \binom{m}{j} - \frac{1}{2} \right)^2 \quad (3.8)$$

Dennoch lohnt sich der Aufwand, eine einfachere Form herzuleiten.

**Satz 3.3** Für  $m = 2, 3, 4, \dots$  gilt

$$\delta_{0,m}^2 = \prod_{\ell=2}^m \frac{2\ell - 1}{2\ell}. \quad (3.9)$$

Somit lassen sich die Werte rekursiv berechnen:

$$\delta_{0,m+1}^2 = \frac{2m + 1}{2m + 2} \delta_{0,m}^2 \quad (3.10)$$

**Beweis:** Induktion über  $m$ . Die Induktionsverankerung ( $m = 2$ ) ist problemlos:

$$\begin{aligned} \delta_{0,2}^2 &\stackrel{(3.8)}{=} 1 - \frac{4}{2} \left[ \left( \frac{\binom{2}{1} + \binom{2}{2}}{4} - \frac{1}{2} \right)^2 + \left( \frac{\binom{2}{2}}{4} - \frac{1}{2} \right)^2 \right] \\ &= 1 - 2 \left[ \left( \frac{3}{4} - \frac{1}{2} \right)^2 + \left( \frac{1}{4} - \frac{1}{2} \right)^2 \right] \\ &= \frac{3}{4} \end{aligned}$$

Für den Induktionsschritt nehmen wir an, (3.9) gelte für ein festes  $m$ . Können wir daraus (3.10) zeigen, so gilt (3.9) auch für  $m + 1$  statt  $m$ .

Wir benötigen folgende Aussagen über Binomialkoeffizienten:

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1} \quad (3.11)$$

$$\sum_{k=0}^n \binom{n}{k} = 2^n \quad (3.12)$$

$$\sum_{k=0}^n \binom{n}{k}^2 = \binom{2n}{n} \quad (3.13)$$

(3.11) und (3.12) sind allgemein bekannt. (3.13) ist eine Konsequenz der Vandermonde-Faltung

$$\sum_{k=0}^c \binom{a}{k} \binom{b}{c-k} = \binom{a+b}{c} \quad (0 \leq c \leq \min\{a, b\}),$$

die man sich mit einer kombinatorischen Überlegung veranschaulicht (Wahl von  $c$  aus  $a + b$  Elementen). Auf einen formellen Beweis wird hier verzichtet.

Wir definieren nun

$$S_m = \sum_{i=1}^m \left( \frac{1}{2^m} \sum_{j=i}^m \binom{m}{j} - \frac{1}{2} \right)^2.$$

Mit den Beziehungen (3.11) bis (3.13) erhalten wir

$$\begin{aligned} S_{m+1} &= \sum_{i=1}^{m+1} \left( \frac{1}{2^{m+1}} \sum_{j=i}^{m+1} \binom{m+1}{j} - \frac{1}{2} \right)^2 \\ &= \sum_{i=1}^m \left( \frac{1}{2^{m+1}} \left[ \sum_{j=i}^m \left( \binom{m}{j-1} + \binom{m}{j} \right) + 1 \right] - \frac{1}{2} \right)^2 + \left( \frac{1}{2^{m+1}} - \frac{1}{2} \right)^2 \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^m \left( \frac{1}{2^m} \sum_{j=i}^m \binom{m}{j} + \frac{1}{2^{m+1}} \binom{m}{i-1} - \frac{1}{2} \right)^2 + \left( \frac{1}{2^{m+1}} - \frac{1}{2} \right)^2 \\
&= S_m + 2 \sum_{i=1}^m \left( \frac{1}{2^m} \sum_{j=i}^m \binom{m}{j} - \frac{1}{2} \right) \left( \frac{1}{2^{m+1}} \binom{m}{i-1} \right) \\
&\quad + \sum_{i=1}^m \left( \frac{1}{2^{m+1}} \binom{m}{i-1} \right)^2 + \frac{1}{2^{2m+2}} - \frac{1}{2^{m+1}} + \frac{1}{4} \\
&= S_m + \frac{1}{2^{2m}} \sum_{i=1}^m \sum_{j=i}^m \binom{m}{i-1} \binom{m}{j} - \frac{1}{2^{m+1}} \sum_{i=1}^m \binom{m}{i-1} \\
&\quad + \frac{1}{2^{2m+2}} \sum_{i=1}^m \binom{m}{i-1}^2 + \frac{1}{2^{2m+2}} - \frac{1}{2^{m+1}} + \frac{1}{4} \\
&= S_m + \frac{1}{2^{2m+1}} \left[ \sum_{0 \leq i, j \leq m} \binom{m}{i} \binom{m}{j} - \sum_{i=0}^m \binom{m}{i}^2 \right] - \frac{2^m - 1}{2^{m+1}} \\
&\quad + \frac{\binom{2m}{m} - 1}{2^{2m+2}} + \frac{1}{2^{2m+2}} - \frac{1}{2^{m+1}} + \frac{1}{4} \\
&= S_m + \frac{1}{2^{2m+1}} \left[ \sum_{i=0}^m \binom{m}{i} \right]^2 - \frac{\binom{2m}{m}}{2^{2m+1}} - \frac{1}{2} + \frac{\binom{2m}{m}}{2^{2m+2}} + \frac{1}{4} \\
&= S_m + \frac{1}{2} - \frac{\binom{2m}{m}}{2^{2m+2}} - \frac{1}{2} + \frac{1}{4} \\
&\stackrel{(3.8)}{=} \frac{m}{4} (1 - \delta_{0,m}^2) - \frac{1}{4} \prod_{\ell=1}^m \frac{2\ell(2\ell-1)}{(2\ell)^2} + \frac{1}{4} \\
&\stackrel{\text{Ind.ann.}}{=} \frac{m}{4} - \frac{m}{4} \delta_{0,m}^2 - \frac{1}{8} \delta_{0,m}^2 + \frac{1}{4} \\
&= \frac{m+1}{4} - \frac{2m+1}{8} \delta_{0,m}^2
\end{aligned}$$

und somit

$$\begin{aligned}
\delta_{0,m+1}^2 &\stackrel{(3.8)}{=} 1 - \frac{4}{m+1} S_{m+1} \\
&= 1 - 1 + \frac{2m+1}{2m+2} \delta_{0,m}^2 \\
&= \frac{2m+1}{2m+2} \delta_{0,m}^2,
\end{aligned}$$

also (3.10). Damit ist der Induktionsschritt gezeigt und Satz 3.3 bewiesen.

Unter Verwendung von Satz 3.3 erhält man unmittelbar die Werte in Tabelle 3.1.



Tabelle 3.1: Werte von  $\delta_{0,m}^2$  für  $m = 2, \dots, 5$ .

$m$	$\delta_{0,m}^2$
2	0.75
3	0.625
4	0.5469
5	0.4922

### 3.5 Ablehnungsbereich von $H_0 : \theta = 0$

Aufgrund von (3.7) ist

$$\left[ \frac{mn}{2} - \frac{1}{2} \sqrt{mn \delta_{0,m}^2} z_{1-\alpha/2}, \frac{mn}{2} + \frac{1}{2} \sqrt{mn \delta_{0,m}^2} z_{1-\alpha/2} \right]$$

ein asymptotischer  $(1-\alpha)$ -Toleranzbereich für  $S_{m,n}^{+RSS}$ , wobei  $z_p$  das  $p$ -Quantil der Standardnormalverteilung ist. Man verwirft folglich  $H_0 : \theta = 0$  zugunsten von  $H_1 : \theta \neq 0$ , wenn

$$\left| S_{m,n}^{+RSS} - \frac{mn}{2} \right| > \frac{1}{2} \sqrt{mn \delta_{0,m}^2} z_{1-\alpha/2}.$$

Dies ergibt einen zweiseitigen Test zum asymptotischen Niveau  $\alpha$ . Analog erhält man einseitige Tests zum asymptotischen Niveau  $\alpha$ .  $H_0 : \theta = 0$  wird zum Beispiel zugunsten von  $H_1 : \theta > 0$  verworfen, wenn

$$S_{m,n}^{+RSS} - \frac{mn}{2} > \frac{1}{2} \sqrt{mn \delta_{0,m}^2} z_{1-\alpha}.$$

### 3.6 Vergleich mit dem SRS-Vorzeichentest: Pitman-Effizienz

Für den Vergleich des RSS-Vorzeichentests mit dem gewöhnlichen Vorzeichentest brauchen wir einige allgemeine Bezeichnungen. Wir beschränken uns hier auf einseitige Alternativen.

Zum Vergleich von verschiedenen Tests  $T^{(1)}, T^{(2)}$  der gleichen Hypothesen  $H_0 : \theta = \theta_0$  gegen  $H_1 : \theta > \theta_0$  bietet sich zunächst die *relative Effizienz von  $T^{(1)}$  gegenüber  $T^{(2)}$*  an. Diese bezeichnet das Verhältnis  $\frac{N^{(2)}}{N^{(1)}}$ , wobei  $N^{(i)}$  derjenige Stichprobenumfang ist, den der Test  $T^{(i)}$  benötigt, um das Niveau  $\alpha$  und die Güte  $1 - \beta$  zu einer bestimmten Alternative  $\theta_1$  zu erreichen.

Um die Abhängigkeit von  $\theta_1$ ,  $\alpha$  und  $\beta$  zu vermeiden, verwendet man folgende asymptotische Betrachtungsweise: Es seien  $\{T_k^{(1)}\}$ ,  $\{T_k^{(2)}\}$  Folgen von Teststatistiken, wobei  $T_k^{(i)}$  auf einem Stichprobenumfang  $N_k^{(i)}$  beruhen soll. Wenn die zugehörigen Niveaus  $\{\alpha_k^{(i)}\}$  für beide Folgen von Teststatistiken gegen einen gemeinsamen Grenzwert  $\alpha$  konvergieren und wenn bezüglich einer Folge  $\{\theta_k\}$  die Güten  $\{1 - \beta_k^{(i)}\}$  beider Folgen gegen  $1 - \beta$  konvergieren, dann bezeichnet man

$$e_{T^{(1)}, T^{(2)}} = \lim_{k \rightarrow \infty} \frac{N_k^{(2)}}{N_k^{(1)}}$$

als *asymptotische relative Effizienz* oder *Pitman-Effizienz* von  $T^{(1)}$  gegenüber  $T^{(2)}$ , sofern das Verhältnis nicht mehr von  $\alpha$ ,  $\beta$  und  $\{\theta_k\}$  abhängt.

Der Beweis für die Existenz der Pitman-Effizienz und die Berechnung ihres Werts lassen sich in vielen Fällen mit den beiden folgenden Sätzen bewerkstelligen:

**Satz 3.4**  $\{T_k^{(1)}\}$ ,  $\{T_k^{(2)}\}$  und die zugehörigen Stichprobenumfänge  $\{N_k^{(i)}\}$  seien wie oben, und es gelte  $N_k^{(i)} \rightarrow \infty$  ( $k \rightarrow \infty$ ) für  $i = 1, 2$ . Die Verteilung von  $T_k^{(i)}$  sei durch den reellen Parameter  $\theta$  bestimmt.

Es gebe für  $i = 1, 2$  eine Funktion  $\mu^{(i)}(\theta)$  sowie eine Folge von Funktionen  $\{\sigma_k^{(i)}(\theta)\}$ , die folgende Bedingungen erfüllen:

(R1)  $\mu^{(i)}(\theta)$  ist differenzierbar bei  $\theta = \theta_0$ , mit Ableitung  $\mu^{(i)'(\theta_0)} \neq 0$ .

(R2) Die  $\sigma_k^{(i)}(\theta_0)$  sind von der Ordnung  $\frac{1}{\sqrt{N_k^{(i)}}}$ .

Der Test  $T_k^{(i)}$  verwerfe die Hypothese  $H_0 : \theta = \theta_0$  zugunsten von  $H_1 : \theta > \theta_0$ , falls

$$\frac{T_k^{(i)} - \mu^{(i)}(\theta_0)}{\sigma_k^{(i)}(\theta_0)} \geq C_k^{(i)}.$$

Für jede Folge  $\{\theta_k\}$  mit  $\theta_k \rightarrow \theta_0$  ( $k \rightarrow \infty$ ) und für  $i = 1, 2$  gelte ausserdem:

(R3)  $\frac{T_k^{(i)} - \mu^{(i)}(\theta_k)}{\sigma_k^{(i)}(\theta_k)} \xrightarrow{d} Z \sim N(0, 1)$  ( $k \rightarrow \infty$ )

(R4)  $\frac{\sigma_k^{(i)}(\theta_k)}{\sigma_k^{(i)}(\theta_0)} \rightarrow 1$  ( $k \rightarrow \infty$ )

Die Niveaus der Testfolgen  $\{T_k^{(i)}\}$  sollen gegen ein gemeinsames Niveau  $\alpha$  konvergieren, und die Güten bezüglich einer bestimmten Folge  $\{\theta_k\}$  von Alternativen sollen gegen den gemeinsamen Wert  $1 - \beta$  konvergieren ( $\alpha < 1 - \beta < 1$ ).

Dann existiert die Pitman-Effizienz von  $T^{(1)}$  gegenüber  $T^{(2)}$ , und es gilt:

$$e_{T^{(1)}, T^{(2)}} = \left( \frac{c^{(1)}}{c^{(2)}} \right)^2,$$

wobei

$$c^{(i)} = \lim_{k \rightarrow \infty} \frac{\mu^{(i)'}(\theta_0)}{\sqrt{N_k^{(i)}} \sigma_k^{(i)}(\theta_0)}$$

die sogenannten Effizienzen sind.

**Satz 3.5** Sei  $i = 1$  oder  $i = 2$ .  $\{T_k^{(i)}\}$  sei eine Folge von Tests mit zugehörigen Stichprobenumfängen  $N_k^{(i)}$  wie oben, die die Voraussetzungen (R1) und (R2) von Satz 3.4 erfüllt. Für jede Folge von Alternativen der Form

$$\theta_k = \theta_0 + \frac{d^{(i)}}{\sqrt{N_k^{(i)}}} + o\left(\frac{1}{\sqrt{N_k^{(i)}}}\right) \quad (d^{(i)} > 0)$$

seien ausserdem (R3) und (R4) erfüllt. Dann gilt

$$P_{\theta_k} \left( \frac{T_k^{(i)} - \mu^{(i)}(\theta_0)}{\sigma_k^{(i)}(\theta_0)} \geq C_k^{(i)} \right) \rightarrow \Phi \left( c^{(i)} d^{(i)} - z_{1-\alpha} \right) \quad (k \rightarrow \infty),$$

wobei wiederum  $c^{(i)}$  die Effizienz der Folge  $\{T_k^{(i)}\}$  und  $C_k^{(i)} \rightarrow z_{1-\alpha}$  ( $k \rightarrow \infty$ ) die kritischen Werte sind ( $\Phi$  sei die Verteilungsfunktion von  $N(0, 1)$ ).

**Beweis der Sätze 3.4 und 3.5:** vgl. Lehmann (1975), pp. 371-375.

Wir zeigen nun, dass der im Abschnitt 3.5 erwähnte einseitige RSS-Vorzeichentest die Regularitätsbedingungen (R1) bis (R4) erfüllt. Dabei wollen wir annehmen, dass die zugrundeliegende Verteilung  $F$  der Stichprobenelemente  $X_{ijk}$  eine Dichte  $f$  habe und dass  $f(0)$  positiv sei. Als Teststatistiken betrachten wir (für ein festes  $m$ )

$$T_n^{\text{RSS}} = \frac{S_{m,n}^{+\text{RSS}}}{n},$$

die auf Stichprobenumfängen  $N_n^{\text{RSS}} = N_n = mn$  beruhen und natürlich zu  $S_{m,n}^{+\text{RSS}}$  äquivalent sind. In unserem Fall ist  $\theta_0 = 0$ .

Wir setzen

$$\begin{aligned}\mu^{\text{RSS}}(\theta) = \mu(\theta) &= \mathbb{E}_\theta [T_n^{\text{RSS}}] \\ &= m [1 - F(-\theta)], \\ \sigma_n^{\text{RSS}}(\theta) = \sigma_n(\theta) &= \sqrt{\text{Var}_\theta [T_n^{\text{RSS}}]} \\ &= \sqrt{\frac{m}{n} F(-\theta) [1 - F(-\theta)] \delta_m^2}.\end{aligned}$$

Aufgrund der getroffenen Annahmen über  $F$  sieht man unmittelbar, dass (R1) erfüllt ist, mit

$$\mu'(0) = mf(-\theta)|_{\theta=0} = mf(0) > 0.$$

Weiter gilt

$$\sigma_n(0) = \sqrt{\frac{m\delta_{0,m}^2}{4n}} = \frac{m\delta_{0,m}}{2} \cdot \frac{1}{\sqrt{mn}},$$

womit auch (R2) erfüllt ist.

$H_0$  wird verworfen, falls

$$\begin{aligned}S_{m,n}^{+\text{RSS}} - \frac{mn}{2} &> \frac{1}{2} \sqrt{mn\delta_{0,m}^2} z_{1-\alpha} \\ \Leftrightarrow \frac{S_{m,n}^{+\text{RSS}}}{n} - \frac{m}{2} &> \frac{1}{2} \sqrt{\frac{m}{n}} \delta_{0,m}^2 z_{1-\alpha} \\ \Leftrightarrow \frac{T_n^{\text{RSS}} - \mu(0)}{\sigma_n(0)} &> z_{1-\alpha} \quad (\equiv C_n^{\text{RSS}}).\end{aligned}$$

Sei  $\{\theta_n\}$  eine Folge von Alternativen mit  $\theta_n \rightarrow 0$  ( $n \rightarrow \infty$ ). Die Verifikation von (R3) ist nicht ganz so trivial, wie man auf den ersten Blick meinen könnte: Zwar gilt wegen (3.3) unter einem festen  $\theta_k$

$$\frac{T_n^{\text{RSS}} - \mu(\theta_k)}{\sigma_n(\theta_k)} \xrightarrow{n \rightarrow \infty} Z \sim N(0, 1). \quad (3.14)$$

Wir müssen aber zeigen, dass diese Konvergenz auch dann noch erhalten bleibt, wenn wir  $k = n$  setzen und mit  $n$  gegen  $\infty$  gehen lassen. Dazu betrachten wir  $S_{m,n}^{+\text{RSS}}$  als Summe von  $n$  i.i.d. Zufallsvariablen mit Verteilung wie  $S_{m,1}^{+\text{RSS}}$ , also mit Erwartungswert  $\mu(\theta_k)$  und Varianz  $\sigma_1^2(\theta_k)$ . Wir wollen annehmen, dass  $\sigma_1(\theta)$  für  $\theta$  aus einer Umgebung  $U$  von 0 grösser oder gleich einem positiven  $\varepsilon$  sei. Man sieht ohne weiteres, dass

$$\mathbb{E} \left[ \left| S_{m,1}^{+\text{RSS}} - \mathbb{E} (S_{m,1}^{+\text{RSS}}) \right|^3 \right] \leq m^3.$$

Der Satz von Berry-Esséen (vgl. z. B. Hettmansperger (1984), p. 305) liefert somit

$$\left| \mathbb{P} \left( \sqrt{n} \frac{T_n^{\text{RSS}} - \mu(\theta_k)}{\sigma_1(\theta_k)} \leq t \right) - \Phi(t) \right| \leq \frac{cm^3}{\sqrt{n} \sigma_1^3(\theta_k)} \leq \frac{cm^3}{\sqrt{n} \varepsilon^3}$$

für  $\theta_k \in U$ , wobei  $c$  unabhängig von  $k$  und kleiner als 1 ist. Dies bedeutet, dass die Konvergenzaussage (3.14) für genügend grosse  $k$  gleichmässig gilt. Daher können wir nun auch  $k = n$  gegen  $\infty$  gehen lassen, und (R3) ist erfüllt.

