Mathematics

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

AIMS Mathematics, 10(2): 2602–2618.
DOI: 10.3934/math.2025121
Received: 24 September 2024
Revised: 07 January 2025
Accepted: 21 January 2025
Published: 13 February 2025

Research article

Saddlepoint approximation for the p-values of some distribution-free tests

Abd El-Raheem M. Abd El-Raheem[1,*] **and Mona Hosny**[2]

[1] Department of Mathematics, Faculty of Education, Ain Shams University, Cairo 11566, Egypt
[2] Department of Mathematics, College of Science, King Khalid University, Abha, 61413, Saudi Arabia

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

Abstract: This article discusses the saddlepoint approximation for the p-values of some distribution-free tests, a signed rank test for bivariate location problems and a dispersion test for scale problems. The statistics of the two considered tests are constructed based on the ratio of two variables. The accuracy of the saddlepoint approximation is compared to traditional asymptotic normal approximation by applying numerical comparisons. Furthermore, the proposed approximations are illustrated by analyzing numerical examples. The results of numerical comparisons indicate that the approximation error resulting from the proposed method is much lower than the traditional method, which is evidence of the superiority of the proposed approximation method over the traditional method. Accordingly, we can say that the saddlepoint approximation method can be a competitive alternative to the traditional method.

Keywords: bivariate signed rank tests; location problems; saddlepoint approximations; scale problems
Mathematics Subject Classification: 62E17, 62G10

1. Introduction

Asymptotic approximations of the saddlepoint type for the p-value of the test statistics of two distribution-free tests are considered. The first test addresses the scale problem and it has been introduced by Mathur and Dolo [1] as a good alternative to some famous scale tests such as the Siegel-Tukey test [2], Levene test [3], Klotz test [4] and Fligner and Killeen test [5]. The second test addresses the location problem and is a bivariate signed-ranked test [6]. This test is characterized by its statistic being independent of the correlation between two variables, making it easy to note the marginal effect of a single variable on the test statistic. Moreover, it is a scale-invariant test, ensuring that the value of the statistic does not change when the scale of the observations has been changed. Furthermore, it is robust to outliers and more robust than its counterparts under the non-normal distribution, even under very small changes in location. Mathur and Sepehrifar [6] have proven that their test is competitive with several analogues, such as the Mardia test [7], Wilcoxon one-sample bivariate rank sum test, [8] and the Peters and Randles test [9]. It should be noted that the statistics of the two tests considered here, whether the scale or the location test, depend on the ratio of the two variables. This technique was also used by many statisticians in the formation of their test statistics, such as Blumen [10], and Sen and Mathur [11].

The saddlepoint approximation is fundamentally a method for approximating a probability density or mass function using its corresponding moment-generating function or cumulant generating function [12]. It is a frequently used statistical approximation method in approximating many statistical and probabilistic functions. The theoretical and applied statistics are full of approximation methods to solve many problems that do not have an exact solution. The accuracy of the approximation method that each method provides is what distinguishes one method from another. In the case of approximating the statistical functions, such as the distribution, mass, and density functions, we can refer to the asymptotic normal approximation method, which depends on the central limit theorem, and to the saddlepoint approximation method, which is considered a generalization to Laplace's method for approximating integrals. The saddlepoint approximation method offers several significant benefits in statistical inference, particularly in hypothesis testing and p-value approximation. One of the primary advantages is their high accuracy, especially for small sample sizes, where traditional asymptotic methods often fall short. Unlike standard approximations, the saddlepoint method produces precise tail probabilities, which are critical in accurately assessing the significance of test statistics. Furthermore, this approximation is highly versatile and can be applied to a wide range of complex distributions, including nonparametric tests and ratio-based statistics. The computational efficiency of the saddlepoint method also makes it a practical alternative to the simulation methods, reducing the need for extensive simulations. Overall, the saddlepoint approximation method is a powerful and flexible tool that can outperform traditional methods, particularly in challenging scenarios where exact solutions are impractical or unavailable. In this article, a comparison is made between the accuracy of the two methods for approximating the p-value of the two statistics of the considered non-parametric tests, clarifying that the saddlepoint approximation method is more accurate than the normal approximation method. The origin of the saddlepoint method in statistics can be traced back to 1954 by the work presented by Daniels [13] as an approximation of the probability density function for the mean of n random variables. Daniels's work served as the starting point for many scholars and statisticians to present many approximations of statistical and probability functions, such as the approximation of the distribution function, the conditional distribution function, and the bivariate distribution function by Lugannani and Rice [14], Skovgaard [15], and Wang [16], respectively. Subsequently, contributions were made to this topic, and its applications have spread to many branches of statistics. In this regard, we can refer to several references, for example, [12, 17–20]. Gatto and Jammalamadaka [21] introduced a saddlepoint approximation for the distribution function of a M statistic, conditioned on another M statistic. Abd-Elfattah and Butler [22, 23] derived the permutation distribution of the weighted log-rank class of tests using the saddlepoint approximation method. Abd-Elfattah [24, 25] proposed the saddlepoint approximation method to approximate p-values for weighted log-rank class of tests considering truncated binomial and randomized block designs, respectively. Abd El-Raheem and Abd-Elfattah [26, 27] extended the results of [22, 25] for the clustered censored data. Abd El-Raheem et al. [28] approximated the tail probabilities for multivariate sign and signed-rank tests using the saddlepoint approximation method. Readers may refer to recent works on the saddlepoint approximation, such as [29–32].

This article aims to enhance the accuracy of p-value calculations in small sample sizes, where traditional methods, such as the normal approximation, can fail to provide precise results. The accuracy of the proposed approach (saddlepoint approximation) is assessed by comparing the approximated p-values to exact p-values obtained by the simulation method (permutation-based, so time-consuming). The relative absolute error is used to evaluate the precision and reliability of the approximation across various scenarios.

