Sinop Üniversitesi Fen Bilimleri Dergisi *Sinop Uni J Nat Sci 5 (2): 125-137(2020)*
ISSN: 2536-4383

Research Article **https://doi.org/10.33484/sinopfbd.641094**

Comparative Study on Synchronized Permutation Tests in Balanced $I \times J$ Designs in Case of Non-Normally Distributed Error Terms

Dilşad YILDIZ KAÇAR, Fikri GÖKPINAR, Hülya BAYRAK*

Gazi University, Faculty of Science, Department of Statistics,06500 Ankara, Turkey

Abstract

ANOVA methodology is one of the important models in the experimental design theory. Synchronized permutation provides exact solutions in factorial designs within a nonparametric framework. The synchronized permutation tests have quadratic form test statistics. In this paper, we study the synchronized permutation tests with absolute value form in the test statistics for main factors and interaction for $I \times J$ balanced factorial designs in case of non-normal distributed error terms. The behavior of synchronized permutation tests with the absolute value form is investigated and the quadratic form in the test statistics in a comparative simulation study with the parametric ANOVA is reported in 3×2 design when the error terms are normally and non-normally distributed.

Keywords: ANOVA, Synchronized permutation tests, Permutation tests.

Hata Terimlerinin Normal Dağılıma Sahip Olmadığı Durumlarda $I \times J$ Dengeli Tasarımlarda Senkronize Permütasyon Testleri Üzerine Karşılaştırmalı Bir Çalışma

Öz

ANOVA metodolojisi, deney tasarım teorisindeki önemli modellerden biridir. Senkronize permütasyon testleri faktöriyel tasarımlarında parametrik olmayan bir çerçevede kesin bir çözüm sunmaktadırlar. Senkronize permütasyon testleri karesel formda test istatistiklerine sahiptir. Bu çalışmada $I \times J$ dengeli faktöriyel tasarımlar için hata terimlerinin normal dağılıma sahip olmadığı durumlarda, ana faktörlerin etkisi ve etkileşim etkisi için önerilen test istatistiklerinin sahip olduğu karesel formun mutlak değer ile değiştirilerek senkronize permütasyon testinin davranışı incelenmiştir. Daha sonra, 3×2 dengeli tasarımı için hataların normal dağılmadığı durumlarda senkronize permütasyon testlerinin mutlak değer formda ve karesel formda test istatistiklerinin davranışını araştırıp, parametrik ANOVA ile karşılaştırmalı bir simülasyon çalışması sunuldu.

Anahtar Kelimeler: ANOVA, Senkronize Permütasyon Testi, Permütasyon Testi.

Introduction

In Design of Experiments, factorial designs are used in many experimental fields, e.g. agriculture, biology, medicine or engineering. Factorial experiments allow

*Corresponding Author: ORCID ID: https://orcid.org/ 0000-0002-6533-8649
e-mail: yldzdilsad@gmail.com
Received: 31.10.2019
Accepted: 22.05.2020

analyzing the main effect of each factor and the response variable in terms of the interactions between factors.

A two-factor factorial design is an experimental design in which data is collected for all possible combinations of the levels of the two factors of interest. The two-factor factorial design is termed as a balanced two-factor factorial design when the number of observations for each possible factor combinations is equal. A widely used parametric method for analyzing the data is the analysis of variance (ANOVA) for factorial designs. ANOVA, the classical F-test, is based on the assumptions of independence of the observations, homogeneity (equal variances) and normality (normally distributed errors). If the assumptions of the variance analysis (normality, homogeneity of variance, independence of the observations) are not satisfied, the parametric approaches may not be appropriate.

Arnold [1] and Mansouri and Chang [2] show that the parametric F test is not robust when it is applied to heavy-tailed or asymmetric distributions. Nonparametric solutions are recommended for these situations. One of these solutions is applying the resampling methods such as permutation tests.

The idea of permutation test was first examined under the name of randomization test by Fisher [3]. The test shows that instead of comparing the actual value of a test statistic with a standard statistical distribution, the data can be compared with the reference distribution obtained from it [4].

In experimental designs, there are several approaches for permutation of raw data. Manly [5] introduced the unrestricted permutation which permutes the observations within all combinations of the two factor levels without any distinction. Good [6] and Edgington [7] introduced the restricted permutation which creates the test statistic distribution via permutations of the observations which are performed only within levels of not-under-testing factors. The synchronized permutation test, which is based on restricted permutation was introduced by Salmaso [8] for 2^k design and by Basso et al [9] for replicated $I \times J$ factorial designs, respectively.

Uncorrelated exact tests for the main effects and for the interaction effect can be obtained with synchronized permutation tests. Thus, synchronized permutation allows us to test main effects and the interaction effect separately.

Synchronized permutation test for replicated $I \times J$ factorial designs were

ISSN: 2536-4383

introduced by Basso et al. [9], which uses the quadratic form in the test statistics for main effects and interaction effects. The aim of this paper is to investigate whether replacing the quadratic form in the test statistic with another even function or absolute value will have an advantage in terms of the power of test when the error terms have a skewed or a symmetric distribution in the framework of Synchronized permutation test introduced by Basso et al. [9]. For this study, we need the level of at least one factor is greater than two. We focus on 3×2 replicated balanced designs throughout the study.