Es bleibt noch (R4) zu zeigen. Wir betrachten

$$\frac{\sigma_n(\theta_n)}{\sigma_n(0)} = \sqrt{\frac{F(-\theta_n) [1 - F(-\theta_n)] \delta_m^2}{F(0) [1 - F(0)] \delta_{0,m}^2}}.$$

Der Zähler unter der Wurzel ist eine stetige Funktion von  $F(-\theta_n)$  und  $\delta_m^2$ .  $\delta_m^2$  ist wiederum eine stetige Funktion von  $F(-\theta_n)$  (man betrachte die Definition von  $\delta_m^2$ , zusammen mit (1.1)). Somit hängt der ganze Zähler stetig von  $F(-\theta_n)$  und wegen der Stetigkeit von  $F$  auch von  $\theta_n$  ab.  $\theta_n$  strebt aber gegen 0, somit strebt der Zähler gegen den Nenner und der ganze Ausdruck auf der rechten Seite gegen 1.

Schliesslich berechnen wir die Effizienz des RSS-Vorzeichentests (bei gegebenem  $m$ ):

$$c_m^{\text{RSS}} = \lim_{n \rightarrow \infty} \frac{\mu'(0)}{\sqrt{mn} \sigma_n(0)} = \frac{mf(0)}{\frac{m}{2} \delta_{0,m}} = \frac{2f(0)}{\delta_{0,m}}$$

Ähnliche Überlegungen kann man für den SRS-Vorzeichentest anstellen: Wir verwenden etwas ungewöhnliche Stichprobenumfänge und Normierungen, was sich später vorteilhaft auswirken wird. Mit

$$\begin{aligned} N_n^{\text{SRS}} &= \left\lfloor \frac{mn}{\delta_{0,m}^2} \right\rfloor, \\ T_n^{\text{SRS}} &= \frac{S_{N_n^{\text{SRS}}}^+}{N_n^{\text{SRS}}}, \\ \mu^{\text{SRS}}(\theta) &= 1 - F(-\theta), \\ \sigma_n^{\text{SRS}}(\theta) &= \sqrt{\frac{F(-\theta) [1 - F(-\theta)]}{N_n^{\text{SRS}}}}, \\ C_n^{\text{SRS}} &\equiv z_{1-\alpha} \end{aligned}$$

erhält man

$$c^{\text{SRS}} = \lim_{n \rightarrow \infty} \frac{\mu^{\text{SRS}'(0)} \sigma_n^{\text{SRS}}(0)}{\sqrt{N_n^{\text{SRS}}} \sigma_n^{\text{SRS}}(0)} = \lim_{n \rightarrow \infty} \frac{f(0)}{\sqrt{F(0) [1 - F(0)]}} = 2f(0).$$

Wir brauchen nun noch die in Satz 3.4 geforderte Folge von Alternativen, die einen gemeinsamen Grenzwert für die Güte ergibt. Wir setzen dazu

$$\begin{aligned}\theta_n &= \frac{1}{\sqrt{n}} = \frac{\sqrt{m}}{\sqrt{mn}} = \frac{d^{\text{RSS}}}{\sqrt{N_n^{\text{RSS}}}} \\ &= \frac{\frac{\sqrt{m}}{\delta_{0,m}}}{\sqrt{\left\lfloor \frac{mn}{\delta_{0,m}^2} \right\rfloor}} + \frac{o(1)}{\sqrt{\left\lfloor \frac{mn}{\delta_{0,m}^2} \right\rfloor}} = \frac{d^{\text{SRS}}}{\sqrt{N_n^{\text{SRS}}}} + \frac{o(1)}{\sqrt{N_n^{\text{SRS}}}}\end{aligned}$$

mit

$$d^{\text{RSS}} = \sqrt{m}, \quad d^{\text{SRS}} = \frac{\sqrt{m}}{\delta_{0,m}}.$$

Satz 3.5 liefert

$$\begin{aligned}\lim_{n \rightarrow \infty} P_{\theta_n} \left( \frac{T_n^{\text{RSS}} - \mu^{\text{RSS}}(0)}{\sigma_n^{\text{RSS}}(0)} \geq z_{1-\alpha} \right) &= \Phi \left( c_m^{\text{RSS}} d^{\text{RSS}} - z_{1-\alpha} \right) \\ &= \Phi \left( \frac{2f(0)\sqrt{m}}{\delta_{0,m}} - z_{1-\alpha} \right) \\ &= \Phi \left( c^{\text{SRS}} d^{\text{SRS}} - z_{1-\alpha} \right) \\ &= \lim_{n \rightarrow \infty} P_{\theta_n} \left( \frac{T_n^{\text{SRS}} - \mu^{\text{SRS}}(0)}{\sigma_n^{\text{SRS}}(0)} \geq z_{1-\alpha} \right).\end{aligned}$$

Mit Hilfe von Satz 3.4 ist somit folgendes Resultat bewiesen:

**Satz 3.6** Die Verteilung  $F$  der  $X_{ijk}$  habe eine Dichte  $f$  mit  $f(0) > 0$ , und es gelte  $F(-\theta) [1 - F(-\theta)] \delta_m^2 \geq \tilde{\varepsilon}$  für alle  $\theta$  aus einer Umgebung von 0, mit  $\tilde{\varepsilon} > 0$ . Dann beträgt die Pitman-Effizienz von  $S_{m,n}^{+\text{RSS}}$  gegenüber  $S_N^{+\text{SRS}}$

$$e_{S_{m,n}^{+\text{RSS}}, S_N^{+\text{SRS}}} = \left( \frac{c_m^{\text{RSS}}}{c^{\text{SRS}}} \right)^2 = \frac{1}{\delta_{0,m}^2}.$$

Dies bedeutet, dass der RSS-Vorzeichentest z. B. für  $m = 5$  asymptotisch etwa halb so viele Messungen benötigt wie der SRS-Vorzeichentest, um das gleiche Niveau und die gleiche Güte zu erreichen.

## Kapitel 4

# Praktische Probleme

In den beiden letzten Kapiteln wurde gezeigt, dass mit Ranked Set Sampling genauere Schätzer bzw. (asymptotisch) effizientere Tests gewonnen werden können als bei der Ziehung einer zufälligen Stichprobe. Dabei wurden jedoch einige Probleme der praktischen Durchführung vernachlässigt, auf die bereits McIntyre (1952) hinweist. Die Betrachtung dieser Probleme soll zeigen, dass unter ungünstigen Umständen zwar der durch die RSS-Methode erzielte Gewinn kleiner ist als im bisherigen, idealisierten Kontext, dass die RSS-Methode aber in den betrachteten Beispielen immer noch mindestens so gut ist wie die entsprechende SRS-Methode.

### 4.1 Rangierungsfehler

Ein offensichtliches Problem bei der Auswahl der zu messenden Elemente besteht darin, dass die Rangierung der Elemente innerhalb eines Sets nicht unbedingt exakt erfolgt, sondern durch ein möglichst einfaches (günstiges) Verfahren, beispielsweise einen Vergleich von blossen Auge durch eine erfahrene Person. Dabei kann es natürlich vorkommen, dass aus dem  $i$ -ten Set nicht wirklich die  $i$ -te Ordnungsstatistik gemessen wird.

In Erweiterung der Notation aus Abschnitt 2.2 bezeichnen wir dasjenige Element aus dem  $i$ -ten Set des  $k$ -ten Durchgangs, das für die  $i$ -te Ordnungsstatistik gehalten wird, als  $X_{[i]k}$  (im Unterschied zur tatsächlichen  $i$ -ten Ordnungsstatistik  $X_{(i)k}$ ). Auch für die allgemeinen Symbole der Ordnungsstatistiken,  $X_{k:n}$ ,  $F_{k:n}$  usw., wollen wir Analogie bereitstellen:  $X_{k:n}$  bezeichne im Sinne einer „Schätzung der Ränge“ dasjenige von  $n$  Elementen, das für die  $k$ -te Ordnungsstatistik gehalten wird, und  $F_{k:n}$  sei dessen Verteilungsfunktion. Weiter kennzeichnen wir Symbole, die sich auf dieses RSS-Verfahren mit Rangierungsfehlern beziehen, mit  $RSS_j$  statt  $RSS$ . Das  $j$  steht dabei für

„judgement ordering“, also die Rangierung gemäss dem Urteil eines Experten.

Wir wollen untersuchen, welche Überlegungen aus den vorhergehenden Kapiteln in diesem neuen Rahmen noch gültig bleiben. Wir beginnen bei den Eigenschaften der Ordnungsstatistiken aus Kapitel 1 und zeigen ein Analogon zu (1.2). Der Einfachheit halber nehmen wir im folgenden an, dass die zugrundeliegende Verteilung der Stichprobenelemente absolut stetig sei, so dass wir von einer eindeutigen Rangierung ausgehen können. Weiter müssen wir als Minimalanforderung an die Rangierung durch den Experten verlangen, dass diese nicht davon abhängt, welcher Rang anschliessend ausgewählt werden soll – sonst kann der Experte die Resultate willkürlich beeinflussen. Für die Praxis heisst dies, dass der Experte erst nach Bekanntgabe seiner Rangierung erfahren sollte, welcher Rang ausgewählt wird.

Wir setzen  $p_{ik}(x) = P(X_{k:n} = X_{i:n} | X_{i:n} \leq x)$  und stellen fest, dass

$$\sum_{k=1}^n p_{ik}(x) = 1$$

gelten muss. Nun erhalten wir

$$\begin{aligned} F_{k:n}(x) &= P(X_{k:n} \leq x) \\ &= \sum_{i=1}^n P(X_{k:n} = X_{i:n}, X_{i:n} \leq x) \\ &= \sum_{i=1}^n p_{ik}(x) P(X_{i:n} \leq x) \\ \sum_{k=1}^n F_{k:n}(x) &= \sum_{k=1}^n \sum_{i=1}^n p_{ik}(x) F_{i:n}(x) \\ &= \sum_{i=1}^n F_{i:n}(x) \\ &\stackrel{(1.2)}{=} nF(x) \end{aligned} \tag{4.1}$$

Daraus folgt für existierende Momente wiederum (analog zu (1.3))

$$\sum_{k=1}^n E[X_{k:n}^r] = n E[X_1^r]. \tag{4.2}$$

Nun zu den Eigenschaften des RSS-Mittelwertschätzers aus Kapitel 2: Mit (4.2) statt (1.3) sieht man, dass auch

$$\hat{\mu}_{m,n}^{\text{RSSj}} = \frac{1}{mn} \sum_{k=1}^n \sum_{i=1}^m X_{[i]k}$$



noch ein erwartungstreuer Schätzer für den Erwartungswert  $E[X_{111}]$  der zugrundeliegenden Verteilung ist.

(2.1) wird zu

$$\text{Var} [\hat{\mu}_{m,n}^{\text{RSSj}}] = \frac{1}{m^2 n} \sum_{i=1}^m \sigma_{i:m}^2, \quad (4.3)$$

und die relative Präzision gegenüber dem SRS-Mittelwertschätzer mit gleichvielen gemessenen Elementen ist nun

$$\frac{\text{Var} [\hat{\mu}_{mn}^{\text{SRS}}]}{\text{Var} [\hat{\mu}_{m,n}^{\text{RSSj}}]} = \frac{m\sigma^2}{\sum_{i=1}^m \sigma_{i:m}^2}. \quad (4.4)$$

Wir möchten zeigen, dass dieses Verhältnis immer noch grösser oder gleich 1 ist. Dazu nehmen wir für einen Moment ohne Einschränkung der Allgemeinheit an, es gelte  $E[X_{111}] = 0$ . (Natürlich ist dies gerade der unbekannte, zu schätzende Parameter, aber momentan interessieren uns nur die Varianzen, die translationsinvariant sind.) Dann gilt

$$\begin{aligned} \sum_{i=1}^m \sigma_{i:m}^2 &= \sum_{i=1}^m E[X_{i:m}^2] - \sum_{i=1}^m [E(X_{i:m})]^2 \\ &\stackrel{(4.2)}{=} m E[X_{111}^2] - \sum_{i=1}^m [E(X_{i:m})]^2 \\ &\leq m E[X_{111}^2] \\ &= m\sigma^2, \end{aligned} \quad (4.5)$$

woraus die Behauptung unmittelbar folgt – die RSS-Methode ist also immer noch mindestens so genau wie die herkömmliche. Die weiteren Aussagen über die relative Präzision aus Kapitel 2 lassen sich aber nicht auf den Fall mit „judgement ordering“ übertragen.

Für den Vorzeichentest aus Kapitel 3 wollen wir zeigen, dass Rangierungsfehler die Pitman-Effizienz von

$$S_{m,n}^{+\text{RSSj}} = \sum_{k=1}^n \sum_{i=1}^m I(X_{[i]k} > 0)$$

gegenüber  $S_N^{+\text{SRS}}$  nicht unter 1 absinken lassen. Mit Hilfe von (4.1) statt (1.2) kann man völlig analog zum Vorgehen in Kapitel 3 zeigen, dass

$$E[S_{m,n}^{+\text{RSSj}}] = mn [1 - G(0)], \quad (4.6)$$

$$\text{Var} [S_{m,n}^{+\text{RSSj}}] = mn G(0) [1 - G(0)] \left[ 1 - \frac{\sum_{i=1}^m [G_{i:m}(0) - G(0)]^2}{m G(0) [1 - G(0)]} \right]. \quad (4.7)$$

Man setzt hier

$$\delta_{j,m}^2 = 1 - \frac{\sum_{i=1}^m [G_{i:m}(0) - G(0)]^2}{m G(0) [1 - G(0)]}$$

und erhält

$$\frac{S_{m,n}^{+RSSj} - mn [1 - G(0)]}{\sqrt{mn G(0) [1 - G(0)] \delta_{j,m}^2}} \xrightarrow{d} Z \sim N(0, 1) \quad (n \rightarrow \infty). \quad (4.8)$$

(4.6), (4.7) und (4.8) entsprechen den Formeln (3.1), (3.2) und (3.3). Nicht übernehmen lässt sich (3.4). Hingegen vereinfachen sich (4.6), (4.7) und (4.8) unter der Nullhypothese wiederum, und zwar zu

$$E_0 [S_{m,n}^{+RSSj}] = \frac{mn}{2}, \quad (4.9)$$

$$\text{Var}_0 [S_{m,n}^{+RSSj}] = \frac{mn}{4} \underbrace{\left[ 1 - \frac{4}{m} \sum_{i=1}^m \left( G_{i:m}(0) - \frac{1}{2} \right)^2 \right]}_{=: \delta_{j,0,m}^2} \quad (4.10)$$

$$\frac{S_{m,n}^{+RSSj} - \frac{mn}{2}}{\sqrt{\frac{mn}{4} \delta_{j,0,m}^2}} \xrightarrow{d} Z \quad (n \rightarrow \infty). \quad (4.11)$$

Der Abschnitt über die Berechnung von  $\delta_{0,m}^2$  lässt sich leider nicht übertragen, da ja kein Analogon zu (3.4) zur Verfügung steht. Die  $\delta_{j,0,m}^2$  sind aber ebenfalls kleiner oder gleich 1.

Wir betrachten nun den einseitigen Test von  $H_0 : \theta = 0$  gegen  $H_1 : \theta > 0$  zum asymptotischen Niveau  $\alpha$ , der  $H_0$  verwirft, falls

$$S_{m,n}^{+RSSj} - \frac{mn}{2} > \frac{1}{2} \sqrt{mn \delta_{j,0,m}^2} z_{1-\alpha}, \quad (4.12)$$

und untersuchen dessen Pitman-Effizienz gegenüber dem SRS-Vorzeichentest. Zur Überprüfung der Voraussetzungen von Satz 3.4 geht man wieder gleich vor. Um (R4) zu zeigen, müssen wir nun zusätzlich annehmen, dass  $G_{i:m}(0)$  stetig von  $\theta$  abhängt (Formel (1.1) steht nicht zur Verfügung). Diese Annahme dürfte aber für eine „vernünftige“ Rangierung wegen der zugrundeliegenden stetigen Verteilung durchaus sinnvoll sein. Die Effizienz ist

$$c_m^{RSSj} = \frac{2f(0)}{\delta_{j,0,m}},$$

und die Pitman-Effizienz gegenüber dem SRS-Vorzeichentest beträgt

$$e_{S_{m,n}^{+RSSj}, S_N^{+SRS}} = \frac{1}{\delta_{j,0,m}^2} \geq 1.$$

Auch der RSS-Vorzeichentest mit Rangierungsfehlern ist somit (asymptotisch) noch mindestens so gut wie der gewöhnliche SRS-Vorzeichentest. Der Nutzen dieses Resultats ist allerdings insofern beschränkt, dass das Ablehnungskriterium (4.12) von  $\delta_{j,0,m}^2$  abhängt. Diese Konstante, die von der Genauigkeit der Rangierung abhängt, ist im allgemeinen nicht bekannt, so dass ein Test zu einem gegebenen asymptotischen Niveau gar nicht angegeben werden kann. Die Aussage lautet also genauer: Wenn man (durch Annahme eines Wertes von  $\delta_{j,0,m}^2$  und Wahl eines Wertes von  $\alpha$ ) kritische Werte festlegt und dadurch Tests zum unbekanntem asymptotischen Niveau  $\tilde{\alpha}$  (i. a.  $\neq \alpha$ ) konstruiert, so benötigen diese Tests asymptotisch weniger Messungen als SRS-Vorzeichentests zum gleichen Niveau  $\tilde{\alpha}$  und mit gleicher Güte.

