

5-1-2012

Empirical Sampling from Permutation Space with Unique Patterns

Justice I. Odiase
University of Benin, Nigeria

Follow this and additional works at: <http://digitalcommons.wayne.edu/jmasm>

 Part of the [Applied Statistics Commons](#), [Social and Behavioral Sciences Commons](#), and the [Statistical Theory Commons](#)

Recommended Citation

Odiase, Justice I. (2012) "Empirical Sampling from Permutation Space with Unique Patterns," *Journal of Modern Applied Statistical Methods*: Vol. 11 : Iss. 1 , Article 10.

DOI: 10.22237/jmasm/1335845340

Available at: <http://digitalcommons.wayne.edu/jmasm/vol11/iss1/10>

This Regular Article is brought to you for free and open access by the Open Access Journals at DigitalCommons@WayneState. It has been accepted for inclusion in Journal of Modern Applied Statistical Methods by an authorized editor of DigitalCommons@WayneState.

Empirical Sampling from Permutation Space with Unique Patterns

Justice I. Odiase
University of Benin,
Benin, Nigeria

The exact distribution of a test statistic ultimately guarantees that the probability of a Type I error is exactly α . Several methods for estimating the exact distribution of a test statistic have evolved over the years with inherent computational problems and varying degrees of accuracy. The unique pattern of permutations resulting from using experimental data to sample within the permutation space without the risk of repeating permutations is identified. The method presented circumvents the theoretical requirements of asymptotic procedures and the computational difficulties associated with an exhaustive enumeration of permutations. Results show that time and space complexities are drastically reduced without compromising accuracy even when enumeration is not exhaustive provided error tolerance is achieved. The exact distribution of the Siegel-Tukey test statistic is examined as an illustration.

Key words: Algorithm, exact test, permutation test, bootstrap, resampling.

Introduction

The first edition of Fisher (1935) contains descriptions of two tests of significance that depend on permutation: Fisher's exact test for analyzing categorical data, and the permutation test for the difference between means. Many studies have been designed to confirm the asymptotic equivalence of permutation and classical tests (Ludbrook, 1994). Fisher wrote that "the statistician does not carry out this very simple and very tedious process, but his conclusions have no justification beyond the fact that they agree with those which could have been arrived at by this elementary method" (1936, p 59). Ernst (2004) noted that with fast computers there is little reason for a statistician not to carry out this very tedious process.

The main problem with permutation tests is that their null distributions are generally very difficult to express in closed form and to calculate exactly. This is because they depend

on a specific data set; thus they vary as the data varies in the sample space, however, for several test statistics involving ranks, the null distributions only need to be computed once. For large sample sizes, direct calculations are practically impossible due to the very large cardinality of associated permutation sample spaces. For example, a data set consisting of four treatments with five observations per treatment, $n_i = 5, i = 1(1)4$, demands as many as

$$\frac{\left(\sum_{i=1}^4 n_i\right)!}{\prod_{i=1}^4 (n_i!)} = \frac{20!}{5!5!5!5!} = 11,732,745,024$$

configurations for an exhaustive enumeration of all permutations. Pesarin (2001) stated that, unless sample sizes are very large, the approximation of such distributions by means of asymptotic arguments is not always appropriate. No general agreement exists regarding how large a sample should be before applying asymptotic approximation (Fahoome, 2000). Pesarin (2001) observed that the algorithms for exact calculations are generally based on direct calculus of upper tail probabilities; a strategy which may become highly impractical, if not

Justice Odiase is a Senior Lecturer in the Department of Mathematics, University of Benin, Benin City, Nigeria. Email him at: justice.odias@uniben.edu.

impossible, in multivariate problems because there is no general computing routine useful to identify the critical regions. This was also observed by Hall and Weierserman (1997). In the early years of research into exact statistical inference, Scheffé (1943) clearly identified the fact that the permutation approach is the only way of constructing the exact distribution of a test statistic.

To avoid the computational difficulty in exact permutation tests, the conditional Monte Carlo (CMC) method was adopted by Pesarin (2001). In CMC replicate resampling is conducted without-replacement on the given data set, which is considered as playing the role of a finite population, provided that sample sizes are finite. According to Opdyke (2003), all existing permutation procedures developed to date can perform conventional Monte Carlo sampling without replacement within a sample but none can avoid the possibility of drawing the same sample more than once. The consideration given by Odiase and Ogbonmwan (2007) is an exception but involves a complete enumeration of all the distinct permutations, which becomes impracticable when the sample size is large.

In this study, the unique pattern of each permutation resulting from experimental data is identified and exploited in sampling from the permutation space without the attendant risk of repeating permutations. The method presented circumvents the elaborate theoretical requirements of asymptotic procedures and the logical and computational difficulties associated with an exhaustive enumeration of permutations.

Exhaustive Permutation Procedures

The process of obtaining permutations begins by choosing the test statistic T and the acceptable significance level α . Let $\pi_1, \pi_2, \dots, \pi_N$ be a set of all distinct permutations or ranks of the observations in the experiment. Compute the test statistic T_i for permutation π_i , that is, $T_i = T(\pi_i)$. Construct an empirical cumulative distribution for T as:

$$p_0 = p(T \leq T_i) = \frac{1}{N} \sum_{i=1}^N \psi(t_0 - T_i),$$

where

$$\psi(\cdot) = \begin{cases} 1, & \text{if } t_0 \geq T_i \\ 0, & \text{if } t_0 < T_i \end{cases}.$$

Under the empirical distribution, if $p_0 \leq \alpha$, reject the null hypothesis.

Paired Permutation

Given two paired samples $X = (x_1, x_2, \dots, x_n)$ and $Y = (y_1, y_2, \dots, y_n)$, suppose a sample of n units from the population distribution F_X is paired with a sample of n units from the population distribution F_Y and are simultaneously tested in an experiment with T as the test statistic. For k distinct values of the test statistic T , the probability distribution of the test statistic $T = (T_1, T_2, \dots, T_k)$ under the null hypothesis $H_0 : F_X = F_Y$ is given by

$$P(T_j = t_0 | H_0) = \frac{f_j}{\sum_{i=1}^k} (2^{-n}) = f_j (2^{-n}),$$

where f_j is the number of occurrences of T_j . For specified value of n and the level of significance α , the critical value c corresponds to a level closest to α . Ordering all the distinct occurrences of T in ascending order of magnitude, and if g is the position of the observed value of T , then the following significance level for the left tail of the distribution of the test statistic is

$$\alpha = P(T_g \leq c | H_0) = \sum_{j=1}^g \sum_{i=1}^{f_j} (2^{-n}) = (2^{-n}) \sum_{j=1}^g f_j$$

and, for the right tail,

$$\alpha = P(T_g \geq c | H_0) = (2^{-n}) \sum_{j=g}^k f_j.$$

For a two-tailed test, the left and right tails are summed. If the distribution of the test statistic is symmetric, then

$$\sum_{j=1}^g f_j = \sum_{j=k-g+1}^k f_j.$$

The distribution of the test statistic is obtained by tabulating the distinct values of the statistic against their probabilities of occurrence in the complete enumeration (see Odiase & Ogbonmwan, 2007a for a detailed description of the implementation of the paired permutation algorithm). Given a balanced two-sample layout as

$$\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_n & y_n \end{pmatrix},$$

where $x_i \in X$ and $y_j \in Y$. If $n = 4$, then, for a two-sample problem, the number of permutations is $\sum_{i=0}^n \binom{n}{i} = 2^n$.

The test statistic is computed for each permutation in the complete enumeration of the distinct permutations. The distribution of the test statistic is obtained by tabulating the distinct values of the statistic against their probabilities of occurrence in the complete enumeration, where all the permutations are equally likely.

