



Research article

Saddlepoint p-values for a class of nonparametric tests for the current status and panel count data under generalized permuted block design

Abd El-Raheem M. Abd El-Raheem^{1,*} and Mona Hosny²

¹ Department of Mathematics, Faculty of Education, Ain Shams University, Roxy, Cairo 11341, Egypt

² Department of Mathematics, Faculty of Science for Girls, King Khalid University, Abha 61413, Saudi Arabia

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

Abstract: Current status and panel count data appear in many applied fields, including medicine, clinical trials, epidemiology, econometrics, demography, engineering and public health. Therefore, in this article, we use the saddlepoint approximation method to approximate the exact p-value of a number of nonparametric tests for the current status and panel count data under a generalized permuted block design. The saddlepoint approximation is referred to as higher-order approximation and it is more accurate than the methods that lead to approximations that are accurate to the first order, such as the asymptotic normal approximation method. To verify the accuracy and efficiency of the saddlepoint approximation method, a simulation study is conducted. The simulation study results confirm that the saddlepoint approximation method is more powerful than the existing approximation method. Furthermore, number of real current status and panel count data sets are analyzed and displayed as illustrative examples.

Keywords: approximating exact p-value; saddlepoint approximation; current status data; panel count data; permuted block design; two-sample tests

Mathematics Subject Classification: 92B15, 62P10, 62E17, 62G10

1. Introduction

In medical and reliability studies, units are inspected for the occurrence of an event of interest within a predefined period. Some units under study may not have experienced the event during the time specified for the study. All we know about the event time of those units is that it exceeds the time specified previously for the end of the study. This type of data is referred to as right-censored data. Another type of data appears when the time of the event of interest cannot be directly observed,

but instead we know that the event of interest occurred within a specific interval. This type of data is referred to as interval censored data and often appears in studies that cannot be monitored continuously over time, such as epidemiology studies, demographic studies, and infectious disease studies, when infection is an unobserved event. In this regard, we can refer to two important references that offer more information and applications for the interval censored data, namely: Huang and Wellner [1] and Sun [2]. Due to time limits, expenses, and other difficulties, a more extreme form of survival censored data, namely current status data may be preferred. In general, the term current state data refers to two types of data, the first of which is also known as case-1 interval censored data, and we refer to it here as CS-I data. In CS-I data, the only available information is whether the event of interest has occurred before the examination time or not. Carcinogenicity studies, partner studies of HIV and studies of non-fatal human disease are examples of studies that contain CS-I data; see Gart et al. [3], Jewell and Shiboski [4] and Keiding et al. [5]. The second type of current status data appears in studies in which subjects acquire a recurring event over time, such as multiple incidental neoplasms. In these studies, each element of the study is examined only once, and the number of events that occurred before the time of the examination is counted. Therefore, the information available for this type of data is the number of repeated events that occurred before the time of the examination. Throughout this article we refer to this type of data as CS-II data. For more information and applications about this type of data and the studies in which it appears, readers can review the articles of Ii et al. [6], Diamond and McDonald [7] and Dinse [8]. Another type of important data that appears in medical follow-up studies and reliability studies which can be considered as a generalization of CS-II data is the panel count data. This type of data differs from the CS-II data in that each element of the study is examined more than once. Accordingly, the panel count data consists of the discrete examination times for each element of the study and the number of events that occurred between the successive examination times. It is possible to refer to a number of essential references that dealt with the panel count data in more detail, such as: Thall and Lachin [9], Sun and Kalbfleisch [10] and Wellner and Zhang [11]. As a result of the importance of the previous data types, nonparametric tests for CS-I data, CS-II data and panel count data were introduced by Sun and Kalbfleisch [12], Sun [13], Sun and Fang [14] and Balakrishnan and Zhao [15]. Such nonparametric tests are presented in more detail in the second section of this article.

The primary goal of clinical trials is to make an accurate treatment comparison. The design of a clinical trial includes choosing a method to assign treatment for patients. This method of allocating treatment to patients must fulfill several conditions, including concealing the allocation from the patient and researcher, so that their judgment on the experimental treatment is more objective and not biased. Moreover, the assignment of the next patient must be unexpected even if some information is available about previous assignments. The simplest procedure to achieve these conditions is to toss a coin to determine which of the two treatments is assigned to the patient. Assigning patients to treatment by tossing a balanced coin is often referred to as simple randomization, SR, or complete randomized design, CRD. The CRD achieves the highest level of randomization but may result in treatment imbalance which may cause selection bias in the trial outcome. The literature includes a number of other designs that ensure equilibrium between treatments, such as the random allocation rule, RAR, truncated binomial design, TBD and permuted block design, PBD. In RAR, assigning patients to treatment is done by selecting a ball at random from an urn containing $n/2$ balls for each treatment, and the withdrawal is without replacement. This process continues until the urn is empty. TBD uses the same technique of CRD until half of the patients are in one of the treatments, and

then the rest go to the other treatment. PBD is the most commonly used randomization design; see McEntegart [16], Berger [17], and Pond et al. [18]. It uses the same technique of RAR within each block. A block is a group of empirical units that are similar in some measure. The way the block elements are similar is expected to have an effect on the response to treatments. This design is applied by dividing the sample into relatively homogeneous blocks of equal sizes, then RAR is applied within each block. The dispersion within each block is less than its counterpart within the entire sample. Therefore, testing the effect of treatment within the block is more effective compared to the whole sample [19–21]. A general form of the PBD is called generalized PBD which enables different block sizes and imbalanced group sizes within each block. As a result of the importance and uses of PBD, we propose in this article an approximation to the lower tail probability of a class of tests for current status and panel count data using the saddlepoint approximation method under generalized PBD.

