Symbolic computation and exact distributions of nonparametric test statistics

Eindhoven University
of Technology

Department of Mathematics and Computing Sciences

Memorandum COSOR 97-17

Symbolic computation and exact distributions of nonparametric test statistics

M.A. van de Wiel
A. Di Bucchianico
P. van der Laan

Eindhoven, October 1997
The Netherlands

Symbolic computation and exact distributions of nonparametric test statistics

M.A. van de Wiel, A. Di Bucchianico and P. van der Laan

Abstract

We show how to use computer algebra for computing exact distributions on nonparametric statistics. We give several examples of nonparametric statistics with explicit probability generating functions that can be handled this way. In particular, we give a new table of critical values of the Jonckheere-Terpstra test that extends tables known in the literature.

Keywords: Computer algebra; generating function; Jonckheere-Terpstra test.

1 Introduction

Nonparametric statistics is a valuable tool of applied statistics. Thus it is important to have correct and extensive tables (on paper or in a digital form) of critical values of nonparametric tests. Many nonparametric tables were computed in the fifties and sixties using recurrences. However, computations with recursions tend to be very time-consuming. Therefore, other ways of computing were developed. The most important contributions in this respect (often for the broader class of permutation tests) are the fast Fourier methods of Pagano and Tritchler (1983), various shift-algorithms (see e.g. Streitberg and Röhmel (1986) and Edgington (1995; pp. 393-398)), and the network algorithms developed by Mehta and co-workers (see Good (1994, chap. 13) for an overview). Baglivo, Pagano and Spino (1993) remark that all these methods can be described as efficient methods to calculate generating functions. It is thus not surprising that the recent availability of computer algebra systems offer new possibilities (see e.g. Baglivo et al. (1993) and Kendall (1993)). It is the purpose of this paper to show that critical values of many nonparametric tests can be computed easily within a computer algebra system at high speed, avoiding the sophisticated approaches mentioned above. The crux is to find expressions for the probability generating function of the test statistic at hand. Since many nonparametric test statistics are of a combinatorial nature (especially those based on ranks), these generating functions can be found in the literature (David and Barton (1962) is a rich source of generating functions, many of which are important for statistics). It is interesting to note that in the statistical literature generating functions of nonparametric statistics are hardly mentioned, or used for other purposes such as deriving recursions (see e.g. Pollicello and Hettmansperger (1976)).

A major advantage of using generating functions and computer algebra systems over other approaches is that one can work directly with mathematical objects like polynomials the way we are used to do as humans, as opposed to representations of these objects in arrays etc., which are suitable for computers only. Another advantage is that computer algebra systems use infinite precision, so that rounding errors during computations do not occur. Examples of computations in Mathematica (a computer algebra system of Wolfram Research) can be found in Section 4. Furthermore, we extend the existing tables for the Jonckheere-Terpstra test. A few words on asymptotics is in order here. We want to show with this paper that with computer algebra, one can compute exact distributions of many nonparametric statistics within reasonable time. Our strategy is to compute exact distributions whenever possible. We found in all cases that when computing exact distributions becomes time-consuming, asymptotic results are sufficiently accurate. We therefore see asymptotic distributions as a useful addendum to exact computations. Also note that now we can compute exact distributions, it is possible to investigate more precisely the convergence of distributions.

This paper is organised as follows. In Section 2 we present generating functions of some rank statistics, in Section 3 we give generating functions for two goodness-of-fit tests. Section 4 contains examples of the use of a generating function in Mathematica. In Section 5 we give a new extended table of critical values of the Jonckheere-Terpstra test.
For more details about the presented tests we refer to Gibbons and Chakraborti (1992). An overview of nonparametric techniques which stresses the analogies with the parametric counterparts can be found in Van der Laan and Verdooren (1987).

We assume, unless stated otherwise, that all distributions function are continuous and that hence, ties do not occur almost surely.

2 Generating functions of rank statistics

In this section we present examples of rank statistics the null distribution of which can be easily computed using generating functions.

