THE INSTITUTE OF STATISTICS

ONE- AND TWO-WAY ANOVA AND RANK TESTS WHEN THE NUMBER OF TREATMENTS IS LARGE

by

Dennis D. Boos and Cavell Brownie

Institute of Statistics Mimeograph Series No. 1993

May 1991

NORTH CAROLINA STATE UNIVERSITY
Raleigh, North Carolina

ABSTRACT

In this paper we consider the popular analysis of variance (ANOVA) F tests, and rank statistic analogs, for testing equality of treatment means in the one-way and two-way experimental layout. The rank-based procedures include the Kruskal-Wallis and Friedman statistics with chi-squared critical values, and the "ANOVA on ranks" or F versions of these procedures. Our focus is on robustness of Type I error rates under nonnormality when the number of treatments, k, is large but the number of replications, n, is small. Two approaches are used to provide insight concerning null performance of the ANOVA F and rank procedures in this large k, small n, situation. The first approach is based on standard central limit theory for the (nonstandard) situation that $k \to \infty$, and the second approach is the classical moment approximation to the permutation distribution of the F statistic. Both approaches confirm robustness of the ANOVA F under nonnormality in the large k situation and also provide justification for the "ANOVA on ranks" procedure. The moment-based adjustments to the degrees of freedom for the usual ANOVA F or for the F on ranks are recommended for routine data analysis purposes.

Key Words: Kruskal-Wallis test, Friedman test, central limit theorem, permutation distribution, Type I error robustness, nonnormality.

1. INTRODUCTION

In a wide range of disciplines, experiments to compare k treatments are carried out using either a one-way or a two-way layout (i.e., a completely randomized or a randomized complete block design). Equality of the k treatment means is tested using the appropriate one-way or two-way analysis of variance (ANOVA) F procedure or, alternatively, the rank-based Kruskal-Wallis or Friedman statistics (often with chi-squared percentiles).

Many studies have investigated the null performance of these procedures because of the importance placed by the experimenter on good agreement between nominal and actual Type I error rates. For example, Box and Andersen (1955) using moment calculations discuss the effect of nonnormality on the ANOVA F. For the rank statistics, using analogous moment calculations, Kendall and Babington Smith (1939) and Wallace (1959) suggested alternatives to the chi-squared percentiles for small sample sizes.

Asymptotic normal and chi-squared approximations are typically derived for the situation where n, the number of replications per treatment, is large (i.e., the standard k fixed, $n \to \infty$ asymptotics). Frequently, however, as in agricultural screening trials, the number of treatments may be large while replication per treatment is extremely limited. In this situation asymptotic results based on n fixed, $k \to \infty$ should provide more useful insight concerning robustness of Type I error rates. Although important in practice, this "small n, large k" situation is seldom addressed explicitly in the literature on robustness of either the ANOVA or the rank procedures. (One exception is Pirie, 1974.)

In this article we therefore focus on $k \to \infty$ asymptotics for the one-way and two-way layouts. Our main objective is to provide a basis for understanding the properties of the familiar ANOVA and rank procedures in the large k setting. A secondary objective is to illustrate the utility of the central limit theorem (CLT) in elucidating the properties of standard procedures in nonstandard but important situations.

In Section 2 we show that under nonnormality the ANOVA F statistics are asymptotically distribution-free as $k \to \infty$, and note that this provides additional justification for robustness of the F percentiles. Also, for readers who teach a course in asymptotic theory, we suggest that the derivation for the one-way F can be used as an exercise for which the motivation is readily apparent. In Section 3, similar asymptotic results are given for the analogous rank statistics. These results support a conjecture of Friedman (1937) for the large k situation and provide new justification for the "F" versions of the rank statistics suggested by Kendall and Babington Smith (1939) and Wallace (1959). Section 4 discusses the example which motivated this study, and the Appendix contains proofs of asymptotic results.

2. ANOVA F STATISTICS

2.1 The one-way layout

We consider first the one-way layout or completely randomized design. For simplicity we assume equal sample sizes n, equal variances, and the null situation H_{01}:all k means equal. That is, we assume

$X_{11},\dots,X_{1n}, X_{21},\dots,X_{2n},\dots,X_{k1},\dots,X_{kn}$ are N=kn iid random variables each with mean μ and variance $\sigma^2 < \infty$. (S1)

The one-way F statistic for detecting differences in the k treatment means is

$$F = \frac{1}{k-1} \sum_{i=1}^{k} n(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})^2 \Big/ S_p^2 \ , \qquad (1)$$

where $\overline{X}_{i\cdot} = n^{-1}\sum_{j=1}^{n} X_{ij}$, $\overline{X}_{\cdot\cdot} = k^{-1}\sum_{i=1}^{k}\overline{X}_{i\cdot}$, and $S_p^2 = k^{-1}\sum_{i=1}^{k} s_i^2$ with $s_i^2 = (n-1)^{-1}\sum_{j=1}^{n}(X_{ij} - \overline{X}_{i\cdot})^2$.

2.1.1 Asymptotics Based on the CLT

The more standard asymptotic situation is to allow the number of observations per treatment to get large. Let χ_ν^2 represent a chi-squared random variable with ν degrees of freedom. The following well-known result is given for completeness.

Lemma 1 (# of trts. fixed). Under (S1), and with k fixed, $(k-1)F \xrightarrow{d} \chi^2_{k-1}$ as $n \to \infty$.

Allowing the number of treatments to get large is less standard in the ANOVA context, but is not uncommon in other contexts such as the analysis of contingency tables (e.g., Santner and Duffy, 1989 , p. 232). Lemma 2 below follows from the CLT but in a different manner from the result of Lemma 1. The proof in the Appendix is straightforward and could be used as an exercise in a course on asymptotic theory. Examples motivating this exercise are easy to find, enhancing the value of the exercise for students with interest in applications.