Considering consecutive number of pairs for a given experiment, the growth rate of the permutations from $n-1$ pairs to n pairs in a two-sample paired permutation experiment is

$$\frac{2^n}{2^{n-1}} = 2, \text{ meaning that it doubles each time a single pair of observations is added.}$$

Independent Samples Experiment

Given a multi-sample experiment with $X_i = (X_{i1}, X_{i2}, \dots, X_{in_i})^T, i = 1(1)p$ and $X_N = (X_1, X_2, \dots, X_p)$. Under the null hypothesis, X_N

is composed of $N = \sum_{i=1}^p n_i$ independent and

identically distributed random variables. An exhaustive permutation of the observations

yields $\frac{N!}{\prod_{i=1}^p [(n_i)!]}$ permutations of the N variates

of p subsets of size $n_i, i = 1(1)p$, which are equally likely and each has the conditional

probability $\left(\frac{N!}{\prod_{i=1}^p [(n_i)!]} \right)^{-1}$. Odiase and

Ogbonmwan (2005a) showed that the number of permutations for a two-sample experiment is $\sum_{i=0}^n \binom{n_1}{i} \binom{n_2}{i}, n = \min(n_1, n_2)$. After obtaining

the permutations of a two sample experiment, find the number of ways to permute any n_3 elements of the combined $(n_1 + n_2 + n_3)$ variates of the three treatments. This yields:

$$\binom{n_1 + n_2 + n_3}{n_3} \sum_{i=0}^n \binom{n_1}{i} \binom{n_2}{i} = \binom{\sum_{r=1}^3 n_r}{n_3} \sum_{i=0}^n \binom{n_1}{i} \binom{n_2}{i}.$$

By following the same procedure as for the case of three treatments, a complete enumeration of the distinct permutations for a four-treatment experiment yields:

$$\begin{aligned} & \binom{\sum_{r=1}^4 n_r}{n_4} \binom{\sum_{r=1}^3 n_r}{n_3} \sum_{i=0}^n \binom{n_1}{i} \binom{n_2}{i} = \\ & \prod_{j=3}^4 \binom{\sum_{r=1}^j n_r}{n_j} \sum_{i=0}^n \binom{n_1}{i} \binom{n_2}{i}. \end{aligned}$$

Continuing in this manner, for $p \geq 3$ treatments, the distinct permutations are enumerated through the expression

$$\prod_{j=3}^p \binom{\sum_{r=1}^j n_r}{n_j} \sum_{i=0}^n \binom{n_1}{i} \binom{n_2}{i},$$

but

$$\sum_{i=0}^n \binom{n_1}{i} \binom{n_2}{i} = \binom{\sum_{r=1}^2 n_r}{n_2}.$$

Therefore, in a p -sample problem, the number of

distinct permutations is $\prod_{j=1}^p \binom{\sum_{r=1}^j n_r}{n_j}$ (Odiase &

Ogbonmwan, 2005b). Observe that, for the balanced case, the number of distinct permutations is

$$\prod_{j=3}^p \binom{jn}{n} \sum_{i=0}^n \binom{n}{i}^2 = \prod_{j=1}^p \binom{jn}{n}.$$

Again, considering consecutive number of treatments for a given experiment, the growth rate of the permutations from $p-1$ treatments to p treatments is

$$\frac{\prod_{j=1}^p \binom{jn}{n}}{\prod_{j=1}^{p-1} \binom{jn}{n}} = \binom{np}{n}.$$

Repeated Measures (Block) Permutation

Repeated measures ANOVA tests the equality of means and is used when all members of a random sample are measured under varying conditions. In the repeated measures design, each trial represents the measurement of the same characteristic under a different condition. Given the layout of a multi-sample ($n \times k$) experiment as

$$\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix},$$

where x_{ij} is an observation in the j^{th} treatment and the i^{th} block and the total number of observations in the experiment is nk . Rank the observations for each row from 1 (smallest x_{ij} on row i) to k (largest x_{ij} on row i). Let the layout of the ranks (r_{ij}) of the observations x_{ij} be

$$L_R = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1k} \\ r_{21} & r_{22} & \cdots & r_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nk} \end{pmatrix}.$$

The data are arranged in k columns (treatments) and n rows (blocks), where each block contains k repeated observations. Obviously, there are $k!$ possible arrangements or permutations of each block and due to multiplication of choices, the entire layout of the $n \times k$ experiment requires $(k!)^n$ permutations of the observations to yield the exact distribution of a test statistic, with the permutations equally likely and each having the conditional probability $(k!)^{-n}$.

The first step in developing permutation algorithm is to formulate an initial configuration of the ranks of the variates of an experiment by taking the trivial configuration

	<i>Tr.1</i>	<i>Tr.2</i>	<i>Tr.3</i>	...	<i>Tr.k</i>
<i>Block1</i>	1	2	3	...	k
<i>Block2</i>	1	2	3	...	k
<i>Block3</i>	1	2	3	...	k
\vdots	\vdots	\vdots	\vdots	...	\vdots
<i>Block n</i>	1	2	3	...	k

because any configuration of the ranks can engender all the distinct permutations. The test statistic is computed for each permutation in the complete enumeration of all the distinct permutations. The distribution of the test statistic is obtained by tabulating the distinct values of the test statistic against their probabilities of occurrence in the complete enumeration.

Considering two consecutive numbers of blocks for a given experiment, the growth rate of the permutations from $n-1$ blocks to n blocks

is $\frac{(k!)^n}{(k!)^{n-1}} = k!$ and the growth rate of the

permutations from $k-1$ treatments to k

treatments is $\frac{(k!)^n}{[(k-1)!]^n} = k^n$ and clearly, $k!$

grows faster than k^n for a fixed n , and for a fixed

k , $k!$ is constant while k^n explodes as n increases. Therefore, the growth rate of the permutations is higher for a unit increase in blocks than a unit increase in treatments for a fixed number of treatments and the reverse is the case when it is the number of blocks that is fixed.

Sampling Permutations with Unique Patterns

Given the layout of a two-sample experiment as

$$\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_{1n_1} & y_{n_2} \end{pmatrix},$$

where x_i and y_i are the i^{th} observations of the independent random samples X and Y respectively, arrange the combined samples in ascending order of magnitude and rank all the $m = n_1 + n_2$ observations from 1 (smallest) to m (largest). Let the layout of the ranks (r_{ij}) of the observations in a two-sample layout be

$$L_R = \begin{pmatrix} r_{11} & r_{21} \\ r_{12} & r_{22} \\ \vdots & \vdots \\ r_{1n_1} & r_{2n_2} \end{pmatrix}.$$

Under the null hypothesis, L_R is composed of m independent and identically distributed random variables and hence conditioned on the observed data set. An exhaustive permutation of the ranks yields $N = \frac{m!}{n_1!n_2!}$ permutations of the m ranks

of the variates of two subsets of sizes n_1 and n_2 , which are equally likely, each having the conditional probability N^{-1} .

The unique pattern of each permutation resulting from experimental data is identified by adopting the first sample, for example, in a two-sample problem. This is carried out by concatenating the ranks or indices of observations in the experiment in a particular manner that makes the pattern unique for every distinct permutation. The unique patterns

obtained are therefore exploited in sampling from the permutation space without the risk of repeating permutations already sampled. The benefit of this approach is that – even when enumeration is not exhaustive – the distribution of a test statistic can be obtained within a reasonable level of accuracy with reduced time and space complexities. This sampling approach therefore circumvents the elaborate theoretical requirements of asymptotic procedures and the logical and computational difficulties associated with an exhaustive enumeration of permutations.

