Taylor & Francis
Taylor & Francis Group

Communications in Statistics - Simulation and Computation

ISSN: 0361-0918 (Print) 1532-4141 (Online) Journal homepage: http://www.tandfonline.com/loi/lssp20

# An Algorithm for Computing the Exact Distribution of the Kruskal–Wallis Test

Won Choi , Jae Won Lee , Myung-Hoe Huh & Seung-Ho Kang

**To cite this article:** Won Choi , Jae Won Lee , Myung-Hoe Huh & Seung-Ho Kang (2003) An Algorithm for Computing the Exact Distribution of the Kruskal–Wallis Test, Communications in Statistics - Simulation and Computation, 32:4, 1029-1040, DOI: 10.1081/SAC-120023876

**To link to this article:** http://dx.doi.org/10.1081/SAC-120023876

MARCEL DEKKER, INC. • 270 MADISON AVENUE • NEW YORK, NY 10016



COMMUNICATIONS IN STATISTICS
Simulation and Computation®
Vol. 32, No. 4, pp. 1029–1040, 2003



# An Algorithm for Computing the Exact Distribution of the Kruskal–Wallis Test

**Won Choi,[1] Jae Won Lee,[1] Myung-Hoe Huh,[1] and Seung-Ho Kang[2,*]**

[1]Department of Statistics, Korea University, Seoul, Korea
[2]Department of Statistics, Ewha Womans University, Seoul, Korea

## ABSTRACT

The Kruskal–Wallis test is a popular nonparametric test for comparing $k$ independent samples. In this article we propose a new algorithm to compute the exact null distribution of the Kruskal–Wallis test. Generating the exact null distribution of the Kruskal–Wallis test is needed to compare several approximation methods. The 5% cut-off points of the exact null distribution which StatXact cannot produce are obtained as by-products. We also investigate graphically a reason that the exact and approximate distributions differ, and hope that it will be a useful tutorial tool to teach about the Kruskal–Wallis test in undergraduate course.

*Correspondence: Seung-Ho Kang, Department of Statistics, Ewha Womans University, 11-1 DaeHyun-dong SeoDaeMun-Ku, Seoul, Korea 120-750; Fax: 82-2-3277-3607; E-mail: seungho@ewha.ac.kr.



DOI: 10.1081/SAC-120023876
0361-0918 (Print); 1532-4141 (Online)

www.dekker.com

MARCEL DEKKER, INC. • 270 MADISON AVENUE • NEW YORK, NY 10016





*Key Words:* Rank; Recursive formula; Nonparametric ANOVA; Permutation.



## 1. INTRODUCTION

Let $X_{11},\ldots,X_{1n_1},\ldots,X_{k1},\ldots,X_{kn_k}$ denote $k$ independent random samples from continuous distributions with the cumulative distribution function $F(x-\theta_1),\ldots,F(x-\theta_k)$, respectively, where $\theta_i$ is the median of the $i$th population for $i=1,\ldots,k$. We consider the problem of testing the null hypothesis $H_0:\theta_1=\cdots=\theta_k$ against the general alternative $H_1:\theta_i\neq\theta_j$ for at least one $i\neq j$. This is often referred to as the one-way layout problem. Let $R_{ij}^*$ be the rank of $X_{ij}$ among $X_{11},\ldots,X_{1n_1},\ldots,X_{k1},\ldots,X_{kn_k}$. That is, $R_{ij}^*$ is the rank of $X_{ij}$ in the pooled sample of $N=\sum_{i=1}^k n_i$ observations. Let $R_i=\sum_{j=1}^{n_i} R_{ij}^*$ be the ranks sum associated with the $i$th population for $i=1,\ldots,k$. Kruskal and Wallis (1952) proposed the following statistic

$$T=\frac{12}{N(N+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i}-3(N+1)$$

and the corresponding $\alpha$-level test which

$$\text{rejects } \mathrm{H}_0 \text{ if } \mathrm{T}\geq t(\alpha)$$

where $t(\alpha)$ is the upper $100\alpha$th percentile for the null distribution of $T$.

To conduct the above Kruskal–Wallis test we need the critical values $t(\alpha)$. This can be accomplished for given integers $k,n_1,\ldots,n_k$ by calculating the values of $T$ for each $N!$ permutations that are possible for the ranks $R_{11}^*,\ldots,R_{kn_k}^*$, and then tabulating the null distribution of $T$ from the fact that each of these permutations is equally likely under the null hypothesis. Such critical values can be found in Table 1 of Lehmann (1975) for $k=3$ and $n_i\leq 5, i=1,2,3$. Additional tables are given by Iman et al. (1975).