Lemma 2 (# of trts. → ∞). Under (S1), and with n fixed, $k^{1/2}(F-1) \xrightarrow{d} N(0,2n/(n-1))$ as $k \to \infty$.

Note that both Lemmas 1 and 2 are true under H_{01} for any type of distribution with finite second moment. That is, the F statistic is asymptotically distribution-free, so that use of the $F(k-1,k(n-1))$ percentiles to obtain critical values yields asymptotically valid tests for either k or n large, or for both k and n large. Here $F(m,n)$ represents the F distribution with numerator and denominator degrees of freedom m and n, respectively.

This Type I error robustness of the F statistic for k or n large was well known among earlier statisticians (c.f., Scheffe, 1959, Chapter 10). It is less evident in current "methods" texts, or is reported only for the large n situation (e.g., Zar, 1984, p. 170), probably because recent work has tended to focus on Type II error robustness and alternatives to the F statistic. It seems instructive, therefore, to briefly review the classic work of Pitman, Welch, and Box and Andersen, which was not based on central limit theory but provided a means for predicting the effect of distribution types other than the normal on the ANOVA F test for moderate k and n. This is done for both the one- and two-way classifications in Section 2.3.

2.2 The Two-way Layout

For the two-way layout or randomized complete block design we assume

$$X_{ij} = \mu + \alpha_i + \gamma_j + \epsilon_{ij}\ , \qquad i=1,...,k,\ j=1,...,n$$

where the ϵ_{ij} are iid with mean 0 and variance $\sigma^2 < \infty$, the α_i are fixed treatment effects, and the γ_j may be fixed or iid mean 0 random variables corresponding to the fixed effects or mixed model, respectively. (S2)

The F statistic for testing equality of treatment means, H_{02}: $\alpha_i = 0$, i=1,...,k, is

$$F = \frac{\frac{1}{k-1}\sum_{i=1}^{k} n(\bar{X}_{i.} - \bar{X}_{..})^2}{\frac{1}{(k-1)(n-1)}\sum_{i=1}^{k}\sum_{j=1}^{n}(X_{ij} - \bar{X}_{i.} - \bar{X}_{.j} + \bar{X}_{..})^2}\ . \tag{2}$$

2.2.1 Asymptotics Based on the CLT

Results for the statistic (2) under H_{02} are analogous to those for the one-way F statistic and are given in Lemmas 3 and 4 below. The proofs, which are again given in the Appendix, involve more steps than those for Lemmas 1 and 2, but should still be suitable for illustrative purposes in a course on asymptotic theory .

Lemma 3 (# of trts. fixed). Under (S2) and H_{02}, with k fixed, $(k-1)F \xrightarrow{d} \chi^2_{k-1}$ as $n \to \infty$.

Lemma 4 (# of trts. → ∞). Under (S2) and H_{02}, with n fixed, $k^{1/2}\,(F-1) \xrightarrow{d} N(0,2n/(n-1))$ as $k \to \infty$.

2.3 Finite Sample Moment Calculations.

For experiments with random allocation of treatments to experimental units, Fisher (see e.g., Welch, 1937, p. 21) suggested that p-values for the (z-transform of) the F statistics (1) and (2) may be obtained, without assuming normality of the data, from the appropriate permutation or randomization distribution. Pitman (1937) and Welch

(1937) computed the moments with respect to the permutation distribution (and hence conditional on the observed sample values) of $W=F/(n-1+F)$, where W is the "beta" version of the statistic (2). To assess the effect of nonnormality on the use of percentiles based on normal theory, the permutation moments of W were compared with those obtained assuming normality. Pitman, on the basis of the first 4 moments, implied that agreement between the normal theory and permutation percentiles should be reasonable if n and k both exceed 4. To improve robustness, Pitman and Welch also suggested obtaining p-values for W from a beta distribution with parameters calculated so that the first two moments agreed exactly with the conditional permutation moments.

Using the relationship between the Beta($\nu_1/2$, $\nu_2/2$) and F(ν_1, ν_2) distributions, Box and Andersen (1955) adapted the Pitman and Welch approximate beta degrees of freedom for W to obtain an approximation to the permutation distribution of the corresponding F statistic. For the one-way ANOVA, this leads to approximating the permutation distribution of the statistic (1) by an F distribution with numerator and denominator degrees of freedom $d_1(k-1)$ and $d_1k(n-1)$ respectively, where $d_1 = 1 + c_2(N+1)/(N-1)(N-c_2)$, $c_2 = k_4/k_2^2$, and k_2, k_4, are the sample k statistics in Box and Andersen(1955, p. 13). Similarly, the permutation distribution of the statistic (2) is approximated by an $F(d_2(k-1), d_2(k-1)(n-1))$ distribution, with

$$d_2 = 1 + \frac{(nk-n+2)V_2 - 2n}{n(k-1)(n-V_2)}, \qquad V_2 = \frac{1}{n-1} \sum_{j=1}^{n} (s_j^2 - \bar{s}^2)^2 \Big/ (\bar{s}^2)^2 ,$$

where s_j^2 are the within block variances, and $\bar{s}^2 = n^{-1} \sum s_j^2$ (Box and Andersen, 1955, pp. 14-15).