Model and Hypotheses

A balanced $I\times J$ factorial design is a factorial design which has *I* number of levels for factor *A*, *J* number of levels for factor *B* and *n* independent replications carried out at each of the $I\times J$ factor (treatment) combinations. A two-factor ANOVA model might be indicated by the equation given below:

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk} \qquad (1)$$

where the first factor has, *I* levels (*i*=1, …, *I*) and the second has *J* levels (*j*=1, …, *J*) and *k* is the index of the replicate (*k*=1, …, *n*). y_{ijk} represents the dependent variables which is the combinations of *i-th* level of factor *A*, *j-th* level of factor *B* and *k-th* observation value, μ is the general mean, α_i is the main factor effects of *i-th* level of factor *A*, β_j is the main effects of *j-th* level of factor *B* and $(\alpha\beta)_{ij}$ is the interaction effects between α_i and β_j, ε_{ijk} refers to exchangeable experimental errors which has an unknown continuous distribution with zero mean and σ^2 variance. Zero-sum constraints for parameters are

$\sum_i \alpha_i = 0$, $\sum_j \beta_j = 0$ and $\sum_i (\alpha\beta)_{ij} = 0 \ \forall j$, $\sum_j (\alpha\beta)_{ij} = 0 \ \forall i$. The hypotheses of interest are given below:

$$H_{0A}: \alpha_i = 0, \forall i, H_{1A}: \exists i: \alpha_i \neq 0$$
$$H_{0B}: \beta_j = 0, \forall j, H_{1B}: \exists j: \beta_j \neq 0$$
$$H_{0AB}: \alpha\beta_{ij} = 0, \forall i,j, H_{1AB}: \exists i,j: \alpha\beta_{ij} \neq 0$$

Synchronized Permutation Tests

Salmaso [8] proposed an exact permutation test called synchronized permutation test, which is based on restricted permutation procedure for hypothesis testing on the main effects of the two factors *A* and *B* and the interaction effect for $I\times J$ factorial designs. Synchronization means that data are permuted within blocks built by one of the factors [9]. Synchronized permutation solution allows us to test main effects and interaction effect separately. There are two

ISSN: 2536-4383

ways to obtain a synchronized permutation: exchanging units in the same positions within blocks (Constrained Synchronized Permutations, CSPs) or exchanging units without considering their original position into blocks (Unconstrained Synchronized Permutations, USPs) [9].

When the sample size n of each block is relatively small, say $n < 5$, the USPs are claimed to be preferred, while if $n \geq 5$, CSPs are claimed to be preferred since the total number of permutations becomes large enough to have a decent permutation distribution [9]. When we focus on the hypothesis testing for the effect of factor *A* with synchronized permutation test, the proposed test statistic becomes

$$T_A^* = \sum_{s<i} \left(\sum_j {}^aT_{si|j}^* \right)^2 \tag{2}$$

where

$$\sum_j {}^aT_{si|j} = \sum_j \left\{ (n-2\nu)(\alpha_s - \alpha_i) + \tilde{e}_{si|j} \right\} \tag{3}$$

and $1 \leq s < i \leq I$. ${}^aT_{si|j}$ is defined as a statistic to compare the effect of levels *i* and *s* of factor *A* at level *j* of factor *B* [8] separately. In the same structure for testing effects of factor *A*, the test statistic for effects of factor *B* is

$$T_B^* = \sum_{j<h} \left(\sum_i {}^bT_{jh|i}^* \right)^2 \tag{4}$$

where

$$\sum_i {}^bT_{jh|i} = \sum_j \left\{ (n-2\nu)(\beta_j - \beta_h) + \tilde{e}_{hj|i} \right\} \tag{5}$$

and $1 \leq j < h \leq J$ [8]. ${}^bT_{jh|i}$ is defined as a statistic to compare the effect of levels *j* and *h* of factor *B* at level *i* of factor *A* separately. "ν" is the number of units, which are randomly exchanged between blocks and "ν" is invariant with respect to the levels of factor which are not under testing. $\tilde{e}_{hj|i}$, $\tilde{e}_{si|j}$ are permutation error terms [8].

There are two alternatives for in order to test the interaction effects. One of them is obtained from factor *A* as follows:

$${}^aT_{AB}^* = \sum_{s<i} \sum_{j<l} \left({}^aT_{si|j}^* - {}^aT_{si|l}^* \right)^2 \tag{6}$$

Other alternative for testing the interaction effects is obtained from factor *B* as follows:

$${}^bT_{AB}^* = \sum_{j<l} \sum_{s<i} \left({}^bT_{jl|i}^* - {}^bT_{jl|s}^* \right)^2 \tag{7}$$

[8]. For all test statistics (2), (4), (6) and (7), square operations are taken in order to avoid some levels of under testing factors and the interaction vanishing [10]. Corain and Salmaso [10] suggested that it is possible to replace the quadratic form in the test statistics with any even function of the partial statistics.