Methodology

Let the initial configuration of the ranks of the variate in a two-sample experiment be L_R . The entire permutation space can be spanned by any of the permutations (configurations) of the observations or ranks of observations. In a two-sample problem, only one of the samples is required to define each permutation because it is obvious that the remaining variates are in the second sample.

In a two-sample experiment

$$\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_{1n_1} & y_{n_2} \end{pmatrix},$$

where x_i and y_i are the i^{th} observations of the independent random samples X and Y respectively, arrange the combined samples in ascending order of magnitude and rank all the $m = n_1 + n_2$ observations from 1 (smallest) to m (largest) to arrive at L_R . For the original permutation and subsequent permutations, sort X such that $x_1 \leq x_2 \leq \dots \leq x_{n_1}$.

The variates are identified by their indices (1, 2, ..., m) or actual ranks, which are employed in obtaining the unique patterns. Attach 0 in front of the first nine indices or ranks (01, 02, ..., 09) to make each number two digits, leaving 10, 11, ..., 99 as they are and treat all the numbers as strings so that it will be possible to manipulate the numbers. Concatenate the indices or ranks of X and store as a single constant value. This now becomes the pattern of the

EMPIRICAL SAMPLING FROM PERMUTATION SPACE WITH UNIQUE PATTERNS

given layout of the observations. (Concatenation is a standard operation in computer programming languages. It is the operation of joining two character strings end to end. In programming languages, string concatenation is a binary operation usually accomplished by putting a concatenation operator between two strings or operands.)

<i>X</i>	<i>Y</i>
01	06
02	07
03	08
04	09
05	10

After a unique pattern is obtained, a resampling without replacement is carried out to obtain a random sample of n_1 variates from the original combined sample of m variates. This is achieved by deleting points already selected at random. Again, sort the resampled n_1 variates and concatenate their indices or ranks to obtain a pattern. Compare this pattern with previously obtained patterns and store it only if it is unique, otherwise, resample without replacement again until a unique pattern is obtained. The chosen test statistic is computed for each unique permutation and the probability distribution of the test statistic is constructed. Finally, compute the cumulative probability distribution of the test statistic, T , under the null hypothesis and obtain the p -values such that the probability of making a Type I error is exactly α .

As an illustration, consider an $n \times k$ experiment with $n = 2$ treatments (X, Y) and $k = 5$ variates in each treatment could have the trivial configuration or permutation of ranks represented as

<i>X</i>	<i>Y</i>
1	6
2	7
3	8
4	9
5	10

Using the methodology presented, the permutation is written as

a total of $\frac{10!}{5! \times 5!} = 252$ permutations are

required for an exhaustive enumeration of all the distinct or unique permutations. The process starts thus: The entire permutation space can be spanned by the trivial permutation (configurations) of the observations or ranks of observations, any other permutation from the permutation space can also be adopted to span all the unique permutations. In a two-sample problem, only one of the samples (X) is required to define each permutation because it is clear that the remaining variates are in the second sample.

The first permutation pattern is 0102030405. Assuming resampling from the permutation space yields $X = \{8, 3, 5, 2, 6\}$, then the second permutation pattern is obtained by first sorting to have $\{2, 3, 5, 6, 8\}$, leading to the pattern 0203050608. Resampling again, given $\{2, 10, 5, 9, 8\}$, the third permutation pattern is 0205080910. The resampling process continues until either all the patterns are enumerated for small samples or the error tolerance is achieved for large samples. See Table 1 for an exhaustive

enumeration of the $\frac{10!}{5! \times 5!} = 252$ permutation

patterns for this illustration.

For very small samples, an exhaustive enumeration of all the unique permutations is achieved with the sampling method described. When sample size is large, enumeration of permutations does not need to be exhaustive. Instead, a subset of the permutation space (for example, 2,000) is obtained and the probability distribution of the test statistic is constructed. Take a second sample of same size and fuse it into the earlier distribution to obtain an updated probability distribution and compare with the earlier distribution. With a given level of error

JUSTICE I. ODIASE

Table 1: Exhaustive Enumeration of Unique Permutation Patterns for $n = 2, k = 5$

0102030405	0304050710	0103050709	0405070809	0304070810	0105060708
0203050608	0206070809	0105060910	0102030408	0203060708	0405060708
0106070809	0102030910	0102040910	0205080910	0102060810	0102030506
0104050809	0102050609	0102060709	0204050708	0304060708	0506080910
0104080910	0102030709	0304050709	0102040608	0206070810	0305060710
0104060708	0205060809	0104060809	0102030510	0204050610	0102070809
0203040708	0104060710	0306070809	0307080910	0102040609	0102040710
0507080910	0103040508	0103040810	0305080910	0103040507	0102040709
0203060810	0306070810	0203040510	0205070810	0105080910	0106080910
0102060710	0103050610	0204070910	0103050810	0203060809	0204050809
0102030409	0102030407	0204060708	0506070809	0102050910	0205070910
0304050910	0405060910	0102060809	0103060910	0104060709	0103040506
0203040508	0205060910	0205060810	0103060709	0204060910	0203040610
0304050708	0104070809	0103050710	0105060710	0102030509	0304050609
0102030607	0406080910	0304050809	0103040708	0203050910	0204080910
0203040810	0104070910	0102040509	0104050910	0203050609	0304070809
0103060710	0103040608	0406070809	0103050609	0107080910	0104070810
0106070810	0204050910	0305060708	0203040609	0102030609	0106070910
0405060710	0304060910	0304050610	0103040609	0102050809	0102040506
0203040506	0102030406	0104050607	0204070810	0103070810	0203040509
0102040610	0304060809	0102040810	0103050708	0405060709	0203060910
0203040607	0104050709	0204050709	0102080910	0206080910	0102030810
0205060710	0103040809	0103060708	0203060710	0405060809	0204060710
0104050609	0102060910	0104050710	0102070910	0103060810	0304080910
0103060809	0203050607	0102040507	0102030708	0203040709	0204050810
0105060709	0207080910	0304070910	0204050710	0203070809	0103040710
0102040508	0305060809	0103050809	0102030410	0103040607	0205060709
0105070810	0103040510	0405070910	0405060810	0607080910	0204060809
0102030710	0102070810	0203080910	0204060810	0406070910	0405070810
0203050809	0203060709	0204060709	0103080910	0105070809	0204050607
0102030507	0103070809	0102050607	0103050910	0104050810	0105070910
0203040608	0204050609	0104060810	0203040809	0103070910	0102040607
0103040610	0103050607	0206070910	0102040809	0506070910	0304060709
0305060810	0306080910	0203040710	0105060809	0102060708	0203050709
0102050710	0203070910	0304060710	0104060910	0304050608	0102030608
0203050610	0102030508	0102030809	0406070810	0103050608	0204070809
0102030610	0203040507	0405080910	0102040708	0105060810	0103040509
0305060910	0205070809	0103040709	0304050607	0203040910	0203050708
0306070910	0305070910	0102040510	0304060810	0305070809	0203070810
0204050608	0205060708	0103040910	0304050810	0407080910	0102050608
0305060709	0104050608	0104050610	0203050810	0506070810	0102050610
0305070810	0104050708	0203050710	0102050708	0102050709	0102050810

tolerance, if the error tolerance is exceeded, another sample is taken and fused into the last update of the probability distribution to obtain another update and again compared with the previous update of the probability distribution. This process is continued until the error tolerance is achieved. Compare for every occurrence of the test statistic in the last two updates using the error tolerance as a guide and proceed to update the probability distribution if the error tolerance is not met. Compute the cumulative probability distribution of the Test Statistic, T , under the null hypothesis and obtain the p-values such that the probability of making a type I error is exactly α .

