韓國數學教育學會誌〈數學教育〉
1987. 12, Vol. XXVI, No. 1

Journal of the Korea Society of Mathematical Education
Dec. 1987, Vol. XXVI, No. 1

A Study on Distribution-Free Rank Test for Ordered Alternatives in Randomized Complete Block Designs

by Euen-Sook Jee

Kwangwoon University

1. Introduction

Let X_{ij} be the measurement on the j^{th} of k treatments in the i^{th} of n blocks. The usual model for a randomized complete block design is of the form

$$X_{ij}=\mu+\beta_i+\tau_j+\epsilon_{ij} \qquad (1)$$

where the β_i's represent block effects subject to the restriction that $\sum\beta_i=0$, the τ_j's represent the main effects of the treatments subject to the restriction that $\sum\tau_j=0$, and the ϵ_{ij}'s are mutually independent random variables from a continuous population with variance σ^2. Over the years much attention has been given to the problem of detecting ordered alternatives, that is,

$$H_1:\ \tau_1\leqq\tau_2\leqq\cdots\cdots\leqq\tau_k, \qquad (2)$$

where at least one inequality is strict, when this model is appropriate. Jonckheere (1954) suggested using a statistic of the form

$$K=\sum_{i=1}^{n}K_i$$

where K_i is Kendall's rank correlation between the observed ordering in the i^{th} block and the hypothesized ordering. Page(1963) suggested rejecting $H_0:\ \tau_1=\cdots\cdots=\tau_k$ in favor of (2) if

$$L=\sum_{j=1}^{n}jR._j$$

is too large, where R_{ij} is the rank of X_{ij} among $X_{i1},\cdots\cdots,X_{ik}$ and $R._j=\sum_{i=1}^{n}R_{ij}$. This test is seen to be equivalent to rejecting H_0 in favor of H_1 if

$$\varphi=\sum_{i=1}^{n}\varphi_i$$

is too large, where φ_i is Spearman's rank correlation coefficient between the observed ordering in the i^{th} block and the hypothesized ordering. Hollander (1967) proposed using the statistic

$$H=\sum_{u<v}\sum T_{uv}$$

where $\quad T_{uv}=\sum_{i=1}^{n}\psi_{uv}^{(i)}\ \ R_{uv}^{(i)},$

$$\psi_{uv}^{(i)} = \begin{cases} 1 \text{ if } X_{iv} - X_{iu} > 0, \\ 0 \text{ otherwise} \end{cases}$$

and $R_{uv}^{(i)}$ is the rank of $|X_{iv}-X_{iu}|$ among $|X_{iv}-X_{iu}|, \cdots\cdots, |X_{nv}-X_{nu}|$. T_{uv} is seen to be the Wilcoxon signed rank statistic for the treatments u and v. Unfortunately, Hollander's statistic, while being intuitively appealing, is difficult to use. In fact, since Page, most procedures for this problem have been progressively more difficult to use. Even the likelihood ratio statistic based on sampling from normal populations is difficult to implement because of the complexity of the statistic and its distribution. Perhaps a reasonable parametric competitor to the preceding nonparametric tests would be

$$N = \frac{1}{n} \sum_{i=1}^{n} \sum_{1<u<v\le k} \sum (X_{iv} - X_{iu}) = \sum_{j=1}^{k} (2j-k-1) \overline{X}_{\cdot j}$$

where $\overline{X}_{\cdot j} = n^{-1} \sum_{i=1}^{n} X_{ij}$. Puri(1965), Hollander(1967), Barlow. et al.(1972), and Skillings and Wolfe(1978), among others, have suggested this statistic or a form of this statistic for detecting the ordered alternatives given in (2) when the ϵ_{ij}'s are normally distributed with mean 0. What I propose is an easy to use distribution-free competitor to $K, L, H,$ and N that is as intuitively appealing as either $K, L,$ or $H,$ but which is as easy to use as L. The statistic is elegant in its simplicity and has desirable asymptotic properties. In Section 2 I describe the small sample test for $k=2, 3,$ or 4 treatments. In Section 3 I establish the large sample properties of my procedure. The limiting distribution of the test statistic, the efficacy of the test, and the consistency of the procedure will be established. Asymptotic relative efficiency comparisons are made with $K, L, H,$ and N

2. The Small Sample Test

Consider the model given in (1), where the ϵ_{ij}'s represent a random sample from an absolutely continuous distribution that is symmetric about zero. In this section I construct the small sample test for $k=2, 3,$ or 4 treatments.

