Journal of the Korean
Statistical Society
Vol. 26, No. 1, 1997

Nonparametric Procedures for Comparing Ordered Treatment Effects with a Control in a Randomized Block Design

Dong Hoon Lim [1]

ABSTRACT

In this paper we are concerned with comparing ordered treatment effects with a control in a randomized block design with multiple observations per cell. Two nonparametric procedures for detecting which treatments are better than the control are proposed and compared. An example is given and the results of a Monte Carlo power study are discussed.

Key Words : Isotonic regression; Monte Carlo study; Nonparametric treatment versus control; Ordered treatment effects; Randomized block design.

1. INTRODUCTION

In experiments typically designed to assess the toxicity or other effects of substances such as drugs, pesticides, and various environmental pollutants on

[1]Department of Statistics, Gyeongsang National University, Chinju, 660-701, Korea.

laboratory animals, several treatments are compared with a control. In nonparametric procedures for a one-way design, Steel (1959) and Dunn (1964) proposed multiple comparison rank tests for the general setting in which no information about the pattern of treatment effects is available. Shirley (1977) suggested a nonparametric version of Williams' (1971, 1972) test procedure for the situation where the experimenter believes a prior that if there was a response to the substance the treatment effects would be monotonically ordered. Williams (1986) further considered a modification that increases the power of Shirley's test. On the other hand, House (1986) proposed a nonparametric version of Williams's test which is appropriate for a randomized block design. In recent, Lim and Wolfe (1997) proposed test procedures for comparing several treatment effects with a control under the more restrictive information about an umbrella pattern of these treatment effects. These procedures have relied on assumption which requires exactly one observation per cell.

Often the experimenter has more than one observation per cell. The purpose of this paper is to give two nonparametric procedures for comparing several treatments with a control in a randomized block design with multiple observations per cell, when the treatment effects have a monotonic ordering. (These procedures will be called Procedure I and Procedure II.)

Suppose $X_{ij\ell}, i = 1, \ldots, m, j = 0, 1, \ldots, k, \ell = 1, \ldots, n$ for each i, j denotes a random sample of size n from $F(x - \beta_i - \theta_j)$, F continuous. We consider $\beta_1, \ldots, \beta_m$ as the nuisance location parameters. We let θ_0 assume the effect of the control and the other $\theta_1, \ldots, \theta_k$ represent treatment effects. The problem of interest is to test the null hypothesis $H_0 : \theta_0 = \theta_1 = \cdots = \theta_k$ against the alternative hypothesis $H_1 : \theta_j > \theta_0$ under the order restriction $\theta_1 \leq \ldots \leq \theta_k$.

Section 2 and Section 3 of this paper describe Procedure I and Procedure II, respectively. Section 4 gives an example, and Section 5 discusses the results of a Monte Carlo simulation investigation of the relative powers of the two procedures for a variety of ordered treatment effects configurations. Section 6 contains some concluding remarks.

2. PROCEDURE I

Let $R_{ij\ell}^{(k)}$ be the rank of $X_{ij\ell}$ among the $N_k = n(k+1)$ observations of the ith block and let $\bar{R}_j^{(k)} = (1/mn)\sum_{i=1}^{m}\sum_{\ell=1}^{n} R_{ij\ell}^{(k)}$ be the average rank of the jth treatment over the m blocks. If ties occur within a block, mean ranks are

used. Let $\hat{R}_1^{(k)} \leq \ldots \leq \hat{R}_k^{(k)}$ be the isotonic regression of $\bar{R}_1^{(k)}, \ldots, \bar{R}_k^{(k)}$ under the restriction $\theta_1 \leq \ldots \leq \theta_k$. The $\hat{R}_j^{(k)}$'s are obtained by the well-known amalgamation procedure (Barlow *et al.* (1972), Robertson *et al.*(1988)).

If there are ties in the data, we propose rejecting H_0 for large values of

$$T_j = (\hat{R}_j^{(k)} - \bar{R}_0^{(k)})[V_k(\frac{2}{mn})]^{-1/2}, \tag{2.1}$$

where

$$V_k = \frac{N_k(N_k+1)}{12} - \frac{\sum_{i=1}^{m}\sum_{\Gamma=2}^{g_i} t_{i\Gamma}(\Gamma^3 - \Gamma)}{12m(N_k - 1)}$$

with g_i the size of the greatest tie in the ith ranking and $t_{i\Gamma}$ the number of ties of size Γ. In the absence of ties, $V_k = N_k(N_k+1)/12$. In particular, the test based on T_j when $m = 1$ is Shirley (1977) test in the one-way design with equal treatment replications.