Unique Permutation Pattern Test Procedure

Let $\pi_1, \pi_2, \dots, \pi_N$ be a set of all distinct permutations of the ranks of the data set in the experiment. The unique permutation pattern test procedure is as follows:

1. Read the original layout of observations.
2. Rank the combined observations of the experiment in (1) to obtain $L_R = \pi_1$ and compute the observed value of T statistic to obtain $T_1 = t_0$.
3. Store pattern of (1) or (2).
4. Obtain a distinct permutation π_i , of the ranks (L_R) from (3) by sampling without replacement.
5. Obtain pattern of (4) and compare with previous patterns. If different, store pattern, if already exists, go to (4).
6. Compute the T statistic $T_i = T(\pi_i)$, for permutation π_i in (5), where $i > 1$; update probability distribution.
7. Perform (4) to (6) for $i = 2, 3, \dots, k \leq N$. If sample size (N) is large,
 - a) Assume a level of error tolerance (0.00001) and take a subset of size $k = 1,000$ of the permutation space to obtain

the probability distribution of the test statistic.

- b) Take another subset of size k from the permutation space and fuse into the earlier probability distribution to obtain an updated distribution. Compare this distribution with the earlier distribution of the test statistic for every value of the test statistic.
 - c) If the error tolerance is exceeded for any value of the test statistic, go back to (b); continue this process until the error tolerance is achieved.
8. Construct the empirical cumulative probability distribution for the distinct values of T and extract critical values.

$$p_0 = p(T \leq T_i) = \frac{1}{k} \sum_{i=1}^k \psi(T_i - t_0),$$

where

$$\psi(T_i - t_0) = \begin{cases} 1, & \text{if } T_i \leq t_0 \\ 0, & \text{if } T_i > t_0 \end{cases}.$$

9. Under the empirical distribution, if $p_0 \leq \alpha$, reject the null hypothesis.

In a two-sample problem, only one of the samples is important in the generation of permutation patterns because it is unique for each permutation, that is,

$$\binom{m}{n_1} = \binom{m}{n_2} = \frac{m!}{n_1! n_2!}, \quad m = n_1 + n_2.$$

To provide exact critical values when ties occur, midranks are assigned as the ranks of tied observations, and the algorithm is implemented with r_{ij} as input, composed of actual ranks containing ties. Tabulated exact critical values of a test statistic are usually provided for experiments with distinct observations, because it will be practically difficult to consider all possible occurrences of ties and create tables of exact critical values for each occurrence of ties

for different sample sizes. This will result in several volumes of tables. In order to arrive at the critical values (see Table 2), the ranks of distinct observations (r_{ij}) were used as input in Algorithm for various sample sizes. See Appendix A for the unique permutation pattern algorithm. This algorithm identifies and compiles the unique permutation patterns of the layout of observations or rank of observations in a two-sample experiment. It is illustratively implemented to produce a table of critical values for the Siegel-Tukey test statistic

Siegel-Tukey Test

The Siegel-Tukey test is similar in procedure to the Wilcoxon Rank Sum test for difference in location. It is based on the logic that if two samples come from populations with the same median, the one with the greater variability will have more extreme scores. The hypotheses for a two-tailed test are:

H_0 : There is no difference in spread
between the two populations

versus

H_1 : There is some difference in spread
between the two populations.

The two samples are combined and ordered in ascending order of magnitude, keeping track of sample membership. For $m = n_1 + n_2$, the ranking proceeds as follows: $\alpha(1) = 1, \alpha(m) = 2, \alpha(m-1) = 3, \alpha(2) = 4, \alpha(3) = 5, \alpha(m-2) = 6, \alpha(m-3) = 7, \alpha(4) = 8$, the ranking continues to alternate from lowest to highest, ranking two scores at each end. It tests for differences in scale between two groups.

Critical Values for the Siegel-Tukey Test Statistic

Figures 1a-1b illustrate that the normal distribution will poorly approximate the exact distribution of the Siegel-Tukey (S-T) test statistic for very small sample sizes. As group sample size increases, the shape of the distribution of the S-T test begins to look more

like the normal distribution as shown in Figures 1c-1d. The critical values of the S-T test statistic shown in Table 2 were obtained from the enumeration of all distinct permutations of the ranks of the observations in an experiment ($m, n < 20$) combined with the idea of resampling while ensuring an error tolerance level ($m, n \geq 20$). These critical values ensure that the probability of a Type I error in decisions arising from the use of the S-T test is exactly α .

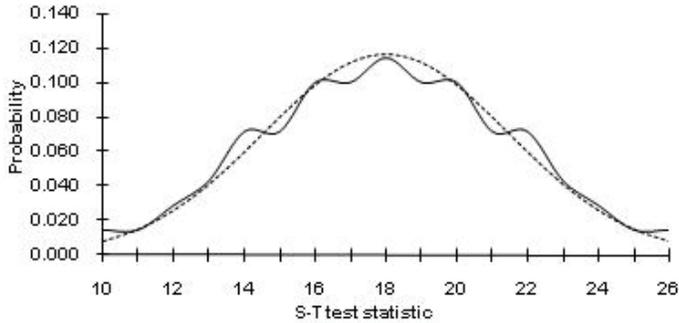
Results obtained from asymptotic procedures and resampling techniques are commonly adopted in several nonparametric tests as alternatives to tabulated exact critical values. Fahoome (2002) conducted a Monte Carlo study and recommended the asymptotic approximation of the S-T test when group sample sizes exceed 25, based on conservative estimates of $0.045 < \text{Type I error rate} < 0.055$ for $\alpha = 0.05$; other authors recommended higher or lower sample sizes.

Conclusion

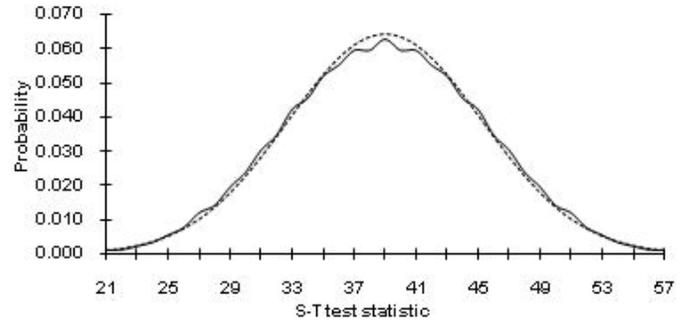
The critical values for a test statistic are determined by cutting off the most extreme $100\alpha\%$ of the theoretical frequency distribution of the test statistic, where α is the level of significance (Siegel & Castellan, 1988). Classical methods require that the theoretical distribution of the test statistic should agree with a mathematically definable frequency distribution, this often leads to a probability of Type I error greater than α , particularly when sample sizes are small. The cost of such an error might be too high to risk. Therefore, the exact permutation paradigm methodology presented in this study is of value because it guarantees that the probability of a Type I error is exactly α with the attendant advantage of no distributional assumptions apart from the exchangeability of the observations. When sample sizes are large and it becomes practically difficult or impossible to construct the probability distribution, permutation sampling becomes very useful because it quickly converges to the actual distribution; Scheffe (1943) opined that this is the only sure way of constructing the exact distribution of a test statistic.