The following notation will be used: Let $[X_{ij}]$, where $i=1, \cdots\cdots, n$ and $j=1, \cdots\cdots, k$, be the data matrix and let $[Y_{ij}]$ be an n x $(k-1)$ dimensional transformation of the data matrix whose componets will be given shortly. Let $C_n = [|Y_{ij}|]$ be the matrix of absolute valuss of the components of $[Y_{ij}]$. Finally, define $\psi_{ij}^* = \text{sign } (Y_{ij})$ for $i=1, \cdots\cdots, n$ a and $j=1, \cdots\cdots, k-1$ and notice that whdn C_n is given, the only random quantities in $[Y_{ij}]$ are the ψ_{ij}^*'s since $Y_{ij} = \psi_{ij}^* |Y_{ij}|$. The following three theorems explain the relationship between the ψ_{ij}^*'s and C_n when H_0 is true. Theorems $1, 2,$ and 3 deal with the cases where $k=2, 3,$ and $4,$ respectively. The first part of each theorem is an immediate consequence of the fact that $(Y_{i1}, \cdots\cdots, Y_{i,k-1})'$ are jointly symmetric (see Definition 1.3.36 in Randles and Wolfe 1979) about the $(k-1)$ dimensional null vector when H_0 is true. The second part can be established using the technique of Lemma 2.4.2 in Randles and Wolfe (1979).

Theorem 1. *When $k=2$, let $Y_{i1} = X_{i2} - X_{i1}$. Then, when H_0 is true.*

$$P(\psi_{i1}^*=1)=P(\psi_{i1}^*=-1)=\frac{1}{2}$$

and ψ_{i1}^* *and* C_n *are stochastically independent for* $i=1,\ldots,n$.

Theorem 2. *When* $k=3$, *Let* $Y_{i1}=X_{i3}-X_{i1}$ *and* $Y_{i2}=2X_{i3}-X_{i1}-X_{i2}$. *Then when* H_o *is true.*

$$P(\psi_{i1}^*=a_{i1},\ \psi_{i2}^*=a_{i2})$$
$$=P(\psi_{i1}^*=a_{i1})\ P(\psi_{i2}^*=a_{i2})=1/4,$$

for $a_{i1}=\pm 1$, $a_{i2}=\pm 1$, *and the random vector* $(\psi_{i1}^*,\ \psi_{i2}^*)$ *and* C_n *are stochastically independent for* $i=1,\ldots,n$.

Theorem 3. *When* $k=4$, *Let* $Y_{i1}=X_{i3}-X_{i1}$, $Y_{i2}=X_{i4}-X_{i2}$, *and* $Y_{i3}=X_{i4}-X_{i3}+X_{i2}-X_{i1}$. *Then, when* H_o *is true,* ψ_{i1}^*, ψ_{i2}^*, *and* ψ_{i3}^* *are mutually (unconditionally) independent with* $P(\psi_{ij}^*=1)=P(\psi_{ij}^*=-1)=\frac{1}{2}$ *and the random vector* $(\psi_{i1}^*,\ \psi_{i2}^*,\ \psi_{i3}^*)$ *and* C_n *are also independent for* $i=1, 2,\ldots, n$.

Before leaving these theorems, note that the transformations used in them were selected because they yield random variables that are jointly symmetric about the null vector when H_o is true and that have nonnegative expectation, if first moments exist, when (2) is true. Because transformations guaranteeing this joint symmetry do not exist, in general, for randomized complete block designs with $k\geqq 5$, the results in this article are valid only when the number of treatments is less than five. This is not a serious limitation because many applications involve fewer than five treatments.

