Statistica Sinica 4(1994), 423-443

BOOTSTRAP ESTIMATES OF THE POWER OF A RANK TEST IN A RANDOMIZED BLOCK DESIGN

Denis Larocque and Christian Léger

Université de Montréal

Abstract: In this article, we use the bootstrap to estimate the power of a distribution-free rank test introduced by Mack and Skillings (1980) for the hypothesis of no treatment effect in a randomized block design, or a two-way analysis of variance model with no interaction. Since ties affect the distribution of rank tests and because the sample size in each cell is usually too small, the resampling must be done from a smoothed version of the empirical distribution function of residuals. The theory will show that the type of smoothing is crucial to attain asymptotic consistency under local alternatives. A small sample simulation shows that a particular implementation of the bootstrap does well.

Key words and phrases: Kernel estimates, sample size determination, testing in two-way ANOVA, rank test, smooth bootstrap, Friedman test, power estimation.

1. Introduction

Many testing procedures based on ranks have been proposed in the literature since the 1930's, with perhaps the Mann-Whitney-Wilcoxon test as the most well-known example. These tests are appealing because they are distribution-free under the null hypothesis, simple to compute, and often have good efficiency compared to the classical testing procedures, even though they do not require strong parametric assumptions.

The power function is often needed in practice. It allows the practitioner to find out the probability of rejection for a given alternative hypothesis of interest, key information when the null hypothesis is not rejected. More importantly, an estimate of the power function, based on a pilot study, can be used to decide what sample size should be used to guarantee a certain power for a given alternative. Unfortunately, just as in classical tests, the power function of rank tests *does* depend on the distribution of the observations. The bootstrap provides a nonparametric estimate of the power function. The resulting combination of a distribution-free test with high asymptotic efficiency and a nonparametric bootstrap procedure to estimate its power should prove very attractive from a practical point of view.

We consider a Friedman-type rank test introduced by Mack and Skillings (1980) for the hypothesis of no treatment effect in a randomized block design. The test can also be used in a two-way analysis of variance *without* interaction. From a practical as well as theoretical point of view, this combination of model and test provides an interesting example for the proper use of the bootstrap in testing. Whereas in the two-sample problem we can usually resample with replacement within each sample, in a two-way ANOVA, the sample size in each cell is usually too small. Instead, as in regression, bootstrap observations are constructed from resampled residuals. Second, since ties adversely affect the distribution of rank tests, the empirical distribution of the residuals must be smoothed. This adds a level of difficulty to the asymptotic theory.

Efron (1979) introduced the bootstrap. Beran (1986) showed how the bootstrap could be used to estimate the critical value of a test as well as how to estimate the power of a test under local alternatives. He proved very general asymptotic results that do not apply directly here. Many others have used the bootstrap to estimate critical values in complex situations and showed asymptotic consistency, most notably Romano (1988), Beran and Millar (1989), Chen and Loh (1991) and Arcones and Giné (1991). The bootstrap has also been used to construct test statistics with smaller level error, e.g., Beran (1988), something that we need not do here since the test is distribution-free. Collings and Hamilton (1988) have used the bootstrap to estimate the power of the two-sample Wilcoxon test, while Hamilton and Collings (1991) used it to estimate the sample size necessary to achieve a given power. Their studies are only empirical. In this paper, we study the two-way ANOVA model without interaction. Fisher and Hall (1990) have also considered that model, but they use the bootstrap to estimate the critical point of F-like tests for which no smoothing is required.

The model, the test, and the bootstrap procedures are introduced in the next section. In Section 3, we demonstrate that, under local alternatives, the bootstrap estimate of the power based on resampling from a kernel estimate computed from the least squares residuals is asymptotically valid. We also discuss a number of generalizations, including other types of residuals, and resampling separately in each block. Section 4 contains the results of a simulation which demonstrate that the procedure works well; and concluding remarks are presented in Section 5. The proofs are deferred to an appendix.

2. Test and the Bootstrap Methodology

Consider the model for a two-way analysis of variance with no interaction given by

$$X_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk}, \quad i = 1, \ldots, I, \ j = 1, \ldots, J, \ k = 1, \ldots, n_{ij} \geq 1, \quad (1)$$

where

$$\sum_{i=1}^{I} \alpha_i = \sum_{j=1}^{J} \beta_j = 0, \quad (2)$$

μ is the global mean, the α_i's and the β_j's are the (fixed) treatment effects of factors A (lines) and B (rows), respectively, and the ϵ_{ijk}'s are independent random variables having the same continuous distribution function F with mean 0, and finite variance σ^2. Also let $N = \sum_{i=1}^{I} \sum_{j=1}^{J} n_{ij}$. While the rank test does not require the existence of moments, our bootstrap procedure requires that we use residuals to estimate F, hence our need for the first two moments. This model also describes a randomized block design where factor A is the blocking factor and factor B is the treatment. The hypothesis of interest is $H_0 : \beta_1 = \cdots = \beta_J = 0$, i.e., no treatment effect for the second factor.

Mack and Skillings (1980) introduced a generalization of the Friedman (1937) rank test for a two-way ANOVA with one observation per cell. Even though the test can be described for general n_{ij}, we restrict ourselves to the case of proportional frequencies, i.e., $n_{ij} = n_{i\cdot}n_{\cdot j}/N$, for all i and j where $n_{i\cdot} = \sum_{j=1}^{J} n_{ij}$ and $n_{\cdot j} = \sum_{i=1}^{I} n_{ij}$. In this case, the test statistic has a simple closed form formula. The test is as follows. Within each level of factor A, rank the observations from smallest to largest. Let R_{ij} denote the sum of the ranks in the cell ij. For each level of factor B, let $R_j = \sum_{i=1}^{I} R_{ij}/n_{ij}$. (Mack and Skillings (1980) used R_j^* instead—which could be confusing in a bootstrap paper—and reserved R_j for a different purpose.) The test statistic is

$$T_N = \frac{12}{N(N+I)} \sum_{j=1}^{J} n_{\cdot j} \left(R_j - \frac{N+I}{2} \right)^2, \quad (3)$$

and H_0 is rejected for large values of T_N.

The test is distribution-free under the null hypothesis. Moreover, Mack and Skillings (1980) have shown that it has the same asymptotic relative efficiency relative to the classical F-test as the Mann-Whitney-Wilcoxon test relative to the two-sample t-test. Hence, the asymptotic efficiency of T_N relative to the F-test is .955 for the Gaussian distribution, at least .864 for all continuous distributions, and it can be larger than or equal to one, as for instance for the uniform (1) and double exponential (1.5) distributions. They have computed selected critical values from the null distribution of T_N for some balanced designs. For other designs, the null distribution can be computed approximately via a simulation using any continuous distribution F (since it is distribution-free) or, for large values of n_{ij} and fixed values of I and J, approximated by a chi-squared distribution with $J-1$ degrees of freedom.

Mack and Skillings (1980) have also obtained the asymptotic distribution of T_N under local alternatives. Let $n_{ij} = Np_{ij}$ for all i and j and suppose that we have proportional frequencies. Consider the sequence of alternative hypotheses given by

$$H_{1N} : \beta_j = \frac{\theta_j}{\sqrt{N}}, \qquad j = 1, \ldots, J, \tag{4}$$

so that

$$X_{ijk} = \mu + \alpha_i + \theta_j/N^{1/2} + \epsilon_{ijk}, \quad i = 1, \ldots, I, \ j = 1, \ldots, J, \ k = 1, \ldots, n_{ij}. \tag{5}$$

The asymptotic distribution of T_N under the alternatives H_{1N} is a noncentral chi-squared distribution with $J-1$ degrees of freedom and noncentrality parameter

$$\lambda_T = 12 \left(\int_{-\infty}^{\infty} f^2(x)dx \right)^2 \sum_{j=1}^{J} p_{\cdot j} \left[\sum_{k=1}^{J} p_{\cdot k}(\theta_j - \theta_k) \right]^2, \tag{6}$$

where $f(x)$ is the density of the errors ϵ_{ijk}.