The saddlepoint approximation method is one of the statistical approximation methods that has importance and various uses in many branches of statistics. It has been described as one of the approximation methods that has a significant impact on the development of high-precision approximations in the various branches of statistics [22]. The most influential and important contribution to this method began with Daniels [23], who presented an accurate approximation of the probability density function. Since the publication of this pioneering article by Daniels, many successive approximations to a number of statistical functions have been appeared, such as the saddlepoint approximation of univariate cumulative distribution function, CDF, [24], conditional CDF [25] and bivariate CDF [26]. After that, the previous approximations were widely used to solve many statistical problems. A limited number of them can be mentioned here, for example, Daniels [27], Davison and Hinkley [28], Butler [29] and Abd-Elfattah and Butler [30, 31]. For a number of recent articles on this topic, readers can see [32–38].

This article consists of five sections, which can be briefly described as follows: The first section is an introduction to the topic of the article. The second section presents a number of nonparametric tests that manipulate the data of interest in this article. The third section is dedicated to displaying the approximation of the exact p-values for tests presented in the second section and similar tests using the saddlepoint approximation method. The fourth section is devoted to clarifying the theoretical results presented in the previous sections and comparing the accuracy of the proposed method with the asymptotic normal approximation method through a simulation study and the analysis of real data sets. The last section of this article presents a summary of the obtained observations and results.

2. Permutation tests for current status and panel count data under PBD

In this section, four permutation tests for current status and panel count data under PBD are displayed. Under PBD a sample of n individuals are divided into b blocks of sizes n_l such that $n = \sum_{l=1}^b n_l$. Within each block, the RAR is applied to randomize q_l , $l = 1, 2, \dots, b$ of the individuals to the treatment group and $n_l - q_l$ to the control group.

The general form of the permutation test statistic for current status and panel count data under PBD to test the tumor prevalence rate is given by

$$T = \sum_{l=1}^b w_l \sum_{j=1}^{n_l} a_{l,j} \rho_{l,j} = \sum_{l=1}^b \sum_{j=1}^{n_l} W_{l,j} \rho_{l,j}, \quad (2.1)$$

where w_l is the weight of the block l , $l = 1, 2, \dots, b$, $W_{l,j} = w_l a_{l,j}$, and $\rho_{l,j}$ is the group indicator for the j th individual belongs to the block l . Furthermore, $a_{l,j}$ is the score function of the test which varies according to the nature and type of data, and this will be addressed in more details in the following subsections for CS-I, CS-II and panel count data. Van Elteren [39] introduced the block weight $w_l = \frac{1}{n_l+1}$ as the optimal block weight. Thus, we consider this optimal block weight through the calculation of this article.

2.1. CS-I data

Consider the observed CS-I data for the j th individual in the l th block be $\{(t_{l,j}, \rho_{l,j}, \delta_{l,j}), j = 1, 2, \dots, n_l, l = 1, 2, \dots, b\}$, where $t_{l,j}$ is the examination time, $\rho_{l,j}$ is the treatment indicator ($\rho_{l,j} = 0$ or 1 if the j th individual belongs to the control or treatment group, respectively), and $\delta_{l,j}$ is an indicator of the occurrence of the event of interest before the examination time $t_{l,j}$. Sun and Kalbfleisch [12] introduced test statistic for testing, H_0 : The prevalence of tumors in the treatment and control groups is equal as follows:

$$T_1 = \sum_{l=1}^b w_l \sum_{j=1}^{n_l} S_{l,j}(\delta_{l,j} - \delta_{l,j}^*)\rho_{l,j}, \quad (2.2)$$

where $S_{l,j}$ is the score function and $\delta_{l,j}^* = \hat{G}(t_{l,j})$ is the isotonic estimate of the CDF, $G(t_{l,j})$, of the common propagation of the tumor. For more details on how to calculate $\hat{G}(t_{l,j})$, see Barlow et al. [40], and Sun and Kalbfleisch [12].

The statistic T_1 is asymptotically normally distributed [12]. It is worth noting that the statistics T_1 include a number of tests, including: Hoel and Walburg [41] statistic when $S_{l,j} = 1$ and Finkelstein [42] statistic when $S_{l,j} = -\log \frac{1-\delta_{l,j}^*}{\delta_{l,j}^*}$.

2.2. CS-II data

Let the observed CS-II data for the j th individual in the l th block be $\{(t_{l,j}, \rho_{l,j}, \psi_{l,j}), j = 1, 2, \dots, n_l, l = 1, 2, \dots, b\}$, where $t_{l,j}$ is the examination time, $\rho_{l,j}$ is the treatment indicator previously defined in Subsection 2.1, and $\psi_{l,j} = \Psi_{l,j}(t_{l,j})$ is the number of events that happened before the examination time $t_{l,j}$. Also, let $\mu_{l,j}(t) = E(\Psi_{l,j}(t)|\rho_{l,j})$ is the conditional mean function, MF, of the individual j in block l . To test the hypothesis of equality of the n_l mean functions within each block, the test statistic of Sun and Kalbfleisch [43] can be modified and used as follows:

$$T_2 = \sum_{l=1}^b w_l \sum_{j=1}^{n_l} (\psi_{l,j} - \psi_{l,j}^*)\rho_{l,j}, \quad (2.3)$$

where $\psi_{l,j}^*$ is the isotonic regression estimate, IRE, of the MF $\mu_{l,j}(t_{l,j})$. One can use the algorithm that was introduced by Barlow et al. [40] to evaluate the IRE of the MF. Sun and Kalbfleisch [43] derived the asymptotic distribution of T_2 , assuming the null hypothesis is true and $n_l \rightarrow \infty$ or $(n \rightarrow \infty)$.