The following theorem, which is established by noting that the signs are mutually independent and that the vector of signs is independent of the vector of ranks, is needed to obtain the H_o distribution of the test statistic. This theorem is interesting in itself, because it shows that a collection of n continuous random variables having possibly different distributions that are jointly symmetric about the n-dimensional null vector is sufficient for constructing a test statistic whose distribution is the same as that of the Wilcoxon signed rank statistic.

Theorem 4. *Let* $(Z_1,\ldots,Z_n)$ *be n continuous random variables that are jointly symmetric about the n-dimensional null vector. Let* $\psi_i^*=sign(Z_i)$ *and* R_i^+ *be the rank of* $|Z_i|$ *among* $|Z_1|,\ldots,|Z_n|$ *is the same as that of the Wilcoxon signed rank statistic based on a random sample of size n from a continuous distribution that is symmetric about zero.*

The test procedure is described when the sample size n is not too large. Simply pool the $n(k-1)$ transformed observations, Y_{ij}, into one large sample. Let R_{ij}^+ *be the rank of* $|Y_{ij}|$ in the pooled collection of $n(k-1)$ absolute values and let.

$$W=\sum_{i=1}^{n}\sum_{j=1}^{k-1}\psi_{ij}^*R_{ij}^+$$

By Theorem 4, the null hypothesis distribution of W is the same as that of the regular Wilcoxon signed rank statistic, and hence is readily available. I see that if H_o is true W should be close to its mean, zero, and if the alternative (2) is true, W should be significantly larger than zero. This latter point will be verified when I discuss the consistency class for this procedure in the next section. I turn my attention now to Section 3 and investigate the large sample properties of the proposed procedure.

3. Large Sample History

Before proceeding to a discussion of the efficacy of the test, several comments are appropriate. As a consequence of Theorem 4, I see that the null hypothesis limiting distribution of $W/\sigma(w)$, where

$$\sigma^2(W) = (1/6)\ [n(k\text{-}1)]\ [n(k\text{-}1)+1]\ [(2n(k\text{-}1)+1]$$

and $k=2, 3, 4$, is the zero mean and unit variance normal distribution. Because W can be written as

$$W = 2[\sum_{j=1}^{k-1} \sum_{i=1}^{n} \psi(Y_{ij}) + \sum_{\alpha=1}^{k-1} \sum_{1 \leq i \leq j \leq n} \psi(Y_{i\alpha}+Y_{j\alpha}) + 2\sum\sum_{1 \leq \alpha \leq \beta \leq k-1} \sum_{i=1}^{n} \sum_{j=1}^{n} \psi(Y_{i\alpha}+Y_{j\beta})$$
$$-(1/2)(k\text{-}1)\ n[(k\text{-}1)n+1], \qquad (3)$$

where $\psi(x)=1,\ 0$ as $x>0, \leqq 0$, it is seen that W is a U statistic. Consequently, $W/\sigma(w)$ has a standard normal limiting distribution under a sequence of Pitman translation alternatives. Finally, since much is known about the nature of my problem when $k=2$, I confine my results in the remainder of this section to the cases where $k=3$ or 4.

I obtain the efficacy of the tests based on W, denoted by $eff(W)$. The efficacy of each of K, L, H, and N, denoted $eff(K)$, $eff(L)$, $eff(H)$, and $eff(N)$, respectively, is well known (see Hollander 1967), so I will be able to make Pitman asymptotic relative efficiency comparisons among the five statistics in this study. Each of these efficiencies will be denoted $ARE(A, E)$, where A and B are two of the statistics of interest, and are obtained when the ε_{ij}'s represent a random sample from each of the following distributions: the uniform distribution on the interval $(-1/2,\ 1/2)$, denoted $U(-1/2,\ 1/2)$; the standard normal distribution, denoted $N(0,\ 1)$; and the double exponential distribution, denoted DE. These three distributions were chosen as representatives of light, medium, and fairly heavy tailed distributions, respectively. The sequence of alternatives converging to $H_0 : \tau_1 = \cdots = \tau_k$, are of the form