Miller (1966) showed that the k-fold vector of pair differences $(\bar{R}_1^{(k)} - \bar{R}_0^{(k)}, \bar{R}_2^{(k)} - \bar{R}_0^{(k)}, \ldots, \bar{R}_k^{(k)} - \bar{R}_0^{(k)})$ follows an asymptotic normal distribution with the same correlation structure as that of $(Z_1 - Z_0, Z_2 - Z_0, \ldots, Z_k - Z_0)$ where the $Z_i, i = 0, \ldots, k$, are independent and normally distributed with zero mean and unit variance. So, the statistic T_k converges in distribution to the statistic

$$\max_{1 \leq u \leq k} \sum_{j=u}^{k} (Z_j - Z_0)/(k - u + 1).$$

Therefore, the distribution of T_j is approximately that of the statistic $\bar{t}_j$ with infinite degrees of freedom, proposed by Williams (1971). The critical values $t(\alpha; j)$ of $\bar{t}_j$ in Table 1 are taken from Williams (1971).

Our procedure starts by comparing T_k with $t(\alpha; k)$. If $T_k \geq t(\alpha; k)$, we conclude that there is a response at treatment level k. Then the statistic T_{k-1} is calculated in the same way and compared to its critical value $t(\alpha; k-1)$. This procedure continues until a nonsignificant T_j is obtained for some treatment j. The conclusion is that there is a significant effect for treatment levels $j+1$ and above and no significant evidence of a treatment effect for levels j and below.

Table 1. Values of $t(\alpha; j)$

d.f.		j =number of treatments (excluding a control)									
ν	α	1	2	3	4	5	6	7	8	9	10
∞	0.01	2.326	2.366	2.377	2.382	2.385	2.386	2.387	2.388	2.389	2.389
∞	0.05	1.645	1.716	1.739	1.750	1.756	1.760	1.763	1.765	1.767	1.768

3. PROCEDURE II

We recall that the observations in Procedure I were replaced by their ranks and these ranks are unchanged as each treatment level is tested. The modification of the Procedure I results from reranking the observations at each stage of the sequential test procedure, after excluding all observations at those higher treatment levels at which significant evidence of an effect has already been established. We begin by defining statistic derived from the first $j+1$ treatments $0, 1, \ldots, j$. Suppose $N_j = n(j+1)$ observations in these treatments of the ith block are ranked in increasing order from 1 to N_j. Let $\hat{R}_1^{(j)} \leq \ldots \leq \hat{R}_j^{(j)}$ be the isotonic regression of $\bar{R}_1^{(j)}, \ldots, \bar{R}_j^{(j)}$ under the order restriction $\theta_1 \leq \ldots \leq \theta_j$.

The modification to the statistic given by (1) now proposed replaces $\hat{R}_j^{(k)}$ by $\hat{R}_j^{(j)}$, $\bar{R}_0^{(k)}$ by $\bar{R}_0^{(j)}$ and V_k by V_j to give

$$\bar{T}_j = (\hat{R}_j^{(j)} - \bar{R}_0^{(j)})[V_j(\frac{2}{mn})]^{-1/2}. \tag{3.1}$$

It is easy to see that $\bar{T}_j$ reduces to House (1986) statistic when all $n = 1$. This modified test procedure still controls at α the probability of declaring an effect at treatment level j when in reality there is no effect at treatments $1, \ldots, j$.

This procedure can be applied in a stepwise manner as follows : At first we apply the statistic $\bar{T}_k$ to all treatments. If $\bar{T}_k \geq t(\alpha; k)$, we conclude that treatment k is better than the control. We then rerank all observations at the remaining treatments after excluding the significant treatment k and apply $\bar{T}_{k-1}$ on the resulting treatments and compared to its critical value $t(\alpha; k-1)$ and so on.

4. EXAMPLE

As an example, consider the artificial data in Table 2 to determine the treatment levels at which there is evidence of a response. The relationship between the dose of a drug that increases a minimum blood pressure and the actual amount in diastolic blood pressure was investigated in a laboratory experiment. The data consist of four blocks and four treatments. Each block is composed of three rabbits drawn from the same litter. Each of three rabbits in each block received in random order four different dose levels of drug, which

a suitable interval between each drug adiministration. The following data on blood pressure are collected and the ranks are shown in parentheses.