Abschliessend noch einige allgemeine, intuitive Überlegungen zum Problem der Rangierungsfehler: Solange man nur kleine Set-Grössen betrachtet, bietet die korrekte Rangierung wenig Probleme. Rangierungsfehler kommen nur dort vor, wo zwei Werte ohnehin praktisch gleich sind, was den Gewinn durch das RSS-Verfahren noch nicht wesentlich beeinträchtigt. Mit zunehmender Set-Grösse würden aber diese kleinen Unterschiede, die nicht richtig erfasst werden können, entscheidend für die unterschiedlichen Verteilungen der verschiedenen Ordnungsstatistiken. Anschaulich lässt sich das Problem etwa so beschreiben: Wenn man zu grosse Sets wählt, liegen die Erwartungswerte zweier benachbarter Ordnungsstatistiken so nahe beieinander, dass man sie nicht mehr unterscheiden kann. Grössere Sets bringen deshalb nur im Fall der fehlerfreien Rangierung noch einen wesentlichen Gewinn. Aus diesem Grund wird man die Set-Grösse nicht beliebig gross wählen. McIntyre (1952) vermutet, dass Sets mit einer Grösse über  $m = 5$  in seiner Anwendung nicht sinnvoll sind, und in der Folge werden in der Literatur (wie auch in dieser Arbeit) hauptsächlich die Fälle  $m = 2, \dots, 5$  betrachtet. Selbstverständlich ist aber der Wert  $m = 5$  keine allgemein gültige Grenze, da je nach Anwendungsgebiet und verwendeter Rangierungsmethode völlig verschiedene Situationen vorliegen können.

## 4.2 Abhängigkeiten innerhalb der Sets

Das „judgement ordering“ bringt noch ein weiteres Problem mit sich: Damit die Rangierung möglichst genau erfolgen kann, wird man die Sets in vielen Fällen nicht als zufällige Stichproben aus der ganzen betrachteten Population ziehen können, sondern beispielsweise räumlich benachbarte Elemente betrachten. Im Beispiel von McIntyre (vgl. Abschnitt 2.1) wird man etwa einige Quadrate in Sichtdistanz zu einem Set zusammenfassen, um die Erträge möglichst gut vergleichen zu können. Die  $m$  Werte innerhalb eines Sets

können deshalb nicht mehr als unabhängig betrachtet werden.

Wir modellieren die Abhängigkeiten innerhalb eines Sets in Anlehnung an Ridout und Cobby (1987) folgendermassen: Es sei

$$X_{ijk} = U_{ik} + V_{ijk} \quad (4.13)$$

mit

$$\begin{aligned} \mathbb{E}[U_{ik}] &\equiv \mathbb{E}[X_{ijk}] \equiv \mathbb{E}[X_{111}], & \text{Var}[U_{ik}] &\equiv \sigma_U^2, \\ \mathbb{E}[V_{ijk}] &\equiv 0, & \text{Var}[V_{ijk}] &\equiv \sigma_V^2. \end{aligned}$$

Dabei seien die  $U_{ik}$  i.i.d., ebenso die  $V_{ijk}$ , und all diese Zufallsvariablen seien unabhängig.  $U_{ik}$  ist also der Erwartungswert für ein bestimmtes Set, und  $V_{ijk}$  die Abweichung eines einzelnen Elements von diesem Erwartungswert. Aufgrund der Annahme der Unabhängigkeit ist  $\text{Var}[X_{ijk}] = \sigma^2 = \sigma_U^2 + \sigma_V^2$ . Weiter gehen wir davon aus, dass die Rangierung translationsinvariant sei, dass also  $X_{[i]k} = U_{ik} + V_{[i]k}$  mit unabhängigen Summanden gelte.

Wieder untersuchen wir die Eigenschaften des Mittelwertschätzers  $\hat{\mu}_{m,n}^{\text{RSSj}}$  unter diesen allgemeineren Voraussetzungen. Man beachte, dass im angegebenen Modell die ausgewählten Elemente  $X_{[i]k}$  immer noch unabhängig und für ein festes  $i$  identisch verteilt sind. Deshalb kann man die gleichen Aussagen über  $\hat{\mu}_{m,n}^{\text{RSSj}}$  machen wie im Abschnitt 4.1: Der Schätzer ist weiterhin erwartungstreu und hat eine Varianz, die kleiner oder gleich derjenigen des SRS-Schätzers ist. Es ändern sich allerdings die  $\sigma_{i:m}^2$ , die in (4.3) die Varianz des Schätzers bestimmen.

Wir leiten einen Ausdruck für die Varianz von  $\hat{\mu}_{m,n}^{\text{RSSj}}$  her, der der Zerlegung in diesem Modell Rechnung trägt:

$$\begin{aligned} \text{Var}[\hat{\mu}_{m,n}^{\text{RSSj}}] &= \frac{1}{m^2 n^2} \sum_{k=1}^n \sum_{i=1}^m \text{Var}[X_{[i]k}] \\ &= \frac{1}{m^2 n^2} \sum_{k=1}^n \sum_{i=1}^m \text{Var}[U_{ik}] + \frac{1}{m^2 n^2} \sum_{k=1}^n \sum_{i=1}^m \text{Var}[V_{[i]k}] \\ &= \frac{\sigma_U^2}{mn} + \frac{1}{m^2 n} \sum_{i=1}^m \sigma_{V,i:m}^2 \end{aligned}$$

Im schlimmsten Fall gilt  $\sigma_U^2 = \sigma^2$ , somit  $\sigma_V^2 = 0$  und auch  $\sigma_{V,i:m}^2 = 0$ . Dann ergibt sich die gleiche Varianz wie für den SRS-Mittelwertschätzer. Dies ist nicht erstaunlich, da ein Set in diesem Fall aus lauter gleichen Elementen besteht und das RSS-Auswahlverfahren demzufolge nichts anderes als eine i.i.d.-Stichprobe vom Umfang  $mn$  liefert.

Auch für den Vorzeichentest lassen sich die Ausführungen des Abschnitts 4.1 unverändert übernehmen. Es gilt wiederum, dass die RSS-Variante asymptotisch mindestens so effizient ist wie die SRS-Variante. (Auch in dieser Situation sind natürlich die Konstanten  $\delta_{j,0,m}^2$  im allgemeinen nicht bekannt.)

Die Veränderung gegenüber Abschnitt 4.1 geschieht hier in den Verteilungen  $G_{\hat{i}m}$ , auf denen die Definition von  $\delta_{j,0,m}^2$  beruht. Um diese Verteilungen im Modell (4.13) passend darstellen zu können, betrachten wir anstelle der Momente der  $U_{ik}$  und  $V_{ijk}$  deren Verteilungsfunktionen: Wir schreiben  $G_U$  für die Verteilungsfunktion von  $U_{ik}$  (mit  $G_U(x) = F_U(x-\theta)$  und  $G_U(\theta) = F_U(0) = \frac{1}{2}$ ) und  $F_V$  für diejenige von  $V_{ijk}$ . Dann gilt für die Verteilungsfunktion  $G$  von  $X_{ijk}$ , dass  $G = G_U * F_V$ , und  $G(x) = F(x-\theta) = (F_U * F_V)(x-\theta)$ . Aufgrund der Translationsinvarianz der Rangierung gilt unter der Nullhypothese

$$G_{\hat{i}m}(0) = \int_{-\infty}^{\infty} F_U(-t) dF_{V,\hat{i}m}(t).$$

Ähnlich wie beim Mittelwertschätzer liegt offensichtlich der schlechteste Fall vor, wenn  $F_V(x)$  und somit auch die  $F_{V,\hat{i}m}(x)$  zur Indikatorfunktion  $I(x \geq 0)$  degeneriert sind – dann gilt  $G_{\hat{i}m}(0) = F_U(0) = \frac{1}{2}$ . Durch Einsetzen in die Definition von  $\delta_{j,0,m}^2$  erhält man unmittelbar  $\delta_{j,0,m}^2 = 1$ , also entspricht dieser Fall dem SRS-Vorzeichentest.

## Kapitel 5

# Kostenüberlegungen

Bisher sind wir stets davon ausgegangen, dass die Kosten für die (ungefähre) Rangierung vernachlässigbar sind und nur die Kosten für die genaue Messung ins Gewicht fallen. Demzufolge haben wir jeweils ein RSS-Verfahren, das  $mn$  Elemente aus  $m^2n$  auswählt, mit einem SRS-Verfahren verglichen, das auf  $mn$  Elementen beruht, also gleichviele Messungen erfordert. Die untersuchten RSS-Verfahren erwiesen sich auch unter ungünstigen Umständen als mindestens so gut wie die SRS-Analoga.

Etwas anders ist die Situation, wenn man den Rangierungsaufwand mitberücksichtigt. Solche Überlegungen findet man bei Dell und Clutter (1972). Wir bezeichnen mit  $\kappa^M$  die „Messkosten“, worunter wir die Kosten zum Ziehen eines Elements und zur Messung des zugehörigen Werts verstehen wollen. Weiter seien  $\kappa_m^R$  die „Rangierungskosten“, definiert als die Kosten, die beim Ziehen von  $m - 1$  zusätzlichen Elementen und der (ungefähren) Rangierung von  $m$  Elementen anfallen.

Aufgrund dieser Definitionen entstehen bei der Durchführung eines RSS-Verfahrens mit Set-Grösse  $m$  und mit  $n$  Durchgängen insgesamt Kosten von  $mn(\kappa^M + \kappa_m^R)$ . Ein SRS-Verfahren, das auf einer Stichprobe des Umfangs  $N$  beruht, verursacht Gesamtkosten von  $N\kappa^M$ .

Wir wollen nun nicht mehr RSS- und SRS-Verfahren mit gleichvielen gemessenen Elementen vergleichen, sondern solche, die die gleichen Gesamtkosten verursachen. Gesucht sind Bedingungen, unter denen das RSS-Verfahren mindestens so gut ist wie das SRS-Verfahren.

Für den Mittelwertschätzer erhalten wir die Bedingungen

$$\frac{\sigma^2}{N} = \text{Var} [\hat{\mu}_N^{\text{SRS}}] \geq \text{Var} [\hat{\mu}_{m,n}^{\text{RSSj}}] = \frac{1}{m^2n} \sum_{i=1}^m \sigma_{i:m}^2$$

und

$$N\kappa^M = mn (\kappa^M + \kappa_m^R).$$

Durch Einsetzen ergibt sich

$$\frac{\sigma^2}{mn \frac{\kappa^M + \kappa_m^R}{\kappa^M}} \geq \frac{1}{m^2 n} \sum_{i=1}^m \sigma_{i:m}^2$$

bzw.

$$\frac{m\sigma^2}{\sum_{i=1}^m \sigma_{i:m}^2} \geq 1 + \frac{\kappa_m^R}{\kappa^M}. \quad (5.1)$$

Die relative Präzision (4.4) sollte also nicht nur grösser oder gleich 1 sein, wie bisher gefordert, sondern noch den Zusatzaufwand fürs Rangieren mit  $\kappa_m^R/\kappa^M$  „abgelten“.

Im Fall der perfekten Rangierung und zufällig gewählter Sets (Kapitel 2) betrug die obere Schranke für die relative Präzision  $\frac{m+1}{2}$ , also ist das RSS-Verfahren höchstens dann interessant, wenn

$$1 + \frac{\kappa_m^R}{\kappa^M} \leq \frac{m+1}{2},$$

d. h. wenn

$$\kappa_m^R \leq \frac{m-1}{2} \kappa^M.$$

Für  $m = 5$  dürfen beispielsweise die Ziehung von 4 zusätzlichen Elementen und die Rangierung eines Sets von 5 Elementen zusammen sicher nicht mehr als doppelt so viel kosten wie die Ziehung und Messung eines Elements. Dies entspricht dem bestmöglichen Fall der Uniformverteilung. Sofern die zugrundeliegende Verteilung bekannt ist und von der Uniformverteilung abweicht, lässt sich eine schärfere Ungleichung angeben.

Ähnliche Überlegungen kann man für den Vorzeichentest anstellen: Die RSS-Variante ist bei gleichen Kosten asymptotisch mindestens so gut wie die SRS-Variante, wenn

$$\frac{mn\kappa^M}{\delta_{j,0,m}^2} = mn\kappa^M e_{S_{m,n}^+ \text{RSS}, S_N^+ \text{SRS}} \geq mn(\kappa^M + \kappa_m^R),$$

was bedeutet, dass

$$\frac{1}{\delta_{j,0,m}^2} \geq 1 + \frac{\kappa_m^R}{\kappa^M}. \quad (5.2)$$

Wir stellen also an das Reziproke von  $\delta_{j,0,m}^2$  genau die gleiche Bedingung wie in (5.1) an die relative Präzision.

Bei perfekter Rangierung und zufälligen Sets sind die  $\delta_{j,0,m}^2 = \delta_{0,m}^2$  aus Abschnitt 3.4 bekannt. Für  $m = 5$  muss etwa  $\kappa_5^R \leq \kappa^M(1/\delta_{0,5}^2 - 1) \approx 1.03 \kappa^M$  gelten – eine wesentlich einschränkendere Bedingung als die oben für den Mittelwertschätzer bei Uniformverteilung angegebene.

## Kapitel 6

# Erkennungsmöglichkeiten für die praktischen Probleme

Durch den Einbezug der Rangierungskosten wird klar, dass die vorgestellten RSS-Verfahren den SRS-Verfahren nicht in jedem Fall überlegen sind. Insbesondere kann die RSS-Methode beim Auftreten der in Kapitel 4 vorgestellten Probleme unrentabel werden – der Aufwand für die Rangierungen würde in diesem Fall besser ins Erheben einiger zusätzlicher Messungen in einem SRS-Verfahren gesteckt.

Wir suchen deshalb nach Methoden zur Überprüfung der ursprünglichen Voraussetzungen in den ersten Kapiteln, also der Voraussetzungen, dass

- (a) die Elemente  $X_{ijk}$  i.i.d. seien (entspricht  $\sigma_U^2 = 0$  in Modell (4.13)) und dass
- (b) die Rangierung korrekt erfolge.

Ziel ist es, wesentliche Verletzungen dieser Voraussetzungen zu erkennen.

### 6.1 Allgemeine Überlegungen

Wir gehen davon aus, dass uns zur Beurteilung die effektiv gemessenen Werte aus der RSS-Auswahlprozedur (mit den zugeordneten Rängen) zur Verfügung stehen, also die  $m \times n$ -Matrix

$$\mathbf{X} = \begin{pmatrix} X_{[1]1} & \cdots & X_{[1]n} \\ \vdots & & \vdots \\ X_{[m]1} & \cdots & X_{[m]n} \end{pmatrix}.$$



Über die  $mn(m-1)$  rangierten, aber nicht gemessenen Elemente liegen uns keine Informationen vor. Wir möchten aufgrund dieser Daten eines RSS-Verfahrens beurteilen, ob für eine weitere, ähnliche Studie wiederum ein RSS-Verfahren zu wählen ist oder ob allenfalls ein SRS-Verfahren vorteilhafter ist.

Zu dieser Problemstellung ist folgendes anzumerken: In der Situation, in der ein Experte die Rangierungen vorgenommen hat, wird einem dieser in der Regel wohl auch Angaben darüber machen können, wie einfach sich die Sets rangieren liessen und ob es wesentliche Unterschiede zwischen den Sets gab. Wir stellen uns also beispielsweise eine der folgenden Situationen vor:

- Die Untersuchung liegt bereits einige Zeit zurück, und der Experte steht nicht mehr zur Verfügung oder kann sich nicht mehr an die Details erinnern.
- Mehrere Experten haben in einer grösseren Studie jeweils einen Teil der Daten rangiert, weshalb keiner alle Sets vergleichen kann.
- Die Rangierung erfolgt durch paarweise Vergleiche, die wirklich nur das Vorzeichen der Differenz zwischen zwei Werten liefern, aber keinerlei Information über den Betrag der Differenz oder die Werte selbst.

Sowohl eine zufällige Rangierung als auch eine Degeneration der Sets zu lauter gleichen Elementen ( $\sigma_V^2 = 0$  in (4.13)) führen dazu, dass  $\mathbf{X}$  einer i.i.d.-Stichprobe entspricht. Wir können deshalb mit den vorliegenden Daten Verletzungen der Voraussetzungen (a) und (b) nicht getrennt untersuchen.

Die Zerlegung der Varianz im Modell (4.13) entspricht einem klassischen Varianzanalyse-Modell. Man beachte aber, dass pro Gruppe (Set) in diesem Fall nur jeweils eine Beobachtung vorliegt. Die Varianz innerhalb der Gruppen kann deshalb nicht geschätzt werden, und ein Vergleich mit der Varianz zwischen den Gruppen ist somit nicht möglich.

Unsere Überlegungen werden im wesentlichen auf folgender Tatsache beruhen: Unter den Bedingungen (a) und (b) ist für  $i < j$  ein  $X_{(i)k}$  tendenziell kleiner als ein  $X_{(j)k}$ , da es ja eine niedrigere Ordnungsstatistik vertritt. (Es braucht allerdings nicht  $X_{(1)k} \leq X_{(2)k} \leq \dots \leq X_{(m)k}$  zu gelten, wie dies bei flüchtiger Betrachtung der Notation angenommen werden könnte und etwa in Hettmansperger/McKean (1998) behauptet wird.)

## 6.2 Grafische Methode

Als einfaches Mittel zur Beurteilung der oben erwähnten Voraussetzungen (a) und (b) untersuchen wir zunächst eine grafische Darstellung der Daten. Dazu zeichnen wir zu jeder Zeile von  $\mathbf{X}$  die empirische Quantilfunktion: Wenn

$r(X_{[i]k})$  den Rang von  $X_{[i]k}$  innerhalb der  $i$ -ten Zeile bezeichnet, so interpolieren wir zwischen den Punkten  $(r(X_{[i]k})/(n+1), X_{[i]k})$  ( $k = 1, \dots, n$ ) linear. Zur Kennzeichnung der  $i$ -ten Zeile beschriften wir die gezeichneten Punkte mit dem Wert von  $i$ . Dieses Vorgehen wiederholen wir für jedes  $i$  von 1 bis  $m$ .

Falls die Voraussetzungen (a) und (b) erfüllt sind, so erhalten wir empirische Quantilfunktionen für die Ordnungsstatistiken einer Stichprobe vom Umfang  $m$  aus der Verteilung  $G$  der  $X_{ijk}$ . Diese empirischen Quantilfunktionen beruhen jeweils auf  $n$  Elementen.

Für die entsprechenden theoretischen Quantilfunktionen  $G_{i:m}^{-1}$  gilt

$$G_{i:m}^{-1}(p) \leq G_{i+1:m}^{-1}(p).$$

Anschaulich ausgedrückt: Die theoretischen Quantilfunktionen liegen der Reihe nach angeordnet übereinander. Auch die empirischen Quantilfunktionen sollten unter (a) und (b) nicht zu sehr von diesem Muster abweichen. Abbildung 6.1 zeigt eine typische Realisierung der Grafik für  $m = 5$  und  $n = 10$ , wobei (a) und (b) erfüllt und die  $X_{ijk}$  standardnormalverteilt sind: Die erwartete Anordnung ist im wesentlichen erfüllt, wie hier ist aber vor allem am Rand mit Überschneidungen zu rechnen.

