Journal of Modern Applied Statistical Methods

Volume 6 | Issue 1 Article 31

5-1-2007

A Comparison of Two Rank Tests for Repeated Measure Designs

Tian Tian
University of Southern California

Rand R. Wilcox
University of Southern California, rwilcox@usc.edu

Recommended Citation

Tian, Tian and Wilcox, Rand R. (2007) "A Comparison of Two Rank Tests for Repeated Measure Designs," *Journal of Modern Applied Statistical Methods*: Vol. 6 : Iss. 1 , Article 31.
DOI: 10.22237/jmasm/1177993800

Journal of Modern Applied Statistical Methods
May, 2007, Vol. 6, No. 1, 331-335

1538 – 9472/07/$95.00

Early Scholars

A Comparison of Two Rank Tests for Repeated Measures Designs

Tian Tian Rand Wilcox
University of Southern California

This article compares the small-sample properties of the Agresti-Pendergast and the ATS rank-based method, as described in Brunner, Domh, and Langer (2002), for comparing J dependent groups. The results indicate that the Type I error of the Agresti-Pendergast method is more conservative when $J = 2$, but under most conditions, the ATS method performs best in terms of both Type I errors and power.

Key words: rank tests, repeated measures

Introduction

The classic rank-based method for comparing J dependent groups is Friedman's test. Consider a random sample of *n* vectors from some *J*-variate distribution. As is well-known, Friedman's test assigns ranks to the values within each vector and is based on a compound symmetry assumption under the hypothesis of no treatment effect (e.g., Brunner, Domhof, & Langer, p. 68). That is, the distribution is assumed to be invariant under all permutations, which implies that the variances and covariances are equal. Two attempts at improving test between Friedman's and are based in part by assigning ranks to the pooled data instead (Iman, 1974; Quade, 1979). Subsequently, Agresti and Pendergast (1986) proposed a rank-based test that was found to provide better control over the probability of a Type I error and better power. (For relevant theoretical results, see Kepner & Robinson, 1988.) Two alternative methods are described by Brunner, Domhof and Langer (2002, section 7.2.2). The first, based on a Wald-type statistic, is known to be rather unsatisfactory when the sample size is relatively small. The second is an ANOVA-type statistic (ATS) that was found to be preferable to the Wald-type statistic, but no results were provided about how it compares to the Agresti-Pendergast technique. The goal in this article is to compare their small-sample properties via simulations. The results indicate that the ATS method performs better than the Agresti-Pendergast technique for most of conditions.

Tian Tian is a graduate student in psychology. She specializes in quantitative methods. Rand Wilcox is Professor of Psychology. He is the author of seven books on statistics. His research interests are robust methods. Email him at rwilcox@almaak-01.usc.edu.

Description of the Methods

Let $\mathbf{X}_1, ..., \mathbf{X}_n$, where $\mathbf{X}_k = (X_{k1}, ..., X_{kJ})'$, $k = 1, \ldots, n$, be a random sample from a *J*-variate distribution with distribution $\mathbf{F} = (F_1, ..., F_J)'$. In the event sampling is from a discrete distribuition, the j^{th} marginal distribution is taken to be $F_j(x) = \frac{1}{2}[F_j^+(x) + F_j^-(x)]$, where F_j^+ and denote F_j^- are the right continuous and the left continuous version of the distribution function, respectively. That is, $F_j^-(x) = P(X_j < x)$ and $F_j^+(x) = P(X_j \le x)$. The total number of observations is $N = n \times J$ and the null hypothesis is $H_0 : F_1 = ... = F_J$.

Agresti-Pendergast Test

Let R_{ij} be the midrank of X_{ij} among all N observations. The midrank is determined by means of the so-called counting functions

$$c^-(x) = \begin{cases} 0, & x \le 0 \\ 1, & x > 0 \end{cases},$$

$$c^+(x) = \begin{cases} 0, & x < 0 \\ 1, & x \ge 0 \end{cases},$$

and

$$c(x) = \frac{1}{2}[c^+(x) + c^-(x)].$$

The midrank of X_{ij} among the N random variables in the i^{th} row and the j^{th} colomn can be written as $R_{ij} = \frac{1}{2} + \sum_{k=1}^{n}\sum_{l=1}^{J} c(X_{ij} - X_{kl})$. Let $\overline{\mathbf{R}}_j = \frac{1}{n}\sum_{i=1}^{n} R_{ij}$ for $1 \le j \le J$. The estimated covariance matrix $\mathbf{S}$ of the ranks, which has entries s_{ij}, is

$$s_{ij} = \frac{1}{n-J+1}\sum_{i=1}^{n}(\mathbf{R}_{ij} - \overline{\mathbf{R}}_j)(\mathbf{R}_{ij} - \overline{\mathbf{R}}_k).$$