Many approximation distributions of $T$ under the null hypothesis were proposed because of the computational difficulty of computing the exact null distribution of $T$. Under the null hypothesis the statistic $T$ has a limiting chi-square distribution with $k-1$ degrees of freedom if $\text{minimum}(n_1,\ldots,n_k)\to\infty$, with $n_i/N\to\lambda_i, 0<\lambda_i<1$, for $i=1,\ldots,k$ (Kruskal and Wallis, 1952). There are several other approximations (Iman and Davenport, 1976; Kruskall and Wallis, 1952; Wallace, 1959). It is desirable to generate the exact null distribution of $T$

MARCEL DEKKER, INC. • 270 MADISON AVENUE • NEW YORK, NY 10016







in order to

1. Compare several approximation methods.
2. Decide appropriate number of replications when Monte Carlo method is used as an approximation (Chow et al., 1974).
3. Obtain the exact cut-off points as by-products when the sample size is small.

Here we would like to emphasize that the problem of generating the exact null distribution is different from that of computing exact $p$-values. Since it is enough to investigate only the tail area of the exact distribution, an elegant method such as the network algorithm (Mehta and Patel, 1983) can be employed. In the network algorithm the reference set is expressed as all paths through a directed acyclic network. For simply calculating $p$-values one can drastically save the computation time by implicit enumeration. However, implicit enumeration is not feasible in generating the exact null distribution where all possible cases should be taken into account. In this article, we propose a new algorithm of computing the exact null distribution of $T$ by improving the algorithm described in Iman et al. (1975).

## 2. ALGORITHMS

There has been the two methods of generating the exact null distribution of $T$. We fix $k = 3$ for simplicity of explanation.

### 2.1. Method 1

For given $n_1, n_2, n_3$ the method generates all possible $(n_1 + n_2 + n_3)!/(n_1!n_2!n_3!)$ cases and generates the exact null distribution of $T$. In order to see the computational difficulty the numbers of all possible cases are calculated when $n_i = n$ for $i = 1, 2, 3$.

In Table 1 we see that the number of all possible cases increases with the exponential speed as $n$ increases. This method requires the huge amount of time even when $n$ is moderate size.

### 2.2. Method 2

Iman et al. (1975) proposed the following algorithm. The exact null distribution of $T$ depends on only $R_i$ and $n_i, i = 1, 2, 3$. Therefore the

***Table 1.*** The number of all possible rank sums $(R_1, R_2, R_3)$ under $H_0$ when $n_1 = n_2 = n_3 = n$.

| $n$ | $(n_1+n_2+n_3)!/(n_1!n_2!n_3!)$ | $n$ | $(n_1+n_2+n_3)!/(n_1!n_2!n_3!)$ |
|---|---|---|---|
| 2 | 90 | 9 | 227,873,431,500 |
| 3 | 1,680 | 10 | 5,550,996,791,340 |
| 4 | 34,650 | 11 | 136,526,995,463,040 |
| 5 | 756,756 | 12 | 3,384,731,762,521,200 |
| 6 | 17,153,136 | 13 | 84,478,098,072,866,400 |
| 7 | 399,072,960 | 14 | 2,120,572,665,910,728,000 |
| 8 | 9,465,511,770 | 15 | 53,494,979,785,374,631,680 |

