Available online at http://www.ajol.info/index.php/njbas/index
Nigerian Journal of Basic and Applied Science (December, 2012), 20(4): 305-314
ISSN 0794-5698

Full Length Research Article

Comparison of a Class of Rank-Score Tests in Two-Factor Designs

A. Danbaba
Department of Mathematics, Usmanu Danfodiyo University, Sokoto
[Corresponding Author: E-Mail: adkoko@udusok.edu.ng; ☏: +234(0)8060747190]

ABSTRACT: Rank score functions are known to be versatile and powerful techniques in factorial designs. Researchers have established the theoretical properties of these methods based on nonparametric hypotheses, but only scanty empirical results are available in the literature on these procedures. In this paper, four types of rank score functions Wilcoxon-scores, Mood-scores, normal-scores and expected normal- scores are studied in the context of two–way factorial designs using asymptotic χ^2 (Wald-Type) and modified Box- approximation (ANOVA-Type) tests. The empirical Type I error rate and power of these test statistics on the rank scores were determined using Monte Carlo simulation to investigate the robustness of the tests. The results show that there are problems of inflation in the Type I error rate using asymptotic χ^2 test for all the rank score functions, especially for small sample sizes and distributions studied. The modified Box- approximation test was found to be robust for both validity and efficiency, especially for Wilcoxon, normal and expected normal score functions. It was concluded that the asymptotic χ^2 test is non-robust for rank score functions in two-factor designs.
Keywords: Rank score functions, Type I error rates, Power, Factorial designs.

INTRODUCTION

When analyzing data from a two-factor design, usually a linear model is assumed and the hypotheses are formulated by the parameters of this model (Brunner and Puri, 2002). If no specific distribution functions are assumed, then there are no parameters to formulate hypotheses. In this situation, artificial parameters are usually introduced to express the hypotheses (Brunner and Puri, 2002). The hypotheses derived from these artificial parameters are called nonparametric hypotheses. Akritas and Arnold (1994) reported the idea to formulate nonparametric hypotheses in factorial designs by contrasts of the distribution functions. However, the nonparametric hypothesis in the one-way layout to higher-way layouts are presented in several studies (Lemmer and Stoker 1967; Rinaman, 1983; Hora and Conover, 1984; Brunner et al., 1995; Brunner and Puri, 2002).

Rank procedures for nonparametric hypotheses based on the distribution functions are derived for score functions with bounded second derivatives (Brunner and Puri, 2002). In this approach data from continuous distributions as well as discrete ordinal data are covered. The results in Brunner and Puri's (2002) paper are presented in a general form such that statistical nonparametric hypotheses in any factorial design can be derived easily from this unified approach. Many rank (score) statistics given in the literature are special cases of the statistics they derived. However, they stated that the procedures are applicable to analyze data for balanced and unbalanced designs, data with continuous distribution functions or data with ties. Furthermore, Brunner and Puri (2002) applied this approach to factorial design using Wilcoxon-scores and Mood-scores. Asymptotic χ^2 and modified Box-approximation (Box, 1954; Brunner *et al*., 1997) are used as tests statistics.

The p-values of the asymptotic χ^2 and modified Box-approximation tests differ (when applied to the same data) for the two types of scores (Brunner and Puri, 2002). A question of whether or not that the tests under Wilcoxon and Mood scores have the same Type I error rates and power for two-factor designs can be raised. However, the p-values of asymptotic χ^2 and modified Box- approximation tests for other scores like normal-scores and expected normal-scores (Sawilowsky, 1990; Conover, 1999) may also differ. In this study, Type I error rates and power comparison of the asymptotic χ^2 and modified Box- approximation tests for Wilcoxon-scores, Mood-scores, normal-scores and expected normal scores were carried out using Monte Carlo simulation.

The purpose of this paper is to determine which of the test statistic (the asymptotic χ^2 test or modified Box-approximation test) on the rank score function has good Type I error rates and power for the nonparametric hypotheses of two-factor designs with independent observations, fixed number of levels and several independent observations per cell (replicates). In addition, the robustness of validity and efficiency of the two test statistics were investigated based on the rank scores.