2.3. Panel count data

As we mentioned earlier, panel count data is a generalization of CS-II data. The panel count data differs from the CS-II data in that each element of the study is examined more than once, and the

number of events that occurred between successive examination times is counted. Whereas in the CS-II data, each element of the study is examined only once, and the number of events that occurred before the time of the examination is counted. Through the relationship between the panel count data and CS-II, Sun and Fang [14] modified the test statistic in (2.3) to take into account the repeated or periodic examination of each element of the study. Since each element of the study is examined several times, the examination times for j th element of block l are as follows: $0 < t_{l,j,1} < t_{l,j,2} < \dots < t_{l,j,m_j}$, $l = 1, 2, \dots, b$, $j = 1, 2, \dots, n_l$, where m_j be the number of examination for the element j . In the framework of the previous description of the difference between the CS-II data and panel count data, the statistic (2.3) becomes

$$T_3 = \sum_{l=1}^b w_l \sum_{j=1}^{n_l} \rho_{l,j} \sum_{k=1}^{m_j} (\psi_{l,j,k} - \psi_{l,j,k}^*), \quad (2.4)$$

where $\psi_{l,j,k}^*$ is IRE of the MF. Sun and Fang [14] obtained the asymptotic normal distribution of T_3 .

While Sun and Fang [14] relied on the IRE method for estimating the MF, Balakrishnan and Zhao [15] suggested using the nonparametric maximum likelihood estimator, NPMLE, for estimating the MF. Thus, the statistic T_3 becomes

$$T_4 = \sum_{l=1}^b \frac{w_l}{\sqrt{n_l}} \sum_{j=1}^{n_l} \rho_{l,j} \left\{ \sum_{k=1}^{m_j-1} \hat{\mu}(t_{l,j,k}) \left[\frac{\Delta\psi_{l,j}(t_{l,j,k+1})}{\Delta\hat{\mu}(t_{l,j,k+1})} - \frac{\Delta\psi_{l,j}(t_{l,j,k})}{\Delta\hat{\mu}(t_{l,j,k})} \right] + \hat{\mu}(t_{l,j,m_j}) \left[1 - \frac{\Delta\psi_{l,j}(t_{l,j,m_j})}{\Delta\hat{\mu}(t_{l,j,m_j})} \right] \right\}, \quad (2.5)$$

where $\Delta\hat{\mu}(t_{l,j,k}) = \hat{\mu}(t_{l,j,k}) - \hat{\mu}(t_{l,j,k-1})$, $\Delta\psi_{l,j}(t_{l,j,k}) = \psi(t_{l,j,k}) - \psi(t_{l,j,k-1})$, and $\hat{\mu}(t)$ is the NPMLE of the common MF $\mu(t)$. Balakrishnan and Zhao [15] introduced the asymptotic normal distribution of T_4 .

3. Approximating the tail probabilities of statistic T

This section presents the proposed procedures for approximating the tail probabilities of statistic T under PBD. As we previously stated that the PBD divides the patients into independent blocks and within each block, patients are randomly distributed using the RAR. Therefore, the PBD is considered as repetition of the RAR. As a result of this, the permutation distribution of the vector of the treatment indicators $(\rho_{1,1}, \rho_{1,2}, \dots, \rho_{1,n_1}, \dots, \rho_{b,1}, \rho_{b,2}, \dots, \rho_{b,n_b})$ is $\prod_{l=1}^b \binom{n_l}{q_l}^{-1}$, where $\binom{n_l}{q_l}^{-1}$ is the permutation distribution of the vector of the treatment indicators of the block l , $(\rho_{l,1}, \rho_{l,2}, \dots, \rho_{l,n_l})$. The RAR ensures that a certain number of patients within block l are assigned to the treatment group, let this number of patients be q_l , i.e. $\sum_{j=1}^{n_l} \rho_{l,j} = q_l$. Thus, the permutation distribution of the vector of the treatment indicators $(\rho_{1,1}, \rho_{1,2}, \dots, \rho_{1,n_1}, \dots, \rho_{b,1}, \rho_{b,2}, \dots, \rho_{b,n_b})$ is equivalent to the distribution of the independent and identically Bernoulli(β_l) random variables, $Y_{1,1}, Y_{1,2}, \dots, Y_{1,n_1}, \dots, Y_{b,1}, Y_{b,2}, \dots, Y_{b,n_b}$ given that $\sum_{j=1}^{n_1} Y_{1,j} = q_1$, $\sum_{j=1}^{n_2} Y_{2,j} = q_2$, ..., $\sum_{j=1}^{n_b} Y_{b,j} = q_b$.

From the foregoing, we conclude that

$$T = \sum_{l=1}^b \sum_{j=1}^{n_l} W_{l,j} \rho_{l,j} \sim \sum_{l=1}^b \sum_{j=1}^{n_l} W_{l,j} Y_{l,j} \mid \sum_{j=1}^{n_1} Y_{1,j} = q_1, \sum_{j=1}^{n_2} Y_{2,j} = q_2, \dots, \sum_{j=1}^{n_b} Y_{b,j} = q_b. \quad (3.1)$$

Let $V = \sum_{l=1}^b \sum_{j=1}^{n_l} W_{l,j} Y_{l,j}$, and $U_l = \sum_{j=1}^{n_l} Y_{l,j}$, $l = 1, 2, \dots, b$, then the p-value of the statistic T can be obtained by evaluating the following probability

$$P_r(T \geq T_0) = P_r(V \geq T_0 | U_1 = q_1, U_2 = q_2, \dots, U_b = q_b), \quad (3.2)$$

where T_0 is the observed value of the statistic T . The probability in the right-hand side of Eq (3.2) can be approximated using the saddlepoint approximation of the conditional CDF [25] as follows:

$$P_r(V \geq T_0 | U_1 = q_1, U_2 = q_2, \dots, U_b = q_b) \approx 1 - \Phi(\hat{\eta}) - \phi(\hat{\eta}) \left(\frac{1}{\hat{\eta}} - \frac{1}{\hat{\nu}} \right), \quad (3.3)$$

where $\Phi()$ and $\phi()$ are the CDF and PDF of the standard normal distribution. Also, $\hat{\eta}$ and $\hat{\nu}$ are given as follows:

$$\hat{\eta} = \text{sgn}(\hat{\tau}) \sqrt{2 \left[\{C(0, \hat{\kappa}_{10}, \dots, \hat{\kappa}_{b0}) - \sum_{l=1}^b q_l \hat{\kappa}_{l0}\} - \{C(\hat{\tau}, \hat{\kappa}_1, \dots, \hat{\kappa}_b) - \sum_{l=1}^b q_l \hat{\kappa}_l - \hat{\tau} T_0\} \right]}, \quad (3.4)$$

and

$$\hat{\nu} = \hat{\tau} \sqrt{\frac{\det(C''(\hat{\tau}, \hat{\kappa}_1, \dots, \hat{\kappa}_b))}{\det(C''_{\kappa}(0, \hat{\kappa}_{10}, \dots, \hat{\kappa}_{b0}))}}, \quad (3.5)$$

where $C(\tau, \kappa_1, \dots, \kappa_b)$ is the joint cumulant generating function of $V, U_1, U_2, \dots, U_{b-1}$ and U_b which can be obtained as

$$C(\tau, \kappa_1, \dots, \kappa_b) = \sum_{l=1}^b \sum_{j=1}^{n_l} \log \left((1 - \beta_l) + \beta_l \exp\{W_{l,j} \tau + \kappa_l\} \right). \quad (3.6)$$

In Eq (3.5), C'' and C''_{κ} are the $(b+1) \times (b+1)$, and $b \times b$ Hessian matrices of $C(\tau, \kappa_1, \dots, \kappa_b)$, and $C(0, \kappa_1, \dots, \kappa_b)$, respectively. The abbreviation $\det(C'')$ refers to determinant of the matrix C'' .

Furthermore, $\hat{\tau}, \hat{\kappa}_1, \dots, \hat{\kappa}_{b-1}$, and $\hat{\kappa}_b$ are the solution of the following system of Eqs (3.7) and (3.8)

$$C'_{\tau}(\hat{\tau}, \hat{\kappa}_1, \dots, \hat{\kappa}_b) = T_0, \quad (3.7)$$

$$C'_{\kappa_l}(\hat{\tau}, \hat{\kappa}_1, \dots, \hat{\kappa}_b) = q_l, \quad l = 1, 2, \dots, b, \quad (3.8)$$

also, $\hat{\kappa}_{10}, \dots, \hat{\kappa}_{b0}$ are the solution of the following system of Eq (3.9)

$$C'_{\kappa_l}(0, \hat{\kappa}_{10}, \dots, \hat{\kappa}_{b0}) = q_l, \quad l = 1, 2, \dots, b. \quad (3.9)$$

It is clear that the permutation distribution of the statistic T does not depend on β_l 's values, for this reason, we can take β_l equals to q_l/n_l . This value of β_l makes $\hat{\kappa}_{l0} = 0$, $l = 1, 2, \dots, b$.

4. Illustrative examples and simulation study

This section explains the importance of the proposed method to approximate the exact p-value by comparing the proposed method with the normal approximation method. This comparison is conducted on two levels: The first level is real data analysis and the second level is a simulation study. During this section, statistics T_2 and T_3 are applied for current status and panel count data, respectively, and statistics T_1 and T_4 can be applied in the same manner. The exact p-value is approximated using the proposed method, which is the saddlepoint approximation method and the asymptotic normal approximation method for the considered tests. To compare the accuracy of the two approximation methods, we need to compute the exact p-value, and unfortunately, it cannot be calculated. Thus, we turn to the simulated method to introduce an accurate approximation to the exact mid-p-value by generating million of randomized PBD sequences for the control and treatment labels. Then, approximating the exact mid-p-value of the statistic T as $\{\sum I(T > T_0) + 0.5 \sum I(T = T_0)\}/10^6$. We refer to this approximated p-value as a reference mid-p-value or simulated mid-p-value. The mid-p-value is used here instead of the ordinary p-value due to its accuracy preference by many statisticians. Especially for hypothesis testing in discrete problems. For views that promote use of the mid-p-value; see Pierce and Peters [44], Routledge [45], Agresti and Gottard [46], and Chapter 7 of Butler [29].

4.1. Illustrative examples

An important way to clarify the purpose of this section and illustrate the suggested procedures is through real illustrative examples. In this regard, two sets of real data are analyzed, representing current status data and panel count data. The current status data set were observed from a tumorigenicity study of laboratory male and female mice by Li et al. [6]. This data is presented in Table 1 of the reference [6], and it represents the number of tumors detected at the time of death or the time of sacrifice for 199 mice, 100 of them are males ($\rho_{l,j} = 1$) and 99 are females ($\rho_{l,j} = 0$). The hypothesis of there is no difference in the rate of tumor development in groups of male and female mice is tested under PBD with nine blocks of sizes $n_1 = 23$, and $n_l = 22$, $l = 2, \dots, 9$ and $q_l = 14, 17, 7, 15, 10, 15, 7, 7, 8$, for $l = 1, \dots, 9$, respectively.

Table 1. P-values for tumorigenicity and bladder tumor data.

Example	Simulated mid-p-value	SPA p-value	ANA p-value	$RAE(SPA)$	$RAE(ANA)$
Example 1	0.032743	0.032798	0.034420	0.001679	0.051217
Example 2	0.014673	0.014309	0.016340	0.024807	0.113610
Example 3	0.000896	0.000809	0.001130	0.097098	0.261160
Example 4	0.353995	0.325680	0.167297	0.079987	0.527402
Example 5	0.364201	0.389976	0.292642	0.070771	0.196482

Table 1 represents the simulated mid-p-value, saddlepoint approximation, SPA, p-value and asymptotic normal approximation, ANA, p-value, for this data set, in the first row in front of the title “Example 1”. As an example of samples of medium and small sizes, two samples of sizes 100 and 32 were taken from the total sample. The results of these two samples are indicated in Table 1 as “Example 2” and “Example 3”, respectively. For the sample of size 100, the number of blocks was 5

with block size $n_l = 20$, and $q_l = 5, 15, 9, 12, 9, l = 1, \dots, 5$. Furthermore, for the sample of size 32, the number of blocks was 4 with block size $n_l = 8$, and $q_l = 5, 5, 2, 4, l = 1, \dots, 4$.