Under general conditions, the asymptotic distribution of $\mathbf{R}' = (\overline{\mathbf{R}}_1,...,\overline{\mathbf{R}}_J)$ is multivariate normal. Let $v = E(\mathbf{R})$ and

$$\mathbf{C}_1 = \begin{pmatrix} 1 & -1 & 0 & ... & 0 & 0 \\ 0 & 1 & -1 & ... & 0 & 0 \\ . & . & . & . & . & . \\ 0 & 0 & 0 & ... & 1 & -1 \end{pmatrix}$$

The null hypothesis $H_0 : F_1 = ... = F_J$ implies that $\mathbf{C}_1 v = 0$ and the test statistic is

$$F = \frac{n}{J-1}(\mathbf{C}_1\mathbf{R})'(\mathbf{C}_1\mathbf{S}\mathbf{C}_1')^{-1}\mathbf{C}_1\mathbf{R},$$

which has, approximately, an F distribution with degrees of freedom $J-1$ and $(J-1)(n-1)$ when null hypothesis is true.

ATS

Following the notation in Brunner et al. (2002), let $\mathbf{I}_J = diag\{1,...,1\}$ be the J-dimensional identity matrix, let $\mathbf{J}_J$ denote the J-by-J matrix of 1s, and let

$$\mathbf{C}_2 = \mathbf{I}_J - \frac{1}{J}\mathbf{J}_J = \begin{pmatrix} 1 & ... & 0 \\ ... & ... & ... \\ 0 & ... & 1 \end{pmatrix}_{J\times J} - \frac{1}{J}\begin{pmatrix} 1 & ... & 1 \\ ... & ... & ... \\ 1 & ... & 1 \end{pmatrix}_{J\times J}.$$

The null hypothesis $H_0 : F_1 = ... = F_J$ is equivalent to

$$H_0 : \mathbf{C}_2\mathbf{F} = \begin{pmatrix} F_1 - \overline{F}_. \\ ... \\ F_J - \overline{F}_. \end{pmatrix} = \begin{pmatrix} 0 \\ ... \\ 0 \end{pmatrix} = 0.$$

Let

$$\mathbf{R}_i = (R_{i1},...,R_{iJ}),\ \overline{\mathbf{R}}_. = \frac{1}{n}\sum_{i=1}^{n}\mathbf{R}_i,$$

and let

$$\hat{\mathbf{V}}_n = \frac{1}{N^2(n-1)}\sum_{i=1}^{n}(\mathbf{R}_i - \overline{\mathbf{R}}_.)(\mathbf{R}_i - \overline{\mathbf{R}}_.)'$$

denote an estimate of the covariance matrix V_n. For the ATS method, the test statistic is

$$F_n(\mathbf{C}_2) = \frac{n}{N^2 trace(\mathbf{C}_2\hat{\mathbf{V}}_n)}\sum_{j=1}^{J}(\overline{\mathbf{R}}_{.j} - \frac{N+1}{2})^2.$$

Under $H_0 : \mathbf{C}_2\mathbf{F} = 0$, the distribution of F_n can be approximated by an F distribution with degrees of freedom $\hat{f} = \frac{[trace(\mathbf{C}_2\hat{\mathbf{V}}_n)]^2}{trace(\mathbf{C}_2\hat{\mathbf{V}}_n\mathbf{C}_2\hat{\mathbf{V}}_n)}$ and ∞.

Simulation Results

This section reports simulation results on the small-sample properties of the Agresti-Pendergast and ATS methods. The simulations were run with MATLAB 7.1. A correlation matrix with a common correlation ρ was used and observations were generated from J-variate normal distribution (J = 2, 3, 4). Because any order preserving transformation of the data does not alter the results, the simulation results apply to a wide range of non-normal distributions. The sample sizes were taken to be n = 10, 20, and 30 and correlations used were ρ = 0, 0.2, 0.5, 0.8 resulting in 36 conditions. A total of 1,000 replications were used to estimate the Type I error probabilities, denoted by $\hat{\alpha}$, and estimated power, which is denoted by $\hat{\gamma}$. When studying power, the mean of the marginal distribution of the first group was increased from zero to one. The results are given in Table 1.