To estimate the power function using a bootstrap approach, we must generate bootstrap errors whose distribution will approach that of the errors, i.e., F, and combine them to form bootstrap random variables which satisfy the alternative hypothesis. So, let $\hat{\epsilon}$ be the vector of least squares residuals obtained from Model (1) and let $\hat{F}_N$ be its empirical distribution function. We cannot generate bootstrap errors from $\hat{F}_N$ as this would lead to ties which create serious distributional problems with rank tests. So we must smooth $\hat{F}_N$. We consider the class of kernel estimates. Let K be a distribution function symmetric about 0 and let

$$\hat{F}_{\lambda_N}(x) = \frac{1}{N} \sum K \left(\frac{x - \hat{\epsilon}_{ijk}}{\lambda_N} \right), \tag{7}$$

where the sum is taken over all indices.

The bootstrap estimate of the power of T_N is computed as follows. Let $\kappa(F; b, \alpha, N)$ be the power of T_N when the alternative hypothesis is that $\beta_j = b_j$, for $j = 1, \ldots, J$, the distribution of the errors is F, the level of the test is α and the total number of observations is N. For simplicity, we have dropped the dependence on the n_{ij}'s. On the other hand, since T_N ranks the observations independently row by row, the power does not depend on the global mean μ or the row effects α_i. The bootstrap estimate of $\kappa(F; b, \alpha, N)$ is $\kappa(\hat{F}_{\lambda_N}; b, \alpha, N)$ and is computed using the following Monte Carlo simulation. Let the N $\hat{\epsilon}^*_{ijk}$ be distributed independently and identically according to $\hat{F}_{\lambda_N}$. Then the bootstrap observations are

$$X^*_{ijk} = b_j + \epsilon^*_{ijk}, \quad i = 1, \ldots, I,\ j = 1, \ldots, J,\ k = 1, \ldots, n_{ij}. \tag{8}$$

Due to ranking done independently in each row, there is no need to estimate the global mean μ or the row effects α_i and, without loss of generality, they are set at 0. Using the X^*_{ijk}'s, T^*_N is computed. Repeating this process a large number of times, $\kappa(\hat{F}_{\lambda_N}; b, \alpha, N)$ is approximated by the proportion of T^*_N's larger than the critical value of the test which is taken from tables or approximated by the α-quantile of a chi-squared distribution with $J - 1$ degrees of freedom.

If more than one alternative hypothesis is of interest, one can estimate the power by using the same bootstrap errors and compute the different T^*_N's in parallel.

Another possible use of the bootstrap observations is to use pilot data to estimate the sample size required to attain a given power for a fixed set of alternatives b_j, $j = 1, \ldots, J$. Hamilton and Collings (1991) have studied this problem for the Wilcoxon two-sample test. The problem is as follows. Let p be the power that we want to attain for the alternatives b_j and let $m_p = IJK_p$ be the smallest total sample size required to attain that power, where, for simplicity, we use K_p observations in each of the IJ cells. The bootstrap estimate of m_p is found iteratively by computing $\kappa(\hat{F}_{\lambda_N}; b, \alpha, m)$ for increasing values of m, that is for increasing values of K, until the estimate of power is larger than p, i.e., $\min\{m : \kappa(\hat{F}_{\lambda_N}; b, \alpha, m) > p\}$. Note that we are estimating the power for a total sample size m different from the total sample size N of the pilot data. In fact, the estimated sample size is usually larger than the sample size of the pilot data set. In computing $\kappa(\hat{F}_{\lambda_N}; b, \alpha, m)$, one generates m rather than N ϵ^*_{ijk} and constructs bootstrap observations following Equation (8) (with $n_{ij} \equiv K$). The larger m is, as compared to N, the less reliable is the estimate.

3. Asymptotic Consistency

In this section, we show that the bootstrap estimate of the power of the Friedman-type test T_N converges to the true power under local alternatives when bootstrapping from $\hat{F}_{\lambda_N}$.

We make the following assumptions on the distribution function of the observations F and on the kernel k, the derivative of K used in defining $\hat{F}_{\lambda_N}$ in (7). These conditions are satisfied by many kernels k, such as, for instance, the normal kernel. First, assume that

(F1) f has mean 0 and a finite variance, (F2) f is bounded and absolutely continuous with finite Fisher information, i.e., $\int_{-\infty}^{\infty}(f'(e)/f(e))^2 f(e)de < \infty$,

(F3) f' is uniformly continuous and bounded.

Also, assume that

(K1) $\int_{-\infty}^{\infty} k(x)dx = 1, \qquad \int_{-\infty}^{\infty} xk(x)dx = 0, \qquad k_2 = \int_{-\infty}^{\infty} x^2 k(x)dx < \infty$,

(K2) k is symmetric about 0,

(a) $\int_{-\infty}^{\infty} |k^{(j)}(x)|dx < \infty$ and $k^{(j)}(x) \to 0$ as $|x| \to \infty$, $j = 0, 1$,

(b) k' is uniformly continuous (with modulus of continuity $w_{k'}$) and of bounded variation,

(K3) (c) $\int_{-\infty}^{\infty} |x \log |x||^{1/2} |dk'(x)| < \infty$,

(d) setting $\gamma(u) = \{w_{k'}(u)\}^{1/2}$, $\int_0^1 \{\log(1/u)\}^{1/2} d\gamma(u) < \infty$,

(e) the Fourier transform of k is not identically 1 in any neighborhood of 0.

Let $G_N(x, F)$ be the distribution function, evaluated at x, of the statistic T_N under the local alternatives Model (5) when the distribution of the errors is F. When resampling from $\hat{F}_{\lambda_N}$, the bootstrap distribution function of T_N^* is $G_N(x, \hat{F}_{\lambda_N})$. Our main result is that the difference between these two distributions converges to 0 in probability.

Theorem 1. *Suppose that F satisfies conditions* (F1) − (F3), *and that the kernel satisfies conditions* (K1) − (K3). *If $\lambda_N \to 0$, $N^{-1}\lambda_N^{-3}\log(1/\lambda_N) \to 0$, and $N^{1/2}\lambda_N^2/\log(N) \to 0$, then*

$$\sup_x |G_N(x, \hat{F}_{\lambda_N}) - G_N(x, F)| \to 0, \qquad \textit{in probability.} \tag{9}$$

Note that the conditions on λ_N are satisfied if $\lambda_N = \lambda N^{-r}$ with $\lambda > 0$ and $1/4 \leq r < 1/3$. The proof is in the appendix.

Remark 1. The conditions on f and k could probably be weakened. We assume the existence of the first two moments because we use least-squares residuals. To show that the derivative of $\hat{f}_{\lambda_N}$ is uniformly bounded, we have imposed conditions (F3) and (K3) so that $\hat{f}'_{\lambda_N}$ converges uniformly to f', a much stronger condition. Condition (K2) on the symmetry of the kernel is only there for convenience.

Remark 2. In a randomized block design, the hypothesis that the distribution of the errors F is the same for all blocks may not be valid. For instance, the variance may differ from block to block. Provided that we have enough observations in each row, we could construct separate kernel estimates of F_i, the distribution of the errors in row i, from the least squares residuals of the one-way ANOVA in each row. The bootstrap resampling would then be done row by row. The asymptotic distribution of T_N would still be chi-squared with $J-1$ degrees of freedom, but the noncentrality parameter would become $\lambda_T = 12(\Sigma_{i=1}^{I} p_{i\cdot} \int_{-\infty}^{\infty} f_i^2(x)dx)^2 \Sigma_{j=1}^{J} p_{\cdot j}[\Sigma_{k=1}^{J} p_{\cdot k}(\theta_j - \theta_k)]^2$, where f_i is the density of F_i. Theorem 1 remains valid so that the bootstrap would be asymptotically justified.

Remark 3. To estimate F, the distribution of the errors, we have used the least squares residuals. This requires the existence of the first two moments, an assumption that we may not want to make, especially when using a distribution-free test. Our result remains valid for many other types of residuals. In particular, our proof relies on Theorem 4.6.2 of Shorack and Wellner (1986) and this theorem is valid for residuals of many types of M-estimates.

Remark 4. The classical test of Friedman (1937) has only one observation per cell and the chi-squared asymptotic distribution is obtained by letting I, the number of blocks, tend to infinity. Other generalizations, such as the tests based on weighted rankings introduced by Quade (1979) and studied by Tardif (1987), also rely on an increasing number of blocks with a fixed number of observations (≥ 1) for each treatment within a block. Our bootstrap approach would not work in this context as the ratio of the number of parameters to the number of observations would not converge to 0, as is required for the appropriate convergence of the empirical distribution function of the residuals to F, see e.g., Bickel and Freedman (1983). This is why we have fixed the number of blocks and assumed instead that the number of observations in each cell tends to infinity. If the number of blocks and the number of observations in each cell both increase at an appropriate rate, the results may remain valid, but our method of proof, which has relied on the assumption of a fixed number of cells, would have to be modified.