2.1 The Wilcoxon-Mann-Whitney test

Let $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ be independent random samples from continuous distributions with finite expectations μ_X and μ_Y, respectively, and with distribution functions $F(x)$ and $G(y) = F(y - \Delta\mu)$, respectively, where $\Delta\mu$ is an unknown shift parameter. In order to test the null hypothesis

$$H_0 : \mu_X = \mu_Y$$

against the alternative hypothesis

$$H_1 : \mu_X \neq \mu_Y,$$

Wilcoxon (1945) introduced the test statistic

$$W_{m,n} = \sum_{i=1}^{m} R_i, \tag{1}$$

where R_i is the rank of X_i in the combined sample $X_1, \ldots, X_m, Y_1, \ldots, Y_n$. Mann and Whitney (1947) introduced the statistically equivalent test statistic

$$M_{m,n} = \sum_{i=1}^{m} \#\{j : Y_j < X_i\}. \tag{2}$$

The generating function of $M_{m,n}$ was already known to Gauss (see, e.g., Andrews (1976), p. 51). A complete overview of recurrences and generating functions for $M_{m,n}$ can be found in Di Bucchianico (1996).

Theorem 2.1 *Under H_0, the probability generating function of the Mann-Whitney test statistic $M_{m,n}$ is given by*

$$\sum_{k=0}^{mn} \Pr(M_{m,n} = k)\, x^k = \frac{1}{\binom{m+n}{m}} \frac{\prod_{i=m+1}^{m+n} (1 - x^i)}{\prod_{i=1}^{n} (1 - x^i)}. \qquad (3)$$

Proof: For a proof based on recurrences we refer to Andrews (1976; Chapter 3), for a proof based on inversions we refer to David and Barton (1962; pp. 203-204). □

2.2 The Jonckheere-Terpstra test

A multi-sample analogue of the Mann-Whitney test is the Jonckheere-Terpstra test. Assume that random samples of size $n_1, \ldots, n_k$, respectively, are given from k populations. Denote by X_{ij} the jth observation in the sample from the ith population, $1 \le i \le k, 1 \le j \le n_i$. Denote by F_i the continuous cumulative distribution function of X_{ij}. Define $\phi(X_{ij})$ to be the number of observations from the first $i - 1$ populations that are smaller than X_{ij}. Let, for $i = 2, \ldots, k$,

$$S_i = \sum_{j=1}^{n_i} \phi(X_{ij})$$

and let

$$S = \sum_{j=2}^{k} S_i.$$

We wish to test the null hypothesis

$$H_0 : F_1(x) = \ldots = F_k(x) \text{ for all } x$$

against the alternative hypothesis

$$H_1 : F_1(x) \le \ldots \le F_k(x) \text{ for all } x,$$

with at least one strict inequality. For this testing problem Terpstra (1952) and Jonckheere (1954) proposed the following test statistic J (nowadays known as the Jockheere-Terpstra statistic):

$$J = 2S - M, \qquad (4)$$

where M is the maximum possible value of S, i.e. $M = \sum_{i=2}^{k} \sum_{j=1}^{i-1} n_i n_j$. Therefore, if we know the distribution of S then we also know the distribution of J.

Theorem 2.2 *Let for $i = 2, \ldots, k, N_i = \sum_{j=1}^{i-1} n_j$ and $M = \sum_{i=2}^{k} n_i N_i$. The probability generating function of S under H_0 is given by*

$$\sum_{\ell=0}^{M} \Pr(S = \ell)\, x^\ell = \prod_{i=2}^{k} \frac{1}{\binom{n_i+N_i}{n_i}} \frac{\prod_{\ell=N_i+1}^{n_i+N_i} (1 - x^\ell)}{\prod_{\ell=1}^{n_i} (1 - x^\ell)} \qquad (5)$$

Proof: It follows from Theorem 1 of Terpstra (1952) or Theorem 3 of Streitberg and Röhmel (1988) that under H_0 the random variables S_i are independent. Further, note that $\Pr(S_i = t) = \Pr(M_{n_i,N_i} = t)$, with $M_{m,n}$ the Mann-Whitney statistic defined by (2). Hence, the probability generating function of S is a product of the probability generating functions of the form (3). □

The trick to reduce the probability generating function of the Jonckheere-Terpstra test to a product of Mann-Whitney type generating functions can also be applied to other tests for partial orders (e.g. the Mack-Wolfe test for umbrella alternatives). See Streitberg and Röhmel (1988) for examples and a characterization of those alternatives for which the corresponding generalization of the Mann-Whitney test can be treated along the same lines as the Jonckheere-Terpstra test.