Table 1. Estimated Type I error probabilities and powers for the ATS and Agresti-Pendergast methods, based on 1,000 replications

n	J		$\rho = 0$		$\rho = 0.2$		$\rho = 0.5$		$\rho = 0.8$	
			$\hat{\alpha}$	$\hat{\gamma}$	$\hat{\alpha}$	$\hat{\gamma}$	$\hat{\alpha}$	$\hat{\gamma}$	$\hat{\alpha}$	$\hat{\gamma}$
10	2	ATS	0.095	0.595	0.085	0.688	0.081	0.843	0.082	0.989
		Agresti-Pendergast	0.064	0.495	0.050	0.593	0.050	0.775	0.053	0.975
	3	ATS	0.062	0.585	0.069	0.669	0.065	0.853	0.062	0.998
		Agresti-Pendergast	0.092	0.581	0.080	0.641	0.084	0.813	0.074	0.994
	4	ATS	0.072	0.579	0.065	0.646	0.066	0.849	0.047	0.996
		Agresti-Pendergast	0.111	0.611	0.116	0.671	0.107	0.814	0.107	0.988
20	2	ATS	0.065	0.862	0.072	0.913	0.059	0.994	0.067	1
		Agresti-Pendergast	0.053	0.833	0.053	0.895	0.046	0.990	0.050	1
	3	ATS	0.058	0.885	0.052	0.949	0.063	0.997	0.066	1
		Agresti-Pendergast	0.065	0.883	0.056	0.936	0.070	0.993	0.067	1
	4	ATS	0.052	0.900	0.056	0.936	0.054	0.993	0.058	1
		Agresti-Pendergast	0.070	0.897	0.088	0.921	0.072	0.992	0.072	1
30	2	ATS	0.060	0.963	0.071	0.975	0.065	1	0.050	1
		Agresti-Pendergast	0.050	0.954	0.060	0.971	0.051	1	0.042	1
	3	ATS	0.058	0.984	0.063	0.995	0.049	0.999	0.060	1
		Agresti-Pendergast	0.062	0.973	0.062	0.988	0.051	0.999	0.064	1
	4	ATS	0.045	0.991	0.058	0.994	0.057	1	0.043	1
		Agresti-Pendergast	0.057	0.986	0.076	0.990	0.068	1	0.059	1

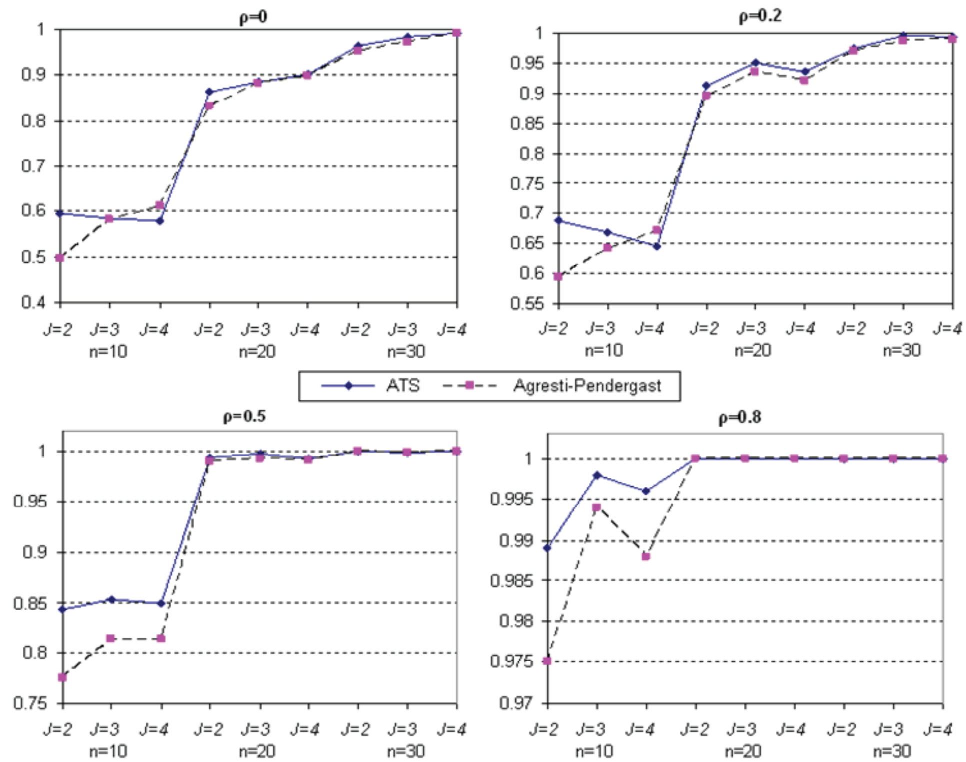

Figure 1 Plots of Power vs. Sample size for ATS test and Agresti-Pendergast test (multinormal)

