Networks and Heterogeneous Media

NHM, 20(4): 1190–1200.
DOI: 10.3934/nhm.2025051
Received: 18 June 2025
Revised: 05 October 2025
Accepted: 21 October 2025
Published: 03 November 2025

https://www.aimspress.com/journal/NHM

Research article

Saddlepoint approximation for the bivariate two-sample tests under the random allocation design

Ibrahim A. A. Shanan[1], Mona Hosny[2] and Abd El-Raheem M. Abd El-Raheem[3,*]

[1] Department of Information Technologies, Management Technical College, Southern Technical University, Basra, Iraq
[2] Department of Mathematics, College of Science, King Khalid University, Abha, 61413, Saudi Arabia
[3] Department of Mathematics, Faculty of Education, Ain Shams University, Cairo, Egypt

* **Correspondence:** Email: abdelrahimmohamed@edu.asu.edu.eg.

Abstract: This article addresses the challenge of accurately estimating p-values in bivariate two-sample tests under random allocation designs, a common setting in clinical and reliability studies. Existing normal approximations often perform poorly in small samples and in the distribution tails, leading to unreliable inference. To overcome this limitation, we propose the use of the saddlepoint approximation as a highly accurate alternative. Through simulation studies and real-data analyses, we demonstrate that the proposed method consistently yields p-values that are closer to the exact permutation values than those obtained from traditional normal approximations, particularly in small-sample settings.

Keywords: bivariate two-sample tests; saddlepoint approximation; random allocation design; nonparametric inference; P-value accuracy.

1. Introduction

In many real-world applications, bivariate analysis is advantageous because it allows for the simultaneous examination of two response variables of equal importance. Bivariate statistical analysis is well established based on the assumption of normality; nevertheless, this assumption is usually disobeyed in reality. Consequently, the non-parametric approaches become essential in bivariate analyses. Testing the equality of two bivariate distributions for two independent groups is a concern in various domains. In a comparative clinical trial, for instance, patients may experience nonfatal signs

or symptoms affecting two body systems (e.g., neurological and dermal). A similar situation occurs in reliability analyses when comparing two systems, each consisting of two components. For more applications of assessing whether two bivariate distributions are identical, see the references [1–4]. Let two independent random samples $\boldsymbol{X}$ and $\mathbf{Y}$ of sizes n_1 and n_2 (such that $n_1+n_2 = N$) are withdrawn from a continuous bivariate populations characterized by cumulative distribution functions $\boldsymbol{F}(x,y)$ and $\boldsymbol{G}(x,y)$, respectively. The issue of determining if two bivariate distributions are statistically identical can be formulated in the following null hypothesis:

$$\boldsymbol{H}_0 = \boldsymbol{F}(x,y) = \boldsymbol{G}(x,y) \tag{1}$$

Various studies have been documented in the literature to address this problem, see for example [5–7]. Mardia [8] extends both the Wilcoxon–Mann–Whitney rank-sum test and Mood's median test to address the problem of comparing location parameters in two bivariate samples. A simple nonparametric procedure for evaluating differences between two populations in a bivariate context was proposed by Williams [9]. A bivariate signed rank test as a statistical method for analyzing the two-sample location problem was introduced by Peters and Randles [10]. Hettmansperger and Oja [11] designed a test to be invariant under affine transformations and are applicable to multivariate data involving multiple samples. Sen and Mathur [12] presented a statistical test designed to address the problem of comparing location parameters in two bivariate samples. Samawi et al. [13] extended the bivariate sign test to ranked set sampling. Mathur [14] proposed a consistent test for differences in locations between two bivariate populations, that is like the Mann-Whitney test and outperforms other statistics when detecting small location shifts. Roland and Herold [15] introduced robust nonparametric tests specifically designed for the two-sample location problem, that compares the central locations of two independent samples.

Although methods for normal approximating are available for approximating the p-values of the bivariate two-sample tests, they suffer from poor accuracy in small samples and at the tails of the distribution. This shortcoming becomes critical in practical applications, such as clinical trials, where incorrect statistical decisions can have significant consequences. The current research presents the saddlepoint approximation as a solution to this practical problem. The approximation used in this work is based on the saddlepoint formulation proposed by Daniels [16,17] and enhanced by Skovgaard [18]. Saddlepoint approximations can be derived to the statistical measure that has a cumulant generating function. Notably, in cases involving small sample sizes, the saddlepoint method is capable of accurately estimating probabilities near the tails of the distribution. Many researchers have extensively utilized these approximations, resulting in significant advancements. Comprehensive discussions of their use in various distributional challenges can be found in the works of Kolassa [19], Butler [20], Abd El-Raheem et al. [21], Abd El-Raheem and Hosny [22], Shanan et al. [23], Abd El-Raheem et al. [24,25], Alhejaili and AlGhamedi, [26], and Abd El-Raheem and Hosny [27].

