

## Research Article

# Statistical Inference of the Class of Nonparametric Tests for the Panel Count and Current Status Data from the Perspective of the Saddlepoint Approximation

Abd El-Raheem M. Abd El-Raheem <sup>1</sup>, Mona Hosny <sup>2</sup>, and Ehab F. Abd-Elfattah <sup>1</sup>

<sup>1</sup>Department of Mathematics, Faculty of Education, Ain Shams University, Roxy, Cairo 11341, Egypt

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

Correspondence should be addressed to Abd El-Raheem M. Abd El-Raheem; [abdelrahimmohamed@edu.asu.edu.eg](mailto:abdelrahimmohamed@edu.asu.edu.eg)

Received 26 October 2022; Revised 21 December 2022; Accepted 23 December 2022; Published 11 January 2023

Academic Editor: Niansheng Tang

Copyright © 2023 Abd El-Raheem M. Abd El-Raheem et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Many statisticians resort to using the asymptotic normal approximation method to carry out statistical inference for many statistical tests, especially nonparametric ones. In this article, the saddlepoint approximation method is proposed as an alternative to the asymptotic normal approximation method to carry out statistical inference for a number of nonparametric tests for an important type of data that appears frequently in many clinical studies such as cancer and tumorigenicity studies. In clinical trials, there are many strategies through which treatments are assigned to patients. Equal allocation of both treatments is a largely prevalent approach in clinical trials to eliminate experimental bias and increase power. Accordingly, the statistical analysis is carried out based on the truncated binomial design, which is one of the designs that achieve a perfect balance between the two treatments. To clarify the accuracy of the proposed approximation method, two sets of real data are analyzed, and for the same purpose, a comprehensive simulation study is carried out.

## 1. Introduction

Medical and epidemiological studies are performed mainly to measure the occurrence of an event. This event of interest is often the onset of a disease, the disappearance of the diseases symptoms, the attainment of a biochemical marker, or death. In many cases, the event of interest may happen repeatedly, for instance, occurrences of the same infection, such as recurrent pyogenic infections, multiple upper respiratory infections, recurrent bacterial infections, and repeated occurrences of certain tumors. Data collected from studies that include such recurrent events can be distinguished according to the way the items under study are monitored. If the items under study are monitored continuously and the times of all occurrences of recurrent events are available, these data are commonly referred to as recurrent event data [1–3]. While if it is not possible to monitor the items under study continuously, and it is only

possible to examine the items under study at prescheduled time points. Therefore, only the numbers of occurrences of the events within consecutively prescheduled times are available. In this case, the data collected from such studies is called panel count data [4–6]. For two main reasons, we focus here on the second type of data, panel count data. The first reason is the continuous observation for the items under study may be too expensive. The second reason, in some cases, is impossible to monitor the items under study continuously. When each item under the study is examined only once, a special case of panel count data arises, and such data are commonly known in literature as current status data [7]. In this case, the data consists of the total number of occurrences of the recurrent event up to the observation time. Panel count data frequently appears in reliability and medical follow-up studies [5, 6, 8–10]. A number of authors have provided some nonparametric tests for the panel count data. Among them are Thall and Lachin [6]; Sun and Fang

[11]; Zhang [12]; Park et al. [13]; and Balakrishnan and Zhao [14]. Current status data appears in tumorigenicity studies that involve the incidence rate of particular tumors, and in these studies, only the number of tumors that occurred prior to the animal's death or sacrifice is known. Demographic studies are another field that commonly generates current status data [15, 16]. Comparison of the rate of development of the recurrent event, in different treatments, has obtained a lot of attention (see Dinse and Lagakos [17]; Finkelstein [18]; Huang [19]; and Sun and Kalbfleisch [20]). Almost all of the approaches presented in the literature demand that all subjects' observation times follow the same distribution. Sun [21] proposed a permutation test that can be applied when the observed time distributions differed between groups.

Clinical trial design involves choosing a strategy for assigning treatments to patients. The strategies used to assign treatments to patients should be completely randomized to avoid selection bias, which may be intentional or unintended. There are many randomization designs, which are the basis for statistical inference. The simplest randomization design is called the complete randomization design, in which a balanced coin is thrown each time a patient has to be assigned to treatment. The hallmark of complete randomization is that it provides the highest level of randomization but may lead to treatment imbalances that may accidentally bias trial results. Thus, equal allocation of both treatments has become a largely dominant approach in clinical trials to eliminate experimental bias and increase power [22]. Random allocation rule (RAR) and truncated binomial design (TBD) are most commonly used designs to ensure balance between treatments. For sample size  $N$ , one can apply the RAR by selecting randomly from all probable sequences of treatment assignments consisting of  $N/2$  patients on each treatment. To use the TBD, toss a balanced coin until one treatment is assigned to  $N/2$  patients; all subsequent patients will receive the opposite treatment. Rukhin [23] pointed out that RAR and TBD turn out to be completely different: the target fulfillment moments for the TBD happen much earlier than for the RAR. In this article, we propose an approximation method to approximate the exact  $p$ -values of a class of nonparametric tests for current status and panel count data under TBD. It is worth noting that Abd-Elfattah [24] approximated the  $p$ -values of a set of nonparametric tests for current status and panel count data under RAR.