The article is organized as follows: In Sections 2 and 3, we provide the saddlepoint results and show the other asymptotic approximations for the considered tests. Finally, Sections 4 and 5 are devoted to the numerical comparisons between the saddlepoint and normal approximation methods using real and simulated data, respectively.

2. Dispersion test

Let y_i, $i = 1, 2, ..., n$ and x_j, $j = 1, 2, ..., m$ be independent random samples from continuous populations with distribution functions $F_Y(y)$ and $F_X(\sigma y)$. To test the hypothesis $H_0 : F_Y(y) = F_X(\sigma y)$, for all y and $\sigma = 1$ against $H_1 : F_Y(y) \neq F_X(\sigma y)$, for all y and $\sigma \neq 1$, $\sigma > 0$. Mathur and Dolo [1] introduced the dispersion test statistic as follows:

$$W = \sum_{k=1}^{nm} \phi_k R(r_k), \tag{2.1}$$

where $r_k = \frac{x_j - M}{y_i - M}$, M is the median of the combined samples, $R(r_k)$ is the rank of r_k, and

$$\phi_k = \begin{cases} 1, & r_k \geq 0, \\ 0, & r_k < 0. \end{cases}$$

The expectation and variance of the test statistic in (2.1) under H_0 are $E(W|H_0) = mn(mn + 1)/4$, and $V(W|H_0) = mn(mn + 1)(2mn + 1)/24$. When sample sizes are large

$$Z = \frac{W - E(W|H_0)}{\sqrt{V(W|H_0)}} \sim N(0, 1).$$

Saddlepoint approximation

In this subsection, the saddlepoint approximation method is applied to approximate the p-value of the statistic in (2.1).

Depending on the permutation distribution of the statistic W in (2.1), then the moment generating function of W is given by:

$$M_W(s) = \prod_{k=1}^{nm} \left(\frac{1}{2} + \frac{1}{2} \exp\{sR(r_k)\} \right),$$

then, the cumulant generating function, CGF, of the statistic W is

$$K_W(s) = \sum_{k=1}^{nm} \log\left(\frac{1}{2} + \frac{1}{2}\exp\{sR(r_k)\}\right). \tag{2.2}$$

The first, second, and third derivatives of the CGF, K_W, in (2.2) are

$$K_W^{'}(s) = \sum_{k=1}^{nm} \frac{R(r_k)\exp\{sR(r_k)\}}{1+\exp\{sR(r_k)\}},$$

$$K_W^{''}(s) = \sum_{k=1}^{nm} \frac{R(r_k)^2\exp\{sR(r_k)\}}{[1+\exp\{sR(r_k)\}]^2},$$

and

$$K_W^{'''}(s) = \sum_{k=1}^{nm} \frac{R(r_k)^3\exp\{sR(r_k)\}\,[1+\exp\{sR(r_k)\}] - 2R(r_k)^3\exp\{2sR(r_k)\}}{[1+\exp\{sR(r_k)\}]^3}.$$

The saddlepoint approximation of the cumulative distribution function of the statistic W, $F_W(w)$ is given by [14]:

$$\hat{F}_W(w) = \begin{cases} \Phi(\tilde{\rho_1}) + \phi(\tilde{\rho_1})\left(\frac{1}{\tilde{\rho_1}} - \frac{1}{\tilde{\gamma_1}}\right) & \text{if } w \neq \mu_W, \\ \frac{1}{2} + \frac{K_W'''(0)}{6\sqrt{2\pi}(K_W''(0))^{3/2}} & \text{if } w = \mu_W, \end{cases}$$

where

$$\tilde{\rho_1} = \operatorname{sgn}(\tilde{s})\sqrt{2[\tilde{s}w - K_W(\tilde{s})]}, \quad \text{and} \quad \tilde{\gamma_1} = \tilde{s}\sqrt{K_W''(\tilde{s})}.$$

Here, Φ and ϕ represent the standard normal distribution function and density function, respectively, and $\mu_W = E(W|H_0)$ is the mean of the distribution. The symbol $\operatorname{sgn}(\tilde{s})$ denotes the sign of $\tilde{s}$. The saddlepoint $\tilde{s}$ is the unique solution to the equation $K_W'(\tilde{s}) = w$.

The saddlepoint approximation of the exact p-value for the statistic W is given by:

$$\hat{P}(W \geq w_0) \simeq 1 - \Phi(\tilde{\rho_1}) - \phi(\tilde{\rho_1})\left(\frac{1}{\tilde{\rho_1}} - \frac{1}{\tilde{\gamma_1}}\right),$$

where w_0 is the observed value of the statistic W.

3. Signed rank test for bivariate location problem

Let a random sample (x_i, y_i) for $i = 1, 2, \ldots, n$, consisting of n independent pairs taken from a bivariate population with a continuous distribution function $F_{X,Y}(x + \mu_1, y + \mu_2)$. Assume the population is elliptically symmetric around its median (μ_1, μ_2). We aim to test the null hypothesis $H_0 : (\mu_1, \mu_2) = (0, 0)$ against the alternative hypothesis $H_1 : (\mu_1 \neq 0, \mu_2 \neq 0)$. For this purpose, assume $\vartheta_i = tan^{-1}(d_i)$, where $d_i = y_i/x_i$ for $i = 1, 2, \ldots, n$ are the tangents of the projected angles corresponding to $S_1, S_2, \ldots, S_n$, respectively, where $S_i = (X_i, Y_i)'$ for $i = 1, 2, \ldots, n$. Sign tests are fundamental in non-parametric methods, and numerous researchers have worked on developing non-parametric tests based on this concept. When the underlying population exhibits elliptical symmetry, it seems intuitive and logical to assess the distance of each observation from the origin. Mathur and Sepehrifar [6] used the ranks of these distances along with the directions of the observations to construct a signed rank statistic for the bivariate location problem as follows:

$$U_1 = \frac{1}{n}\sum_{i=1}^{n} \delta_i R(d_i) \cos\left(\frac{i\pi}{n}\right), \tag{3.1}$$

and

$$U_2 = \frac{1}{n}\sum_{i=1}^{n} \delta_i R(d_i) \sin\left(\frac{i\pi}{n}\right), \tag{3.2}$$

where

$$\delta_i = \begin{cases} 1, & \text{if } Y_i \text{ is negative in the } (i+1) \text{ ordered slope,} \\ -1, & \text{if } Y_i \text{ is positive in the } (i+1) \text{ ordered slope.} \end{cases}$$

