Advances and Applications in Statistics
© 2014 Pushpa Publishing House, Allahabad, India
Published Online: March 2014
Available online at http://pphmj.com/journals/adas.htm
Volume 39, Number 1, 2014, Pages 1-23

A TEST FOR MAIN EFFECTS WHEN OBSERVATIONS ARE RANDOMLY RIGHT CENSORED

A. Shanubhogue and D. P. Raykundaliya*

Department of Statistics
Sardar Patel University
Vallabh Vidyanagar 388 120, India
e-mail: dp_raykundaliya@spuvvn.edu

Abstract

In this article, we generalize the Gore-test (Gore [8]), along the lines of Breslow [3] and Patel [16], for randomized block design when the observations are subject to arbitrary right censorship. The distributions of censoring variables are allowed to vary from block to block. The asymptotic distribution of the proposed statistic under the null hypotheses as well as Pitman type alternative is shown to be chi-square. The asymptotic power study is made when error distributions are logistic, Laplace, normal, Pareto of first kind, exponential, generalized exponential and Weibull. It is seen that the test performed well for most of the positive error distributions.

1. Introduction

Analysis of variance is a common method for analyzing continuous

Received: August 24, 2013; Revised: November 20, 2013; Accepted: December 10, 2013

2010 Mathematics Subject Classification: 62N05, 62K99.

Keywords and phrases: randomized block design, censoring, Pitman type alternative hypothesis, asymptotic chi-square, logistic, Laplace, normal, Pareto of first kind, exponential, generalized exponential, Weibull.

*Corresponding author

Communicated by Kuan-Fu Shen

interval-scaled data in block designs but it is not appropriate when observations are ordinally ranked, the relation between observed data and their importance for the observational unit is not linear, or block effects are not additive. Because the effects are apparently not additive, only the order of the data within the block can be meaningfully interpreted, differences in expectation are less relevant than differences in tendency. Friedman [5] has considered block ranks as a basis for analysis. Wittkowski [18] considered the test procedure for the global hypothesis of no tendency in treatment effects in two-way layouts with arbitrary tied and missing observations. Several non-parametric procedures have been proposed in the literature for testing a null hypothesis about main effects in two-way layouts by Bhapkar and Gore [2], Brown and Mood [4], Hodges and Lehmann [10] and Lehmann [11].

Let us consider a linear model

$$X_{ijl}^{0} = \mu + \alpha_i + \beta_j + \varepsilon_{ijl}, \tag{1.1}$$

where μ is the overall effect, α_i $(i = 1, 2, ..., r)$ effect of ith treatment, β_j $(j = 1, 2, ..., c)$ effect of jth block and ε_{ijl} $(l = 1, 2, ..., n_{ij}, n_{ij} \geq 1)$ is a random error component. We assume that ε_{ijl} (for all i, j and l) are independent, identically distributed random variables with common distribution function $F_{ij}(x)$, $l = 1, 2, ..., n_{ij}$; $i = 1, 2, ..., r$, $j = 1, 2, ..., c$ with median zero. Let (X_{ijl}, δ_{ijl}) denote lth observation on ith treatment in jth block, where $X_{ijl} = \min(X_{ijl}^{0}, Z_{ijl})$ and $\delta_{ijl} = 1$ if $X_{ijl} = X_{ijl}^{0}$ and zero otherwise. The uncensored observations in (i, j)th cell, i.e., X_{ijl}^{0}, $l = 1, 2, ..., n_{ij}$ are assumed to be distributed like a continuous random variable with distribution function $F_{ij}^{0}(x)$ and the censoring variables Z_{ijl}, $l = 1, 2, ..., n_{ij}$, $j = 1, 2, ..., c$ assumed to be governed by a common distribution $I_j(x)$, $j = 1, 2, ..., c$ need not be identical and may vary from block to block. Further, Z_{ijl}'s are assumed to distributed independently of X_{ijl}'s.

Without loss of generality, we assume that $\sum_{i=1}^{r} \alpha_i = \sum_{j=1}^{c} \beta_j = 0.$ Under this setup, we wish to test

$$H_0 : F_{1j}^0(x) = F_{2j}^0(x) = \cdots = F_{rj}^0(x) = F_j^0(x),\ \forall x,\ 1 \leq j \leq c, \tag{1.2}$$

where $F_j^0(x)$'s belong to the class of continuous distribution functions and may be arbitrarily different from block to block. Equivalently, $H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \cdots = \alpha_r$, i.e., $H_0 :$ the hypothesis of no "treatment effect".

Now alternative hypothesis may be stated as

$H_1 :$ At least one of the treatments is stochastically better than the rest, in all the blocks. (1.3)

In the case, censoring is not operating in a randomized block design with one observation per treatment in each block; the Friedman [5] test is most commonly used for the problem of testing equality of treatment effects. Gore [8] generalized the Friedman test to take care of the situation of multiple observations per cell. Several authors suggested the application of rank procedures to test for main effects. However, there exist fewer rank tests for detecting interaction effects in multi-factor designs. Aligned tests perform alignment of the data by subtracting the estimates of the row effects and column effects; see, for example, Sen [17] and Mansouri and Govindarajulu [13]. Further, Mansouri [14] discussed on aligned rank transform tests in linear models. However, this requires the preprocessing of the data and an estimation of the effects. Gao and Alvo [6] proposed unified non-parametric approach to perform hypothesis testing arbitrary unbalanced designs with and without interaction. Under this framework pure rank statistics can be constructed to test for main effect. Magel [12] discussed a non-parametric test for ordered alternatives of the treatment effects when the data follow a two-way layout. The standard approach for testing the treatment effects on survival study is the log rank test. Moore and Vander Lann [15] studied the estimation of parameters in randomized control trials that the outcome is time-to event in nature and subject to right censorship.

A generalization of Wilcoxon's statistic for comparing two populations has been proposed by Gehan [7] for use when observations are subject to arbitrary right censorship. Breslow [3] has extended Gehan [7] test procedure for right censorship observations to the comparison of k populations. Patel [16] generalized Friedman test for randomized block design when observations are subject to arbitrary right censorship. In this article, we generalize Gore [8] test for testing a null hypothesis of main effects in two-way layouts when observations are subject to arbitrary right censorship.