EMPIRICAL SAMPLING FROM PERMUTATION SPACE WITH UNIQUE PATTERNS

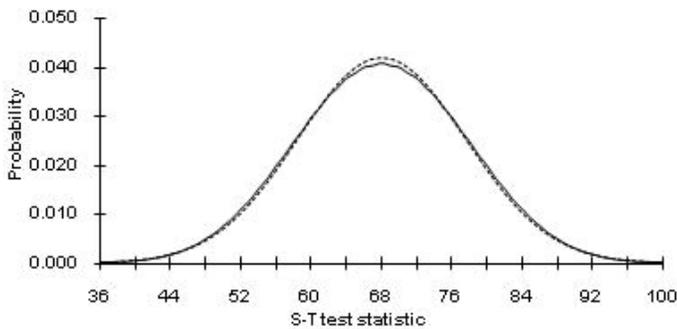
Figure 1: Exact Distribution of Siegel-Tukey Test Statistic (S-T) for Different Sample Sizes with their Large Sample Approximations (Z)



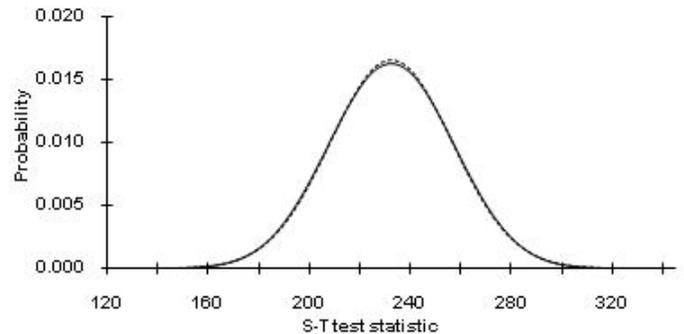
(a)
 $n_1 = 4, n_2 = 4$



(b)
 $n_1 = 6, n_2 = 6$

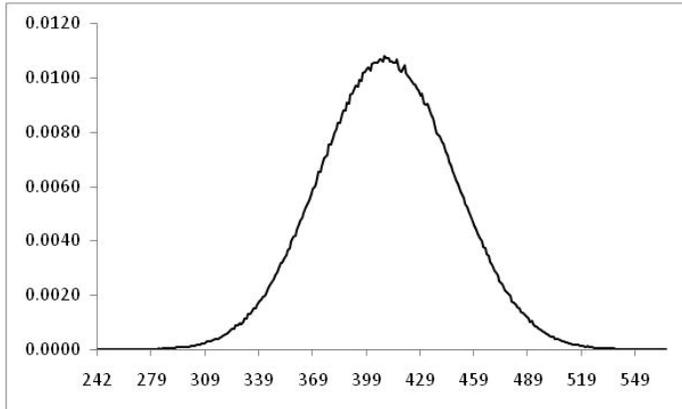


(c)
 $n_1 = 8, n_2 = 8$

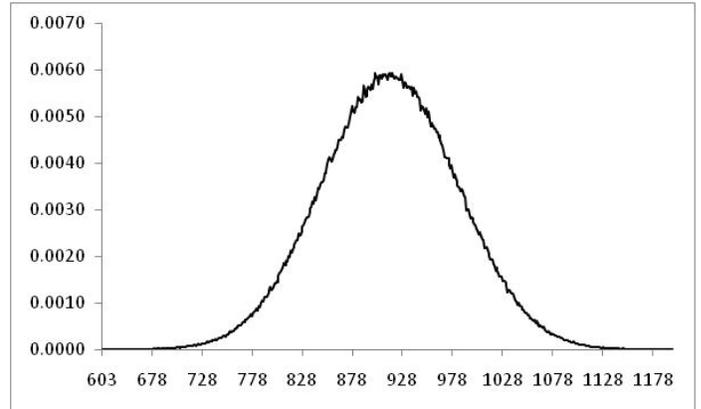


(d)
 $n_1 = 15, n_2 = 15$

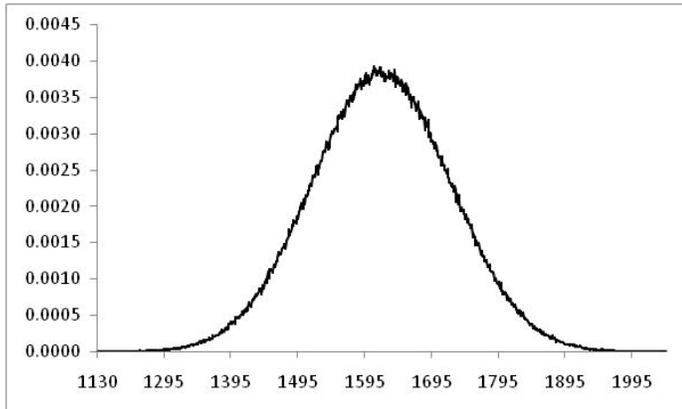
Figure 2: Distribution of Siegel-Tukey Test Statistic with Error Tolerance = 0.00001 for Different Sample Sizes



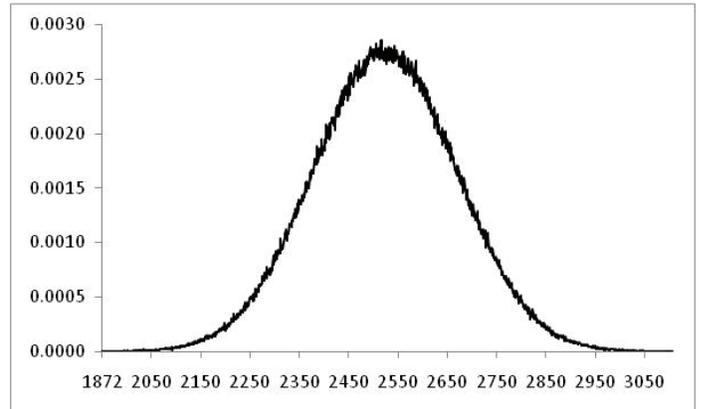
(a)
 $n_1 = 20, n_2 = 20$



(b)
 $n_1 = 30, n_2 = 30$



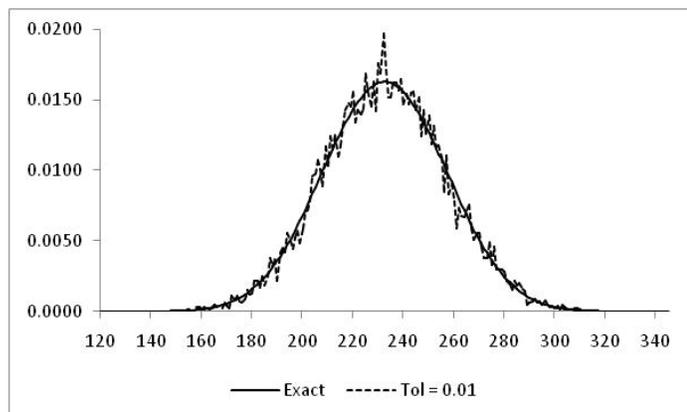
(c)
 $n_1 = 40, n_2 = 40$



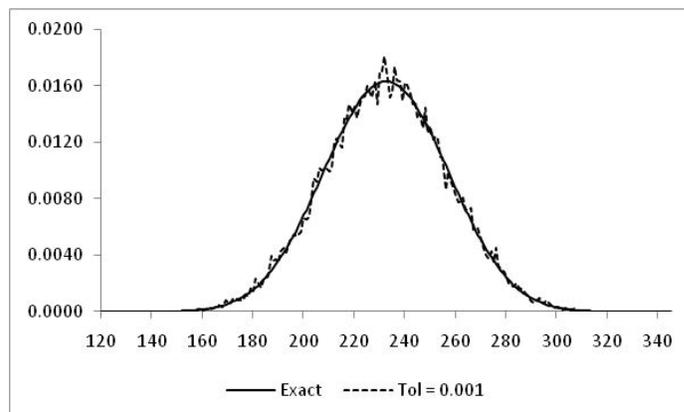
(d)
 $n_1 = 50, n_2 = 50$

EMPIRICAL SAMPLING FROM PERMUTATION SPACE WITH UNIQUE PATTERNS

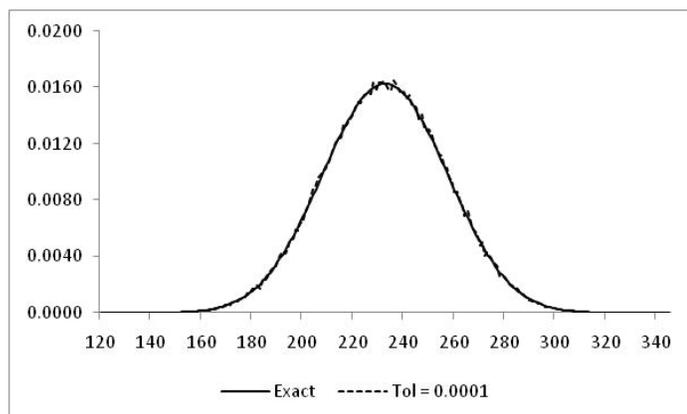
Figure 3: Distribution of Siegel-Tukey Test Statistic for Different Levels of Error Tolerance ($n_1 = 15, n_2 = 15$)



