Rank Transformations as a Bridge Between Parametric and Nonparametric Statistics

W.J. CONOVER AND RONALD L. IMAN*

Many of the more useful and powerful nonparametric procedures may be presented in a unified manner by treating them as rank transformation procedures. Rank transformation procedures are ones in which the usual parametric procedure is applied to the ranks of the data instead of to the data themselves. This technique should be viewed as a useful tool for developing nonparametric procedures to solve new problems.

KEY WORDS: Nonparametric; Rank transformation; Regression; Experimental designs; Discriminant analysis; Multiple comparisons

1. INTRODUCTION

A problem that applied statisticians have been confronted with virtually since the inception of parametric statistics is that of fitting real world problems into the framework of normal statistical theory when many of the data they deal with are clearly nonnormal. From such problems have emerged two distinct approaches or schools of thought: (a) transform the data to a form more closely resembling a normal distribution framework or (b) use a distribution free procedure. The first method may include the log transformation, square root transformation, arcsin transformation, and so forth, and may even be broad enough to include robust procedures that tend to give small weights to outliers, that is, to observations that may contribute greatly to the nonnormal form of the data. The second method includes a large body of methods based on the ranks of the data.

There is a way of combining these two methods by presenting many nonparametric methods as parametric methods applied to transformed data. Simply replace the data with their ranks, then apply the usual parametric t test, F test, and so forth, to the ranks. We call this the rank transformation (RT) approach. This approach results in a class of nonparametric methods that includes the Wilcoxon-Mann-Whitney test, the Kruskal-Wallis test, the Wilcoxon signed ranks test, the Friedman test, Spearman's rho, and others. The rank transformation approach also furnishes useful methods in multiple regression, discriminant analysis, cluster analysis, analysis of experimental designs, and multiple comparisons.

Of course, there are several ways in which ranks can be assigned to observations. We suggest the following types.

RT-1. The entire set of observations is ranked from smallest to largest, with the smallest observation having rank 1, the second smallest rank 2, and so on. Average ranks are assigned in case of ties.

RT-2. The observations are partitioned into subsets and each subset is ranked within itself independently of the other subsets.

RT-3. This rank transformation is RT-1 applied after some appropriate reexpression of the data.

RT-4. The RT-2 type is applied to some appropriate reexpression of the data.

The rank transformation approach provides a useful pedagogical technique for introducing these nonparametric methods as an integral part of an introductory course in statistics, instead of isolating the methods in a separate unit that may appear to the student to be disconnected from the general flow of the course. Also, it allows the practitioner to make full use of existing statistical packages that may not have suitable nonparametric programs by simply entering the ranks of the data into the programs for the parametric analysis. And finally, this approach may be viewed as a useful tool for developing new nonparametric methods in situations where satisfactory parametric procedures exist.

2. TWO INDEPENDENT SAMPLES

Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ represent two independent random samples. To test the hypothesis that $E(X) = E(Y)$ the parametric procedure employs the two-sample t statistic

$$t = \frac{\bar{X} - \bar{Y}}{\left[\left(\sum (X_i - \bar{X})^2 + \sum (Y_i - \bar{Y})^2\right) \dfrac{N}{nm(N-2)}\right]^{1/2}} \tag{2.1}$$

where $N = n + m$, and compares t with quantiles from the t distribution with $N - 2$ degrees of freedom (df). The nonparametric Wilcoxon-Mann-Whitney two-sample test requires replacing the data by the ranks R_i from 1 to N, and uses the statistic, in its standardized form with the adjustment for ties incorporated,

$$T = \frac{S - n(N+1)/2}{\left[\dfrac{nm}{N(N-1)} \sum_{i=1}^{N} R_i^2 - \dfrac{nm(N+1)^2}{4(N-1)}\right]^{1/2}} \tag{2.2}$$

* W.J. Conover is Horn Professor of Statistics and Associate Dean for Research and Graduate Studies in the College of Business Administration, Texas Tech University, Lubbock, TX 79409. Ronald L. Iman is a Statistical Consultant, Division 1223, Sandia National Laboratories, Albuquerque, NM 87185.

where $S = \sum_{i=1}^{n} R_i$ is the sum of the ranks of the X's. The statistic T is compared with the standard normal distribution or, if there are no ties and the sample sizes are less than 20, exact tables may be used for S (Conover 1980).