2.3 The Kendall rank correlation test

Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be a sample of n pairs of observations. A nonparametric correlation test is the Kendall rank correlation test. The rank correlation coefficient τ of Kendall is defined as

$$\tau = 1 - \frac{2I}{\binom{n}{2}}, \tag{6}$$

where I is the number of inversions, i.e. the number of pairs $\{(X_i, Y_i), (X_j, Y_j)\}$ such that $X_i < X_j$ and $Y_i > Y_j$ for $i < j, i = 1, \ldots n-1$ and $j = 2, \ldots, n$. The probability generating function of I has the following simple form:

Theorem 2.3 *The probability generating function of the number of inversions I is*

$$\sum_{k=0}^{\binom{n}{2}} \Pr(I = k)\, x^k = \frac{1}{n!} \prod_{k=1}^{n} \frac{x^k - 1}{x - 1}. \tag{7}$$

Proof: See Kendall and Stuart (1977; pp. 505-506). □

Recently, generating functions for the null distribution of Kendall's rank correlation statistic when ties are present in both ranks have been derived (see Valz et al. (1995) for details).

2.4 The Wilcoxon signed rank test

The Wilcoxon signed-rank test is used to test whether the median of a random sample $X_1, \ldots, X_m$ from a symmetric distribution equals m_0. Under the null hypothesis the differences $D_i = X_i - m_0, i = 1, \ldots, m$, are symmetrically distributed around zero. Ranks $\{1, \ldots, m\}$ are assigned to the absolute values of the D_i's from small to large and the rank of $|D_i|$ is denoted by R_i. The test statistic is

$$T_m = \sum_{i=1}^{m} R_i\, Z_i, \tag{8}$$

where

$$Z_i = \begin{cases} 1 & \text{if } D_i > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

Theorem 2.4 *Under H_0, the probability generating function of T_m is*

$$\sum_{i=0}^{\binom{m+1}{2}} \Pr(T_m = i)\, x^i = \frac{1}{2^m} \prod_{i=1}^{m} (1 + x^i). \tag{10}$$

Proof: The generating function is an easy consequence of the fact that under H_0, T_m has the same distribution as $U = \sum_{i=1}^{m} U_i$, where $U_i = 0$ or i, both with probability $\frac{1}{2}$. □

Computations with this generating function are so fast, that existing algorithms as in Castagliola (1996) become obsolete. Mitic (1996) reports that existing tables contain many errors.

2.5 Other one-sample rank tests

Instead of assigning ranks $\{1, \ldots, m\}$ to the $|D_i|$'s as in the Wilcoxon signed-rank test, one can also assign rank scores $a(i)$ to the $|D_i|$'s, where $a : \{1, \ldots, m\} \to \mathbb{R}$. We can now define the following test statistic:

$$T_m^* = \sum_{i=1}^{m} a(i)\, Z_i, \tag{11}$$

with Z_i as in (9). A similar argument as for T_m yields the generating function of T_m^* under H_0:

$$\sum_{i=1}^{M_a} \Pr(T_m^* = i)\, x^i = \frac{1}{2^m} \prod_{i=1}^{m} (1 + x^{a(i)}), \tag{12}$$

where $M_a = \sum_{i=1}^{m} a(i)$.

Examples of such scores include

- $a(i) = \max[0, i - \frac{m+1}{2}], i = 1, \ldots, m$. These are the scores proposed in Randles and Hogg (1973) for light-tailed distributions.
- $a(i) = \min[2i, m + 1], i = 1, \ldots, m$. These are the scores proposed in Pollicello and Hettmansperger (1976) for heavy-tailed distributions.
- $a(i) = \Phi^{-1}\left(\frac{1}{2} + \frac{i}{2(m+1)}\right), i = 1, \ldots m$, where Φ^{-1} is the inverse of the standard normal cumulative distribution function. These are the inverse normal scores. Note that the scores in this case are not rational and that exact computations are not possible unless we approximate the scores by rational numbers.