METHODS

The Expected normal

A normal distribution was sampled randomly, ordered, recorded, and replaced, and this process was repeated many number of times. The average of each position of N is the expected normal score (Harter, 1961 and Royston, 1981). In a sample of size N the expected value of the r^{rh} largest order statistic is given by

$$E(r, N) = -\Phi^{-1}\left(\frac{r-\theta}{N-2\theta+1}\right),$$

where E(r, N) is the expected normal score for an observation, r is the rank for that observation, $\Phi(.)$ is the quantile from the standard normal distribution and $\theta = 0.375$ (Harter, 1961; Royston, 1981). The expected value of the r^{th} smallest observation is given by the same expression preceded by a minus sign.

The Normal-Score

The data were ranked from 1 to sample size, N. The ranked observations (r_{ijk}) were then replaced by their normal scores ($\Phi^{-1}\left(\frac{r_{ijk}}{N+1}\right)$),

where Φ is the cumulative distribution function of a standard normal distribution (Conover, 1999).

The Wilcoxon-Score and Mood-Score

Brunner and Puri (2002) defined the Wilcoxon-scores as $a_{ijk}^{w} = \frac{1}{N}\left(R_{ijk} - \frac{1}{2}\right)$,

where R_{ijk} is the rank of the original observations from one to sample size, N. They also defined the Mood scores as $a_{ijk}^{m} = \left(a_{ijk}^{w} - \frac{1}{2}\right)^2$.

Model and Nonparametric Hypothesis

For a two-factor design with fixed levels *a* and *b*, the k^{th} observation from (i, j) is modeled as

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk},$$

where i = 1, …, *a;* j = 1, …, *b;* k = 1, …, n; μ is the overall mean, α_i is the effect of the *i*th level of factor *A*, β_j is the effect of the *j*th level of factor *B*, $(\alpha\beta)_{ij}$ is the effect of the interaction between the *i*th level of factor *A* and the *j*th level of factor *B*, ε_{ijk} is the random error associated with the *k*th replicate in cell (*i, j*), and Y_{ijk} is the kth observations in cell (i, j).

The hypotheses usually tested by the two-way factorial ANOVA for the A main effect, B main effect, and interaction are, respectively,

Ho: $\alpha_i = \mu_{i.} - \mu = 0$ for all $i = 1,\ldots, a$

Ho: $\beta_j = \mu_{.j} - \mu = 0$ for all $j = 1,\ldots, b$

Ho: $\alpha\beta_{ij} = \mu_{ij} - \mu_i - \mu_j + \mu = 0$ for all i, j'.

Brunner and Puri (2002) claimed that the rank procedures might test hypotheses where rank mean counterparts are substituted for the appropriate μ's in the above hypotheses, but in reality they test truly nonparametric hypotheses. In this situation, independent random variables Y_{ijk} have distribution function

$$F_{ij}(y) = F(y - \mu_{ij}),$$

where $\mu_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$.

The nonparametric hypotheses are given as a function of the cumulative distribution for each cell, $F_{ij}(y)$ (Brunner et al. 1997 and Akritas *et al.,* 1997). $Fi.$ is the average of the $F_{ik}(y)$ across the *b* levels of B, $F_{.j.}$ is the average of the $F_{ij}(y)$ across the *a* levels of A, and $F..$ is the average of the $F_{ij}(y)$ across the *ab* cells. Then the hypotheses tested by these nonparametric methods for the A main effect, B main effect, and interaction, are, respectively,

Ho: $\alpha_i = Fi. - F.. = 0$ for all $i = 1,\ldots, a$

Ho: $\beta_j = F.j - F.. = 0$ for all $j = 1,\ldots, b$

Ho: $\alpha\beta_{ij} = Fij(x) - Fi. - Fj + F.. = 0$ for all $i = 1,\ldots, a$, for all $j = 1,\ldots, b$