In this study, we replace the quadratic form in the test statistics with absolute value in order to investigate the behavior of two different forms of test statistics when the error terms are distributed symmetric and skewed under the same synchronized permutation test

ISSN: 2536-4383

procedure for balanced replicated $I \times J$ factorial designs. We then use the absolute value function in the test statistics to avoid some levels of under testing factors and the interaction vanishing. The quadratic form in the test statistics (2), (4), (6) and (7) is replaced with their absolute value and the test statistics for testing factor effects and interaction effects are obtained as given below:

$$U_A^* = \sum\nolimits_{s<i} \left| \left(\sum\nolimits_j {}^aT_{si|j}^* \right) \right| \quad (8)$$

$$U_B^* = \sum\nolimits_{j<h} \left| \left(\sum\nolimits_i {}^bT_{jh|i}^* \right) \right| \quad (9)$$

$${}^aU_{AB}^* = \sum\nolimits_{s<i} \sum\nolimits_{j<l} \left| \left({}^aT_{si|j}^* - {}^aT_{si|l}^* \right) \right| \quad (10)$$

where the test statistics is obtained over factor *A*,

$${}^bU_{AB}^* = \sum\nolimits_{j<l} \sum\nolimits_{s<i} \left| \left({}^bT_{jl|i}^* - {}^bT_{jl|s}^* \right) \right| \quad (11)$$

where the test statistics is obtained over factor *B*.

The absolute value function in the test statistics (8), (9), (10) and (11) might lead them to be less sensitive to skewed distributions and less outlier than using square operation in test statistics (2), (4), (6) and (7). Therefore, in this study, we investigate that if absolute vale function is used instead of squaring, it will give a more proper solution for skewed distribution and it will give parallel results for symmetrical distributions.

The quadratic expressions in the test statistics (2), (4), (6) and (7) produce a symmetry, for example, when testing factor *A*, which we can express as a row effect, there are two different permutations that give the same value of the test statistic (2). The absolute value operator, which we put in the place of the quadratic expression in the statistics in our article, does not disrupt this symmetrical structure of the test statistics. As to the permutation approach discussed in our article, when the absolute value operator is used in the test statistics, each value of the test statistics has the same probability in the null distribution of the test statistics.

We provide comparative power simulation study among synchronized permutation tests (with the quadratic form and the absolute value form of test statistics) and ANOVA to investigate their behavior under different symmetric distributed and skewed distributed error terms with homoscedastic variances.

Simulation study

In this section, we report a performed power simulation study to investigate the behavior of the synchronized permutation tests with the quadratic form of the test statistics, with the absolute value form of test statistics and ANOVA. For the

ISSN: 2536-4383

error terms, different symmetric and skewed distributions are used:

Symmetrical distributions: Normal (0,1), Logistic (0,1)

Skewed distributions: Log-Normal (where scale parameter value is 0, shape parameters values are 1.5 and 2), Inverse Gaussian (where scale parameter values are 0.005, 0.01 and shape parameter value is 1), Weibull (where scale parameter values are 0.25, 0.50 and shape parameter value is 1). All skewed distributions mentioned above are standardized to achieve an expected value of 0 for the error terms. For the simulation study, we focus on the main effect *A* and the interaction effect in 3×2 fixed effect balanced factorial design. For power simulation study, we set the number of replications as n=5.

In order to conduct the simulation study, different effects are set with $\mu = 0$ while generating the data. For factor *A* and the interaction effect *AB*, the main effect is set on six values which are 0, 0.20, 0.40, 0.60, 0.80 and 1, where the main effect for factor *B* is set on the value 0. While studying synchronize permutation test, we use CSP testing procedure for the main factors and the interaction effect. In tables and figures CSP and ABS_CSP correspond to that we performed constraint synchronize permutation procedure with test statistics in the quadratic form and with test statistics in the absolute value form, respectively and F correspond to parametric F test. The number of simulations is 5000 and all simulations are performed at 5% level of significance. All algorithms for the synchronized permutation test and the simulation study is written via Oracle Java programming language and Apache Commons Math 3.6 Application Programming Interface is used for drawing data from the statistical distributions focused in the article.

Figure 1 shows that when we carry out hypothesis testing on the effects of factor *A*, the power of the synchronized permutation test with the quadratic form of the statistic and the absolute value form of the test statistics are very close to each other and also that of the ANOVA test in case of both standard normal and Logistic distributed error terms. Figure 2 and Figure 3 show similar solutions with Figure 1 when we carry out hypothesis testing on the effects of factor AB even if we obtain the test statistics over factor *A* or factor *B*.

From Figure 4 to Figure 12, the figures refer to the skewed distributions selected for the simulation study. Figure 4 shows that when we carry out hypothesis testing on the effects of factor *A* with Log-

ISSN: 2536-4383

Normal (0,1.5) and Log-Normal (0,2) distributed errors, the synchronized permutation test with the absolute value form of the test statistic gives higher power than the synchronized permutation test with the quadratic form of the test statistic. On the other hand, ANOVA test seems to have a very lower power than the other two alternatives. Figure 5. and Figure 6. show that when we carry out hypothesis testing on the effects of factor *AB* even if we obtain the test statistics over factor *A* or factor *B* with Log-Normal (0,1.5) and Log-Normal (0,2) distributed errors, the synchronized permutation test with the absolute value form of the test statistic and the synchronized permutation test with the quadratic form of the test statistic give*s* proximate power to each other.