$$H_{1n} : \tau_{1n} \leqq \tau_{2n} \leqq \cdots \leqq \tau_{kn},$$

where $\tau_{in} = \tau_i \theta_n$ for $j=1, \ldots, k$, $\Sigma_1^k \tau_j = 0$, $\theta_n = n^{-\frac{1}{2}}$ and at least one of the inequalities is strict. Then, using (3), I find that for $k=3$,

$$eff(W) = \tfrac{3}{2}[(\tau_2-\tau_1) \int g_1^2(x)dx + (2\tau_3-\tau_1-\tau_2) \int g_2^2(x)\,dx$$
$$+2(\tau_3-\tau_1) \int g_1(x)\ g_2(x)\,dx]^2 \qquad (4)$$

where $g_1(\cdot)$ (respectively, $g_2(\cdot)$) is the probaility density function (pdf) of $\varepsilon_{12}-\varepsilon_{11}$ (respectively, $2\varepsilon_{13}-\varepsilon_{11}-\varepsilon_{12}$). Similarly, When $k=4$,

$$eff(W) = 4/9[2(\tau_4+\tau_3-\tau_2-\tau_1) \int g_1^2(x)dx + (\tau_4-\tau_3+\tau_2-\tau_1) \int g_3^2(x)\,dx$$
$$+(3\tau_4-\tau_3+\tau_2-3\tau_1) \int g_1(x)g_3(x)\,dx]^2 \qquad (5)$$

where $g_1(\cdot)$ is as before and $g_3(\cdot)$ is the pdf of $\varepsilon_{14}-\varepsilon_{13}$.

The desired ARE's, which are complicated continuous functions of the treatment effects, can now be computed using(4) and (5) and the efficacies for the other statistics as given in Hollander(1967). To help interpret the ARE's, Table 1 contains a summary of their ranges. The minimum and maximum values, which are respectively, the left and right endpoints of the ranges, were obtained by straightforward calculus calculations subject to the constraints given in (2) and $\Sigma\tau_j=0$. These computations show that $ARE(W, L)$ is a minimum when $k=3$ *if* $\tau_1=\tau_2=-\delta/2$ and $\tau_3=\delta$, and is a maximum when $\tau_1=-\delta$ and $\tau_2=\tau_3=\delta/2$. Similar computations show $ARE(W, L)$ is a minimum when $k=4$, if $\tau_1=\tau_2$ and $\tau_3=\tau_4$ and is a maximum when $k=4$ if $\tau_1=\tau_2$ and $\tau_3=\tau_4$ and is a maximum when $\tau_2=\tau_3$. Since $ARE(W, N)$ and $ARE(W, H)$ are linear functions of $ARE(W, L)$ for $k=3$ or 4, the minimum and maximum values of these ARE's occur for precisely the same configurations of alternatives hypothesis treatment effects. Also, since the $ARE(L, K)>1$ for $k>2$, only the efficiencies involving Page's L are summarized.

By inspecting Table I see that if the user knows the nature of the distribution of the error terms, when for a certain configuration of the treatment effects, it is possible to choose a statistic from among those in this study that is asymptotically best in the Pitman sense. Unfortunately, a practitioner rarely can be this explicit, and so it can be difficult to choose the "best" statistic from among those in this study. At first glance, Hollander's H is an intuitively appealing statistic with good overall Pitman ARE's when compared to W, L, and N. However, because the variance of H is a function of $P(\varepsilon_1<\varepsilon_2+\varepsilon_3-\varepsilon_4,\ \varepsilon_1<\varepsilon_5+\varepsilon_6-\varepsilon_7)$, where the ε_i's may be considered independent and identically distributed error terms in my model (1), a small sample distribution-free test is not possible and a large sample test based on a standardized form of H would be inconvenient to use. Therefore, the user probably needs to select from among W, L, and N. From Table I see that under certain conditions N is a reasonable choice; however, logical considerations may indicate that n should not be used. In these cases, we suggest that W is a viable alternative to L. The small sample distribution-free test based on W is easy to construct and tables for the null distribution of W are readily available even for moderate sample sizes.