These hypotheses are respectively equivalent to

$H_o{:}\mathbf{C}_A F = 0$, $H_o{:}\mathbf{C}_B F = 0$, $H_o{:}\mathbf{C}_{AB} F = 0$,

where

$\mathbf{C}_A = \mathbf{P}_a \otimes \frac{1}{b}\mathbf{J}_b$, $\mathbf{C}_B = \frac{1}{a}\mathbf{J}_a \otimes \mathbf{P}_b$, $\mathbf{C}_{AB} = \mathbf{P}_a \otimes \mathbf{P}_a$, $\mathbf{P}_a = \mathbf{I}_a - \frac{1}{a}\mathbf{J}_a$, $\mathbf{P}_b = \mathbf{I}_b - \frac{1}{b}\mathbf{J}_b$, 1_d is the $d \times 1$ summing vector, $J_d = 1_d 1_d^1$ and $I_d = diag\{1,\ldots,1\}$.

$\hat{P} = \left(\bar{Q}_{11}.,\bar{Q}_{12}.,\ ...,\bar{Q}_{ab}.\right)'$ is used to test these hypotheses using

$Q_i.. = \frac{1}{b}\sum_{j=1}^{b} Q_{ij}.$ and $Q_{ij}. = \frac{1}{n}\sum_{k=1}^{n} Q_{ijk}$,

where Q_{ijk} (i= 1, 2, …, a; j = 1, 2, …, b; k = 1, 2, …, n) are the rank scores from normal, expected normal, Mood or Wilcoxon scores. See Brunner and Puri (2002) for more details.

Asymptotic χ^2 test (Wald-Type test)
The Wald -Type tests for factor A, B and AB interaction are respectively as follows:

$$Q_A = N.\hat{P}'\left\{C_A'(C_A.\hat{V}.C_A')^- C_A\right\}.\hat{P} \sim \chi^2_{a-1}$$

$$Q_B = N.\hat{P}'\left\{C_B'(C_B.\hat{V}.C_B')^- C_B\right\}.\hat{P} \sim \chi^2_{b-1}$$

$$Q_{AB} = N.\hat{P}'\left\{C_{AB}'(C_{AB}.\hat{V}.C_{AB}')^- C_{AB}\right\}.\hat{P} \sim \chi^2_{(a-1)(b-1)}$$

Where $\hat{P} = (\bar{Q}_{11.},\bar{Q}_{12.},\cdots,\bar{Q}_{ab.})'$,

$$\hat{s}^2_{ij1} = \frac{1}{(n-1)}\sum_{k=1}^{n}(Q_{ijk} - \bar{Q}_{ij.})^2,$$

$$\hat{V} = \frac{N}{n}\times diag(\hat{s}^2_{111},...,\hat{s}^2_{ab1})$$ and N = *abn*

Modified Box- approximation test (ANOVA-Type test)
The ANOVA-Type tests for factor A, B and AB interaction are respectively as follows:

$$T_A = \frac{N}{m\times tr(\hat{V})}\hat{P}'\left\{C_A'[C_A.C_A']^-\right\}\hat{P} \sim F_{a-1,\,f_o}$$

$$T_B = \frac{N}{m\times tr(\hat{V})}\hat{P}'\left\{C_B'[C_B.C_B']^-\right\}\hat{P} \sim F_{b-1,\,f_o}$$

$$T_{AB} = \frac{N}{m\times tr(\hat{V})}\hat{P}'\left\{C_{AB}'[C_{AB}.C_{AB}']^-\right\}\hat{P} \sim F_{(a-1)(b-1),\,f_o}$$

where $f_o = \frac{[tr(\hat{V})]^2}{tr[\hat{V}(\Lambda_d - I_d)^{-1}]}$, *m* is any element in diag(**C**(**CC**$'$)$^-$ **C**) and Λ_d= diag(n, …, n)