The statistical problem considered in this paper arises in clinical trials comparing several treatments under different conditions, where the observations on each patient are often time to failure or censoring.

The article is organized as follows: In Section 2, the new test statistic is proposed for main effects in two-way layouts when observations per cell are equal or unequal and are subject to arbitrary right censorship. In Section 3, limiting distribution of test statistic under both the null hypothesis and Pitman alternatives is derived. The efficacy expression of proposed statistic is also derived. Consistency of the proposed test is also discussed. The proposed test statistic is reduced to Patel's when there is only one observation per treatment per block in that way it is a generalized version of the test proposed by Patel [16]. In Section 4, test statistic proposed by Gore [8] and Patel [16] is given. In Section 5, Monte Carlo simulations are conducted to verify the small sample performance of proposed test for different cell sizes and different error distributions. Some concluding remarks are given in Section 6.

2. Test Statistics

Let us define for $i \neq i'$ a score function

$$\psi_{ii'j}(X_{ijl}, \delta_{ijl}; X_{i'jl'}, \delta_{i'jl'}) = \begin{cases} -1 & \text{if } X_{ijl} < X_{i'jl'} \text{ and } \delta_{ijl} = 1 \\ 1 & \text{if } X_{i'jl'} < X_{ijl} \text{ and } \delta_{i'jl'} = 1 \\ 0 & \text{otherwise,} \end{cases} \quad (2.1)$$

$$U_{ii'j} = \frac{1}{n_{ij} n_{i'j}} \sum_{k=1}^{n_{ij}} \sum_{l=1}^{ni'j} \psi_{ii'j}(X_{ijl}, \delta_{ijl}; X_{i'jl'}, \delta_{i'jl'}) \quad (2.2)$$

and

$$T_i = \sum_{\substack{i'=1 \\ i' \neq i}}^{r} \sum_{j=1}^{c} U_{ii'j}, \quad (2.3)$$

where $i = 1, 2, ..., r,$ $j = 1, 2, ..., c.$

For the case of equal number of observations per cell, i.e., $n_{ij} = n$ for all i and j, we propose the following statistic for H_0:

$$S_1^* = \frac{N}{r^3 c\hat{h}(\underline{F})} \sum_{i=1}^{r} T_i^2, \quad (2.4)$$

where

$$\hat{h}(\underline{F}) = \frac{1}{n^3 r^3} \sum_{j=1}^{c} \sum_{i'=1}^{r} \sum_{l'=1}^{n} \delta_{i'jl'} \left[\sum_{i=1}^{r} \sum_{l=1}^{n} \varepsilon(X_{ijl} - X_{i'jl'}) \right]^2 \quad (2.5)$$

and

$$\varepsilon(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x \leq 0. \end{cases}$$

In case of number of observations per cell is not the same, let $n_{ij} = Np_{ij}$, where $\sum_{i=1}^{r}\sum_{j=1}^{c} n_{ij} = N$, $0 < p_{ij} < 1$, $\sum_{i=1}^{r}\sum_{j=1}^{c} p_{ij} = 1$.

Further, let $q_{ij} = \frac{1}{p_{ij}} \int (1 - F_j(x))^2 d\tilde{F}_j(x)$, $q_{i.} = \sum_{j=1}^{c} q_{ij}$, $q_{..} = \sum_{i=1}^{r} q_{i.}$ and $q_{..}^{+} = \sum_{i=1}^{r} q_{i.}^{-}$, where $F_j(x) = [X_j \le x]$ and $\tilde{F}_j(x) = P[X_j \le x, \delta_j = 1]$.

For this more general set up, the proposed statistic is:

$$S_2^* = \frac{N}{r^2}\left\{ \sum_{i=1}^{r} \frac{T_i^2}{\hat{q}_{i.}} - \frac{1}{\hat{q}_{..}^{+}} \left[\sum_{i=1}^{r} \frac{T_i}{\hat{q}_{i.}} \right]^2 \right\}, \tag{2.6}$$

where

$$\hat{q}_{ij} = \frac{1}{p_{ij} n_{.j}^3} \sum_{i'=1}^{r} \sum_{l=1}^{n_{i'j}} \delta_{i'jl} \left[\sum_{i=1}^{r} \sum_{l=1}^{n_{ij}} \varepsilon(X_{ijl} - X_{i'jl'}) \right]^2 \tag{2.7}$$

and accordingly $\hat{q}_{i.}$ and $\hat{q}_{..}^{+}$ are defined.

The test procedure based on S_1^* and S_2^* is to reject H_0 in favour of H_1 if test statistic exceeds the appropriate α% critical points $s_1^*(\alpha)$ and $s_2^*(\alpha)$, respectively. For large *N*, the distribution of S_1^* and S_2^*, under H_0, may be approximated by chi-square distribution with $(r - 1)$ degrees of freedom.

3. Asymptotic Distributions of the Test Statistic

3.1. Asymptotic distribution of the test statistics under the null hypothesis

It can be easily seen that, under H_0,

$E(U_{ii'j}) = 0$ and

$$Var_0(U_{ii'j}) = \frac{n_{ij} + n_{i'j}}{n_{ij}n_{i'j}} \int_{-\infty}^{\infty} (1 - F_j^0(x))^2 (1 - I_j(x))^2 \, dF_j^0(x)$$

$$+ \frac{2}{n_{ij}n_{i'j}} \int_{-\infty}^{\infty} (1 - F_j^0(x)) F_j^0(x) \, d\widetilde{F}_j^0(x),$$

$$Cov_0(U_{ii'j}, U_{ii''j}) = \frac{1}{n_{ij}} \int_{-\infty}^{\infty} (1 - F_j^0(x))^2 (1 - I_j(x))^3 \, dF_j^0(x),$$

