Original Research Article

TWO SAMPLE TEST FOR CENSORED DATA BASED ON SUB-SAMPLE MEDIANS

ABSTRACT

In hypothesis testing, the purpose of two-sample problem is to determine the statistical significance of the difference between two independent populations. This paper proposes a class of nonparametric test statistic based on U-statistic, useful for two-sample scale problem when dealing with randomly right censored data. Proposed test concentrates on comparing the medians of sub-samples from two populations. A simulation study is given for critical values and power of the proposed test statistic at various sample sizes and considering different lifetime distributions. Asymptotic relative efficiencies show that the proposed test is more efficient in comparison to some existing tests available in literature. A real-life data application is also given.

Keywords: Two-sample scale, relative efficiency, U-statistic, random right censoring.

1. INTRODUCTION

In survival analysis, the researchers mainly focus on the time-lapsed to the outcome of a certain event. The subjects of study might be censored at distinct time periods. Statistical techniques for evaluation of treatments based on survival data often involve convolutions due to incompleteness of the available data as a result of dropout or loss to follow-up. For example, in clinical trials, researchers usually came across the situations that require testing for homogeneity of survival distributions across study of some trial. This problem attracted substantial attention, and a number of procedures have already been developed, such as Wilcoxon [1] and Mann-Whitney [2] tests, which have been extended by Gehan [3], Efron [4], and Ayushee et al. [5] and others, to accommodate censored survival data.

In this paper, we propose a new nonparametric method using medians of sub-samples for testing the equality of survival distributions from randomly right-censored data. Several procedures for testing the equality of medians have been proposed for two-sample problem. The median test is a well-known and useful non-parametric test procedure, for comparative studies between two treatments ([6]).

In case of uncensored data median test was modified by many authors, such as Mathisen HC. [7], Chatterjee and Sen [8], Hettmansperger [9], Kumar [10], Kumar et al. [11]. Brookmeyer and Crowley [12] introduced a method for comparing medians of several survival distributions for right censored data. Some other authors, such as Prentice [13], Gastwirth and Wang [14], Park and Desu [15] and Park [16] considered extensions of the median test in the presence of censored data.

Testing for similitude of survival distributions is not equivalent to testing for equality of survival medians, because homogeneity of survival distributions of two groups implies that they do not differ in their median survival times. But the converse is not true, i.e., if median survival times are same in both the groups, then one cannot conclude that their survival distributions will also be same. Such situation generally arises when the survival curves cross each other.

A class of nonparametric tests is proposed in this paper for the two-sample scale problem, where the data is randomly censored. By two-sample scale problem, we mean to test whether there is any difference in the scale parameters of two populations or not. Suppose that two samples X and Y with $n_1,\ n_2$ independent observations are drawn from two independent populations having cumulative distribution functions $F_1(x)$ and $F_2(x)$ respectively. It is assumed that these two populations differ from each other only in terms of their scale parameters. Statistically, our problem is to test the null hypothesis:

$$H_0 : F_1(t) = F_2(t),$$

against the alternative

$$H_1 : F_1(t) = F_2(\theta t), \quad (\theta \neq 1).$$

Let X_i, be the observed survival time for i^{th} individual $(i = 1, 2, \ldots, n_1)$, when observations are subject to randomly right censorship and the period of follow-up for the i^{th} individual is restricted by the censoring time T_i. The observed survival time is $X_i = \min(X_i^0, T_i)$, where X_i^0 is the true but often unobserved survival time. We can observe randomly censored sample $(X_1, \delta_1), \ldots, (X_{n_1}, \delta_{n_1})$, where δ_i is an indicator function which indicates whether X_i is censored or not; thus, if $X_i < X_i^0$ the observation is said to be censored and we set $\delta_i = 0$. On the other hand, if $X_i = X_i^0$ it is an observed death and we set $\delta_i = 1$. Similarly, let Y_j, be the observed survival time for j^{th} individual $(j = 1, 2, \ldots, n_2)$ from sample Y and the observations are randomly censored with period of follow-up is restricted by the censoring time V_j. Then we can observe randomly censored sample $(Y_1, \varepsilon_1), \ldots, (Y_{n_2}, \varepsilon_{n_2})$, where $Y_j = \min\left(Y_j^0, V_j\right)$ and Y_j^0 is the true survival time associated with indicator function ε_j which indicates whether Y_j is censored or not; for values $\varepsilon_j = 0$ and $\varepsilon_j = 1$ respectively. The objective of the proposal of this test is to have more efficiency with respect to the existing two-sample scale problems.

The newly proposed test statistic is defined in Section 2. Section 3 gives the evaluated expression for mean and variance of the test statistic. Section 4 includes the tables for Critical points of the proposed test statistic at different sample sizes and percentage censorings. Section 5 deals with asymptotic relative efficiency comparisons. Further, an illustration based on a real-life data for the proposed statistic is given in Section 6. In Section 7, Statistical power of the proposed test is given at various sample sizes and percentage censoring.