Lots of approximations are routinely used in statistics. These approximations are often approximations to some intractable integrals. Such approximations are typically developed under the assumption that the sample size becomes infinitely large. In general, asymptotic approaches lead to approximations that are accurate to the first order. Saddlepoint methods provide higher-order approximations, and the precision holds even for extremely small sample sizes. Furthermore, saddlepoint approximations are particularly precise in the distribution's tails and are frequently employed as tail-probability approximations. Daniels [25] introduced the most fundamental saddlepoint approximation to the statistical community, which is simply a method for approximating the density or the mass function of a statistic from its cumulant generating function, CGF. Daniels

explained that the saddlepoint approximation is a substantial improvement over the approximations provided by both the central limit theorem and Edgeworth expansions. Daniels' highly influential article on developing high-order approximations in statistics has led several authors to provide approximations for a number of statistical functions and probability quantities. For example, Lugannani and Rice [26] derived the saddlepoint approximation for the tail probability  $P(X \leq x)$ . Skovgaard [27] proposed an approximation for conditional probabilities of the form  $P(X \leq x | Y = y)$ . The saddlepoint approximation for the conditional probabilities is known as a double saddlepoint approximation. Wang [28] developed a saddlepoint approximation for the bivariate probabilities of the form  $P(X \leq x, Y \leq y)$ . For the saddlepoint approximations of multivariate probabilities, see Kolassa [29]. A number of statisticians have applied these different saddlepoint approximations to solve some statistical problems including Daniels [30]; Robinson [31]; Daniels [32]; Daniels [33]; Davison and Hinkley [34]; Daniels [35]; and Booth and Butler [36]. After that, saddlepoint applications have spread in many statistical fields such as survival analysis [37, 38], regression analysis [39, 40], reliability analysis [41, 42], bootstrapping [43, 44], resampling analysis [28, 34], nonparametric statistics Abd-Elfattah and Butler [45, 46]; Abd El-Raheem and Abd-Elfattah [47, 48]; Kamal et al. [49, 50]. Here, we have mentioned few applications and references of the saddlepoint applications for rigorous proofs and more applications, see Butler [51].

Abd-Elfattah [52] applied double saddlepoint approximation technique to approximate the tail probability of the linear rank class, for ordinary survival data under TBD. Here, we extend the results of Abd-Elfattah [52] to an important type of data, current status and panel count data, that appear frequently in demographic studies, acquired immunodeficiency syndrome studies, tumorigenicity experiments, and human immune deficiency virus studies.

The second section contains a number of nonparametric tests for the panel count and current status data, which are finally reformatted as a linear combination of the treatment indicator and score function. The proposed method for approximating the tail probability for statistics of nonparametric tests for the panel count and current status data is presented in the third section. The fourth section is devoted to clarifying the accuracy of the proposed method compared to the normal approximation method, by analyzing a number of real data sets. Moreover, the accuracy of the proposed method is confirmed by conducting a comprehensive simulation study. Finally, an adequate conclusion is presented for the results of this article.

## 2. Nonparametric Tests for Current Status and Panel Count Data

This section contains the statistics of a number of nonparametric tests for the panel count and current status data, which are finally reformatted as a linear combination of the treatment indicator and score function.

2.1. Nonparametric Tests for Current Status Data

2.1.1. *Nonparametric Test Procedure I.* Suppose a tumorigenicity study with  $N$  independent patients divided into two groups: control and treatment. Such a tumorigenicity study aims to test the null hypothesis that there is no difference in the tumor propagation rate in the two groups. The observed data for the  $j$ th patient is  $(\delta_j, x_j, \xi_j)$ , where  $\delta_j$  is the treatment indicator,  $x_j$  is the observed time, and  $\xi_j$  is the tumor indicator at  $x_j$ . Let  $G(x) = P(X \leq x)$  denote the cumulative distribution function of the common propagation of the tumor,  $\hat{G}(x)$  be the maximum likelihood estimate, MLE, of  $G(x)$ , and  $\{\tau_1, \tau_2, \dots, \tau_{m_N}\}$  represent the distinct ordered times of  $\{x_1, x_2, \dots, x_N\}$ . Under the premise that the tumor is not fatal, i.e., that the time of observation and time to event are independent, a log-rank type test statistic for testing null hypothesis,  $H_0$ , is given by the following equation [20]:

$$S_1 = \sum_{j=1}^N \psi_j (\xi_j - \xi_j^*) \delta_j, \quad (1)$$

where  $\psi_j$  is a prespecified weight,  $\xi_j^* = \hat{G}(\tau_i)$  at  $x_j = \tau_i$ , and  $\hat{G}(\tau_i)$  is the weighted isotonic regression of  $\{\bar{\xi}_1/w_1, \bar{\xi}_2/w_2, \dots, \bar{\xi}_{m_N}/w_{m_N}\}$  with weights  $\{w_1, w_2, \dots, w_{m_N}\}$ . Thus,

$$\hat{G}(\tau_i) = \frac{\sum_{\ell \in T_k} \bar{\xi}_\ell}{\sum_{\ell \in T_k} w_\ell}, \quad i \in T_k, \quad (2)$$

and  $k = 1, 2, \dots, M_N$  [53]. In the previous expression,  $w_i$  is the number of subjects who died or were sacrificed at time  $\tau_i$ ,  $\bar{\xi}_i$  is the number of tumors discovered at the time  $\tau_i$ ,  $M_N$  is the number of the blocks of the isotonic estimate, and  $\{T_1, T_2, \dots, T_{M_N}\}$ , a partition of  $\{1, \dots, m_N\}$ , the blocks of the isotonic estimate of  $\hat{G}(x)$ . The statistic  $S_1$  is asymptotically normally distributed with mean 0 and standard deviation as follows:

$$V_1 = \left[ \hat{\sigma}_1^2 \sum_{j=1}^N (\xi_j - \xi_j^*)^2 \right]^{1/2}, \quad (3)$$

where  $\hat{\sigma}_1^2 = (1/N) \sum_{j=1}^N (\psi_j \delta_j - (1/N) \sum_{j=1}^N \psi_j \delta_j)^2$ . It should be stated that the statistic  $S_1$  with  $\psi_j = 1$  reduces to the statistics of Hoel and Walburg [54] and to the statistic of Finkelstein [18] when  $\psi_j = -\log(1 - \xi_j^*/\xi_j^*)$ .