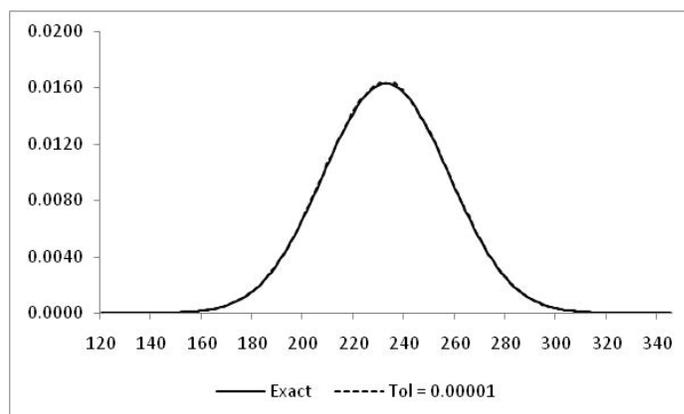
(a)
2 Iterations



(b)
5 Iterations

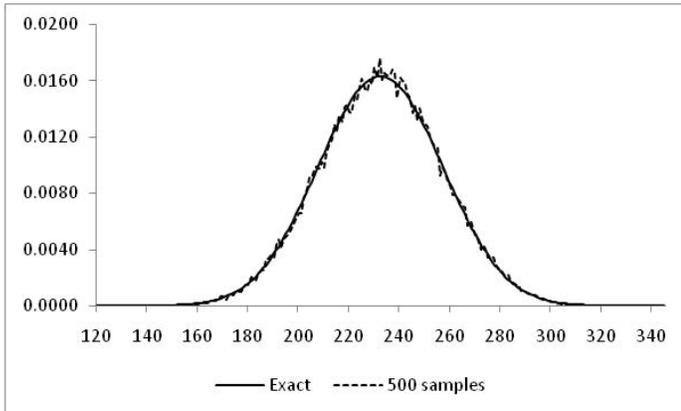


(c)
36 Iterations

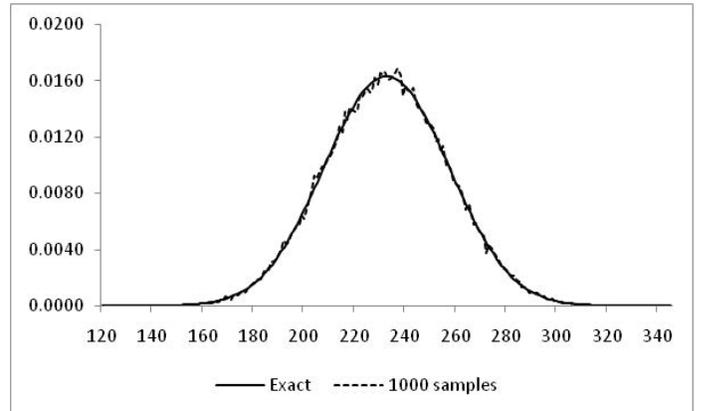


(d)
339 Iterations

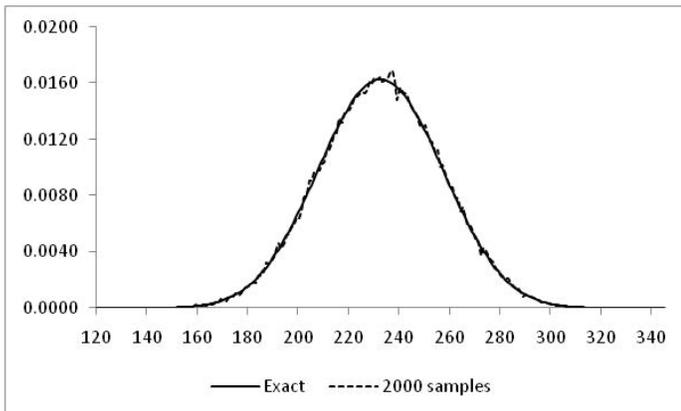
Figure 4: Distribution of Siegel-Tukey Test Statistic for Different Number Samples ($n_1 = 15, n_2 = 15$)



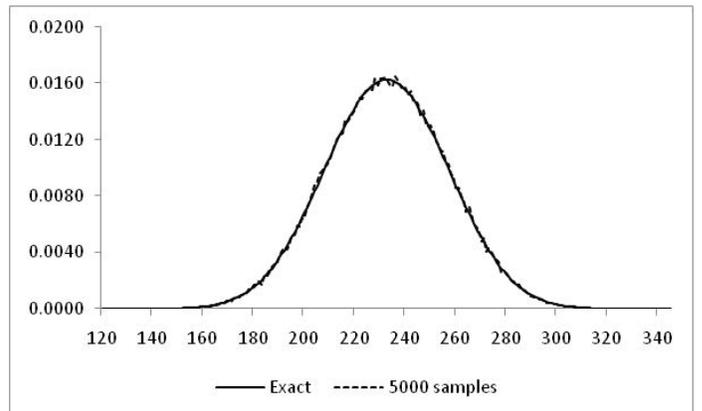
(a)
103 Iterations



(b)
82 Iterations



(c)
50 Iterations



(d)
36 Iterations

EMPIRICAL SAMPLING FROM PERMUTATION SPACE WITH UNIQUE PATTERNS

Table 2: Lower and Upper Critical Values for the Siegel-Tukey Test Statistic

Sample Size n_1 n_2	$ST_{0.9000}$	$ST_{0.9500}$	$ST_{0.9750}$	$ST_{0.9900}$	$ST_{0.9950}$	$ST_{0.9975}$	$ST_{0.9990}$
3 2	6 12						
3 3	7 14	6 15					
4 2	11 17						
4 3	12 20	11 21					
4 4	13 23	12 24	11 25				
5 2	17 23	16 24					
5 3	18 27	17 28	16 29				
5 4	20 30	18 32	17 33	16 34			
5 5	21 34	20 35	18 37	17 38	16 39	16 39	
6 2	23 31	22 32					
6 3	25 35	24 36	23 37				
6 4	27 39	25 41	24 42	23 43	22 44		
6 5	29 43	27 45	25 47	24 48	23 49	22 50	
6 6	31 47	29 49	27 51	25 53	24 54	23 55	
7 2	30 40	29 41					
7 3	33 44	31 46	30 47	29 48			
7 4	35 49	33 51	32 52	30 54	29 55		
7 5	37 54	35 56	34 57	32 59	30 61	29 62	
7 6	40 58	37 61	35 63	33 65	32 66	31 67	29 69
7 7	42 63	40 65	37 68	35 70	33 72	32 73	30 75
8 2	39 49	38 50	37 51				
8 3	42 54	40 56	39 57	37 59			
8 4	44 60	42 62	41 63	39 65	38 66	37 67	
8 5	47 65	45 67	43 69	41 71	39 73	38 74	37 75