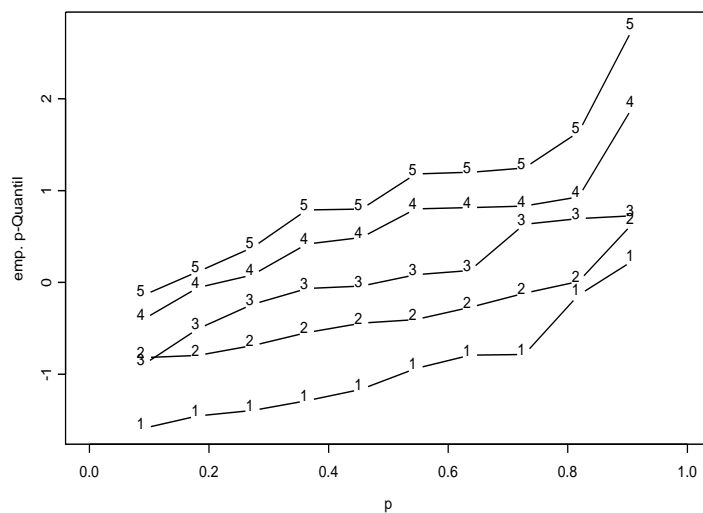


Abbildung 6.1: Beispiel für die empirischen Quantilfunktionen unter den Voraussetzungen (a) und (b) für  $m = 5$  und  $n = 10$  ( $X_{ijk}$  standardnormalverteilt).

Sind die Elemente von  $\mathbf{X}$  i.i.d., so erhalten wir  $m$  empirische Quantilfunktionen für die  $X_{ijk}$  selbst. Die entsprechenden theoretischen Quantilfunktionen sind alle gleich (nämlich  $G^{-1}$ ), und die  $m$  empirischen Quantilfunktionen werden nahe beieinander liegen. Wenn überhaupt eine Anordnung erkennbar ist, so wird diese jede Permutation mit gleicher Wahrscheinlichkeit annehmen. In Abbildung 6.2 ist eine Realisierung dieses Falles zu sehen, wobei die  $X_{[i]k}$  i.i.d. standardnormalverteilt sind.

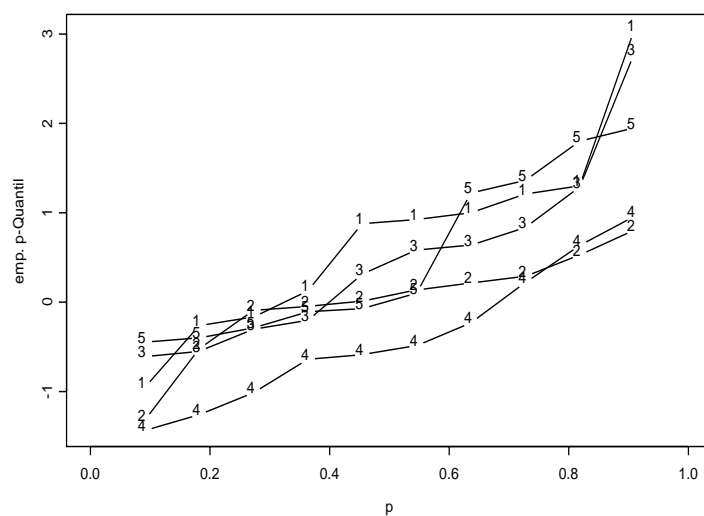


Abbildung 6.2: Beispiel für die empirischen Quantilfunktionen mit i.i.d. standardnormalverteilten  $X_{[i]k}$  für  $m = 5$  und  $n = 10$ .

Die beschriebene grafische Methode hat den Vorteil, dass sie die Daten praktisch ohne Informationsverlust darstellt (lediglich die Zugehörigkeit zu einem bestimmten Durchgang geht verloren). Eine „objektive“ Interpretation ist hingegen kaum möglich – welche Muster sollen noch toleriert werden, welche nicht mehr? Ausserdem hängen die entstehenden Muster von  $m$  und  $n$  sowie auch relativ stark von der zugrundeliegenden Verteilung ab; unter Umständen müsste die Skala für die beobachteten Werte angepasst werden. Aufgrund dieser Überlegungen wäre eine Statistik wünschenswert, die die Abweichung von den Voraussetzungen in eine Zahl fasst und allenfalls sogar einen Test erlaubt. Solche Statistiken suchen wir in den folgenden Kapiteln.

## Kapitel 7

### Statistik $R_{m,n}(\mathbf{X})$

Eine einfache Statistik zur Überprüfung der Voraussetzungen (a) und (b) zählt, wie oft innerhalb jedes Durchgangs eine „ $i$ -te Ordnungsstatistik“ grösser als die „ $(i + 1)$ -te Ordnungsstatistik“ ist. Wir normieren diese Anzahl mit der maximal möglichen Anzahl solcher Ereignisse und erhalten folgende Statistik:

$$R_{m,n}(\mathbf{X}) = \frac{1}{(m-1)n} \sum_{k=1}^n \sum_{i=1}^{m-1} \mathbf{I}(X_{[i]k} > X_{[i+1]k})$$

(Die Beschränkung auf Vergleiche innerhalb der Durchgänge vereinfacht die Abhängigkeitsstruktur und wird später fallengelassen. Falls die Zugehörigkeit zu den verschiedenen Durchgängen nicht mehr bekannt ist, kann man zufällig je einen Vertreter jeder Ordnungsstatistik zu einem „Durchgang“ zusammenfassen – die Elemente eines Durchgangs stehen in keiner besonderen Beziehung zueinander.)

Wenn wir unsere Fragestellung als Testproblem darstellen wollen, so lautet die Nullhypothese, dass die  $X_{[i]k}$  aus dem RSS-Auswahlverfahren mit perfekter Rangierung stammen, welches auf  $m^2n$  i.i.d. Stichprobenelementen  $X_{ijk}$  basiert (Voraussetzungen (a) und (b)). Diese Formulierung der Nullhypothese entspricht unserem Ziel, wesentliche Abweichungen von (a) und (b) zu erkennen. Als extreme Alternative betrachten wir wiederum insbesondere den Fall, dass die  $X_{[i]k}$  eine i.i.d.-Stichprobe mit  $mn$  Elementen bilden (tritt ein für zufällige Rangierungen oder  $\sigma_V^2 = 0$ ).

Für dieses wie auch für die folgenden Kapitel gehen wir davon aus, dass die Stichprobenelemente  $X_{ijk}$  einer absolut stetigen Verteilung entstammen.

## 7.1 Verhalten unter der Nullhypothese

**Satz 7.1** *Unter  $H_0$  gilt*

$$E_0 [R_{m,n}(\mathbf{X})] = \frac{1}{m-1} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{i \binom{m}{i} \binom{m}{j}}{(i+j) \binom{2m}{i+j}}, \quad (7.1)$$

$$\text{Var}_0 [R_{m,n}(\mathbf{X})] = \frac{1}{(m-1)^2 n} \left[ \sum_{i=1}^{m-1} \text{Var}_0 (J_{i,k}) + 2 \sum_{i=1}^{m-2} \text{Cov}_0 (J_{i,k}, J_{i+1,k}) \right]. \quad (7.2)$$

Dabei ist  $J_{i,k} = \mathbb{I}(X_{(i)k} > X_{(i+1)k})$  ( $i = 1, \dots, m-1$ ), mit

$$\text{Var}_0 (J_{i,k}) = \left[ \sum_{j=i+1}^m \frac{i \binom{m}{i} \binom{m}{j}}{(i+j) \binom{2m}{i+j}} \right] \left[ 1 - \sum_{j=i+1}^m \frac{i \binom{m}{i} \binom{m}{j}}{(i+j) \binom{2m}{i+j}} \right], \quad (7.3)$$

$$\begin{aligned} \text{Cov}_0 (J_{i,k}, J_{i+1,k}) &= \sum_{j=i+2}^m \sum_{\ell=i+j+1}^{2m} \frac{i(i+1) \binom{m}{i} \binom{m}{i+1} \binom{m}{j} \binom{2m}{\ell}}{(i+j+1)(i+\ell) \binom{2m}{i+j+1} \binom{3m}{i+\ell}} \\ &\quad - \left[ \sum_{j=i+1}^m \frac{i \binom{m}{i} \binom{m}{j}}{(i+j) \binom{2m}{i+j}} \right] \left[ \sum_{j=i+2}^m \frac{(i+1) \binom{m}{i+1} \binom{m}{j}}{(i+j+1) \binom{2m}{i+j+1}} \right]. \end{aligned} \quad (7.4)$$

**Beweis:** Wir benötigen vorerst folgendes Resultat:

$$\begin{aligned} &\int_{-\infty}^y F_{k_1:n_1}(x) f_{k_2:n_2}(x) dx \\ &\stackrel{(1.1), (1.4)}{=} \int_{-\infty}^y \left[ \sum_{j=k_1}^{n_1} \binom{n_1}{j} [F(x)]^j [1-F(x)]^{n_1-j} \right] \\ &\quad \cdot \frac{n_2!}{(k_2-1)!(n_2-k_2)!} [F(x)]^{k_2-1} f(x) [1-F(x)]^{n_2-k_2} dx \\ &\stackrel{(1.4)}{=} \sum_{j=k_1}^{n_1} \frac{k_2 \binom{n_1}{j} \binom{n_2}{k_2}}{(j+k_2-1)!(n_1+n_2-j-k_2)!} \int_{-\infty}^y f_{j+k_2:n_1+n_2}(x) dx \\ &= \sum_{j=k_1}^{n_1} \frac{k_2 \binom{n_1}{j} \binom{n_2}{k_2}}{(j+k_2) \binom{n_1+n_2}{j+k_2}} F_{j+k_2:n_1+n_2}(y) \end{aligned} \quad (7.5)$$

Indem wir  $y$  gegen  $\infty$  gehen lassen, erhalten wir weiter

$$\int_{-\infty}^{\infty} F_{k_1:n_1}(x) f_{k_2:n_2}(x) dx = \sum_{j=k_1}^{n_1} \frac{k_2 \binom{n_1}{j} \binom{n_2}{k_2}}{(j+k_2) \binom{n_1+n_2}{j+k_2}}. \quad (7.6)$$

Nun betrachten wir die Bernoulli-Zufallsvariablen  $J_{i,k}$ . Für diese gilt

$$\begin{aligned}
 E_0(J_{i,k}) &= P_0(X_{(i)k} > X_{(i+1)k}) \\
 &= \int_{-\infty}^{\infty} F_{i+1:m}(x) f_{i:m}(x) dx \\
 &\stackrel{(7.6)}{=} \sum_{j=i+1}^m \frac{i \binom{m}{i} \binom{m}{j}}{(i+j) \binom{2m}{i+j}}
 \end{aligned} \tag{7.7}$$

und

$$\begin{aligned}
 \text{Var}_0(J_{i,k}) &= [P_0(X_{(i)k} > X_{(i+1)k})] [1 - P_0(X_{(i)k} > X_{(i+1)k})] \\
 &= \left[ \sum_{j=i+1}^m \frac{i \binom{m}{i} \binom{m}{j}}{(i+j) \binom{2m}{i+j}} \right] \left[ 1 - \sum_{j=i+1}^m \frac{i \binom{m}{i} \binom{m}{j}}{(i+j) \binom{2m}{i+j}} \right],
 \end{aligned}$$

also (7.3). Ähnlich berechnen wir die Kovarianzen zwischen benachbarten  $J_{i,k}$ :

$$\begin{aligned}
 E_0(J_{i,k} J_{i+1,k}) &= P_0(X_{(i)k} > X_{(i+1)k} > X_{(i+2)k}) \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{x_1} F_{i+2:m}(x_2) f_{i+1:m}(x_2) dx_2 f_{i:m}(x_1) dx_1 \\
 &\stackrel{(7.5)}{=} \int_{-\infty}^{\infty} \sum_{j=i+2}^m \frac{(i+1) \binom{m}{i+1} \binom{m}{j}}{(i+j+1) \binom{2m}{i+j+1}} F_{i+j+1:2m}(x_1) f_{i:m}(x_1) dx_1 \\
 &\stackrel{(7.6)}{=} \sum_{j=i+2}^m \frac{(i+1) \binom{m}{i+1} \binom{m}{j}}{(i+j+1) \binom{2m}{i+j+1}} \sum_{\ell=i+j+1}^{2m} \frac{i \binom{m}{i} \binom{2m}{\ell}}{(i+\ell) \binom{3m}{i+\ell}} \\
 &= \sum_{j=i+2}^m \sum_{\ell=i+j+1}^{2m} \frac{i(i+1) \binom{m}{i} \binom{m}{i+1} \binom{m}{j} \binom{2m}{\ell}}{(i+j+1)(i+\ell) \binom{2m}{i+j+1} \binom{3m}{i+\ell}}
 \end{aligned} \tag{7.8}$$

Um (7.4) zu zeigen, müssen wir nur noch (7.8) und (7.7) in

$$\text{Cov}_0(J_{i,k}, J_{i+1,k}) = E_0(J_{i,k} J_{i+1,k}) - E_0(J_{i,k}) E_0(J_{i+1,k})$$

einsetzen.

Nach diesen Vorarbeiten erhalten wir leicht Erwartungswert und Varianz von  $R_{m,n}(\mathbf{X})$  unter der Nullhypothese, dass (a) und (b) gelten:

$$E_0[R_{m,n}(\mathbf{X})] = E_0 \left[ \frac{1}{(m-1)n} \sum_{k=1}^n \sum_{i=1}^{m-1} J_{i,k} \right]$$

$$\begin{aligned}
&\stackrel{(7.7)}{=} \frac{1}{m-1} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{i \binom{m}{i} \binom{m}{j}}{(i+j) \binom{2m}{i+j}} \\
\text{Var}_0 [R_{m,n}(\mathbf{X})] &= \text{Var}_0 \left[ \frac{1}{(m-1)n} \sum_{k=1}^n \sum_{i=1}^{m-1} J_{i,k} \right] \\
&= \frac{1}{(m-1)^2 n} \left[ \sum_{i=1}^{m-1} \text{Var}_0 (J_{i,k}) + 2 \sum_{i=1}^{m-2} \text{Cov}_0 (J_{i,k}, J_{i+1,k}) \right]
\end{aligned}$$

Für die letzte Gleichung haben wir die Unabhängigkeit von  $J_{i,k}$  und  $J_{j,\ell}$  für  $k \neq \ell$  oder  $|i-j| > 1$  benützt. Damit ist der Beweis von Satz 7.1 vollständig.

Die Erwartungswerte und Varianzen für  $m = 2, \dots, 5$  und  $n = 1$  sind in Tabelle 7.1 angegeben. Für andere  $n$  gelten die gleichen Erwartungswerte; die Varianzen ergeben sich mittels Division durch  $n$ .

Tabelle 7.1: Erwartungswert und Varianz von  $R_{m,1}(\mathbf{X})$  unter (a) und (b) ( $m = 2, \dots, 5$ ).

$m$	$E_0 [R_{m,1}(\mathbf{X})]$	$\text{Var}_0 [R_{m,1}(\mathbf{X})]$
2	0.1666667	0.13888889
3	0.2000000	0.06476190
4	0.2238095	0.04106576
5	0.2420635	0.02964419

Mit Hilfe des zentralen Grenzwertsatzes erhalten wir asymptotische Toleranzbereiche für  $n \rightarrow \infty$  (man beachte, dass  $R_{m,n}(\mathbf{X})$  als Summe von  $n$  i.i.d. Zufallsvariablen mit Verteilung wie  $R_{m,1}(\mathbf{X})$  dargestellt werden kann). Für grosse  $n$  können wir also die Nullhypothese zum approximativen Niveau  $\alpha$  verwerfen, falls

$$|R_{m,n}(\mathbf{X}) - E_0 [R_{m,1}(\mathbf{X})]| > \sqrt{\frac{1}{n} \text{Var}_0 [R_{m,1}(\mathbf{X})]} z_{1-\alpha/2}. \quad (7.9)$$

In diesem Fall müssen wir davon ausgehen, dass zumindest eine der Voraussetzungen (a) und (b) verletzt ist. Dies braucht allerdings noch nicht zu heissen, dass kein RSS-Verfahren mehr angewendet werden soll – aufgrund des festen Niveaus kann es durchaus geschehen, dass der Test „zu gut“ wird, also bei grossen Stichproben bereits kleine Abweichungen von (a) und (b) erkennt, die die Wirksamkeit des RSS-Verfahrens kaum beeinträchtigen. Wir werden im Abschnitt 7.3 ein alternatives Entscheidungskriterium betrachten.

Vorerst wollen wir aber noch die Verteilungen der  $R_{m,n}(\mathbf{X})$  für kleine  $m$  und  $n$  untersuchen, um in diesem Fall ohne Normalverteilungsapproximation argumentieren zu können. Wir müssen für jedes  $m$  nur die Verteilung von  $R_{m,1}(\mathbf{X})$  berechnen, die übrigen ergeben sich durch Faltung.

Für  $m = 2$  ist dies kein Problem –  $R_{2,1}(\mathbf{X})$  nimmt nur die Werte 0 und 1 an, mit

$$\begin{aligned} P_0 [R_{2,1}(\mathbf{X}) = 1] &= E_0 [R_{2,1}(\mathbf{X})] \\ &\stackrel{(7.1)}{=} \frac{1}{6}. \end{aligned} \quad (7.10)$$

Für  $m = 3$  und eine beliebige Permutation  $(i_1, i_2, i_3)$  von  $(1, 2, 3)$  berechnen wir

$$\begin{aligned} &P_0 (X_{(i_1)1} < X_{(i_2)1} < X_{(i_3)1}) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{x_3} F_{i_1:3}(x_2) f_{i_2:3}(x_2) dx_2 f_{i_3:3}(x_3) dx_3 \\ &\stackrel{(7.5)}{=} \int_{-\infty}^{\infty} \sum_{j=i_1}^3 \frac{i_2 \binom{3}{i_2} \binom{3}{j}}{(i_2+j) \binom{6}{i_2+j}} F_{i_2+j:6}(x_3) f_{i_3:3}(x_3) dx_3 \\ &\stackrel{(7.6)}{=} \sum_{j=i_1}^3 \frac{i_2 \binom{3}{i_2} \binom{3}{j}}{(i_2+j) \binom{6}{i_2+j}} \sum_{\ell=i_2+j}^6 \frac{i_3 \binom{3}{i_3} \binom{6}{\ell}}{(i_3+\ell) \binom{9}{i_3+\ell}}. \end{aligned}$$

Mit einigem Rechenaufwand erhalten wir

$$\begin{aligned} P_0 (X_{(1)1} < X_{(2)1} < X_{(3)1}) &= \frac{64}{105}, \\ P_0 (X_{(1)1} < X_{(3)1} < X_{(2)1}) &= P_0 (X_{(2)1} < X_{(1)1} < X_{(3)1}) = \frac{143}{840}, \\ P_0 (X_{(2)1} < X_{(3)1} < X_{(1)1}) &= P_0 (X_{(3)1} < X_{(1)1} < X_{(2)1}) = \frac{17}{840}, \\ P_0 (X_{(3)1} < X_{(2)1} < X_{(1)1}) &= \frac{1}{105}. \end{aligned}$$

Im ersten Fall gilt  $R_{3,1}(\mathbf{X}) = 0$ , im letzten  $R_{3,1}(\mathbf{X}) = 1$ , und in allen übrigen gilt  $R_{3,1}(\mathbf{X}) = 0.5$ . Somit haben wir unter der Nullhypothese folgende Verteilung gefunden:

$$\begin{aligned} P_0 [R_{3,1}(\mathbf{X}) = 0] &= \frac{64}{105}, \\ P_0 [R_{3,1}(\mathbf{X}) = 0.5] &= \frac{8}{21}, \\ P_0 [R_{3,1}(\mathbf{X}) = 1] &= \frac{1}{105}. \end{aligned}$$



Analog kann man die Verteilungen für grössere  $m$  berechnen. Die Wahrscheinlichkeiten für die verschiedenen Anordnungen von  $X_{(1)1}, \dots, X_{(m)1}$  sind für  $m = 2, \dots, 5$  in Tabelle 7.2, die daraus resultierenden Verteilungen von  $R_{m,1}(\mathbf{X})$  in Tabelle 7.3 zusammengestellt. Diese Verteilungen erlauben die Konstruktion exakter Toleranzbereiche für  $R_{m,1}(\mathbf{X})$ . Mittels Faltung können solche auch für  $R_{m,n}(\mathbf{X})$  ( $n > 1$ ) konstruiert werden.