The second application example represents panel count data observed from a bladder tumors study [47, 48]. The study began by removing all tumors from 72 patients and then distributing the patients into two groups, one of which was a treatment group and the other a control group. At the monthly follow-up visit for each patient during the year, the number of bladder tumors that were detected is recorded. The hypothesis of there is no difference in the rate of tumor development in treatment and control groups is tested under PBD with six blocks of sizes $n_l = 12$, and $q_l = 9, 4, 6, 7, 6, 6$, for $l = 1, \dots, 6$, respectively. The three p-values for such data set are presented in Table 1 in front of the title “Example 4”. As an example of a sample of small size, a sample of size 50 was taken from the total sample. For this sample, the number of blocks was 5 with block size $n_l = 10$, and $q_l = 7, 4, 4, 4, 6, l = 1, \dots, 5$. Also, the three p-values for such data set are presented in Table 1 in front of the title “Example 5”. Furthermore, Table 1 represents the relative absolute errors, *RAE*, of the saddlepoint and asymptotic normal approximation methods with respect to the simulated method which is considered here as the reference method.

By comparing the three columns of the p-values in Table 1, we can note that, in all the proposed examples, the SPA p-values are closer to the simulated p-values than the ANA p-values. Accordingly, we can say that the proposed method is more accurate than ANA method.

4.2. Simulation study

It is not possible to completely rely on the accuracy that appeared through the illustrative examples, although it was a good indicator of the accuracy of the proposed method. Therefore, a simulation study is conducted to verify that accuracy and form a clear conclusion about the accuracy of the proposed method. In this context, clinical trials are simulated to yield current status and panel count data. First, the current status data for the block $l, l = 1, 2, \dots, b$ is generated from Poisson($\mu_{l,j}(t_{l,j})e^{\alpha\rho_{l,j}}$) processes, where α is regression coefficient, the MF $\mu_{l,j}(t_{l,j}) = \lambda t_{l,j}$, λ is constant and $t_{l,j} = j/n_l, j = 1, 2, \dots, n_l$. Secondly, the panel count data for the block $l, l = 1, 2, \dots, b$ is generated assuming that we have a clinical trial in which patients are examined monthly for half a year. Accordingly, the number of examination for the patient j in block l , follows Uniform distribution $U\{1, 2, \dots, 6\}$. As in the current status data, the number of the repeated event that found out at each examination time follows a Poisson process with MF, $\mu(t_{l,j,k}) \exp\{\alpha\rho_{l,j}\}$, where $\mu(t_{l,j,k}) = \lambda t_{l,j,k}$, and $t_{l,j,k} = k/6, k = 1, 2, \dots, 6, j = 1, 2, \dots, n_l, l = 1, 2, \dots, b$. 1,000 current status and panel count data sets are generated according to four scenarios, SC, namely SC1, SC2, SC3 and SC4 as follows:

- SC1: $n = 30, b = 5, n_l = 6, q_l = 3, l = 1, 2, \dots, 5$.
- SC2: $n = 40, b = 5, n_l = 8, q_l = 4, l = 1, 2, \dots, 5$.
- SC3: $n = 60, b = 6, n_l = 10, q_l = 5, l = 1, 2, \dots, 6$.
- SC4: $n = 80, b = 8, n_l = 10, q_l = 5, l = 1, 2, \dots, 8$.

The treatment indicators $\rho_{l,j}$ are chosen by simulating a PBD sequence for each simulated data set. Since the aim of this subsection is to emphasize the accuracy of the proposed method, which appeared clearly in the examples illustrated in the previous subsection. Accordingly, three criteria are proposed to compare the saddlepoint and asymptotic normal approximation methods. The first criterion expresses the percentage of the number of times that the proposed approximation was closer

to the reference method, which is the simulation method, which needs a lot of calculations, as we explained how to calculate it at the beginning of this section. We refer to this criterion as $PER(SPA)$. The second and third criteria both measure the amount of error resulting from both methods to approximate the exact p-value of the considered class of tests. These two criteria are the mean relative absolute error and mean square error for both methods. We refer to them as $RAE(SPA)$, $RAE(ANA)$, $MSE(SPA)$, and $MSE(ANA)$. The $RAE(SPA)$ can be obtained as $RAE(SPA) = \frac{1}{N} \sum_{i=1}^N |P_i - \hat{P}_i|/P_i$, where P_i and \hat{P}_i are the simulated and SPA p-values, respectively. Similarly, the $RAE(ANA)$ can be calculated by replacing SPA p-value by ANA p-value. Furthermore, the $MSE(SPA)$ can be obtained as $MSE(SPA) = \frac{1}{N} \sum_{i=1}^N (P_i - \hat{P}_i)^2$. The $MSE(ANA)$ also can be calculated by replacing SPA p-value by ANA p-value in the last formula, where $N = 1,000$. The three comparison criteria are summarized in Table 2 of the current status data and in Table 3 for panel count data.

Table 2. The results of the three comparison criteria for the current status data.