JUSTICE I. ODIASE

Table 2 (continued): Exact Critical Values for Siegel-Tukey Test Statistic

Sample Size n_1 n_2	$ST_{0.9000}$	$ST_{0.9500}$	$ST_{0.9750}$	$ST_{0.9900}$	$ST_{0.9950}$	$ST_{0.9975}$	$ST_{0.9990}$
8 6	50 70	47 73	45 75	43 77	41 79	40 80	38 82
8 7	53 75	50 78	47 81	44 84	43 85	41 87	39 89
8 8	56 80	52 84	50 86	46 90	44 92	43 93	41 95
9 2	48 60	47 61	46 62				
9 3	51 66	49 68	48 69	47 70	46 71		
9 4	55 71	52 74	50 76	49 77	47 79	46 80	
9 5	58 77	55 80	53 82	51 84	49 86	48 87	47 88
9 6	61 83	58 86	56 88	53 91	51 93	50 94	48 96
9 7	64 89	61 92	58 95	55 98	53 100	51 102	49 104
9 8	68 94	64 98	61 101	57 105	55 107	53 109	51 111
9 9	71 100	67 104	63 108	60 111	57 114	55 116	53 118
10 10	88 122	83 127	79 131	75 135	72 138	69 141	66 144
11 11	107 146	101 152	97 156	92 161	88 165	85 168	82 171
12 12	128 172	121 179	116 184	110 190	106 194	103 197	99 201
13 13	150 201	143 208	137 214	131 220	126 225	122 229	118 233
14 14	175 231	167 239	161 245	153 253	148 258	144 262	138 268
15 15	201 264	193 272	185 280	177 288	172 293	167 298	161 304
20 20	362 458	349 471	338 482	325 495	316 504	308 512	298 522
30 30	828 1002	804 1027	783 1047	759 1071	743 1088	728 1102	710 1121
40 40	1487 1754	1449 1791	1417 1824	1380 1861	1354 1886	1331 1910	1303 1937
50 50	2339 2711	2286 2764	2241 2809	2189 2861	2153 2898	2122 2930	2082 2970

References

- Agresti, A. (1992). A survey of exact inference for contingency tables. *Statistical Science*, 7, 131-177.
- Bayarri, M. J., & Berger, J. O. (2004). The interplay of Bayesian and frequentist analysis. *Statistical Science*, 19, 58-80.
- Darwin, C. (1878). *The Effects of Cross and self-fertilization in the Vegetable Kingdom*, 2nd Ed. London: John Murray.
- Devore, J. L. (1982). *Probability and statistics for engineering and the sciences*. California: Brooks/Cole Publishing Company.
- Efron, B. (1982). *The jackknife, the bootstrap, and other resampling plans*. Philadelphia, PA: SIAM.
- Ernst, M. D. (2004). Permutation methods: A basis for exact inference. *Statistical Science*, 19, 676-685.
- Fahome, G. (2002). Twenty nonparametric statistics and their large sample approximations. *Journal of Modern Applied Statistical Methods*, 1, 248-268.
- Fisher, R. A. (1935). *The design of experiments*. Edinburgh: Oliver and Boyd.
- Fisher, R. A. (1936). The coefficient of racial likeness and the future of Craniometry. *Journal of the Royal Anthropological Institute*, 66, 57-63.
- Fisher, R. A. (1971). *The design of experiments* (8th Ed.). New York: Hafner Publishing.
- Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, 32(200), 675-701.
- Friedman, M. (1939). A correction: The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, 34(205), 109.
- Friedman, M. (1940). A comparison of alternative tests of significance for the problem of m rankings. *The Annals of Mathematical Statistics*, 11(1), 86-92.
- Good, P. (2000). *Permutation tests: a practical guide to resampling methods for testing hypotheses* (2nd Ed.). New York: Springer-Verlag.
- Hall, P. and I. Weissman (1997). On the estimation of extreme tail probabilities. *The Annals of Statistics*, 25, 1311-1326.
- Headrick, T. C. (2003). An algorithm for generating exact critical values for the Kruskal-Wallis one-way ANOVA. *Journal of Modern Applied Statistical Methods*, 2, 268-271.
- Ludbrook, J., & Dudley, H. (1998). Why permutation tests are superior to t and F tests in biomedical research. *The American Statistician*, 52, 127-132.
- Odiase, J. I., & Ogbonmwan, S. M. (2005a). An algorithm for generating unconditional exact permutation distribution for a two-sample experiment. *Journal of Modern Applied Statistical Methods*, 4, 319-332.
- Odiase, J. I., & Ogbonmwan, S. M. (2005b). Exact permutation critical values for the Kruskal-Wallis one-way ANOVA. *Journal of Modern Applied Statistical Methods*, 4(2), 609-620.
- Odiase, J. I., & Ogbonmwan, S. M. (2007a). Exact permutation of paired observations: The challenge of R. A. Fisher. *Journal of Mathematics and Statistics*, 3(3), 116-121.
- Odiase, J. I., & Ogbonmwan, S. M. (2007b). Exact permutation paradigm in multisample experiments. *Advances and Applications in Statistics*, 7(3), 417-434.
- Opdyke, J. D. (2003). Fast permutation tests that maximize power under conventional Monte Carlo sampling for pairwise and multiple comparisons. *Journal of Modern Applied Statistical Methods*, 2, 27-49.
- Scheffe, H. (1943) Statistical inference in the nonparametric case. *The Annals of Mathematical Statistics*, 14, 305-332.
- Siegel, S., & Castellan, N. J. (1989). *Nonparametric statistics for the behavioural sciences* (3rd Ed.). New York: McGraw-Hill.
- Spiegelhalter, D. J. (2004). Incorporating Bayesian ideas into health-care evaluation. *Statistical Science*, 19, 156-174.
- Wald, A., & Wolfowitz, J. (1944). Statistical tests based on permutation of the observations. *The Annals of Mathematical Statistics*, 15, 358-372.

Appendix A: Unique Permutation Pattern Algorithm

This computer algorithm identifies and compiles the unique permutation patterns of the layout of observations or rank of observations in a two-sample experiment. It was implemented to produce a table of critical values for the Siegel-Tukey test statistic.

Algorithm (Permutation)

Two-Sample problem: Unique Permutation Patterns
 Assign ranks to the observations based on sample size

- 1: $N \leftarrow N1 + N2$
- 2: **for** $I \leftarrow 1, N$ **do**
- 3: $X_I \leftarrow I$
- 4: **if** $I \leq N1$ **then**
- 5: $X_{2I} \leftarrow X_I$
- 6: **end if**
- 7: **end for**
- 8: $K \leftarrow 1$
- Determine the number of permutations to be sampled
- 9: $P1 \leftarrow 1 : P2 \leftarrow 1 : P3 \leftarrow 1$
- 10: **for** $I \leftarrow 1, N$ **do**
- 11: **if** $I \leq N1$ **then**
- 12: $P1 \leftarrow P1 * I$
- 13: **end if**
- 14: **if** $I > N1$ **then**
- 15: $P2 \leftarrow P2 * (I - N1)$
- 16: **end if**
- 17: $P3 \leftarrow P3 * I$
- 18: **end for**
- 19: $P \leftarrow P3 / (P1 * P2)$
- 20: Call *Pattern*
- 21: Compute Test Statistic PP and update frequency $FREQ$
 Resample without replacement
- 22: **for** $K \leftarrow 2, P$ **do**
- 23: $CHECK \leftarrow 1$
- 24: **while** $CHECK \leftarrow 1$ **do**
- 25: **for** $I \leftarrow 1, N$ **do**
- 26: $X_{1I} \leftarrow X_I$
- 27: **end for**
- 28: **for** $I \leftarrow 1, N1$ **do**
- 29: $NT \leftarrow N - I + 1$
- 30: $S \leftarrow INT(NT * RND + 1)$
- 31: $X_{2I} \leftarrow X_{1S}$
- 32: **for** $J \leftarrow S, NT$ **do**
- 33: $X_{1J} \leftarrow X_{1J+1}$
- 34: **end for**

