Discussion Papers in Statistics and Quantitative Economics Diskussionsbeiträge zur Statistik und Quantitativen Ökonomik Editors / Herausgeber: Harry Hauptmann and Wolf Krumbholz

> "Ein Algorithmus zur Bestimmung der exakten Verteilung der Kruskal-Wallis Teststatistik"

> > von

Jutta Arrenberg

Nr. 33

Mai 1988

Editorial Board / Mitherausgeber:

Siegfried Heiler (Konstanz)
Karl Clemens Mosler (Hamburg)
Martin Schader (Hamburg)
Friedrich Schmid (Hamburg)
Wolfgang Sendler (Trier)

Fachbereich Wirtschafts- und Organisationswissenschaften Universität der Bundeswehr Hamburg Postfach 70 08 22, D-2000 Hamburg 70

O. Zusammenfassung

Die exakte Verteilung der Kruskal-Wallis Teststatistik kann mit Hilfe eines Algorithmus berechnet werden. Gleiche Stichprobenwerte werden dabei berücksichtigt und brauchen nicht aus der Stichprobe entfernt zu werden. Der Algorithmus wurde in APL programmiert.

1. Einleitung

Streitberg und Röhmel haben 1983 [4] einen "Stift"-Algorithmus vorgestellt, mit dessen Hilfe die exakte Verteilung für eine große Klasse nichtparametrischer Tests berechnet werden kann. Es stellt sich die Frage, inwieweit
der obige Algorithmus für die Berechnung der exakten Verteilung der Kruskal-Wallis Teststatistik benutzt werden
kann.

Beim Kruskal-Wallis Test basiert die Teststatistik auf der gewichteten Summe aller Stichprobenwerte. Für Stichproben ohne Bindungen ist die exakte Verteilung für Stichproben-umfänge kleiner gleich 4 und höchstens 5 Stichproben zwar tabelliert [2]. Treten jedoch Bindungen auf, so ist die Verteilung der Kruskal-Wallis Teststatistik eine bedingte Verteilung, die abhängig ist von der Anzahl und der Rang-ordnung der Bindungen. Für diesen Fall wären Tabellen für die exakte Verteilung zu umfangreich, da die Anzahl des Auftretens und die Größe von Bindungen zu viele verschiedene Bedingungen stellen. Dieses Problem löst ein Algorithmus, der in der vorliegenden Arbeit vorgestellt wird.

In Abschnitt 2 wird der bekannte Kruskal-Wallis Test kurz dargestellt, und zwar ohne die Voraussetzung der Stetigkeit der Verteilungen. In Abschnitt 3 wird dann der Algorithmus zur Berechnung der exakten Verteilung der Kruskal-Wallis Teststatistik entwickelt. Als Anwendung des Algorithmus aus Abschnitt 3 werden in den Abschnitten 4 und 5 zwei numerische Beispiele mit und ohne Bindungen gerechnet.

2. Der Kruskal-Wallis Test

Der Kruskal-Wallis Test ist ein nichtparametrischer Test auf Lage. Seien $X_1,\ X_2,\ldots,X_k,\ (k\ge 2)$ unabhängige Zufalls-variablen mit den Verteilungsfunktionen $F_1,\ F_2,\ldots,F_k$ resp. Aufgrund von k unabhängigen Stichproben $(X_{i\,1},\ldots,X_{i\,n\,i})$ vom Umfang n_i , $i=1,\ldots,k$ resp. soll auf Gleichheit der Lage der Verteilungen getestet werden:

$$H_0: F_1 = F_2 = \dots = F_k$$

gegen

 $H_1: F_i(x) = F_j(x-\Theta)$ für alle x und für mindestens ein Paar (i,j), i‡j und $\Theta \neq 0$.

Werden die Verteilungsfunktionen F_1 , F_2 ,..., F_k als stetig vorausgesetzt, so hat das Ereignis, gleiche Beobachtungs-werte in der Stichprobe (Bindungen) zu haben, die Wahrscheinlichkeit Null. In Anwendungen des Tests tauchen i. d. R. jedoch Bindungen auf. Der vorgestellte Algorithmus berücksichtigt diese Situation.