SC	n	α	λ	$PER(SPA)$	$RAE(SPA)$	$RAE(ANA)$	$MSE(SPA)$	$MSE(ANA)$
SC1	30	0.2	6	81.9	0.0337	0.1417	3.8274×10^{-5}	1.0536×10^{-4}
			9	83.3	0.0331	0.1391	3.4174×10^{-5}	1.0564×10^{-4}
			13	82.4	0.0326	0.0947	3.9993×10^{-5}	1.2438×10^{-4}
SC2	40	0.3	6	85.1	0.0198	0.0942	8.8331×10^{-6}	2.9461×10^{-5}
			9	83.9	0.0185	0.0872	8.9298×10^{-6}	3.1368×10^{-5}
			13	82.4	0.0215	0.1647	8.6034×10^{-6}	3.3513×10^{-5}
SC3	60	0.4	6	82.0	0.0144	0.1108	1.0042×10^{-6}	4.8999×10^{-6}
			9	85.4	0.0149	0.1309	9.4446×10^{-7}	5.0233×10^{-6}
			13	84.0	0.0177	0.1738	8.0937×10^{-7}	4.3561×10^{-6}
SC4	80	0.6	6	85.4	0.0177	0.1510	2.9212×10^{-7}	1.4843×10^{-6}
			9	88.1	0.0191	0.1776	2.5418×10^{-7}	1.3690×10^{-6}
			13	86.1	0.0194	0.2185	2.3148×10^{-7}	1.3115×10^{-6}

Table 3. The results of the three comparison criteria for the panel count data.

SC	n	α	λ	$PER(SPA)$	$RAE(SPA)$	$RAE(ANA)$	$MSE(SPA)$	$MSE(ANA)$
SC1	30	0.3	2	95.6	0.0866	0.8150	6.0992×10^{-4}	0.0590
			4	96.5	0.0912	0.8370	6.1364×10^{-4}	0.0621
			5	95.4	0.1039	1.5151	5.9696×10^{-4}	0.0583
SC2	40	0.4	2	93.4	0.0921	0.7561	7.3496×10^{-4}	0.0619
			4	95.3	0.1051	1.0024	7.1451×10^{-4}	0.0587
			5	94.8	0.0999	0.9172	7.1403×10^{-4}	0.0573
SC3	60	0.5	2	95.6	0.1009	1.0071	8.1288×10^{-4}	0.0658
			4	93.0	0.1131	1.1161	8.7497×10^{-4}	0.0708
			5	92.6	0.1176	0.9039	9.1274×10^{-4}	0.0580
SC4	80	0.8	2	92.3	0.1122	1.1069	8.7906×10^{-4}	0.0656
			4	90.9	0.1350	1.7249	0.0011	0.0730
			5	90.8	0.1294	1.0177	0.0012	0.0656

All the results presented in Tables 2 and 3 confirm the observation drawn from the illustrative examples, which is the proposed method is more accurate than the traditional method in approximating the exact p-value of the proposed class of tests. For example, we find that the lowest percentage of the proposed method approaching the reference method is $PER(SPA) = 81.9\%$, while the highest percentage exceeds 96%. Moreover, if we take the results presented in the first row in Table 2, we note that the mean relative absolute error for the proposed and traditional methods are 3.37% and 14.17%, respectively. In this case, the mean relative absolute error resulting from the traditional method is approximately three times the corresponding error resulting from the proposed method. To illustrate this in another way, we plot the relative absolute errors for both approximation methods, SPA and ANA methods, for the aforementioned case in Figure 1.

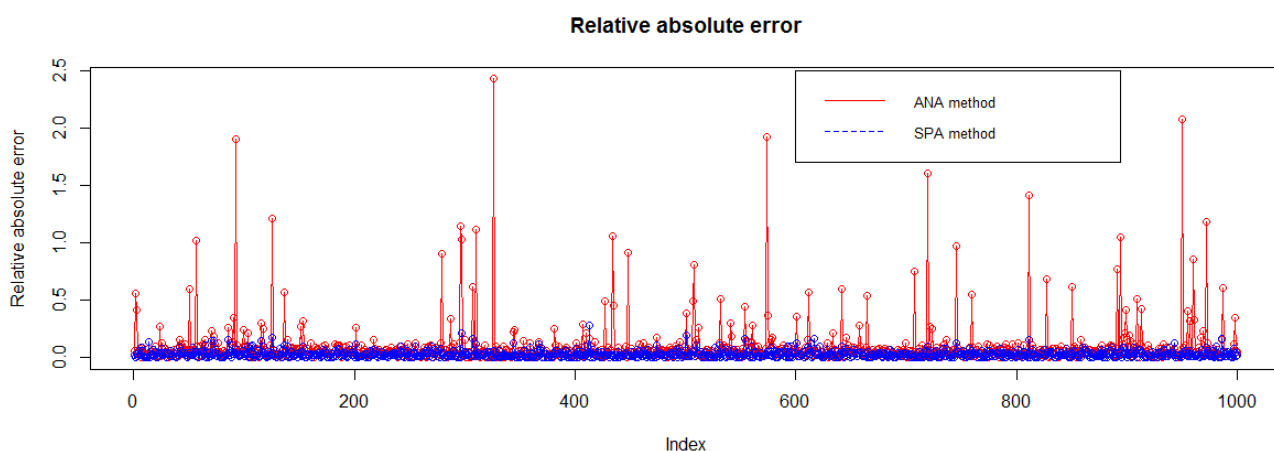


Figure 1. Relative absolute errors for both SPA and ANA methods.

We would like to point out that the exact p-value of the considered class of nonparametric tests can be approximated using three approximation methods: The asymptotic normal approximation method, the simulation method, the saddlepoint approximation method, and the last one is the method proposed in this article. We have shown later that the saddlepoint approximation method is more accurate than the asymptotic normal approximation method. Therefore, we can use the saddlepoint approximation method as an alternative to the asymptotic normal approximation method for approximating the exact p-value. But there remains an important question here that may interest readers. This question is why we may use the saddlepoint method as an alternative to the simulation method. The answer is that the saddlepoint method is computationally less demanding than the simulation method. To make this matter more clear to the readers, the computing times for saddlepoint and simulation methods were recorded and these results are presented in Tables 4 and 5 for the two data types used in this article. From Tables 4 and 5, we note that to find the average value of 1,000 p-values that were calculated using the saddlepoint method, we need approximately a minute for all the imposed sample sizes. While we need 15 to 23 hours to calculate the corresponding value using the simulation method. Thus, we can say that the proposed method saves a lot of time compared to the simulation method and it is more accurate than the asymptotic normal approximation. Thus, we can suggest it as a good alternative to both asymptotic normal and simulated methods.