$E(T_i) = 0,$

$$Var_0(T_i) = \sum_{j=1}^{c} \left\{ \frac{(r-1)^2}{n_{ij}} + \Sigma_{i \neq i'} \frac{1}{n_{i'j}} \right\} \int_{-\infty}^{\infty} (1 - F_j^0(x))^2 (1 - I_J(x))^3 \, dF_j^0(x)$$

$$+ 2 \sum_{j=1}^{c} \frac{1}{n_{ij}} \Sigma_{i' \neq i} \frac{1}{n_{i'j}} \int_{-\infty}^{\infty} (1 - F_j^0(x)) F_j^0(x) (1 - I_J(x))^2 \, dF_j^0(x)$$

and

$$Cov_0(T_i, T_k) = \sum_{j=1}^{c} \left\{ \Sigma_{i' \neq i, k} \frac{1}{n_{i'j}} - \frac{(r-2)}{n_{ij}} - \frac{(r-2)}{n_{kj}} - \frac{(n_{ij} + n_{kj})}{n_{ij}n_{kj}} \right\}$$

$$\cdot \int_{-\infty}^{\infty} (1 - F_j^0(x))^2 (1 - I_J(x))^3 \, dF_j^0(x)$$

$$- \sum_{j=1}^{c} \frac{1}{n_{ij}n_{kj}} \int_{-\infty}^{\infty} (1 - F_j^0(x)) F_j^0(x) \, \widetilde{F}_j(x),$$

where $i, i', k = 1, 2, ..., r$ and $j = 1, 2, ..., c.$

Since the computations involved are routine, the details are omitted. Now, we shall obtain the asymptotic distribution of S_1^* and S_2^*, by invoking the properties of U statistics. Define

$$\underline{U_j} = (U_{12j}, U_{13j}, ..., U_{1rj}; U_{21j}, U_{23j}, ..., U_{2rj}; ...; U_{r(r-1)j})',$$

$$j = 1, 2, ..., c.$$

Since $U_{ii'j}$ for each i, i' and j is a two sample generalized U-statistic, the limiting distribution of $\sqrt{N}(\underline{U_j})$, under, H_0 as $N \to \infty$ is $r(r-1)$-variate normal with null mean vector and appropriate covariance matrix. Therefore, limiting distribution $\sqrt{N}(\underline{U}) = \sqrt{N} \sum_{j=1}^{c} (\underline{U_j})$ is also multivariate normal. Let $\underline{T} = (T_1, T_2, ..., T_r)'$. Notice that $\underline{T}$ can be written as $\underline{T} = A\underline{U}$ for suitable choice of the matrix A. Therefore, the limiting distribution of $\sqrt{N}(\underline{T})$, under H_0, is r-variate normal with null mean vector and covariance matrix $\Lambda = ((\sigma_{ik}))$.

The elements of Λ can be easily computed and are given by

$$\sigma_{ii} = \sum_{j=1}^{c} \left\{ (r-1)^2 q_{ij} + \sum\nolimits_{i \neq i'} q_{i'j} \right\}, \tag{3.1}$$

$$\sigma_{ik} = \sum_{j=1}^{c} \{ q_{.j} - r(q_{ij} + q_{kj}) \}, \quad i \neq k. \tag{3.2}$$

It is seen that the rows of Λ add up to zero. This is expected, since $\sum_{i=1}^{r} T_i = 0$. Let $\underline{T} = (\underline{T_0'}, Tr)$ and Λ_0 denote the limiting covariance matrix of $\sqrt{N}(\underline{T_0})$. Then it is seen that Λ_0 can be expressed as

$$\Lambda_0 = r^2 D_0 - r(\underline{\Pi_0}\, \underline{J_0'} + \underline{J_0}\, \underline{\Pi_0'}) + q_{..}\, \underline{J_0}\, \underline{J_0'}, \tag{3.3}$$

where $\underline{J_0}$ is a column vector of $(r-1)$ unit elements, $D_0 = diag(q_{1.}, q_{2.}, ..., q_{(r-1).})$ and $\underline{\Pi_0'} = (q_{1.}, q_{2.}, ..., q_{(r-1).})$. Therefore, $\underline{T_0'}\Lambda^{-1}\underline{T_0}$ can be simplified using $\sum_{i=1}^{r} T_i = 0,$ to get

$$N \underline{T'_0} \Lambda_0^{-1} \underline{T_0} = \frac{N}{r^2} \left\{ \sum_{i=1}^{r} \frac{T_i^2}{\hat{q}_{i.}} - \frac{1}{\hat{q}_{..}^{+}} \left[\sum_{i=1}^{r} \frac{T_i}{\hat{q}_{i.}} \right]^2 \right\}. \tag{3.4}$$

Notice that q_{ij}'s depend upon the unknown distributions F_j, $j = 1, 2, ..., c$. Hence, they are replaced by the consistent estimators and thus we get the statistic S_2^*. If $n_{ij} = n$ for all i, j, then $p_{ij} = 1/rc$,

$$q_{i.} = rc \sum_{j=1}^{c} \int_{-\infty}^{\infty} (1 - F_j(x))^2 d_{\tilde{F}_j}(x),$$

$$q_{..} = r^2 c \sum_{j=1}^{c} \int_{-\infty}^{\infty} (1 - F_j(x))^2 d_{\tilde{F}_j}(x)$$

and the limiting covariance matrix of $\underline{T_0}$ is given by

$$\Lambda_0 = r^2 ch(\underline{F})[rI_{(r-1)} - \underline{J_0}\, \underline{J'_0}], \tag{3.5}$$

where

$$h(\underline{F}) = \sum_{j=1}^{c} \int_{-\infty}^{\infty} (1 - F_j(x))^2 d\tilde{F}_j(x).$$

Therefore,

$$N_{\underline{T'_0}} \Lambda_0^{-1} \underline{T_0} = \frac{N}{r^3 ch(\underline{F})} \sum_{i=1}^{r} T_i^2. \tag{3.6}$$

Since $h(\underline{F})$ depends on the unknown distributions, it is consistently estimated as in (2.5), which yields the statistic S_1^*. We have thus proved the following theorem.

Theorem 3.1. *If* H_0 *is true and* $n_{ij} = Np_{ij}$, *where* $0 < p_{ij} < 1$ *and* $\sum_{i=1}^{r}\sum_{j=1}^{c} p_{ij} = 1$, *then both the statistics* S_1^* *and* S_2^* *defined by* (2.4) *and* (2.6), *respectively*, *have the limiting chi-square distributions with* $(r-1)$ *degree of freedom*, *as* $N \to \infty$.