Monte Carlo Simulation
For two-way table with *a* levels of factor A, *b* levels of factor B, n >1 observations per cell, and level of significanceα. A set of data x_{ijk} (i = 1, 2, …, a; j =1, 2, …, b; k = 1, 2, …, n) were obtained from the probability distribution. From these data (x_{ijk}), test statistics were computed, and used to determine whether to accept or reject their corresponding null hypotheses. In this study *a* = *b* = 4, and n = 5, 10, 15, 20, 25, 30 replicates were used. The power of the tests for main effects were obtained for only factor A at α_1=0.5, α_3= – 0.5, α_2 = α_4=0. However, the power for tests of interaction was first generated when main effects are null and then when main effects are non-null (α_1 = 0.5, α_3 = – 0.5, α_2 = α_4 = β_1=β_4= 0, β_2= 0.5,β_3= –0.5) using the interaction effects $\alpha\beta_{11}$= 1, $\alpha\beta_{13}$=0.5, $\alpha\beta_{22}$= –0.5, $\alpha\beta_{33}$ = –1, $\alpha\beta_{42}$ = 0.5, $\alpha\beta_{44}$ = –0.5, the remaining $\alpha\beta_{ij}$ = 0. In addition, the probability distributions used for the study are N(μ, σ^2), exponential, lognormal and mixed normal [0.75×N(μ, σ^2)+ 0.25× N(10+μ, σ^2)], where σ^2 = 1. Data generated from each distribution are converted to rank score (y_{ijk}). The estimate of Type I error rate for a particular test is obtain by plugging y_{ijk} in the two-way table, computing

$$C_r = \begin{cases} 0, & \text{if the true null hypothesis is accepted} \\ 1, & \text{if the true null hypothesis is rejected} \end{cases}$$

and $T = \frac{\sum_{r=1}^{G} C_r}{G}$.

Then T is the required Type I error rate. Similarly, the power of the test is obtained by computing

$$D_r = \begin{cases} 0, & \text{if the false null hypothesis is accepted} \\ 1, & \text{if the false null hypothesis is rejected} \end{cases}$$

and $P = \frac{\sum_{r=1}^{G} D_r}{G}$.

P is the required power of the test, where G =1000.

Robustness
Empirical Type I error rates (π) within the confidence interval

$$\alpha - Z_\alpha\sqrt{\frac{\alpha(1-\alpha)}{G}} \le \pi \le \alpha + Z_\alpha\sqrt{\frac{\alpha(1-\alpha)}{G}}$$

for a test is considered robust for validity, where G and α are the number of replications and level of significance, respectively (Lin and Myers, 2006). This criterion is used with G=1000 and α =0.05. That is, a test is robust for validity if 0.036 ≤ π ≤ 0.064.

A test with empirical power (P_a) is considered robust for efficiency over another test with empirical power (P_b) if $|P_a - P_b| \geq 2 \times Z_{\alpha/2} \times SE(\hat{P})$, where the quantity $2 \times Z_{\alpha/2} \times SE(\hat{P})$ is the difference between the upper and lower limits of the confidence interval for *p*, $SE(\hat{P}) = \sqrt{\frac{P(1-P)}{G}}$ and *p* denote independent Bernoulli trial probability of success (Steidl and Thomas, 2000). At G = 1000 and *p* = 0.5, two tests with empirical power difference (from the same population) within ±0.062 are considered equal.

RESULTS

In Table 1 through Table 5, QW, QM, QN, QE are Wald Type tests for Wilcoxon-score, Mood-score, normal-score, and expected normal-score respectively, while AW, AM, , AN, AE are ANOVA Type tests for Wilcoxon-score, Mood-score, normal-score, expected normal-score respectively. Table 1 and Table 2 show the Type I error rates for factor A and interaction tests, respectively. Table 3 shows the power for tests of factor A, while Tables 4 and 5 show the power for tests of interaction. The power for QN and QE are similar and therefore only the power for QN is reported. Similarly, the power for AN and AE are similar and only the power for AN is reported.

Table 1 shows the Type I error rates for tests of factor A. The bolded values are the Type I error rates outside the interval of robustness (0.036 0.064). At n = 5, the Type I error rates of QW, QM, QN and QE are outside the interval (0.036 0.064) while for the remaining sample sizes, the rates are within the interval. The Type I error rates for AW, AM, AN and AE are within the interval for all sample sizes and populations studied.