2.1.2. *Nonparametric Test Procedure II.* Suppose an experiment in which  $N$  independent individuals are exposed to recurrent events over time. Let  $\Lambda_j(x)$  be the number of events that happened for the  $j$ th subject at time  $x$  and  $\mu_j(x) = E(\Lambda_j(x) | \delta_j)$  be the conditional expectation function of  $\Lambda_j(x)$  given the indicator variable  $\delta_j$  of individual  $j$ ,  $j = 1, 2, \dots, N$ . The observed data for the  $j$ th individual is  $(\delta_j, x_j, \lambda_j)$ , where  $x_j$  denote the observed time for the  $j$ th individual and  $\lambda_j = \Lambda_j(x_j)$  is the number of events that occurred

at time  $x_j$ . Supposing that the observed times are independent of  $\Lambda_j(x)$  given  $\delta_j$ , and follow the same distribution for both treatment groups, a statistic  $S_2$  for testing the null hypothesis  $H_0$ , that  $\mu_j(x)$  is independent of  $\delta_j$ , is given by the following equation [7]:

$$S_2 = \sum_{j=1}^N \delta_j (\lambda_j - \lambda_j^*), \quad (4)$$

where  $\lambda_j^*$  is the isotonic regression estimate (IRE) of  $\mu_j(x_j)$ , which can be evaluated using the pool-adjacent-violators algorithm [53].

It should be noted that if  $\Lambda_j(x_j)$ ,  $j = 1, 2, \dots, N$  are independent Poisson processes with mean function (MF), then

$$\mu_j(x) = \mu(x) \exp(\eta \delta_j), \quad (5)$$

where  $\mu(x)$  is the baseline MF and  $\eta$  is regression coefficient, then the statistic  $S_2$  in equation (4) seems as a score test for  $\eta = 0$  from the conditional likelihood about  $\eta$ . The statistic  $S_2$  is asymptotically normally distributed with mean 0 and standard deviation as follows:

$$V_2 = \left[ \hat{\sigma}_2^2 \sum_{j=1}^N (\lambda_j - \lambda_j^*)^2 \right]^{1/2}, \quad (6)$$

where  $\hat{\sigma}_2^2 = (1/N) \sum_{j=1}^N (\delta_j - \bar{\delta})^2$  and  $\bar{\delta} = (1/N) \sum_{j=1}^N \delta_j$ .

2.2. Nonparametric Comparison for Panel Count Data

2.2.1. *Nonparametric Comparison Procedure I.* Suppose a recurrent event study with  $N$  independent individuals produce only panel count data. Let  $0 < x_{j,1} < x_{j,2} < \dots < x_{j,m_j}$  be the observation times for the  $j$ th individual. As in Section 2.1.2, let  $\delta_j$  be the treatment indicator of  $j$ th individual, and  $\lambda_{j,k} = \Lambda_j(x_{j,k})$ ,  $k = 1, \dots, m_j$ ,  $j = 1, \dots, N$  be the observed value of  $\Lambda_j(x)$  at time  $x_{j,k}$ . Assume that all individuals are drawn from two populations or receive one of two different treatments, and the purpose is to determine whether there is a treatment difference based on the observed panel count data. Let  $\mu_A(x)$  and  $\mu_B(x)$  represent the MF of  $\Lambda_j(x)$ 's corresponding to the individuals receive treatments A and B, respectively. Therefore, the null hypothesis can be formulated as  $H_0: \mu_A(x) = \mu_B(x) = \mu(x) \forall x$ . Then, to test  $H_0$ , a log-rank type test statistic is given by the following equation [11]:

$$S_3 = \sum_{j=1}^N \delta_j \sum_{k=1}^{m_j} (\lambda_{j,k} - \lambda_{j,k}^*), \quad (7)$$

where  $\lambda_{j,k}^*$  is IRE of  $\mu(x)$  under the null hypothesis  $H_0$ . Statistic  $S_3$  is a generalization of the statistic  $S_2$  for current status data. Statistic  $N^{(-1/2)} S_3$ , asymptotically, has a normal distribution with mean 0 and variance that can be estimated by the following equation:

$$\hat{\sigma}_3^2 = \frac{1}{N} \sum_{j=1}^N \left[ (\delta_j - \bar{\delta}) \sum_{k=1}^{m_j} (\lambda_{j,k} - \lambda_{j,k}^*) \right]^2. \quad (8)$$