Under the the null hypothesis H_0, $P(\delta_i = -1) = P(\delta_i = 1) = 1/2$, the statistic U_1 has normal distribution with $E(U_1|H_0) = 0$ and $V(U_1|H_0) = \frac{1}{3}\sum_{i=1}^{n} \cos^2\left(\frac{i\pi}{n}\right)$, and the U_2 has normal distribution with $E(U_2|H_0) = 0$ and $V(U_2|H_0) = \left(n - \sum_{i=1}^{n} \cos^2\left(\frac{i\pi}{n}\right)\right)/3$. Thus, the bivariate signed rank statistic is given by:

$$U = U_1 + U_2 = \sum_{i=1}^{n} \delta_i C_i, \tag{3.3}$$

where $C_i = \frac{1}{n} R(d_i)\left(\cos\left(\frac{i\pi}{n}\right) + \sin\left(\frac{i\pi}{n}\right)\right)$. The statistic U in (3.3) is normally distributed with mean $E(U|H_0) = 0$ and variance $V(U|H_0) = 1/3$. Let $\eta_i = \frac{\delta_i + 1}{2}$, then $\eta_i = 0$ or $\eta_i = 1$. Thus, the statistic in (3.3) becomes

$$U = \sum_{i=1}^{n} 2\eta_i C_i - \sum_{i=1}^{n} C_i. \tag{3.4}$$

Saddlepoint approximation

This subsection applies the saddlepoint approximation method to approximate the p-value of the statistic in (3.4).

The moment generating function of the statistic U in (3.4) is determined by its permutation distribution and is given by:

$$M_U(s) = e^{-s\sum_{i=1}^{n} C_i} \prod_{i=1}^{n} \left(\frac{1}{2} + \frac{1}{2} e^{2sC_i}\right),$$

then, the CGF of the statistic U is

$$K_U(s) = -s\sum_{i=1}^{n} C_i + \sum_{i=1}^{n} \log\left(\frac{1}{2} + \frac{1}{2} e^{2sC_i}\right). \tag{3.5}$$

The first, second, and third derivatives of the CGF, K_U, in (3.5) are

$$K_U'(s) = -\sum_{i=1}^{n} C_i + \sum_{i=1}^{n} \frac{C_i e^{2sC_i}}{\frac{1}{2} + \frac{1}{2}e^{2sC_i}},$$

$$K_U''(s) = \sum_{i=1}^{n} \frac{C_i^2 e^{2sC_i}}{\left(\frac{1}{2} + \frac{1}{2}e^{2sC_i}\right)^2},$$

$$K_U'''(s) = \sum_{i=1}^{n} \frac{C_i^3 e^{2sC_i}\left[1 - e^{2sC_i}\right]}{\left(\frac{1}{2} + \frac{1}{2}e^{2sC_i}\right)^3}.$$

The saddlepoint approximation of the cumulative distribution function of the statistic U, $F_U(u)$, is given by [14]:

$$\hat{F}_U(u) = \begin{cases} \Phi(\tilde{\rho}_2) + \phi(\tilde{\rho}_2)\left(\frac{1}{\tilde{\rho}_2} - \frac{1}{\tilde{\gamma}_2}\right) & \text{if } u \neq \mu_U, \\ \frac{1}{2} + \frac{K_U'''(0)}{6\sqrt{2\pi}(K_U''(0))^{3/2}} & \text{if } u = \mu_U, \end{cases}$$

where

$$\mu_U = E(U|H_0), \quad \tilde{\rho}_2 = \text{sgn}(\tilde{s})\sqrt{2[\tilde{s}u - K_U(\tilde{s})]}, \quad \text{and} \quad \tilde{\gamma}_2 = \tilde{s}\sqrt{K_U''(\tilde{s})}.$$

The saddlepoint $\tilde{s}$ is the unique solution to the equation $K_U'(\tilde{s}) = u$.

The saddlepoint approximation of the exact p-value for the statistic U is given by:

$$\hat{P}(U \geq u_0) \simeq 1 - \Phi(\tilde{\rho}_2) - \phi(\tilde{\rho}_2)\left(\frac{1}{\tilde{\rho}_2} - \frac{1}{\tilde{\gamma}_2}\right),$$

where u_0 is the observed value of the statistic U.

4. Numerical comparisons

Analyzing real data is essential for validating the saddlepoint approximation and comparing it to the normal approximation. By applying the saddlepoint approximation to real data, we can assess its accuracy in approximating the exact p-value and its effectiveness across various scenarios, thereby confirming its robustness and reliability.

4.1. Numerical comparisons for the dispersion test

Example 1: Prehistoric Native Americans used pipes for ceremonial purposes, which were typically made of carved stone or clay ceramics. Clay pipes were easier to produce, while stone pipes required careful drilling using a hollow bone and special stone drills. According to one anthropologist, the easier manufacturing process of clay pipes resulted in greater variation in their construction. The diameters of ceramic pipe bowls and stone pipes (cm) from the Wind Mountain archaeological area were measured to evaluate this claim. These data are presented in Table 1, and the source of these data is the reference [33]. The dispersion test is used to evaluate the null hypothesis, which states no difference in variance, against the alternative hypothesis supporting the anthropologist's claim that clay pipes exhibit greater variance. The approximated p-value using the simulation (permutation-based method), saddlepoint, and normal approximation methods are calculated and listed in Table 2.