Table 2 shows the Type I error rates for the interaction tests. The Type I error rates of QW, QM, QN and QE are outside the interval (0.036 0.064) in most of the sample sizes and populations studied. The Type I error rates for AW, AM, AN and AE are within the interval for all sample sizes and populations studied.

In Table 3, the results indicate that QM and AM have low power for all sample sizes and populations studied. The tests QW, AW, QN and AN are powerful and have similar power for all sample sizes and populations used in the study.

When main effects are null, the power of interaction tests are given in Table 4. The results show that QW has some power advantage over other tests, especially for small sample size. The test AN has smaller power than other tests.

The powers of interaction tests when main effects are non-null are given in Table 5. The results show that QN and QW have some power advantage over other tests for small samples.

Table 1: Type I error rate for test of factor A

Test Statistic	Population	n					
		5	10	15	20	25	30
QW	Normal	**0.080**	0.058	0.066	0.049	0.058	0.046
	Exponential	**0.075**	0.061	0.052	0.054	0.059	0.062
	Lognormal	**0.085**	0.062	0.053	0.052	0.054	0.052
	Mixed normal	**0.078**	0.059	0.052	0.050	0.046	0.047
QM	Normal	**0.089**	0.062	0.066	0.045	0.050	0.059
	Exponential	**0.076**	**0.066**	0.064	0.063	0.049	0.063
	Lognormal	**0.087**	0.059	0.062	0.045	0.048	0.049
	Mixed normal	**0.089**	0.063	0.052	0.061	0.051	0.053
AW	Normal	0.048	0.046	0.057	0.046	0.047	0.050
	Exponential	0.040	0.048	0.047	0.052	0.054	0.057
	Lognormal	0.060	0.050	0.051	0.044	0.046	0.052
	Mixed normal	0.045	0.046	0.047	0.044	0.040	0.068
AM	Normal	0.058	0.051	0.056	0.042	0.049	0.052
	Exponential	0.043	0.049	0.054	0.063	0.048	0.060
	Lognormal	0.053	0.046	0.043	0.043	0.042	0.042
	Mixed normal	0.054	0.051	0.048	0.055	0.052	0.048
QN	Normal	**0.071**	0.063	0.056	0.042	0.049	0.043
	Exponential	**0.074**	0.061	0.056	0.060	0.054	0.062
	Lognormal	**0.082**	0.064	0.055	0.050	0.048	0.050
	Mixed normal	**0.068**	0.059	0.055	0.057	0.041	0.063
QE	Normal	**0.070**	0.063	0.056	0.041	0.049	0.042
	Exponential	**0.074**	0.061	0.055	0.060	0.055	0.064
	Lognormal	**0.082**	0.063	0.055	0.050	0.048	0.051
	Mixed normal	**0.070**	0.058	0.056	0.055	0.040	0.064
AN	Normal	0.048	0.047	0.045	0.039	0.043	0.044
	Exponential	0.037	0.051	0.052	0.054	0.052	0.058
	Lognormal	0.057	0.053	0.055	0.052	0.047	0.055
	Mixed normal	0.046	0.050	0.048	0.048	0.045	0.062
AE	Normal	0.047	0.048	0.043	0.039	0.045	0.044
	Exponential	0.038	0.052	0.052	0.054	0.053	0.058
	Lognormal	0.057	0.053	0.054	0.053	0.047	0.055
	Mixed normal	0.046	0.049	0.048	0.049	0.046	0.063