To quantify the effect of nonnormality on the null distributions of (1) and (2), Box and Andersen then approximate the expectation, taken with respect to a given distribution type, of the sample quantities d_1 and d_2, respectively. This leads to a simple characterization of null performance in terms of the kurtosis β_2 of the data, where $\beta_2 = E(X-EX)^4/(\mathrm{Var}\ X)^2$ and, for example, $\beta_2 = 3$ if X has a normal distribution. Thus for data from a distribution with kurtosis β_2, each of the ANOVA F statistics (1) and (2) has approximately an F distribution with the usual degrees of freedom multiplied by a factor d, where

$$d = 1 + \frac{\beta_2 - 3}{N} + O(N^{-2}) . \quad (3)$$

From (3) it is readily seen that for distributions with kurtosis greater than the normal, $\beta_2 > 3$, use of the normal theory F percentiles (without adjustment) will lead to conservative error rates, while for distributions with $\beta_2 < 3$ the usual ANOVA F tests will be liberal. Equation (3) also shows that, in agreement with the asymptotic results in Lemmas 1-4, the ANOVA tests will be approximately correct provided N=kn is large.

3. RANK STATISTICS

3.1 The One-way Layout

For the one-way layout with balanced data, (S1), let R_{ij} be the rank of X_{ij} in the ordered sample $X_{(1)} \le X_{(2)} \le \cdots \le X_{(N)}$. Also let

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{k} n\left(\bar{R}_{i\cdot} - \frac{N+1}{2}\right)^2, \text{ where } \bar{R}_{i\cdot} = \frac{1}{n}\sum_{j=1}^{n} R_{ij},\ i=1,\ldots,k .$$

The Kruskal-Wallis k-sample rank procedure rejects the null hypothesis H_{01} for large values of H (Kruskal and Wallis,1952). For small n and k, critical values are obtained from the exact null (i.e., permutation) distribution of H. For large n, critical values are obtained from the χ^2_{k-1} distribution, since as is well known, $H \xrightarrow{d} \chi^2_{k-1}$ as $n \to \infty$ under H_{01} with k fixed (e.g., Kruskal, 1952).

An alternative procedure, proposed by Wallace (1959), is to compute the F statistic (1) on the ranks R_{ij} and obtain critical values from the F distribution. We write F_R to denote an F statistic on ranks, so for the one-way situation we have

$$F_R = \frac{\sum_{i=1}^{k} n\left(\bar{R}_{i\cdot} - \frac{N+1}{2}\right)^2 \Big/ (k-1)}{\sum_{i=1}^{k}\sum_{j=1}^{n} (R_{ij} - \bar{R}_{i\cdot})^2 \Big/ (N-k)} = \frac{(N-k)H}{(k-1)(N-1-H)} . \quad (4)$$

Obtaining degrees of freedom for F_R will be discussed further in Section 3.1.2, but one possibility is to use the ANOVA degrees of freedom k-1 and $k(n-1)$ (c.f. Wallace, 1959).

Numerical studies (e.g., Iman and Davenport, 1976) have shown that for intermediate values of n, agreement between actual and nominal Type I error rates is better for F_R with $F(k-1,k(n-1))$ percentiles, than for H with χ^2_{k-1} percentiles. The F_R procedure is, however, slightly liberal, while the latter is conservative, particularly for n small and k large. Iman and Davenport give little explanation for the greater accuracy of the ANOVA on ranks procedure, nor do they provide a basis for predicting performance in the large k, small n situation. We will see below that results in Lemmas 5 and 6 for the asymptotic behavior of F_R help to explain the empirical findings and provide justification for this procedure.

3.1.1 Asymptotics Based on the CLT

In stating asymptotic results about the statistic F_R, for simplicity we have restricted attention to continuous data. Note, however, that the more general results in Appendix Lemmas A5 - A8 can be applied to the situation of data with ties as in Conover (1973). In fact an appealing feature of F_R used with mid-ranks (average ranks) is that no additional computations are needed to account for ties. This can be seen more clearly in Section 3.1.2 below where the F distribution with adjusted degrees of freedom is used to approximate the permutation distribution of F_R.

Note also that the finite moment assumption in (S1) is not required for Lemmas 5 and 6, and that the result in Lemma 6 appears to be new.

Lemma 5 (# of trts. fixed). For F_R in (4), given (S1) with X_{ij} continuous and k fixed, $(k-1)F_R \xrightarrow{d} \chi^2_{k-1}$ as $n \to \infty$.

Lemma 6 (# of trts. $\to \infty$). For the statistic F_R in (4), given (S1) with X_{ij} continuous and n fixed, $k^{1/2}(F_R - 1) \xrightarrow{d} N(0,2n/(n-1))$ as $k \to \infty$.

Several comments concerning Lemmas 5 and 6 are worth making. First, the asymptotic distributions of the statistics F in (1) and F_R in (4) are identical. It is not possible, however, to obtain the results in Lemmas 5 and 6 directly from Lemmas 1 and 2 because the R_{ij} are not independent. On the other hand, correlations between the R_{ij} have essentially no impact on the F_R procedure, even for small samples, because the R_{ij} are equicorrelated both between and within treatment groups. (For intuition note that given normal data under H_{01}, the F statistic (1) has exactly the $F(k-1, k(n-1))$ distribution if all N random variables are equicorrelated.) More importantly, equivalence of the asymptotic distributions of the statistics F and F_R justifies the "ANOVA on ranks" procedure if either n or k is large.

Second, Lemma 6 indirectly provides an explanation for why the performance of H with χ^2_{k-1} percentiles becomes increasingly conservative as k increases for n small and fixed (see Monte Carlo results in Iman and Davenport, 1976). Using Lemma 6 and $F_R = (N-k)H/[(k-1)(N-1-H)]$, we can show that, for $k \to \infty$, $H/(k-1)$ is asymptotically normal with mean 1 and variance $2(n-1)/(kn)$. Comparing H to the χ^2_{k-1} distribution corresponds, however, to assuming an asymptotic ($k \to \infty$) normal distribution for $H/(k-1)$ with mean 1 and variance $2/(k-1)$. Since $2(n-1)/(kn) < 2/(k-1)$, when k is large but n is small the χ^2_{k-1} percentiles will be conservative because the corresponding normal approximation has a variance that is too large. Lemma 6, of course, also explains why in the same situation of n small and fixed, and k increasing, accuracy of the F_R procedure increases.