Kruskal und Wallis [3] schlugen folgendes Testverfahren vor, das auf Ranginformationen beruht.

Zunächst werden die N := $\sum_{i=1}^{K} n_i$ Stichprobenwerte der Größe i=1

nach geordnet; sodann werden ihnen Ränge zugewiesen:

Rang
$$(X_{ij}) = Rang \ von \ X_{ij} \ in \ Ordnungs-Statistik \ (X_{(1)}, \dots, X_{(N)})$$

$$von \ (X_{11}, \dots, X_{kn_k}).$$

Treten Bindungen in der Stichprobe auf, so werden diese Stichprobenwerte mit mittleren Rängen versehen. Wir schreiben kurz: $a_j := Rang \ X_{(j)}$.

Bezeichnen

 $N(j) := Anzahl der Beobachtungswerte, die kleiner als <math>X_{(j)}$ sind

und

L(j) := Anzahl des Auftretens des Beobachtungswertes $X_{(j)} in (X_{(1)}, ..., X_{(N)}),$

so ist

$$a_{j} = N(j) + \frac{\frac{1}{2} L(j)(L(j)+1)}{L(j)}$$

$$= N(j) + \frac{1}{2} (L(j)+1) ; j=1,...,N$$

Insbesondere ist $a_j = j$, falls L(j) = 1 ist, d. h., falls keine Bindungen auftreten.

Für jede der k Stichproben lassen sich die Rangsummen R_1, R_2, \ldots, R_k berechnen:

$$R_i := \sum_{j=1}^{n_i} Rang(X_{ij}); i=1,...,k$$

Unter der Nullhypothese H_0 ist jede Permutation des

N-Tupels
$$(a_1, \ldots, a_N)$$

gleichwahrscheinlich für $(X_{11}, \dots, X_{kn_k})$.

Der Erwartungswert einer Rangsumme R_i unter H_0 ist:

$$E(R_i) = n_i \frac{N+1}{2}.$$

Die Teststatistik basiert auf den quadrierten Abweichungen der beobachteten Rangsummen R $_i$ von den erwarteten Rangsummen n $_i$ $\frac{N+1}{2}$. Statt der Summe

$$S := \sum_{i=1}^{K} (R_i - \frac{n_i(N+1)}{2})^2$$

schlugen Kruskal und Wallis jedoch als Teststatistik T die mit den jeweiligen Stichprobenumfängen gewichteten Summanden aus S vor, um unter H_0 mit der zusätzlichen Voraussetzung der Stetigkeit von F_1,\ldots,F_k für hinreichend große n_i eine Approximation von T durch die χ^2 -Verteilung zu gewährleisten:

$$T := \frac{12}{N(N+1)} \sum_{i=1}^{K} \frac{1}{n_i} (R_i - \frac{n_i(N+1)}{2})^2.$$

Große Werte von T führen zur Ablehnung der Nullhypothese H₀:

Annahme von H_0 , falls $T_{empirisch} < k_{\alpha}$, Ablehnung von H_0 , falls $T_{empirisch} \geq k_{\alpha}$, mit P_{H_0} (T < k_{α}) = 1- α , α = Testniveau.

3. Ein Algorithmus für die exakte Verteilung der Kruskal-Wallis Teststatistik \overline{T}

Um den Wertebereich der Statistik T zu bestimmen, reicht es, den Wertebereich der Statistik

$$R := \sum_{i=1}^{k} \frac{R_i^2}{n_i}$$

zu kennen, da gilt:

$$T = \left(\frac{12}{N(N+1)} \frac{k}{i=1} \frac{R_{i}^{2}}{n_{i}}\right) - 3(N+1) .$$

Mit Hilfe des Streitberg/Röhmel Algorithmus [4] lassen sich die einzelnen Verteilungen der Statistiken R_1,\ldots,R_k berechnen. Quadriert man die Werte des Wertebereichs von

 R_1, \dots, R_k respektive und gewichtet sie jeweils mit $\frac{1}{n_1}, \dots, \frac{1}{n_k}$ respektive, so erhält man auch die einzelnen Verteilungen der Statistiken $\frac{R_1^2}{n_1}, \dots, \frac{R_k^2}{n_k}$.