The primary theoretical advantage of the saddlepoint approximation over the normal approximation lies in its ability to achieve high relative accuracy, even in the tails of the distribution and with small sample sizes [20]. While the normal approximation relies on a first-order Taylor expansion (central limit theorem), the saddlepoint approximation is based on an exponential tilt of the underlying distribution and can be viewed as an optimal second-order approximation. This allows it to capture skewness and kurtosis more effectively, which is crucial for the permutation distributions of nonparametric test statistics, which are often discrete and non-normal, especially in small samples.

In clinical trials, randomization designs are employed to assign patients to one of two groups: A control group or a treatment group. This article focuses on the random allocation design, which is applied to assign experimental units or patients to these groups while maintaining balance. Specifically,

it is a forced-balance procedure that ensures equal group sizes for both the control and treatment arms. For a detailed discussion, see Rosenberger and Lachin [28]. Using the random allocation design, the permutation distribution of the considered bivariate two-sample tests, $\binom{N}{n_1}^{-1}$, is derived as a conditional distribution of Bernoulli random variables. This enables the use of the saddlepoint approximation method to compute the significance level for the bivariate two-sample tests under consideration.

The aim of this article is to develop and assess a more accurate method for estimating p-values in bivariate two-sample tests, specifically under a random allocation design commonly used in clinical trials.

The framework of the following sections is as follows: Section 2 introduces bivariate two-sample tests. Section 3 discusses the saddlepoint approximation of p-values for these tests under a random allocation design. Sections 4 and 5 present simulation studies and numerical examples that contrast the efficiency of the saddlepoint method with that of the asymptotic approach.

2. Non-parametric bivariate two-sample tests

In this section, we present some bivariate statistical tests for two samples, let $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_{n_1}$ and $\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_{n_2}$ are two separate random samples drawn from continuous bivariate populations with joint cumulative distribution functions $\boldsymbol{F}(\boldsymbol{x} - \boldsymbol{\theta})$ and $\boldsymbol{F}(\boldsymbol{y} - \boldsymbol{\theta} - \boldsymbol{\rho})$ respectively. The distribution function F has unique center, 0, (the multivariate Oja median). Hettmansperger and Oja [13] proposed multivariate two-sample location test to test the null hypothesis $\boldsymbol{H_0}: \boldsymbol{\rho} = \boldsymbol{0}$. In this section, we present the bivariate case of the statistic of Hettmansperger and Oja [13] as a special case of multivariate case. Define the aligned observations by $\boldsymbol{w_i} = \boldsymbol{x_i} - \widehat{\boldsymbol{\theta}}$ for $i = 1, \dots n_1$ and $\boldsymbol{w_i} = \boldsymbol{y_{i-n_1}} - \widehat{\boldsymbol{\theta}}$ for $i = n_1 + 1, \dots, N = n_1 + n_2$, where $\widehat{\boldsymbol{\theta}}$ is the median of the merged sample. Thus, the test statistic can be written as follows:

$$U_{HO} = \sum_{i=1}^{N} a_i \boldsymbol{q_i}, \tag{2}$$

where $a_i = 1$ for $i = 1, \dots n_1$ and $a_i = -1$ for $i = n_1 + 1, \dots, N$,

$$\boldsymbol{q_i} = \boldsymbol{q}_N(\boldsymbol{w_i}) = \frac{1}{N-1} \sum_{i=1}^{N} sgn\left\{det \begin{pmatrix} w_{i1} & w_{j1} \\ w_{i2} & w_{j2} \end{pmatrix}\right\} \begin{pmatrix} \boldsymbol{B}_1(j) \\ \boldsymbol{B}_2(j) \end{pmatrix},$$

where $\boldsymbol{B}_1(j) = -$ cofactor of w_{i1} from the matrix $\begin{pmatrix} w_{i1} & w_{j1} \\ w_{i2} & w_{j2} \end{pmatrix}$, and $\boldsymbol{B}_2(j) = -$cofactor of w_{i2} from the matrix $\begin{pmatrix} w_{i1} & w_{j1} \\ w_{i2} & w_{j2} \end{pmatrix}$, such that the vector $(\boldsymbol{B}_1(j), \boldsymbol{B}_2(j))^T = \boldsymbol{w}_j^*$ which is the vector $\boldsymbol{w}_j$ rotated 90 degrees counterclockwise. Thus,

$$\boldsymbol{q}_i = \frac{1}{N-1} \sum_{i=1}^{N} sgn\left\{det \begin{pmatrix} w_{i1} & w_{j1} \\ w_{i2} & w_{j2} \end{pmatrix}\right\} \boldsymbol{w}_j^* .$$