Figure 5 shows that when we perform hypothesis testing on the effects of factor *A* with Log-Normal (0,1.5) and Log-Normal (0,2) distributed errors, the synchronized permutation test with the absolute value form of the test statistic gives higher power than the synchronized permutation test with the quadratic form of the test statistic and ANOVA test seems to have a very lower power than the other two alternatives. Figure 5 and Figure 6 demonstrate that when we carry out hypothesis testing on the effects of factor *AB* even if we obtain the test statistics over factor *A* or factor *B* with Log-Normal (0,1.5) and Log-Normal (0,2) distributed errors, the synchronized permutation test with the absolute value form of the test statistic and the synchronized permutation test with the quadratic form of the test statistic give close power to each other. Figure 7 shows that when we carry out hypothesis testing on the effects of factor *A* with Inverse Gaussian (0.005,1) and Inverse Gaussian (0.01,1) distributed errors the synchronized permutation test with the absolute value form of the test statistic gives higher power than the synchronized permutation test with the quadratic form of the test statistic.

In addition, Figure 10. shows similar results with Figure 4. when we carry out hypothesis testing on the effects of factor *A* with Weibull (0.005,1) and Weibull (0.01,1) distributed errors. Figure 11. and Figure 12. show that when we carry out hypothesis testing on the effects of factor AB even if we obtain the test statistics over factor *A* or factor *B* with Weibull (0.005,1) and Weibull (0.01,1) distributed. errors, the synchronized permutation test with the absolute value form of the test statistic gives greater power than the synchronized permutation test with the quadratic form of the test statistic.

ISSN: 2536-4383

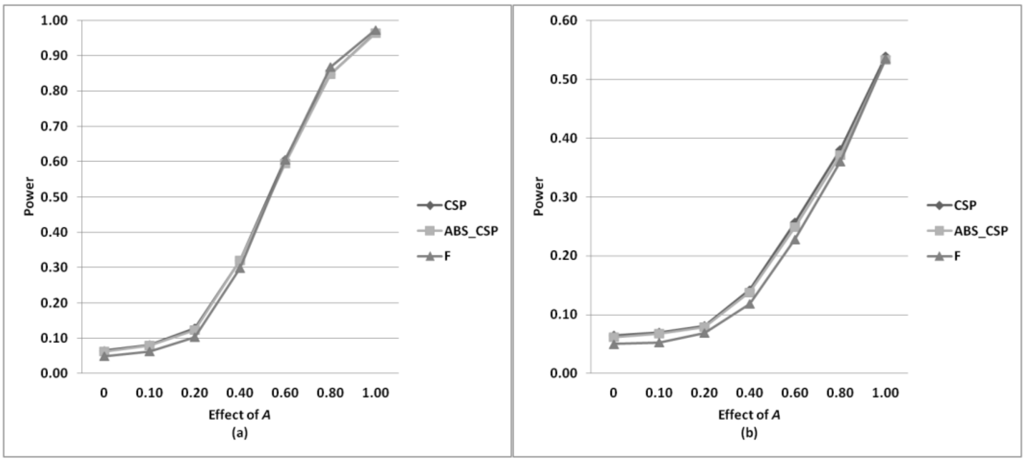

Figure 1. *Power comparison by increasing the effect of the main factor A and the standard normal distributed errors (a) and by increasing the effect of the main factor A and the Logistic distributed error terms (b).*

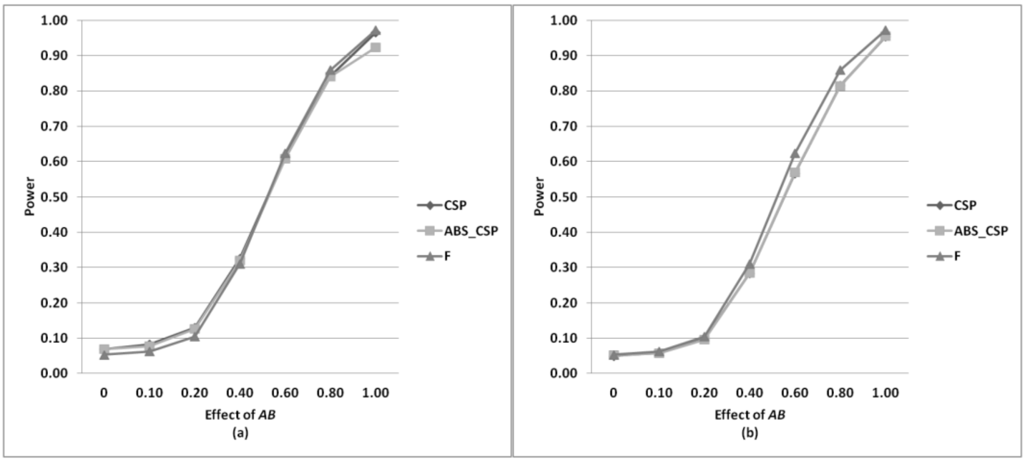

Figure 2. *Power comparison by increasing the effects of the interaction AB obtained over the main effect A with the standard normal distributed error terms (a) and by increasing the effects of the interaction AB obtained over the main effect B with the standard normal distributed error terms (b).*