2.2.2. Nonparametric Comparison Procedure II.

$$S_4 = \frac{1}{\sqrt{N}} \sum_{j=1}^N \delta_j \left\{ \sum_{k=1}^{m_j-1} \hat{\mu}(x_{j,k}) \left[ \frac{\Delta\Lambda_j(x_{j,k+1})}{\Delta\hat{\mu}(x_{j,k+1})} - \frac{\Delta\Lambda_j(x_{j,k})}{\Delta\hat{\mu}(x_{j,k})} \right] + \hat{\mu}(x_{j,m_j}) \left[ 1 - \frac{\Delta\Lambda_j(x_{j,m_j})}{\Delta\hat{\mu}(x_{j,m_j})} \right] \right\}, \quad (9)$$

where  $\Delta\mu(x_{j,k}) = \mu(x_{j,k}) - \mu(x_{j,k-1})$ ,  $\Delta\Lambda(x_{j,k}) = \Lambda(x_{j,k}) - \Lambda(x_{j,k-1})$ , and  $\hat{\mu}(x)$  is the NPML of the common MF  $\mu(x)$  under the null hypothesis  $H_0$ . Statistic  $S_4$  in equation (9) has

$$\hat{\sigma}_4^2 = \frac{1}{N} \sum_{j=1}^N \left\{ (\delta_j - \bar{\delta}) \left[ \sum_{k=1}^{m_j-1} \hat{\mu}(x_{j,k}) \left( \frac{\Delta\Lambda_j(x_{j,k+1})}{\Delta\hat{\mu}(x_{j,k+1})} - \frac{\Delta\Lambda_j(x_{j,k})}{\Delta\hat{\mu}(x_{j,k})} \right) + \hat{\mu}(x_{j,m_j}) \left( 1 - \frac{\Delta\Lambda_j(x_{j,m_j})}{\Delta\hat{\mu}(x_{j,m_j})} \right) \right] \right\}^2. \quad (10)$$

It may be appropriate to include the statistics mentioned above in a class of tests that takes the linear form as follows:

$$S = \sum_{j=1}^N \delta_j \varphi_j, \quad (11)$$

where  $\delta_j$  is the treatment indicator and  $\varphi_j$  is a score function which varies according to the test.

3. Saddlepoint P Value Procedures

In this section, the saddlepoint approximation method is applied to approximate the exact  $p$  value of the class of nonparametric tests for panel count and current status data (11) under TBD.

$$\hat{P}(s_0) = 1 - \sum_{i=n}^{N-1} \binom{i-1}{i-n} 2^{-i} \left[ P \left( \sum_{l=1}^{i-1} \varphi_l z_l \leq s_0 - \varphi_i \mid \sum_{l=1}^{i-1} z_l = n-1 \right) + P \left( \sum_{l=1}^{i-1} \varphi_l z_l \leq s_0 - \sum_{l=i+1}^N \varphi_l \mid \sum_{l=1}^{i-1} z_l = i-n \right) \right]. \quad (13)$$

The conditional probabilities in equation (13) can be approximated using Skovgaard's saddlepoint approximation [27] as follows:

$$P \left( \sum_{l=1}^i \varphi_l z_l \leq h \mid \sum_{l=1}^i z_l = n \right) = \Phi(\hat{\gamma}) + \phi(\hat{\gamma}) \left( \frac{1}{\hat{\gamma}} - \frac{1}{\hat{\beta}} \right), \quad (14)$$

where  $\Phi(\cdot)$  and  $\phi(\cdot)$  are the standard normal distribution function and density function, respectively,

Balakrishnan and Zhao [14] developed test procedures for the null hypothesis described in Section 2.2.1 by replacing the IRE with the nonparametric maximum likelihood estimator, NPML. The statistic of Balakrishnan and Zhao [14] is given by the following equation:

an asymptotic normal distribution with mean 0 and estimated variance as follows:

Let  $\delta_1, \delta_2, \dots, \delta_N$  be the random treatment indicators of TBD, where  $\delta_j = 0$  or 1 if control or treatment are assigned, respectively,  $j = 1, \dots, N$ . Blackwell and Hodges [55] defined the stopping rule of the TBD as follows:

$$\chi = \min \left\{ l: \sum_{i=1}^l z_i \geq n \text{ or } \sum_{i=1}^l (1 - z_i) \geq n \right\}, \quad (12)$$

where  $z_i, i = 1, 2, \dots$  are independent Bernoulli trials until one treatment is assigned to  $n = N/2$  patients.

Let  $s_0$  be the observed value of statistic  $S$  in equation (11) and  $P(s_0) = P[S \geq s_0]$  is the  $p$  value of the  $S$ , then the saddlepoint approximation of  $P[S \geq s_0]$  under TBD is given by the following equation [52]:

$$\hat{\gamma} = \text{sgn}(\hat{\alpha}) \sqrt{2 \left[ \{C(\hat{\theta}_0, 0) - \hat{\theta}_0 n\} - \{C(\hat{\theta}, \hat{\alpha}) - \hat{\theta} n - \hat{\alpha} h\} \right]},$$

$$\hat{\beta} = \hat{\alpha} \sqrt{\frac{|C''(\hat{\theta}, \hat{\alpha})|}{C''_{\theta}(\hat{\theta}_0, 0)}}, \quad \hat{\alpha} \neq 0, \quad (15)$$

and  $C(\theta, \alpha) = \sum_{i=1}^i [\log(1/2) e^{\theta + \alpha \varphi_i} + (1/2)]$  is the cumulant generating function of  $(\sum_{l=1}^i \varphi_l z_l, \sum_{l=1}^i z_l)$ . The saddlepoints

TABLE 1:  $P$  values for tumorigenicity and bladder tumor data.