Remark. We have

$$\int_{-\infty}^{\infty} (1 - F_j(x))^2 d\widetilde{F}_j(x) \leq \int_{-\infty}^{\infty} (1 - F_j(x))^2 dF_j(x) = \frac{1}{3}.$$

This suggests evaluating

$$S_1^{**} = \frac{3N}{r^3 c} \sum_{i=1}^{r} T_i^2, \tag{3.7}$$

$$S_2^{**} = \frac{3N}{r^2} \left\{ \sum_{i=1}^{r} \frac{T_i^2}{q_{i.}^*} - \frac{1}{q_{..}^{++}} \left[\sum_{i=1}^{r} \frac{T_i}{q_{i.}^*} \right]^2 \right\}, \tag{3.8}$$

where

$$q_{i.}^* = \sum_{j=1}^{c} \frac{1}{p_{ij}}, \quad q_{..}^{++} = \sum_{i=1}^{r} \frac{1}{q_{i.}^*},$$

as lower bound for S_1^* and S_2^*, respectively, in order to check computations. Also, tests based on S_1^{**} and S_2^{**} themselves would be conservative.

3.2. Asymptotic distribution of S_1^* and S_2^* under translation type alternative

Theorem 3.2. *Consider a sequence of Pitman alternatives given by*

$$H_N : X_{ijl} = \mu + N^{-1/2}\alpha_i + \beta_j + \varepsilon_{ijl}, \quad \forall i, j, l, \tag{3.9}$$

where not all α_i*'s are equal and* ε*'s behave as in* (1.1). *Also, assume without loss of generality that* $\sum_{i=1}^{r}\alpha_i = 0$.

(a) *The limiting distribution of* $\underline{T_0}$, *under* H_N *as* $N(\Sigma_i\Sigma_j n_{ij}) \to \infty$ *in such a way that* $\frac{n_{ij}}{N} = p_{ij} > 0$ *for* $i = 1, 2, ..., r$, $j = 1, 2, ..., c$ *remained fixed and* $\sum_{i=1}^{r}\sum_{j=1}^{c} p_{ij} = 1$, *is multivariate normal with mean* $\underline{\eta_0} = (\eta_1, \eta_2, ..., \eta_r)'$, *where*

$$\eta_i = -2N^{-\frac{1}{2}} r(\alpha_i - \alpha^*)\Gamma, \tag{3.10}$$

where

$$\Gamma = \sum_{j=1}^{c}\left\{\int f^0(y)[1 - I_j(y)]^2 dF^0(y) + \int i_j(y)[1 - I_j(y)][1 - F^0(y)]dF^0(y)\right\} \tag{3.11}$$

and $\alpha^* = \sum_{i=r}^{r}\alpha_i / r$ *and its covariance matrix* Λ_0 *whose elements are given in equations* (3.7) *and* (3.8), *provided following conditions hold*:

(i) $F(y)$ *is absolutely continuous with derivative* $f(y)$

(ii) $\left|\frac{1}{h}F(y + h) - F(y)\right| < g(y)$ *for small h and* $\int_{-\infty}^{\infty} yg(y)dF(y) < \infty$.

(b) *Under* H_N *and condition defined above,* S_2^* *follows in limit as* $N \to \infty$, *a non-central chi-square distribution with* $(r - 1)$ *degree of freedom and the non-centrality parameter* λ_0 *is given by*

$$\lambda_0 = \underline{\eta_0'}\Lambda_0^{-1}\underline{\eta_0} = \frac{4\sum_{i=1}^{r}(\alpha_i - \alpha^*)^2}{rch(\underline{F})}\Gamma^2, \tag{3.12}$$

where Γ *is defined in equation* (3.11).

Proof. (a) To prove this theorem, we use the following lemma:

$$\int \prod_{j=1}^{n} G_i(x+h_j)\,dF(x) = \int \prod_{j=1}^{n} G_j(x)\,dF(x)$$

$$+ \sum_{i=1}^{n} h_i \int g_i(x) \prod_{\substack{j=1 \\ j\neq i}}^{n} G_j(x)\,dF(x). \quad (3.13)$$

Under Pitman sequence of alternative

$$\lim_{N\to\infty} N^{\frac{1}{2}} E(T_i/H_n) = \lim_{N\to\infty} N^{\frac{1}{2}} E\left(\sum_{\substack{i'=1 \\ i'\neq i}}^{r} \sum_{j=1}^{c} U_{ii'j}\right)$$

$$= \lim_{N\to\infty} \frac{N^{\frac{1}{2}}}{n^2} n^2 E\left(\sum_{\substack{i'=1 \\ i'\neq i}}^{r} \sum_{j=1}^{c} \psi_{ii'j}/H_N\right), \quad (3.14)$$

$$E\{\psi_{ii'j}/H_N\} = Prob.\{\widetilde{X}_{i'jl'} < X_{ijl}/H_N\} - Prob.\{\widetilde{X}_{ijl} < X_{i'jl'}/H_N\}$$

$$= \int [1 - F_{ij}(x)]\,d\widetilde{F}_{i'j}(x)/H_n - \int [1 - F_{i'j}(x)]\,d\widetilde{F}_{ij}(x)/H_N$$

$$= \int [1 - F_{ij}^{0}(x)][1 - I_j(x)]^2\,dF_{i'j}^{0}(x)/H_n$$

$$- \int [1 - F_{i'j}^{0}(x)][1 - I_j(x)]^2\,dF_{ij}^{0}(x)/H_n$$

$$= \int [1 - F^0(x - \mu - N^{-1/2}\alpha_i - \beta_j)]$$

$$\cdot [1 - I_j(x)]^2\,dF^0(x - \mu - N^{-1/2}\alpha_i' - \beta_j)$$

$$- \int [1 - F^0(x - \mu - N^{-1/2}\alpha_{i'} - \beta_j)]$$

$$\cdot [1 - I_j(x)]^2\,dF^0(x - \mu - N^{-1/2}\alpha_i - \beta_j)$$

$$= \int \left[1 - F^0\left(y + N^{-\frac{1}{2}}(\alpha_{i'} - \alpha_i)\right)\right]$$

$$\cdot [1 - I_j(y + \mu + N^{-1/2}\alpha_{i'} + \beta_j)]^2 \, dF^0(y)$$

$$- \int \left[1 - F^0\left(y + N^{-\frac{1}{2}}(\alpha_i - \alpha_{i'}) \right) \right]$$

$$\cdot [1 - I_j(y + \mu + N^{-1/2}\alpha_i + \beta_j)]^2 \, dF^0(y). \tag{3.15}$$