The term $sgn\{\det(\boldsymbol{w}_i \boldsymbol{w}_j)\}$ equals 1 if the vector $\boldsymbol{w}_i$ lies in the same half-space defined by $\boldsymbol{w}_j$ and occupied by $\boldsymbol{w}_j^*$ and is -1 otherwise. The statistic U_{HO} has conditional mean $\boldsymbol{\mu} = E(U_{HO}|\boldsymbol{H_1}) = \frac{n_1 - n_2}{N} \sum_1^N \boldsymbol{q_i}$ and its covariance matrix is $\boldsymbol{\sigma} = \frac{4 n_1 n_2}{N(N-1)} \sum_1^N \boldsymbol{q_i}\, \boldsymbol{q_i^T}$. When the samples are of equal size, $n_1 = n_2$, the conditional expectation is **0.**

The test statistic in Eq (2) is bivariate and its two components are independent and normally distributed. Following the linear form derivation common in permutation tests [29], the statistic can be reformulated as

$$U_{HO}^{*} = \sum_{i=1}^{N} a_i(\boldsymbol{q_{1i}} + \boldsymbol{q_{2i}}). \tag{3}$$

Let $b_i = \frac{a_i+1}{2}$, then $b_i = 0$ or 1. Thus, the statistic U_{HO}^{*} can be represented as follows:

$$U_{HO}^{*} = \sum_{i=1}^{N} 2b_i(\boldsymbol{q_{1i}} + \boldsymbol{q_{2i}}) - \sum_{i=1}^{N}(\boldsymbol{q_{1i}} + \boldsymbol{q_{2i}}). \tag{4}$$

The second bivariate two-sample test statistic was proposed by Mathur [14]. Let (x_{1i}, y_{1i}), $i = 1, \dots, n_1$ and (x_{2j}, y_{2j}), $j = 1, \dots, n_2$ are two independent random samples drawn from continuous bivariate populations, each defined by its cumulative distribution function $\boldsymbol{F}(x, y)$ and $\boldsymbol{G}(x, y)$, respectively. Mathur [14] introduced bivariate two-sample test to test the null hypothesis $\boldsymbol{H_0} = \boldsymbol{F}(x, y) = \boldsymbol{G}(x, y)$ for all (x, y) against the alternative hypothesis $\boldsymbol{H_1} = \boldsymbol{F}(x, y) = \boldsymbol{G}(x + \alpha_1, y + \alpha_2)$, where $(\alpha_1, \alpha_2) \neq (0,0)$.

Consider a point, (X, Y) represented in polar coordinates, corresponding to a specific location in the cartesian coordinate system. Any change in the values of x and y will alter the point's position within the coordinate system. Let d_1 and d_2 represent the linear distance from the origin to the point in sample 1 and the point in sample 2, respectively. Based on the null hypothesis, it is expected that the distances d_1 and d_2 are equal. Mathematically, this can be expressed as follows:

Let $d_1^2 = x_{1i}^2 + y_{1i}^2, i = 1, \dots, n_1$ and $d_2^2 = x_{2j}^2 + y_{2j}^2, j = 1, \dots, n_2$, Mathur [14] defined the test statistic U_M as the number of instances where d_{1i}^2 precedes d_{2j}^2 in the ranking of the two independent squared distances, which are concatenated into a single sequence of $N = n_1 + n_2$ variables in either decreasing or increasing magnitude. Thus, the test statistic is given by:

$$U_M = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} D_{ij}, \tag{5}$$

where

$$D_{ij} = \begin{cases} 1, & d_{1i}^2 > d_{2j}^2 \\ 0, & d_{1i}^2 \leq d_{2j}^2 \end{cases} \qquad i = 12, \dots . n_1;\ j = 1,2, \dots, n_2.$$

The conditional expectation and variance of the statistic U_M are $E(U_M|\boldsymbol{H_0}) = n_1 n_2 p$, and

$$Var(U_M|H_0) = n_1 n_2[p - p^2(N - 1) + (n_1 - 1)p_1 + (n_2 - 1)p_2],$$

where $N = n_1 + n_2$, and under $\boldsymbol{H_0}$: $p_1 = p_1 = 1/3\ and\ p = 1/2$.