Table 2: Type I error rate for test of interaction

Test Statistic	Population	n					
		5	10	15	20	25	30
QW	Normal	**0.245**	**0.113**	**0.098**	**0.075**	**0.067**	0.061
	Exponential	**0.211**	**0.115**	**0.104**	**0.079**	**0.068**	**0.065**
	Lognormal	**0.225**	**0.116**	**0.090**	**0.076**	**0.067**	0.059
	Mixed normal	**0.211**	**0.134**	**0.094**	**0.073**	**0.065**	0.059
QM	Normal	**0.230**	**0.100**	**0.091**	**0.069**	**0.067**	**0.068**
	Exponential	**0.204**	**0.115**	**0.089**	**0.097**	**0.077**	**0.081**
	Lognormal	**0.218**	**0.123**	**0.089**	**0.072**	**0.091**	**0.065**
	Mixed normal	**0.236**	**0.141**	**0.089**	**0.072**	**0.065**	**0.076**
AW	Normal	0.050	0.045	0.048	0.051	0.047	0.043
	Exponential	0.042	0.044	0.064	0.050	0.044	0.037
	Lognormal	0.047	0.053	0.046	0.046	0.051	0.044
	Mixed normal	0.043	0.064	0.052	0.045	0.042	0.044
AM	Normal	0.038	0.046	0.049	0.038	0.048	0.051
	Exponential	0.037	0.045	0.043	0.055	0.047	0.045
	Lognormal	0.046	0.045	0.048	0.045	0.064	0.049
	Mixed normal	0.044	0.058	0.045	0.046	0.044	0.063
QN	Normal	**0.213**	**0.101**	**0.085**	**0.079**	0.064	0.060
	Exponential	**0.184**	**0.098**	**0.105**	**0.077**	**0.066**	0.062
	Lognormal	**0.196**	**0.113**	**0.085**	**0.071**	0.061	0.058
	Mixed normal	**0.198**	**0.116**	**0.088**	**0.074**	**0.067**	0.053
QE	Normal	**0.211**	**0.100**	**0.084**	**0.079**	0.063	0.060
	Exponential	**0.182**	**0.097**	**0.103**	**0.078**	**0.065**	0.063
	Lognormal	**0.192**	**0.115**	**0.084**	**0.072**	0.060	0.056
	Mixed normal	**0.192**	**0.114**	**0.088**	**0.073**	**0.066**	0.053
AN	Normal	0.054	0.048	0.048	0.049	0.047	0.046
	Exponential	0.042	0.037	0.060	0.049	0.048	0.050
	Lognormal	0.048	0.053	0.044	0.045	0.048	0.050
	Mixed normal	0.045	0.060	0.048	0.044	0.043	0.041
AE	Normal	0.053	0.048	0.046	0.048	0.046	0.044
	Exponential	0.042	0.036	0.060	0.050	0.048	0.050
	Lognormal	0.046	0.053	0.044	0.046	0.048	0.050
	Mixed normal	0.044	0.059	0.048	0.046	0.040	0.040

.

Table 3: Power for test of factor A at $\alpha_1 = 0.5$, $\alpha_2 = 0$, $\alpha_3 = -0.5$, $\alpha_4 = 0$

Test Statistic	Population	n					
		5	10	15	20	25	30
QW	Normal	0.748	0.973	0.996	1.000	1.000	1.000
	Exponential	0.145	0.203	0.251	0.334	0.400	0.458
	Lognormal	0.765	0.972	0.996	0.998	1.000	1.000
	Mixed normal	0.840	0.989	1.000	1.000	1.000	1.000
QM	Normal	0.096	0.123	0.130	0.181	0.199	0.223
	Exponential	0.045	0.058	0.074	0.085	0.090	0.110
	Lognormal	0.115	0.123	0.139	0.157	0.224	0.241
	Mixed normal	0.099	0.137	0.178	0.191	0.233	0.275
AW	Normal	0.697	0.968	0.997	1.000	1.000	1.000
	Exponential	0.099	0.170	0.234	0.315	0.385	0.432
	Lognormal	0.709	0.968	0.997	0.999	1.000	1.000
	Mixed normal	0.794	0.986	1.000	1.000	1.000	1.000
AM	Normal	0.050	0.097	0.120	0.169	0.185	0.218
	Exponential	0.040	0.044	0.062	0.078	0.091	0.108
	Lognormal	0.062	0.088	0.098	0.144	0.198	0.228
	Mixed normal	0.058	0.112	0.157	0.178	0.221	0.263
QN	Normal	0.749	0.980	1.000	1.000	1.000	1.000
	Exponential	0.147	0.203	0.264	0.346	0.421	0.474
	Lognormal	0.781	0.976	0.997	1.000	1.000	1.000
	Mixed normal	0.842	0.989	1.000	1.000	1.000	1.000
AN	Normal	0.719	0.978	0.999	1.000	1.000	1.000
	Exponential	0.096	0.174	0.247	0.328	0.407	0.460
	Lognormal	0.733	0.976	0.997	1.000	1.000	1.000
	Mixed normal	0.807	0.989	1.000	1.000	1.000	1.000