Wir zeigen nun, wie sich durch eine Abwandlung des Streitberg/ Röhmel Algorithmus auch die Verteilung von R berechnen läßt. Ausgangspunkt ist das Produkt:

Betrachtet man bei diesem Produkt nur die Terme mit genaun, $1 \le n \le N$, δ_i -Produkten:

und transponiert man diese Terme derart, daß die Potenzen von x bei diesen Termen quadriert und gewichtet werden:

so erhält man eine "erzeugende" Funktion H_N :

$$H_{N}(x,n,\delta_{1},\ldots,\delta_{N}) := \sum_{\substack{1\\1,\ldots,1\\n} \in I < n > } x^{\frac{1}{n}(\sum_{j=1}^{n} a_{1j})^{2}} n_{\substack{1\\j=1}} \delta_{1j};$$

dabei bezeichnet I<n> die Menge aller geordneten n-Tupel
aus {1,2,...,N}, d. h.:

$$I < n > := \{(l_1, ..., l_n); l_1 < ... < l_n, \{l_1, ..., l_n\} \underline{\mathbf{c}} \{1, 2, ..., N\}\},$$

 $f \text{ ur } 1 \le n \le N.$

Für die Verteilungen der Statistiken $\frac{R_1^2}{n_1}, \dots, \frac{R_k^2}{n_k}$ gilt der folgende Satz:

Satz 1

Sei i ϵ {1,2,...,k} fest. Dann ist

$$\binom{N}{n_{i}} \cdot P_{H_{0}} \left(\frac{R_{i}^{2}}{n_{i}} = r \right) = \# \left\{ \binom{1}{1}, \dots, \binom{1}{n_{i}} \right\} \in I < n_{i} > :$$

$$\frac{1}{n_{i}} \left(\sum_{j=1}^{n} a_{1j} \right)^{2} = r \right\}.$$

Beweis:

Der Beweis ergibt sich aus der Definition der erzeugenden Funktion H_N wie folgt:

Seien i ϵ {1,2,...,k} und n_i fest, und sei

Nullhypothese H_0 :

$$\binom{N}{n_{i}}$$
 $\binom{N}{H_{0}}$ $\binom{N}{H_{0}}$

Da $\frac{R_i^2}{n_i}$ eine streng monotone Funktion von R_i ist, folgt die Behauptung.

Folgerung aus Satz 1

Für i ε {1,2,...,k} ist

$$P_{H_0} \left(\frac{R_i^2}{n_i} = r \right) = \frac{1}{\binom{N}{n_i}} \cdot \left\{ \begin{array}{l} \text{Häufigkeit von r als Exponent} \\ \text{von x in } H_N \left(x, n_i, \delta_1, \dots, \delta_N \right) \end{array} \right\}$$

Um unter der Nullhypothese H_0 aus den Verteilungen der Statistiken $\frac{R_i^2}{n_1}, \dots, \frac{R_k^2}{n_k}$ die Verteilung von $R = \sum\limits_{i=1}^k \frac{R_i^2}{n_i}$ bestimmen zu können, müssen jeweils genau k Werte der k Werteberreiche von $\frac{R_1^2}{n_1}, \dots, \frac{R_k^2}{n_k}$ respektive addiert werden.

Multipliziert man alle k erzeugenden Funktionen

 H_N $(x,n_1,\delta_1,\ldots,\delta_N)$,..., H_N $(x,n_k,\delta_1,\ldots,\delta_N)$ miteinander, so erhält man:

$$H_{N}$$
 $(x; n_{1}, ..., n_{k}; \delta_{1}, ..., \delta_{N})$

$$\vdots = \prod_{i=1}^{k} H_{N} (x, n_{i}, \delta_{1}, ..., \delta_{N})$$

$$= \sum_{\substack{(1_{11}, \dots, 1_{1n_1}) \in I < n_1 > \\ }} \sum_{i=1}^{n_1} (\sum_{j=1}^{n_1} a_{1_{1j}})^2 \cdots \delta_{i=1}^{n_1} \cdots \delta$$