Table 4. Computing times in seconds for approximating the exact p-values using saddlepoint and simulation methods for the current status data.

SC	n	α	λ	Method of approximation	
				SPA	Simulation
SC1	30	0.2	6	53.68	55214.76
			9	54.18	55397.42
			13	54.28	55133.70
SC2	40	0.3	6	54.53	56227.81
			9	55.53	56336.91
			13	54.58	56616.22
SC3	60	3*0.4	6	68.44	67432.92
			9	66.49	67584.30
			13	70.34	67665.36
SC4	80	0.6	6	70.85	86348.34
			9	69.44	87901.32
			13	71.50	85761.00

Table 5. Computing times in seconds for approximating the exact p-values using saddlepoint and simulation methods for the panel count data.

SC	n	α	λ	Method of approximation	
				SPA	Simulation
SC1	30	0.3	2	65.86	55472.85
			4	64.32	56250.42
			5	66.79	55516.63
SC2	40	0.4	2	65.44	56310.87
			4	64.44	56826.64
			5	64.19	57470.86
SC3	60	0.5	2	73.75	69500.82
			4	74.00	67675.92
			5	71.85	67915.20
SC4	80	0.8	2	78.46	87364.08
			4	76.60	92884.86
			5	77.25	88322.82

5. Conclusions

In this article, a number of nonparametric tests for the current status and panel count data which appear in many studies such as clinical trials, epidemiology, and demography are formulated in the form of a class of linear rank tests. The saddlepoint approximation method is proposed to approximate the exact distribution of the proposed class of tests under PBD. The proposed method of approximation showed high accuracy compared to the normal approximation method which is widely used with these types of tests. Thus, we can conclude that the proposed method is a good alternative to the normal

approximation method, and the proposed method does not require a lot of calculations that take a lot of time, like the simulated method. Therefore, we can say that the proposed method has the required accuracy and does not require much time in its calculation.

Use of AI tools declaration

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

Acknowledgments

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 RGP2/310/44.

Conflict of interest

The authors declare no conflict of interest.

References

1. J. Huang, J. Wellner, *Interval censored survival data: A review of recent progress*, In: Proceedings of the first seattle symposium in biostatistics, Springer, 1997, 123–169.
2. J. G. Sun, *The statistical analysis of interval-censored failure time data*, Springer, 2006.
3. J. J. Gart, D. Krewski, P. N. Lee, R. E. Tarone, J. Wahrendorf, *Statistical methods in cancer research. volume III—the design and analysis of long-term animal experiments*, IARC Scientific Publications, **79** (1986), 1–219. <https://doi.org/10.2307/2290099>
4. N. P. Jewell, S. C. Shiboski, Statistical analysis of HIV infectivity based on partner studies, *Biometrics*, **46** (1990), 1133–1150. <https://doi.org/10.2307/2532454>
5. N. Keiding, K. Begtrup, T. H. Scheike, G. Hasibeder, Estimation from current-status data in continuous time, *Lifetime Data Anal.*, **2** (1996), 119–129. <https://doi.org/10.1007/BF00128570>
6. Y. H. Ii, R. Kikuchi, K. Matsuoka, Two-dimensional (time and multiplicity) statistical analysis of multiple tumors, *Math. Biosci.*, **84** (1987), 1–21. [https://doi.org/10.1016/0025-5564\(87\)90040-X](https://doi.org/10.1016/0025-5564(87)90040-X)
7. I. D. Diamond, J. W. McDonald, *The analysis of current status data*, In: J. Trussell, R. Hankinson, and J. Tilton (eds) Demographic Applications of Event History Analysis, Oxford University Press, 1991.
8. G. E. Dinse, A comparison of tumour incidence analyses applicable in single-sacrifice animal experiments, *Stat. Med.*, **13** (1994), 689–708. <https://doi.org/10.1002/sim.4780130530>
9. P. F. Thall, J. M. Lachin, Analysis of recurrent events: Nonparametric methods for random-interval count data, *J. Am. Stat. Assoc.*, **83** (1988), 339–347. <https://doi.org/10.1080/01621459.1988.10478603>
10. J. Sun, J. D. Kalbfleisch, Estimation of the mean function of point processes based on panel count data, *Stat. Sinica*, **5** (1995), 279–289. <https://doi.org/10.1007/BF01192198>