Table 4: Power for tests of interaction when main effects are null

Test Statistic	Population	n					
		5	10	15	20	25	30
QW	Normal	0.909	0.938	0.995	0.998	1.000	1.000
	Exponential	0.851	0.886	0.950	0.977	0.989	0.999
	Lognormal	0.916	0.938	0.992	0.997	1.000	1.000
	Mixed normal	0.902	0.944	0.991	0.999	1.000	1.000
QM	Normal	0.668	0.722	0.754	0.794	0.798	0.858
	Exponential	0.693	0.731	0.793	0.814	0.838	0.880
	Lognormal	0.711	0.718	0.747	0.808	0.816	0.860
	Mixed normal	0.699	0.700	0.755	0.787	0.792	0.859
AW	Normal	0.156	0.372	0.769	0.910	0.983	0.993
	Exponential	0.088	0.257	0.461	0.661	0.835	0.919
	Lognormal	0.140	0.376	0.712	0.911	0.982	0.993
	Mixed normal	0.157	0.408	0.769	0.918	0.988	0.996
AM	Normal	0.039	0.049	0.088	0.158	0.191	0.297
	Exponential	0.026	0.068	0.107	0.196	0.247	0.320
	Lognormal	0.031	0.080	0.114	0.137	0.196	0.317
	Mixed normal	0.023	0.063	0.090	0.165	0.224	0.310
QN	Normal	0.574	0.639	0.667	0.683	0.741	0.773
	Exponential	0.439	0.477	0.544	0.620	0.729	0.768
	Lognormal	0.585	0.620	0.645	0.728	0.750	0.805
	Mixed normal	0.619	0.650	0.691	0.702	0.767	0.836
AN	Normal	0.018	0.018	0.060	0.083	0.143	0.168
	Exponential	0.016	0.021	0.023	0.025	0.027	0.028
	Lognormal	0.013	0.026	0.042	0.072	0.093	0.190
	Mixed normal	0.014	0.031	0.044	0.077	0.137	0.200

Table 5: Power for tests of interaction when main effects are non-null

Test Statistic	Population	n					
		5	10	15	20	25	30
QW	Normal	0.973	0.999	1.000	1.000	1.000	1.000
	Exponential	0.898	0.950	0.987\	0.998	1.000	1.000
	Lognormal	0.984	1.000	1.000	1.000	1.000	1.000
	Mixed normal	0.981	1.000	1.000	1.000	1.000	1.000
QM	Normal	0.814	0.838	0.870	0.946	0.964	0.989
	Exponential	0.832	0.777	0.820	0.872	0.923	0.966
	Lognormal	0.821	0.822	0.870	0.925	0.965	0.984
	Mixed normal	0.819	0.839	0.879	0.918	0.966	0.989
AW	Normal	0.408	0.917	0.996	1.000	1.000	1.000
	Exponential	0.146	0.454	0.734	0.902	0.975	0.996
	Lognormal	0.420	0.932	0.999	1.000	1.000	1.000
	Mixed normal	0.446	0.962	0.999	1.000	1.000	1.000
AM	Normal	0.055	0.170	0.289	0.448	0.591	0.731
	Exponential	0.059	0.124	0.216	0.349	0.451	0.605
	Lognormal	0.041	0.147	0.307	0.461	0.587	0.752
	Mixed normal	0.052	0.179	0.307	0.438	0.616	0.767
QN	Normal	0.950	0.997	1.000	1.000	1.000	1.000
	Exponential	0.745	0.672	0.679	0.715	0.757	0.827
	Lognormal	0.971	0.997	1.000	1.000	1.000	1.000
	Mixed normal	0.978	0.998	1.000	1.000	1.000	1.000
AN	Normal	0.239	0.758	0.970	0.997	1.000	1.000
	Exponential	0.024	0.057	0.077	0.116	0.160	0.223
	Lognormal	0.250	0.761	0.969	0.998	1.000	1.000
	Mixed normal	0.309	0.860	0.984	1.000	1.000	1.000