A rank transformation procedure is based on computing t on the ranks R_i to get the statistic

$$t_R = \frac{1}{n} S - \frac{1}{m}\left(\frac{N(N+1)}{2} - S\right) \div \left[\left(\sum_{i=1}^{N} R_i^2 - \frac{1}{n} S^2 - \frac{1}{m}\left(\frac{N(N+1)}{2} - S\right)^2\right)\frac{N}{nm(N-2)}\right]^{1/2} \tag{2.3}$$

and using the t tables as with (2.1). This is an example of an RT-1 type procedure. A little algebra reveals an important relationship between t_R and T:

$$t_R = \frac{T}{\left[\frac{N-1}{N-2} - \frac{1}{N-2} T^2\right]^{1/2}}, \tag{2.4}$$

which shows that t_R is a monotonically increasing function of T. Since T contains the correction for ties, so does t_R because it is a function of T.

Let us consider the implication of (2.4). When T is in its upper α level tail region, then t_R is in its upper α level tail region also. The same can be said for the lower α level tail regions of each. For example, when $n = 14$ and $m = 18$ the upper five percent value for S (Conover 1980, Table A.7) is 274 if there are no ties. Substitution of S into (2.2) and (2.3) reveals the exact upper five percent values for T and t_R to be 1.633 and 1.681, respectively. When T is compared with the upper five percent quantile 1.645 from a standard normal distribution, a slightly conservative test results. The t distribution with 30 df gives an upper five percent critical value of 1.697, also resulting in a slightly conservative test for t_R. Because t_R is a monotonic function of T, the two tests are equivalent when the exact critical values are used. The normal distribution and the t distribution provide two different approximations, which have been compared by Iman (1976). The Wilcoxon-Mann-Whitney test, with all of its good properties, may be performed using t_R as a test statistic instead of T if desired. In fact t_R may be preferred, as computing routines are generally readily available for the t statistic; and also it may be simpler to teach this procedure to someone who understands the t test and transformations, but who does not have a working knowledge of nonparametric statistics.

3. *k* INDEPENDENT SAMPLES

Consider k independent random samples, $(X_{11}, \ldots, X_{1n_1}), \ldots, (X_{k1}, \ldots, X_{kn_k})$ and let

$$SSA = \sum_{i=1}^{k} n_i(\bar{X}_i - \bar{X}_{..})^2 \tag{3.1}$$

and

$$SSE = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (\bar{X}_{ij} - \bar{X}_i)^2 \tag{3.2}$$

represent the analysis of variance sums of squares. The usual parametric test of the hypothesis of equal means compares the statistic

$$F = (SSA/(k-1))/(SSE/(N-k)) \tag{3.3}$$

with the F distribution, $k - 1$ and $N - k$ df.

For the Kruskal-Wallis test the data are replaced by their ranks $R(X_{ij})$ from 1 to $N = \sum n_i$ and the statistic, incorporating the correction for ties, is given as

$$H = \frac{\sum_{i=1}^{k} R_i^2/n_i - N(N+1)^2/4}{(\sum_i \sum_j R^2(X_{ij}) - N(N+1)^2/4)/(N-1)} \tag{3.4}$$

where

$$R_i = \sum_{j=1}^{n_i} R(X_{ij}).$$

The statistic H is compared with the chi-squared distribution, $k - 1$ df. The most extensive tables of the exact distribution of H, applicable only if there are no ties, are by Iman, Quade, and Alexander (1975).