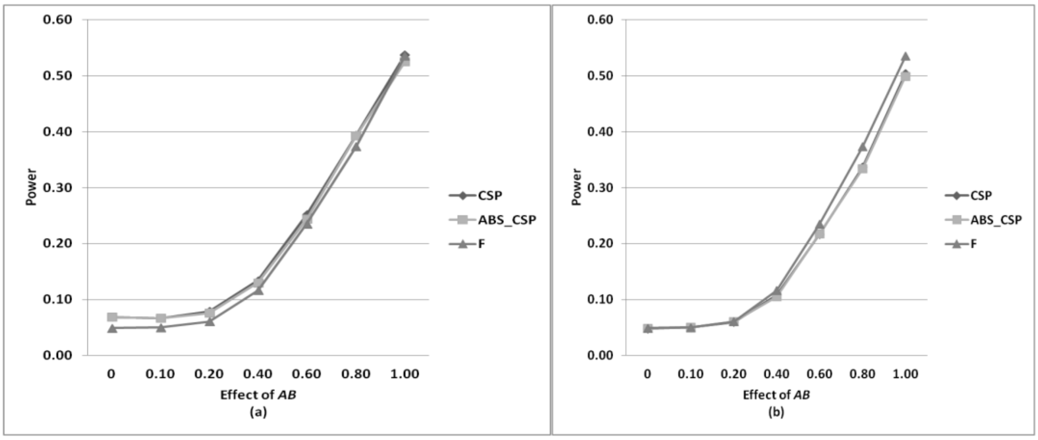

Figure 3. *Power comparison by increasing the effects of the interaction AB obtained over the main effect A with the Logistic distributed error terms (a) and by increasing the effects of the interaction AB obtained over the main effect B with the Logistic distributed error terms (b).*

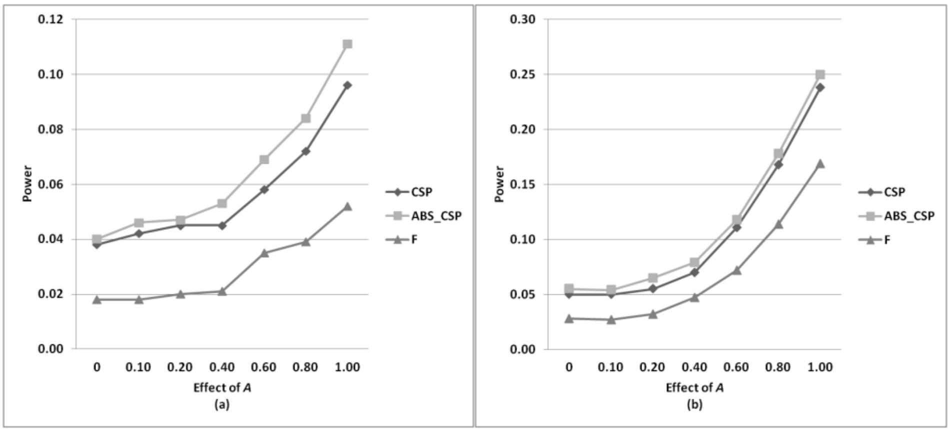

Figure 4. *Power comparison by increasing the effect of the main factor A with the Log-Normal (scale parameter is 0 shape parameter is 2) distributed error terms (a) and by increasing the effect of the main factor A and the Log-Normal (scale parameter is 0 shape parameter is 1.5) distributed error terms (b).*

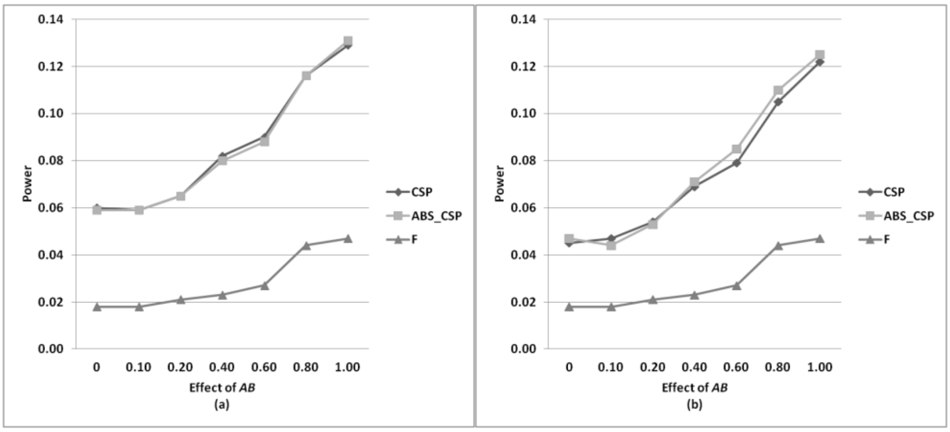

Figure 5. *Power comparison by increasing the effects of the interaction AB obtained over the main effect A with the Log-Normal (scale parameter is 0 shape parameter is 2) distributed error terms (a) and by increasing the effects of the interaction AB obtained over the main effect B with the Log-Normal (scale parameter is 0 shape parameter is 2) distributed error terms (b).*