$$\sum_{\substack{X \\ (l_{k1}, \dots, l_{kn_k}) \in I < n_k > \\ i = 1}} \sum_{\substack{X \\ (l_{k1}, \dots, l_{kn_k}) \in I < n_k > \\ k}} \sum_{\substack{X \\ (l_{i1}, \dots, l_{kn_k}) \\ i = 1}} \sum_{\substack{X \\ (l_{i1}, \dots, l_{kn_k}) \\ (l_{i1}, \dots, l_{in_i}) \in I < n_i > \text{ für alle } i = 1, \dots, k}} \sum_{\substack{X \\ (l_{i1}, \dots, l_{kn_k}) \\ (l_{i1}, \dots, l_{in_i}) \in I < n_i > \text{ für alle } i = 1, \dots, k}}$$

Diese Summe wird zerlegt in zwei Summanden

also - in kurzer Notation -:

$$H_{N}(x;n_{1},...,n_{k};\delta_{1},...,\delta_{N}) =$$

$$\sum_{(l_{11},...,l_{kn_{k}})}^{\Sigma} \min_{\{l_{11},...,l_{kn_{k}}\}} \neq \{1,...,N\}$$

$$+ H_{red}(x,n_{1},...,n_{k},\delta_{1},...,\delta_{N}),$$

wobei die "reduzierte" Funktion H_{red} wie folgt lautet:

$$H_{\text{red.}}(x,n_{1},\ldots,n_{k},\delta_{1},\ldots,\delta_{N}) := \sum_{x_{i=1}^{k}} x_{i}^{n_{i}} (\sum_{j=1}^{n_{i}} a_{1}_{i,j})^{2} \delta_{1}^{n_{i}} \cdots \delta_{N}^{n_{k}} (1_{11},\ldots,1_{kn_{k}})^{n_{i}} \delta_{1}^{n_{i}} \cdots \delta_{N}^{n_{i}} \delta_{N}^{n_{i}} \delta_{1}^{n_{i}} \cdots \delta_{N}^{n_{i}} \delta_{N}^$$

mit $h(r) := H \ddot{a} u f i g k e i t von r als Exponent von x in H_{red.}(x,n_1,...,n_k, 1,...,1), h(r) <math>\geq 1$.

Wir haben folgendes Ergebnis:

Satz 2

- (1) r aus dem Wertebereich von R \leftarrow r Exponent von x in \leftarrow \leftarrow (x, \leftarrow 1,..., \leftarrow n_k, 1,...,1)
- $(2) \quad \frac{N!}{n_1! \cdot \cdot \cdot \cdot n_k!} P_{H_0} (R=r) = h(r) .$

4. Ein numerisches Beispiel (ohne Bindungen)

$$n_1 = 3$$
, $n_2 = 2$, $n_3 = 1$, $N = 6$

Die Stichprobe $(x_{11}, x_{12}, x_{13}, x_{21}, x_{22}, x_{31})$ sei ohne Bindungen. Der Rang der Ordnungsstatistik ist:

$$(a_1, \ldots, a_6) = (1, 2, 3, 4, 5, 6)$$

und das Produkt

$$\frac{1}{1} (1 + x^{a_{i}} \delta_{i}) = \frac{6}{1} (1 + x^{i} \delta_{i})$$

$$= 1 + x \delta_{1} + x^{2} \delta_{2} + \dots + x^{6} \delta_{6}$$

$$+ x^{3} \delta_{1} \delta_{2} + \dots + x^{11} \delta_{5} \delta_{6}$$

$$\vdots$$

$$+ x^{15} \delta_{1} \delta_{2} \delta_{3} \delta_{4} \delta_{5} + \dots + x^{20} \delta_{2} \delta_{3} \delta_{4} \delta_{5} \delta_{6}$$

$$+ x^{21} \delta_{1} \delta_{2} \delta_{3} \delta_{4} \delta_{5} \delta_{6}$$

Daraus ergeben sich für $i=1,\ldots,k$ die erzeugenden Funktionen H_N $(x,n_i,\delta_1,\ldots,\delta_N)$ wie folgt (nach Transformation der Potenzwerte von x):