Table 2. Artificial Data

	dosage			
	0.0	0.2	0.4	0.6
block				
	70 (2)	78 (5)	94 (8)	100 (11)
1	72 (3.5)	80 (6)	96 (10)	107 (12)
	67 (1)	72 (3.5)	87 (7)	95 (9)
	82 (5)	90 (6)	100 (10)	115 (12)
2	62 (1)	65 (3)	96 (8.5)	105 (11)
	64 (2)	68 (4)	95 (7)	96 (8.5)
	62 (1)	67 (2)	80 (7)	92 (9)
3	75 (5)	69 (3)	82 (8)	98 (11)
	73 (4)	78 (6)	95 (10)	106(12)
	77 (5)	85 (6.5)	93 (11)	102 (12)
4	77 (4)	88 (8)	90 (9)	92 (10)
	75 (1)	58 (2)	70 (3)	85 (6.5)

First the Procedure I based on the statistic (2.1) is applied to the data of Table 2. We compute the average ranks, obtaining $\bar{R}_0^{(3)}$ =2.875, $\bar{R}_1^{(3)}$ =4.583, $\bar{R}_2^{(3)}$ =8.208 and $\bar{R}_3^{(3)}$ =10.333. Note that for the data the correction for ties is

$$3(2^3 - 2)/(12 \cdot 4 \cdot 11) = 0.0341.$$

To test for an effect at dosage 0.6, we obtain

$$\begin{aligned} T_3 &= (10.333 - 2.875)[\{(12 \cdot 13)/12 - 0.0341\}\{2/(4 \cdot 3)\}]^{-1/2} \\ &= 5.074. \end{aligned}$$

The approximate 1% and 5% significance critical values of T_3 are, respectively, $t(0.01;3)$ =2.377 and $t(0.05;3)$=1.739, obtained from Table 1. There is a significant effect at dosage 0.6. Now examining dosage 0.4, we obtain

$$\begin{aligned} T_2 &= (8.208 - 2.875)[\{(12 \cdot 13)/12 - 0.0341\}\{2/(4 \cdot 3)\}]^{-1/2} \\ &= 3.628. \end{aligned}$$

The critical value at the 5% level is 1.716, so there is evidence of an effect at dosage 0.4. To test for an effect at dosage 0.2, we obtain

$$\begin{aligned} T_1 &= (4.583 - 2.875)[\{(12 \cdot 13)/12 - 0.0341\}\{2/(4 \cdot 3)\}]^{-1/2} \\ &= 1.162. \end{aligned}$$

The 5% critical value for T_1 is 1.645, so there is not significant at dosage 0.2. Hence we may claim evidence for an effect at the dosages greater than or equal to dosage 0.4.

The same data are used for the Procedure II based on (3.1) in order to determine which treatment levels are better than the zero-control. To test an effect at dosage 0.6, the statistic (3.1) is the same as the statistic (2.1), so the Procedure II has the same result as the Procedure I. Examining dosage 0.4, we rerank the data for the three lowest dosages. This gives average ranks of $\bar{R}_0^{(2)}$ =2.875, $\bar{R}_1^{(2)}$ =4.458 and $\bar{R}_2^{(2)}$ =7.667 for the dosage 0.0, 0.2 and 0.4, respectively. The correction for ties is

$$(2^3 - 2)/(12 \cdot 4 \cdot 8) = 0.0156$$

and

$$\begin{aligned} \bar{T}_2 &= (7.667 - 2.875)[\{(9 \cdot 10)/12 - 0.0156\}\{2/(4 \cdot 3)\}]^{-1/2} \\ &= 5.093. \end{aligned}$$

The 5% critical value for $\bar{T}_2$ is 1.716, so there is a significant effect at dosage 0.4. Reranking the two lowest dosages give average ranks of $\bar{R}_0^{(1)}$ =2.708 and $\bar{R}_1^{(1)}$ =4.292 for the dosage 0.0 and 0.2, respectively. The correction for ties is

$$(2^3 - 2)/(12 \cdot 4 \cdot 5) = 0.025$$

and

$$\begin{aligned} \bar{T}_1 &= (4.292 - 2.708)[\{(6 \cdot 7)/12 - 0.025\}\{2/(4 \cdot 3)\}]^{-1/2} \\ &= 2.081, \end{aligned}$$

which is a significant effect at the 5% level. Thus, there are significant effects at all three treatment levels at 5% level.