Data set	NAM	SPAM	Exact
Data set 1	0.001443	0.002400	0.002825
Data set 2	0.000084	0.000041	0.000058
Data set 3	< 0.0001	0.030772	0.035047
Data set 4	< 0.0001	0.016159	0.024969

$(\hat{\theta}, \hat{\alpha})$  and  $\hat{\theta}_0$  solve the system of equations  $C'(\hat{\theta}, \hat{\alpha}) = (n, h)$  and  $C'_{\theta}(\hat{\theta}_0, 0) = n$ , respectively.

#### 4. Illustrative Examples and Some Numerical Results

This section is dedicated to clarifying the proposed procedures for approximating the exact  $p$  value of the class of nonparametric tests for panel count and current status data under TBD that were discussed in the previous sections. Firstly, two sets of real current status and panel count data are analyzed to illustrate the proposed procedures. After that, the accuracy of the proposed method, saddlepoint approximation, is compared with the traditional method, normal approximation, by carrying out a simulation study. Two nonparametric procedures were presented in Section 2 for current status data and another two for panel count data. In this section, one statistic is selected for each data type and applied to it, and the application to other statistics will be the same. These statistics are the statistic in equation (4) for the current status data and the statistic in equation (7) for the panel count data.

*4.1. Illustrative Examples.* We illustrate the proposed procedures for approximating the exact  $p$  value discussed in the previous section using two sets of real current status and panel count data. The first data set arising from a tumorigenicity study is taken from Ii et al. [56]. The data are the number of tumors detected at the time of death for two groups of male ( $\delta_j = 1$ ) and female ( $\delta_j = 0$ ) mice. The goal of the study was to compare tumor development rates in male and female groups. Two random sample of size  $N = 32$  and 100 are selected from the data of Ii et al. [56] as examples of small and large sized samples. In Table 1, the results for these data sets are referred as data set 1 and data set 2, respectively. In the table, the approximated  $p$  values are calculated using the normal approximation method, NAM, and the saddlepoint approximation method, SPAM. Moreover, the table contains the simulated exact mid- $p$ -value, Exact. The simulated exact mid- $p$ -values are computed by simulating million truncated binomial sequences for the control and treatment indicators and then calculating the ratio of times that  $S$  is greater than  $s_0$  plus half the ratio of times it reaches  $s_0$ .

The second illustration is for panel count data collected from a study of superficial bladder tumors [8, 57]. All tumors detected at the start of the study were removed and the patients were randomly divided into a treatment or control group. In follow-up visits, patients were examined, the new tumors were removed, and treatment continued. Two

portion of this data set as small ( $N = 30$ ) and intermediate ( $N = 72$ ) size data sets are used to illustrate the proposed procedures. In Table 1, the results for these data sets are referred as data set 3 and data set 4, respectively. The results of Table 1 show a clear superiority of the saddlepoint approximation method over the normal approximation method to approximate the exact  $p$  value of the tests.

*4.2. Some Numerical Results.* A simulation study is carried out under various situations to compare the accuracy of the proposed method and the commonly used approximation method. For the current status data, we generate  $\lambda_j = \Lambda_j(x_j)$  from Poisson ( $\mu_j(x_j)e^{\eta\delta_j}$ ) processes, where  $\mu(x) = \sigma x$ ,  $\sigma$  is constant and  $x_j = j/N$ ,  $j = 1, 2, \dots, N$ . For each sample size  $N = 2n = 30, 50, 70, 90$ , one thousand data sets are generated from this Poisson process with  $\sigma = 6, 10, 12$ . For each data set, the treatment labels,  $\delta_j$ , are chosen by simulating a TBD sequence.

In the same way, one thousand panel count data sets of sizes,  $N = 2n = 30, 50, 70, 90$ , are generated by considering a follow-up study in which patients are followed-up quarterly. The number of visits for each patient is distributed uniformly over the 3 visits. For the number of observations,  $m_j$ , we generate  $m_j$ 's from a discrete uniform distribution  $U\{1, 2, 3\}$ , which are integers selected to give specific percentages of missing observations. Let the number of tumors discovered at each follow-up visit follow a Poisson process with MF  $\mu(x_{j,k})\exp\{\eta\delta_j\}$ , where  $\mu(x_{j,k}) = \sigma x_{j,k}$  and  $x_{j,k} = k/3$ ,  $k = 1, 2, 3$ ,  $j = 1, 2, \dots, N$ . For the purpose of comparing the accuracy of the saddlepoint approximation and normal approximation methods, the simulated exact mid- $p$ -value, normal approximation  $p$  value, and saddlepoint approximation  $p$  value are calculated for each simulated data set. After that, the percentage of closeness of the saddlepoint approximation method to the simulated method is calculated as one of the comparison criteria between the two methods. This criterion is included in the Tables 2 and 3 under the heading "Sad.Perc." In addition, two other criteria are used to compare the two approximation methods, namely the mean square error, Mse, and mean relative absolute error, Rae. We can define the Mse of the saddlepoint approximation  $p$  values ( $\hat{P}$ ) and normal approximation  $p$  values ( $\tilde{P}$ ) as  $\text{Mse.Sad} = (1/M) \sum_{i=1}^M (P_i - \hat{P}_i)^2$  and  $\text{Mse.Nor} = (1/M) \sum_{i=1}^M (P_i - \tilde{P}_i)^2$ , respectively, where  $M = 1,000$  and  $P_i$  is the simulated exact mid- $p$ -value. Also, the Rae of the saddlepoint approximation  $p$  values and normal approximation  $p$  values can be defined as  $\text{Rae.Sad} = (1/M) (\sum_{i=1}^M |P_i - \hat{P}_i|/P_i)$ , and  $\text{Rae.Nor} = (1/M) (\sum_{i=1}^M |P_i - \tilde{P}_i|/P_i)$ , respectively. The results of the simulation study are presented in Tables 2 and 3.