Both statistics in Eqs (4) and (5) can be written as the following linear equation:

$$U = \sum_{i=1}^{N} \gamma_i \boldsymbol{z_i} + e, \tag{6}$$

where z_i's are weights and e is constant, γ_i are indicators for the treatment sample, the statistics U_{HO}^{*} in Eq (4) are equivalent to the Eq (6) when $\boldsymbol{z_i} = 2(\boldsymbol{q_{1i}} + \boldsymbol{q_{2i}})$ and $e = -\sum_{i=1}^{N}(\boldsymbol{q_{1i}} + \boldsymbol{q_{2i}})$. Also U_M in Eq (5) are equivalent to the Eq (6) when $\boldsymbol{z_i} = \boldsymbol{r_i}$ and $e = \frac{n_1(n_2+1)}{2}$ where $\boldsymbol{r_i}$ is the rank of the magnitudes of d_i^2 within the combined dataset.

In other words:

- In the test one U_{HO}^{*}, $\boldsymbol{z_i} = 2(\boldsymbol{q_{1i}} + \boldsymbol{q_{2i}})$ and $e = -\sum_{i=1}^{N}(\boldsymbol{q_{1i}} + \boldsymbol{q_{2i}})$.

- In the test two U_M, $\boldsymbol{z_i} = \boldsymbol{r_i}$ and $e = n_1(n_2 + 1)/2$.

The double saddlepoint approximation for the underlying permutation distribution of the statistic U, as specified by Eq (6), is derived in the next section.

3. Saddlepoint approximation under the allocation design

For assessing the p-value of the tests given in Eq (6), the saddlepoint approximation is utilized as a more refined substitute for the asymptotic normal approximation. The underlying permutation distribution of the statistic U in Eq (6) may be expressed as the conditional distribution of the random variable $\sum_{i=1}^{N} \varepsilon_i \boldsymbol{z_i} + e$ given that $\sum_{i=1}^{N} \varepsilon_i = n_1$, where the random variables $\{z_1, z_2, \dots, z_N\}$ are independent identically distributed, i.i.d., Bernoulli (θ) variables, where θ belongs to the interval $(0, 1)$. Hence, an equivalent formulation of the distribution of the statistic U in Eq (6) is

$$U = \sum_{i=1}^{N} \gamma_i \boldsymbol{z_i} + e \sim \sum_{i=1}^{N} \varepsilon_i \boldsymbol{z_i} + e \;\Big|\; \sum_{i=1}^{N} \varepsilon_i = n_1 ,$$

This approximation relies on the joint cumulant generating function of $\sum_{i=1}^{N} \varepsilon_i \boldsymbol{z_i} + e$ and $\sum_{i=1}^{N} \varepsilon_i$ corresponding to the equation above, which is given by:

$$K(\alpha, \beta) = \sum_{i=1}^{N} \log\{(1 - \theta) + \theta \exp(\alpha + \beta \boldsymbol{z_i})\} + \beta e.$$

Let us assume that the actual value of Eq (6) is U_0. Thus, we obtain the midp-value of the statistic U at U_0 as

$$mid - p(U_0) = P(U > U_0) + \frac{1}{2} P(U = U_0).$$

Here, we use mid-p values from the permutation tests to reduce the conservative bias inherent in exact p-values, a common practice that provides a better approximation to the continuous sampling distribution. In particular, the mid-p adjustment effectively corrects the discreteness of permutation distributions, yielding significance levels that are less biased and closer to the nominal level.

The $mid - p(U_0)$ may be approximated using the saddlepoint approach developed by Skovgaard [18], specifically as the conditional distribution of $P(\sum_{i=1}^{N} \varepsilon_i \boldsymbol{z_i} + e \geq U_0 \big| \sum_{i=1}^{N} \varepsilon_i = n_1)$ as follows:

$$mid - p(U_0) \simeq 1 - \Phi(\hat{u}_1) - \varphi(\hat{u}_1) \left(\frac{1}{\hat{u}_1} - \frac{1}{\hat{u}_2} \right), \tag{7}$$

where

$$\hat{u}_1 = sgn(\hat{\beta}) \left[2 \left\{ (K(\hat{\alpha}_0, 0) - n_1 \hat{\alpha}_0) - \left(K\left(\hat{\alpha}, \hat{\beta}\right) - n_1 \hat{\alpha} - U_0 \hat{\beta} \right) \right\} \right]^{1/2}$$

$$\hat{u}_2 = \hat{\beta} \left[\left| K''\left(\hat{\alpha}, \hat{\beta}\right) \right| / K''_{ss}(\hat{\alpha}_0, 0) \right]^{1/2}.$$