exact null distribution of $T$ can be obtained from the distribution of rank totals $(R_1, R_2, R_3)$. Let $(r_1, r_2, r_3)$ be an observed vector of rank totals $(R_1, R_2, R_3)$ for given sample size $(n_1, n_2, n_3)$ with $n_1 + n_2 + n_3 = N$. Since each possible case of rank configuration to produce $(r_1, r_2, r_3)$ is equally likely under the null hypothesis, we may interpret the number of such cases as frequencies and can obtain the distribution of $(R_1, R_2, R_3)$. Let $W(r_1, r_2, r_3|n_1, n_2, n_3)$ denote the number of all possible rank configurations which produces the rank totals $(R_1 = r_1, R_2 = r_2, R_3 = r_3)$ with the sample size $(n_1, n_2, n_3)$. Suppose that we know the values of $W(r_1, r_2, r_3|n_1, n_2, n_3)$ with the sample size of $n_1 + n_2 + n_3 = N - 1$ for all possible cases of $(r_1, r_2, r_3)$ and we would like to obtain $W(r_1, r_2, r_3|n_1, n_2, n_3)$ with the sample size of $n_1 + n_2 + n_3 = N$. The rank $N$ will belong to only one of the three populations. If the subject with the rank $N$ is in the first population, the sample sizes except that the subject with the rank $N$ would be $(n_1 - 1, n_2, n_3)$ and the rank totals would be $(r_1 - N, r_2, r_3)$. Hence $W(r_1, r_2, r_3|n_1, n_2, n_3)$ with the sample size $N$ can be determined from the assumption that $W(r_1, r_2, r_3|n_1, n_2, n_3)$ with the sample size $N - 1$ for all possible cases of $(r_1, r_2, r_3)$ are known. Similarly, the number of cases of $(R_1 = r_1, R_2 = r_2, R_3 = r_3)$ can be obtained if the subject with the rank $N$ belongs to either the second or third population. Therefore, we have the following recursive formula.

$$\begin{aligned} W(r_1, r_2, r_3|n_1, n_2, n_3) = {} & W(r_1 - N, r_2, r_3|n_1 - 1, n_2, n_3) \\ & + W(r_1, r_2 - N, r_3|n_1, n_2 - 1, n_3) \\ & + W(r_1, r_2, r_3 - N|n_1, n_2, n_3 - 1) \end{aligned}$$

This method is faster, but requires much more storage space than the Method 1 (Iman et al., 1975).

MARCEL DEKKER, INC. • 270 MADISON AVENUE • NEW YORK, NY 10016





### 2.3. Disadvantages of Method 2

A simple example is provided to show a disadvantage of the Method 2. We consider the case of $n_1 = 1, n_2 = 1,$ and $n_3 = 1$ as an example. The $W$'s with the sample size $N = 2$ is as follows.

$(n_1, n_2, n_3) = (1, 1, 0)$

| $r_1$ | $r_2$ | $r_3$ | $W(r_1, r_2, r_3 \mid 1, 1, 0)$ |
|---|---|---|---|
| 1 | 2 | 0 | 1 |
| 2 | 1 | 0 | 1 |

$(n_1, n_2, n_3) = (1, 1, 0)$

| $r_1$ | $r_2$ | $r_3$ | $W(r_1, r_2, r_3 \mid 1, 0, 1)$ |
|---|---|---|---|
| 1 | 0 | 2 | 1 |
| 2 | 0 | 1 | 1 |

$(n_1, n_2, n_3) = (0, 1, 1)$

| $r_1$ | $r_2$ | $r_3$ | $W(r_1, r_2, r_3 \mid 0, 1, 1)$ |
|---|---|---|---|
| 0 | 1 | 2 | 1 |
| 0 | 1 | 2 | 1 |

Then we can use the recursive formula.

$$\begin{aligned} W(1,2,3 \mid n_1 = 1, n_2 = 1, n_3 = 1) &= W(1-3,2,3 \mid n_1 = 1-1, n_2 = 1, n_3 = 1) \\ &\quad + W(1,2-3,3 \mid n_1 = 1, n_2 = 1-1, n_3 = 1) \\ &\quad + W(1,2,3-3 \mid n_1 = 1, n_2 = 1, n_3 = 1-1) \\ &= 0 + 0 + 1 \end{aligned}$$

In order to finish the above computation we need to search $W(1, 2, 3-3 \mid 1, 1, 0)$ in the table of $W(r_1, r_2, r_3 \mid 1, 1, 0)$. We do not need to search the values of $W(1-3, 2, 3 \mid 0, 1, 1)$ and $W(1, 2-3, 3 \mid 1, 0, 1)$, because some values of the rank totals are negative. But, in general we should search the three tables. If the maximum number of rows among the three tables of $W(r_1, r_2, r_3 \mid n_1, n_2, n_3)$ with the sample size $n_1 + n_2 + n_3 = N - 1$ are $m$ and the number of rows in the table with the sample size $N$ is $cm$ where $c$ is a constant ($1 \le c \le 3$), the Method 2 needs $3m \times cm$ operations in the worst case to make a table with the

sample size of $N$ from the three tables with the sample size of $N-1$. Later we will show that the Method 3 needs less computations.