Remark 5. When we estimate the sample size necessary to attain a given power, we estimate the power for a total sample size different from that of our sample. Let N be the sample size of the pilot study and let m be the sample size for which we want to estimate the power. Then $\hat{F}_m = \hat{F}_{\lambda_N}$, i.e., the estimate of F is still based on N observations and not m. Therefore $\{\hat{F}_m\}$ must satisfy conditions (10), (11), and (12), given in the appendix. Only Condition (10) imposes restrictions on the relationship between m and N. By looking at the proof of Theorem 1, we see that if $\sqrt{m/N}/\log(m) \to 0$, $N^{-1}\lambda_N^{-3}\log(1/\lambda_N) \to 0$, and $\sqrt{N}\lambda_N^2 = O(1)$ then the bootstrap estimate of the power under local alternatives remains valid. For instance that condition remains valid if $N = cm$ where c is a constant, or if $N = m/\log(m)$; but the condition is not satisfied if $N = m^p$, for $0 < p < 1$. Therefore we can estimate the power under local alternatives for a sample size

m larger than N, the sample size of our experiment, but m cannot be "much" larger than N.

4. Simulation Study

In this section, we give some empirical evidence that the bootstrap estimate of power works well by evaluating its performance via a simulation. We begin by a description of the simulation study.

We used a complete factorial design with four factors: the number of blocks $I = 5, 10$, the number of treatments $J = 3, 5$, an equal number of observations in each cell of $n_{ij} \equiv K = 1, 3, 5$, and three distributions for the errors. The distributions were the standard normal, the double exponential with scale parameter $1/\sqrt{2}$ and the uniform on the interval $(-1.732, 1.732)$ so that each distribution has mean 0 and variance 1. The tests were performed at the 5% level. The treatment effects for the alternative hypotheses were equally spaced and symmetric about 0; for instance, when $J = 5$, they are of the form $\beta_1 = -2c$, $\beta_2 = -c$, $\beta_3 = 0$, $\beta_4 = c$ and $\beta_5 = 2c$ where c is a positive constant. For each of the 36 combinations, seven sets of alternatives were selected by choosing values of c so that the true power of the test is close to 0.05 ($c = 0$), 0.10, 0.25, 0.50, 0.75, 0.90 and 0.95. To find these constants, a preliminary simulation of size 40,000 was carried out. The estimated power for this pilot simulation will be referred to as the "true power". Then, for each combination, the bootstrap procedure was repeated 500 times with 1,000 bootstrap samples. Each sample of bootstrap residuals was used to compute bootstrap observations for each of the seven alternatives. We used a normal kernel with a bandwidth of $\lambda_N = N^{-.3}$. The computations were performed on a Sun Sparc station 2 using FORTRAN 77 and functions in the NAG library.

For each of the 36 combinations, we computed the mean, the median, the standard error, and the 5th and 95th percentiles of the 500 estimations for each of the seven alternatives. We analyzed the results over the seven alternatives case by case, and also over the 36 combinations for each of the seven alternatives. The median bias will not be discussed here as it is in good agreement with the mean bias. We begin with the analysis for each fixed alternative.

Table 1. Ratio of the number of parameters to the number of observations

I	J	K	ratio	I	J	K	ratio	I	J	K	ratio
5	3	1	0.467	5	5	3	0.120	10	3	5	0.080
5	3	3	0.156	5	5	5	0.072	10	5	1	0.280
5	3	5	0.093	10	3	1	0.400	10	5	3	0.093
5	5	1	0.360	10	3	3	0.133	10	5	5	0.056

In Remark 4, we noted that if the ratio of the number of parameters to the number of observations does not tend to zero, the bootstrap may not work. We have found that using this ratio as a dependent variable explains most of the variability in the mean bias and in the standard error of the 36 cases that we have considered. Table 1 provides the value of the ratio for the 12 combinations of number of blocks, treatments, and cell sample size. Figures 1 and 2 show plots of the mean bias and of the standard error versus the ratio of parameters to observations when the true power is 50% and 90% respectively. Each of these four plots contains 36 points corresponding to the 36 cases studied. The plotting character identifies the distribution: N for normal, D for double exponential and U for uniform. The same plots when the true power is 10%, 25%, 75% and 95% are not included as each of them is similar to one of the two presented. When the true power is 5% we simply have noise since the test is distribution-free under H_0. Looking at these two figures, we find that, as expected, the performance of the bootstrap procedure gets better as the ratio decreases; except perhaps for the mean bias of the double exponential distribution which does not improve in absolute value. When the true power is 50% (Figure 1), the fit for each distribution is almost linear both for the mean bias and the standard error. Looking at the upper plots, the mean bias curve of the double exponential distribution is always below the other two which are similar. The plots of standard error show that the uniform distribution has the smallest standard errors when the ratio is small. When the ratio is large, no distribution dominates. In summary, for the designs and distributions considered here, the mean bias (in absolute value) is less than 0.04 and the standard error less than 0.10 everywhere along the power curve when the ratio of parameters to observations is below 1/10.

We also analyzed the results case by case by plotting the mean, and the 5th and 95th percentiles of the difference between the estimated power and the true power for each of the seven alternatives versus the true power. Figures 3 and 4 show four typical examples. These figures were inspired by Figures 5 through 10 of Collings and Hamilton (1988). In Figures 3 and 4, the distributions are normal and double exponential, respectively, while the results for the uniform are usually similar to those of the normal. In both cases, the upper and lower plots correspond to designs with a small and a large ratio, respectively. The two upper plots of Figures 3 and 4, with a design of 75 observations and 7 parameters, compare well with Figures 7, 8 and 9 of Collings and Hamilton (1988) which correspond to 40 observations and 1 parameter (the shift parameter). These plots give a very good idea of the general pattern when the ratio varies and the procedure seems reasonably reliable when the ratio is below 1/10.