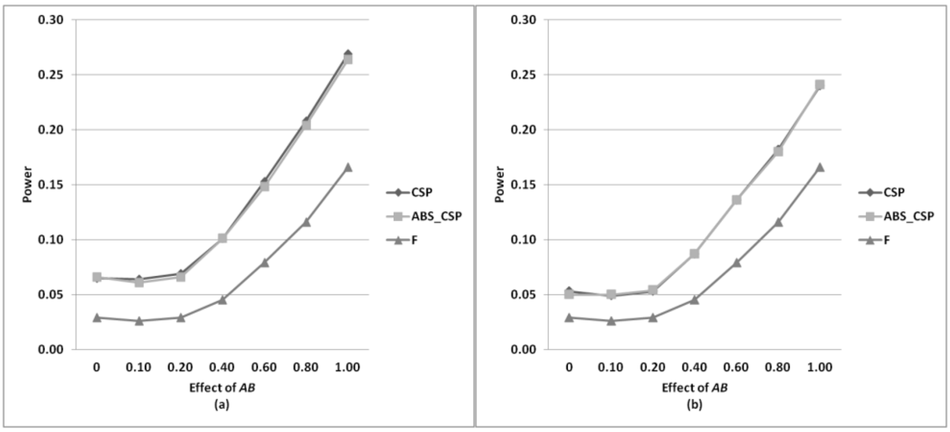

Figure 6. *Power comparison by increasing the effects of the interaction AB obtained over the main effect A with the Log-Normal (scale parameter is 0 shape parameter is 1.5) distributed error terms (a) and by increasing the effects of the interaction AB obtained over the main effect B with the Log-Normal (scale parameter is 0 shape parameter is 1.5) distributed error terms (b)*

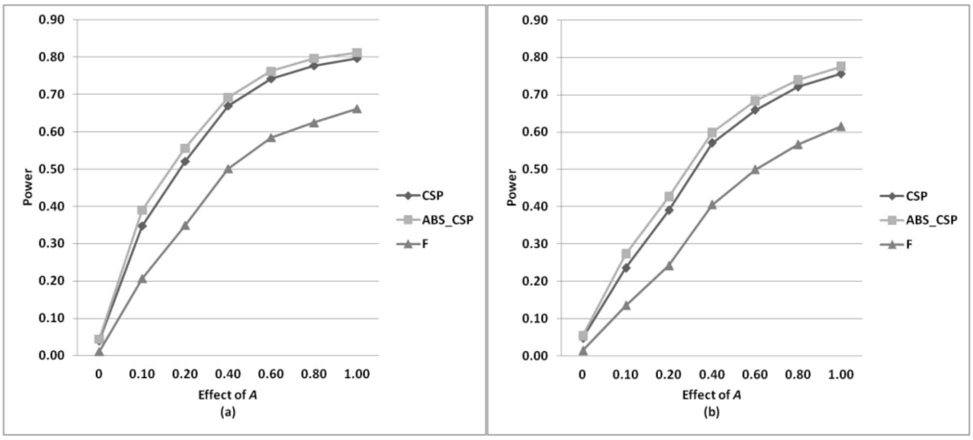

Figure 7. *Power comparison by increasing the effect of the main factor A with Inverse Gaussian (shape parameter is 0.005 parameter is 1) distributed error terms (a) and by increasing the effect of the main factor A and the Inverse Gaussian (shape parameter is 0.0,1 mean is 1) distributed error terms (b).*

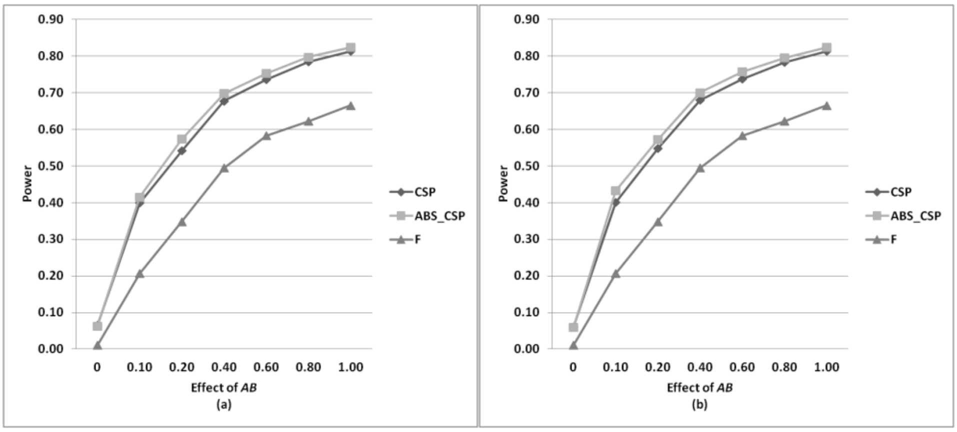

Figure 8. *Power comparison by increasing the effects of the interaction AB obtained over the main effect A Inverse Gaussian (shape parameter is 0.005 and parameter is 1) distributed error terms (a) and by increasing the effects of the interaction AB obtained over the main effect B Inverse Gaussian (shape parameter is 0.005, mean is 1) distributed error terms (b)*