Third, as we show in the Appendix, Lemmas 5 and 6 are easily generalized to rank statistic analogs of F_R which employ not the Wilcoxon scores, but other scores such as, for example, normal scores. The finite sample permutation moment calculations suggest that F_R with normal scores should be more closely approximated by an $F(k-1, k(n-1))$ random variable than is the F_R with Wilcoxon scores.

3.1.2 Finite Sample Moment Calculations.

For X_{ij} satisfying (S1), the null distribution of F_R in (4) is a special case of the permutation distributions considered by Box and Andersen (1955) for the one-way layout. Wallace (1959) pointed out that for F_R on ranks, in the expression for d_1 in Section 2.3, c2 has the value $-6/5$, so that the adjustment to degrees of freedom is exactly $d_1 = 1-(6/5)(N+1)(N-1)^{-1}(N+6/5)^{-1}$. For moderately large N=kn, $d_1 \doteq 1-1.2/N$, and comparison of F_R to percentiles from the $F(k-1,k(n-1))$ distribution will result in a liberal test. As N increases, however, the $F(k-1,k(n-1))$ percentiles will become increasingly accurate. Thus the classical work using moment-based approximations explains, as does the CLT-based asymptotic theory, why the F_R procedure is accurate for large k or large n. In addition, the classical approach provides an explanation of why the F_R procedure is liberal for moderate values of N. Since standard statistical packages typically include a function to compute F probabilities with non-integer degrees of freedom, we recommend F_R be compared to $F(d_1(k-1), d_1k(n-1))$ percentiles with d_1 as above.

3.2 The Two-Way Classification.

We now consider the rank procedure proposed by Friedman (1937) for testing equality of treatment means in the two-way layout given by (S2). For each j, let R_{ij} be the rank of X_{ij} in the ordered set $X_{(i)j} \leq \cdots \leq X_{(k)j}$. That is, the R_{ij} are the ranks of the k observations **within** each block. The Friedman statistic is

$$T = \frac{12n}{k(k+1)} \sum_{i=1}^{k} \left\{ \bar{R}_{i\cdot} - \tfrac{1}{2}(k+1) \right\}^2 .$$

Under H_{02} $T \xrightarrow{d} \chi^2_{k-1}$ as $n \to \infty$ (e.g., Friedman, 1937), and H_{02} is rejected for large values of T, often on the basis of critical values from the χ^2_{k-1} distribution. As tables of the exact null distribution of T are necessarily cumbersome and limited, there have been several suggestions concerning approximations that will produce critical values that are more accurate than the χ^2_{k-1} percentiles, for smaller n. Kendall and

Babington Smith (1939), citing Pitman (1937), thus proposed a beta-transform of T with approximate degrees of freedom (see Section 3.2.2). Iman and Davenport (1980) studied the corresponding F statistic, which is just the statistic (2) applied to the ranks, but did not mention the degrees of freedom approximation based on the beta-transform. Conover (1980, p. 300) also recommends that H_{02} be tested using this (unadjusted) two-way ANOVA on ranks approach, corresponding to comparison of F_R to the $F(k-1,(k-1)(n-1))$ distribution, where

$$F_R = \frac{(n-1)T}{n(k-1)-T} \ . \tag{5}$$

Iman and Davenport (1980) present empirical results comparing the Type I error rates of T with χ^2_{k-1} critical values and the two-way F_R procedure. Patterns seen are similar to those described in Section 3.1 for the analogous H and F_R procedures in the one-way situation. Thus Iman and Davenport remark that "The χ^2 approximation falls off as k increases for fixed b", where b is our n, and that "The F approximation improves as k increases and is liberal but still dominates the χ^2 approximation..." Again the asymptotic theory (see Lemmas 7 and 8) and the classical approach based on moment calculations (Section 3.2.2) explain these comments. Also, Lemma 8 shows why a procedure suggested by Iman and Davenport based on averaging the percentiles of a χ^2 and an F will not be asymptotically correct when $k \to \infty$.

3.2.1 Asymptotics based on the CLT

Similar to the remarks at the beginning of Section 3.1.1, continuous data are again assumed for simplicity, and the finite moment assumption of (S2) is not required for Lemmas 7 and 8.

Lemma 7 (# of trts. fixed). For the statistic F_R in (5), given (S2) with X_{ij} continuous and k fixed, then under H_{02}, $(k-1)F_R \xrightarrow{d} \chi^2_{k-1}$ as $n \to \infty$.

Lemma 8 (# of trts. $\to \infty$). For the statistic F_R in (5), given (S2) with X_{ij} continuous and n fixed, then under H_{02}, $k^{1/2}(F_R - 1) \xrightarrow{d} N(0, 2n/(n-1))$ as $k \to \infty$.

Friedman (1937, pp. 694,695) considered the large k, fixed n situation and conjectured that as $k \to \infty$, T is asymptotically normal with mean $k-1$ and variance $2(n-1)(k-1)/n$. A proof of this conjecture can be obtained from that for F_R in Lemma 8, since from (5) $T = n(k-1)F_R/(n-1+F_R)$.