As can be seen, for $n \geq 20$, the $\hat{\alpha}$ values are reasonably close to the nominal value of 0.05 for both the Agresti-Pendergast and ATS methods. With a fixed *n* and *ρ*, $\hat{\alpha}$ decreases with *J* increasing when using ATS, while $\hat{\alpha}$ increases for the Agresti-Pendergast test. For instance, when $\rho = 0.2$ and *n* = 30, the $\hat{\alpha}$ values are 0.071 (*J* = 2), 0.062 (*J* = 3), and 0.058 (*J* = 4) for ATS, and for the Agresti-Pendergast test $\hat{\alpha}$ values are 0.060 (*J* = 2), 0.062 (*J* = 3), and 0.076 (*J* = 4),. For $n = 10$ and *J* = 2, the ATS method can be unsatisfactory in terms of Type I errors, the estimate exceeding .075. Otherwise, ATS is generally preferable to the Agresti-Pendergast test. Also, for $n = 10$ and $J > 2$, now the Agresti-Pendergast method performs poorly in terms of Type I errors; the ATS method is preferable.

Table 2 gives the basic descriptive statistics of estimated the Type I errors and power for the two methods. As can be seen from the table, the Type I errors for ATS have smaller variances.

Table 2 Descriptive statistics of estimated Type I errors and powers for Agresti-Pendergast test and ATS

	Type I error $\hat{\alpha}$		Power $\hat{\gamma}$	
	mean	STD	Mean	STD
ATS	0.0625	0.0112	0.9061	0.1370
Agresti-Pendergast	0.0681	0.0192	0.8931	0.1485

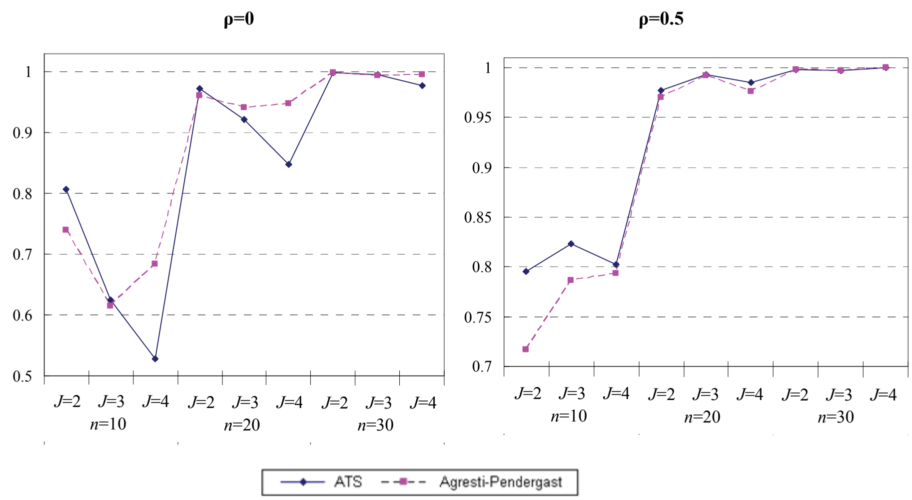

Figure 2 Plots of Power vs. Sample size for Brunner-Puri test and Agresti-Pendergast test (Bin(10, 0.4))

Figure 1 contains the estimated powers for all of the conditions. To make it clear, the four ρs are listed separately. As indicated, ATS is generally preferable.

The discrete case, where tied values occur, was also considered. For the goal of creating a reasonable number of tied values, the distribution used here is Binomial (10, 0.4). Figure 2 gives the plots of power vs. sample size in this case. As can be seen, ATS has higher power than Agresti-Pendergast for $\rho = 0.5$. For the independent case, the choice of method is less clear, with the Agresti-Pendergast offering a bit of an advantage in some instances.

Conclusion

In summary, the simulations show that in many situations, there is little separating ATS and Agresti-Pendergast. However, there are situations where ATS is preferable to Agresti-Pendergast in terms of both Type I errors and power. The main exception is the case $J = 2$ and $n = 10$, where the Agresti-Pendergast performs reasonably well in terms of Type I errors, while ATS does not.

References

Agresti, A. & Pendergast, J. (1986). Comparing mean ranks for repeated measures data. *Communications in Statistics: Theory and Method*, *15*, 1417-1433.

Brunner, E., Domhof, S. & Langer, F., (2002). *Nonparametric analysis of longitudinal data in factorial experiments*. New York: John Wiley & Sons.

Iman, R. L. (1974). Use of a t-statistic as an approximation to the exact distribution of the Wilcoxon signed rank statistic. *Communications in Statistics*, *3*, 795 -806.

Kepner, J. L. & Robinson, D. H. (1988). Nonparametric methods for detecting treatment effects in repeated-measures designs. *Journal of the American Statistical Association*, *83*, 456-461.

Quade, D. (1979). Using weighted rankings in the analysis of complete blocks with additive block effects. *Journal of the American Statistical Association*, *74*, 680–683.