A rank transformation procedure is based on computing F on the ranks to get

$$F_R = \frac{\left[\sum_{i=1}^{k} \frac{R_i^2}{n_i} - \frac{N(N+1)^2}{4}\right] \Big/ (k-1)}{\sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(R(X_{ij}) - \frac{R_i}{n_i}\right)^2 \Big/ (N-k)} \tag{3.5}$$

as a test statistic. This is another example of an RT-1 type procedure. Elementary algebra reveals F_R is a monotonic function of H,

$$F_R = (H/(k-1))/((N-1-H)/(N-k)) \tag{3.6}$$

so the two tests are equivalent. The upper α level critical value from the chi-squared distribution, when substituted for H in (3.6), results in a slightly different critical value for F_R than that obtained from the appropriate tables of the F distribution. Both methods, however, merely provide approximations to the true critical value. Iman and Davenport (1976) compare these and other approximations and show that the F approximation should be preferred to the chi-squared in most cases.

4. THE ONE-SAMPLE OR MATCHED-PAIRS PROBLEM

Let $D_1, \ldots, D_n$ represent independent random variables with a common mean where, in the case of matched pairs (X_i, Y_i), D_i equals $X_i - Y_i$. To test the hypothesis $E(D) = 0$ the one-sample t statistic

$$t = \frac{\sum D_i}{\left[\frac{n}{n-1} \sum D_i^2 - \frac{1}{n-1} (\sum D_i)^2\right]^{1/2}} \tag{4.1}$$

is compared with the t distribution, $n - 1$ df, in the parametric test valid for normally distributed D's.

For the Wilcoxon signed ranks test the D_i's are replaced by the signed ranks R_i, where

$$R_i = (\text{sign } D_i) \times (\text{rank of } |D_i| \text{ among } |D_1|, \ldots, |D_n|). \quad (4.2)$$

The hypothesis is rejected when the test statistic

$$T = (\sum R_i)/(\sqrt{\sum R_i^2}) \quad (4.3)$$

is too large or too small, as measured by the normal approximation. If n is small and there are no ties, exact tables may be used (cf. Conover 1980, Table A.13). The statistic T is equivalent to the form for this test, which appears in most textbooks and is based on the sum of the positive ranks only. This form is simpler to use in the presence of ties, since the correction for ties is incorporated into (4.3).

Alternatively, the one-sample t statistic may be computed on the signed ranks, resulting in

$$t_R = \frac{\sum R_i}{\left[\dfrac{n}{n-1} \sum R_i^2 - \dfrac{1}{n-1} (\sum R_i)^2\right]^{1/2}} \quad (4.4)$$

which is compared with the t distribution, $n - 1$ df, as an approximation. Since D_i represents a reexpression of the data (X_i, Y_i) and may be considered to be the product of (sign D_i) and $|X_i - Y_i|$, this is an example of an RT-3 type procedure.

Note that t_R is also expressible as

$$t_R = \frac{T}{\left[\dfrac{n}{n-1} - \dfrac{1}{n-1} T^2\right]^{1/2}}, \quad (4.5)$$

which is a monotonic function of T. Thus the test that uses t_R is equivalent to the test that uses T. A comparison of the normal approximation, the student's t approximation, and the exact distribution is given by Iman (1974a).

Suppose t in (4.1) is applied directly to RT-1 type ranks; that is, the X's and Y's are replaced by their corresponding ranks 1 to $2n$ and D_i is the difference in those ranks. This application of the rank transformation approach does not yield the Wilcoxon signed ranks test, but rather introduces a new procedure. This new procedure is conditionally distribution free given the ranks in the blocks and asymptotically distribution free by virtue of the central limit theorem. Properties of this test are reported by Iman and Conover (1980a).

5. THE RANDOMIZED COMPLETE BLOCK DESIGN

In the randomized complete block design with one observation per cell, let X_{ij} be the random variable for block i, treatment j, $i \le b$, and $j \le k$. If $\bar{X}_{i\cdot}$, $\bar{X}_{\cdot j}$, and $\bar{X}_{\cdot\cdot}$ represent the block, treatment, and grand means respectively, then

$$SST = b \sum_{j=1}^{k} (\bar{X}_{\cdot j} - \bar{X}_{\cdot\cdot})^2 \quad (5.1)$$

and

$$SSE = \sum_{i=1}^{b} \sum_{j=1}^{k} (X_{ij} - \bar{X}_{\cdot j} - \bar{X}_{i\cdot} + \bar{X}_{\cdot\cdot})^2 \quad (5.2)$$

are the analysis of variance treatment and error sums squares. The parametric test of equal treatment effects compares the statistic

$$F = (SST/(k - 1))/(SSE/(b - 1)(k - 1)) \quad (5.3)$$

with the F distribution, $k - 1$ and $(b - 1)(k - 1)$ df.