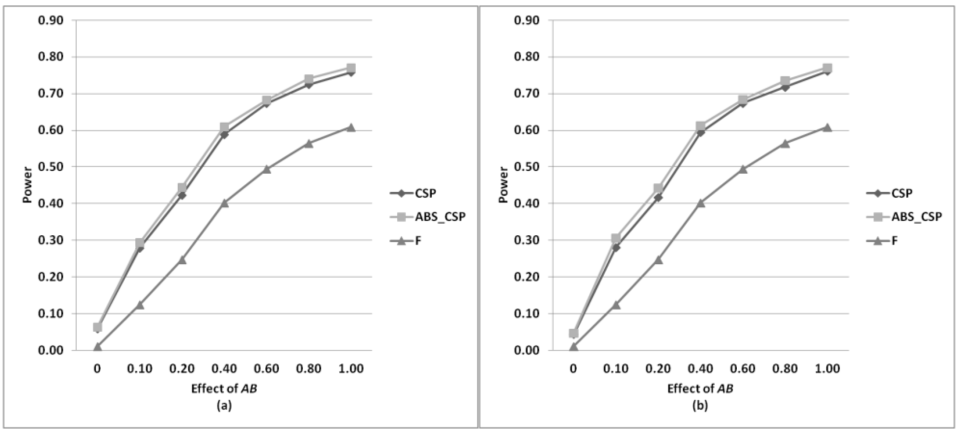

Figure 9. *Power comparison by increasing the effects of the interaction AB obtained over the main effect A Inverse Gaussian (shape parameter is 0.01, parameter is 1) distributed error terms (a) and by increasing the effects of the interaction AB obtained over the main effect B Inverse Gaussian (shape parameter is 0.01, mean is 1) distributed error terms (b).*

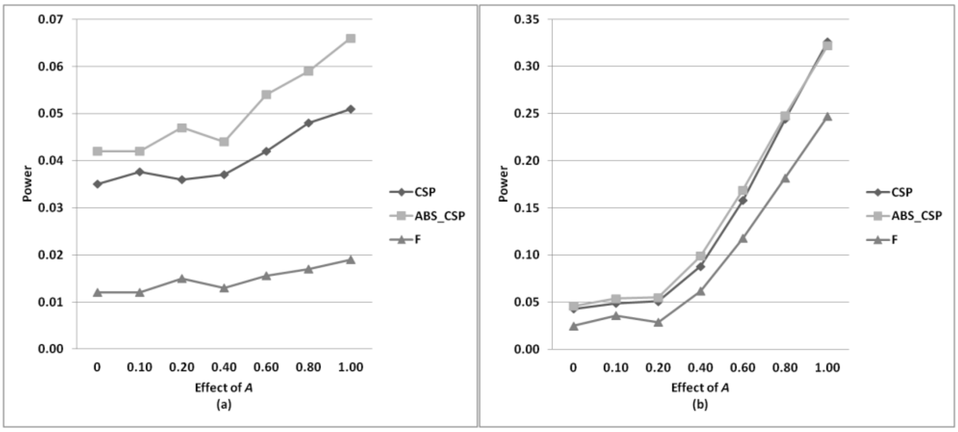

Figure 10. *Power comparison by increasing the effect of the main factor A with Weibull (shape parameter is 0.25 scale parameter is 1) distributed error terms (a) and by increasing the effect of the main factor A and the Weibull (shape parameter is 0.50 and scale parameter is 1) distributed error terms (b).*