The null mean and variance of W are easy to compute by their usual formulas so that it is easy to standardize W to obtain an asymptotically $N(0, 1)$ statistic. W is asymptotically "better"

Table. Asymptotic Relative Efficiencies for Three or Four Treatments.

		U(-1/2, 1/2)	N(0, 1)	DE
	ARE (W, L)	(.945, 1.289)	(1.100, 1.492)	(.861, 1.171)
k=3	ARE (W, N)	(.708, 967)	(.788, 1.068)	(.968, 1.318)
	ARE (W, H)	(.793, 1.083)	(.818, 1.109)	(.815, 1.109)
	ARE (W, L)	(.821, 1.234)	(.918, 1.368)	(.703, 1.018)
k=4	ARE (W, N)	(.657, .987)	(.701, 1.045)	(.843, 1.222)
	ARE (W, H)	(.733, 1.102)	(.724, 1.079)	(.704, 1.020)

(again in the pitman sense) than L except when $k=4$ and the error terms are from the double exponential distribution. In particular, if $k=3$ and $\tau_1=-\delta$, $\tau_2=0$, and $\tau_3=\delta$, then $ARE(W, L)$

$=1.110$ if the error terms are from $U(-1/2, 1/2)$,

$=1.288$ if the error terms are from $N(0, 1)$,

$= 1.010$ if the error terms are from DE.

If $k=4$ and $\tau_1=-\delta$, $\tau_2=-\delta/3$, $\tau_3=\delta/_3$ and $\tau_4=\delta$ then

$ARE(W, L)$

$= 1.059$ if the error terms are from $U(-1/2, 1/2)$,

$= 1.177$ if the error terms are from $N(0, 1)$,

$= .885$ if the error terms are from DE.

4. Conclusion

A composite Wilcoxon signed rank statistic, based on a set of transformed variables possessing joint symmetry, provides a viable distribution-free competitor to Page's L (as well as other statistics) in the randomized complete block design where the user wishes to detect ordered alternatives. The statistic has several advantages that enhance its attractiveness to the user: ease of computation, tables for the small sample test that are readily available, a well-known large sample normal approximation, and good power properties in both the small smaple and large sample settings. Also, the distribution of the statistic is unchanged if the block effects are taken to be random, since they cancel out when the contrasts are formed. The primary disadvantages of the statistic are that it cannot be applied in designs involving more than four treatments and that its power function depends in a complicated way on the magnitude and placement of the treatment effects.

References

1. Barlow, R. E., Bartholomew, D. J., Bremner, J. M., and Brunk H. D. (1972), *Statistical Inference Under Order Restrictions,* London: John Wiley.

2. Hollander, M. (1967), "Rank Tests for Randomized Blocks When the Alternatives Have a Prior Ordering," *Annals of Mathematical Statistics,* 38, 867-877.

3. Jonckheere, A. R. (1954), "A Test of Significance for the Relation Between m Rankings and k Ranked Categories," *British Journal of Statistica Psychology 7,* 93-100.

4. Lehmann, E. L. (1951), "Consistency and Unbiasedness of Certain Nonparametric Tests," *Annals of Mathematical Statistics* 22, 165-179.

5. Odeh, R. E. (1977), "The Exact Distribution of Page's L-Statistic in the Two-Way Layout," *Communications in Statistics,* Ser. B, 6, 49-62.

6. Page, E. B. (1963), "Ordered Hypotheses for Multiple Treatments: A Significance Test for Linear Ranks, *Journal of the American Statistical Association,* 58, 216-230.

7. Puri, M. L. (1965), "Some Distribution-free k-Sample Rank Tests of Homogeneity Against Ordered Alternatives," *Communications on Pure and Applied Mathematics,* 18, 51-63.

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

9. Skillings, J. H., and Wolfe, D. A. (1978), "Distribution-Free Tests for Ordered Alternatives in a Randomized Block Design," *Journal of the American Statistical Association,* 73, 427-431.