## 7.2 Verhalten unter der speziellen Alternative

Nun interessieren uns aber natürlich auch die Eigenschaften von  $R_{m,n}(\mathbf{X})$  unter unserer „extremen Alternative“, die besagt, dass die Elemente von  $\mathbf{X}$  i.i.d. sind. Wir bezeichnen diese Alternative mit  $H_1^*$ , die entsprechenden Erwartungswerte, Wahrscheinlichkeiten usw. mit  $E_{1^*}$ ,  $P_{1^*}$  usw.

**Satz 7.2** *Unter  $H_1^*$  gilt*

$$E_{1^*} [R_{m,n}(\mathbf{X})] = \frac{1}{2}, \quad (7.11)$$

$$\text{Var}_{1^*} [R_{m,n}(\mathbf{X})] = \frac{m+1}{12(m-1)^2n}. \quad (7.12)$$

**Beweis:** Erneut betrachten wir zunächst die Indikatorfunktionen  $J_{i,k}$ , die wir in diesem Fall als

$$J_{i,k} = \mathbf{I}(X_{[i]k} > X_{[i+1]k}) \quad (i = 1, \dots, m-1)$$

definieren, um mit der Notation wieder eventuellen Rangierungsfehlern Rechnung zu tragen. Da nun die  $X_{[1]k}, \dots, X_{[m]k}$  i.i.d. sind, werden die Berechnungen wesentlich einfacher als unter (a) und (b):

$$E_{1^*}(J_{i,k}) = P_{1^*}(X_{[i]k} > X_{[i+1]k}) = \frac{1}{2}$$

$$\text{Var}_{1^*}(J_{i,k}) = \frac{1}{2} \left(1 - \frac{1}{2}\right) = \frac{1}{4}$$

$$E_{1^*}(J_{i,k} J_{i+1,k}) = P_{1^*}(X_{[i]k} > X_{[i+1]k} > X_{[i+2]k}) = \frac{1}{6}$$

$$\text{Cov}_{1^*}(J_{i,k}, J_{i+1,k}) = \frac{1}{6} - \left(\frac{1}{2}\right)^2 = -\frac{1}{12}$$

Tabelle 7.2: Wahrscheinlichkeiten für die Anordnungen von  $X_{(1)1}, \dots, X_{(m)1}$  unter (a) und (b) für  $m = 2, \dots, 5$  sowie entsprechende Werte der Statistiken  $R_{m,1}$  und  $\tilde{R}_{m,1}$  (vgl. Kapitel 8).

$m = 2$	$i_1$	$i_2$		$P_0(X_{(i_1)1} < X_{(i_2)1})$	$R_{2,1}(\mathbf{X})$	$\tilde{R}_{2,1}(\mathbf{X})$	
	1	2		0.8333333	0.0000	0.0000	
	2	1		0.1666667	1.0000	1.0000	
$m = 3$	$i_1$	$i_2$	$i_3$	$P_0(X_{(i_1)1} < \dots < X_{(i_3)1})$	$R_{3,1}(\mathbf{X})$	$\tilde{R}_{3,1}(\mathbf{X})$	
	1	2	3	0.6095238	0.0000	0.0000	
	1	3	2	0.1702381	0.5000	0.3333	
	2	1	3	0.1702381	0.5000	0.3333	
	2	3	1	0.0202381	0.5000	0.6667	
	3	1	2	0.0202381	0.5000	0.6667	
	3	2	1	0.0095238	1.0000	1.0000	
$m = 4$	$i_1$	$i_2$	$i_3$	$i_4$	$P_0(X_{(i_1)1} < \dots < X_{(i_4)1})$	$R_{4,1}(\mathbf{X})$	$\tilde{R}_{4,1}(\mathbf{X})$
	1	2	3	4	0.4028293	0.0000	0.0000
	1	2	4	3	0.1268577	0.3333	0.1667
	1	3	2	4	0.1627741	0.3333	0.1667
	1	3	4	2	0.0239891	0.3333	0.3333
	1	4	2	3	0.0254943	0.3333	0.3333
	1	4	3	2	0.0127508	0.6667	0.5000
	2	1	3	4	0.1268577	0.3333	0.1667
	2	1	4	3	0.0387933	0.6667	0.3333
	2	3	1	4	0.0254943	0.3333	0.3333
	2	3	4	1	0.0019616	0.3333	0.5000
	2	4	1	3	0.0041328	0.6667	0.5000
	2	4	3	1	0.0012119	0.6667	0.6667
	3	1	2	4	0.0239891	0.3333	0.3333
	3	1	4	2	0.0031338	0.3333	0.5000
	3	2	1	4	0.0127508	0.6667	0.5000
	3	2	4	1	0.0009188	0.6667	0.6667
	3	4	1	2	0.0006315	0.3333	0.6667
	3	4	2	1	0.0003841	0.6667	0.8333
	4	1	2	3	0.0019616	0.3333	0.5000
	4	1	3	2	0.0009188	0.6667	0.6667
	4	2	1	3	0.0012119	0.6667	0.6667
	4	2	3	1	0.0003366	0.6667	0.8333
	4	3	1	2	0.0003841	0.6667	0.8333
	4	3	2	1	0.0002319	1.0000	1.0000

Tabelle 7.2 (Fortsetzung)

$m = 5$	$i_1$	$i_2$	$i_3$	$i_4$	$i_5$	$P_0(X_{(i_1)1} < \dots < X_{(i_5)1})$	$R_{5,1}(\mathbf{X})$	$\bar{R}_{5,1}(\mathbf{X})$
	1	2	3	4	5	0.2450499	0.0000	0.0000
	1	2	3	5	4	0.0819477	0.2500	0.1000
	1	2	4	3	5	0.1134602	0.2500	0.1000
	1	2	4	5	3	0.0186571	0.2500	0.2000
	1	2	5	3	4	0.0204599	0.2500	0.2000
	1	2	5	4	3	0.0105929	0.5000	0.3000
	1	3	2	4	5	0.1134602	0.2500	0.1000
	1	3	2	5	4	0.0370130	0.5000	0.2000
	1	3	4	2	5	0.0289218	0.2500	0.2000
	1	3	4	5	2	0.0026345	0.2500	0.3000
	1	3	5	2	4	0.0054134	0.5000	0.3000
	1	3	5	4	2	0.0016989	0.5000	0.4000
	1	4	2	3	5	0.0289218	0.2500	0.2000
	1	4	2	5	3	0.0043191	0.2500	0.3000
	1	4	3	2	5	0.0165630	0.5000	0.3000
	1	4	3	5	2	0.0014281	0.5000	0.4000
	1	4	5	2	3	0.0010302	0.2500	0.4000
	1	4	5	3	2	0.0006418	0.5000	0.5000
	1	5	2	3	4	0.0030149	0.2500	0.3000
	1	5	2	4	3	0.0014804	0.5000	0.4000
	1	5	3	2	4	0.0019444	0.5000	0.4000
	1	5	3	4	2	0.0005821	0.5000	0.5000
	1	5	4	2	3	0.0006648	0.5000	0.5000
	1	5	4	3	2	0.0004113	0.7500	0.6000
	2	1	3	4	5	0.0819477	0.2500	0.1000
	2	1	3	5	4	0.0271583	0.5000	0.2000
	2	1	4	3	5	0.0370130	0.5000	0.2000
	2	1	4	5	3	0.0059691	0.5000	0.3000
	2	1	5	3	4	0.0064863	0.5000	0.3000
	2	1	5	4	3	0.0033333	0.7500	0.4000
	2	3	1	4	5	0.0204599	0.2500	0.2000
	2	3	1	5	4	0.0064863	0.5000	0.3000
	2	3	4	1	5	0.0030149	0.2500	0.3000
	2	3	4	5	1	0.0001595	0.2500	0.4000
	2	3	5	1	4	0.0005470	0.5000	0.4000
	2	3	5	4	1	0.0001111	0.5000	0.5000
	2	4	1	3	5	0.0054134	0.5000	0.3000
	2	4	1	5	3	0.0007565	0.5000	0.4000
	2	4	3	1	5	0.0019444	0.5000	0.4000
	2	4	3	5	1	0.0000972	0.5000	0.5000
	2	4	5	1	3	0.0001184	0.5000	0.5000
	2	4	5	3	1	0.0000501	0.5000	0.6000
	2	5	1	3	4	0.0005470	0.5000	0.4000
	2	5	1	4	3	0.0002608	0.7500	0.5000

Tabelle 7.2 (Fortsetzung)

$m = 5$	$i_1$	$i_2$	$i_3$	$i_4$	$i_5$	$P_0(X_{(i_1)1} < \dots < X_{(i_5)1})$	$R_{5,1}(\mathbf{X})$	$\bar{R}_{5,1}(\mathbf{X})$
	2	5	3	1	4	0.0002403	0.5000	0.5000
	2	5	3	4	1	0.0000462	0.5000	0.6000
	2	5	4	1	3	0.0000824	0.7500	0.6000
	2	5	4	3	1	0.0000345	0.7500	0.7000
	3	1	2	4	5	0.0186571	0.2500	0.2000
	3	1	2	5	4	0.0059691	0.5000	0.3000
	3	1	4	2	5	0.0043191	0.2500	0.3000
	3	1	4	5	2	0.0003656	0.2500	0.4000
	3	1	5	2	4	0.0007565	0.5000	0.4000
	3	1	5	4	2	0.0002289	0.5000	0.5000
	3	2	1	4	5	0.0105929	0.5000	0.3000
	3	2	1	5	4	0.0033333	0.7500	0.4000
	3	2	4	1	5	0.0014804	0.5000	0.4000
	3	2	4	5	1	0.0000750	0.5000	0.5000
	3	2	5	1	4	0.0002608	0.7500	0.5000
	3	2	5	4	1	0.0000517	0.7500	0.6000
	3	4	1	2	5	0.0010302	0.2500	0.4000
	3	4	1	5	2	0.0000743	0.2500	0.5000
	3	4	2	1	5	0.0006648	0.5000	0.5000
	3	4	2	5	1	0.0000302	0.5000	0.6000
	3	4	5	1	2	0.0000171	0.2500	0.6000
	3	4	5	2	1	0.0000118	0.5000	0.7000
	3	5	1	2	4	0.0001184	0.5000	0.5000
	3	5	1	4	2	0.0000323	0.5000	0.6000
	3	5	2	1	4	0.0000824	0.7500	0.6000
	3	5	2	4	1	0.0000151	0.7500	0.7000
	3	5	4	1	2	0.0000126	0.5000	0.7000
	3	5	4	2	1	0.0000087	0.7500	0.8000
	4	1	2	3	5	0.0026345	0.2500	0.3000
	4	1	2	5	3	0.0003656	0.2500	0.4000
	4	1	3	2	5	0.0014281	0.5000	0.4000
	4	1	3	5	2	0.0001095	0.5000	0.5000
	4	1	5	2	3	0.0000743	0.2500	0.5000
	4	1	5	3	2	0.0000452	0.5000	0.6000
	4	2	1	3	5	0.0016989	0.5000	0.4000
	4	2	1	5	3	0.0002289	0.5000	0.5000
	4	2	3	1	5	0.0005821	0.5000	0.5000
	4	2	3	5	1	0.0000268	0.5000	0.6000
	4	2	5	1	3	0.0000323	0.5000	0.6000
	4	2	5	3	1	0.0000134	0.5000	0.7000
	4	3	1	2	5	0.0006418	0.5000	0.5000
	4	3	1	5	2	0.0000452	0.5000	0.6000
	4	3	2	1	5	0.0004113	0.7500	0.6000
	4	3	2	5	1	0.0000181	0.7500	0.7000

Tabelle 7.2 (Fortsetzung)

$m = 5$	$i_1$	$i_2$	$i_3$	$i_4$	$i_5$	$P_0(X_{(i_1)1} < \dots < X_{(i_5)1})$	$R_{5,1}(\mathbf{X})$	$\tilde{R}_{5,1}(\mathbf{X})$
	4	3	5	1	2	0.0000101	0.5000	0.7000
	4	3	5	2	1	0.0000069	0.7500	0.8000
	4	5	1	2	3	0.0000171	0.2500	0.6000
	4	5	1	3	2	0.0000101	0.5000	0.7000
	4	5	2	1	3	0.0000126	0.5000	0.7000
	4	5	2	3	1	0.0000051	0.5000	0.8000
	4	5	3	1	2	0.0000058	0.5000	0.8000
	4	5	3	2	1	0.0000039	0.7500	0.9000
	5	1	2	3	4	0.0001595	0.2500	0.4000
	5	1	2	4	3	0.0000750	0.5000	0.5000
	5	1	3	2	4	0.0000972	0.5000	0.5000
	5	1	3	4	2	0.0000268	0.5000	0.6000
	5	1	4	2	3	0.0000302	0.5000	0.6000
	5	1	4	3	2	0.0000181	0.7500	0.7000
	5	2	1	3	4	0.0001111	0.5000	0.5000
	5	2	1	4	3	0.0000517	0.7500	0.6000
	5	2	3	1	4	0.0000462	0.5000	0.6000
	5	2	3	4	1	0.0000083	0.5000	0.7000
	5	2	4	1	3	0.0000151	0.7500	0.7000
	5	2	4	3	1	0.0000061	0.7500	0.8000
	5	3	1	2	4	0.0000501	0.5000	0.6000
	5	3	1	4	2	0.0000134	0.5000	0.7000
	5	3	2	1	4	0.0000345	0.7500	0.7000
	5	3	2	4	1	0.0000061	0.7500	0.8000
	5	3	4	1	2	0.0000051	0.5000	0.8000
	5	3	4	2	1	0.0000035	0.7500	0.9000
	5	4	1	2	3	0.0000118	0.5000	0.7000
	5	4	1	3	2	0.0000069	0.7500	0.8000
	5	4	2	1	3	0.0000087	0.7500	0.8000
	5	4	2	3	1	0.0000035	0.7500	0.9000
	5	4	3	1	2	0.0000039	0.7500	0.9000
	5	4	3	2	1	0.0000027	1.0000	1.0000

Tabelle 7.3: Verteilungen von  $R_{m,1}(\mathbf{X})$  unter (a) und (b) für  $m = 2, \dots, 5$ .

	$r$	$P_0(R_{m,1}(\mathbf{X}) = r)$	$P_0(R_{m,1}(\mathbf{X}) \leq r)$
$m = 2$	0	0.8333333	0.8333333
	1	0.1666667	1.0000000
$m = 3$	0	0.6095238	0.6095238
	0.5	0.3809524	0.9904762
	1	0.0095238	1.0000000
$m = 4$	0	0.4028293	0.4028293
	0.3333	0.5231447	0.9259740
	0.6667	0.0737941	0.9997681
	1	0.0002319	1.0000000
$m = 5$	0	0.2450499	0.2450499
	0.25	0.5501239	0.7951739
	0.5	0.1963510	0.9915249
	0.75	0.0084724	0.9999973
	1	0.0000027	1.0000000

Damit erhalten wir

$$\begin{aligned}
 E_{1^*}[R_{m,n}(\mathbf{X})] &= E_{1^*}\left[\frac{1}{(m-1)n} \sum_{k=1}^n \sum_{i=1}^{m-1} J_{i,k}\right] \\
 &= \frac{1}{2}, \\
 \text{Var}_{1^*}[R_{m,n}(\mathbf{X})] &= \text{Var}_{1^*}\left[\frac{1}{(m-1)n} \sum_{k=1}^n \sum_{i=1}^{m-1} J_{i,k}\right] \\
 &= \frac{1}{(m-1)^2 n} \left[ \sum_{i=1}^{m-1} \text{Var}_{1^*}(J_{i,k}) + 2 \sum_{i=1}^{m-2} \text{Cov}_{1^*}(J_{i,k}, J_{i+1,k}) \right] \\
 &= \frac{1}{(m-1)^2 n} \left[ \frac{m-1}{4} - \frac{2(m-2)}{12} \right] \\
 &= \frac{m+1}{12(m-1)^2 n},
 \end{aligned}$$

wie in Satz 7.2 behauptet.

Wiederum können wir mit dem zentralen Grenzwertsatz die Verteilung von  $R_{m,n}(\mathbf{X})$  approximieren.

Auch unter der speziellen Alternative  $H_1^*$  wollen wir aber für kleine  $m$  und  $n$  die effektiven Verteilungen bestimmen. Da jeweils jede Anordnung

von  $X_{[1]1}, \dots, X_{[m]1}$  die gleiche Wahrscheinlichkeit ( $\frac{1}{m!}$ ) hat, brauchen wir nur die Anzahl der Anordnungen zu zählen, die einen bestimmten Wert von  $R_{m,1}(\mathbf{X})$  ergeben (vgl. etwa Tabelle 7.2). Wir erhalten die Verteilungen in Tabelle 7.4.

Tabelle 7.4: Verteilungen von  $R_{m,1}(\mathbf{X})$  unter der speziellen Alternative  $H_1^*$ , dass die Elemente von  $\mathbf{X}$  i.i.d. sind, für  $m = 2, \dots, 5$ .