2. THE PROPOSED TEST STATISTIC

The proposed test statistics is based on the symmetrical kernel $\mathrm{U_{ij}}$, which is defined as:

$$U_{ij} = \begin{cases} 1, \ \text{if } M_d(X_{i_1}, X_{i_2}, X_{i_3}) > M_d(Y_{j_1}, Y_{j_2}, Y_{j_3}) & \text{with } \delta_i = \varepsilon_j = 1 \ \forall \ (i, j) \\ \text{or } M_d(X_{i_1}, X_{i_2}, X_{i_3}) > M_d(Y_{j_1}, Y_{j_2}, Y_{j_3}) & \text{for } \max(X_{i_1}, X_{i_2}, X_{i_3}) \text{ associated } \delta = 0 \\ & \text{and for others, } \varepsilon_j = \delta_i = 1 \\ -1, \ \text{if } M_d(X_{i_1}, X_{i_2}, X_{i_3}) < M_d(Y_{j_1}, Y_{j_2}, Y_{j_3}) & \text{with } \delta_i = \varepsilon_j = 1 \ \forall \ (i, j) \\ \text{or } M_d(X_{i_1}, X_{i_2}, X_{i_3}) < M_d(Y_{j_1}, Y_{j_2}, Y_{j_3}) & \text{for } \max(Y_{j_1}, Y_{j_2}, Y_{j_3}) \text{ associated } \varepsilon = 0 \\ & \text{and for others, } \varepsilon_j = \delta_i = 1 \\ 0, \ \text{elsewhere} \end{cases} \quad (1)$$

where, $M_d(X_{i_1}, X_{i_2}, X_{i_3})$ and $M_d(Y_{j_1}, Y_{j_2}, Y_{j_3})$ are representing the medians of the sub-samples $(X_{i_1}, X_{i_2}, X_{i_3})$ and $(Y_{j_1}, Y_{j_2}, Y_{j_3})$ of size three each, taken from random samples X and Y respectively.

Here, we have compared the medians of sub-samples from each sample to grab more information from the samples. When uncensored sub-samples are chosen to be compared and the median of the sub-sample $(X_{i_1}, X_{i_2}, X_{i_3})$ from random sample X, find out to be greater than the median of the subsample $(Y_{j_1}, Y_{j_2}, Y_{j_3})$ from random sample Y, we assign 1 to the kernel U_{ij}, contrariwise, -1 is assigned to the kernel U_{ij}.

In case, when one observation is censored in any one of the chosen sub-samples and this censored observation comes out to be greater among all three observations, if this happens in the sub-sample chosen from X sample, the observed sub-sample would be $(X_{i_1}, X_{i_2}, X_{i_3})$ with

$$\delta = \begin{cases} 0, & \text{only for } \max(X_{i_1}, X_{i_2}, X_{i_3}) \\ 1, & \text{for other observations} \end{cases}$$

and $(Y_{j_1}, Y_{j_2}, Y_{j_3})$ with $(\varepsilon_j = 1 \ \forall \ j)$. Now, we assign 1 to the kernel U_{ij} if the median of the subsample from X sample with one censored observation, is find out to be greater than the median of the sub-sample from Y sample with all uncensored observations. Similarly, if the censored observation comes in the sub-sample from Y sample, then the observed sub-sample would be $(Y_{j_1}, Y_{j_2}, Y_{j_3})$ with

$$\varepsilon = \begin{cases} 0, & \text{only for } \max(Y_{j_1}, Y_{j_2}, Y_{j_3}) \\ 1, & \text{for other observations} \end{cases}$$

and $(X_{i_1}, X_{i_2}, X_{i_3})$ with $(\delta_i = 1 \ \forall \ i)$ and we assign -1 to the kernel U_{ij} if the median of the subsample from Y sample with one censored observation, is find out to be greater than the median of the sub-sample from X sample with all uncensored observations. We assign zero to the kernel U_{ij}, for remaining cases.

On the basis of proposed kernel U_{ij}, the test statistic is defined as:

$$W = \sum_{i,j} U_{ij}. \quad (2)$$

where the sum is extended over all n_1, n_2 combinations.