TABLE 2: Numerical results for the current status data.

$N$	$\eta$	$\sigma$	Sad.Perc	Rae.Sad	Rae.Nor	Mse.Sad	Mse.Nor
30	-1.2	6	84.9	0.00177	0.00355	$2.34192 \times 10^{-5}$	$4.23413 \times 10^{-5}$
		10	91.2	0.00123	0.00217	$5.77630 \times 10^{-6}$	$1.82808 \times 10^{-5}$
		12	90.5	0.00105	0.00175	$1.82512 \times 10^{-5}$	$1.65949 \times 10^{-5}$
50	-0.9	6	90.0	0.00441	0.04163	$4.16125 \times 10^{-5}$	$5.84732 \times 10^{-3}$
		10	92.2	0.00033	0.00083	$1.90184 \times 10^{-6}$	$4.84102 \times 10^{-6}$
		12	92.8	0.00026	0.00074	$1.28425 \times 10^{-6}$	$7.12743 \times 10^{-6}$
70	-0.6	6	75.3	0.00191	0.00263	$1.31064 \times 10^{-5}$	$2.52658 \times 10^{-5}$
		10	87.3	0.00066	0.00125	$3.26490 \times 10^{-6}$	$1.00190 \times 10^{-5}$
		12	90.5	0.00045	0.00092	$2.26435 \times 10^{-6}$	$5.71027 \times 10^{-6}$
90	-0.4	6	72.1	0.00188	0.00268	$7.80175 \times 10^{-6}$	$1.75748 \times 10^{-5}$
		10	77.4	0.00126	0.00202	$4.82269 \times 10^{-6}$	$1.31863 \times 10^{-5}$
		12	82.4	0.00103	0.00176	$3.94121 \times 10^{-6}$	$1.17254 \times 10^{-5}$

TABLE 3: Numerical results for the panel count data.

$N$	$\eta$	$\sigma$	Sad.Perc	Rae.Sad	Rae.Nor	Mse.Sad	Mse.Nor
30	-1.2	6	90.3	0.01560	0.09170	0.00038	0.01517
		10	90.2	0.01828	0.09467	0.00046	0.01363
		12	91.0	0.02046	0.10060	0.00054	0.01419
50	-0.9	6	90.0	0.00441	0.04163	$4.16125 \times 10^{-5}$	$5.84241 \times 10^{-3}$
		10	90.9	0.00620	0.05282	$6.59715 \times 10^{-5}$	$7.80276 \times 10^{-3}$
		12	93.7	0.00639	0.05189	$6.81304 \times 10^{-5}$	$6.77170 \times 10^{-3}$
70	-0.6	6	90.3	0.00195	0.02483	$1.20194 \times 10^{-5}$	$3.3617 \times 10^{-3}$
		10	90.4	0.00331	0.03128	$2.32985 \times 10^{-5}$	$4.03523 \times 10^{-3}$
		12	93.0	0.00399	0.03632	$3.08520 \times 10^{-5}$	$4.76476 \times 10^{-3}$
90	-0.4	6	88.1	0.00071	0.01382	$2.18075 \times 10^{-5}$	$1.41973 \times 10^{-3}$
		10	90.8	0.00138	0.02135	$5.02310 \times 10^{-6}$	$2.58790 \times 10^{-3}$
		12	91.2	0.00178	0.02652	$7.19949 \times 10^{-6}$	$3.73650 \times 10^{-3}$

The results of Tables 2 and 3 show the superiority of the saddlepoint approximation method over the normal approximation method in approximating the exact  $p$ -value of the class of nonparametric tests for panel count and current status data under TBD in terms of Mse and Rae. For example in the case of sample size  $N = 50$  with  $\sigma = 12$  of Table 3, the saddlepoint  $p$ -values were closer to the simulated exact mid- $p$ -values 93.7% of the time and the mean relative absolute error for the saddlepoint approximation method was 0.639% versus 5.189% for the normal approximation method. When averaged over all simulations, the saddlepoint approximation  $p$ -values were closer to 85.5% and 90.8% of the time to the simulate exact  $p$  values for current status and panel count, respectively.

## 5. Conclusion

In this paper, the saddlepoint approximation method is proposed as a more accurate alternative than normal approximation method to approximate the exact  $p$  value of the class of nonparametric tests for panel count and current status data under TBD. Such data appears frequently in clinical and tumorigenicity studies. A TBD was applied to ensure perfect equilibrium between the two treatment groups, which is one of the requirements of many clinical studies. The accuracy of the proposed method was verified by analyzing a number of real data sets and carrying out a comprehensive simulation study.

## Data Availability

The data sets generated and analyzed in this study are simulated data created by using R program.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## 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 R. G. P. 2/144/43.

## References

- [1] M. H. Gail, "Evaluating serial cancer marker studies in patients at risk of recurrent disease," *Biometrics*, vol. 37, no. 1, pp. 67–78, 1981.
- [2] M. S. Pepe and J. Cai, "Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates," *Journal of the American Statistical Association*, vol. 88, no. 423, pp. 811–820, 1993.
- [3] R. J. Cook and J. F. Lawless, *The Statistical Analysis of Recurrent Events*, Springer, New York, NY, USA, 2007.