3 Generating functions for goodness-of-fit tests

3.1 The Kolmogorov one-sample test

The Kolmogorov one-sample test is used to test whether the sample $X_1, \ldots, X_m$ comes from a certain distribution function. The null hypothesis is

$$H_0 : F(x) = F_0(x) \quad \text{for all } x,$$

where $F(x)$ is the continuous distribution function of the observations and $F_0(x)$ is a given continuous distribution function. The two-sided alternative is

$$H_1 : F(x) \neq F_0(x) \quad \text{for at least one } x\,.$$

The test statistic is

$$D_m = \sup_x |F_m(x) - F_0(x)|\,,$$

where $F_m(x)$ denotes the empirical distribution function defined by $F_m(x) := \frac{1}{m}\#\{\ell : X_\ell \leq x, \ell = 1, \dots, m\}$.
For the one-sided alternative hypotheses $H_1 : F(x) \leq F_0(x)$ and $H_1 : F(x) \geq F_0(x)$, for all x and with strict inequality for at least one x, the test statistics are

$$D_m^+ = \sup_x\{F_0(x) - F_m(x)\} \quad \text{and} \quad D_m^- = \sup_x\{F_m(x) - F_0(x)\}, \tag{13}$$

respectively.
Kemperman (1957) (see also Niederhausen (1981)) gives the following implicit generating function, which holds under H_0:

$$\sum_{k=0}^{\infty} \Pr\left(-r < k\,D_k^- < s\right)\frac{(kx)^k}{k!} = \frac{Q_r(x)\,Q_s(x)}{Q_{r+s}(x)}, \tag{14}$$

for $x < 1/e$, where $Q_t(x) = \sum_{i=0}^{\lfloor t \rfloor}(i-t)^i\,x^i/i!$, and $\lfloor t \rfloor$ denotes the largest integer not exceeding t. Under H_0, D_m^+ is distributed as $-D_m^-$. We also know that $D_m = \max(D_m^+, D_m^-)$. Therefore,

$$\Pr(D_m \geq \frac{s}{k}) = 1 - \Pr(\max(-D_m^-, D_m^-) < \frac{s}{k}) = 1 - \Pr\left(-s < k\,D_m^- < s\right) \tag{15}$$

So in order to compute the exact distribution of D_m we need the coefficient of x^m of the right-hand side of (14). This can be done by expanding the right-hand side of (14) by hand, which yields an explicit expression for the null distribution of D_m (cf. Kemperman (1957)). Alternatively, we may ask Mathematica to compute the coefficient of x^m of the right-hand side of (14). Critical values can be computed using a numerical procedure for root finding. We refer to Section 4 for further details. Tail probabilities for D_m^+ or D_m^- can be obtained by choosing $r = m$ or $s = m$.

For the corresponding two-sample Smirnov test analogous generating functions exist for sample sizes that are not relatively prime (Kemperman (1957) and Niederhausen (1981)). For a combinatorial explanation of the influence of relative primeness of the sample sizes on this statistic, see Di Bucchianico and Loeb (1997).

3.2 The Kuiper test

Kuiper (1960) suggested

$$K_m = D_m^+ + D_m^-,$$

where D_m+ and D_m^- are defined in the previous subsection, as a Kolmogorov-type test statistic on a circle. It has the property that if the observations are circular data, its value does not depend on the choice of the origin for measuring x. From Niederhausen (1981) we obtain that the generating function of K_m has the same form as (14):

$$\sum_{k=0}^{\infty} \Pr\left(K_k < \frac{s}{k}\right) \frac{(kx)^k}{k!\,k} = x\,\frac{Q_{s-1}(x)Q_1(x)}{Q_s(x)}, \tag{16}$$

with $Q_s(x)$ as in (14). Thus we can compute tail probabilities in the same way as for the Kolmogorov test.