$$H_{6} (x,n_{1},\delta_{1},...,\delta_{6}) = x^{12}\delta_{1}\delta_{2}\delta_{3} + x^{16\frac{1}{3}}\delta_{1}\delta_{2}\delta_{4} + x^{21\frac{1}{3}}\delta_{1}\delta_{3}\delta_{4}$$

$$+ x^{27}\delta_{2}\delta_{3}\delta_{4} + x^{21\frac{1}{3}}\delta_{1}\delta_{2}\delta_{5} + x^{27}\delta_{1}\delta_{3}\delta_{5}$$

$$+ x^{3\frac{1}{3}}\delta_{1}\delta_{4}\delta_{5} + x^{40\frac{1}{3}}\delta_{2}\delta_{4}\delta_{5} + x^{48}\delta_{3}\delta_{4}\delta_{5}$$

$$+ x^{27}\delta_{1}\delta_{2}\delta_{6} + x^{3\frac{1}{3}}\delta_{1}\delta_{3}\delta_{6} + x^{40\frac{1}{3}}\delta_{2}\delta_{3}\delta_{6}$$

$$+ x^{40\frac{1}{3}}\delta_{1}\delta_{4}\delta_{6} + x^{48}\delta_{2}\delta_{4}\delta_{6} + x^{56\frac{1}{3}}\delta_{3}\delta_{4}\delta_{6}$$

$$+ x^{48}\delta_{1}\delta_{5}\delta_{6} + x^{56\frac{1}{3}}\delta_{2}\delta_{5}\delta_{6} + x^{56\frac{1}{3}}\delta_{3}\delta_{5}\delta_{6}$$

$$+ x^{75}\delta_{4}\delta_{5}\delta_{6} + x^{3\frac{1}{3}}\delta_{2}\delta_{3}\delta_{5}$$

$$+ x^{75}\delta_{4}\delta_{5}\delta_{6} + x^{18}\delta_{2}\delta_{4} + x^{12\frac{1}{2}}\delta_{2}\delta_{3}$$

$$+ x^{12\frac{1}{2}}\delta_{1}\delta_{4} + x^{18}\delta_{2}\delta_{4} + x^{24\frac{1}{2}}\delta_{3}\delta_{4}$$

$$+ x^{18}\delta_{1}\delta_{5} + x^{24\frac{1}{2}}\delta_{2}\delta_{5} + x^{32}\delta_{3}\delta_{5}$$

$$+ x^{40\frac{1}{2}}\delta_{4}\delta_{5} + x^{24\frac{1}{2}}\delta_{1}\delta_{6} + x^{32}\delta_{2}\delta_{6}$$

$$+ x^{40\frac{1}{2}}\delta_{4}\delta_{5} + x^{24\frac{1}{2}}\delta_{1}\delta_{6} + x^{32}\delta_{2}\delta_{6}$$

$$+ x^{40\frac{1}{2}}\delta_{3}\delta_{6} + x^{50}\delta_{4}\delta_{6} + x^{60\frac{1}{2}}\delta_{5}\delta_{6}$$

$$H_6(x,n_3,\delta_1,...,\delta_6) = x \delta_1 + x^4 \delta_2 + x^9 \delta_3 + x^{16} \delta_4 + x^{25} \delta_5 + x^{36} \delta_6$$

Für die 1. Summe des Produkts $\frac{k}{11}H_N$ ergibt sich für $(\delta_1, \dots, \delta_6) = (1, \dots, 1)$: $H_{\text{red.}}(x, n_1, n_2, n_3, 1, \dots, 1) = 4x + 4x^{75} + 2x^{76} + 2x^{11} + 4x^{11} +$

Daraus ergibt sich die Verteilung der Statistik R unter H_0 wie folgt:

r	73 <u>5</u>	74 1	75	76 1	76 <u>1</u>	77 <u>5</u>	78 1
P _{H0} (R=r)	4/60	2/60	4/60	2/60	6/60	6/60	2/60
0						87	
	2/60	6/60	10/60	4/60	4/60	2/60	6/60