5. MONTE CARLO POWER STUDY

We conduct a Monte Carlo study to compare the estimates of power of the Procedure I based on the statistic T_j (2.1) with the Procedure II based on the statistic $\bar{T}_j$ (3.1) for a variety of different ordered treatment effects. In this study, we consider three randomized block designs with k=3, m=4 n=3, and k=4, m=5, n=3, and k=5, m=6, n=4.

For each of these settings, the normal and exponential random variates were generated using the IMSL routine GGNML and GGEXN, respectively. In each case, we used 10,000 replications in obtaining the various estimates of power for the two tests considered in this study are presented in Table 3. The estimated power π_{j0} represents the probability of declaring a significant effect at treatment j at $\alpha = 0.05$. The designated alternative configurations correspond to values of $\theta_{10} = \theta_1 - \theta_0, \ldots, \theta_{k0} = \theta_k - \theta_0$.

We observe from the simulation results that the Procedure II is generally preferable to the Procedure I. In particular, the Procedure II is superior to the Procedure I in the extreme case $\theta_{10} = \ldots = \theta_{j-10} \leq \theta_{j0} \leq \ldots \leq \theta_{k0}$. This is not surprising since in the Procedure I, V_k in the statistic (2.1) is an overestimate of the variance of the ranks contributing to the numerator of (1). However, it is less powerful at detecting the lowest treatment effects for the case $\theta_{10} < \theta_{20} < \ldots < \theta_{k0}$.

6. CONCLUDING REMARKS

In this paper we are concerned with comparing ordered treatment effects with a control in a randomized block design with multiple observations per cell. A few comparions of the Porcedure I and the Procedure II for the problem of detecting which treatments are better than the control, are ordered as follows.

(i) The Procedure I is easier to use than the Procedure II, as illustrated by the example in Section 4.

(ii) For the problem of testing an effect at the highest treatment level, these procedures are the same, but at the low treatment levels, the Procedure II will usually be greater than the Procedure I because V_j in (3.1) is substantially less than V_k in (2.1).

(iii) The main finding of this simulation study is that the Procedure II is generally more powerful than the Procedure I, although it is more cumbersome to apply.

Table 3. Monte Carlo power estimates for k=3,m=4 and n=3

(a) Normal

θ_{10}	θ_{20}	θ_{30}		T_j	$\bar{T}_j$
0.0	0.5	1.0	π_{10}	–	–
			π_{20}	0.244	0.262
			π_{30}	0.715	0.715
0.0	1.0	1.0	π_{10}	–	–
			π_{20}	0.625	0.666
			π_{30}	0.775	0.775
0.5	0.75	1.0	π_{10}	0.227	0.162
			π_{20}	0.473	0.497
			π_{30}	0.739	0.739

(b) Exponential

θ_{10}	θ_{20}	θ_{30}		T_j	$\bar{T}_j$
0.0	0.5	1.0	π_{10}	–	–
			π_{20}	0.398	0.446
			π_{30}	0.892	0.892
0.0	1.0	1.0	π_{10}	–	–
			π_{20}	0.829	0.849
			π_{30}	0.897	0.897
0.5	0.75	1.0	π_{10}	0.371	0.368
			π_{20}	0.691	0.737
			π_{30}	0.913	0.913

Table 3. Monte Carlo power estimates for k=4,m=5 and n=3 (continued)

(a) Normal

θ_{10}	θ_{20}	θ_{30}	θ_{40}		T_j	T_j
0.0	0.0	0.5	1.0	π_{10}	–	–
				π_{20}	–	–
				π_{30}	0.296	0.315
				π_{40}	0.791	0.791
0.0	0.0	1.0	1.0	π_{10}	–	–
				π_{20}	–	–
				π_{30}	0.725	0.763
				π_{40}	0.846	0.846
0.0	0.5	0.75	1.0	π_{10}	–	–
				π_{20}	0.264	0.234
				π_{30}	0.562	0.589
				π_{40}	0.829	0.829
0.25	0.5	0.75	1.0	π_{10}	0.100	0.031
				π_{20}	0.286	0.255
				π_{30}	0.569	0.597
				π_{40}	0.830	0.830

(b) Exponential

θ_{10}	θ_{20}	θ_{30}	θ_{40}		T_j	T_j
0.0	0.0	0.5	1.0	π_{10}	–	–
				π_{20}	–	–
				π_{30}	0.484	0.538
				π_{40}	0.947	0.947
0.0	0.0	1.0	1.0	π_{10}	–	–
				π_{20}	–	–
				π_{30}	0.917	0.933
				π_{40}	0.956	0.956
0.0	0.5	0.75	1.0	π_{10}	–	–
				π_{20}	0.439	0.465
				π_{30}	0.790	0.826
				π_{40}	0.958	0.958
0.25	0.5	0.75	1.0	π_{10}	0.180	0.150
				π_{20}	0.480	0.513
				π_{30}	0.810	0.846
				π_{40}	0.964	0.964