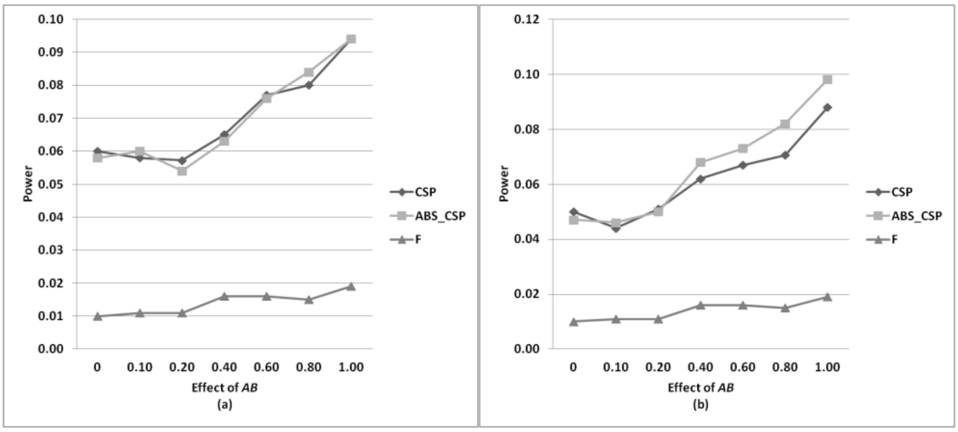

Figure 11. *Power comparison by increasing the effects of the interaction AB obtained over the main effect A Weibull (shape parameter is 0.25 scale parameter is 1) distributed error terms (a) and by increasing the effects of the interaction AB obtained over the main effect B Weibull (shape parameter is 0.25 scale parameter is 1) distributed error terms (b).*

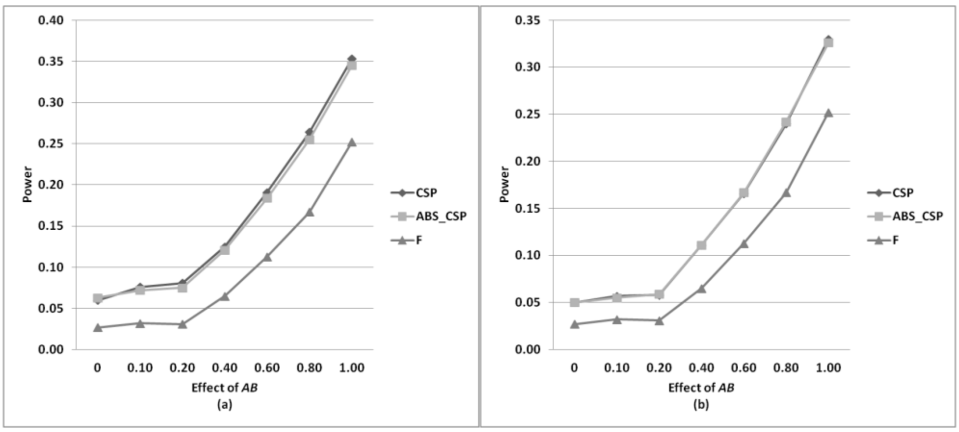

Figure 12. *Power comparison by increasing the effects of the interaction AB obtained over the main effect A Weibull (shape parameter is 0.50 scale parameter is 1) distributed error terms (a) by increasing the effects of the interaction AB obtained over the main effect B Weibull (shape parameter is 0.50, scale parameter is 1) distributed error terms (b).*

Conclusion

It can be concluded for balanced replicated $I \times J$ designs that the synchronized permutation test, which is mentioned as an exact test in the literature, is a good alternative to ANOVA when the required assumptions for ANOVA are not provided. When the error terms symmetrically distributed as normal and the Logistic in the simulation study, we observe a similar behavior in power with respect to the classical F test and synchronized permutation test, with the quadratic form of the test statistic and with the absolute value form of the test statistics. A comparative power simulation study for the examined skewed distributions demonstrates that when the distribution is highly skewed, the synchronized permutation test with the absolute value form of the test statistics generally results in a higher power compared to the quadratic form.

References

[1] Arnold HJ, 1964. Permutation Support for Multivariate Techniques. Biometrika 51: 65–70.

[2] Mansouri H, Chang GH, 1995. A Comparative Study of Some Rank Tests for Interaction. Computational Statistics & Data Analysis, 19: 85–96.

[3] Fisher R, 1935. The Design of Experiments. Oliver & Boyd Edinburgh, Scotland.

[4] Anderson MJ, Legendre P, 1999. An Empirical Comparison of Permutation Methods for Tests of Partial Regression Coefficients in a Linear Model. Journal of Statistical Computation and Simulation, 62(3):271-303.

[5] Manly B, 1997. Randomization, Bootstrap and Monte Carlo Methods in Biology. Chapman and Hall, Landon.

[6] Good P, 1994. Author Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses. Springer.

[7] Edgington ES, 1995. Randomization Tests. 3rd ed. New York: Marcel Dekker

[8] Salmaso L, 2003. Synchronized Permutation Tests in 2^k Factorial Designs. Communications in Statistic – Theory and Methods, 32(7): 1419–1437.

[9] Basso D, Chiarandini M, Salmaso L, 2006. Synchronized permutation Tests in $I \times J$ Designs. Journal of Statististical Planning and Inference, 137(8): 2564–2578.

[10] Corain L, Salmaso L, 2007. A Critical Review and a Comparative study on Conditional Permutation Tests for Two-Way ANOVA. Communications in Statistics - Simulation and Computation, 36(4): 791-805

[11] Hahn S, Konietschke F, Salmaso L. 2014. A Comparison of Efficient Permutation Tests for Unbalanced ANOVA in Two by Two Designs and Their Behavior

Under Heteroscedasticity. In: Melas V., Mignani S., Monari P., Salmaso L. (eds) Topics in Statistical Simulation. Springer Proceedings in Mathematics & Statistics, vol 114. Springer, New York, NY.