Table 1. Pipe bowl diameters for ceramic and stone pipes.

Ceramic pipe bowl diameters (cm)	Stone pipe bowl diameters (cm)
1.7	1.6
5.1	2.1
1.4	3.1
0.7	1.4
2.5	2.2
4.0	2.1
3.8	2.6
2.0	3.2
3.1	3.4
5.0	
1.5	

Table 2. The approximated p-values using the simulation, saddlepoint, and normal approximation methods for the dispersion test.

Example	Simulation	Saddlepoint	Normal
Example 1	0.020465	0.020576	0.020758
Example 2	0.284281	0.283228	0.282564

Example 2: A key indicator of a company's productivity is the relative annual return to its total assets. This metric shows the return generated from assets over a year, thoroughly assessing the company's financial efficiency and profitability. It is useful for comparing competing firms. Table 3 displays the percentage returns based on assets for a random sample of prominent companies from France and Germany. The source of these data is the reference [34]. The dispersion test is used to evaluate the claim that there is a difference in the population variance of percentage yields between leading companies in France and Germany. The approximated p-value using the simulation, saddlepoint, and normal approximation methods are obtained and listed in Table 2.

Table 3. The percentage returns based on assets for a random sample of prominent companies from France and Germany.

Sample from France	Sample from Germany
2.5	2.3
2.0	3.2
4.5	3.6
1.8	1.2
0.5	3.6
3.6	2.8
2.4	2.3
0.2	3.5
1.7	2.8
1.8	
1.4	
5.4	
1.1	

Table 2 shows that the saddlepoint approximation technique offers higher accuracy and reliability than the traditional method, as it is significantly closer to that obtained using the simulation method.

4.2. Numerical comparisons for the bivariate signed rank test

Example 1: In 2010, the average minimum temperatures recorded for counties Roscommon and Meath in Ireland showed notable variations throughout the year. These temperatures, measured in degrees Celsius (°C), reflect the coldest daily temperatures, averaged over the year. Table 4 shows the average minimum temperatures (°C) recorded for counties Roscommon and Meath, Ireland 2010; see [35] for more details. We aim to test the null hypothesis, $H_0 : (\mu_1, \mu_2) = (0, 0)$ against the alternative hypothesis $H_1 : (\mu_1 \neq 0, \mu_2 \neq 0)$.

Table 4. The average minimum temperatures (°C) recorded for counties Roscommon and Meath, Ireland 2010.

County	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Roscommon	-1.6	-1.4	0.6	3	4.3	9	11.7	8.3	8.8	4.5	0.8	-5.4
Meath	-2	-1.4	0.6	3.4	5	9.7	11.9	9.2	9.1	5.9	1.5	-4.6

Example 2: The data were collected to study the development of the immune system in 36 HIV-positive newborns. These children were given Ritonavir therapy, and their CD45RA and CD45RO T cell counts were measured at birth and again after 24 weeks of treatment. Table 5 shows the difference between CD45RA T cells and CD45RO T cells at 24 weeks of treatment and at birth. The source of these data is the reference [36]. The bivariate signed rank test was conducted to determine whether these cell counts were significantly changed over the treatment period. The p-value of this test is calculated using simulation, saddlepoint, and normal approximation methods and presented in Table 6.

Table 5. The difference between CD45RA T cells and CD45RO T cells at 24 weeks of treatment and at birth.

CD45RA T	242	569	270	-25	309	22	-42	-233	206	-106	55	85
	30	194	-87	159	29	89	-9	158	76	15	3	93
	160	66	180	237	105	16	167	-10	-16	-7	15	160
CD45RO T	1708	569	757	499	231	338	26	119	163	-186	54	48
	50	525	-110	148	102	364	36	234	122	24	36	71
	44	128	155	85	76	6	364	-18	-21	-2	32	188

Table 6. The approximated p-values using the simulation, saddlepoint, and normal approximation methods for the bivariate signed rank test.

Example	Simulation	Saddlepoint	Normal
Example 1	0.260260	0.260738	0.255757
Example 2	0.034412	0.034437	0.031085

Table 6 demonstrates that the saddlepoint approximation technique provides greater accuracy and reliability than the traditional method, as its results are much closer to those obtained using the simulation method.

5. Simulation study

When analytical solutions are intractable or impractical, statistical approximation methods, such as asymptotic normal approximation or saddlepoint approximation methods, are often employed to make inferences about complex models or datasets. Understanding the relative performance of these methods is crucial for selecting the most appropriate technique for a given context. This section uses simulation studies to evaluate and compare the performance of different approximation methods under various conditions.

5.1. Simulation study for the dispersion test

For the dispersion test, a simulation study is performed to assess the consistency of the saddlepoint p-value approximation across various sample sizes, distributions, and location parameter values. The simulations involved two distributions: logistic and extreme value. For each distribution, 1,000 data sets are generated with total sample sizes of $N = 16, 24, 32$, and 42, where $m = n = N/2$. The simulated mid-p-value for each of the 1,000 data sets is calculated using 10^6 randomized sequences for the indicators ϕ_k. Let $\sigma = \sigma_x/\sigma_y$ represent the dispersion parameter, where σ_x and σ_y are the scale parameters for the populations X and Y, respectively. Let M_x and M_y be the medians of the two populations X and Y, respectively. The data sets are generated with $M_y = 0$, $\sigma_y = 1$, $M_x = c\sigma_y$, where $c = 0, 1$, and 2, and σ_x is selected to ensure that the mean of the simulated mid-p-values for the 1,000 data sets is approximately 0.05. To compare the saddlepoint and normal approximation methods, we calculate the following quantities: saddlepoint approximation proportion (Sap.Prop.) refers to the proportion of the saddlepoint method to the simulation method, relative absolute error of saddlepoint (Rel.Abs.Err.Sap.) indicated the accuracy of the saddlepoint method compared to the simulation method, and relative absolute error of normal (Rel.Abs.Err.Nor.) represented the accuracy of the normal method compared to the simulation method. The mathematical definition of the quantities Sap.Prop., Rel.Abs.Err.Sap., and Rel.Abs.Err.Nor. are given by:

$$Sap.Prop. = 100 * \frac{\sum_{i=1}^{M} I\left(|P_i(Sap) - P_i(Sim)| < |P_i(Nor) - P_i(Sim)|\right)}{M},$$

where $I()$ denotes the indicator function, $P_i(Sap)$ represents the saddlepoint p-value, $P_i(Nor)$ represents the normal approximation p-value, and $P_i(Sim)$ represents the simulated p-value.

$$Rel.Abs.Err.Sap. = \frac{1}{M}\sum_{i=1}^{M}\frac{|P_i(Sap) - P_i(Sim)|}{P_i(Sim)},$$

and

$$Rel.Abs.Err.Nor. = \frac{1}{M}\sum_{i=1}^{M}\frac{|P_i(Nor) - P_i(Sim)|}{P_i(Sim)}.$$

Results of the simulation study for comparing saddlepoint and normal approximation techniques are presented in Table 7. It also shows that the average of the Sap.Prop. is approximately 87.56% (this value is the average of all fourth-column values in Table 7). This high percentage indicates that the saddlepoint approximation method was more accurate in 87.56% of the considered cases. Furthermore, The average of the Rel.Abs.Err.Sap. is approximately 0.049 (this value is the average of all fifth-column values in Table 7), and the corresponding value for the normal approximation is approximately 0.1193 (this value is the average of all sixth-column values in Table 7). The large difference between the relative absolute error of the two methods also shows that the saddlepoint method is more accurate than the normal approximation method. We can illustrate the superiority of the saddlepoint approximation method over the normal approximation method by plotting the relative absolute error of the two methods. Figures 1 and 2 display the relative absolute error of the saddlepoint and normal approximation methods for logistic and extreme value distributions when $N = 16$ and $c = 0$. It is clearly evident that the relative absolute error resulting from the saddlepoint method is much less than that resulting from the normal approximation method.

Table 7. Comparison of accuracy and efficiency of saddlepoint and normal approximation methods.

Distribution	N	c	Sap.Prop.	Rel.Abs.Err.Sap.	Rel.Abs.Err.Nor.
Logistic	16	0	92.6	0.0185	0.1186
		1	92.2	0.0193	0.1238
		2	90.6	0.0190	0.1208
	24	0	85.7	0.1217	0.2555
		1	84.3	0.0664	0.1422
		2	91.1	0.1153	0.2616
	32	0	82.8	0.1223	0.2189
		1	83.2	0.0978	0.2033
		2	92.3	0.0611	0.1601
	42	0	80.1	0.1923	0.2640
		1	78.9	0.0568	0.0594
		2	90.0	0.0110	0.0175
Extreme value	16	0	89.5	0.0175	0.1081
		1	90.1	0.0164	0.1010
		2	90.2	0.0159	0.0996
	24	0	89.6	0.0659	0.1898
		1	93.0	0.03703	0.1364
		2	93.1	0.0197	0.0918
	32	0	84.3	0.0341	0.0752
		1	87.5	0.0127	0.0319
		2	89.8	0.0006	0.0022
	42	0	80.2	0.0001	0.0023
		1	82.0	0.0104	0.0165
		2	88.4	0.0446	0.0628

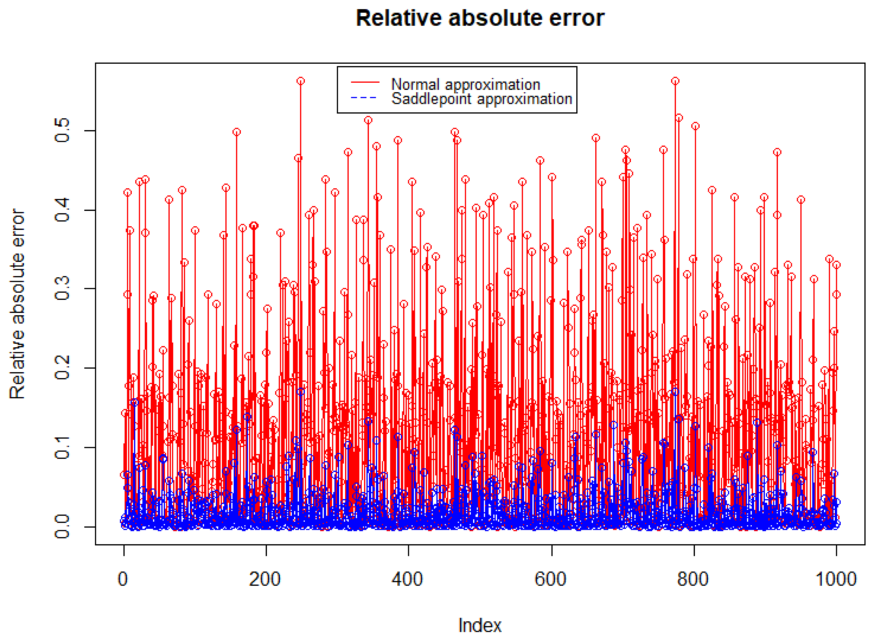

Figure 1. The relative absolute error of the saddlepoint and normal approximation methods for approximating the p-value of the dispersion test for data generated from the logistic distribution with $N = 16$ and $c = 0$.