	$r$	$P_{1^*}(R_{m,1}(\mathbf{X}) = r)$	$P_{1^*}(R_{m,1}(\mathbf{X}) \leq r)$
$m = 2$	0	0.5000000	0.5000000
	1	0.5000000	1.0000000
$m = 3$	0	0.1666667	0.1666667
	0.5	0.6666667	0.8333333
	1	0.1666667	1.0000000
$m = 4$	0	0.0416667	0.0416667
	0.3333	0.4583333	0.5000000
	0.6667	0.4583333	0.9583333
	1	0.0416667	1.0000000
$m = 5$	0	0.0083333	0.0083333
	0.25	0.2166667	0.2250000
	0.5	0.5500000	0.7750000
	0.75	0.2166667	0.9916667
	1	0.0083333	1.0000000

### 7.3 Minimax-Test

Mit Hilfe der Verteilungen unter der speziellen Alternative  $H_1^*$  können wir nun ein anderes Kriterium für die Verwerfung der Nullhypothese aufstellen. Wie bereits erwähnt, ist es nicht unbedingt wünschenswert, für verschiedene  $n$  die Fehlerwahrscheinlichkeit erster Art festzuhalten, da die RSS-Verfahren auch bei leichten Verletzungen der Voraussetzungen noch nützlich sind. Wenn wir nur entscheiden wollen, ob eher die Nullhypothese oder die spezielle Alternative  $H_1^*$  zutrifft, können wir leicht Minimax-Tests (innerhalb einer ausgewählten Familie nichtrandomisierter Tests) konstruieren. Diese ergeben asymptotisch gleiche Fehlerwahrscheinlichkeiten erster und zweiter Art, für kleinere  $n$  zumindest solche ähnlicher Grössenordnung. Im wesentlichen behandeln wir also  $H_0$  und  $H_1^*$  gleich.

Da  $E_0 [R_{m,n}(\mathbf{X})] < E_{1^*} [R_{m,n}(\mathbf{X})] = \frac{1}{2}$  gilt, suchen wir (einseitige) Tests, die  $H_0$  zugunsten von  $H_1^*$  verwerfen, falls  $R_{m,n}(\mathbf{X}) > r$ . Die Minimax-Bedingung verlangt, dass wir

$$\max \{P_0 [R_{m,n}(\mathbf{X}) > r], P_{1^*} [R_{m,n}(\mathbf{X}) \leq r]\} \quad (7.13)$$

minimieren.

Weil  $P_0 [R_{m,n}(\mathbf{X}) > r]$  monoton fallend in  $r$  und  $P_{1^*} [R_{m,n}(\mathbf{X}) \leq r]$  monoton steigend in  $r$  ist, ist (7.13) sicher minimal, wenn die beiden Werte gleich sind. Dies können wir wegen der Sprungstellen für endliche  $n$  im allgemeinen nicht fordern, wohl aber für  $n \rightarrow \infty$ . Dann können wir aber auch die Verteilungen von  $R_{m,n}(\mathbf{X})$  durch Normalverteilungen approximieren, und es gilt asymptotisch:

$$\begin{aligned} P_0 [R_{m,n}(\mathbf{X}) > r] &= P_{1^*} [R_{m,n}(\mathbf{X}) \leq r] \\ &\stackrel{n \rightarrow \infty}{\Leftrightarrow} 1 - \Phi \left( \frac{r - E_0 [R_{m,1}(\mathbf{X})]}{\sqrt{\frac{1}{n} \text{Var}_0 [R_{m,1}(\mathbf{X})]}} \right) = \Phi \left( \frac{r - E_{1^*} [R_{m,1}(\mathbf{X})]}{\sqrt{\frac{1}{n} \text{Var}_{1^*} [R_{m,1}(\mathbf{X})]}} \right) \\ &\stackrel{(7.11), (7.12)}{\Leftrightarrow} \frac{E_0 [R_{m,1}(\mathbf{X})] - r}{\sqrt{\text{Var}_0 [R_{m,1}(\mathbf{X})]}} = \frac{r - \frac{1}{2}}{\sqrt{\frac{m+1}{12(m-1)^2}}} \\ &\Leftrightarrow r = \frac{\sqrt{\frac{m+1}{12(m-1)^2}} E_0 [R_{m,1}(\mathbf{X})] + \frac{1}{2} \sqrt{\text{Var}_0 [R_{m,1}(\mathbf{X})]}}{\sqrt{\frac{m+1}{12(m-1)^2}} + \sqrt{\text{Var}_0 [R_{m,1}(\mathbf{X})]}} \end{aligned}$$

$r$  kann mit (7.1) und (7.2) bzw. mit den Angaben aus Tabelle 7.1 bestimmt werden. Dies ergibt die Werte  $r = r_{m,\infty}^{\text{Minimax}}$  in Tabelle 7.5.

Tabelle 7.5: Kritische Werte für den Minimax-Test mit  $R_{m,n}(\mathbf{X})$  (asymptotisch für  $n \rightarrow \infty$ ).

$m$	$r_{m,\infty}^{\text{Minimax}}$
2	0.3090
3	0.3406
4	0.3578
5	0.3693

Mit den oben berechneten Verteilungen von  $R_{m,n}(\mathbf{X})$  unter  $H_0$  und  $H_1^*$  lassen sich selbstverständlich auch genaue Verwerfungsgrenzen für kleine  $n$  bestimmen:



**Beispiel 7.1** Wir wollen im Fall  $m = 3$ ,  $n = 5$  mit dem beschriebenen Minimax-Test zwischen  $H_0$  und  $H_1^*$  entscheiden. Mittels Faltung erhalten wir aus den Tabellen 7.3 und 7.4 die Verteilungen von  $R_{3,5}(\mathbf{X})$ . Diese sind in Tabelle 7.6 angegeben.

Tabelle 7.6: Verteilungen von  $R_{3,5}(\mathbf{X})$  unter  $H_0$  und  $H_1^*$ .

$r$	$P_0(R_{3,5}(\mathbf{X}) = r)$	$P_0(R_{3,5}(\mathbf{X}) > r)$	$P_{1^*}(R_{3,5}(\mathbf{X}) = r)$	$P_{1^*}(R_{3,5}(\mathbf{X}) \leq r)$
0.0	0.084130	0.915870	0.000129	0.000129
0.1	0.262908	0.652962	0.002572	0.002701
0.2	0.335207	0.317754	0.021219	0.023920
0.3	0.221828	0.095926	0.092593	0.116512
0.4	0.079797	0.016129	0.227623	0.344136
0.5	0.014827	0.001302	0.311728	0.655864
0.6	0.001247	0.000055	0.227623	0.883488
0.7	0.000054	0.000001	0.092593	0.976080
0.8	0.000001	0.000000	0.021219	0.997299
0.9	0.000000	0.000000	0.002572	0.999871
1.0	0.000000	0.000000	0.000129	1.000000

(7.13) wird für  $r = 0.3 = r_{3,5}^{\text{Minimax}}$  minimal. Somit wird man  $H_0$  verwerfen, falls  $R_{3,5}(\mathbf{X})$  einen Wert von mindestens 0.4 annimmt – man geht dann davon aus, dass das RSS-Verfahren wenig Nutzen bringt. Die Fehlerwahrscheinlichkeiten betragen 9.6% bzw. 11.7%. In diesem Beispiel ergibt sich mit der asymptotischen Grenze  $r_{3,\infty}^{\text{Minimax}}$  übrigens die gleiche Entscheidungsregel.

Schliesslich wollen wir zum Vergleich noch die Fehlerwahrscheinlichkeiten unter Verwendung der Normalverteilungsapproximation bestimmen. Dabei verwenden wir den Wert 0.35 statt 0.3, um der diskreten Verteilung Rechnung zu tragen.

$$\begin{aligned}
 P_0(R_{3,5}(\mathbf{X}) > 0.3) &\approx 1 - \Phi\left(\frac{0.35 - E_0[R_{3,1}(\mathbf{X})]}{\sqrt{\frac{1}{5}\text{Var}_0[R_{3,1}(\mathbf{X})]}}\right) \\
 &\approx 1 - \Phi(1.318) \\
 &\approx 0.0938 \quad (\text{statt } 0.0959) \\
 P_{1^*}(R_{3,5}(\mathbf{X}) \leq 0.3) &\approx \Phi\left(\frac{0.35 - E_{1^*}[R_{3,1}(\mathbf{X})]}{\sqrt{\frac{1}{5}\text{Var}_{1^*}[R_{3,1}(\mathbf{X})]}}\right) \\
 &\approx \Phi(-1.162) \\
 &\approx 0.1226 \quad (\text{statt } 0.1165)
 \end{aligned}$$

Die Abweichungen von den exakten Werten sind in Anbetracht der kleinen Anzahl Durchgänge relativ bescheiden.

## Kapitel 8

### Statistik $\tilde{R}_{m,n}(\mathbf{X})$

Die bisher verwendete Statistik hat den Vorteil, dass jedes  $X_{[i]k}$  höchstens zweimal für einen Vergleich benützt wird und dementsprechend in den Berechnungen nur wenige Abhängigkeiten berücksichtigt werden müssen. Als Nachteil zeigen sich aber bei kleinen Umfängen in den Verteilungsfunktionen grosse Sprünge, die für die Konstruktion von Toleranzbereichen und Tests unerwünscht sind. Wir betrachten deshalb folgende Statistik, die zu tendenziell kleineren Sprüngen in der Verteilungsfunktion führt, da mehr verschiedene Werte angenommen werden:

$$\tilde{R}_{m,n}(\mathbf{X}) = \frac{1}{\binom{m}{2}n} \sum_{k=1}^n \sum_{i=1}^{m-1} \sum_{j=i+1}^m \mathbf{I}(X_{[i]k} > X_{[j]k})$$

Nach wie vor vergleichen wir also nur Elemente, die zum gleichen (effektiven oder „künstlichen“) Durchgang gehören. Innerhalb jedes Durchgangs betrachten wir aber nicht mehr nur benachbarte Elemente, sondern jedes mögliche Paar von Elementen. (Für  $m = 2$  ergibt dies natürlich nichts Neues.)

### 8.1 Verhalten unter der Nullhypothese

Im Prinzip können wir unter der üblichen Nullhypothese, dass die Bedingungen (a) und (b) erfüllt seien, mit den gleichen Methoden wie für  $R_{m,n}(\mathbf{X})$  Erwartungswerte, Varianzen sowie für bestimmte  $m$  exakte Verteilungen berechnen.

**Satz 8.1** *Für den Erwartungswert erhalten wir unter  $H_0$*

$$E_0 [\tilde{R}_{m,n}(\mathbf{X})] = \frac{1}{\binom{m}{2}} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{i(j-i) \binom{m}{i} \binom{m}{j}}{(i+j) \binom{2m}{i+j}}. \quad (8.1)$$

**Beweis:** Weitgehend analog zum entsprechenden Teil des Beweises von Satz 7.1, wobei man Indikatorfunktionen

$$\tilde{J}_{i,j,k} = \mathbb{I}(X_{(i)k} > X_{(j)k}) \quad (i < j)$$

betrachtet.

Die direkte Berechnung der Varianz würde für diese Statistik noch um einiges komplizierter als für  $R_{m,n}(\mathbf{X})$ . Wir wählen deshalb einen andern Zugang, der uns allerdings nur die Varianz für  $m = 2, \dots, 5$  liefert.

Wie die Verteilungen von  $R_{m,1}(\mathbf{X})$  sind auch diejenigen von  $\tilde{R}_{m,1}(\mathbf{X})$  durch  $P_0(X_{(i_1)1} < \dots < X_{(i_m)1})$  bestimmt. Wir haben schon gesehen, wie sich diese Wahrscheinlichkeiten berechnen lassen. Daraus und aus den Werten, die  $\tilde{R}_{m,1}(\mathbf{X})$  für bestimmte Permutationen von  $X_{(1)1}, \dots, X_{(m)1}$  annimmt (vgl. Tabelle 7.2), bestimmen wir die Verteilungen von  $\tilde{R}_{m,1}(\mathbf{X})$  unter  $H_0$ . Diese Verteilungen sind in Tabelle 8.1 zusammengestellt.

Weil wir die Verteilungen nun kennen, können wir natürlich daraus die Varianzen von  $\tilde{R}_{m,1}(\mathbf{X})$  für  $m = 2, \dots, 5$  berechnen. Die Erwartungswerte von  $\tilde{R}_{m,n}(\mathbf{X})$  für  $n > 1$  entsprechen denjenigen für  $n = 1$  in Tabelle 8.2, die Varianzen für  $n > 1$  erhält man wiederum mittels Division durch  $n$  aus denjenigen für  $n = 1$ . Somit haben wir die nötigen Angaben, um die Verteilung von  $\tilde{R}_{m,n}(\mathbf{X})$  für grosse  $n$  durch eine Normalverteilung zu approximieren – wiederum können wir den zentralen Grenzwertsatz verwenden, da  $\tilde{R}_{m,n}(\mathbf{X})$  die Summe von  $n$  i.i.d. Zufallsvariablen mit Verteilung wie  $\tilde{R}_{m,1}(\mathbf{X})$  ist.

Zu bemerken ist noch, dass bei dieser Statistik (im Gegensatz zu  $R_{m,n}(\mathbf{X})$ ) der Erwartungswert unter der Nullhypothese mit steigendem  $m$  abnimmt. Während sich die mittels  $R_{m,n}(\mathbf{X})$  verglichenen Elemente für grosse  $m$  in ihrer Verteilung kaum noch unterscheiden, vergleicht man mit Hilfe von  $\tilde{R}_{m,n}(\mathbf{X})$  auch für grosse  $m$  noch Elemente, deren Verteilungen sich wesentlich unterscheiden.

## 8.2 Verhalten unter der speziellen Alternative

Unter der speziellen Alternative  $H_1^*$  ist die Herleitung der Formeln für den Erwartungswert und die Varianz von  $\tilde{R}_{m,n}(\mathbf{X})$  im Bereich des Zumutbaren und wird durch recht einfache Ausdrücke belohnt:

**Satz 8.2** *Unter  $H_1^*$  gilt*

$$E_{1^*} [\tilde{R}_{m,n}(\mathbf{X})] = \frac{1}{2}, \quad (8.2)$$

Tabelle 8.1: Verteilungen von  $\tilde{R}_{m,1}(\mathbf{X})$  unter (a) und (b) für  $m = 2, \dots, 5$ .

	$r$	$P_0(\tilde{R}_{m,1}(\mathbf{X}) = r)$	$P_0(\tilde{R}_{m,1}(\mathbf{X}) \leq r)$
$m = 2$	0	0.8333333	0.8333333
	1	0.1666667	1.0000000
$m = 3$	0	0.6095238	0.6095238
	0.3333	0.3404762	0.9500000
	0.6667	0.0404762	0.9904762
	1	0.0095238	1.0000000
$m = 4$	0	0.4028293	0.4028293
	0.1667	0.4164895	0.8193188
	0.3333	0.1377601	0.9570789
	0.5	0.0366914	0.9937703
	0.6667	0.0048930	0.9986632
	0.8333	0.0011048	0.9997681
	1	0.0002319	1.0000000
$m = 5$	0	0.2450499	0.2450499
	0.1	0.3908159	0.6358659
	0.2	0.2372620	0.8731279
	0.3	0.0934234	0.9665513
	0.4	0.0254875	0.9920388
	0.5	0.0060582	0.9980970
	0.6	0.0015867	0.9996837
	0.7	0.0002394	0.9999231
	0.8	0.0000594	0.9999825
	0.9	0.0000148	0.9999973
1	0.0000027	1.0000000	

Tabelle 8.2: Erwartungswert und Varianz von  $\tilde{R}_{m,1}(\mathbf{X})$  unter (a) und (b) ( $m = 2, \dots, 5$ ).

$m$	$E_0[\tilde{R}_{m,1}(\mathbf{X})]$	$\text{Var}_0[\tilde{R}_{m,1}(\mathbf{X})]$
2	0.1666667	0.13888889
3	0.1500000	0.04284392
4	0.1380952	0.02015220
5	0.1289683	0.01150770

$$\text{Var}_{1^*} [\tilde{R}_{m,n}(\mathbf{X})] = \frac{2m+5}{18m(m-1)n}. \quad (8.3)$$

**Beweis:** Die Ideen für den direkten Beweis entsprechen denjenigen bei Satz 7.2. Die Behauptungen für  $n = 1$  werden sich ausserdem als Spezialfall von Satz 9.3 ergeben ( $\tilde{R}_{m,1}(\mathbf{X}) = R_{m,1}^*(\mathbf{X})$ ).

Wiederum berechnen wir auch die Verteilungen von  $\tilde{R}_{m,1}(\mathbf{X})$  für  $m = 2, \dots, 5$  (Tabelle 8.3).

Tabelle 8.3: Verteilungen von  $\tilde{R}_{m,1}(\mathbf{X})$  unter der speziellen Alternative  $H_1^*$ , dass die Elemente von  $\mathbf{X}$  i.i.d. sind, für  $m = 2, \dots, 5$ .

	$r$	$P_{1^*}(\tilde{R}_{m,1}(\mathbf{X}) = r)$	$P_{1^*}(\tilde{R}_{m,1}(\mathbf{X}) \leq r)$
$m = 2$	0	0.5000000	0.5000000
	1	0.5000000	1.0000000
$m = 3$	0	0.1666667	0.1666667
	0.3333	0.3333333	0.5000000
	0.6667	0.3333333	0.8333333
	1	0.1666667	1.0000000
$m = 4$	0	0.0416667	0.0416667
	0.1667	0.1250000	0.1666667
	0.3333	0.2083333	0.3750000
	0.5	0.2500000	0.6250000
	0.6667	0.2083333	0.8333333
	0.8333	0.1250000	0.9583333
	1	0.0416667	1.0000000
$m = 5$	0	0.0083333	0.0083333
	0.1	0.0333333	0.0416667
	0.2	0.0750000	0.1166667
	0.3	0.1250000	0.2416667
	0.4	0.1666667	0.4083333
	0.5	0.1833333	0.5916667
	0.6	0.1666667	0.7583333
	0.7	0.1250000	0.8833333
	0.8	0.0750000	0.9583333
	0.9	0.0333333	0.9916667
	1	0.0083333	1.0000000

### 8.3 Minimax-Test

In Tabelle 8.4 sind asymptotische kritische Werte für einen Minimax-Test analog zu demjenigen mit  $R_{m,n}(\mathbf{X})$  angegeben. Sie berechnen sich nach der Formel

$$\tilde{r}_{m,\infty}^{\text{Minimax}} = \frac{\sqrt{\frac{2m+5}{18m(m-1)}} E_0 [\tilde{R}_{m,1}(\mathbf{X})] + \frac{1}{2} \sqrt{\text{Var}_0 [\tilde{R}_{m,1}(\mathbf{X})]}}{\sqrt{\frac{2m+5}{18m(m-1)}} + \sqrt{\text{Var}_0 [\tilde{R}_{m,1}(\mathbf{X})]}}.$$

Tabelle 8.4: Kritische Werte für den Minimax-Test mit  $\tilde{R}_{m,n}(\mathbf{X})$  (asymptotisch für  $n \rightarrow \infty$ ).