DISCUSSION

The results show that QW, QM, QN and QE tests for both main effect and interaction are not robust for validity, especially for small sample sizes. The results also show that the Type I error rates for AW, AM, AN and AE tests are within the interval (0.036 0.064) for all sample sizes, factor effects and populations studied. Therefore, AW, AM, AN and AE tests are robust for validity.

The tests QW, AW, QN and AN are found to be powerful in testing main effect or interaction. Low power was observed for QM and AM tests for small sample sizes. A slit power advantage was observed for QN and QW tests over other tests for interaction in small sample sizes. In terms of power, QW, AW, QN and AN tests are robust for efficiency. However, only AW and AN tests are found to be robust for both validity and efficiency.

CONCLUSION

Monte Carlo simulation was performed to compare Wald–Type test and ANOVA–Type test for two-factor designs using rank score functions Wilcoxon-score, Mood-score, normal-score and expected normal- score. The results show that ANOVA–Type test on Wilcoxon-score, normal-score and expected normal-score is robust for both validity and efficiency, while the Wald-Type test on the score functions is non-robust.

REFERENCES

Akritas, M.G. and Arnold, S.F., (1994). Fully nonparametric hypotheses for factorial designs I: multivariate repeated measures designs. *Journal* *of the American Statistical Association*, **89:** 336–343.

Akritas, M.G., Arnold, S.F. and Brunner, E. (1997). Nonparametric Hypothesis and rank Statistics for Unbalanced Factorial Designs. *Journal of the American Statistical Association*, **92:** 258 – 265.

Box, G.E.P. (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Effect of inequality of variance in the one-way classification. *The Annals of Mathematical Statistics*, **25:** 290–302.

Brunner, E., Dette, H. and Munk, A. (1997). Box-type approximations in nonparametric factorial designs. *Journal of the American Statistical Association*, **92:** 1494– 1502

Brunner, E and Puri, M.L. (2002). A class of rank-score tests in factorial designs. *Journal of Statistical Planning and interference,* **103:** 331 –360.

Brunner, E., Puri, M.L., Sun, S. (1995). Nonparametric methods for stratified two-sample designs with application to multi clinic trials. *Journal of the American Statistical Association*, **90**, 1004–1014.

Conover, W.J. (1999). *Practical Nonparametric Statistics*. Wiley, New York

Harter,H. L. (1961). Expected values of normal order Statistics. *Biometrika*, **48(1 & 2):** 151 – 165.

Hora, S.C., Conover, W.J., (1984). The F-statistic in the two-way layout with rank score transformed data. *Journal of the American Statistical Association*, **79:** 668–673.

Lemmer, H.H., Stoker, D.J. (1967). A distribution-free analysis of variance for the two-way classification. *South African Statistical Journal,* **1:** 67–74.

Lin, H. and Myers, L. (2006). Power and Type I error rates of goodness-of-fit statistics for binomial generalized estimating equations models. *Computational Statistics and Data Analysis,* **50:** 3432 – 3448.

Rinaman, W.C. Jr., (1983). On distribution-free rank tests for two-way layouts. *Journal of the American Statistical Association*, **78:** 655–659.

Royston,J. P. (1981). Expected Normal Order Statistics (Exact and Approximate) As 177. *Applied Statistics*, **31:** 161 – 165.

Sawilowsky, S. S. (1990). Nonparametric Tests of Interaction in Experimental Design. *Review of Educational Research*, **60**(1), 91 – 129.

Steidl, R.J. and Thomas, L. (2001). *Power analysis and experimental design*. Pp 14-36 In Scheiner, S.M. and J. Gurevitch (eds) Design and Analysis of Ecological Experiments. 2nd Edition. Oxford University Press, New York.