Using the lemma given in equation (3.13), equation (3.15) can be written as

$$\int \left[1 - F^0\left(y + N^{-\frac{1}{2}}(\alpha_{i'} - \alpha_i) \right) \right]$$

$$\cdot \left[\left(1 - I_j\left(y + \mu + N^{-\frac{1}{2}}\alpha_{i'} + \beta_j \right) \right) \right]^2 dF^0(y)$$

$$= \int [1 - F^0(y)][1 - I_j(y)]^2 \, dF^0(y)$$

$$- N^{-\frac{1}{2}}(\alpha_{i'} - \alpha_i) \int f^0(y)[1 - I_j(y + \mu + N^{-1/2}\alpha_{i'} + \beta_j)]^2 \, dF^0(y)$$

$$- 2(\mu + N^{-1/2}\alpha_{i'} + \beta_j) \int i_j(y) \left[1 - I_j\left(y + \mu + N^{-\frac{1}{2}}\alpha_{i'} + \beta_j \right) \right]$$

$$\cdot [1 - F^0(y)] dF^0(y) \tag{3.16}$$

and

$$\int \left[1 - F^0\left(y + N^{-\frac{1}{2}}(\alpha_i - \alpha_{i'}) \right) \right]$$

$$\cdot [1 - I_j(y + \mu + N^{-1/2}\alpha_i + \beta_j)]^2 \, dF^0(y)$$

$$= \int [1 - F^0(y)][1 - I_j(y)]^2 \, dF^0(y)$$

$$- N^{-\frac{1}{2}}(\alpha_i - \alpha_{i'})\int f^0(y)[1 - I_j(y + \mu + N^{-1/2}\alpha_i + \beta_j)]^2 dF^0(y)$$

$$- 2(\mu + N^{-1/2}\alpha_i + \beta_j)\int i_j(y)\left[1 - I_j\left(y + \mu + N^{-\frac{1}{2}}\alpha_i + \beta_j\right)\right]$$

$$\cdot [1 - F^0(y)]dF^0(y). \tag{3.17}$$

Using equations (3.16) and (3.17), (3.15) can be simplified as

$$-N^{-\frac{1}{2}}(\alpha_{i'} - \alpha_i)\int f^0(y)[1 - I_j(y + \mu + N^{-1/2}\alpha_{i'} + \beta_j)]^2 dF^0(y)$$

$$- 2(\mu + N^{-1/2}\alpha_{i'} + \beta_j)$$

$$\cdot \int i_j(y)\left[1 - I_j\left(y + \mu + N^{-\frac{1}{2}}\alpha_{i'} + \beta_j\right)\right][1 - F^0(y)]dF^0(y)$$

$$+ N^{-\frac{1}{2}}(\alpha_i - \alpha_{i'})\int f^0(y)[1 - I_j(y + \mu + N^{-1/2}\alpha_i + \beta_j)]^2 dF^0(y)$$

$$+ 2(\mu + N^{-1/2}\alpha_i + \beta_j)$$

$$\cdot \int i_j(y)\left[1 - I_j\left(y + \mu + N^{-\frac{1}{2}}\alpha_i + \beta_j\right)\right][1 - F^0(y)]dF^0(y) \tag{3.18}$$

Using Taylor's series for first order of approximation in equation (3.18), we get

$$= -2N^{-\frac{1}{2}}(\alpha_{i'} - \alpha_i)\int f^0(y)[1 - I_j(y)]^2 dF^0(y)$$

$$- 2N^{-\frac{1}{2}}(\alpha_{i'} - \alpha_i)\int i_j(y)[1 - I_j(y)][1 - F^0(y)]dF^0(y) + O(N^{-\frac{1}{2}})$$

$$= -2N^{-\frac{1}{2}}(\alpha_{i'} - \alpha_i)\Gamma_j,$$

where

$$\Gamma_j = \int f^0(y)[1 - I_j(y)]^2 dF^0(y)$$

$$+ \int i_j(y)[1 - I_j(y)][1 - F^0(y)]dF^0(y),$$

$$\therefore E(T_i / H_N) = \frac{1}{n^2}(-2N^{-\frac{1}{2}})n^2 \sum_{\substack{i'=1 \\ i' \neq i}}^{r} (\alpha_{i'} - \alpha_i) \sum_{j=1}^{c} \Gamma_j$$

$$= -2N^{-\frac{1}{2}} r(\alpha_i - \alpha^*)\Gamma. \tag{3.19}$$

(b) As $\underline{T_0}$ has multivariate normal distribution with mean $\underline{\eta_0}$ and variance-covariance matrix Λ_0. It is obvious that $N\underline{T_0'}\Lambda_0^{-1}\underline{T_0}$ has non-central χ^2 distribution with $(r-1)$ degree of freedom and its non-central parameter is

$$\underline{\eta'}_0 \Lambda_0^{-1} \underline{\eta_0} = \frac{4r^2}{cr^3 h(\underline{F})} \Gamma^2 \sum_{i=1}^{r} \alpha_i^2,$$

$$\therefore \underline{\eta'}_0 \Lambda_0^{-1} \underline{\eta_0} = \frac{4}{crh(\underline{F})} \Gamma^2 \sum_{i=1}^{r} \alpha_i^2.$$

3.3. Consistency of the tests

Here, let us assume that X_{ijl}^0 can be written as per equation (1.1).