There is another disadvantage of the Method 2. In order to use the recursive formula we should know the all possible vectors of rank totals $(R_1, R_2, R_3)$ with the sample size $n_1+n_2+n_3=N$. But, the problem is the difficulty of knowing those vectors. Probably the following range of rank totals $(R_1, R_2, R_3)$ can be obtained easily.

$$n_1(n_1+1)/2 \le R_1 \le N(N+1)/2-(n_2+n_3)(n_2+n_3+1)/2$$
$$n_2(n_2+1)/2 \le R_2 \le N(N+1)/2-(n_1+n_3)(n_1+n_3+1)/2$$
$$n_3(n_3+1)/2 \le R_3 \le N(N+1)/2-(n_1+n_2)(n_1+n_2+1)/2$$

with constraint $R_1+R_2+R_3=N(N+1)/2$. However, $(R_1, R_2, R_3)$ does not take on all values in the above ranges. For example, when $n_1=4$, $n_2=3$, and $n_3=1$, $(R_1, R_2, R_3)=(11,20,5),(25,7,4)$ exist in the above range, but we cannot produce rank totals $(11,20,5)$ and $(25,7,4)$ using the positive integers from 1 to 8.

### 2.4. Method 3

We propose the new method, called Method 3, by improving the Method 2. In the Method 2, first we decide the vectors of rank totals $(R_1, R_2, R_3)$ with the sample size of $N$ whose frequency we would like to count. Second, from the recursive formula, the frequency is expressed as the sum of frequencies of rank totals $(R_1, R_2, R_3)$ with the sample size of $N-1$. In other words, we start the stages with the sample size of $N$ and move to the stages with smaller sample size until we reach the stages with the sample size of one. On the other hand, in the Method 3, the stages are processed in the opposite way. We begin with the stages with the sample size of one and build up the stages with larger sample size until we construct the stages with the sample size of $N$.

The new proposed method uses the following theorem. For simplicity let $W(n_1,\ldots,n_k)=W(r_1,\ldots,r_k|n_1,\ldots,n_k)$.

**Symmetry Theorem.** *$W(n_1,\ldots,n_k)$ is equal to $W(n'_1,\ldots,n'_k)$ if $(n'_1,\ldots,n'_k)$ is a permutation of $(n_1,\ldots,n_k)$.*

The proof is straightforward from the definition of $W(n_1,\ldots,n_k)$. The Method 2 does not use the above symmetry. So for example, when $k=3$, we need $W(n_1=0,n_2=1,n_3=1)$, $W(n_1=1,n_2=0,n_3=1)$, and $W(n_1=1,n_2=1,n_3=0)$ to compute $W(n_1=1,n_2=1,n_3=1)$.

MARCEL DEKKER, INC. • 270 MADISON AVENUE • NEW YORK, NY 10016







However, with the symmetry it is enough to have only $W(n_1 = 0, n_2 = 1, n_3 = 1)$. When $n_1 = \cdots = n_k = n$, the Method 2 requires $n^3$ $W$'s while the Method 3 needs only $n(n+1)(n+2)/6$ $W$'s, because

$$\sum_{i=1}^{n}\sum_{j=1}^{i}\sum_{k=1}^{j} 1 = n(n+1)(n+2)/6.$$

Next we describe an improvement in data structure. The case of constructing $W(n_1 = 1, n_2 = 1, n_3 = 1)$ is considered as an example. Note that 3 is the largest rank when $n_1 = n_2 = n_3 = 1$.

1. Add 3 to $r_3$ in the table $(n_1 = 1, n_2 = 1, n_3 = 0)$.

| $r_1$ | $r_2$ | $r_3$ | $W(r_1, r_2, r_3 \mid 1, 1, 0)$ |
|---|---|---|---|
| 1 | 2 | $0+3$ | 1 |
| 2 | 1 | $0+3$ | 1 |

2. Add 3 to $r_2$ in the table $(n_1 = 1, n_2 = 0, n_3 = 1)$.

| $r_1$ | $r_2$ | $r_3$ | $W(r_1, r_2, r_3 \mid 1, 0, 1)$ |
|---|---|---|---|
| 1 | $0+3$ | 2 | 1 |
| 2 | $0+3$ | 1 | 1 |

3. Add 3 to $r_1$ in the table $(n_1 = 0, n_2 = 1, n_3 = 1)$.

| $r_1$ | $r_2$ | $r_3$ | $W(r_1, r_2, r_3 \mid 0, 1, 1)$ |
|---|---|---|---|
| $0+3$ | 1 | 2 | 1 |
| $0+3$ | 2 | 1 | 1 |