4 Generating functions in Mathematica

In this section we show how we use the generating funtions to obtain tail probabilities using Mathematica. We give examples for the Mann-Whitney test and the Kolmogorov test. One can deal with the other tests in the same way.

4.1 Implementation of the Mann-Whitney test

```
MannWhitneyGf[m_,n_,x_:x]:=
Module[{i,mini = Min[m,n],maxi=Max[m,n]}, Expand[Factor[Product[
1-x^i,{i,maxi+1,maxi+mini}]/Product[1-x^i,i,1,mini]]]]

MannWhitneyFrequencies[m_,n_]:=
Module[{x,i,mini = Min[m,n],maxi=Max[m,n]}, Drop[FoldList[
Plus,0,CoefficientList[ MannWhitneyGf[m,n],x]],1]]

MannWhitneyRightTail[m_,n_,k_]:=
N[1-(Part[MannWhitneyFrequencies[m,n],k]/Binomial[m+n,n])]

MannWhitneyRightCriticalValue[m_,n_,alpha_]:=
Module[help= Length[Select[ MannWhitneyFrequencies[m,n],# < (1-1/2 * alpha)*
Binomial[m+n,n]&]]+1; If[m * n/2 <= help && help <= m*n,help,"*"]]

MannWhitneyGf[2,3]
```

$1 + x^1 + 2\,x^2 + 2\,x^3 + 2\,x^4 + x^5 + x^6$

```
MannWhitneyFrequencies[2,3]
```

$\{1, 2, 4, 6, 8, 9, 10\}$

```
MannWhitneyRightTail[2,3,5]
0.2

MannWhitneyRightCriticalValue[3,2,0.2]
6

Timing[MannWhitneyRightCriticalValue[25,20,0.05]]
{13.46 Second,337}
```

Explanation
The MannWhitneyGf[m_,n_] function factors (Factor) and expands (Expand) formula (3) without the constant $1/\binom{m+n}{m}$. With the aid of the local variables $q, i, mini, maxi$ (Module) the function MannWhitneyFrequencies[m_,n_] generates a list $\{c_1, \ldots, c_{mn}\}$ of coefficients (CoefficientList) and transforms it into $\{0, c_1, c_1 + c_2, \ldots, \sum_{i=1}^{mn} c_i\}$ (FoldList). Finally, it drops the '0' in this list (Drop). Thus, the function MannWhitneyFrequencies[m_,n_] generates up to a factor $1/\binom{m+n}{m}$ the cumulative distribution of the statistic $M_{m,n}$. The function MannWhitneyRightTail[m_,n_,k_] takes the kth part of MannWhitneyFrequencies[m_,n_] and divides it by $\binom{m+n}{m}$. This quotient is subtracted from 1 to obtain the right-tail probability of k.

The function MannWhitneyRightCriticalValue[m_,n_,alpha_] computes the right critical value corresponding to the two-sided confidence level α. By using Select it selects all frequencies with right-tail probability larger or equal to $\frac{1}{2}\alpha$. We note that for every $k, k = 0, \ldots, mn$ the probability that the Mann-Whitney statistic equals k is positive. Therefore, the length (Length) of the list that results after applying Select equals the largest value for which the right-tail probability is larger or equal to $\frac{1}{2}\alpha$. We add one to this number; the result (help) is the right critical value if (If) $\frac{1}{2}mn \leq$ help $\leq mn$.

The example shows that the right-tail probability in the case $m = 2, n = 3$ equals 0.2 for $k = 5$. The right critical value for this case with $\alpha = 0.2$ equals 6. To give an indication of the speed of this method we use the Timing function for obtaining the time needed for computing the right critical value for the case $m = 20, n = 25, \alpha = 0.05$. We see that the right critical value equals 337 and that the computing time is 13.46 CPU seconds on a SunSPARC10.

Remark:
In order to obtain right-tail probabilities one can also use the formula:

$$\sum_{k=0}^{\infty} \Pr(T > k)x^k = \frac{1 - \sum_{k=0}^{\infty} \Pr(T = k)\, x^k}{1 - x}, \tag{17}$$

where T is an arbitrary rank statistic.
The advantage of using (17) is that we immediately know the right-tail probability of all k. However, expanding (17) is more time consuming than our method since it involves division.