3.3.2 Finite Sample Moment Calculations.

Pitman (1937) and Welch (1937) both provide moments for the permutation distribution of the beta-transform W of F_R, given by $W=T/(n(k-1))$. Kendall and Babington Smith (1939) suggested using the "Fisher z-transform" of W with degrees of freedom obtained from the Pitman and Welch beta approximation to perform tests of significance. The F statistic analog of this test can be obtained directly from Kendall and Babington Smith or from the Box and Andersen (1955) expression for d_2 given in Section 2.3. Taking the latter approach, note that V_2 in the expression for d_2 is identically 0 for the rank statistic F_R in (5) (assuming no ties). Thus under H_{02}, the distribution of F_R in (5) is approximated by an $F(d_2(k-1), d_2(k-1)(n-1))$ distribution with $d_2 = 1 - 2/[n(k-1)]$. (For data containing ties, the value of V_2 must be calculated to obtain d_2.)

The expression for d_2 (with $V_2 = 0$) shows clearly that for small n and k, the two-way ANOVA on ranks procedure, i.e., F_R with the usual degrees of freedom, will result in a liberal test. As either n or k increases, and hence $n(k-1)$ increases, null performance of this test will improve, as is evident in the results of Iman and Davenport (1980). Whenever possible, we suggest that F_R be compared to the percentiles of an $F(d_2(k-1), d_2(k-1)(n-1))$ distribution with d_2 as above.

4. EXAMPLE

A screening trial was carried out to assess k=35 crepe myrtle cultivars for resistance to aphid infestation. The field layout utilized a randomized complete block design with n=4 blocks. The response recorded for each plant was the sum of the number of aphids on the three most heavily infested leaves and the percent of foliage covered with sooty mold. The resulting data are presented in Table 1.

Inspection of the data suggests that the error distribution is markedly nonnormal, being highly skewed with a long right tail. It is interesting, therefore, to consider how the results in Sections 2 and 3 should affect the choice of procedure for testing equality of the cultivar means. Our results indicate that for k (or n) large, both the usual ANOVA F (possibly after a transformation) and the F on ranks are robust with respect to Type I error performance. This leads to the important conclusion that the choice of procedure should be based on power considerations. The rank procedure is therefore preferred to the usual ANOVA because the former procedure will have better power for long tailed data such as these (see e.g., Table 4.1.7 of Randles and Wolfe, 1979).

As the data in Table 1 contain a large number of ties, we remind the reader that the reason for restricting to continuous data in Lemma 8 was to simplify the statement of results and that the F_R statistic based on mid-ranks automatically adjusts for ties, although the correction factor d_2 does depend on the tie structure. We expect the Friedman statistic (corrected for ties) to be conservative if compared to the χ^2_{34} distribution. Also, the ANOVA on ranks procedure should provide accurate p values since the Box and Andersen adjustment to degrees of freedom (see Section 3.2.2) is $d_2 = 0.985$.

Results for the rank procedures for the data in Table 1 are as follows. The Friedman statistic (corrected for ties) yielded $T = 47.30$ with p-value $p = .064$ based on the χ^2_{34} distribution. The ANOVA on ranks gave $F_R = 1.60$, with $p = .038$ for degrees of freedom 34 and 102, and p=.039 with the adjusted degrees of freedom 33.5 and 100.5. As expected, the p-value is smaller for the F_R procedures than for the Friedman with the χ^2_{34} distribution, though either approach suggests there are differences between the cultivars with respect to their resistance to aphids.

ACKNOWLEDGEMENT

We thank Dr. James Baker (North Carolina State University, Dept. of Entomology) for permission to use the data in the Example of Section 4.

Table 1. Data from a screening trial involving 35 crepe myrtle varieties. Values are a measure of susceptibility to aphid damage and are the sum of (a) the number of aphids on the three most infested leaves and (b) the percentage of foliage covered with sooty mold.

Variety	Block #1	Block #2	Block #3	Block #4
1	1	9	2	4
2	0	14	2	5
3	1	390	0	49
4	8	0	0	0
5	0	0	0	1
6	0	0	0	0
7	0	24	2	4
8	93	0	10	2
9	78	3	0	2
10	5	2	0	0
11	1	180	3	0
12	0	0	2	1
13	21	0	3	47
14	1	3	3	1
15	1	9	140	52
16	12	9	3	0
17	2	3	12	67
18	2	1	0	0
19	9	2	26	3
20	1	2	0	4
21	0	3	3	2
22	0	2	0	1
23	0	29	0	55
24	2	0	11	5
25	0	1	0	0
26	0	0	0	0
27	53	11	0	0
28	0	2	5	33
29	0	0	9	115
30	0	0	0	0
31	93	145	0	0
32	4	405	10	11
33	7	0	3	1
34	3	3	52	0
35	22	0	0	1

APPENDIX - PROOFS

Proof of Lemma 1. The numerator of $(k-1)F$ is $\sum_{i=1}^{k} n(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})^2 = Z_n^T(I_k - k^{-1}1_k 1_k^T)Z_n$, where $Z_n^T = n^{1/2}(\overline{X}_{1\cdot} - \mu, ..., \overline{X}_{k\cdot} - \mu)$, I_k is the k-dimensional identity matrix, and 1_k is a vector of ones. By the central limit theorem (CLT) $Z_n \xrightarrow{d} N(0, \sigma^2 I_k)$ as $n \to \infty$. Since $(I_k - k^{-1}1_k 1_k^T)$ is idempotent with trace $= k-1$, Corollary 1.7 and Theorem 3.5 of Serfling (1980) yield that $\Sigma n(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})^2 \xrightarrow{d} \sigma^2 \chi^2_{k-1}$ as $n \to \infty$. Also $S_p^2 \xrightarrow{p} \sigma^2$ as $n \to \infty$ since each individual $s_i^2 \xrightarrow{p} \sigma^2$ as $n \to \infty$ (Serfling, 1980, Theorem 2.2.3A). Lemma 1 then follows by Slutsky's theorem (Serfling, 1980, p. 19).