In the above formulas, K'' represents the 2×2 Hessian matrix of $K(\alpha, \beta)$, whereas K''_{ss} denotes the $\partial^2/\partial\alpha^2$ component of this Hessian matrix. The numerator saddlepoint $\left(\hat{\alpha}, \hat{\beta}\right)$ is the solution to the equations:

$$K'_{\alpha}(\hat{\alpha},\hat{\beta}) = \sum_{i=1}^{N} \frac{\exp(\hat{\alpha}+\hat{\beta}\boldsymbol{z_i})}{(1-\theta)/\theta + \exp(\hat{\alpha}+\hat{\beta}\boldsymbol{z_i})} = n_1,$$

$$K'_{\beta}(\hat{\alpha},\hat{\beta}) = \sum_{i=1}^{N} \frac{\boldsymbol{z_i}\exp(\hat{\alpha}+\hat{\beta}\boldsymbol{z_i})}{(1-\theta)/\theta + \exp(\hat{\alpha}+\hat{\beta}\boldsymbol{z_i})} + e = U_0,$$

while the denominator saddlepoint $\hat{\alpha}_0$ solves

$$K'_{\alpha}(\hat{\alpha}_0, 0) = \sum_{i=1}^{N} \frac{\theta \exp(\hat{\alpha}_0)}{(1-\theta)+\theta \exp(\hat{\alpha}_0)} = n_1. \tag{8}$$

Given that the computations for the permutation distribution of statistic (6) are not influenced by the specific value of θ employed, the value $\theta = n_1/N$ has been applied. This leads to a direct solution for Eq (8) as $\hat{\alpha}_0 = 0$, which leads to a simplification of the computations.

4. Simulation study

An extensive simulation study is performed to evaluate the performance of the saddlepoint approximation for approximating the p-value of the two specific tests, namely the Hettmansperger and Oja [13] test (U^*_{HO}) and the Mathur [14] test (U_M). The data is simulated using three distributions: the standard bivariate normal distribution, the bivariate exponential distribution, and the bivariate gamma distribution (skewed). A total of one thousand data sets are generated, with sample sizes of $N = 10, 20, 30, 50, 80,$ and 100 from the three considered distributions. The data set is randomly partitioned into two arms, namely treatment and control, where the sizes of the arms are equal, such that $n_1 = n_2$. The results for the three distributions are presented in Tables 1–3. Every table presented in this study contains the following set of information: Sad.P.' indicates the fraction of the 1,000 datasets where the saddlepoint p-value more closely approximated the simulated midp-value compared to the asymptotic p-value. On the other hand, the term "E.Sad." denotes the average relative absolute error that exists between the saddlepoint p-value and the simulated midp-value. The remaining items pertain to the normal approximation of the same assessments.

Table 1. Performance assessment using simulated data from the bivariate exponential distribution.

Test	U^*_{HO}			U_M		
N	Sad.P	E.Sad	E.Nor	Sad.P	E.Sad	E.Nor
10	77.5	0.05186	0.29713	98.5	0.00331	0.02184
20	84.9	0.02362	0.25897	96.4	0.00142	0.01423
30	85.4	0.01894	0.19353	93.8	0.00134	0.00903
50	86.3	0.01679	0.26434	86.2	0.00134	0.00515
80	91.2	0.00347	0.07641	84.5	0.00736	0.05409
100	99.0	0.00445	0.05840	87.0	0.00331	0.01946

To illustrate the results contained in the tables, we take Table 3 as an example, where the sample size $N = 10$ and the test statistic is referred to as U_M. The saddlepoint p-values had a higher degree of proximity to the simulated values, this is evident from the value of Sad.P. $= 98.5\%$. The relative absolute error for the saddlepoint p-values is 0.31%, whereas the asymptotic normal approximation

had a higher relative absolute error of 2.21%. Across all tests and conditions, the saddlepoint approximation yielded an average accuracy of 86.5% and 93.5% for the test statistics U_{HO} and U_M, respectively. In all considered scenarios, it was consistently observed that the saddlepoint approximation exhibited an elevated level of accuracy and outperformed the conventional approximation.

Table 2. Performance assessment using simulated data from the bivariate normal distribution.

Test	U_{HO}^*			U_M		
N	Sad.P	E.Sad	E.Nor	Sad.P	E.Sad	E.Nor
10	81.0	0.04080	0.32455	98.6	0.00331	0.02273
20	85.5	0.01537	0.19649	95.2	0.00148	0.01389
30	93.6	0.01417	0.24155	92.6	0.00127	0.00790
50	96.4	0.01362	0.28629	87.0	0.00135	0.00489
80	99.0	0.00236	0.16505	85.4	0.00191	0.02198
100	98.0	0.00458	0.48701	90.0	0.00679	0.00848

Table 3. Performance assessment using simulated data from the bivariate gamma distribution.