3. THE MEAN AND VARIANCE OF TEST STATISTIC

Let us suppose that both the samples have same variation, i.e., the null hypothesis, H_0 is true. We consider the conditional mean and variance of V, denoted by $E(W|P, H_0)$ and $var(W|P, H_0)$ respectively, under H_0 and P be the observational pattern, observed on rank ordering the data. The expectation has been taken over the possible number of samples $\binom{n_1 + n_2}{n_1}$ that are equally likely and follow the observational pattern P. Due to symmetry, we see:

$$E(W|P, H_0) = E\left(\sum_{i,j} U_{ij} \middle| P, H_0\right) = 0. \quad (3)$$

The variance of W under H_0 and restricted to the observational pattern P, can be defined as:

$$var(W|P, H_0) = E\left(\sum_{i,j} U_{ij} - E\left(\sum_{i,j} U_{ij}\right) \middle| P, H_0\right)^2. \quad (4)$$

Using eq. (3), we have:

$$var(W|P, H_0) = E\left(\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} U_{ij}^2 + \sum_{i\neq i'=1}^{n_1}\sum_{j}^{n_2} U_{ij}U_{i'j} + \sum_{i=1}^{n_1}\sum_{j\neq j'=1}^{n_2} U_{ij}U_{ij'} + \sum_{i\neq i'=1}^{n_1}\sum_{j\neq j'=1}^{n_2} U_{ij}U_{i'j'} \middle| P, H_0\right). \quad (5)$$

The final expression of variance of the statistic can be obtained by evaluating each term of eq. (5), as:

$$var(W|P, H_0) = \frac{144\binom{n_1 + n_2 - 6}{n_1 - 3}}{\binom{n_1 + n_2}{n_1}} K_1 + \left\{\frac{\binom{n_1 + n_2 - 9}{n_1 - 6}}{\binom{n_1 + n_2}{n_1}} + \frac{\binom{n_1 + n_2 - 9}{n_2 - 6}}{\binom{n_1 + n_2}{n_1}}\right\} K_2. \quad (6)$$

In eq. (6), the quantity K_1 is the proportion of times a specific pair of observations (i, j) to turn up in X and Y samples. Similarly, the quantity K_2 is the proportion of times a specific pair $(i, i' \text{and } i \neq i')$ to turn up in any one of the samples with observation j from the other sample. Here

$$
\begin{aligned}
K_1 = \sum_{i=1}^{s} 10 \times \Bigg[& \binom{m_i}{1}\binom{M_{i-1}}{5} + \binom{m_i}{1}\binom{n_1 + n_2 - M_i - L_{i-1}}{5} \\
& + \binom{m_i}{1}\binom{M_{i-1}}{4}\binom{n_1 + n_2 - M_i - L_{i-1}}{1} + \binom{m_i}{1}\binom{M_{i-1}}{3}\binom{n_1 + n_2 - M_i - L_{i-1}}{2} \\
& + \binom{m_i}{1}\binom{M_{i-1}}{2}\binom{n_1 + n_2 - M_i - L_{i-1}}{3} + \binom{m_i}{1}\binom{M_{i-1}}{1}\binom{n_1 + n_2 - M_i - L_{i-1}}{4} \\
& + \binom{l_i}{1}\binom{M_i}{5}\Bigg] + \binom{l_i}{2}\binom{M_i}{4}, \qquad (7)
\end{aligned}
$$

where, first and second terms in eq. (7), representing the total no. of ways of pairing any failed observation at i^{th} rank with any five observations at lesser rank and any five observations of rank greater than i respectively. The third represents the total no. of ways of pairing any failed observation at i^{th} rank with any four observations at lesser rank and one observation at rank greater than i, whereas sixth term representing the vice-versa case of third term i.e., pairing a failed observation at i^{th} rank with one observation at lesser rank and any four observations at rank greater than i. Similarly, fourth and fifth terms also reverse cases of each other. The fourth term represents the total no. of ways of pairing any failed observation at i^{th} rank with any three observations at lesser rank and any two observations at rank greater than i, whereas fifth term represents the total no. of ways of pairing a failed observation at i^{th} rank with any two observations at lesser rank and any three observations at rank greater than i. Second last term represents the no. of ways of pairing a censored observation immediately after i^{th} rank with any five observations that have failed earlier, and the last term represents the no. of ways of pairing any two censored observations immediately after i^{th} rank with any four observations that have failed earlier. Similarly,