- [4] J. D. Kalbfleisch and J. F. Lawless, "The analysis of panel data under a Markov assumption," *Journal of the American Statistical Association*, vol. 80, no. 392, pp. 863–871, 1985.
- [5] P. F. Thall and J. M. Lachin, "Analysis of recurrent events: nonparametric methods for random-interval count data," *Journal of the American Statistical Association*, vol. 83, no. 402, pp. 339–347, 1988.
- [6] J. Sun and J. Kalbfleisch, "Estimation of the mean function of point processes based on panel count data," *Statistica Sinica*, vol. 5, no. 1, pp. 279–289, 1995.
- [7] J. Sun and J. D. Kalbfleisch, "The analysis of current status data on point processes," *Journal of the American Statistical Association*, vol. 88, no. 424, pp. 1449–1454, 1993.
- [8] D. 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*, pp. 363–370, Springer, New York, NY, USA, 1980.
- [9] L.-J. Wei, D. Y. Lin, and L. Weissfeld, "Regression analysis of multivariate incomplete failure time data by modeling marginal distributions," *Journal of the American Statistical Association*, vol. 84, no. 408, pp. 1065–1073, 1989.
- [10] J. A. Wellner and Y. Zhang, "Two estimators of the mean of a counting process with panel count data," *Annals of Statistics*, vol. 28, no. 3, pp. 779–814, 2000.
- [11] J. Sun and H.-B. Fang, "A nonparametric test for panel count data," *Biometrika*, vol. 90, no. 1, pp. 199–208, 2003.
- [12] Y. Zhang, "Nonparametric k-sample tests with panel count data," *Biometrika*, vol. 93, no. 4, pp. 777–790, 2006.
- [13] D.-H. Park, J. Sun, and X. Zhao, "A class of two-sample nonparametric tests for panel count data," *Communications in Statistics - Theory and Methods*, vol. 36, no. 8, pp. 1611–1625, 2007.
- [14] N. Balakrishnan and X. Zhao, "A nonparametric test for the equality of counting processes with panel count data," *Computational Statistics & Data Analysis*, vol. 54, no. 1, pp. 135–142, 2010.
- [15] I. D. Diamond, J. W. McDonald, and I. H. Shah, "Proportional hazards models for current status data: application to the study of differentials in age at weaning in Pakistan," *Demography*, vol. 23, no. 4, pp. 607–620, 1986.
- [16] I. Diamond and J. McDonald, "The analysis of current status data," in *Demographic Applications of Event History Analysis*, J. Trussell, R. Hankinson, and J. Tilton, Eds., Oxford University Press, Oxford, UK, 1991.
- [17] G. E. Dinse and S. W. Lagakos, "Regression analysis of tumour prevalence data," *Applied Statistics*, vol. 32, no. 3, pp. 236–248, 1983.
- [18] D. M. Finkelstein, "A proportional hazards model for interval-censored failure time data," *Biometrics*, vol. 42, no. 4, pp. 845–854, 1986.
- [19] J. Huang, "Efficient estimation for the proportional hazards model with interval censoring," *Annals of Statistics*, vol. 24, no. 2, pp. 540–568, 1996.
- [20] J. Sun and J. D. Kalbfleisch, "Nonparametric tests of tumor prevalence data," *Biometrics*, vol. 52, no. 2, pp. 726–731, 1996.
- [21] J. Sun, "A nonparametric test for current status data with unequal censoring," *Journal of the Royal Statistical Society: Series B*, vol. 61, no. 1, pp. 243–250, 1999.
- [22] J. N. Matthews, *Introduction to Randomized Controlled Clinical Trials*, Chapman and Hall/CRC, New York, NY, USA, 2006.
- [23] A. L. Rukhin, "Stopping rules in balanced allocation problems: exact and asymptotic distributions," *Sequential Analysis*, vol. 27, no. 3, pp. 277–292, 2008.
- [24] E. F. Abd-Elfattah, "Saddlepoint p-values and confidence intervals for a class of two sample permutation tests for current status and panel count data," *Lifetime Data Analysis*, vol. 17, no. 3, pp. 461–472, 2011.
- [25] H. E. Daniels, "Saddlepoint approximations in statistics," *The Annals of Mathematical Statistics*, vol. 25, no. 4, pp. 631–650, 1954.
- [26] R. Lugannani and S. Rice, "Saddlepoint approximation for the distribution of the sum of independent random variables," *Advances in Applied Probability*, vol. 12, no. 2, pp. 475–490, 1980.
- [27] I. M. Skovgaard, "Saddlepoint expansions for conditional distributions," *Journal of Applied Probability*, vol. 24, no. 4, pp. 875–887, 1987.
- [28] S. Wang, "Saddlepoint approximations for bivariate distributions," *Journal of Applied Probability*, vol. 27, no. 3, pp. 586–597, 1990.
- [29] J. E. Kolassa, "Multivariate saddlepoint tail probability approximations," *Annals of Statistics*, vol. 31, no. 1, pp. 274–286, 2003.
- [30] H. E. Daniels, "Exact saddlepoint approximations," *Biometrika*, vol. 67, no. 1, pp. 59–63, 1980.
- [31] J. Robinson, "Saddlepoint approximations for permutation tests and confidence intervals," *Journal of the Royal Statistical Society: Series B*, vol. 44, no. 1, pp. 91–101, 1982.
- [32] H. E. Daniels, "Saddlepoint approximations for estimating equations," *Biometrika*, vol. 70, no. 1, pp. 89–96, 1983.
- [33] H. E. Daniels, "Tail probability approximations," *International Statistical Review/Revue Internationale de Statistique*, vol. 55, pp. 37–48, 1987.
- [34] A. C. Davison and D. V. Hinkley, "Saddlepoint approximations in resampling methods," *Biometrika*, vol. 75, no. 3, pp. 417–431, 1988.
- [35] H. E. Daniels, "A class of nearly exact saddlepoint approximations," in *The Art of Statistical Science*, K. Mardia, Ed., Wiley, New York, NY, USA, 1991.
- [36] J. G. Booth and R. W. Butler, "Randomization distributions and saddlepoint approximations in generalized linear models," *Biometrika*, vol. 77, no. 4, pp. 787–796, 1990.
- [37] R. Gatto and S. R. Jammalamadaka, "A saddlepoint approximation for testing exponentiality against some increasing failure rate alternatives," *Statistics & Probability Letters*, vol. 58, no. 1, pp. 71–81, 2002.
- [38] S. Huzurbazar and A. V. Huzurbazar, "Survival and hazard functions for progressive diseases using saddlepoint approximations," *Biometrics*, vol. 55, no. 1, pp. 198–203, 1999.
- [39] J. E. Kolassa, "Infinite parameter estimates in logistic regression, with application to approximate conditional inference," *Scandinavian Journal of Statistics*, vol. 24, no. 4, pp. 523–530, 1997.
- [40] R. H. Spady, "Saddlepoint approximations for regression models," *Biometrika*, vol. 78, no. 4, pp. 879–889, 1991.
- [41] X. Du and A. Sudjianto, "First order saddlepoint approximation for reliability analysis," *AIAA Journal*, vol. 42, no. 6, pp. 1199–1207, 2004.
- [42] K.-V. Yuen, J. Wang, and S.-K. Au, "Application of saddlepoint approximation in reliability analysis of dynamic systems," *Earthquake Engineering and Engineering Vibration*, vol. 6, no. 4, pp. 391–400, 2007.
- [43] C. Field, J. Robinson, and E. Ronchetti, "Saddlepoint approximations for multivariate m-estimates with applications