Before proceeding to the proof of Lemma 2, we first give two algebraic identities which are quite useful. Let $Y_1, ..., Y_n$ be a sample of size n and let μ and σ^2 be arbitrary constants. Then Algebraic Identity 1 is

$$s^2 - \sigma^2 - \frac{1}{n}\sum_{i=1}^{n}\left[(Y_i - \mu)^2 - \sigma^2\right] = \frac{-2}{n(n-1)}\sum_{i<j}(Y_i - \mu)(Y_j - \mu), \qquad \text{(AI1)}$$

where $s^2 = (n-1)^{-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2$. Algebraic Identity 2 is

$$(\overline{Y} - \mu)^2 - \frac{s^2}{n} = \frac{2}{n(n-1)}\sum_{i<j}(Y_i - \mu)(Y_j - \mu) \,. \qquad \text{(AI2)}$$

When $Y_1, ..., Y_n$ are iid with mean μ and variance σ^2, the right-hand sides of (AI1) and (AI2) both have expectation zero and variance $2\sigma^4/n(n-1)$. (AI1) is commonly used to approximate $s^2 - \sigma^2$ in the proof of asymptotic normality of s^2.

Proof of Lemma 2. The statistic of interest may be written as

$$k^{1/2}(F-1) = k^{1/2}\left[\frac{1}{k-1}\sum_{i=1}^{k}(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})^2 - \frac{\sigma^2}{n} - \frac{1}{k}\sum_{i=1}^{k}\left(\frac{s_i^2}{n} - \frac{\sigma^2}{n}\right)\right] \Big/ (S_p^2/n).$$

Noting that $(k-1)^{-1}\sum_{i=1}^{k}(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})^2$ is an "s^2" in the means $\overline{X}_{i\cdot}$, (AI1) may be used to replace the numerator of $k^{1/2}(F-1)$ by

$$k^{-1/2}\sum_{i=1}^{k}\left[(\overline{X}_{i\cdot} - \mu)^2 - \frac{\sigma^2}{n} - \left(\frac{s_i^2}{n} - \frac{\sigma^2}{n}\right)\right] + \text{Rem} \,,$$

where $\mathrm{Var(Rem)} = 2[\sigma^2/n]^2 / (k-1)$. Using the CLT and Slutsky's theorem, the numerator of $k^{1/2}(F-1) \xrightarrow{d} N(0, \mathrm{Var}[(\overline{X}_{1\cdot} - \mu)^2 - s_1^2/n])$ as $k \to \infty$. But by (AI2) this latter asymptotic variance is $\mathrm{Var}[(\overline{X}_{1\cdot} - \mu)^2 - s_1^2/n] = 2\sigma^4/n(n-1)$. Finally, by the weak law of large numbers $S_p^2/n \xrightarrow{p} \sigma^2/n$ as $k \to \infty$, so that Slutsky's theorem yields $k^{1/2}(F-1) \xrightarrow{d} N(0, 2n/(n-1))$.

Proof of Lemma 3. Using the linear model $X_{ij} = \mu + \alpha_i + \gamma_j + \epsilon_{ij}$ with $\overline{\alpha} = k^{-1} \sum_{i=1}^k \alpha_i = 0$, we have under H_0: $\alpha_1 = \cdots = \alpha_k$ that $\sum_{i=1}^k n(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})^2 = Z_n^T(I_k - k^{-1}1_k1_k^T)Z_n$, where $Z_n = n^{1/2}(\overline{\epsilon}_{1\cdot},\ldots,\overline{\epsilon}_{k\cdot})$ and $\overline{\epsilon}_{i\cdot} = n^{-1} \sum_{j=1}^n \epsilon_{ij}$. Similar to Lemma 1 we get $Z_n \xrightarrow{d} N(0, \sigma^2 I_k)$ and $\sum_{i=1}^k(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})^2 \xrightarrow{d} \sigma^2\chi^2_{k-1}$ as $n \to \infty$. By ANOVA partitioning of sums of squares we may directly verify that

$$\frac{1}{(k-1)(n-1)} \sum_{i=1}^{k} \sum_{j=1}^{n} (X_{ij} - \overline{X}_{i\cdot} - \overline{X}_{\cdot j} + \overline{X}_{\cdot\cdot})^2 \xrightarrow{p} \sigma^2 \quad \text{as} \quad n \to \infty .$$

Slutsky's theorem then gives $(k-1)F \xrightarrow{d} \chi^2_{k-1}$ as $n \to \infty$.

Proof of Lemma 4. Using the linear model representation, we have under H_0: $\alpha_1 = \cdots = \alpha_k$ that

$$k^{1/2}(F-1) = k^{1/2}\left[\frac{1}{k-1} \sum_{i=1}^{k} n(\overline{\epsilon}_{i\cdot} - \overline{\epsilon}_{\cdot\cdot})^2 - D\right] \Big/ D ,$$

where the denominator D is given by

$$D = \frac{1}{(k-1)(n-1)} \sum_{i=1}^{k} \sum_{j=1}^{n} (\epsilon_{ij} - \overline{\epsilon}_{i\cdot} - \overline{\epsilon}_{\cdot j} + \overline{\epsilon}_{\cdot\cdot})^2 .$$