Test	U_{HO}^*			U_M		
N	Sad.P	E.Sad	E.Nor	Sad.P	E.Sad	E.Nor
10	75.3	0.04951	0.33657	98.5	0.00313	0.02212
20	87.3	0.01944	0.27900	96.6	0.00133	0.01297
30	91.1	0.01434	0.20367	93.3	0.00117	0.00791
50	93.9	0.01096	0.22636	86.3	0.00123	0.00437
80	100	0.00289	0.06239	90.0	0.00506	0.00840
100	98.1	0.00410	0.22953	81.2	0.00124	0.01947

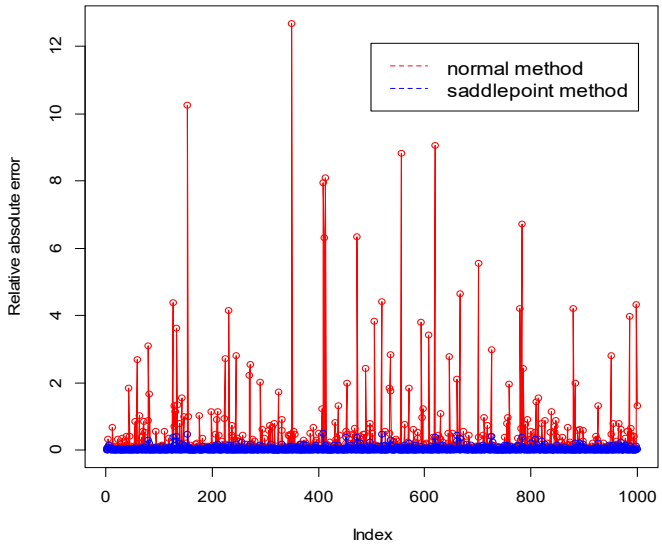

Figure 1. Relative absolute errors of saddlepoint approximation and normal approximation for test statistics U_{HO} with sample size $n = 10$ generated from bivariate normal distribution.

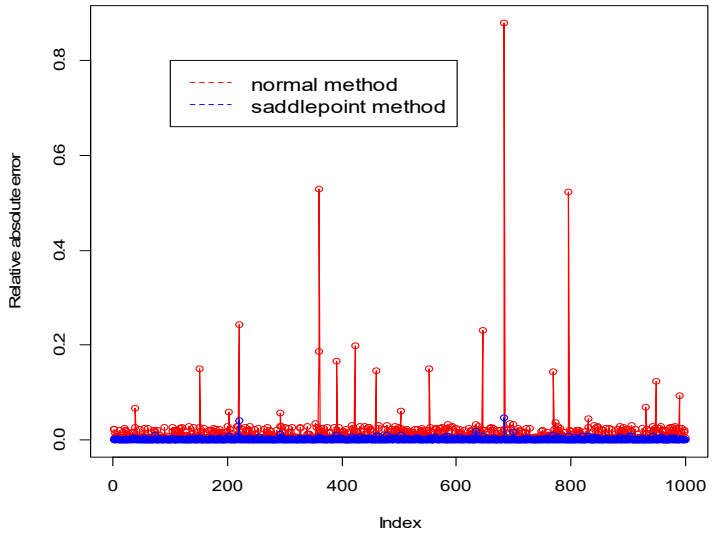

Figure 2. Relative absolute errors of saddlepoint approximation and normal approximation for test statistics U_M with sample size $n = 30$ generated from bivariate gamma distribution.

Figures 1 and 2 show the relative absolute errors of the saddlepoint approximation and the normal approximation for two simulation study instances in order to help visualize the simulation findings.

Figures 1 and 2 confirm the superiority of the saddlepoint approximation over the conventional normal approximation for estimating p-values in bivariate two-sample tests. Across both symmetric (normal) and skewed (gamma) distributions, the saddlepoint approach delivers markedly smaller relative errors, emphasizing its precision and reliability in small-sample and non-normal scenarios. These visual results reinforce the findings from the simulation tables that the saddlepoint method provides a more accurate and stable approximation to the true permutation p-values than the standard asymptotic method.

5. Numerical examples

To assess the effectiveness of the proposed approximation method in realistic applications, three real data sets with treatment–control or comparative structures are analyzed. Although these datasets originate from different fields, each reflects the essence of random allocation or comparative clinical-trial settings, where two groups are compared based on bivariate outcomes. Such situations are common in medical and reliability studies, where accurate p-value estimation is crucial for decision making.

The first dataset concerns the numbers of students in Australian higher education from 1981 to 1988 (Hand et al., [30]), divided into two age-based groups, students below and above 24 years, representing a comparison between two distinct population groups, analogous to treatment and control arms in a trial.