Aus der Äquivalenz der Statistik R und der Teststatistik T, T = $\frac{12}{N(N+1)}$ R - 3(N+1) folgt die Verteilung von T unter H_0 :

t	$P_{H_0}(T \geq t)$
4.285714286	0.1
3.257142857	0.1333333333
3,523809524	0.2
3.095238095	0.2666666667
2.380952381	0.4333333333
2.142857143	0.5333333333
1.952380952	0.5666666667
1.330952381	0.6
1.238095238	0.7
0.8571428571	0.8
0.8095238095	0.8333333333
0.4285714286	0.9
0.2380952381	0.9333333333
0.09523809524	1

5. Ein numerisches Beispiel (mit Bindungen)

 $n_1 = 3$, $n_2 = 2$, $n_3 = 1$, N = 6

Die Stichprobe $(x_{11}, x_{12}, x_{13}, x_{21}, x_{22}, x_{31}) = (1, 2, 2, 3, 4, 9).$

Dann beträgt der Rang der Ordnungsstatistik:

$$(a_1, \dots, a_6) = (1, 2.5, 2.5, 4, 5, 6)$$

und das Produkt

Daraus ergeben sich für $i=1,\ldots,k$ die erzeugenden Funktionen H_N $(x,n_i,\delta_1,\ldots,\delta_N)$ wie folgt (nach Transformation der Potenzwerte von x):

H₆
$$(x, n_1, \delta_1, \dots, \delta_6) = x^{12} \delta_1 \delta_2 \delta_3 + x^{18\frac{3}{4}} \delta_1 \delta_2 \delta_4 + x^{18\frac{3}{4}} \delta_1 \delta_3 \delta_4 + x^{27} \delta_2 \delta_3 \delta_4 + x^{24\frac{1}{12}} \delta_1 \delta_2 \delta_5 + x^{24\frac{1}{12}} \delta_1 \delta_3 \delta_5 + x^{33\frac{1}{3}} \delta_1 \delta_4 \delta_5 + x^{44\frac{1}{12}} \delta_2 \delta_4 \delta_5 + x^{44\frac{1}{12}} \delta_3 \delta_4 \delta_5 + x^{30\frac{1}{12}} \delta_1 \delta_3 \delta_6 + x^{30\frac{1}{12}} \delta_1 \delta_3 \delta_6 + x^{40\frac{1}{3}} \delta_2 \delta_3 \delta_6$$