11. J. A. Wellner, Y. Zhang, Two estimators of the mean of a counting process with panel count data, *Ann. Stat.*, **28** (2000), 779–814. <https://doi.org/10.2307/2674053>
12. J. G. Sun, J. D. Kalbfleisch, Nonparametric tests of tumor prevalence data, *Biometrics*, **52** (1996), 726–731. <https://doi.org/10.2307/2532912>
13. J. G. Sun, A nonparametric test for current status data with unequal censoring, *J. R. Stat. Soc. B*, **61** (1999), 243–250. <https://doi.org/10.1111/1467-9868.00174>
14. J. G. Sun, H. B. Fang, A nonparametric test for panel count data, *Biometrika*, **90** (2003), 199–208. <https://doi.org/10.1093/biomet/90.1.199>
15. N. Balakrishnan, X. Q. Zhao, A nonparametric test for the equality of counting processes with panel count data, *Comput. Stat. Data Anal.*, **54** (2010), 135–142. <https://doi.org/10.1016/j.csda.2009.07.015>
16. D. J. McEntegart, The pursuit of balance using stratified and dynamic randomization techniques: An overview, *Drug Inform. J.*, **37** (2003), 293–308. <https://doi.org/10.1177/009286150303700305>
17. V. W. Berger, Varying the block size does not conceal the allocation, *J. Crit. Care*, **2** (2006), 229. <https://doi.org/10.1016/j.jcrc.2006.01.002>
18. G. R. Pond, P. A. Tang, S. A. Welch, E. X. Chen, Trends in the application of dynamic allocation methods in multi-arm cancer clinical trials, *Clin. Trials*, **7** (2010), 227–234. <https://doi.org/10.1177/1740774510368301>
19. J. P. Matts, J. M. Lachin, Properties of permuted-block randomization in clinical trials, *Control. Clin. Trials*, **9** (1988), 327–344. [https://doi.org/10.1016/0197-2456\(88\)90047-5](https://doi.org/10.1016/0197-2456(88)90047-5)
20. J. M. Lachin, J. P. Matts, L. J. Wei, Randomization in clinical trials: Conclusions and recommendations, *Control. Clin. Trials*, **9** (1988), 365–374. [https://doi.org/10.1016/0197-2456\(88\)90049-9](https://doi.org/10.1016/0197-2456(88)90049-9)
21. J. Efrid, Blocked randomization with randomly selected block sizes, *Int. J. Env. Res. Pub. He.*, **8** (2011), 15–20. <https://doi.org/10.3390/ijerph8010015>
22. R. L. Strawderman, Higher-order asymptotic approximation: Laplace, saddlepoint, and related methods, *J. Am. Stat. Assoc.*, **95** (2000), 1358–1364. <https://doi.org/10.1080/01621459.2000.10474348>
23. H. E. Daniels, Saddlepoint approximations in statistics, *Ann. Math. Stat.*, **25** (1954), 631–650. <https://doi.org/10.1214/aoms/1177728652>
24. R. Lugannani, S. Rice, Saddlepoint approximation for the distribution of the sum of independent random variables, *Adv. Appl. Prob.*, **12** (1980), 475–490. <https://doi.org/10.1017/S0001867800050278>
25. I. M. Skovgaard, Saddlepoint expansions for conditional distributions, *J. Appl. Prob.*, **24** (1987), 875–887. <https://doi.org/10.2307/3214212>
26. S. J. Wang, Saddlepoint approximations for bivariate distributions, *J. Appl. Prob.*, **27** (1990), 586–597. <https://doi.org/10.1017/S0021900200039139>
27. H. E. Daniels, Exact saddlepoint approximations, *Biometrika*, **67** (1980), 59–63. <https://doi.org/10.1093/biomet/67.1.59>

28. A. C. Davison, D. V. Hinkley, Saddlepoint approximations in resampling methods, *Biometrika*, **75** (1988), 417–431. <https://doi.org/10.1093/biomet/75.3.417>
29. R. W. Butler, *Saddlepoint approximations with applications*, Cambridge University Press, UK, 2007.
30. 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>
31. 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.
32. A. M. Abd El-Raheem, E. F. Abd-Elfattah, Weighted log-rank tests for clustered censored data: Saddlepoint p-values and confidence intervals, *Stat. Meth. Med. Res.*, **29** (2020), 2629–2636. <https://doi.org/10.1177/0962280220908288>
33. A. 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>
34. 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.-Theor. Meth.*, **52** (2023), 4103–4113.
35. K. S. Kamal, A. M. Abd El-Raheem, E. F. Abd-Elfattah, Weighted log-rank tests for left-truncated data under weis urn design: Saddlepoint p-values and confidence intervals, *J. Biopharm. Stat.*, **32** (2022), 641–651.
36. A. M. Abd El-Raheem, E. F. Abd-Elfattah, Weighted log-rank tests for censored data under wei's urn design: Saddlepoint approximation and confidence intervals, *J. Biopharm. Stat.*, In press, 2023.
37. 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** (2023), 1–8.
38. A. 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.*, In press, 2023.
39. P. H. V. Elteren, On the combination of independent two sample tests of wilcoxon, *Bull. Int. Stat. Inst.*, **37** (1960), 351–361. <https://doi.org/10.1016/j.chemosphere.2013.08.020>
40. R. E. Barlow, D. J. Bartholomew, J. M. Bremner, H. D. Brunk, *Statistical inference under order restrictions*, John Wiley, New York, 1972.
41. D. G. Hoel, H. E. Walburg, Statistical analysis of survival experiments, *J. Natl. Cancer I.*, **49** (1972), 361–372. <https://doi.org/10.1093/jnci/49.2.361>
42. D. M. Finkelstein, A proportional hazards model for interval-censored failure time data, *Biometrics*, **42** (1986), 845–854. <https://doi.org/10.2307/2530698>
43. J. G. Sun, J. D. Kalbfleisch, The analysis of current status data on point processes, *J. Am. Stat. Assoc.*, **88** (1993), 1449–1454. <https://doi.org/10.1080/01621459.1993.10476432>

44. D. A. Pierce, D. Peters, Practical use of higher order asymptotics for multiparameter exponential families, *J. Roy. Stat. Soc. Ser. B*, **54** (1992), 701–725. <https://doi.org/10.1111/j.2517-6161.1992.tb01445.x>
45. R. D. Routledge, Practicing safe statistics with the mid-p, *Can. J. Stat.*, **22** (1994), 103–110. <https://doi.org/10.2307/3315826>
46. A. Agresti, A. Gottard, Comment: Randomized confidence intervals and the mid-p approach, *Stat. Sci.*, **20** (2005), 367–371.
47. D. P. Byar, *The veterans administration study of chemoprophylaxis for recurrent stage i bladder tumours: Comparisons of placebo, pyridoxine and topical thiotepa*, In *Bladder tumors and other topics in urological oncology*, Springer, 1980, 363–370.
48. C. S. Davis, L. J. Wei, Tnonparametric methods for analyzing incomplete nondecreasing repeated measurements, *Biometrics*, **44** (1988), 1005–1018. <https://doi.org/10.2307/2531731>



AIMS Press

© 2023 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)