4.2 Implementation of the Kolmogorov test

```
Q[s_]:=
Module[{i}, Sum[(i-s)^i*(x^i)/i!,{i,0,Floor[s]}]]

F[s_,m_]:=
Normal[Series[Simplify[(Q[s]^2)/Q[2*s]],{x,0,m}]]

F[3.5,10]
```

$1 + x + 2\,x^2 + 4.5\,x^3 + 10.6615\,x^4 + 25.8969\,x^5 + 63.6287\,x^6 + 157.128\,x^7 + 388.858\,x^8 +$
$963.186\,x^9 + 2386.57\,x^{10}$

```
Kolmogorov[d_,m_]:=
N[1-Coefficient[F[m*d,m],x,m]*m!/(m^m)]
```

```
Kolmogorov[0.35,10]
0.866039

Timing[Kolmogorov[0.23,40]]
{0.96 Second, 0.035}
```

Explanation

Floor[s] represents $\lfloor s \rfloor$. The function F[s_,m_] simplifies (Simplify) the right-hand side of (14) and then expands it into a Taylor polynomial (Series) of degree m, including an order term which is removed by applying Normal. The function Kolmogorov [d_,m_] first computes the coefficient of x^m (Coefficient) in F[m*d,m] and multiplies this by $\frac{m!}{m^m}$ for obtaining $\Pr(-d < D_m^- < d)$, where D_m^- is the Kolmogorov statistic as in (13). From equality (15) we know that subtracting this result from one gives us $\Pr(D_m \geq d)$. The function N provides a numerical result instead of a fraction. The example shows F[m*d,m] for $d = 0.35$ and $m = 10$ and it gives the right-tail probability for these values of d and m. We use the Timing function to show that this method is very fast (computation on a SunSPARC10), even for $m = 40$.

5 Table for the Jonckheere-Terpstra test

With the generating function (5) we extended the existing tables for the Jonckheere-Terpstra test statistic. In Odeh (1971) the following cases were tabulated: $k = 3, 2 \leq n_1 \leq n_2 \leq n_3 \leq 8; k = 4, 5, 6, n_1 = \ldots = n_k = 2(1)6$. We tabulated the cases $n_i = n_j, i \neq j, i, j = 1, \ldots k$. Our tables are 2 to 5 times larger than the existing ones. Tabulation of the cases $n_i \neq n_j$ requires a lot of space, because there are so many different cases. For these cases we recommend to use our Mathematica packages for computing tail probabilities. A star denotes that a critical value does not exist for this case.

k	n	0.2	0.1	0.05	0.025	0.01	0.005
3	2	9	10	11	12	*	*
	3	18	20	22	23	25	25
	4	31	34	36	38	40	42
	5	47	51	54	57	60	62
	6	66	71	75	79	83	86
	7	88	95	100	105	110	114
	8	113	121	128	134	140	145
	9	142	152	160	166	174	180
	10	173	185	194	202	212	218
	11	208	222	232	242	252	260
	12	246	261	274	284	297	305
	13	287	304	318	330	344	353
	14	332	351	366	380	395	406
	15	379	400	418	432	450	461
	16	430	453	472	488	507	520
	17	483	509	530	548	568	582
	18	540	568	591	610	633	648
	19	600	630	655	676	701	717
	20	663	696	722	745	772	790
	21	729	764	793	818	846	865
	22	799	836	867	893	924	944
	23	871	911	944	972	1005	1027
	24	947	989	1024	1054	1089	1113
	25	1025	1071	1108	1140	1177	1202
	26	1107	1155	1194	1228	1268	1294
	27	1192	1243	1284	1320	1362	1390
	28	1280	1333	1377	1415	1459	1489
	29	1371	1427	1474	1514	1560	1591
	30	1465	1524	1573	1615	1664	1697
	31	1562	1624	1676	1720	1771	1806
	32	1662	1728	1782	1828	1882	1918
	33	1766	1834	1891	1939	1996	2034
	34	1872	1944	2003	2054	2113	2153
	35	1982	2057	2118	2171	2233	2275
	36	2095	2173	2237	2292	2356	2400
	37	2210	2292	2358	2416	2483	2528
	38	2329	2414	2483	2543	2613	2660
	39	2451	2539	2611	2674	2746	2795
	40	2576	2667	2742	2807	2883	2934
4	2	16	18	19	21	22	23
	3	34	37	40	42	44	45
	4	58	63	67	70	73	76