The second dataset records the length of stay in a psychiatric observation ward for patients admitted either compulsorily or voluntarily, with male and female categories representing the bivariate outcomes. This setting closely parallels a clinical trial structure, where the admission type corresponds to treatment allocation, and the bivariate responses capture multiple aspects of patient outcomes.

The third dataset, from Samawi and Pararai [31], compares the ages of couples practicing daily walking versus non-daily walking a structure reflecting intervention and control groups in behavioral or lifestyle intervention trials.

For each dataset, the saddlepoint p-values associated with the test statistic (6) are computed and compared with the traditional normal approximations. As shown in Table 4, in all cases the saddlepoint method provides p-values that are closer to the simulated mid-p-values, confirming its superior accuracy and reinforcing its suitability for clinical-trial-type analyses involving bivariate outcomes.

Table 4. The approximated p-values for the considered data sets.

Test statistic	Data set	Simulation	Saddlepoint	Normal
U_{HO}^{*}	Data set 1	0.230016	0.230473	0.304105
	Data set 2	0.063890	0.063895	0.063601
	Data set 3	0.186084	0.184927	0.179766
U_{M}	Data set 1	0.000185	0.000180	0.000346
	Data set 2	0.128075	0.128075	0.123997
	Data set 3	0.257129	0.257217	0.254769

6. Conclusions

This paper explores strategies for handling bivariate data, its frequent occurrences in clinical trials, and reliability research. Additionally, it suggests the saddlepoint approximation method as a precise technique for estimating the p-value of certain bivariate two-sample tests within the framework of random allocation design. A comparison between the suggested and traditional methods for approximating the p-value for the considered tests is performed by analyzing real data sets and performing a simulation study. Analyzing the real data and simulation study results show that saddlepoint approximation attains an elevated accuracy level in all scenarios compared to the normal approximation.

Use of AI tools declaration

The authors declare that they have not used Artificial Intelligence tools in the creation of this article.

Acknowledgment

The second author extends her appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through a research groups program under grant RGP2/522/46.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

A. M. Abd El-Raheem: Conceptualization, Methodology, Software, Software, Writing—review & editing; I. A. A. Shanan: Visualization, Resources, Investigation, Writing—original draft; M. Hosny: Funding acquisition, Project administration.

References

1. L. Guo, R. Modarres, Testing the equality of matrix distributions, *Stat. Methods Appl.*, **29** (2020), 289–307. https://doi.org/10.1007/s10260-019-00477-7
2. B. Sürücü, Testing for censored bivariate distributions with applications to environmental data, *Environ. Ecol. Stat.*, **22** (2015), 637–649. https://doi.org/10.1007/s10651-015-0323-x
3. H. Wendt, P. Abry, G. Didier, Wavelet domain bootstrap for testing the equality of bivariate self-similarity exponents, *IEEE Stat. Signal Process. Workshop (SSP)*, (2018), 563–567. https://doi.org/10.1109/SSP.2018.8450710
4. R. Modarres, Testing the equality and exchangeability of bivariate distributions, *Commun. Stat. Theory Methods*, **41** (2012), 3750–3758. https://doi.org/10.1080/03610926.2011.566973
5. G. K. Bhattacharyya, R. A. Johnson, A layer rank test for ordered bivariate alternatives, *Ann. Math. Stat.*, **41** (1970), 1296–1310. https://doi.org/10.1214/aoms/1177696904
6. R. A. Johnson, K. G. Mehrotra, Nonparametric tests for ordered alternatives in the bivariate case, *J. Multivar. Anal.*, **2** (1972), 219–229. https://doi.org/10.1016/0047-259x(72)90029-2
7. E. J. Dietz, Multivariate generalization of Johckheere's test for ordered alternatives, *Commun. Stat. Theory Methods*, **18** (1989), 3763–3783. https://doi.org/10.1080/03610928908830122
8. K. V. Mardia, A non-parametric test for the bivariate two-sample location problem, *J. R. Stat. Soc. Ser. B (Methodol.)*, **29** (1967), 320–342. https://doi.org/10.1111/j.2517-6161.1967.tb00699.x
9. J. S. Williams, Quick test for comparing two populations with bivariate data, *Biometrics*, **30** (1977), 215–219. https://doi.org/10.2307/2529314
10. D. Peters, R. H. Randles, A bivariate signed rank test for the two-sample location problem, *J. R. Stat. Soc. Ser. B (Methodol.)*, **53** (1991), 493–504. https://doi.org/10.1111/j.2517-6161.1991.tb01841.x
11. T. P. Hettmansperger, H. Oja, Affine invariant multivariate multisample sign tests, *J. R. Stat. Soc. Ser. B (Stat. Methodol.)*, **56** (1994), 235–249. https://doi.org/10.1111/j.2517-6161.1994.tb01974.x
12. K. Sen, S. K. Mathur, A test for bivariate two-sample location problem, *Commun. Stat. Theory Methods*, **29** (2000), 417–436. https://doi.org/10.1080/03610920008832492
13. H. Samawi, M. F. Al-Saleh, O. Al-Saidy, The matched pair sign test using bivariate ranked set sampling for different ranking-based schemes, *Stat. Methodol.*, **6** (2009), 397–407. https://doi.org/10.1016/j.stamet.2009.02.002
14. S. K. Mathur, A new nonparametric bivariate test for two-sample location problem, *Stat. Methods Appl.*, **18** (2009), 375–388. https://doi.org/10.1007/s10260-008-0095-7
15. F. Roland, D. Herold, Robust nonparametric tests for the two-sample location problem, *Stat. Methods Appl.*, **20** (2011), 409–422. https://doi.org/10.1007/s10260-011-0164-1
16. H. E. Daniels, Saddlepoint approximation in statistics, *Ann. Math. Stat.*, **25** (1954), 631–650.
17. H. E. Daniels, Tail probability approximations, *Int. Stat. Rev.*, **55** (1987), 37–48. https://doi.org/10.1515/9783112319086-033
18. I. M. Skovgaard, Saddlepoint expansions for conditional distributions, *J. Appl. Probab.*, **24** (1987), 875–887. https://doi.org/10.1017/s0021900200116754
19. J. E. Kolassa, *Series Approximation Methods in Statistics*, New York: Springer, 2006. https://doi.org/10.1007/0-387-32227-2
20. R. W. Butler, *Saddlepoint Approximations with Applications*, Cambridge: Cambridge Univ. Press, 2007.