4. Merge the three tables and sort by $r_1, r_2$, and $r_3$. If there are same $(r_1, r_2, r_3)$'s, then sum the corresponding $W$'s up, remaining one $(r_1, r_2, r_3)$ and delete the others.

| $r_1$ | $r_2$ | $r_3$ | $W(r_1, r_2, r_3 \mid 1, 1, 1)$ |
|---|---|---|---|
| 1 | 2 | 3 | 1 |
| 1 | 3 | 2 | 1 |
| 2 | 1 | 3 | 1 |
| 2 | 3 | 1 | 1 |
| 3 | 1 | 2 | 1 |
| 3 | 2 | 1 | 1 |

MARCEL DEKKER, INC. • 270 MADISON AVENUE • NEW YORK, NY 10016





The Method 3 is very simple and does not need to know the ranges of $r_1, r_2$, and $r_3$. Let $m$ denote the maximum number of rows among the three tables of $W(r_1, r_2, r_3|n_1, n_2, n_3)$ with the sample size $n_1 + n_2 + n_3 = N - 1$. Then the maximum possible number of rows obtained in the above step 4 is $3m$. The Method 3 has only $3m\log(3m)$ operations in the worst case, because the algorithm search only the combined table just once and a well-developed sorting algorithm (e.g., mergesort) has only $m\log(m)$ operations when there are $m$ elements.

## 3. THE COMPARISON BETWEEN THE EXACT AND THE CHI-SQUARE APPROXIMATE NULL DISTRIBUTION OF $T$

We investigate a reason that the two null distributions of $T$ differ. The investigation is done by comparing the exact and the normal approximate distribution of $R_i$. The chi-square approximation is obtained from the normal approximations of $R_i$. Figure 1 is the exact distribution of $(R_1, R_2)$ when $k = 3$ and $n_1 = n_2 = n_3 = 5$. $R_3$ is determined from the constraint $R_1 + R_2 + R_3 = 120$. Figure 2 is the bivariate normal distribution used as an approximate distribution. Figures 3 and 4 are contour plots of Figs. 1 and 2, respectively. The contours in Fig. 1 are hexagons because of the constraint $R_1 + R_2 + R_3 = 120$, while the contours in Fig. 2 are ellipsoids.

For $k = 3, 4$ and a given sample size $(n_1, \ldots, n_k)$, we also examine the maximum and the minimum errors of the approximation. Let

$$EXA = \frac{\text{number of cases whose value of } T \geq t}{\text{number of all possible cases}} \text{ and}$$

$$CHI = P(\chi^2_{k-1} \geq t)$$

where $\chi^2_{k-1}$ follows a chi-square distribution with $k - 1$ degrees of freedom. The error is defined by

$$ERROR = EXA - CHI.$$

The values of $EXA$ and $CHI$ are plotted against the values of $t$ in Fig. 3 when $n_1 = n_2 = n_3 = n_4 = 3$. The solid and broken lines represent $EXA$ and $CHI$, respectively. Figure 4 shows how the error changes as the value of $t$ moves when $n_1 = n_2 = n_3 = n_4 = 3$. When the value of $EXA$ is equal to 0.05, the value of $CHI$ is around 0.07, which means that the chi-square approximation is conservative in this particular case. The maximum and minimum errors in the range $EXA < 0.1$ are computed