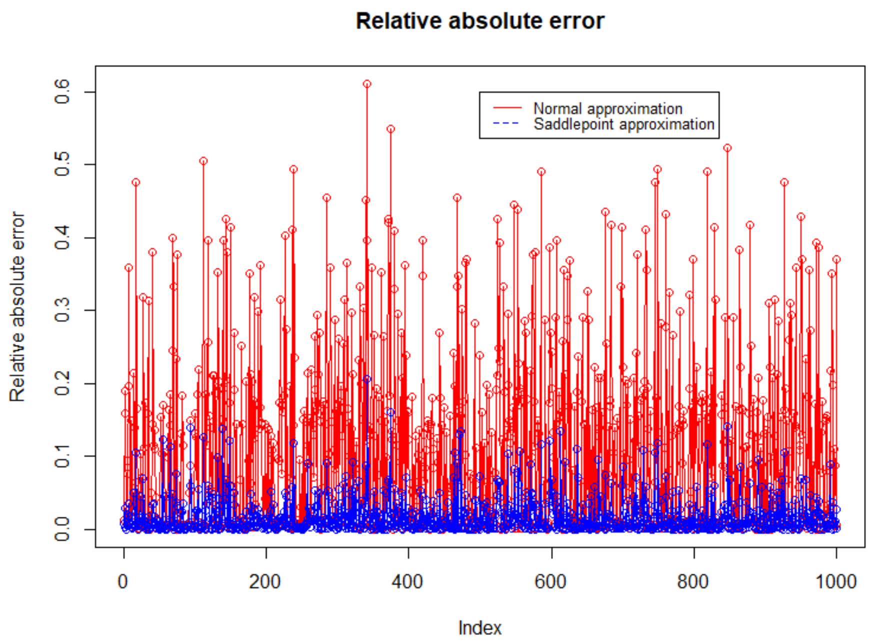

Figure 2. The relative absolute error of the saddlepoint and normal approximation methods for approximating the p-value of the dispersion test for data generated from the extreme value distribution with $N = 16$ and $c = 0$.

5.2. Simulation study for the bivariate signed rank test

Bivariate data are generated from the bivariate normal, logistic, and extreme value distributions to compare the different approximation methods used to approximate the exact p-value of the bivariate signed rank test statistic. 1,000 samples are generated from each distribution. The p-value is calculated using the three approximation methods for each sample of data, and then we calculate the average of the thousand p-values of each approximation method. Results of the simulation study for comparing saddlepoint and normal approximation techniques are displayed in Table 8.

It also shows that the average of the Sap.Prop. is approximately 97%. This high percentage indicates that the saddlepoint approximation method was more accurate in 97% of the considered cases. Furthermore, the average of the Rel.Abs.Err.Sap. is approximately 0.0258, and the corresponding value for the normal approximation is approximately 0.6145. The large difference between the relative absolute error of the two methods also shows that the saddlepoint method is more accurate than the normal approximation method. Figures 3 and 4 illustrate the relative absolute error of approximating the p-value of the bivariate signed rank test using the normal approximation (shown in red) and the saddlepoint approximation (shown in blue). From Figures 3 and 4, it is evident that the saddlepoint approximation consistently achieves lower error rates across all sample indices compared to the normal approximation. The red lines corresponding to the normal approximation show frequent and large spikes in error, indicating that the normal approximation tends to produce larger deviations from the true p-values. In contrast, the blue lines representing the saddlepoint approximation remain much closer to zero, with minimal variation, highlighting its superior accuracy.

Table 8. Comparison of accuracy and efficiency of saddlepoint and normal approximation methods.

Distribution	Sample size	Sap.Prop.	Rel.Abs.Err.Sap.	Rel.Abs.Err.Nor.
Normal	16	97.4	0.0164	0.5477
	24	98.0	0.0206	1.5034
	32	97.5	0.0469	1.3640
	42	94.8	0.1118	1.9154
Logistic	16	96.9	0.0062	0.1441
	24	98.1	0.0059	0.2399
	32	98.0	0.0129	0.3834
	42	96.3	0.0577	0.9086
Extreme value	16	96.5	0.0081	0.2042
	24	96.9	0.0060	0.2334
	32	97.5	0.0059	0.1998
	42	97.7	0.0115	0.2576

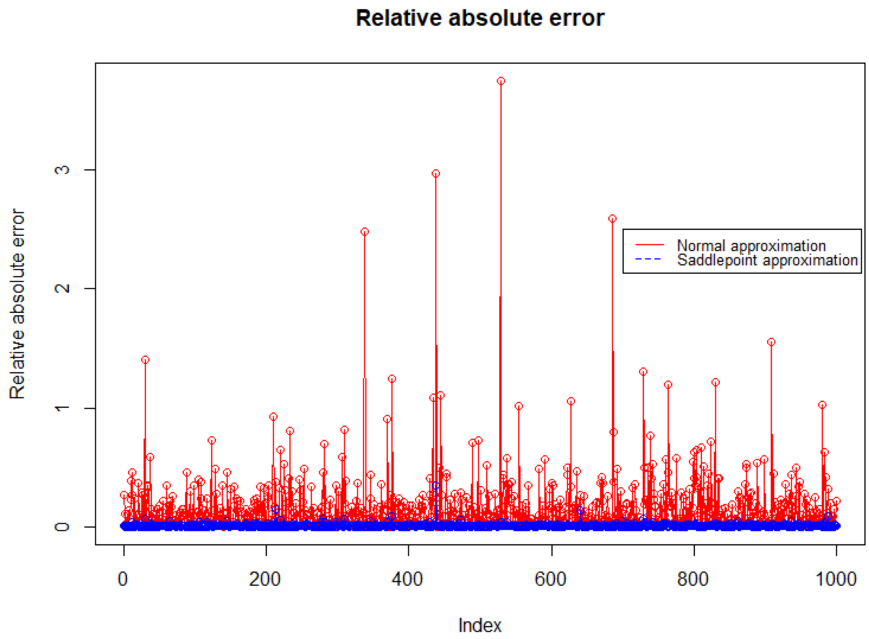

Figure 3. The relative absolute error of the saddlepoint and normal approximation methods for approximating the p-value of the bivariate signed rank test for data generated from the logistic distribution with sample size $n = 32$.