Using (AI1) with $\mu = 0$ and $\sigma^2 = 0$, we replace $(k-1)^{-1} \sum_{i=1}^k n(\overline{\epsilon}_{i\cdot} - \overline{\epsilon}_{\cdot\cdot})^2$ by $k^{-1}\sum_{i=1}^k \overline{\epsilon}_{i\cdot}^2$ and D by $[k(n-1)]^{-1}\sum_{i=1}^k(\epsilon_{ij} - \overline{\epsilon}_{i\cdot})^2$ to write the numerator of $k^{1/2}(F-1)$ as

$$k^{-1/2} \sum_{i=1}^{k} \left[n\overline{\epsilon}_{i\cdot}^2 - \frac{1}{n-1} \sum_{j=1}^{n} (\epsilon_{ij} - \overline{\epsilon}_{i\cdot})^2\right] + \mathrm{Rem} ,$$

where $\mathrm{Rem} \xrightarrow{p} 0$ as $k \to \infty$. By (AI2) the random variables inside the brackets of this last expression are iid with the form

$$\frac{2}{n-1}\sum_{j<k} \epsilon_{ij}\epsilon_{ik} .$$

Thus by the CLT and Slutsky's theorem the numerator of $k^{1/2}(F-1) \xrightarrow{d}$ $N(0,2n\sigma^2/(n-1))$ as $k \to \infty$. As for $n \to \infty$ in the proof of Lemma 3, we have $D \xrightarrow{p} \sigma^2$ as $k \to \infty$, and thus $k^{1/2}(F-1) \xrightarrow{d} N(0, 2n/(n-1))$.

For Lemmas 5 and 6 on rank statistics in the one-way setup, let $\phi_{ij} = \phi(R_{ij}/(N+1))$, where R_{ij} is the rank of X_{ij} in the combined sample and ϕ is a score function. We shall assume that ϕ is the difference of two increasing functions and call such functions "square integrable score functions" if $\int_0^1 \phi^2(u)du < \infty$ as in Randles and Wolfe (1980, p. 272). The most popular scores are "Wilcoxon scores" given by $\phi(u) = u$ and "normal scores" $\phi(u) = \Phi^{-1}(u)$, where $\Phi(u)$ is the standard normal distribution function. Let F_ϕ be the F statistic in (1) with ϕ_{ij} replacing X_{ij} and note that $F_R = F_\phi$ when $\phi(u) = u$. The following lemma is more general than Lemma 5.

Lemma A5. If ϕ is a square integrable score function and the conditions of Lemma 5 hold, then $(k-1)F_\phi \xrightarrow{d} \chi^2_{k-1}$ as $n \to \infty$.

Proof. Let U_{ij}, $i = 1,...,k$, $j = 1,...,n$, be $N = kn$ independent uniform (0,1) random variables and set $\phi^*_{ij} = \phi(U_{ij})$. Following the method of approximation found in Randles and Wolfe (1980, p. 285), we may replace $\phi_{ij} - \overline{\phi}_{..}$ by $\phi^*_{ij} - \overline{\phi}^*_{..}$ and then follow the proof of Lemma 1.

The following general version of Lemma 6 with $k \to \infty$ is harder to prove than Lemma A5 because simple approximation of the numerator of F_ϕ as in Lemma A5 is not sufficient.

Lemma A6. If ϕ and ϕ^2 are square integrable score functions and the conditions of Lemma 6 hold, then $k^{1/2}(F_\phi - 1) \xrightarrow{d} N(0, 2n/(n-1))$ as $k \to \infty$.

Proof.

$$k^{1/2}(F_\phi - 1) = \frac{k^{1/2}\left[\frac{1}{k-1}\sum_{i=1}^{k} n(\overline{\phi}_{i.} - \overline{\phi}_{..})^2 - \frac{1}{k}\sum_{i=1}^{k} s_{\phi i}^2\right]}{\frac{1}{k}\sum_{i=1}^{k} s_{\phi i}^2},$$

where $s_{\phi i}^2 = (n-1)^{-1}\sum_{j=1}^{n}[\phi_{ij} - \overline{\phi}_{i.}]^2$. Since $k^{-1}\sum_{i=i}^{k} s_{\phi i}^2 \xrightarrow{p} \sigma_\phi^2$ as $k \to \infty$, where $\sigma_\phi^2 = \int_0^1 (t - \mu_\phi)^2 dt$ with $\mu_\phi = \int_0^1 \phi(t)dt$, we can concentrate on the numerator of the last display. Replacing $(k-1)^{-1}\sum_{i=1}^{k}(\overline{\phi}_{i.} - \overline{\phi}_{..})^2$ by $k^{-1}\sum_{i=1}^{k}(\overline{\phi}_{i.} - \overline{\phi}_{..})^2$ leads to

$$\begin{aligned} A &= k^{1/2}\left[\frac{1}{k}\sum_{i=1}^{k}\left(n(\overline{\phi}_{i.} - \overline{\phi}_{..})^2 - s_{\phi i}^2\right)\right] \\ &= k^{1/2}\left[\frac{1}{k}\sum_{i=1}^{k}\left(\frac{2}{n-1}\sum_{j<\ell}(\phi_{ij} - \overline{\phi}_{..})(\phi_{i\ell} - \overline{\phi}_{..})\right)\right], \end{aligned}$$

where this last step uses (AI2). Now define A^* to be A with ϕ_{ij} replaced by ϕ_{ij}^* as in Lemma A5. Using the method of Randles and Wolfe (1980, p. 285) along with $\int \phi^4(t)dt < \infty$ and quite a bit of algebra yields $E(A - A^*)^2 \xrightarrow{p} 0$ as $k \to \infty$. Finally, we can replace $\overline{\phi}_{..}^*$ in A^* by μ_ϕ and obtain $A^* \xrightarrow{d} N(0, 2\sigma_\phi^4 n/(n-1))$ as $k \to \infty$ by the CLT. The conclusion of Lemma A6 then follows by various applications of Slutsky's theorem.