k	n	0.2	0.1	0.05	0.025	0.01	0.005
4	5	89	95	100	105	110	114
	6	126	134	141	147	154	158
	7	169	179	188	196	204	210
	8	218	231	242	251	262	269
	9	274	290	302	313	326	334
	10	336	354	369	382	397	407
	11	404	425	443	457	474	486
	12	479	503	522	539	559	572
	13	560	587	609	628	650	665
	14	647	677	701	723	747	764
	15	740	773	801	824	851	870
	16	839	876	906	932	962	983
	17	945	985	1018	1047	1080	1102
	18	1057	1101	1137	1168	1203	1228
	19	1175	1222	1261	1295	1334	1360
	20	1299	1350	1392	1429	1471	1499
5	2	26	28	30	32	33	35
	3	54	59	62	65	69	71
	4	94	100	106	110	116	119
	5	144	153	160	167	174	179
	6	204	216	226	235	244	251
	7	275	290	303	313	325	334
	8	357	375	390	403	418	428
	9	448	470	488	504	522	534
	10	550	576	597	615	636	650
	11	663	693	717	738	762	778
	12	786	819	847	871	899	917
	13	919	957	988	1015	1046	1067
	14	1062	1105	1140	1170	1205	1228
	15	1216	1263	1302	1335	1374	1400
6	2	37	40	43	45	47	49
	3	80	85	90	94	98	101
	4	138	147	154	160	167	171
	5	212	224	234	242	252	259
	6	302	317	330	342	354	363
	7	407	427	443	457	474	484
	8	528	552	572	589	609	623
	9	664	693	717	738	761	777
	10	816	850	878	902	930	949
	11	984	1023	1055	1083	1115	1136
	12	1167	1211	1248	1279	1316	1341

Table 1: Right critical values for the Jonckheere-Terpstra test, $n_i = n_j = n$

k	n	0.2	0.1	0.05	0.025	0.01	0.005
7	2	51	55	58	61	64	66
	3	109	117	122	127	133	137
	4	190	201	210	218	227	233
	5	293	308	321	332	344	352
	6	418	438	454	468	484	495
	7	564	589	610	628	648	662
	8	732	763	788	810	835	852
	9	922	959	989	1015	1045	1065
	10	1133	1176	1211	1242	1277	1301
8	2	66	71	75	78	82	85
	3	144	153	160	166	173	178
	4	251	264	275	285	296	303
	5	387	406	421	434	449	460
	6	552	577	597	614	634	648
	7	746	777	802	824	849	867
	8	969	1007	1038	1064	1095	1116
	9	1221	1266	1303	1335	1371	1396
9	2	84	90	95	99	103	106
	3	183	194	202	210	218	224
	4	320	336	349	360	373	382
9	5	494	516	535	550	568	581
	6	705	734	759	779	803	819
	7	954	990	1021	1047	1077	1097
	8	1239	1284	1321	1353	1390	1415
10	2	104	111	116	121	126	130
	3	227	239	249	258	268	274
	4	397	416	431	444	460	470
	5	614	640	661	680	701	716
	6	877	911	939	964	992	1011
	7	1186	1229	1265	1295	1331	1355
11	2	126	134	140	145	152	156
	3	276	290	301	311	323	330
	4	483	504	522	537	555	567
	5	746	777	801	823	847	864
	6	1067	1106	1139	1167	1199	1221
12	2	150	159	166	172	179	184
	3	329	345	358	369	383	391
	4	576	601	621	639	659	672
	5	892	926	954	979	1007	1026

Table 2: Right critical values for the Jonckheere-Terpstra test, $n_i = n_j = n$

References

Andrews, G. E. (1976), *The Theory of Partitions, Encyclopedia of Mathematics and its Applications,* Addison-Wesley, Reading.