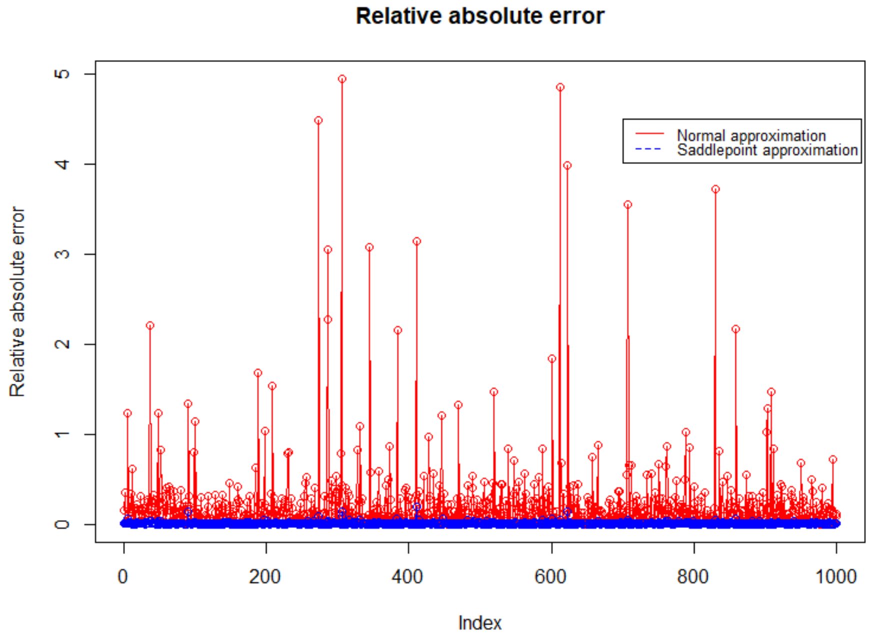

Figure 4. The relative absolute error of the saddlepoint and normal approximation methods for approximating the p-value of the bivariate signed rank test for data generated from the extreme value distribution with sample size $n = 32$.

6. Conclusions

The article highlights the effectiveness of using the saddlepoint approximation for calculating p-values in distribution-free tests, particularly for the signed rank test and a nonparametric scale test. The study compares the saddlepoint approximation method to the traditional asymptotic normal approximation method, showing that the saddlepoint approximation consistently provides lower error rates in p-value approximation. Through numerical comparisons and practical examples, the findings demonstrate that the proposed method offers greater accuracy and can serve as a reliable alternative to traditional approaches in nonparametric statistics. This suggests that the saddlepoint approximation has practical advantages in improving the precision of statistical tests.

Author contributions

A. M. Abd El-Raheem: Conceptualization, Methodology, Investigation, Software, Writing – review & editing, Visualization, Resources, Software, Writing – original draft; M. Hosny: review & editing, Funding acquisition, Project administration. All the authors have agreed and given their consent for the publication of this research paper.

Use of Generative-AI tools declaration

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

Acknowledgments

The second author thanks the Deanship of Scientific Research and Graduate Studies at King Khalid University for funding this work through a Large Research Project under grant number RGP2/398/45.

Conflict of interest

The authors declare no conflicts of interest.

References

1. S. Mathur, S. Dolo, A nonparametric test for scale in univariate population two-sample setup, *Model Assisted Stat. Appl.*, **2** (2007), 145–152.
2. S. Siegel, J. W. Tukey, A nonparametric sum of ranks procedure for relative spread in unpaired samples, *J. Am. Stat. Assoc.*, **55** (1960), 429–445. https://doi.org/10.1080/01621459.1960.10482073
3. H. Levene, Robust tests for equality of variances, In: I. Olkin, *Contributions to probability and statistics*, Stanford University Press, Palo Alto, 1960, 278–292.
4. J. Klotz, Nonparametric tests for scale, *Ann. Math. Statist.*, **33** (1962), 498–512. https://doi.org/10.1214/aoms/1177704576
5. M. A. Fligner, T. J. Killeen, Distribution-free two-sample tests for scale, *J. Am. Stat. Assoc.*, **71** (1976), 210–213. https://doi.org/10.1080/01621459.1976.10481517
6. S. Mathur, M. B. Sepehrifar, A new signed rank test based on slopes of vectors for bivariate location problems, *Stat. Methodol.*, **10** (2013), 72–84. https://doi.org/10.1016/j.stamet.2012.07.001
7. K. V. Mardia, A non-parametric test for the bivariate two-sample location problem, *J. Royal Stat. Soc.: Ser. B (Methodological)*, **29** (1967), 320–342. https://doi.org/10.1111/j.2517-6161.1967.tb00699.x
8. T. P. Hettmansperger, J. W. McKean, Statistical inference based on ranks, *Psychometrika*, **43** (1978), 69–79. https://doi.org/10.1007/BF02294090
9. D. Peters, R. H. Randles, A multivariate signed-rank test for the one-sample location problem, *J. Am. Stat. Assoc.*, **85** (1990), 552–557. https://doi.org/10.1080/01621459.1990.10476234
10. I. Blumen, A new bivariate sign test, *J. Am. Stat. Assoc.*, **53** (1958), 448–456. https://doi.org/10.1080/01621459.1958.10501451
11. 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