Table 3. Monte Carlo power estimates for k=5,m=6 and n=4 (continued)

(a) Normal

θ_{10}	θ_{20}	θ_{30}	θ_{40}	θ_{50}		T_j	$\bar{T}_j$
0.0	0.0	0.0	0.5	1.0	π_{10}	–	–
					π_{20}	–	–
					π_{30}	–	–
					π_{40}	0.440	0.454
					π_{50}	0.930	0.930
0.0	0.0	0.0	1.0	1.0	π_{10}	–	–
					π_{20}	–	–
					π_{30}	–	–
					π_{40}	0.912	0.926
					π_{50}	0.963	0.963
0.0	0.0	0.5	0.75	1.0	π_{10}	–	–
					π_{20}	–	–
					π_{30}	0.421	0.350
					π_{40}	0.772	0.784
					π_{50}	0.950	0.950
0.0	0.25	0.5	0.75	1.0	π_{10}	–	–
					π_{20}	0.140	0.047
					π_{30}	0.428	0.354
					π_{40}	0.771	0.786
					π_{50}	0.951	0.951
0.25	0.5	0.75	1.0	1.5	π_{10}	0.125	0.017
					π_{20}	0.398	0.231
					π_{30}	0.747	0.720
					π_{40}	0.949	0.955
					π_{50}	0.999	0.999

(b) Exponential

θ_{10}	θ_{20}	θ_{30}	θ_{40}	θ_{50}		T_j	$\bar{T}_j$
0.0	0.0	0.0	0.5	1.0	π_{10}	–	–
					π_{20}	–	–
					π_{30}	–	–
					π_{40}	0.689	0.743
					π_{50}	0.996	0.996
0.0	0.0	0.0	1.0	1.0	π_{10}	–	–
					π_{20}	–	–
					π_{30}	–	–
					π_{40}	0.991	0.994
					π_{50}	0.997	0.997
0.0	0.0	0.5	0.75	1.0	π_{10}	–	–
					π_{20}	–	–
					π_{30}	0.639	0.647
					π_{40}	0.938	0.955
					π_{50}	0.997	0.997
0.0	0.25	0.5	0.75	1.0	π_{10}	–	–
					π_{20}	0.248	0.182
					π_{30}	0.674	0.683
					π_{40}	0.944	0.961
					π_{50}	0.997	0.997
0.25	0.5	0.75	1.0	1.5	π_{10}	0.205	0.169
					π_{20}	0.584	0.646
					π_{30}	0.904	0.945
					π_{40}	0.992	0.997
					π_{50}	1.000	1.000

REFERENCES

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

(2) Dunn,O.J.(1964). Multiple Comparisons using Rank Sums. *Technometrics,* **6**, 241-252.

(3) House,D.E.(1986). A Nonparametric Version of Williams' Test for a Randomized Block Design, *Biometrics,* **42**, 187-190.

(4) Lim,D.H. and Wolfe,D.A.(1997). Nonparametric Tests for Comparing Umbrella Pattern Treatment Effects with a Control in a Randomized Block Design, to appear in *Biometrics.*

(5) Miller,R.G.,Jr.(1966).*Simultaneous Statistical Inference,* New York : McGraw-Hill.

(6) Robertson,T., Wright,F.T., and Dykstra,R.L.(1988). *Order Restricted Statistical Inference*, New York : Wiley

(7) Shirley,E.(1977). A Nonparametric Equivalent of Williams's Test for Contrasting Increasing Dose Levels of a Treatment. *Biometrics,* **33**, 386-389.

(8) Steel,R.G.D.(1959). A Multiple Comparison Rank Sum Test: Treatments versus Control. *Biometrics,* **15**, 560-572.

(9) Williams,D.A.(1971). A Test for Differences Between Treatment Means when Several Dose Levels are Compared with a Zero-dose Control. *Biometrics,* **27**, 103-117.

(10) Williams,D.A.(1972). The Comparison of Several Dose Levels with a Zero-dose Control. *Biometrics,* **28**, 519-531.

(11) Williams,D.A.(1986). A Note on Shirley's Nonparametric Test for Comparing Several Dose Levels with a Zero-dose Control. *Biometrics,* **42**, 183-186.