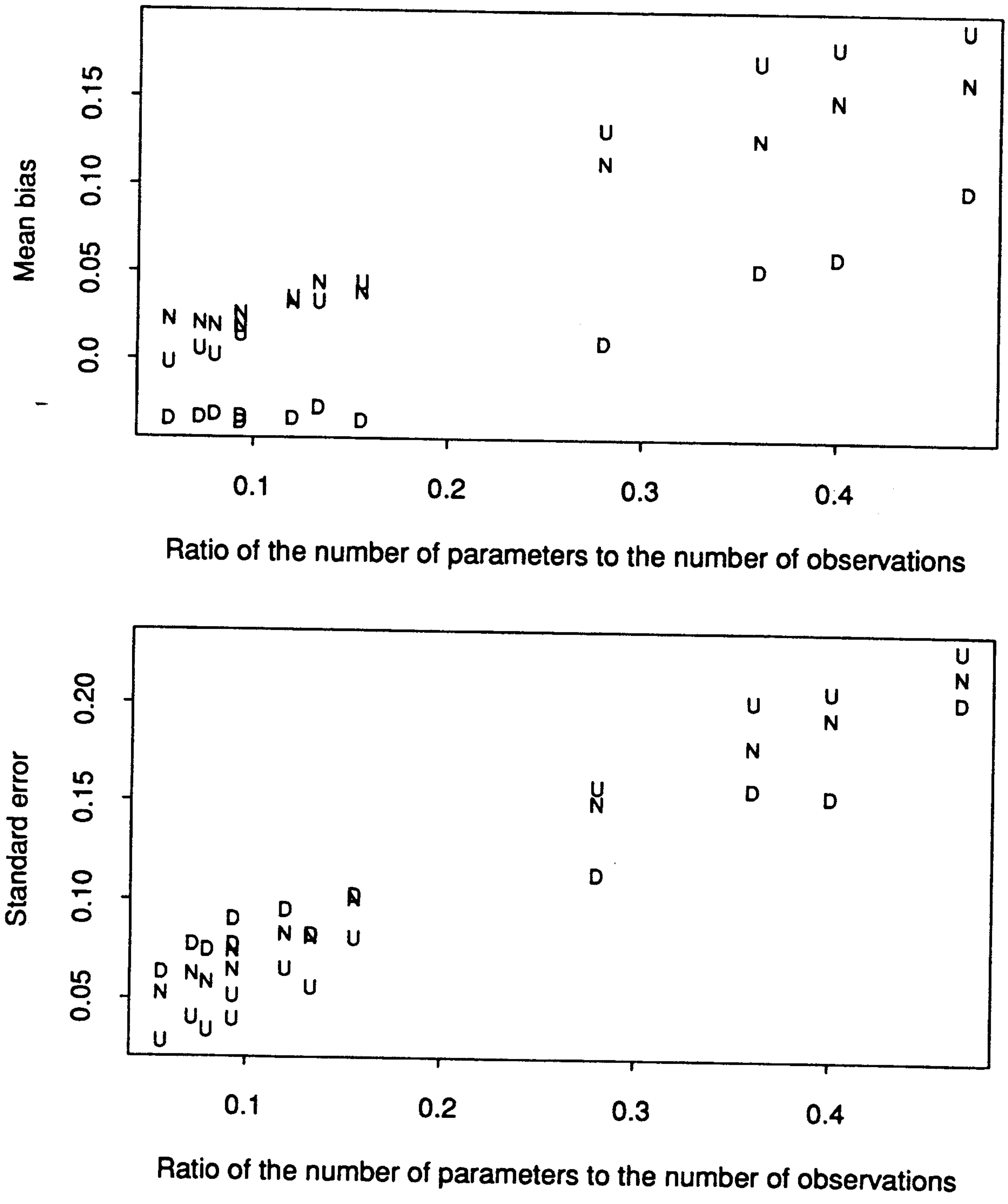

Figure 1. Mean bias and standard error versus the ratio of the number of parameters to the number of observations when the true power is 50%. The letter N stands for normal, D for double exponential and U for uniform random errors.

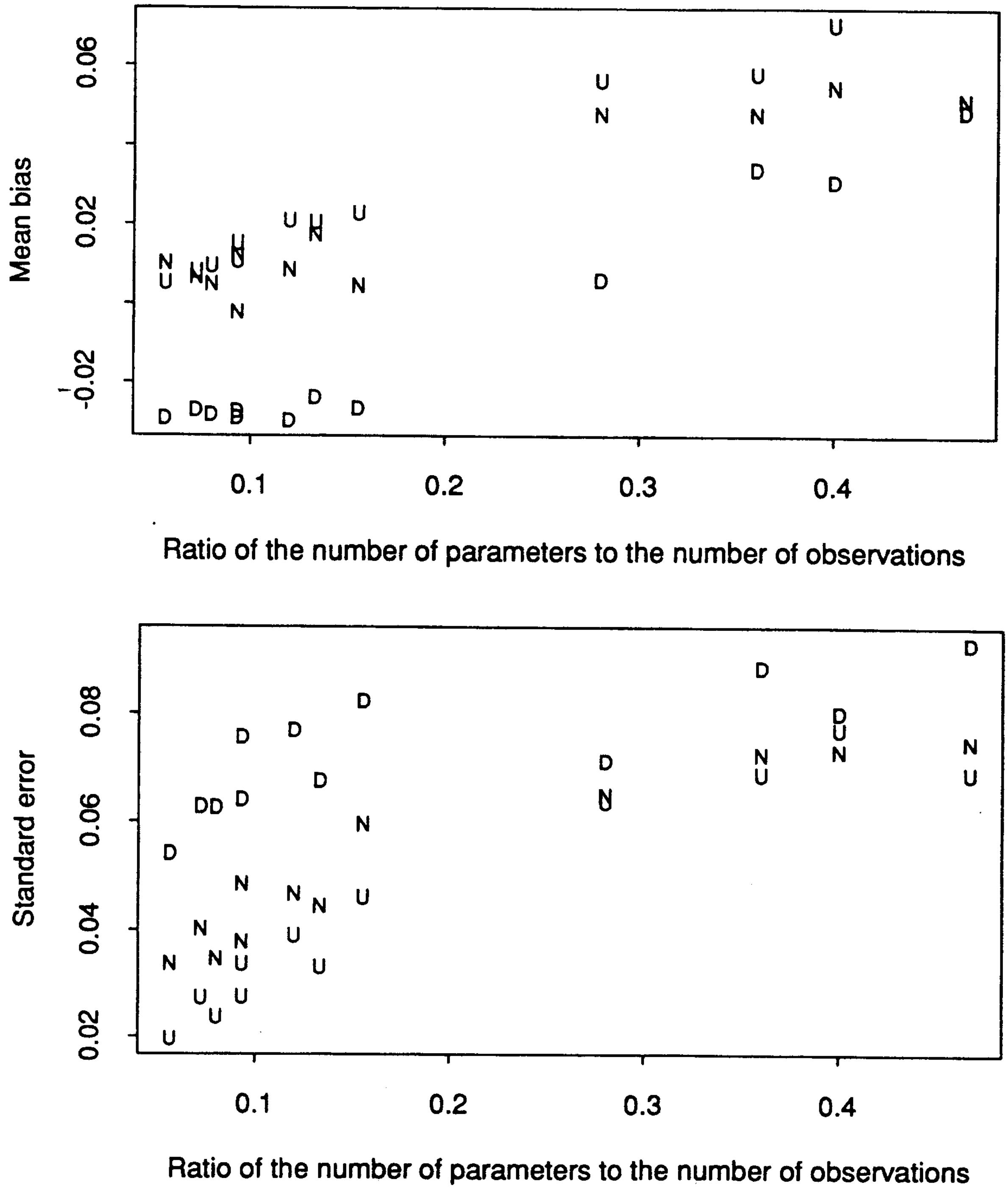

Figure 2. Mean bias and standard error versus the ratio of the number of parameters to the number of observations when the true power is 90%. The letter N stands for normal, D for double exponential and U for uniform random errors.