12. R. W. Butler, *Saddlepoint approximations with applications*, Cambridge University Press, UK, 2007.
13. H. E. Daniels, Saddlepoint approximations in statistics, *Ann. Math. Statist.*, **25** (1954), 631–650.
14. R. Lugannani, S. Rice, Saddlepoint approximation for the distribution of the sum of independent random variables, *Adv. Appl. Probabil.*, **12** (1980), 475–490. https://doi.org/10.2307/1426607
15. I. M. Skovgaard, Saddlepoint expansions for conditional distributions, *J. Appl. Probabil.*, **24** (1987), 875–887. https://doi.org/10.2307/3214212
16. S. Wang, Saddlepoint approximations for bivariate distributions, *J. Appl. Probabil.*, **27** (1990), 586–597. https://doi.org/10.2307/3214543
17. S. Broda, M. S. Paolella, Saddlepoint approximations for the doubly noncentral t distribution, *Comput. Stat. Data Anal.*, **51** (2007), 2907–2918. https://doi.org/10.1016/j.csda.2006.11.024
18. R. W. Butler, Reliabilities for feedback systems and their saddlepoint approximation, *Statist. Sci.*, **15** (2000), 279–298. https://doi.org/10.1214/ss/1009212818
19. R. W. Butler, D. A. Bronson, Bootstrapping survival times in stochastic systems by using saddlepoint approximations, *J. Royal Stat. Soc.: Ser. B (Stat. Methodol.)*, **64** (2002), 31–49. https://doi.org/10.1111/1467-9868.00323
20. R. W. Butler, S. Huzurbazar, Saddlepoint approximations for the generalized variance and wilks' statistic, *Biometrika*, **79** (1992), 157–169. https://doi.org/10.1093/biomet/79.1.157
21. R. Gatto, S. R. Jammalamadaka, A conditional saddlepoint approximation for testing problems, *J. Am. Stat. Assoc.*, **94** (1999), 533–541. https://doi.org/10.1080/01621459.1999.10474148
22. E. F. Abd-Elfattah, R. W. Butler, The weighted log-rank class of permutation tests: p-values and confidence intervals using saddlepoint methods, *Biometrika*, **94** (2007), 543–551. https://doi.org/10.1093/biomet/asm060
23. E. F. Abd-Elfattah, R. W. Butler, Log-rank permutation tests for trend: saddlepoint p-values and survival rate confidence intervals, *Can. J. Stat.*, **37** (2009), 5–16. https://doi.org/10.1002/cjs.10002
24. E. F. Abd-Elfattah, The weighted log-rank class under truncated binomial design: saddlepoint p-values and confidence intervals, *Lifetime Data Anal.*, **18** (2012), 247–259. https://doi.org/10.1007/s10985-011-9206-0
25. E. F. Abd-Elfattah, Saddlepoint p-values and confidence intervals for the class of linear rank tests for censored data under generalized randomized block design, *Comput. Stat.*, **30** (2015), 593–604. https://doi.org/10.1007/s00180-014-0551-9
26. A. E. M. Abd El-Raheem, E. F. Abd-Elfattah, Weighted log-rank tests for clustered censored data: Saddlepoint p-values and confidence intervals, *Stat. Methods Med. Res.*, **29** (2020), 2629–2636. https://doi.org/10.1177/0962280220908288
27. A. E. M. Abd El-Raheem, E. F. Abd-Elfattah, Log-rank tests for censored clustered data under generalized randomized block design: saddlepoint approximation, *J. Biopharm. Stat.*, **31** (2021), 352–361. https://doi.org/10.1080/10543406.2020.1858310
28. A. E. M. Abd El-Raheem, I. A. 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

29. A. E. M. Abd El-Raheem, M. Hosny, E. F. Abd-Elfattah, Statistical inference of the class of non-parametric tests for the panel count and current status data from the perspective of the saddlepoint approximation, *J. Math.*, **2023** (2023), 1–8. https://doi.org/10.1155/2023/9111653

30. A. E. M. Abd El-Raheem, K. S. Kamal, E. F. Abd-Elfattah, P-values and confidence intervals of linear rank tests for left-truncated data under truncated binomial design, *J. Biopharm. Stat.*, **34** (2024), 127–135. https://doi.org/10.1080/10543406.2023.2171431

31. K. S. Kamal, A. M. Abd El-Raheem, E. F. Abd-Elfattah, Weighted log-rank tests for left-truncated data: saddlepoint p-values and confidence intervals, *Commun. Stat. Theory Methods*, **52** (2023), 4103–4113. https://doi.org/10.1080/03610926.2021.1986534

32. K. S. Kamal, A. E. M. Abd El-Raheem, E. F. Abd-Elfattah, Weighted log-rank tests for left-truncated data under wei's urn design: saddlepoint p-values and confidence intervals, *J. Biopharm. Stat.*, **32** (2022), 641–651. https://doi.org/10.1080/10543406.2021.2010091

33. A. I. Woosley, A. J. McIntyre, *Mimbres mogollon archaeology: Charles C. Di Peso's excavations at Wind Mountain*, Amerind Foundation; University of New Mexico Press, 1996.

34. C. H. Brase, C. P. Brase, *Understandable statistics: concepts and methods*, Houghton Mifflin Company, 2009.

35. Met Éireann, Historical data, n.d. Available from: `https://www.met.ie/climate/available-data/historical-data`.

36. J. W. Sleasman, R. P. Nelson, M. M. Goodenow, D. Wilfret, A. Hutson, M. Baseler, et al., Immunoreconstitution after ritonavir therapy in children with human immunodeficiency virus infection involves multiple lymphocyte lineages, *J. Pediatr.*, **134** (1999), 597–606. https://doi.org/10.1016/S0022-3476(99)70247-7