21. A. M. Abd El-Raheem, M. Hosny, E. F. Abd-Elfattah, Statistical inference of the class of nonparametric tests for the panel count and current status data from the perspective of the saddlepoint approximation, *J. Math.*, **2023** (2022), 9111653. https://doi.org/10.1155/2023/9111653
22. A. M. Abd El-Raheem, M. Hosny, Saddlepoint p-values for a class of nonparametric tests for the current status and panel count data under generalized permuted block design, *AIMS Math.*, **8** (2023), 18866–18880. https://doi.org/10.3934/math.2023960
23. I. A. Shanan, E. F. Abd-Elfattah, A. M. Abd El-Raheem, A new approach for approximating the p-value of a class of bivariate sign tests, *Sci. Rep.*, **13** (2023), 19133. https://doi.org/10.1038/s41598-023-45975-7
24. A. M. Abd El-Raheem, I. A. Shanan, M. Hosny, Saddlepoint approximation of the p-values for the multivariate one-sample sign and signed-rank tests, *AIMS Math.*, **9** (2024), 25482–25493. https://doi.org/10.3934/math.20241244
25. A. M. Abd El-Raheem, H. N. Mohamed, E. F. Abd-Elfattah, Saddlepoint p-values for a class of location-scale tests, *J. Biopharm. Stat.*, **35** (2025), 641–660. https://doi.org/10.1038/s41598-024-53451-z
26. A. D. Alhejaili, A. A. AlGhamedi, A review of saddle-point approximation: Theory and applications, *Adv. Appl. Stat.*, **92** (2025), 343–390. https://doi.org/10.17654/0972361725015
27. A. M. Abd El-Raheem, M. Hosny, On the distribution of the random sum and linear combination of independent exponentiated exponential random variables, *Symmetry*, **17** (2025), 200. https://doi.org/10.3390/sym17020200
28. W. F. Rosenberger, J. M. Lachin, *Randomization in Clinical Trials: Theory and Practice*, Hoboken: John Wiley & Sons, 2015.
29. E. F. Abd-Elfattah, Saddlepoint p-values for two-sample bivariate tests, *J. Stat. Plann. Inference*, **171** (2016), 92–98. https://doi.org/10.1016/j.jspi.2015.11.002
30. D. J. Hand, F. Daly, K. McConway, D. Lunn, E. Ostrowski, *A Handbook of Small Data Sets*, Boca Raton: CRC Press, 1993.
31. H. M. Samawi, M. Pararai, An optimal bivariate ranked set sample design for the matched pairs sign test, *J. Stat. Theory Pract.*, **3** (2009), 393–406. https://doi.org/10.1080/15598608.2009.10411932

AIMS Press