$m$	$\tilde{r}_{m,\infty}^{\text{Minimax}}$
2	0.3090
3	0.2877
4	0.2708
5	0.2568

**Beispiel 8.1** Wir nehmen nochmals die Situation von Beispiel 7.1 auf ( $m = 3$ ,  $n = 5$ ), wollen nun aber einen Minimax-Test durchführen, der auf  $\tilde{R}_{3,5}(\mathbf{X})$  basiert. Aus Tabelle 8.5 finden wir  $\tilde{r}_{3,5}^{\text{Minimax}} = 0.2667$  und verwerfen  $H_0$  zugunsten von  $H_1^*$ , falls  $\tilde{R}_{3,5}(\mathbf{X}) \geq 0.3333$ . Da wir mehr Information verwenden als mit der Statistik  $R_{3,5}(\mathbf{X})$ , erhalten wir kleinere Fehlerwahrscheinlichkeiten von 6.3% bzw. 8.1%. Wiederum erhalten wir die gleiche Entscheidungsregel, wenn wir  $\tilde{r}_{3,\infty}^{\text{Minimax}}$  anstelle von  $\tilde{r}_{3,5}^{\text{Minimax}}$  verwenden.

Tabelle 8.5: Verteilungen von  $\tilde{R}_{3,5}(\mathbf{X})$  unter  $H_0$  und  $H_1^*$ .

$r$	$P_0(\tilde{R}_{3,5}(\mathbf{X}) = r)$	$P_0(\tilde{R}_{3,5}(\mathbf{X}) > r)$	$P_{1^*}(\tilde{R}_{3,5}(\mathbf{X}) = r)$	$P_{1^*}(\tilde{R}_{3,5}(\mathbf{X}) \leq r)$
0	0.084130	0.915870	0.000129	0.000129
0.0667	0.234974	0.680896	0.001286	0.001415
0.1333	0.290444	0.390452	0.006430	0.007845
0.2	0.215624	0.174828	0.021219	0.029064
0.2667	0.111648	0.063180	0.051440	0.080504
0.3333	0.044319	0.018862	0.096708	0.177212
0.4	0.014152	0.004709	0.145319	0.322531
0.4667	0.003714	0.000995	0.177469	0.500000
0.5333	0.000817	0.000178	0.177469	0.677469
0.6	0.000151	0.000027	0.145319	0.822788
0.6667	0.000023	0.000003	0.096708	0.919496
0.7333	0.000003	0.000000	0.051440	0.970936
0.8	0.000000	0.000000	0.021219	0.992155
0.8667	0.000000	0.000000	0.006430	0.998585
0.9333	0.000000	0.000000	0.001286	0.999871
1	0.000000	0.000000	0.000129	1.000000

## Kapitel 9

### Statistik $R_{m,n}^*(\mathbf{X})$

Eine naheliegende Modifikation der Statistik  $\tilde{R}_{m,n}(\mathbf{X})$  besteht darin, die Vergleiche nicht mehr auf die einzelnen Durchgänge zu beschränken, sondern jedes  $X_{[i]k}$  mit jedem  $X_{[j]\ell}$  ( $i < j$ ) zu vergleichen. Dies ergibt die Statistik

$$R_{m,n}^*(\mathbf{X}) = \frac{1}{\binom{m}{2}n^2} \sum_{k=1}^n \sum_{\ell=1}^n \sum_{i=1}^{m-1} \sum_{j=i+1}^m \mathbf{I}(X_{[i]k} > X_{[j]\ell}).$$

#### 9.1 Verhalten unter der Nullhypothese

**Satz 9.1** Für die so definierte Statistik gilt unter den Bedingungen (a) und (b)

$$\mathbb{E}_0 [R_{m,n}^*(\mathbf{X})] = \frac{1}{\binom{m}{2}} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{i(j-i) \binom{m}{i} \binom{m}{j}}{(i+j) \binom{2m}{i+j}}. \quad (9.1)$$

**Beweis:**  $X_{(j)\ell}$  hat die gleiche Verteilung wie  $X_{(j)k}$ , und beide sind für  $i \neq j$  unabhängig von  $X_{(i)k}$ . Daher gilt

$$\begin{aligned} \mathbb{E}_0 [R_{m,n}^*(\mathbf{X})] &= \mathbb{E}_0 \left[ \frac{1}{\binom{m}{2}n^2} \sum_{k=1}^n \sum_{\ell=1}^n \sum_{i=1}^{m-1} \sum_{j=i+1}^m \mathbf{I}(X_{(i)k} > X_{(j)\ell}) \right] \\ &= \mathbb{E}_0 \left[ \frac{1}{\binom{m}{2}n} \sum_{k=1}^n \sum_{i=1}^{m-1} \sum_{j=i+1}^m \mathbf{I}(X_{(i)k} > X_{(j)k}) \right] \\ &= \mathbb{E}_0 [\tilde{R}_{m,n}(\mathbf{X})], \end{aligned}$$

und die Behauptung folgt aus Satz 8.1.



Die Herleitung eines Ausdrucks für die Varianz wäre nochmals komplizierter als für  $\tilde{R}_{m,n}(\mathbf{X})$ . Wir unternehmen deshalb vorerst eine empirische Untersuchung. Erschwerend ist dabei, dass es nicht (wie für  $R_{m,n}(\mathbf{X})$  und  $\tilde{R}_{m,n}(\mathbf{X})$ ) genügt, den Fall  $n = 1$  zu betrachten. Aufgrund der Vergleiche über die Grenzen der Durchgänge hinweg ist nämlich  $\text{Var}_0[R_{m,n}^*(\mathbf{X})]$  nicht von der Form  $\frac{cm}{n}$ .

Wir schätzen die gesuchte Varianz aufgrund von Simulationen separat für  $m = 2, \dots, 5$  und  $n = 1, \dots, 10$ . Die Werte in Tabelle 9.1 sind solche geschätzte Varianzen, basierend auf je 100 000 Simulationen, in denen jeweils auf  $m^2 n$  i.i.d. „Zufallszahlen“ das RSS-Auswahlverfahren angewandt und  $R_{m,n}^*(\mathbf{X})$  berechnet wurde.

Tabelle 9.1: Geschätzte Varianz von  $R_{m,n}^*(\mathbf{X})$  unter (a) und (b).

$n$	$m$			
	2	3	4	5
1	0.140	0.0429	0.0202	0.0115
2	0.0545	0.0169	0.00805	0.00460
3	0.0328	0.0103	0.00493	0.00285
4	0.0235	0.00735	0.00357	0.00205
5	0.0180	0.00570	0.00272	0.00157
6	0.0146	0.00465	0.00224	0.00130
7	0.0124	0.00392	0.00189	0.00110
8	0.0106	0.00343	0.00163	0.000944
9	0.00942	0.00298	0.00144	0.000836
10	0.00839	0.00267	0.00130	0.000745

Diese Schätzungen wurden mit den Resultaten einer zweiten, gleichen Reihe von Simulationen verglichen. Aufgrund dieses Vergleichs ist anzunehmen, dass jeweils etwa zwei der drei angegebenen signifikanten Stellen brauchbar sind.

Ausser für  $n = 1$  sind die so geschätzten Varianzen von  $R_{m,n}^*(\mathbf{X})$  alle kleiner als die entsprechenden (exakten) Varianzen von  $\tilde{R}_{m,n}(\mathbf{X})$ , die man aus Tabelle 8.2 berechnen kann.

Nun wollen wir doch noch versuchen, uns auf theoretischem Weg an die Varianzen heranzutasten:

**Satz 9.2** *Es gibt für jedes  $m \geq 2$  Konstanten  $a_{1,m}, a_{2,m}, a_{3,m}$ , so dass*

$$\text{Var}_0 [R_{m,n}^*(\mathbf{X})] = \frac{a_{1,m}}{n} + \frac{a_{2,m}}{n^2} + \frac{a_{3,m}}{n^3}. \quad (9.2)$$

**Beweis:** Wir verwenden die abgekürzte Schreibweise

$$A_{k,\ell} = \frac{1}{\binom{m}{2}} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \mathbf{I}(X_{(i)k} > X_{(j)\ell}),$$

so dass

$$\begin{aligned} & \text{Var}_0 [R_{m,n}^*(\mathbf{X})] \\ &= \frac{1}{n^4} \text{Var}_0 \left[ \sum_{k=1}^n \sum_{\ell=1}^n A_{k,\ell} \right] \\ &= \frac{1}{n^4} \sum_{k=1}^n \sum_{\ell=1}^n \sum_{\tilde{k}=1}^n \sum_{\tilde{\ell}=1}^n \text{Cov}_0 [A_{k,\ell}, A_{\tilde{k},\tilde{\ell}}] \\ &= \frac{1}{n^4} \left[ n(n-1)(n-2)(n-3) \text{Cov}_0 [A_{1,2}, A_{3,4}] \right. \\ &\quad + n(n-1)(n-2) \left( \text{Cov}_0 [A_{1,2}, A_{3,3}] + \text{Cov}_0 [A_{1,2}, A_{3,2}] \right. \\ &\quad \quad + \text{Cov}_0 [A_{1,2}, A_{3,1}] + \text{Cov}_0 [A_{1,2}, A_{2,3}] \\ &\quad \quad \left. \left. + \text{Cov}_0 [A_{1,2}, A_{1,3}] + \text{Cov}_0 [A_{1,1}, A_{2,3}] \right) \right. \\ &\quad + n(n-1) \left( \text{Cov}_0 [A_{1,1}, A_{2,2}] + \text{Cov}_0 [A_{1,2}, A_{1,2}] + \text{Cov}_0 [A_{1,2}, A_{2,1}] \right. \\ &\quad \quad + \text{Cov}_0 [A_{1,1}, A_{1,2}] + \text{Cov}_0 [A_{1,1}, A_{2,1}] \\ &\quad \quad \left. \left. + \text{Cov}_0 [A_{1,2}, A_{1,1}] + \text{Cov}_0 [A_{2,1}, A_{1,1}] \right) \right. \\ &\quad \left. + n \text{Cov}_0 [A_{1,1}, A_{1,1}] \right] \\ &= \frac{1}{n^4} \left[ (n^3 - 3n^2 + 2n)c_{1,m} + (n^2 - n)c_{2,m} + nc_{3,m} \right] \\ &= \frac{c_{1,m}}{n} + \frac{-3c_{1,m} + c_{2,m}}{n^2} + \frac{2c_{1,m} - c_{2,m} + c_{3,m}}{n^3}. \end{aligned}$$

Dabei haben wir die Summe über alle  $(k, \ell, \tilde{k}, \tilde{\ell})$  nach der Anzahl verschiedener Indices zerlegt und verwendet, dass  $\text{Cov}_0 [A_{1,2}, A_{3,4}] = 0$ . Man beachte, dass die Umformungen auch für  $n < 4$  gültig bleiben, da vor den Kovarianzen, die für ein bestimmtes  $n$  gar nicht vorkommen, ein Faktor 0 steht. Die Behauptung folgt mit passender Definition der Konstanten  $a_{i,m}$ .

Im Prinzip könnte man nun für ein festes  $m$  die Varianzen von  $R_{m,n}^*(\mathbf{X})$  für drei Werte von  $n$  berechnen ( $n = 1$  ist von  $\tilde{R}_{m,n}(\mathbf{X})$  her bekannt, nahe- liegenderweise würde man noch  $n = 2, 3$  betrachten) und könnte daraus die Konstanten  $a_{i,m}$  bestimmen. Dies ergäbe die Varianz für jedes  $n$ . Die exakte Berechnung der dazu benötigten Varianzen ist allerdings nicht wesentlich weniger aufwendig als im allgemeinen Fall, so dass uns dieser Ansatz auch nicht

weiterbringt. Eine Schätzung der  $a_{i,m}$  mittels Regression aufgrund der simulierten Werte in Tabelle 9.1 scheint nicht ratsam, da diese Werte doch recht ungenau sind. Satz 9.2 liefert uns also letztlich nicht viel mehr als die Aussage, dass unter  $H_0$  für  $n \rightarrow \infty$  die Varianz von  $R_{m,n}^*(\mathbf{X})$  von der Ordnung  $\frac{1}{n}$  ist.

Leider lässt sich für  $R_{m,n}^*(\mathbf{X})$  auch mit ungefährender Kenntnis der Varianzen nicht ohne weiteres eine Aussage über asymptotische Verteilungen machen – der zentrale Grenzwertsatz in seiner klassischen Form genügt nicht mehr, da sich die Statistik nicht in  $n$  i.i.d. Zufallsvariablen zerlegen lässt. Dennoch kann man vermuten, dass die Verteilung asymptotisch einer Normalverteilung entspricht, mit Erwartungswert gemäss Satz 9.1 und Varianz  $\frac{c_{1,m}}{n}$ , mit  $c_{1,m}$  wie im Beweis von Satz 9.2. Diese Vermutung wird durch Simulationen gestützt, wobei die Konvergenz wegen der Schiefe der Verteilung relativ langsam erfolgt.

## 9.2 Verhalten unter der speziellen Alternative

**Satz 9.3** *Unter der speziellen Alternative, dass die  $X_{[i]k}$  i.i.d. sind, gilt*

$$E_{1^*} [R_{m,n}^*(\mathbf{X})] = \frac{1}{2}, \quad (9.3)$$

$$\text{Var}_{1^*} [R_{m,n}^*(\mathbf{X})] = \frac{2mn + 2n + 3}{18m(m-1)n^2}. \quad (9.4)$$

**Beweis:** Da die  $X_{[i]k}$  i.i.d. sind, ist  $E_{1^*} [\mathbf{I}(X_{[i]k} > X_{[j]\ell})] = \frac{1}{2}$ . Somit ist (9.3) trivial. Die Varianz schreiben wir als

$$\begin{aligned} & \text{Var}_{1^*} [R_{m,n}^*(\mathbf{X})] \\ &= \frac{1}{\binom{m}{2}^2 n^4} \text{Var}_{1^*} \left[ \sum_{k=1}^n \sum_{\ell=1}^n \sum_{i=1}^{m-1} \sum_{j=i+1}^m \mathbf{I}(X_{[i]k} > X_{[j]\ell}) \right] \\ &= \frac{1}{\binom{m}{2}^2 n^4} \left( \sum_{i < j} \text{Var}_{1^*} \left[ \sum_{k=1}^n \sum_{\ell=1}^n \mathbf{I}(X_{[i]k} > X_{[j]\ell}) \right] \right. \\ & \quad \left. + \sum_{\substack{i < j, \tilde{i} < \tilde{j} \\ i \neq \tilde{i} \text{ oder } j \neq \tilde{j}}} \text{Cov}_{1^*} \left[ \sum_{k=1}^n \sum_{\ell=1}^n \mathbf{I}(X_{[i]k} > X_{[j]\ell}), \sum_{\tilde{k}=1}^n \sum_{\tilde{\ell}=1}^n \mathbf{I}(X_{[\tilde{i}]\tilde{k}} > X_{[\tilde{j}]\tilde{\ell}}) \right] \right), \end{aligned} \quad (9.5)$$

wobei hier wie im folgenden  $i, \tilde{i}, j$  und  $\tilde{j}$  immer von 1 bis  $m$  laufen und *zusätzlich* die angegebenen Summationsbedingungen erfüllen sollen.

Für die Summanden der ersten Summe in (9.5) finden wir (mit  $i < j$ )

$$\begin{aligned}
& \text{Var}_{1^*} \left[ \sum_{k=1}^n \sum_{\ell=1}^n \mathbf{I}(X_{[i]k} > X_{[j]\ell}) \right] \\
&= \mathbf{E}_{1^*} \left[ \sum_{k=1}^n \sum_{\ell=1}^n \sum_{\tilde{k}=1}^n \sum_{\tilde{\ell}=1}^n \mathbf{I}(X_{[i]k} > X_{[j]\ell}) \mathbf{I}(X_{[i]\tilde{k}} > X_{[j]\tilde{\ell}}) \right] \\
&\quad - \left( \mathbf{E}_{1^*} \left[ \sum_{k=1}^n \sum_{\ell=1}^n \mathbf{I}(X_{[i]k} > X_{[j]\ell}) \right] \right)^2 \\
&\stackrel{(*)}{=} n^2(n-1)^2 \cdot \frac{1}{4} + 2n^2(n-1) \cdot \frac{1}{3} + n^2 \cdot \frac{1}{2} - \left( \frac{n^2}{2} \right)^2 \\
&= \frac{n^4}{4} - \frac{2n^3}{4} + \frac{n^2}{4} + \frac{2n^3}{3} - \frac{2n^2}{3} + \frac{n^2}{2} - \frac{n^4}{4} \\
&= \frac{n^3}{6} + \frac{n^2}{12}. \tag{9.6}
\end{aligned}$$

Dabei steht nach dem (\*) der erste Summand für die  $n^2(n-1)^2$  Fälle mit  $k \neq \tilde{k}$  und  $\ell \neq \tilde{\ell}$ , der zweite für diejenigen mit genau einer Ungleichung und einer Gleichung und der dritte für diejenigen mit  $k = \tilde{k}$  und  $\ell = \tilde{\ell}$ .

Die zweite Summe in (9.5) zerlegen wir mittels folgender Partition der Indexmenge:

$$\begin{aligned}
& \{(i, j, \tilde{i}, \tilde{j}) : i < j, \tilde{i} < \tilde{j}; i \neq \tilde{i} \text{ oder } j \neq \tilde{j}\} \\
&= \{(i, j, \tilde{i}, \tilde{j}) : i < j, \tilde{i} < \tilde{j}, i \neq \tilde{i}, j \neq \tilde{j}, i \neq \tilde{j}, \tilde{i} \neq j\} \\
&\quad \cup \{(i, j, i, \tilde{j}) : i < j, i < \tilde{j}, j \neq \tilde{j}\} \cup \{(i, j, \tilde{i}, j) : i < j, \tilde{i} < j, i \neq \tilde{i}\} \\
&\quad \cup \{(i, j, j, \tilde{j}) : i < j < \tilde{j}\} \cup \{(i, j, \tilde{i}, i) : \tilde{i} < i < j\} \tag{9.7}
\end{aligned}$$

Wir berechnen nun die gesuchte Kovarianz für die verschiedenen Fälle:

- Für  $i < j, \tilde{i} < \tilde{j}, i \neq \tilde{i}, j \neq \tilde{j}, i \neq \tilde{j}, \tilde{i} \neq j$  gilt

$$\text{Cov}_{1^*} \left[ \sum_{k=1}^n \sum_{\ell=1}^n \mathbf{I}(X_{[i]k} > X_{[j]\ell}), \sum_{\tilde{k}=1}^n \sum_{\tilde{\ell}=1}^n \mathbf{I}(X_{[i]\tilde{k}} > X_{[j]\tilde{\ell}}) \right] = 0,$$

da kein  $X_{[i]k}$  in beiden Summen vorkommt und die verschiedenen  $X_{[i]k}$  unabhängig sind.