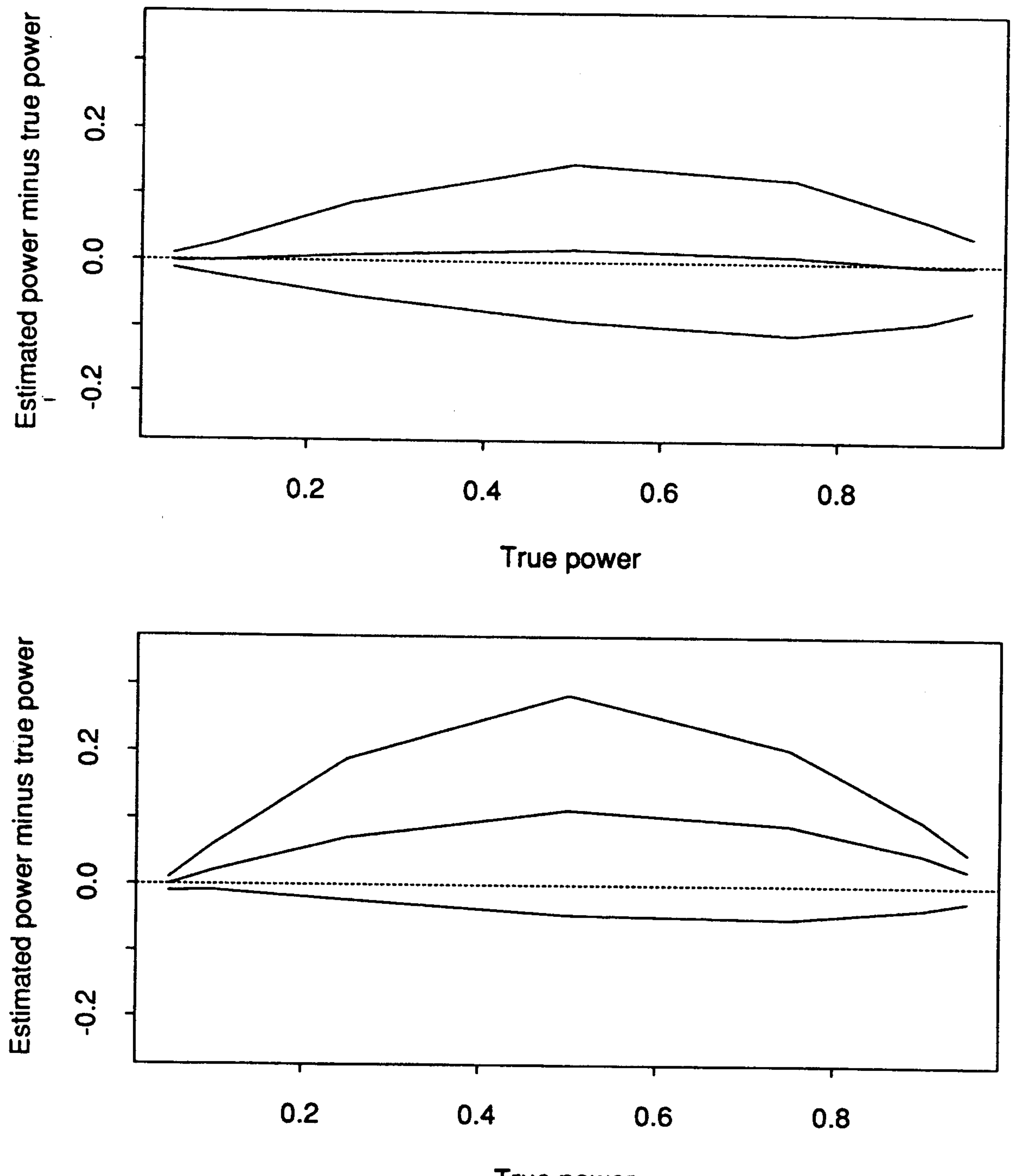

Figure 3. Mean curve and 90% envelope curves for the deviation of the estimated power from the true power. The distribution of the random errors is normal and the design is 5 (10) blocks, 3 (5) treatments and 5 (1) observations per cell for the upper (lower) plot.

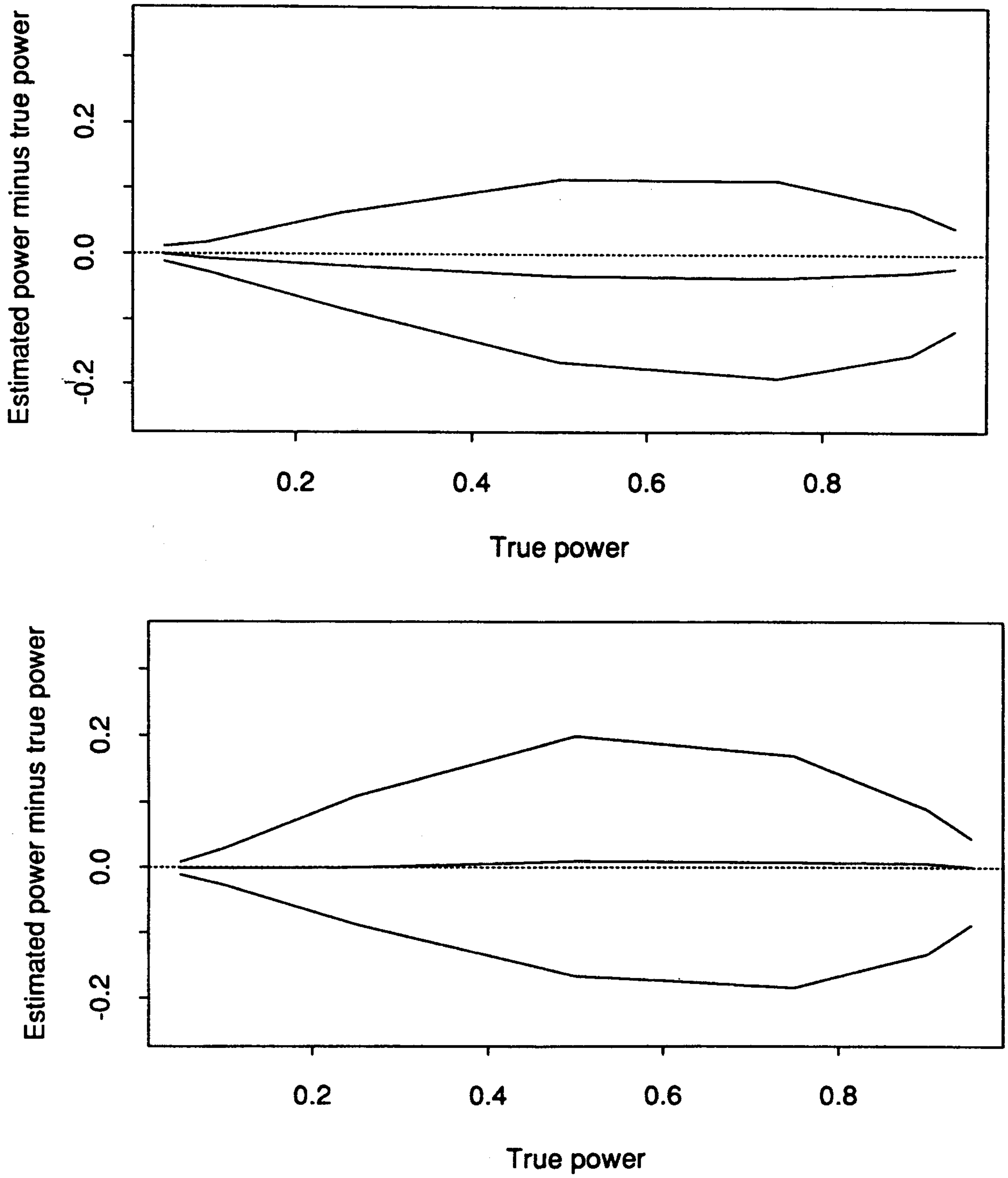

Figure 4. Mean curve and 90% envelope curves for the deviation of the estimated power from the true power. The distribution of the random errors is double exponential and the design is 5 (10) blocks, 3 (5) treatments and 5 (1) observations per cell for the upper (lower) plot.

5. Conclusion

We have used the bootstrap to estimate the power of a rank test introduced by Mack and Skillings (1980) to test the hypothesis of no treatment effect in a randomized block design. To use the bootstrap, bootstrap errors must be simulated from a smoothed version of the empirical distribution function of the residuals. The bootstrap estimate of power is consistent under local alternatives provided that the estimator of the distribution of the errors is smooth enough. A kernel estimate is an example. Results from a simulation study are very good, especially when the ratio of the number of parameters to the number of observations is less than 1/10. In that case, for all powers, the absolute mean bias is less than 0.04 and the standard error less than 0.10.

Appendix

This appendix contains the proof of Theorem 1. Since it is possible to use many estimates of F in applying the bootstrap, we will use the approach of Beran (1984) to show the asymptotic consistency.

Let C_F be the class of sequences of fixed distribution functions $\{F_N\}$ with densities $\{f_N\}$ and its first derivatives $\{f'_N\}$ such that the conditions

$$\sup_x |F_N(x) - F(x)| = O(N^{-1/2}\ln N), \tag{10}$$

$$\sup_N \sup_x |f'_N(x)| = O(1), \tag{11}$$

and

$$\int_{-\infty}^{\infty} f_N^2(x)dx \to \int_{-\infty}^{\infty} f^2(x)dx \qquad \text{as } N \to \infty \tag{12}$$

are satisfied where f is the density of F.

We begin with a lemma. Let R_{ijk} be the rank of X_{ijk} in $X_{i11}, X_{i12}, \ldots, X_{iJn_{iJ}}$, i.e., its rank in the ith row. For now, let the index i be fixed and consider the vector $R_N^{(i)} = (R_{N,1}^{(i)}, R_{N,2}^{(i)}, \ldots, R_{N,J}^{(i)})'$ where

$$R_{N,j}^{(i)} = \frac{\sqrt{12}}{\sqrt{N}}\left(\frac{1}{n_{i\cdot}}R_{ij} - \frac{p_{\cdot j}(n_{i\cdot}+1)}{2} - \frac{\xi_j^{(i)} b_{ij}(n_{i\cdot}+1)}{n_{i\cdot}}\right),$$

with R_{ij} the sum of the ranks in the ijth cell, i.e., $R_{ij} = \Sigma_{k=1}^{n_{ij}} R_{ijk}$, and

$$\xi_j^{(i)} = \frac{\sqrt{p_{i\cdot}p_{\cdot j}}}{\sqrt{1-p_{\cdot j}}}\left(\theta_j - \sum_{l=1}^{J}\theta_l p_{\cdot l}\right)\int_{-\infty}^{\infty} f^2(x)dx, \tag{13}$$

$$b_{ij} = \sqrt{n_{i\cdot}p_{\cdot j}(1-p_{\cdot j})}, \tag{14}$$

with f the density of F. The following lemma shows that $R_N^{(i)}$ is asymptotically normal.