Baglivo, J., Pagano, M. and Spino, C. (1993), "Symbolic computation of permutation distributions," *Proceedings of the Statistical Computing Section ASA*, 218-223.

Castagliola, P. (1996), "An optimized algorithm for computing Wilcoxon's T_n^+ statistic when n is small", *Computational Statistics,* 11, 1-10.

David, F. N. and Barton, D.E. (1962), *Combinatorial Chance*, Charles Griffin & Co., London.

Di Bucchianico, A. (1996), " Combinatorics, computer algebra, and the Wilcoxon-Mann-Whitney test", *COSOR Memorandum 96-24*, Eindhoven University of Technology, The Netherlands.

Di Bucchianico, A. and Loeb, D.E. (1997), "Dominance refinements of the Smirnov two-sample test," *Journal of Statistical Planning and Inference*, to appear.

Edgington, E.S. (1995), *Randomization tests*, Marcel Dekker, New York.

Gibbons, J. D. and Chakraborti, S. (1992), *Nonparametric Statistical Inference*, Marcel Dekker, New York.

Good, P. (1994), *Permutation tests: a practical guide to resampling methods for testing hypotheses,* Springer, Berlin.

Jonckheere, A. R. (1954), "A distribution-free k-sample test against ordered alternatives," *Biometrika*, 41, 133-145.

Kemperman, J. H. B. (1957), "Some exact formulae for the Kolmogorov-Smirnov distributions," *Indagationes Mathematicae*, 19, 535-540.

Kendall, M. and Stuart, A. (1977), *The Advanced Theory of Statistics,* (Vol. 2), Charles Griffin & Company Ltd., London.

Kendall, W. S. (1993), "Computer algebra in probability and statistics," *Statistica Neerlandica*, 47, 9-25.

Kuiper, N. H. (1960), "Tests concerning random points on a circle," *Koninklijke Nederlandse Academie van Wetenschappen Proceedings Serie A*, 63, 38-47.

Mann, H. B. and Whitney, D. R. (1947), "On a test of whether one of two random variables is stochastically larger than the other," *Annals of Mathematical Statistics*, 18, 50-60.

Mitic, P. (1996), "Critical values for the Wilcoxon signed rank statistic," *The Mathematica Journal,* 6, 73-77.

Niederhausen, H. (1981), "Sheffer polynomials for computing exact Kolmogorov-Smirnov and Rényi type distributions," *Annals of Statistics*, 923-944.

Odeh, R. E. (1971), "On Jonckheere's k-sample test against ordered alternatives," *Technometrics*, 13, 912-918.

Pagano, M. and Tritchler, D. (1983), "Obtaining permutation distributions in polynomial time," *Journal of the American Statistical Association*, 78, 435-441.

Pollicello, G. E. and Hettmansperger, T. P. (1976), "Adaptive robust procedures for the one-sample location problem," *Journal of American Statistical Association*, 71, 624-633.

Randles, R. H. and Hogg, R. V. (1973), "Adaptive distribution-free tests," *Communications in Statistics*, 2, 337-356.

Streitberg, B. and Röhmel, J. (1986), "Exact distributions for permutation and rank tests: an introduction to some recently published algorithms," *Statistical Software Newsletter*, 12, 10-17.

Streitberg, B. and Röhmel, J. (1988), "Exact nonparametrics for partial order tests," *Computational Statistics Quarterly*, 1, 23-41.

Terpstra, T. J. (1952), "The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking," *Indagationes Mathematicae*, 14, 327-333.

Valz, P. D., McLeod, A. I. and Thompson, M. E. (1995), "Cumulant generating function and tail probability approximations for Kendall's score with tied rankings," *Annals of Statistics*, 23, 144-160.

Van der Laan, P. and Verdooren, L.R. (1987), "Classical analysis of variance methods and nonparametric counterparts," *Biometrical Journal*, 29, 635-665.

Wilcoxon, F. (1945), "Individual comparisons by ranking methods," *Biometrics* 1, 80-83.