The usual nonparametric test involves ranking the observations from 1 to k within each block, making no interblock comparisons. The Friedman test uses the statistic, corrected for ties,

$$T = \frac{(k-1) \sum_{j=1}^{k} [R_j - b(k+1)/2]^2}{\sum_i \sum_j R^2(X_{ij}) - bk(k+1)^2/4} \quad (5.4)$$

where R_j is the sum of the ranks $R(X_{ij})$ for treatment j. The chi-squared distribution with $k - 1$ df is used as an approximation to the distribution of T.

Another way of considering the Friedman test is to compute the F statistic in (5.3) on the intrablock ranks that are used in the Friedman test. This is a type RT-2 procedure. The result is a statistic F_R that is a monotonic function of the Friedman statistic

$$F_R = (T/(k - 1))/((b(k - 1) - T)/(b - 1)(k - 1)). \quad (5.5)$$

Comparison of F_R with the F distribution provides a more accurate approximation (Iman and Davenport 1980) than the chi-squared approximation used with (5.4).

Suppose F is applied directly to RT-1 type ranks where all of the observations are ranked together, from 1 to bk in this case. This type of ranking takes advantage of both between and within block information. The result is a test which is conditionally distribution free, given the partitioning of ranks into blocks. This procedure, with the F distribution as an approximation, compares favorably with the RT-2 type Friedman test (Iman and Conover 1980a), and even Fisher's randomization test (Conover and Iman 1980a) in terms of robustness and power.

It is easy to extend the RT-1 type procedure to other experimental designs. This approach is robust and powerful in the two-way layout with interaction (Iman 1974b), in a test for interaction when replication effects are present (Conover and Iman 1976), and in a test for main effects in the presence of replication and interaction effects (Iman and Conover 1976).

The advantage of ranking all of the observations together is that any analysis of variance procedure may be applied to the ranks, with the resulting tests for main effects, interactions, or whatever, following immediately. Other rank tests that involve a separate ranking for each test of hypothesis become difficult or impossible to apply. The same may be said for aligned ranks tests, in which the appropriate means are first subtracted from each observation before ranking. This resembles an RT-3 type procedure. In addition, for aligned ranks tests some power may be lost in the process (Conover and Iman 1976).

6. CORRELATION

One of the earliest applications of a rank transformation involves computing Pearson's product moment correlation coefficient

$$r = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{[\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2]^{1/2}} \tag{6.1}$$

on ranks to obtain Spearman's rho

$$\rho = \frac{\sum \left(R(X_i) - \frac{n+1}{2}\right)\left(R(Y_i) - \frac{n+1}{2}\right)}{\left[\sum \left(R(X_i) - \frac{n+1}{2}\right)^2 \sum \left(R(Y_i) - \frac{n+1}{2}\right)^2\right]^{1/2}} \tag{6.2}$$

for paired data $(X_1, Y_1), \ldots, (X_n, Y_n)$. Since the observations within the subset $\{X_i\}_{i=1}^n$ are ranked within themselves, and the same is true for the subset $\{Y_i\}_{i=1}^n$, this is an example of an RT-2 type procedure. Just as r is a measure of linearity of the relationship between X and Y, so is ρ a measure of the linearity between the ranks of X and the ranks of Y, which translates as a measure of monotonicity in the relationship between X and Y.

The direct extension to multiple correlation is obvious. The observations on each component X_{ij} of $\mathbf{X}_j = (X_{1j}, \ldots, X_{kj}), j = 1, \ldots, n$ are ranked separately from 1 to n. Multiple correlations, partial correlations, and the like may be computed on the ranks just as they would be computed on the data.