Lemma 1. *Let F have a density f with bounded derivative f'. Let $\{F_N\}$ be in C_F. Consider the local alternatives Model* (5) *with ϵ_{ijk} independently and identically distributed from F_N, and $n_{ij} = Np_{ij} = Np_{i\cdot}p_{\cdot j} > 0$, for all i and j. Then the vector $R_N^{(i)}$ has a multivariate asymptotic normal distribution with mean* 0 *and covariance matrix $p_{i\cdot}\Sigma$, where $\Sigma_{k,k} = p_{\cdot k}(1 - p_{\cdot k})$, $k = 1, \ldots, J$, and $\Sigma_{k,l} = -p_{\cdot k}p_{\cdot l}$ for $k \neq l$.*

Proof. We begin by showing the asymptotic normality of a vector $S_N^{(i)}$, apply a linear transformation to it, modify its mean, before concluding that $R_N^{(i)}$ is asymptotically normal. Consider the simple linear rank statistic $S_{N,j}^{(i)}$ given by

$$S_{N,j}^{(i)} = \frac{1}{b_{ij}(n_{i\cdot} + 1)} \left(\sum_{k=1}^{n_{ij}} R_{ijk} - p_{\cdot j} \frac{n_{i\cdot}(n_{i\cdot} + 1)}{2} \right) = \sum_{l=1}^{J} \sum_{k=1}^{n_{il}} c_{lk}^{j} a(R_{ilk}), \tag{15}$$

where

$$c_{lk}^{j} = \begin{cases} (1 - p_{\cdot j})/b_{ij}, & \text{if } l = j, \\ -p_{\cdot j}/b_{ij}, & l \neq j, \end{cases}$$

and the scores $a(k)$ are generated as follows

$$a(k) = \phi\left(\frac{k}{n_{i\cdot} + 1}\right),$$

with $\phi(u) = u$, and b_{ij}^2 defined in (14). Now let $S_N^{(i)} = \sqrt{12}(S_{N,1}^{(i)} - \mu_{N,1}^{(i)}, \ldots, S_{N,J}^{(i)} - \mu_{N,J}^{(i)})'$, where

$$\mu_{N,j}^{(i)} = \sum_{l=1}^{J} \sum_{k=1}^{n_{il}} c_{lk}^{(j)} \int_{-\infty}^{\infty} \left(\sum_{m=1}^{J} \sum_{n=1}^{n_{im}} \frac{F_N(x - \theta_m/\sqrt{N})}{n_{i\cdot}} \right) f_N(x - \theta_l/\sqrt{N}) dx.$$

We now show that $S_N^{(i)}$ is asymptotically normal with mean 0 and covariance matrix V with $V_{k,k} = 1$, $k = 1, \ldots, J$ and $V_{k,l} = -\sqrt{p_{\cdot k}p_{\cdot l}}/\sqrt{(1 - p_{\cdot k})(1 - p_{\cdot l})}$ for $k \neq l$.

Let $\lambda' = (\lambda_1, \ldots, \lambda_J)$ be any fixed vector in $\mathbb{R}^J$ and let $Z' = (Z_1, \ldots, Z_J)$ be a multivariate normal random vector with mean 0 and covariance matrix V. Using Theorem C of Serfling (1980, p.303), we now show that $\lambda' S_N^{(i)}$ converges weakly to a normal random variable with mean 0 and variance $\lambda' V \lambda$. Since ϕ has a bounded second derivative, $\Sigma\Sigma c_{lk}^{j} = 0$, $\Sigma\Sigma(c_{lk}^{j})^2 = 1$, and $\max(c_{lk}^{j})^2 = O(N^{-1/2})$, then the conditions of Serfling's Theorem on c_{lk}^{j} and ϕ for $S_{N,j}^{(i)}$ are satisfied. Now consider $E_i = \Sigma_{j=1}^{J} \lambda_j S_{N,j}^{(i)}$. We can write $E_i = \Sigma_{l=1}^{J} \Sigma_{k=1}^{n_{il}} d_{lk} a(R_{ilk})$ where $d_{lk} = \Sigma_{j=1}^{J} \lambda_j c_{lk}^{j}$. Because the sum of the c_{lk}^{j} for fixed j is 0, then the sum of the d_{lk} is also 0. Also, since $\max(c_{lk}^{j})^2 = O(N^{-1/2})$, the same is true of the d_{lk}. Let $\Sigma_{l=1}^{J} \Sigma_{k=1}^{n_{il}} d_{lk}^2 = e^2 \geq 0$. Before applying Serfling's Theorem, we must first show that the following condition is satisfied:

$$\sup_{1 \leq i,j \leq J;\, x} |F_{N_i}(x) - F_{N_j}(x)| = O(N^{-1/2} \log N), \tag{16}$$

where $F_{N_i}(x) = F_N(x - \theta_i/N^{1/2})$. Now

$$\begin{aligned}\sup_{1\le i,j\le J;\,x} |F_{N_i}(x) - F_{N_j}(x)| \le &\sup_{1\le i,j\le J;\,x} |F_{N_i}(x) - F(x-\theta_i/N^{1/2})| \\ &+ \sup_{1\le i,j\le J;\,x} |F(x-\theta_i/N^{1/2}) - F(x-\theta_j/N^{1/2})| \\ &+ \sup_{1\le i,j\le J;\,x} |F(x-\theta_j/N^{1/2}) - F_{N_j}(x)|.\end{aligned}$$

The first and third terms are $O(N^{-1/2}\ln N)$ by Condition (10) of C_F while the second one is $O(N^{-1/2})$ under the assumption that f' is bounded, using a Taylor series expansion of F. Therefore, applying Serfling's Theorem and some algebra we find that $\sqrt{12}(E_i - \Sigma_{j=1}^J \mu_{N,j}^{(i)}\lambda_j)$ converges weakly to a normal distribution with mean 0 and variance e^2, since $\sigma_\phi^2 = \int_0^1 [t-1/2]^2 dt = 1/12$. Through simple algebra, it can be shown that $\lambda' V \lambda = e^2$, so that $\sqrt{12}(E_i - \Sigma_{j=1}^J \mu_{N,j}^{(i)}\lambda_j)$ converges weakly to $\Sigma_{j=1}^J \lambda_j Z_j$. Hence the vector $S_N^{(i)}$ converges weakly to a multivariate normal distribution with mean 0 and covariance matrix V.

Next, we develop the term $\mu_{N,j}^{(i)}$. First, start with the integral in $\mu_{N,j}^{(i)}$.

$$\begin{aligned}&\int_{-\infty}^{\infty} \left(\sum_{m=1}^{J}\sum_{n=1}^{n_{im}} \frac{F_N(x-\theta_m/\sqrt{N})}{n_{i\cdot}}\right) f_N(x-\theta_l/\sqrt{N})dx \\ &= \sum_{m=1}^{J} p_{\cdot m} \int_{-\infty}^{\infty} F_N(x-\theta_m/\sqrt{N}) f_N(x-\theta_l/\sqrt{N})dx,\end{aligned}$$

since $n_{im}/n_{i\cdot} = Np_{i\cdot}p_{\cdot m}/(Np_{i\cdot}) = p_{\cdot m}$ by the proportional frequencies hypothesis. If $m = l$, then using integration by parts we obtain $\int_{-\infty}^{\infty} F_N(x-\theta_m/\sqrt{N}) f_N(x - \theta_l/\sqrt{N})dx = 1/2$. If $m \neq l$, making the change of variable $y = x - \theta/\sqrt{N}$, then

$$\begin{aligned}&\int_{-\infty}^{\infty} F_N(x-\theta_m/\sqrt{N}) f_N(x-\theta_l/\sqrt{N})dx \\ &= \int_{-\infty}^{\infty} F_N(y + (\theta_l - \theta_m)/\sqrt{N}) f_N(y)dy \\ &= \int_{-\infty}^{\infty} [F_N(y) + (\theta_l-\theta_m)/\sqrt{N} f_N(y)] f_N(y)dy + O(1/N) \\ &= 1/2 + (\theta_l - \theta_m)/\sqrt{N} \int_{-\infty}^{\infty} f_N^2(x)dx + O(1/N),\end{aligned}$$

using a Taylor series expansion and Condition (11) of C_F. So,

$$\begin{aligned}\mu_{N,j}^{(i)} = &\frac{1}{2}\sum_{l=1}^{J}\sum_{k=1}^{n_{il}} c_{lk}^{(j)} \sum_{m=1}^{J} p_{\cdot m} + \int f_N^2 \sum_{l=1}^{J} \theta_l/\sqrt{N} \sum_{k=1}^{n_{il}} c_{lk}^{(j)} \sum_{m=1;m\neq l}^{J} p_{\cdot m} \\ &- \int f_N^2 \sum_{l=1}^{J}\sum_{k=1}^{n_{il}} c_{lk}^{(j)} \sum_{m=1;m\neq l}^{J} \frac{p_{\cdot m}\theta_m}{\sqrt{N}} + \sum_{l=1}^{J}\sum_{k=1}^{n_{il}} c_{lk}^{(j)} \sum_{m=1;m\neq l}^{J} p_{\cdot m} O(1/N) \\ = &A + B + C + D.\end{aligned}$$

Since the sum of the $c_{lk}^{(j)}$ is 0, it follows that $A = 0$.