MARCEL DEKKER, INC. • 270 MADISON AVENUE • NEW YORK, NY 10016







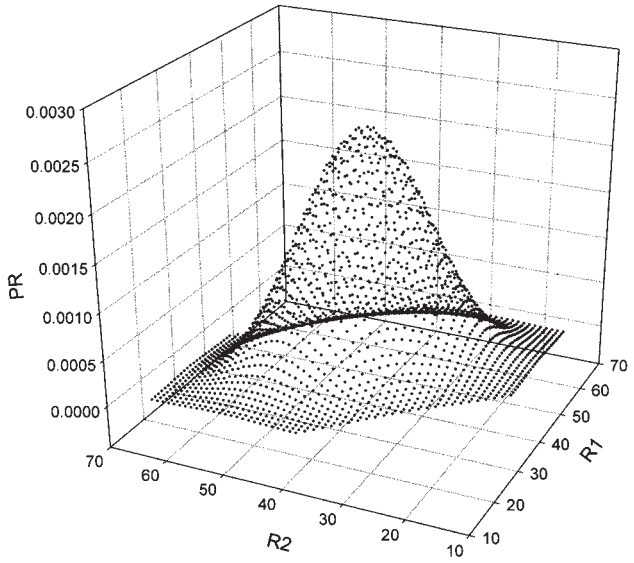


***Figure 1.*** The exact distribution of $(R_1, R_2)$ when $k = 3$ and $n_1 = n_2 = n_3 = 5$.

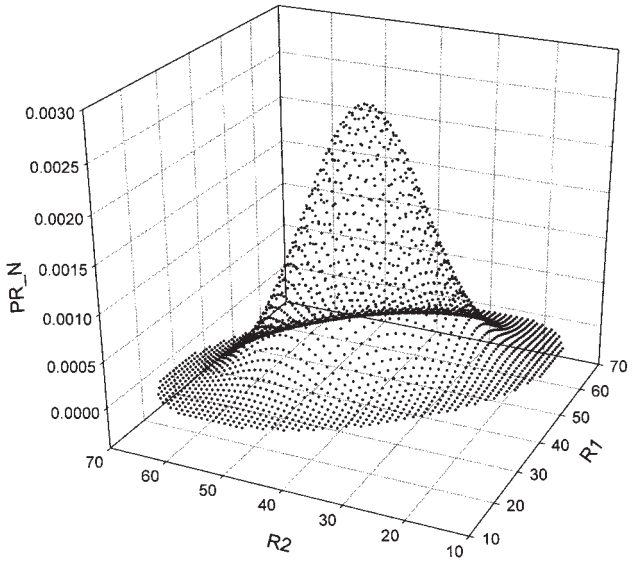


***Figure 2.*** The bivariate normal distribution of $(R_1, R_2)$ when $k = 3$ and $n_1 = n_2 = n_3 = 5$.

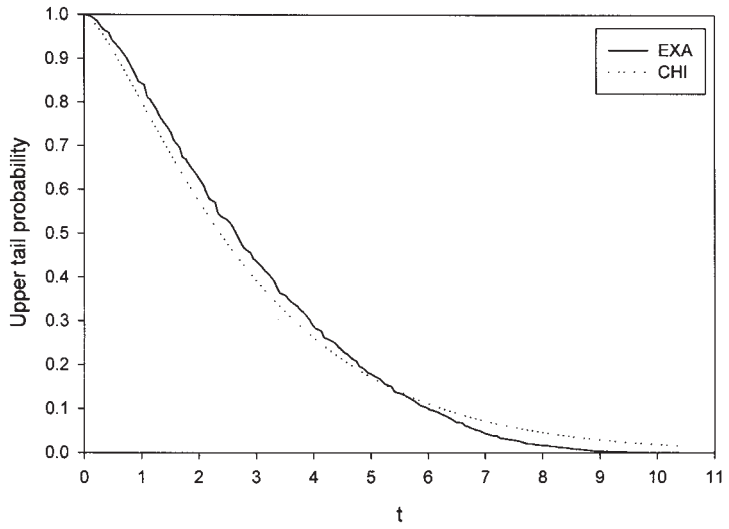


***Figure 3.*** The upper tail probabilities obtained from the exact method (solid line) and chi-square approximation (broken line).