To test for independence between X and Y the statistic

$$t = r\sqrt{n-2}/\sqrt{1-r^2} \tag{6.3}$$

is compared with the student's t distribution with $n - 2$ df, in a parametric test valid with bivariate normal distributions. The nonparametric test statistic takes the form

$$Z = \rho\sqrt{n-2}, \tag{6.4}$$

which is compared with the standard normal distribution.

The computation of t on the rank transformed observations results in the statistic

$$t_R = \rho\sqrt{n-2}/\sqrt{1-\rho^2}, \tag{6.5}$$

which is compared with the same distribution used with (6.3). This approximation was suggested by Pitman (1937). A comparison of the normal approximation on Z with the student's t approximation on t_R by Iman and Conover (1978) shows the latter approximation to be slightly better.

7. REGRESSION

The adaptation of correlation procedures to the ranks of the data immediately suggests adapting regression methods to RT-2 type data. Least squares, forward or backward stepwise regression, or any other regression method may be applied to the ranks of the observations (Iman and Conover 1979). The result is that the rank of the dependent variable is predicted using the ranks of the independent variables. The predicted rank $\hat{R}(Y_i)$ may be transformed back into a predicted value $\hat{Y}$ of Y, by linear interpolation between the two values of Y that have ranks bracketing $\hat{R}(Y_i)$.

Another way of using a rank transformation to develop nonparametric methods in regression is to assume the regression equation is linear and use ordinary least squares to obtain an estimated regression equation $Y = a + bx$. The parametric method of testing the hypothesis $\beta = \beta_0$ for slope uses the statistic

$$t = \frac{(b - \beta_0)[(n-2) \sum (X_i - \bar{X})^2]^{1/2}}{[\sum (Y_i - \bar{Y} - b(X_i - \bar{X}))^2]^{1/2}}, \tag{7.1}$$

which has a t distribution under certain assumptions including normality. It is easier to see how a rank transformation leads to a nonparametric test if t is rearranged as follows:

$$t = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y} - \beta_0(X_i - \bar{X}))\sqrt{n-2}}{[\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y} - b(X_i - \bar{X}))^2]^{1/2}}. \tag{7.2}$$

Note that, except for a multiplicative constant, the numerator is the sample covariance of X and the hypothesized residuals, and the denominator has the sample variances of X and the least squares residuals. Replace the covariance with the covariance between the ranks of X and the ranks of the hypothesized residuals, which is an RT-4 type transformation. Use the variances of those ranks in the denominator, and one has the nonparametric test for slope proposed by Hogg and Randles (1975).

A natural extension of the rank transformation approach to the general linear model consists of ranking each quantitative variable separately in the general linear model and applying usual parametric procedures. This is an RT-2 type application. The result is new tests for equal slopes, analysis of covariance (Conover and Iman 1980b), and any experimental designs covered by the general linear model. These tests are in general not distribution free, except perhaps in some asymptotic sense, but they may be more robust and powerful than their competitors in nonnormal situations. Each procedure, however, needs to be evaluated on its own merits. Another RT-2 type procedure for analysis of covariance is given by Quade (1967).

8. DISCRIMINANT ANALYSIS

The usual linear discriminant function (LDF) and quadratic discriminant function (QDF) have been applied to the RT-2 type data with good results. In an extensive simulation study (Conover and Iman 1980c) these rank transformation methods showed an ability to discriminate between populations that equaled the LDF and QDF procedures with normal populations, and surpassed them with nonnormal populations. Unlike the nonparametric competitors, these methods can be used with small samples. Also no new computer programs are required, since routines with LDF and QDF are readily available. All that is required is an RT-2 type transformation of the data before entering the data into the computer.

9. MULTIPLE COMPARISONS

Any of the popular multiple comparisons techniques, including Scheffé's, Tukey's, Duncan's, and Fisher's methods, as well as others, may be applied to the RT-1 type data with good results. The power with normal populations is about the same whether the analysis is done on the data or on the ranks. With nonnormal populations the multiple comparisons procedures are more robust and have more power when rank transformed data are used (Iman and Conover 1980b).