We consider the following alternative hypothesis:

$$H_1 : \alpha_{i'} > \alpha_i, \;\; i' \neq i, \, i' = 1, 2, ..., r.$$

Further, we shall assume that the censoring distributions are same as defined the null hypothesis. Now, under H_1, the expected value of $U_{ii'j}$ can be written as

$$\eta^{(i,i')} = E(U_{ii'j} \mid H_1)$$

$$= \int_{-\infty}^{\infty} [1 - F(x + (\alpha_{i'} - \alpha_i))][1 - I_j(x + \mu + \alpha_i + \beta_j)]^2 dF(x)$$

$$- \int_{-\infty}^{\infty} [1 - F(x + (\alpha_i - \alpha_{i'}))][1 - I_j(x + \mu + \alpha_i + \beta_j)]^2 dF(x).$$

It is easily seen that under H_1, $\eta^{(i,i')} < 0$. Therefore, $\eta^{(i)} = E(T_i \mid H_1) < 0$. Also, by asymptotic normality of U-statistic, it follows $U_{ii'j}$ converges in probability to $\eta^{(i,i')}$ for $i' \neq i$, $i, i' = 1, 2, ..., r$; $j = 1, 2, ..., c$. Therefore, T_i converges in probability to $\eta^{(i)}$ for $i = 1, 2, ..., r$. Now, using Lemma 4.1 of Bhapkar [1], it follows that tests based on S_1^{**} and S_2^{**} are consistent against the alternative H_1. This implies that the tests based on S_1^{*} and S_2^{*} are consistent against a wider class of alternatives for which $\eta^{(i)} \neq 0$ for at least one i.

4. Test Statistics by Gore [8] and Patel [16]

4.1. Test statistic by Gore [8]

Let $\varphi(t) = 1$ if $t > 0$, $\frac{1}{2}$ if $t = 0$ and $= 0$ otherwise.

Define

$$U_{ii'j} = \sum_{k=1}^{n_{ij}} \sum_{l=1}^{n_{i'j}} \varphi(X_{ijk} - X_{i'jk}) / n_{ij} n_{i'j}. \tag{4.1}$$

Note that due to the assumption about continuity of F, ties occur only with zero probability and can be ignored. Let

$$U_i = \sum_{\substack{i'=1 \\ i' \neq i}}^{r} \sum_{j=1}^{c} U_{ii'j}. \tag{4.2}$$

For the case $n_{ij} = n$ for all the *i* and *j* the proposed statistic for testing H_0,

$$S_1 = \frac{12n}{r^2} \sum_{i=1}^{r} \left(U_i - \frac{(r-1)c}{2} \right)^2. \tag{4.3}$$

More generally, if the number of observations per cell is not the same, then let $n_{ij} = Np_{ij}$, $0 < p_{ij} < 1$, $\sum_{i,j} p_{ij} = 1$ (clearly, when $n_{ij} = n$, $p_{ij} = \frac{1}{rc}$). Further, let $q_{ij} = \frac{1}{p_{ij}}$ and $q_{i.} = \sum_{j=1}^{c} p_{ij}^{-1}$, $q_{..} = \sum_{i=1}^{r} q_i$, $q_{i.}^{*} = \sum_{i=1}^{r} q_{i.}^{-1}$

The statistic proposed for the setup is

$$S_2 = \frac{12N}{r^2} \sum_{i=1}^{r} \frac{\left(U_i - \frac{(r-1)c}{2} \right)^2}{q_i} - \frac{\left\{ \sum_{i=1}^{r} \frac{\left(U_i - \frac{(r-1)c}{2} \right)}{q_i} \right\}^2}{q_{..}^{*}}, \tag{4.4}$$

where $N = \sum_i \sum_j n_{ij}$.

The test based on $S_1(S_2)$ consists of rejecting H_0 at a level of significance α if $S_1(S_2)$ exceeds a predetermined constant $S_{1\alpha}(S_{2\alpha})$. We claim that S_k and $S_{2\alpha}$ are free of F under H_0 and here the tests are distribution-free.

4.2. Test statistic by Patel [16]

Define the scoring function $\phi(\cdot)$ for comparing two observations x_{ij} and $x_{i'j}$ in the *j*th block by

$$\phi(x_{ij}, \delta_{ij}; x_{i'j}, \delta_{i'j}) = \begin{cases} -1 & \text{if } x_{ij} < x_{i'j}, \delta_{ij} = 1, \\ +1 & \text{if } x_{i'j} < x_{ij}, \delta_{i'j} = 1, \\ 0 & \text{otherwise, } i \neq i'. \end{cases} \tag{4.5}$$

The total score for the observation x_{ij} by

$$w_{ij} = \sum_{\substack{i'=1 \\ i' \neq i}}^{r} \phi(x_{ij}, \delta_{ij}; x_{i'j}, \delta_{i'j}). \tag{4.6}$$

Corresponding test statistic is

$$Q_c = \frac{c}{r\hat{\sigma}^2} \sum_{i=1}^{r} T_i^2, \tag{4.7}$$

where $T_i = \sum_{j=1}^{c} \frac{W_{ij}}{c}$ and $\hat{\sigma}^2$ is a consistent estimator of σ^2 and is given by

$$\hat{\sigma}^2 = \sum_{j=1}^{c} \sum_{i=1}^{r} \frac{W_{ij}}{rc(c-1)}. \tag{4.8}$$

5. Methods and Results of Comparisons

In this section, we present the results of the simulation study for power analysis of proposed test statistic when error distributions are logistic, Laplace, normal, Pareto of first kind, exponential, generalized exponential and Weibull. Five Thousands (5000) values of each statistic under each experimental situation were simulated and the proportions of rejection of hypothesis at the nominal five percent levels were recorded. The simulated proportions obtain under the translation alternative represent the estimates of the power function. For study, we consider the following distribution of the observation (X^0):

I. Standard normal distribution.

II. Standard logistic distribution.

III. Standard double exponential (Laplace) distribution.

IV. Pareto distribution of first kind (location parameter =1; shape parameters $p = 0.5$ and 1.5) with median zero.

V. Standard generalized exponential distribution suggested by Gupta and

Kundu [9] (shape parameters $p = 1.5$ and 2.5) with median zero, i.e., GE (1.5) and GE (2.5) with median zero.