- Für  $i < j, i < \tilde{j}, j \neq \tilde{j}$  gilt

$$\begin{aligned} & \text{Cov}_{1^*} \left[ \sum_{k=1}^n \sum_{\ell=1}^n \mathbf{I}(X_{[i]k} > X_{[j]\ell}), \sum_{\tilde{k}=1}^n \sum_{\tilde{\ell}=1}^n \mathbf{I}(X_{[i]\tilde{k}} > X_{[\tilde{j}]\tilde{\ell}}) \right] \\ &= \mathbf{E}_{1^*} \left[ \sum_{k=1}^n \sum_{\ell=1}^n \sum_{\tilde{k}=1}^n \sum_{\tilde{\ell}=1}^n \mathbf{I}(X_{[i]k} > X_{[j]\ell}) \mathbf{I}(X_{[i]\tilde{k}} > X_{[\tilde{j}]\tilde{\ell}}) \right] - \frac{n^4}{4} \\ &= n^3(n-1) \cdot \frac{1}{4} + n^3 \cdot \frac{1}{3} - \frac{n^4}{4} = \frac{n^3}{12}. \end{aligned}$$

- Analog findet man für  $i < j, \tilde{i} < j, i \neq \tilde{i}$

$$\text{Cov}_{1^*} \left[ \sum_{k=1}^n \sum_{\ell=1}^n \mathbf{I}(X_{[i]k} > X_{[j]\ell}), \sum_{\tilde{k}=1}^n \sum_{\tilde{\ell}=1}^n \mathbf{I}(X_{[\tilde{i}]\tilde{k}} > X_{[j]\tilde{\ell}}) \right] = \frac{n^3}{12}.$$

- Weiter gilt für  $i < j < \tilde{j}$

$$\begin{aligned} & \text{Cov}_{1^*} \left[ \sum_{k=1}^n \sum_{\ell=1}^n \mathbf{I}(X_{[i]k} > X_{[j]\ell}), \sum_{\tilde{k}=1}^n \sum_{\tilde{\ell}=1}^n \mathbf{I}(X_{[j]\tilde{k}} > X_{[\tilde{j}]\tilde{\ell}}) \right] \\ &= \mathbf{E}_{1^*} \left[ \sum_{k=1}^n \sum_{\ell=1}^n \sum_{\tilde{k}=1}^n \sum_{\tilde{\ell}=1}^n \mathbf{I}(X_{[i]k} > X_{[j]\ell}) \mathbf{I}(X_{[j]\tilde{k}} > X_{[\tilde{j}]\tilde{\ell}}) \right] - \frac{n^4}{4} \\ &= n^3(n-1) \cdot \frac{1}{4} + n^3 \cdot \frac{1}{6} - \frac{n^4}{4} = -\frac{n^3}{12}. \end{aligned}$$

- Wiederum analog erhält man für  $\tilde{i} < i < j$

$$\text{Cov}_{1^*} \left[ \sum_{k=1}^n \sum_{\ell=1}^n \mathbf{I}(X_{[i]k} > X_{[j]\ell}), \sum_{\tilde{k}=1}^n \sum_{\tilde{\ell}=1}^n \mathbf{I}(X_{[\tilde{i}]\tilde{k}} > X_{[i]\tilde{\ell}}) \right] = -\frac{n^3}{12}.$$

Indem wir (9.6), die Partition (9.7) der Indexmenge sowie die berechneten Kovarianzen in (9.5) einsetzen, erhalten wir

$$\begin{aligned} & \text{Var}_{1^*} [R_{m,n}^*(\mathbf{X})] \\ &= \frac{1}{\binom{m}{2}^2 n^4} \left( \sum_{i < j} \left[ \frac{n^3}{6} + \frac{n^2}{12} \right] \right. \\ & \quad \left. + \sum_{i < j, i < \tilde{j}, j \neq \tilde{j}} \frac{n^3}{12} + \sum_{i < j, \tilde{i} < j, i \neq \tilde{i}} \frac{n^3}{12} + \sum_{i < j < \tilde{j}} \left[ -\frac{n^3}{12} \right] + \sum_{\tilde{i} < i < j} \left[ -\frac{n^3}{12} \right] \right) \end{aligned}$$

$$\begin{aligned}
&= \frac{4}{m^2(m-1)^2n^4} \left( \frac{m(m-1)}{2} \cdot \frac{2n^3+n^2}{12} \right. \\
&\quad \left. + 2 \frac{m(m-1)(m-2)}{3} \cdot \frac{n^3}{12} - 2 \frac{m(m-1)(m-2)}{6} \cdot \frac{n^3}{12} \right) \\
&= \frac{6n^3 + 3n^2 + 2mn^3 - 4n^3}{18m(m-1)n^4} \\
&= \frac{2mn + 2n + 3}{18m(m-1)n^2},
\end{aligned}$$

was zu beweisen war.

Aufgrund dieses Resultats und des Satzes 8.2 ist leicht zu sehen, dass unter  $H_1^*$  die Varianz von  $R_{m,n}^*(\mathbf{X})$  für  $n \geq 2$  kleiner als diejenige von  $\tilde{R}_{m,n}(\mathbf{X})$  ist (für  $n = 1$  besteht trivialerweise Gleichheit).

Zur Bestimmung der asymptotischen Verteilung von  $R_{m,n}^*(\mathbf{X})$  lässt sich (wie schon unter  $H_0$ ) der zentrale Grenzwertsatz nicht anwenden. Wir sehen jedoch im folgenden Abschnitt, wie wir der Literatur eine Normalverteilungsapproximation entnehmen können.

### 9.3 Exkurs: Jonckheere-Test

Die Statistik  $R_{m,n}^*(\mathbf{X})$  entspricht weitgehend der Teststatistik, die Terpstra und (unabhängig von ihm) Jonckheere (1954) vorgeschlagen haben und die als „Jonckheere-Test“ für geordnete Alternativen bekannt geworden ist. Wir stellen diesen Test vor und zeigen den Zusammenhang mit unserer Statistik  $R_{m,n}^*(\mathbf{X})$ . Dabei passen wir die Notation von Jonckheere (1954) etwas an unsere bisherigen Konventionen an.

Gegeben seien Stichproben  $(X_{i1}, \dots, X_{in_i})$ ,  $i = 1, \dots, m$ , wobei  $X_{ij}$  der absolut stetigen Verteilung  $F_i$  entstamme. Wir untersuchen in diesem Abschnitt

$$H_0 : F_1(x) = \dots = F_m(x), \forall x$$

gegen

$$H_1 : F_1(x) > \dots > F_m(x), \forall x.$$

Zu bemerken ist erstens, dass Jonckheere die Alternative mit den umgekehrten Ungleichungen schreibt. Aus seinen Erörterungen wird aber klar, dass auch er Trends der Form  $X_{ik} < X_{j\ell}$  ( $i < j$ ) sucht, was den angegebenen Ungleichungen in den Verteilungsfunktionen entspricht. Zweitens ist festzuhalten, dass die Alternative wohl vorsichtiger geschrieben werden sollte, etwa als  $F_1(x) \geq \dots \geq F_m(x), \forall x; \exists (i, j, x_0) : i < j, F_i(x_0) > F_j(x_0)$ .

Jonckheere untersucht nun die Eigenschaften der Statistik

$$S = 2 \sum_{i=1}^{m-1} \sum_{j=i+1}^m \sum_{k=1}^{n_i} \sum_{\ell=1}^{n_j} \mathbf{I}(X_{ik} < X_{j\ell}) - \sum_{i=1}^{m-1} \sum_{j=i+1}^m n_i n_j$$

unter  $H_0$ . Insbesondere zeigt er, dass  $S$  für proportional wachsende Gruppen (d. h.  $\sum n_i \rightarrow \infty$  und  $\frac{n_i}{\sum n_i} \rightarrow a_i > 0$ ) asymptotisch normalverteilt ist.

Jonckheere betrachtet auch den Spezialfall  $n_1 = \dots = n_m = n$ . In diesem Fall ist  $S$  asymptotisch ( $n \rightarrow \infty$ ) normalverteilt mit Erwartungswert 0 und Varianz  $\frac{1}{18} [m(m-1)n^2(2n(m+1) + 3)]$ .  $S$  vereinfacht sich zu

$$S = 2 \sum_{i=1}^{m-1} \sum_{j=i+1}^m \sum_{k=1}^n \sum_{\ell=1}^n \mathbf{I}(X_{ik} < X_{j\ell}) - \binom{m}{2} n^2,$$

und wenn wir  $X_{ik} = X_{[i]k}$  einsetzen, erhalten wir

$$\begin{aligned} S &= 2 \sum_{i=1}^{m-1} \sum_{j=i+1}^m \sum_{k=1}^n \sum_{\ell=1}^n [1 - \mathbf{I}(X_{[i]k} > X_{[j]\ell})] - \binom{m}{2} n^2 \\ &= \binom{m}{2} n^2 [1 - 2R_{m,n}^*(\mathbf{X})] \end{aligned}$$

bzw.

$$R_{m,n}^*(\mathbf{X}) = \frac{1}{2} - \frac{S}{2 \binom{m}{2} n^2}. \quad (9.8)$$

Wir sehen leicht, dass unter Jonckheeres Nullhypothese auch  $R_{m,n}^*(\mathbf{X})$  asymptotisch normalverteilt ist, und zwar mit Erwartungswert  $\frac{1}{2}$  und Varianz  $\frac{2mn+2n+3}{18m(m-1)n^2}$ , was unseren Berechnungen im letzten Abschnitt entspricht.

Nun müssen wir noch die Hypothesen des Jonckheere-Tests mit denjenigen vergleichen, die wir im Rest dieses Kapitels betrachten: Unter  $H_0$  des Jonckheere-Tests sind die  $X_{[i]k}$  i.i.d., was wir ausserhalb dieses Abschnitts mit  $H_1^*$  bezeichnen. Die Alternative des Jonckheere-Tests beinhaltet unter anderem den Fall, dass  $X_{[i]k}$  gemäss  $F_{i:m}$  verteilt ist, also das, was wir sonst als Nullhypothese (Bedingungen (a) und (b)) bezeichnet haben.

Wir können also Jonckheeres Ergebnisse folgendermassen auf unser Problem übertragen (wir kehren nun endgültig zu den früheren Hypothesen zurück): Unter  $H_1^*$  gilt

$$\frac{R_{m,n}^*(\mathbf{X}) - \frac{1}{2}}{\sqrt{\frac{2mn+2n+3}{18m(m-1)n^2}}} \xrightarrow{d} Z \quad (n \rightarrow \infty), \quad (9.9)$$

wobei  $Z$  standardnormalverteilt ist.

## 9.4 Minimax-Test

Unter  $H_0$  steht uns keine Approximation der Verteilung von  $R_{m,n}^*(\mathbf{X})$  für grosse  $n$  zur Verfügung. Wir können demzufolge keine asymptotischen kritischen Werte berechnen und betrachten nur das konkrete Beispiel mit  $m = 3$  und  $n = 5$  (vgl. Bsp. 7.1, 8.1).

**Beispiel 9.1** *Da wir die exakte Verteilung von  $R_{m,n}^*(\mathbf{X})$  unter  $H_0$  ebenfalls nicht kennen, verwenden wir die empirische Verteilungsfunktion von  $R_{m,n}^*(\mathbf{X})$  aus 1 000 000 Simulationen. Unter  $H_1^*$  können wir die exakte Verteilung der Testgrösse von Jonckheere aus der Literatur (z. B. Jonckheere 1954, Table 3;  $k = 3, m = 5$ ) auf unser Problem übertragen, indem wir die Beziehung (9.8) verwenden. Die beiden so gewonnenen Verteilungen sind in Tabelle 9.2 auszugsweise angegeben.*

Tabelle 9.2: Verteilungen von  $R_{3,5}^*(\mathbf{X})$  unter  $H_0$  (geschätzt) und  $H_1^*$ .

$r$	$\widehat{P}_0(R_{3,5}^*(\mathbf{X}) > r)$	$P_{1^*}(R_{3,5}^*(\mathbf{X}) \leq r)$
0.2	0.2123	0.0087
0.2133	0.1710	0.0120
0.2267	0.1362	0.0161
0.24	0.1073	0.0214
0.2533	0.0835	0.0279
0.2667	0.0645	0.0359
0.28	0.0491	0.0456
0.2933	0.0370	0.0572
0.3067	0.0278	0.0708
0.32	0.0207	0.0867
0.3333	0.0154	0.1049
0.3467	0.0113	0.1256
0.36	0.0082	0.1489

Daraus können wir entnehmen, dass  $H_0$  nach dem Minimax-Test mit  $R_{3,5}^*(\mathbf{X})$  zu verwerfen ist, falls  $R_{3,5}^*(\mathbf{X})$  grösser als  $r_{3,5}^{*\text{Minimax}} = 0.28$  ist. Die Fehlerwahrscheinlichkeiten betragen nun 4.9% bzw. 4.6%, sind also nochmals deutlich tiefer als im Beispiel 8.1.

Schliesslich vergleichen wir  $P_{1^*}(R_{3,5}^*(\mathbf{X}) \leq r_{3,5}^{*\text{Minimax}})$  noch mit dem Wert, den wir unter Verwendung der Normalverteilungsapproximation (9.9) erhalten (wir verwenden wieder eine Stetigkeitskorrektur):

$$P_{1^*}(R_{3,5}^*(\mathbf{X}) \leq 0.28) \approx \Phi\left(\frac{0.28667 - 0.5}{\sqrt{\frac{2 \cdot 3 \cdot 5 + 2 \cdot 5 + 3}{18 \cdot 3 \cdot 2 \cdot 5^2}}}\right)$$



$$\begin{aligned} &\approx \Phi(-1.6905) \\ &\approx 0.0455 \end{aligned}$$

*Die Approximation ist in diesem Fall also sehr gut.*

Die deutlich kleineren Fehlerwahrscheinlichkeiten im Beispiel erstaunen nicht – sowohl unter  $H_0$  als auch unter  $H_1^*$  hat ja  $R_{m,n}^*(\mathbf{X})$  den gleichen Erwartungswert wie  $\tilde{R}_{m,n}(\mathbf{X})$ , aber eine kleinere Varianz. Demzufolge dürfte sich die Verwendung von  $R_{m,n}^*(\mathbf{X})$  lohnen, auch wenn zum Teil die theoretischen Grundlagen fehlen.

# Kapitel 10

## Ausblick

In den letzten drei Kapiteln wurden einfache Minimax-Tests empfohlen, um zu überprüfen, wie schwerwiegend die praktischen Probleme in einem konkreten Fall sind. Solange man annimmt, dass die zugrundeliegende Verteilung der Stichprobenelemente stetig ist, beeinflusst diese die Verteilung der Teststatistiken unter den betrachteten Hypothesen nicht, da nur Bernoulli-Variablen der Form  $I(X_{[i]k} > X_{[j]\ell})$  betrachtet werden. Die Minimax-Tests sind also in diesem Sinne nichtparametrisch.

Die Tests sind aber insofern unbefriedigend, dass sie die Kostenüberlegungen aus Kapitel 5 völlig ausser Acht lassen – schliesslich möchte man wissen, ob *bei gleichen Kosten* eher ein RSS- oder ein SRS-Verfahren besser ist. Als Fortsetzung dieser Arbeit wäre es deshalb interessant, Tests zu entwickeln, die diese Aspekte mitberücksichtigen.

Die hier vorgestellten Statistiken dürften grundsätzlich auch für solche Tests geeignet sein. Allerdings müssten Verwerfungsgrenzen bestimmt werden, die ausser von den Parametern  $m$  und  $n$  auch vom anzuwendenden Verfahren (z. B. Vorzeichentest), von den Mess- und Rangierungskosten sowie (je nach Verfahren) von der zugrundeliegenden Verteilung abhängen. Auch müssten wohl recht weitreichende Modellannahmen bezüglich der Rangierungsfehler und der Struktur der Sets getroffen werden.

Angesichts dieser Schwierigkeiten wird man wohl in den meisten Fällen auf einen solchen Test verzichten müssen. Statt sich der Illusion eines „perfekten“ Tests hinzugeben, wird man realistischerweise eine der vorgestellten Statistiken berechnen, diese mit den Erwartungswerten unter  $H_0$  und  $H_1^*$  sowie mit dem kritischen Wert des Minimax-Tests vergleichen und schliesslich unter Einbezug der Kostenaspekte abwägen, ob man in einer weiteren Studie nochmals ein RSS-Verfahren anwenden soll.

## Zusammenfassung

In der vorliegenden Arbeit wurden zunächst zwei Anwendungsbeispiele von Ranked Set Sampling aus der Literatur vorgestellt und wichtige Resultate dazu nachvollzogen. Ebenso wurden die Auswirkungen der bekannten Probleme in der praktischen Anwendung der Methode aufgezeigt, wobei auch die Kosten für die Durchführung des Verfahrens mit denjenigen bei Verwendung gewöhnlicher Stichproben verglichen wurden.

Hauptziel der Arbeit war aber die Bereitstellung von Methoden, mit denen die praktischen Probleme erkannt werden können, und zwar nur aufgrund der zur genauen Erhebung ausgewählten Stichprobenelemente. Neben einer einfachen grafischen Methode wurden zu diesem Zweck drei nichtparametrische Statistiken untersucht. Diese erlauben zwar mit zunehmender Verfeinerung sicherere Aussagen, bieten dafür aber auch mehr und mehr Probleme bei der theoretischen Behandlung. So konnten für die letzte dieser Statistiken, die in enger Beziehung zur bekannten Testgrösse von Jonckheere steht, teilweise nur noch empirische Untersuchungen angestellt werden. Dennoch scheint gerade diese letzte Statistik zumindest eine recht sichere Erkennung extremer Abweichungen von den Voraussetzungen zu ermöglichen.

Für die Anwendung der Erkennungsmethoden mussten gewisse Vorbehalte angebracht werden, da die Frage der Kosten nicht berücksichtigt werden konnte.

## Literaturverzeichnis

- [1] Barabesi, L. (2001). The unbalanced ranked-set sample sign test. *Journal of Nonparametric Statistics* 13, 279-289.
- [2] Dell, T. R. und Clutter, J. L. (1972). Ranked set sampling theory with order statistics background. *Biometrics* 28, 545-555.
- [3] Hettmansperger, T. P. (1984). *Statistical inference based on ranks*, Wiley.
- [4] Hettmansperger, T. P. (1995). The ranked-set sample sign test. *Journal of Nonparametric Statistics* 4, 263-270.
- [5] Hettmansperger, T. P. und McKean, J. W. (1998). *Robust nonparametric statistical methods* (Kendall's library of statistics, 5), Arnold.
- [6] Jonckheere, A. R. (1954). A distribution-free  $k$ -sample test against ordered alternatives. *Biometrika* 41, 133-145.
- [7] Lehmann, E. L. (1975). *Nonparametrics: Statistical methods based on ranks*, Holden-Day/McGraw-Hill.
- [8] McIntyre, G. A. (1952). A method for unbiased selective sampling, using ranked sets. *Australian Journal of Agricultural Research* 3, 385-390.
- [9] Ridout, M. S. und Cobby, J. M. (1987). Ranked set sampling with non-random selection of sets and errors in ranking. *Journal of the Royal Statistical Society – Series C, Applied Statistics* 36, 145-152.
- [10] Takahasi, K. und Wakimoto, K. (1968). On unbiased estimates of the population mean based on the sample stratified by means of ordering. *Annals of the Institute of Statistical Mathematics* 20, 1-31.