10. BIOASSAY

Williams (1971, 1972) developed a procedure designed to compare the toxicity of various dose levels of a drug against the zero dose level. This same procedure is used on the RT-1 type data by Shirley (1977) in a nonparametric test.

11. CLUSTER ANALYSIS

Scott and Knott (1974) developed a method of partitioning means in the analysis of variance into two clusters. The same procedure is applied to the RT-1 type data to obtain a distribution free procedure by Worsley (1977).

12. DISCUSSION

Nonparametric methods should be among the working tools of any statistician. The rank transformation approach provides a vehicle for presenting both the parametric and nonparametric methods in a unified manner. This should enable novice statisticians to understand the differences and similarities of the two types of analysis. Also the rank transformation approach leads to easier computational methods, since it is often more convenient to enter ranks into a program for parametric analysis than it is to find or write a program for a nonparametric analysis. Most existing programs for nonparametric methods do not incorporate corrections for ties, while rank transformation procedures all automatically make the required corrections for ties. The user merely uses average ranks whenever ties occur.

Other scores may be used instead of ranks, if desired, to obtain nonparametric tests that are equivalent to tests such as the van der Waerden test, Capon test, median test, McNemar test, and others. Our experience indicates that the ranks themselves provide scores that are difficult to improve upon for general all-around use.

Some limitations of the rank transformation approach should be noted here also. The rank transformation procedures lead to distribution free tests in some cases, while in other cases the resulting tests may be only conditionally distribution free, asymptotically distribution free, or neither. An example of the latter case arises when the ratio of two sample variances is computed on RT-1 type data and compared with tables of the F distribution as a test for equal variances in two independent samples. Simulation results indicate a severe lack of robustness for this procedure, even when the population means are equal. This is probably related to the fact that the parametric F test for this problem is notoriously sensitive to normality, and the central limit theorem does not apply in this situation. A referee pointed out that a similar situation exists when Welch's t test is used on ranks of the data in the hopes of solving the nonparametric Behrens-Fisher problem.

[*Received February 1980. Revised March 1981.*]

REFERENCES

CONOVER, W.J. (1980), *Practical Nonparametric Statistics* (2nd ed.), New York: John Wiley.

CONOVER, W.J., and IMAN, RONALD L. (1976), "On Some Alternative Procedures Using Ranks for the Analysis of Experimental Designs," *Communications in Statistics*, Ser. A, 5, 1348–1368.

——— (1980a), "Small Sample Efficiency of Fisher's Randomization Test when Applied to Experimental Designs," unpublished manuscript presented at the annual meeting of the American Statistical Association, Houston, August 1980.

——— (1980b), "Analysis of Covariance Using the Rank Transformation," unpublished manuscript.

——— (1980c), "The Rank Transformation as a Method of Discrimination with Some Examples," *Communications in Statistics*, Ser. A, 9, 465–487.