Using the definition of $c_{lk}^{(j)}$ and some algebra, we obtain that

$$B = \int f_N^2 \frac{\sqrt{p_{\cdot j} p_{i\cdot}}}{\sqrt{1-p_{\cdot j}}} \left(\theta_j - p_{\cdot j}\theta_j - \sum_{l=1}^{J} \theta_l p_{\cdot l} + \sum_{l=1}^{J} \theta_l p_{\cdot l}^2 \right),$$

$$C = -\int f_N^2 \frac{\sqrt{p_{\cdot j} p_{i\cdot}}}{\sqrt{1-p_{\cdot j}}} \left(-p_{\cdot j}\theta_j + \sum_{l=1}^{J} \theta_l p_{\cdot l}^2 \right),$$

and $D = O(1/N)$. Hence

$$\mu_{N,j}^{(i)} = \int f_N^2 \frac{\sqrt{p_{\cdot j} p_{i\cdot}}}{\sqrt{1-p_{\cdot j}}} \left(\theta_j - \sum_{l=1}^{J} \theta_l p_{\cdot l} \right) + O(1/\sqrt{N}).$$

Since $\int f_N^2(x)dx \to \int f^2(x)dx$ by Condition (12) of C_F, and $S_N^{(i)} = \sqrt{12}(S_{N,1}^{(i)} - \mu_{N,1}^{(i)}, \ldots, S_{N,J}^{(i)} - \mu_{N,J}^{(i)})'$ is asymptotically normal with mean 0 and covariance matrix V, we can replace $\mu_{N,j}^{(i)}$ by $\xi_j^{(i)}$ of (13). But the vector $R_N^{(i)}$ is $(n_{i\cdot}+1)/n_{i\cdot}$ times the product of the diagonal matrix D with the vector $\sqrt{12}(S_{N,1}^{(i)} - \xi_j^{(i)}, \ldots, S_{N,J}^{(i)} - \xi_J^{(i)})$, where the vector on the diagonal of D is $(b_{i1}/\sqrt{N}, \ldots, b_{iJ}/\sqrt{N})$. Therefore $R_N^{(i)}$ is asymptotically normal with mean 0 and covariance matrix DVD. But $DVD = p_{i\cdot}\Sigma$, which concludes the proof of the lemma.

In the next lemma, we sum up independent asymptotically normal vectors to obtain the asymptotic normality of the vector based on which the test statistic T_N is a quadratic form. Since all observations are independent and the ranking is done independently in each row, the proof is straightforward and not included.

Lemma 2. *Let $U_N = (U_{N,1}, \ldots, U_{N,J})'$ where $U_{N,j} = \sqrt{12/N}p_{\cdot j}[R_j - (N+I)/2]$ and $R_j = \sum_{i=1}^{I} R_{ij}/n_{i\cdot}$. Under the conditions of Lemma 1, U_N is asymptotically multivariate normal with mean vector γ and covariance matrix Σ of Lemma 1 where $\gamma_j = \sqrt{12}\int f^2(x)p_{\cdot j}(\theta_j - \Sigma_{l=1}^{J}\theta_l p_{\cdot l})dx$.*

Consider the diagonal matrix C with vector $(1/p_{\cdot 1}, \ldots, 1/p_{\cdot J})'$ on its diagonal. Then $T_N = N/(N+I)U_N' C U_N$. It can easily be shown that $C\Sigma$ is idempotent and that $trace(C\Sigma) = J - 1$. Hence, using Lemma 2, T_N is asymptotically distributed according to a chi-squared distribution with $J-1$ degrees of freedom and noncentrality parameter $\gamma' C \gamma = \lambda_T$ of Equation (6), proving the following Lemma.

Lemma 3. *Suppose that F has a density f and that its derivative f' is bounded. Let $\{F_N\}$ be in C_F. Consider the local alternatives Model (5) with ϵ_{ijk} independently and identically distributed from F_N, and $n_{ij} = Np_{ij} = Np_{i\cdot}p_{\cdot j} > 0$, for all i and j. Then the asymptotic distribution of T_N is a noncentral chi-squared* *distribution with $J-1$ degrees of freedom and noncentrality parameter given by* (6).

Note that Lemma 3 is a generalization of the result of Mack and Skillings (1980) which was for $F_N \equiv F$. Let us now prove Theorem 1.

Proof of Theorem 1. We must show that the conditions (10)–(12) are appropriately satisfied by the random sequence $\{\hat{F}_{\lambda_N}\}$.

For Condition (10), we write $\sqrt{N}(\hat{F}_{\lambda_N}(x) - F(x))$ as $\sqrt{N}(\hat{F}_{\lambda_N}(x) - E\hat{F}_{\lambda_N}(x) + E\hat{F}_{\lambda_N}(x) - F(x))$. Following the proof of Theorem 23.2.1 of Shorack and Wellner (1986) (which will be referred to as S&W), p.765, we get

$$\begin{aligned}\sqrt{N}(\hat{F}_{\lambda_N}(x) - F(x)) &= \int_{-\infty}^{\infty} K\left(\frac{x-y}{\lambda_N}\right) d[\sqrt{N}(\hat{F}_N - F)](y) \\ &\quad + \sqrt{N}\left[\int_{-\infty}^{\infty} K\left(\frac{x-y}{\lambda_N}\right) dF(y) - F(x)\right] \\ &= R_{1,N}(x) + R_{2,N}(x),\end{aligned}$$

where $\hat{F}_N$ is the empirical distribution function of the residuals. Using integration by parts, $R_{1,N}(x) = -\int_{-\infty}^{\infty} \sqrt{N}[\hat{F}_N(x - s\lambda_N) - F(x - s\lambda_N)]dK(s)$. Since conditions (F1) and (F2) are satisfied, we can use Theorem 4.6.2 of S&W, p.198, and their special construction of Theorem 3.1.1, p.93, to find a version of $\hat{F}_N$ and a Brownian bridge $U(x)$ with uniformly continuous sample paths such that

$$\sup_x \left|\sqrt{N}(\hat{F}_N(x) - F(x)) - \left\{U(F(x)) + f(x)\int_0^1 F^{-1}(x)dU(x)\right\}\right| \to 0, \quad \text{in probability.}$$

Now, since $\sup_x |U(F(x)) + f(x)\int_0^1 F^{-1}(x)dU(x)|$ is $O_P(1)$, then $\sup_x |\sqrt{N}(\hat{F}_N(x) - F(x))|$ is also $O_P(1)$. Hence $\sup_x |R_{1,N}(x)|$ is $O_P(1)$. Moreover, from the proof of Theorem 23.2.1 of S&W, $\sup_x |R_{2,N}(x)|$ is $O(\sqrt{N}\lambda_N^2)$ provided that Condition (K1) is satisfied. Therefore, since $N^{1/2}\lambda_N^2/\log(N) \to 0$,

$$\frac{\sqrt{N}}{\log(N)} \sup_x |\hat{F}_{\lambda_N}(x) - F(x)| \to 0, \qquad \text{in probability.} \tag{17}$$