For Lemmas 7 and 8 we give more general versions using notation similar to that of Lemmas A5 and A6. Here, though, R_{ij} is just the rank of X_{ij} within the jth block. Let F_ϕ be the statistic F in (2) with X_{ij} replaced by $\phi_{ij} = \phi(R_{ij}/(k+1))$.

Lemma A7. If ϕ is a square integrable score function and the conditions of Lemma 7 hold, then $(k-1)F_\phi \xrightarrow{d} \chi_{k-1}^2$ as $n \to \infty$.

Proof. The denominator of $(k-1)F_\phi$ can easily be shown to converge in probability to $\sigma_k^2 k(k-1)^{-1}$ as $n \to \infty$, where $\sigma_k^2 = k^{-1}\sum_{i=1}^{k}[\phi(i/(k+1)) - \overline{\phi}_{..}]^2$. The numerator of $(k-1)F_\phi$ is $Z_n^T Z_n$, where $Z_n = n^{1/2}(\overline{\phi}_{1\cdot} - \overline{\phi}_{\cdot\cdot},\ldots,\overline{\phi}_{k\cdot} - \overline{\phi}_{\cdot\cdot})$. Since Z_n is an average of iid vectors with mean zero and covariance matrix $\Sigma_k = \sigma_k^2 k(k-1)^{-1}[I_k - k^{-1}1_k 1_k^T]$, $Z_n \xrightarrow{d} N(0,\Sigma_k)$ by the CLT and $Z_n^T Z_n \xrightarrow{d} \sigma_k^2 k(k-1)^{-1}\chi^2_{k-1}$ as $n \to \infty$. Slutsky's theorem then gives the result.

Lemma A8. If ϕ is a square integrable score function and the conditions of Lemma 8 hold, then $k^{1/2}(F_\phi - 1) \xrightarrow{d} N(0, 2n/(n-1))$ as $k \to \infty$.

Proof. The proof is very similar to that of Lemma 4 and parts of Lemma A6. One can show that the denominator of $k^{1/2}(F_\phi - 1)$ converges in probability to σ_ϕ^2 and that the numerator is exactly

$$A = k^{1/2}\left[\frac{1}{k-1}\sum_{i=1}^{k}\left(\frac{2}{n-1}\sum_{j<\ell}(\phi_{ij} - \overline{\phi}_{i\cdot})(\phi_{i\ell} - \overline{\phi}_{i\cdot})\right)\right].$$

Similar to the proof of Lemma A6 we define A^* to be A with $\phi(U_{ij}) - \mu_\phi$ in place of $\phi_{ij} - \overline{\phi}_{i\cdot}$. Showing $E(A - A^*)^2 \to 0$ is simpler here than in Lemma A6 because of the independence between blocks and requires only $\int\phi^2(u)du < \infty$ compared to the $\int\phi^4(u)du < \infty$ used in Lemma A6. Slutsky's theorem and the CLT complete the proof.

REFERENCES

Box, G. E. P., and Andersen, S. L. (1955), "Permutation Theory in the Derivation of Robust Criteria and the Study of Departures from Assumption," *Journal of the Royal Statistical Society, Ser. B*, 17, 1-26.

Conover, W. J. (1973), "Rank Tests for One Sample, Two Samples, and k Samples without the Assumption of a Continuous Distribution Function," *Annals of Statistics*, 1, 1105-1125.

Conover, W. J. (1980), *Practical Nonparametric Statistics*, New York, John Wiley.

Friedman, Milton (1937), "The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance," *Journal of the American Statistical Association*, 217, 929-932.

Iman, R. L., and Davenport, J. M. (1976), "New Approximations to the Exact Distribution of the Kruskal-Wallis Test Statistic," *Communications in Statistics - Theory and Methods,* A5, 1335-1348.

Iman, R. L., and Davenport, J. M. (1976), "Approximations of the Critical Region of the Friedman Statistic," *Communications in Statistics - Theory and Methods,* A9, 571-595.

Kendall, M. G., and Babington Smith, B. (1939), "The Problem of m Rankings," *Annals of Mathematical Statistics*, 10, 275-287.

Kruskal, W. H. (1952), "A Nonparametric Test for the Several Sample Problem," *Annals of Mathematical Statistics*, 23, "525-540.

Kruskal, W. H., and Wallis, W. A. (1952), "Use of Ranks in One-Criterion Variance Analysis," *Journal of the American Statistical Association*, 47, 583-621.

Pirie, W. R. (1974), "Comparing Rank Tests for Ordered Alternatives in Randomized Blocks," *Annals of Statistics*, 2, 374-382.

Pitman, E. J. G. (1937), "Significance Tests Which May be Applied to Samples from any Populations: III. The Analysis of Variance Test," *Biometrika*, 28, 322-621.

Randles, R. H., and Wolfe, D. A. (1979), *Introduction to the Theory of Nonparametric Statistics*, New York, John Wiley.

Santner, T. J., and Duffy, D. E. (1989), *The Statistical Analysis of Discrete Data*, New York: Springer-Verlag.

Scheffe, Henry (1959), *The Analysis of Variance*, New York, John Wiley.

Serfling, R. J. (1980), *Approximation Theorems of Mathematical Statistics*, New York: John Wiley.

Wallace, D. L. (1959), "Simplified Beta-Approximations to the Kruskal-Wallis H Test," *Journal of the American Statistical Association*, 54, 225-230.

Welch, B. L. (1937), "On the z-Test in Randomized Blocks and Latin Squares," *Biometrika*, 29, 21-52.

Zar, J. H. (1984), *Biostatistical Analysis*, New Jersey, Prentice-Hall.