- to bootstrap accuracy,” *Annals of the Institute of Statistical Mathematics*, vol. 60, no. 1, pp. 205–224, 2008.
- [44] S. Wang, “General saddlepoint approximations in the bootstrap,” *Statistics & Probability Letters*, vol. 13, no. 1, pp. 61–66, 1992.
- [45] E. F. Abd-Elfattah and R. W. Butler, “The weighted log-rank class of permutation tests: P-values and confidence intervals using saddlepoint methods,” *Biometrika*, vol. 94, no. 3, pp. 543–551, 2007.
- [46] E. F. Abd-Elfattah and R. W. Butler, “Log-rank permutation tests for trend: saddlepoint p-values and survival rate confidence intervals,” *Canadian Journal of Statistics*, vol. 37, no. 1, pp. 5–16, 2009.
- [47] A. M. Abd El-Raheem and E. F. Abd-Elfattah, “Weighted log-rank tests for clustered censored data: saddlepoint p-values and confidence intervals,” *Statistical Methods in Medical Research*, vol. 29, no. 9, pp. 2629–2636, 2020.
- [48] A. M. Abd El-Raheem and E. F. Abd-Elfattah, “Log-rank tests for censored clustered data under generalized randomized block design: saddlepoint approximation,” *Journal of Biopharmaceutical Statistics*, vol. 31, no. 3, pp. 352–361, 2021.
- [49] K. S. Kamal, A. M. Abd El-Raheem, and E. F. Abd-Elfattah, “Weighted log-rank tests for left-truncated data: saddlepoint p-values and confidence intervals,” *Communications in Statistics - Theory and Methods*, 2021.
- [50] K. S. Kamal, A. M. Abd El-Raheem, and E. F. Abd-Elfattah, “Weighted log-rank tests for left-truncated data under weis urn design: saddlepoint p-values and confidence intervals,” *Journal of Biopharmaceutical Statistics*, vol. 32, no. 5, pp. 641–651, 2023.
- [51] R. W. Butler, *Saddlepoint Approximations with Applications*, Cambridge University Press, Cambridge, UK, 2007.
- [52] E. F. Abd-Elfattah, “The weighted log-rank class under truncated binomial design: saddlepoint p-values and confidence intervals,” *Lifetime Data Analysis*, vol. 18, no. 2, pp. 247–259, 2012.
- [53] R. E. Barlow, D. J. Bartholomew, J. M. Bremner, and H. D. Brunk, *Statistical Inference under Order Restrictions*, John Wiley, New York, NY, USA, 1972.
- [54] D. G. Hoel and H. E. Walburg, “Statistical analysis of survival experiments,” *Journal of the National Cancer Institute*, vol. 49, no. 2, pp. 361–372, 1972.
- [55] D. Blackwell and J. L. Hodges, “Design for the control of selection bias,” *The Annals of Mathematical Statistics*, vol. 28, no. 2, pp. 449–460, 1957.
- [56] Y. Ii, R. Kikuchi, and K. Matsuoka, “Two-dimensional (time and multiplicity) statistical analysis of multiple tumors,” *Mathematical Biosciences*, vol. 84, no. 1, pp. 1–21, 1987.
- [57] C. S. Davis and L. J. Wei, “Nonparametric methods for analyzing incomplete nondecreasing repeated measurements,” *Biometrics*, vol. 44, no. 4, pp. 1005–1018, 1988.