To consider conditions (11) and (12), first note that

$$\hat{F}_{\lambda_N}(x) = \frac{1}{N}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n_{ij}} K\left(\frac{x - \hat{\epsilon}_{ijk}}{\lambda_N}\right) = \sum_{i=1}^{I}\sum_{j=1}^{J} p_{ij}\hat{F}_{\lambda_N}^{(ij)}(x),$$

where $\hat{F}_{\lambda_N}^{(ij)}(x)$ is the kernel estimate constructed from the residuals in the ijth cell. It is sufficient to show that each $\hat{F}_{\lambda_N}^{(ij)}(x)$ satisfies the two conditions. Fix the cell ij and note that $\hat{f}_{\lambda_N}^{(ij)\prime}(x) = \tilde{f}_{\lambda_N}^{(ij)\prime}(x + \hat{\tau}_{ij})$ where $\tilde{f}_{\lambda_N}^{(ij)\prime}(x)$ is the derivative of the kernel density estimate computed from the i.i.d $X_{ij1}, \ldots, X_{ijn_{ij}}$, and $\hat{\tau}_{ij} = \hat{\mu} + \hat{\alpha}_i + \hat{\theta}_j / N^{1/2}$ is computed from the least squares estimates of the model (5). So $\sup_{N,x} |\hat{f}_{\lambda_N}^{(ij)\prime}(x)| = \sup_{N,x} |\tilde{f}_{\lambda_N}^{(ij)\prime}(x)|$ and we need only prove that the derivative of a kernel density estimate based on i.i.d. observations $\tilde{f}_{\lambda_N}'$ is uniformly bounded over x and N. Assuming that f' is uniformly continuous, that $\int_{-\infty}^{\infty} k(x)dx = 1$ (satisfied under conditions (F3) and (K1)), that Condition (K3) is satisfied, and provided that $N^{-1}\lambda_N^{-3}\log(1/\lambda_N) \to 0$, Silverman (1978) has shown that $\sup_x |\tilde{f}_{\lambda_N}'(x) - f'(x)| \to 0$ almost surely. Since it is assumed that f' is bounded under (F3), then for almost all sample paths, $\tilde{f}_{\lambda_N}'$ is uniformly bounded. Therefore, for almost all sample paths of $\hat{F}_{\lambda_N}$, Condition (11) is satisfied.

Now for Condition (12), note that $\int_{-\infty}^{\infty} [\hat{f}_{\lambda_N}^{(ij)}(x)]^2 dx = \int_{-\infty}^{\infty} [\tilde{f}_{\lambda_N}^{(ij)}(x + \hat{\tau}_{ij})]^2 dx$ which by a change of variable is $\int_{-\infty}^{\infty} [\tilde{f}_{\lambda_N}^{(ij)}(x)]^2 dx$. Hence, it is sufficient to consider a kernel estimate from i.i.d. observations. But Hall and Marron (1987) have shown that $\int_{-\infty}^{\infty} [\tilde{f}_{\lambda_N}^{(ij)}(x)]^2 dx \to \int_{-\infty}^{\infty} [f(x)]^2 dx$ in probability under general conditions which are satisfied if the kernel $k(x)$ is symmetric, of order 2, and vanishes at $\pm\infty$ and if the derivative of the density f is uniformly bounded, these conditions being satisfied under (K1), (K2), (K3)(a) and (F3).

Finally, to show (9), all we need do is prove that for any subsequence there is a further subsequence for which (9) holds almost surely (e.g., Theorem 20.5 of Billingsley (1986)). Let N_i be any subsequence. Since $\int_{-\infty}^{\infty} \hat{f}_{\lambda_N}^2(x)dx$ converges to $\int_{-\infty}^{\infty} f^2(x)dx$ in probability, there exists a further subsequence $N_{i(j)}$ such that $\int_{-\infty}^{\infty} \hat{f}_{\lambda_{N_{i(j)}}}^2(x)dx \to \int_{-\infty}^{\infty} f^2(x)dx$ almost surely. Moreover, by (17), there exists a further subsequence $N_{i(j)[k]}$ such that $\sup_x \sqrt{N_{i(j)[k]}}/\log(N_{i(j)[k]})|\hat{F}_{\lambda_{N_{i(j)[k]}}}(x) - F(x)| \to 0$ almost surely. Hence the subsequence $\{\hat{F}_{\lambda_{N_{i(j)[k]}}}\}$ satisfies conditions (10), (11), and (12) almost surely. Thus

$$
\begin{aligned}
&\sup_x |G_{N_{i(j)[k]}}(x, \hat{F}_{\lambda_{N_{i(j)[k]}}}) - G_{N_{i(j)[k]}}(x, F)| \\
&\leq \sup_x |G_{N_{i(j)[k]}}(x, \hat{F}_{\lambda_{N_{i(j)[k]}}}) - H(x; \lambda_T, J-1)| \\
&\quad + \sup_x |G_{N_{i(j)[k]}}(x, F) - H(x; \lambda_T, J-1)| \\
&\to 0, \qquad \text{almost surely,}
\end{aligned}
$$

using Lemma 3 twice and where $H(x; \lambda, J)$ is the distribution function of a noncentral chi-squared distribution with J degrees of freedom and noncentrality parameter λ. This completes the proof of Theorem 1.

Acknowledgements

The authors wish to thank the Natural Science and Engineering Research Council of Canada and the Fonds pour la formation de chercheurs et l'aide à la recherche of the Province of Québec for their financial support. We also wish to thank Joseph P. Romano and Luc Devroye for helpful discussions and two anonymous referees whose suggestions led to an improved presentation.

References

Arcones, M. A. and Giné, E. (1991). Some bootstrap tests of symmetry for univariate continuous distributions. *Ann. Statist.* **19**, 1496-1511.

Beran, R. (1984). Bootstrap methods in statistics. *Jahresber. Deutsch. Math. Verein* **86**, 14-30.

Beran, R. (1986). Simulated power functions. *Ann. Statist.* **14**, 151-173.

Beran, R. (1988). Prepivoting test statistics: A bootstrap view of asymptotic refinements. *J. Amer. Statist. Assoc.* **83**, 687-697.

Beran, R. and Millar, P. W. (1989). A stochastic minimum distance test for multivariate parametric models. *Ann. Statist.* **17**, 125-140.

Bickel, P. J. and Freedman, D. A. (1983). Bootstrapping regression models with many parameters. In *A festschrift for Erich Lehmann* (Edited by P. J. Bickel, K. Doksum, and J. L. Hodges), 28-48, Wadsworth, Belmont CA.

Billingsley, P. (1986). *Probability and Measure.* John Wiley, New York.

Chen, H. and Loh, W.-Y. (1991). Consistency of the bootstrap for the transformed two-sample *t*-test. *Comm. Statist. Theory Methods* **20**, 997-1014.

Collings, B. J. and Hamilton, M. A. (1988). Estimating the power of the two-sample Wilcoxon test for location shift. *Biometrics* **44**, 847-860.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann. Statist.* **7**, 1-26.

Fisher, N. I. and Hall, P. (1990). On bootstrap hypothesis testing. *Austral. J. Statist.* **32**, 177-190.

Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Amer. Statist. Assoc.* **32**, 675-701.

Hall, P. and Marron, J. S. (1987). Estimation of integrated squared density derivatives. *Statist. Probab. Lett.* **6**, 109-115.

Hamilton, M. A. and Collings, B. J. (1991). Determining the appropriate sample size for nonparametric tests for location shift. *Technometrics* **33**, 327-337.

Mack, G. A. and Skillings, J. H. (1980). A Friedman-type rank test for main effects in a two-factor ANOVA. *J. Amer. Statist. Assoc.* **75**, 947-951.

Quade, D. (1979). Using weighted rankings in the analysis of complete blocks with additive block effects. *J. Amer. Statist. Assoc.* **74**, 680-683.

Romano, J. P. (1988). A bootstrap revival of some nonparametric distance tests. *J. Amer. Statist. Assoc.* **83**, 698-708.

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

Shorack, G. R. and Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics.* John Wiley, New York.

Silverman, B. W. (1978). Weak and strong uniform consistency of the kernel estimate of a density and its derivatives. *Ann. Statist.* **6**, 177-184.

Tardif, S. (1987). Efficiency and optimality results for tests based on weighted rankings. *J. Amer. Statist. Assoc.* **82**, 637-644.

Département de mathématiques et de statistique, Université de Montréal, Montréal, PQ H3C 3J7, Canada.

Département d'informatique, et de recherche opérationnelle, Université de Montréal, Montréal, PQ H3C 3J7, Canada.

(Received February 1993; accepted March 1994)