VI. Standard exponential distribution with median zero.

VII. Standard Weibull distribution (shape parameters $p = 0.5$ and 1.5) with median zero.

Also, the distribution of Z is exponential and its CDF is given by

$$I_j(z) = 1 - e^{-(z-\beta)}, \text{ where } z > \beta.$$

In our study, we observe the main effect of treatments, by fixing the block effects. Therefore, block parameters $(\beta_j;\ j = 1.\ 2, \ldots, c)$ remain identical for both null as well as alternative hypothesis. The block parameters are considered [(2, 1), (2, 1, 2.5) and (2, 1, 2.5, 3)] for different number of blocks [2, 3 and 4], respectively. Further, it is noted that our study is restricted up to identical exponential censoring with location parameter 3 for all blocks. The critical values S_1^* calculated under the null hypothesis when error distributions are logistic, Laplace, normal, Pareto of first kind (location parameter $= 1;$ shape parameters $p = 0.5$ and 1.5), exponential, generalized exponential (shape parameters $p = 1.5$ and 2.5) and Weibull (shape parameters $p = 0.5$ and $p = 1.5)$ are given in Table 5.1. Alternative to null hypothesis given in equation (1.1) can be written as

$$H_0 : \alpha_i = 1;\ i = 1, 2, \ldots, r \text{ vs } H_1 : \text{at least one pair differs significantly.}$$

The empirical powers proposed test statistics (S_1^*) calculate under translation type alternative are given in Table 5.2. For simulation study, we consider treatment parameters [(0.8, 1, 1.2), (0.8, 1,1.2, 1.4), (0.6, 0.8, 1, 1.2, 1.4) and (0.6, 0.8, 1, 1.2, 1.4, 1.6)] for different number of treatments [3, 4, 5 and 6], respectively. Table 5.1 and Table 5.2 give the critical values and power of proposed test statistic (S_1^*), with number of observations per cell (n) is taken 2 and 4 for different combination of numbers of treatments (r) and blocks (c), respectively.

From Table 5.2, it can be seen that the value of simulated power of test is increasing function of (r, c, n) for all distributions.

Table 5.1. **Empirical critical values of test statistic S_1^*** when error distributions are Standard Normal, Standard Logistic, Standard Double Exponential (Laplace), Pareto distribution of first kind (location parameter =1;shape parameter $p = 0.5$ and 1.5), Standard Generalized Exponential distribution (shape parameter $p = 1.5$ and 2.5), Standard Exponential and Standard Weibull with shape parameters $p = 0.5$ and $p = 1.5$

r	c	n	Normal	Logistic	Double Exponential	Pareto ($p = 0.5$)	Pareto ($p = 1.5$)	*GE(1.5)*	*GE(2.5)*	Exponential	Weibull ($p = 0.5$)	Weibull ($p = 1.5$)
3	2	2	8.457	8.313	8.547	8.229	8.457	8.509	8.457	8.509	8.509	8.509
		4	7.287	7.19	7.295	7.023	7.311	6.930	7.150	7.058	7.16	7.104
	3	2	8.429	8.714	8.586	8.68	8.229	8.727	8.550	8.641	8.871	8.776
		4	7.227	7.316	7.390	7.217	7.158	7.189	7.163	6.905	7.319	7.113
	4	2	8.619	8.376	8.890	8.444	8.803	8.566	8.883	8.566	8.599	8.658
		4	7.229	7.260	7.244	7.276	7.517	7.072	7.227	7.286	7.353	7.213
4	2	2	9.797	9.90	9.798	9.543	9.723	9.835	9.914	9.826	9.557	9.848
		4	8.842	9.033	9.009	8.759	8.636	8.663	9.005	8.529	8.915	8.666
	3	2	10.045	9.983	10.187	10.027	10.194	10.222	10.135	10.079	10.183	9.904
		4	8.876	8.984	8.978	8.587	8.658	8.829	8.986	9.106	8.608	9.181
	4	2	10.206	10.325	12.227	10.265	10.017	10.180	10.268	10.498	10.108	10.179
		4	8.932	8.730	9.106	8.757	8.909	9.016	8.895	8.741	8.686	9.056
5	2	2	11.188	11.252	11.250	10.932	11.346	11.233	11.225	11.086	11.115	11.124
		4	10.218	10.338	10.098	10.276	10.355	10.412	10.477	10.551	10.243	10.491
	3	2	11.393	11.493	11.594	11.368	11.451	11.324	11.487	11.639	11.542	11.508
		4	10.141	10.625	10.353	10.412	10.329	10.378	10.238	10.363	10.493	10.551
	4	2	11.654	11.716	11.302	11.545	11.600	11.626	11.608	11.574	11.466	11.504
		4	10.54	10.755	10.655	10.549	10.399	10.263	10.701	10.201	10.393	10.401
6	2	2	12.267	12.689	12.250	12.403	12.158	12.419	12.465	12.328	12.348	12.369
		4	11.992	11.781	11.769	11.778	11.864	11.803	12.000	11.538	11.938	11.759
	3	2	12.854	12.816	12.588	12.921	12.898	12.734	12.558	12.855	12.768	12.837
		4	11.705	11.959	11.936	11.863	11.893	11.699	11.861	12.059	11.919	11.916
	4	2	13.181	12.756	12.939	12.699	13.015	13.346	13.195	12.994	12.92	12.913
		4	11.781	11.892	11.936	11.939	12.055	11.859	12.213	11.835	12.068	12.123

Table 5.2. **Empirical power of test statistic S_1^*** when error distributions are Standard Normal, Standard Logistic, Standard Double Exponential (Laplace) , Pareto distribution of first kind (location parameter =1;shape parameter $p=0.5$ and 1.5) , Standard Generalized Exponential distribution (shape parameter $p=1.5$ and 2.5) , Standard Exponential and Standard Weibull with shape parameters $p=0.5$ and $p=1.5$