HOGG, ROBERT V., and RANDLES, RONALD H. (1975), "Adaptive Distribution-Free Regression Methods and Their Applications," *Technometrics*, 17, 399–407.
IMAN, RONALD L. (1974a), "Use of a t-statistic as an Approximation to the Exact Distribution of the Wilcoxon Signed Ranks Test Statistic," *Communications in Statistics*, 3, 795–806.
——— (1974b), "A Power Study of a Rank Transform for the Two Way Classification Model When Interaction May be Present," *Canadian Journal of Statistics*, 2, 227–239.
——— (1976), "An Approximation to the Exact Distribution of the Wilcoxon-Mann-Whitney Rank Sum Test Statistic," *Communications in Statistics*, Ser. A, 5, 587–598.
IMAN, RONALD L., and CONOVER, W.J. (1976), "A Comparison of Several Rank Tests for the Two-Way Layout," Technical Report SAND76-0631, Sandia Laboratories, Albuquerque, New Mexico.
——— (1978), "Approximations of the Critical Region for Spearman's Rho With and Without Ties Present," *Communications in Statistics*, Ser. B, 7, 269–282.
——— (1979), "The Use of the Rank Transform in Regression," *Technometrics*, 21, 499–509.
——— (1980a), "A Comparison of Distribution Free Procedures for the Analysis of Complete Blocks," unpublished manuscript presented at the annual meeting of the American Institute of Decision Sciences, Las Vegas, November 1980.
——— (1980b), "Multiple Comparisons Procedures Based on the Rank Transformation," unpublished manuscript presented at the annual meeting of the American Statistical Association, Houston, August 1980.
IMAN, RONALD L., and DAVENPORT, JAMES M. (1976), "New Approximations to the Exact Distribution of the Kruskal-Wallis Test Statistic," *Communications in Statistics*, Ser. A, 5, 1335–1348.
——— (1980), "Approximations of the Critical Region of the Friedman Statistic," *Communications in Statistics*, Ser. A, 9, 571–595.
IMAN, RONALD L., QUADE, DANA, and ALEXANDER, DOUGLAS (1975), "Exact Probability Levels for the Kruskal-Wallis Test," in *Selected Tables in Mathematical Statistics* (Vol. 3), eds. H.L. Harter and D.B. Owen, Providence, R.I.: American Mathematical Society.
PITMAN, E.J.G. (1937), "Significance Tests Which May be Applied to Samples from any Populations: II. The Correlation Coefficient Test," *Journal of the Royal Statistical Society*, Suppl. 4, 225–232.
QUADE, DANA (1967), "Rank Analysis of Covariance," *Journal of the American Statistical Association*, 62, 1187–1200.
SCOTT, A.J., and KNOTT, M. (1974), "A Cluster Analysis Method for Grouping Means in the Analysis of Variance," *Biometrics*, 30, 507–512.
SHIRLEY, E. (1977), "A Non-Parametric Equivalent of Williams' Test for Contrasting Increasing Dose Levels of Treatment," *Biometrics*, 33, 386–389.
WILLIAMS, D.A. (1971), "A Test for Differences Between Treatment Means When Several Dose Levels are Compared With a Zero Dose Control," *Biometrics*, 27, 103–117.
——— (1972), "A Comparison of Several Dose Levels With a Zero Dose Control," *Biometrics*, 28, 519–531.
WORSLEY, K.J. (1977), "A Non-Parametric Extension of a Cluster Analysis Method by Scott and Knott," *Biometrics*, 33, 532–535.

Comment

GOTTFRIED E. NOETHER*

To gain greater generality, statisticians have applied standard normal theory procedures to the ranks of observations for a long time. Conover and Iman have now systematized this approach in a series of papers. The present survey provides a readable, nontechnical account of this work.

The authors call attention to three potential uses of the method of rank transformations: (a) as a pedagogical technique for incorporating nonparametrics in introductory statistics courses; (b) as a device for adapting normal-theory computer packages to the computation of nonparametric statistics; and (c) as a tool for developing new nonparametric procedures in situations where parametric procedures already exist. My concern is primarily with the first of these three suggested uses.

During the last 30 years or so the role played by nonparametrics has undergone considerable change. In the 1950's, and even the 1960's, it was quite common to characterize nonparametric methods as "rough and ready." Few practicing statisticians were inclined to accept them into their regular arsenal of statistical tools. This attitude is clearly changing. Now practically every introductory statistics text contains a chapter concerning nonparametric, or distribution-free, or rank methods. As Conover and Iman point out, "Nonparametric methods should be among the working tools of any statistician."

The question is no longer whether but how nonparametrics should be incorporated in an introductory statistics course. Conover and Iman advocate that nonparametric methods be taught within the normal-theory framework. I should like to give three reasons for opposing this point of view.

1. It unnecessarily restricts the perceived applicability of the nonparametric procedure.
2. It further emphasizes the widely held misconception that the nonparametric approach is essentially restricted to hypothesis testing.

* Gottfried E. Noether is Professor and Head of the Department of Statistics at the University of Connecticut, Storrs, CT 06268.