Appendix A (continued): Unique Permutation Pattern Algorithm

```

35:     end for
36:     Call Pattern
        Ensure no repetition of permutations
37:      $CHECK \leftarrow 0$ 
38:     for  $I1 \leftarrow 1, K - 1$  do
39:         if  $PAT_K \leftarrow PAT_{I1}$  then
40:              $CHECK \leftarrow 1$ 
41:         end if
42:     end for
43:     if  $CHECK \leftarrow 1$  then
44:          $PAT_K \leftarrow ""$ 
45:     end if
46: end while
47: Compute Test Statistic  $PP$  and update frequency  $FREQ$ 
48: end for
49: Sort  $PP_I$  in ascending order for  $I \leftarrow 1, COUNT1$ 
50: PRINT  $N1, N2$ 
51:  $PV \leftarrow 0$ 
52: for  $I \leftarrow 1, COUNT1$  do
53:      $PDM \leftarrow FREQ_I/P$ 
54:      $PV \leftarrow PV + PDM$ 
55:     PRINT  $PP_I, FREQ_I, PDM, PV$ 
56: end for

```

Algorithm (Pattern)

```

1: Sort  $X2(I)$  for  $I \leftarrow 1, N1$  to have a unique pattern
2: for  $J \leftarrow 1, N1$  do
3:     if  $X2_J < 10$  then
4:          $PAT_K \leftarrow PAT_K + "0" + LTRIM(RTRIM(STR(X2_J)))$ 
5:     else  $PAT_K \leftarrow PAT_K + LTRIM(RTRIM(STR(X2_J)))$ 
6:     end if
7: end for

```

Appendix A (continued): Unique Permutation Pattern Algorithm

Algorithm (Permutation-Convergence)

Unique Permutation Patterns: Iterate to convergence
Assign ranks to the observations based on sample size

- 1: $N \leftarrow N1 + N2$
- 2: **for** $I \leftarrow 1, N$ **do**
- 3: $X_I \leftarrow I$
- 4: **if** $I \leq N1$ **then**
- 5: $X_{2I} \leftarrow X_I$
- 6: **end if**
- 7: **end for**
- 8: $BB \leftarrow 0$
- 9: $CK \leftarrow 1$
- 10: Ensure error tolerance is achieved
- 11: **while** $CK \leftarrow 1$ **do**
- 12: **for** $C0 \leftarrow 1, COUNT1$ **do**
- 13: $FREQ_{C0} \leftarrow 0$
- 14: **end for**
- 15: $COUNT1 \leftarrow 0$
- 16: **for** $K \leftarrow 1, NN$ **do**
- 17: $PAT_K \leftarrow ""$
- 18: **end for**
- 19: $BB \leftarrow BB + 1$
Resample without replacement
- 20: **for** $K \leftarrow 1, NN$ **do**
- 21: $CHECK \leftarrow 1$
Ensure pattern is unique
- 22: **while** $CHECK \leftarrow 1$ **do**
- 23: **for** $I \leftarrow 1, N$ **do**
- 24: $X_{1I} \leftarrow X_I$
- 25: **end for**
- 26: **for** $I \leftarrow 1, N1$ **do**
- 27: $NT \leftarrow N - I + 1$
- 28: $S \leftarrow INT(NT * RND + 1)$
- 29: $X_{2I} \leftarrow X_{1S}$
- 30: **for** $J \leftarrow S, NT$ **do**
- 31: $X_{1J} \leftarrow X_{1J+1}$
- 32: **end for**
- 33: **end for**
- 34: Sort X_{2I} to have a unique pattern

Appendix A (continued): Unique Permutation Pattern Algorithm

```

35:         Call Pattern
           Ensure no repetition of permutations
36:          $CHECK \leftarrow 0$ 
37:         for  $I_1 \leftarrow 1, K - 1$  do
38:             if  $PAT_K \leftarrow PAT_{I_1}$  then
39:                  $CHECK \leftarrow 1$ 
40:             end if
41:         end for
42:         if  $CHECK \leftarrow 1$  then
43:              $PAT_K \leftarrow ""$ 
44:         end if
45:     end while
46:     Compute Test Statistic  $PP$  and update frequency  $FREQ$ 
47: end for
48: Sort  $PP_I$  in ascending order for  $I \leftarrow 1, COUNT1$ 
   Merge and Update pdf
49: if  $BB > 1$  then
50:      $NP \leftarrow COUNT2 + COUNT1$ 
51:     for  $I \leftarrow 1, NP$  do
52:         if  $I \leq COUNT2$  then
53:              $UPD_I \leftarrow UPD_{1I}$ 
54:              $FREQ_I \leftarrow FREQ_{1I}$ 
55:         else
56:              $UPD_I \leftarrow PP_{I-COUNT2}$ 
57:              $FREQ_I \leftarrow FREQ_{2I-COUNT2}$ 
58:         end if
59:     end for
60:     for  $J \leftarrow 1, COUNT2$  do
61:         for  $I \leftarrow COUNT2 + 1, NP$  do
62:             if  $UPD_J \leftarrow UPD_I$  then
63:                  $FREQ_J \leftarrow FREQ_J + FREQ_I$ 
64:                 for  $T \leftarrow I, NP - 1$  do
65:                      $UPD_T \leftarrow UPD_{T+1}$ 
66:                      $FREQ_T \leftarrow FREQ_{T+1}$ 
67:                 end for
68:                  $NP \leftarrow NP - 1$ 
69:             end if
70:         end for
71:     end for
72:     Sort the update  $UPD$  in ascending order
73: else

```

JUSTICE I. ODIASE

Appendix A (continued): Unique Permutation Pattern Algorithm

```

74:      $NP \leftarrow COUNT1$ 
75:     for  $I2 \leftarrow 1, NP$  do
76:          $UPD_{I2} \leftarrow PP_{I2}$ 
77:          $FREQ_{I2} \leftarrow FREQ2_{I2}$ 
78:     end for
79: end if
80: Print Number of trials  $BB$ , Permutation Sample size  $NN$ , Sample1
     $N1$ , Sample2  $N2$ 
81:  $PV \leftarrow 0$ 
82: for  $I \leftarrow 1, NP$  do
83:      $PDM_I \leftarrow FREQ_I / NN * BB$ 
84:      $PV \leftarrow PV + PDM_I$ 
85:     PRINT  $UPD_I, FREQ_I, PDM_I, PV$ 
86: end for
    Check error tolerance
87: if  $BB > 1$  then
88:      $CK \leftarrow 0$ 
89:     for  $I \leftarrow 1, COUNT2$  do
90:         for  $II \leftarrow 1, NP$  do
91:             if  $UPD1_I \leftarrow UPD_{II}$  And  $ABS(PDM1_I - PDM_{II}) >$ 
                 $TOL$  then
92:                  $CK \leftarrow 1$ 
93:             end if
94:         end for
95:     end for
96: end if
    Store last iteration
97: for  $I \leftarrow 1, NP$  do
98:      $UPD1_I \leftarrow UPD_I$ 
99:      $FREQ1_I \leftarrow FREQ_I$ 
100:     $PDM1_I \leftarrow PDM_I$ 
101: end for
102:  $COUNT2 \leftarrow NP$ 
103: end while

```