$$\begin{array}{c} + x^{40\frac{1}{3}} \delta_{1} \delta_{4} \delta_{6} + x^{52\frac{1}{12}} \delta_{2} \delta_{4} \delta_{6} + x^{52\frac{1}{12}} \delta_{3} \delta_{4} \delta_{6} \\ + x^{48} \delta_{1} \delta_{5} \delta_{6} + x^{60\frac{3}{4}} \delta_{2} \delta_{5} \delta_{6} + x^{60\frac{3}{4}} \delta_{3} \delta_{5} \delta_{6} \\ + x^{75} \delta_{4} \delta_{5} \delta_{6} + x^{61\frac{3}{4}} \delta_{2} \delta_{3} \delta_{5} \\ + x^{75} \delta_{4} \delta_{5} \delta_{6} + x^{61\frac{3}{4}} \delta_{2} \delta_{3} \delta_{5} \\ + x^{75} \delta_{4} \delta_{5} \delta_{6} + x^{61\frac{3}{4}} \delta_{2} \delta_{3} \delta_{5} \\ + x^{12\frac{1}{2}} \delta_{1} \delta_{4} + x^{12\frac{1}{4}} \delta_{2} \delta_{4} + x^{12\frac{1}{4}} \delta_{3} \delta_{4} \\ + x^{18} \delta_{1} \delta_{5} + x^{21\frac{1}{8}} \delta_{2} \delta_{4} + x^{18\frac{1}{8}} \delta_{3} \delta_{4} \\ + x^{18} \delta_{1} \delta_{5} + x^{21\frac{1}{8}} \delta_{2} \delta_{5} + x^{28\frac{1}{8}} \delta_{3} \delta_{5} \\ + x^{18} \delta_{1} \delta_{5} + x^{21\frac{1}{2}} \delta_{1} \delta_{6} + x^{21\frac{1}{8}} \delta_{3} \delta_{5} \\ + x^{18} \delta_{3} \delta_{6} + x^{50} \delta_{4} \delta_{6} + x^{16\frac{1}{4}} \delta_{2} \delta_{5} \delta_{6} \\ + x^{61\frac{1}{4}} \delta_{3} \delta_{6} + x^{50} \delta_{4} \delta_{6} + x^{16\frac{1}{4}} \delta_{3} + x^{16} \delta_{4} \\ + x^{25} \delta_{5} + x^{36} \delta_{6} \\ \\ \text{Für die 1. Summe des Produkts} \quad \begin{array}{c} k \\ 1 \\ 1 \\ 1 \end{array} \quad \text{H}_{N} \text{ ergibt sich für} \\ (\delta_{1}, \ldots, \delta_{6}) = (1, \ldots, 1) : \\ \\ \text{Hred.} \quad (x, n_{1}, n_{2}, n_{3}, 1, \ldots, 1) = \\ x^{73.83} + 2x^{74.2083} + x^{74.3} + 2x^{74.7083} + 2x^{74.83} \\ + 2x^{75.375} + 2x^{75.7083} + 4x^{76.2083} + 2x^{76.3} + 2x^{80.3} + x^{81} \\ + 6x^{81.2083} + 3x^{81.83} + 6x^{82.875} \\ + 2x^{83.2083} + x^{84.3} + 2x^{85.5} + 2x^{86.2083} + x^{87} \\ + 2x^{87.375} + 4x^{88.5} \end{array}$$

Daraus ergibt sich die Verteilung der Statistik R unter H_{Ω} wie folgt:

r	73.83	74.2083	74.3	74.7083	74.83
P _{H0} (R=r)	1/60	2/60	1/60	2/60	2/60
	75.375	75.708	3 76.2	083 76.3	76.5
	2/60	2/60	4/60	2/60	2/60
	76.83	77.83	79.5 7	9.875 80	.3
	2/60	2/60	2/60 2	/60 2/	60
	81 8	1.2083	81.83	82.875 8	3.2083
	1/60 6	/60	3/60	6/60 2	/60
	84.3	85.5 86	.2083	87 87.3	75 88.5
	1/60	2/60 2/	60	1/60 2/60	4/60

Aus der Äquivalenz der Statistik R und der Teststati- stik T, T = $\frac{12}{N(N+1)}$ R - 3(N+1) folgt die Verteilung von T unter H₀. Aufgeführt sind nur die oberen kritischen Werte für α = 0.1:

t	$P_{H_0}(T \ge t)$
4.285714286	0.0666666667
3.964285714	0.1
3.857142857	0.1166666667

Ich danke Herrn Professor Dr. Fritz Pokropp für wertvolle Anregungen zu diesem Aufsatz.

Literaturverzeichnis:

- [1] Gibbons, J.D. (1985): Nonparametric Statistical Inference, Marcel Dekker, Inc.
- [2] Iman, R.L., Quade, D. and Alexander, D.A. (1975):
 Exact Probability Levels for the Kruskal-Wallis Test.
 In Institute of Mathematical Statistics, ed.,
 Selected Tabels in Mathematical Statistics, Vol. III,
 American Mathematical Society, Providence, RI, 329-384.
- [3] Kruskal, W.H. and Wallis, W.A. (1952): Use of ranks in one criterion variance analysis. J. Amer. Statis. Assoc. 47, 583-621; errata (1953) 48, 907-911.
- [4] Streitberg, B., Röhmel, J. (1984): Exact Nonparametrics in APL. Proceedings of the APL-Conference 84, 313-325, ACM, 0-89791-137-7/84/006/0313.
- [5] Streitberg, B., Röhmel, J. (1987): Exakte Verteilungen für Rang- und Randomisierungstests im allgemeinen c-Stichprobenproblem. EDV in Medizin und Biologie 18 (1), 12-19.