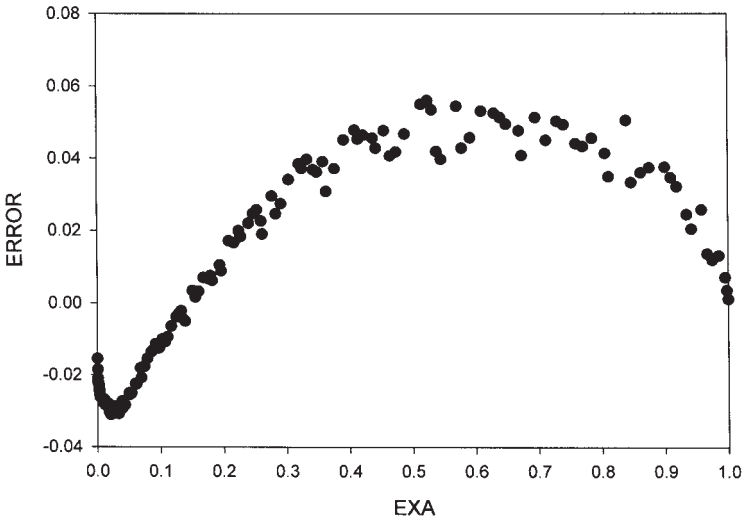


***Figure 4.*** The errors of chi-square approximation.

***Table 2.*** The maximum and the minimum errors of the chi-square approximation.

| $(n_1, n_2, n_3)$ | Min. error | Max. error | $(n_1, n_2, n_3, n_4)$ | Min. error | Max. error |
|---|---|---|---|---|---|
| (3, 3, 3) | −0.0295951 | 0.0063204 | (3, 3, 3, 3) | −0.0311761 | −0.0113788 |
| (4, 4, 4) | −0.0143748 | 0.0013821 | (4, 4, 4, 4) | −0.0170958 | −0.0027489 |
| (5, 5, 5) | −0.0096366 | −0.0011896 | (5, 5, 5, 5) | −0.0121449 | −0.0004707 |
| (6, 6, 6) | −0.0078869 | 0.0006242 | (6, 5, 5, 5) | −0.0116046 | −0.0003055 |
| (7, 7, 7) | −0.0067288 | −0.0001351 | (6, 6, 5, 5) | −0.0109391 | −0.0001962 |
| (8, 8, 8) | −0.0054769 | −0.0000357 | (6, 6, 6, 5) | −0.0102774 | −0.0001249 |
| (9, 9, 9) | −0.0050573 | −0.0000009 | (6, 6, 6, 6) | −0.0096640 | −0.0000790 |

and displayed in Table 2. An alternative way of computing the errors is to evaluate $ERROR$ at some values of $T$ where the exact cumulative probabilities are 0.1, 0.05, 0.01, 0.005, and 0.001. However, since the exact distribution is discrete, it may not be possible to examine $ERROR$ at a value of $T$ where the exact cumulative probability is exactly equal to a given value of significance level. Therefore, we compute the error for all the cases in the range $EXA < 0.1$, and then take the maximum and the minimum. In this computation 0.1 is chosen because, we think, it covers a useful tail area of the exact distribution in most practical situations. From Table 2 the errors decrease less than 1% with $n_i \geq 5$ and $k = 3$. When $k = 4$, $n_i \geq 6$ is needed to achieve the same error.

## 4. CONCLUDING REMARKS

In this article we propose a new algorithm to compute the exact null distribution of the Kruskal–Wallis test by improving the algorithm of Iman et al. (1975). The improvements are obtained from the symmetry theorem and modifying the way of using the recursive formula. Using the new algorithm and the increased computing power, we tabulate some exact probability levels of the Kruskal–Wallis test which are not covered in Iman et al. (1975). Those levels will be helpful when an asymptotic method is dubious. Especially, when a difference between an asymptotic $p$-value and a nominal level is less than a corresponding maximum error in Table 2, the exact inference is strongly recommended. The tables and the program to compute those levels is available from authors upon request.

Currently, commercial software StatXact version 4.01 can compute either exact $p$-values or exact null distributions of the Kruskal–Wallis

test. Its maximum range is about $n_1 = n_2 = n_3 \leq 4$ when $k = 3$ and $n_1 = n_2 = n_3 = n_4 \leq 3$ when $k = 4$. The ranges of tables in Appendix is beyond those of StatXact.

## REFERENCES

Chow, B., Dickinson, P., Champgne, Q. (1974). An approximation of the critical values of the Kruskal–Wallis H-test using a Monte Carlo sampling technique. *Journal of Quality Technology* 6(2):95–97.

Iman, R. L., Davenport, J. M. (1976). New approximations to the exact distribution of the Kruskal–Wallis test statistic. *Communications in Statistics: Theory and Methods* 5(14):1335–1348.

Iman, R. L., Quade, D., Alexander, D. (1975). Exact probability levels for the Kruskal–Wallis test. In: Harter, H. L., Owen, D.B., eds. *Selected Tables in Mathematical Statistis*. Providence: American Mathematical Society.

Kruskall, W. H., Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association* 47:583–621.

Lehmann, E. L. (1975). *Nonparametrics*. San Francisco: Holden-Day Inc.,

Mehta, C. R., Patel, N. R. (1983). A network algorithm for performing Fisher's exact test in $r \times c$ contingency tables. *Journal of the American Statistical Association* 78:427–434.

Wallace, D. L. (1959). Simplified beta-approximations to the Kruskall–Wallis H-test. *Journal of the American Statistical Association* 54:225–230.