r	c	n	Normal	Logistic	Double Exponential	Pareto ($p=0.5$)	Pareto ($p=1.5$)	*GE(1.5)*	*GE(2.5)*	Exponential	Weibull ($p=0.5$)	Weibull ($p=1.5$)
3	2	2	0.0614	0.0560	0.0602	0.0556	0.1012	0.0698	0.0754	0.0964	0.0992	0.1094
		4	0.083	0.0658	0.0788	0.0614	0.1504	0.1370	0.1006	0.1578	0.1542	0.2008
	3	2	0.0868	0.0578	0.0758	0.0566	0.1496	0.0952	0.0884	0.1326	0.1264	0.1576
		4	0.109	0.0656	0.0872	0.0592	0.2638	0.1706	0.1370	0.249	0.2546	0.295
	4	2	0.0786	0.0654	0.0722	0.0560	0.1612	0.1272	0.0938	0.1546	0.1758	0.1884
		4	0.1072	0.0746	0.1066	0.0576	0.3184	0.2390	0.1612	0.297	0.328	0.3504
4	2	2	0.0854	0.0624	0.0748	0.0550	0.1414	0.1092	0.0946	0.1398	0.1522	0.1684
		4	0.1394	0.0720	0.1090	0.0644	0.2890	0.2018	0.1500	0.2962	0.2364	0.3732
	3	2	0.0964	0.0670	0.0950	0.0567	0.2136	0.1428	0.1148	0.2052	0.2094	0.2642
		4	0.1698	0.0902	0.1588	0.0678	0.4826	0.3210	0.2318	0.4488	0.4602	0.556
	4	2	0.1114	0.0688	0.1012	0.065	0.2792	0.2030	0.1564	0.2602	0.2664	0.326
		4	0.2162	0.1182	0.1704	0.0952	0.5628	0.4138	0.3304	0.5754	0.5612	0.6702
5	2	2	0.1092	0.0638	0.0942	0.0652	0.1840	0.1448	0.1266	0.2168	0.1834	0.2846
		4	0.2098	0.1008	0.1898	0.0754	0.4202	0.3182	0.2412	0.4158	0.364	0.6052
	3	2	0.1642	0.0876	0.1198	0.0706	0.3190	0.2484	0.1770	0.2976	0.288	0.4578
		4	0.316	0.1202	0.2560	0.0832	0.6762	0.5320	0.4078	0.6808	0.614	0.8354
	4	2	0.1808	0.0926	0.1724	0.0788	0.4170	0.3290	0.2700	0.4456	0.3922	0.5796
		4	0.378	0.1386	0.2950	0.0962	0.8020	0.6956	0.5174	0.833	0.7556	0.9256
6	2	2	0.1564	0.0804	0.1426	0.0702	0.3066	0.2192	0.1764	0.3006	0.2384	0.4588
		4	0.3078	0.1340	0.2462	0.0714	0.5944	0.4704	0.3666	0.6608	0.5064	0.8442
	3	2	0.2198	0.1046	0.1942	0.0738	0.4582	0.3950	0.3086	0.486	0.414	0.6654
		4	0.4906	0.1830	0.3830	0.0980	0.8562	0.7726	0.5950	0.8754	0.7992	0.9742
	4	2	0.2676	0.1226	0.2046	0.094	0.5892	0.5968	0.4672	0.6278	0.54	0.7904
		4	0.6006	0.2300	0.4596	0.1316	0.9328	0.8820	0.7566	0.9524	0.8998	0.9902

6. Concluding Remarks

In this article, the goal is to provide test procedures for testing equality of main effects in two-way layouts using *U*-statistics based on right censorship. Further, we derive the asymptotic distributions under the null hypothesis and Pitman alternative for equal and unequal sample sizes. We simulate the cut off point for the test statistic and obtain power of the test under normal, logistic, double exponential, Pareto of first kind, generalized exponential, exponential and Weibull distributions. It is seen that the test is not performing well for thick tail distributions like logistic and double exponential (Laplace) compared to normal distribution. The performance is seen to be well for positive valued distributions like Pareto of first kind, exponential, generalized exponential and Weibull. The use of test procedure is recommended for positive valued error random variable.

Note. We have written a subroutine to compute the test statistics in C++. The source code is available on request.

Acknowledgement

The authors would like to thank the anonymous referee for his valuable comments and suggestions which helped them to improve the earlier version of this paper. The second author would like to thank University Grant Commission (UGC) for financial support. This research is funded by University Grant Commission (UGC) for Minor Research Project (F.37-549/2009(SR)).

References

[1] V. P. Bhapkar, A nonparametric test for the problem of several samples, Ann. Math. Stat. 32 (1961), 1108-1117.

[2] V. P. Bhapkar and A. P. Gore, A distribution-free test for symmetry in Hierarchical data, J. Multivariate Anal. 3 (1973), 483-489.

[3] N. Breslow, A generalized Kruskal-Wallis test for comparing *k*-samples subject to unequal pattern of censorship, Biometrika 57(3) (1970), 576-594.

[4] G. W. Brown and A. M. Mood, On median tests for linear hypotheses, Proc. 2nd Berkeley Symp. Math. Stat. Prob. (J. Neymann, ed.), Vol. 1, 1950, pp. 159-166.

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

[6] X. Gao and M. Alvo, A unified nonparametric approach for unbalanced factorial designs, J. Amer. Statist. Assoc. 100(471), (2005), 926-941.

[7] E. Gehan, A generalized Wilcoxon test for comparing arbitrarily single censored samples, Biometrika 52 (1965), 203-223.

[8] A. P. Gore, Some nonparametric tests and selection procedures for the main effects in two-way layouts, Ann. Inst. Statist. Math. 27(3) (1975), 487-500.

[9] R. D. Gupta and D. Kundu, Generalized exponential distribution, Aust. N. Z. J. Stat. 41(20) (1999), 173-188.

[10] J. L. Hodges and E. L. Lehmann, Rank methods for combination of independent experiments in analysis of variance, Ann. Math. Statist. 33 (1962), 482-497.

[11] E. L. Lehmann, Asymptotically nonparametric inference in some linear models with one observation per cell, Ann. Math. Statist. 35 (1964), 726-734.

[12] R. C. Magel, A nonparametric test for ordered alternatives in two-way layouts with censored data, Biom. J. 31(5) (1989), 609-618.

[13] H. Mansouri and Z. Govindarajulu, A class of rank tests for interaction in two-way layouts, J. Appl. Stat. 17 (1990), 417-426.

[14] H. Mansouri, Aligned rank transform tests in linear models, J. Statist. Plann. Inference 79 (1999), 141-155.

[15] K. L. Moore and M. J. Vander Lann, Increasing power in randomized trains with right censored outcomes through covariate adjustment, J. Biopharm. Stat. 19(6) (2009), 1099-1131.

[16] K. M. Patel, A generalized Friedman test for the randomized block design when observations are subject to arbitrary right censorship, Commum. Statist. 4(4) (1975), 389-395.

[17] P. K. Sen, On a class of aligned rank order tests in two-way layouts. The Annals of Mathematical Statistics 39 (1968), 1115-1124.

[18] K. M. Wittkowski, Friedman-type statistics and consistent multiple comparisons for unbalanced designs with missing data, J. Amer. Statist. Assoc. 83(404) (1988), 1163-1170.