$$
\begin{aligned}
K_2 = 9 \times \sum_{i=1}^{s} \Bigg[& 8064\binom{m_i}{1}\binom{M_{i-1}}{8} + 11160\binom{m_i}{1}\binom{n_1 + n_2 - M_i - L_{i-1}}{8} \\
& + 9936\binom{m_i}{1}\binom{M_{i-1}}{7}\binom{n_1 + n_2 - M_i - L_{i-1}}{1} \\
& + 7920\binom{m_i}{1}\binom{M_{i-1}}{6}\binom{n_1 + n_2 - M_i - L_{i-1}}{2} \\
& + 4932\binom{m_i}{1}\binom{M_{i-1}}{5}\binom{n_1 + n_2 - M_i - L_{i-1}}{3} \\
& - 5184\binom{m_i}{1}\binom{M_{i-1}}{4}\binom{n_1 + n_2 - M_i - L_{i-1}}{4} \\
& + 3960\binom{m_i}{1}\binom{M_{i-1}}{3}\binom{n_1 + n_2 - M_i - L_{i-1}}{5} \\
& + 8856\binom{m_i}{1}\binom{M_{i-1}}{2}\binom{n_1 + n_2 - M_i - L_{i-1}}{6} \\
& + 9864\binom{m_i}{1}\binom{M_{i-1}}{1}\binom{n_1 + n_2 - M_i - L_{i-1}}{7} \\
& + 7704\binom{l_i}{1}\binom{M_i}{8} + 4704\binom{l_i}{2}\binom{M_i}{7} + 1872\binom{l_i}{3}\binom{M_i}{6}\Bigg] \qquad (8)
\end{aligned}
$$

The first and second terms in eq. (8), represents the total ways of pairing any failed observation at i^{th} rank with any eight observations of lesser rank and any eight observations of rank greater than i respectively. The third and ninth term represents the reverse scenario of each other i.e., if third term gives the total no. of ways of pairing any failed observation at i^{th} rank with any seven observations at lesser rank and one observation at rank greater than i, then ninth term gives the no. of ways of pairing a failed observation at i^{th} rank with one observation at lesser rank and any seven observations at rank greater than i. In the same way, fourth and eighth terms are contrary to each other, and fifth term is contrary to seventh term. The fourth term gives the no. of ways of pairing any failed observation at i^{th} rank with any six observations at lesser rank and any two observations at rank greater than i and eighth term gives the no. of ways of pairing a failed observation at i^{th} rank with any two observations at lesser rank and any six observations at rank greater than i. The fifth term represents the total no. of ways of pairing any failed observation at i^{th} rank with any five observations at lesser rank and any three observations at rank greater than i and seventh term gives the no. of ways of pairing a failed observation at i^{th} rank with any three observations at lesser rank and any five observations at rank greater than i, and

$$M_j = \sum_{i=1}^{j} m_i, \quad M_0 = 0,$$
$$L_j = \sum_{i=1}^{j} l_i \ , \quad L_0 = 0.$$

Here, m_i's be the total no. of uncensored observations at i^{th} rank with unalike values and l_i's be the total no. of randomly right censored observations with values greater than the observational value at i^{th} rank but should be smaller than the observational value at $(i+1)^{th}$ rank, when we rank order the data.

4. CRITICAL POINTS

In testing of hypothesis, critical values are those points on the test distribution that are compared to the test statistic, to dictate whether to reject or do not reject the null hypothesis H_0. One can conclude the statistical significance and reject H_0, if the numerical value of the test statistic is found to be greater than the critical value. These values correspond to the level of significance α i.e., these values become fixed for a particular α. Considering this concept, for our proposed test statistic W critical points are found using simulation. Here, we consider three different lifetime distributions (Exponential, Lindley and Weibull) for generating time to failure observations and Exponential distribution for generating time to censored observations. Two samples, each of size n are generated from these distributions and then the standardized test statistic value is calculated using the formula:

$$Z = \frac{W - E(W|P, H_0)}{\sqrt{var(W|P, H_0)}} \tag{9}$$

where, $E(W|P, H_0)$ and $var(W|P, H_0)$ are given in eq. (3) and (6). Further, we find Z_α such that it satisfies $P(Z > Z_\alpha) = 0.025$. The critical points are found as the average of ten thousand simulated values of Z_α. Critical points are given in the Tables $1-3$ for various sample sizes $(n_1 = n_2 = n)$ and censoring percentage $(pcens)$ for each considered distribution.

Table 1. Critical points of the proposed test, when both time to failure and time to censoring distributions are Exponential

n \ pcens	0.1	0.125	0.15	0.175	0.2	0.225	0.25
10	3.43368	3.31217	2.50262	1.95716	1.74136	1.52072	1.27348
15	2.96611	2.42388	2.01962	1.64266	1.37709	1.20409	1.05835
20	2.86519	2.24159	1.89780	1.63547	1.38460	1.17276	1.08101
25	2.87919	2.22975	1.90395	1.62831	1.40741	1.24239	1.12524
30	2.98026	2.31803	1.93517	1.67972	1.47543	1.35427	1.15257

Table 2. Critical points of the proposed test, when the time to failure distribution is Lindley and time to censoring distribution is Exponential

n \ $pcens$	0.1	0.125	0.15	0.175	0.2	0.225	0.25
10	3.69569	3.11290	2.50841	2.03812	1.70947	1.43429	1.21169
15	3.00529	2.45004	1.98007	1.61619	1.41912	1.22157	1.02593
20	2.83092	2.16999	1.79993	1.52519	1.31193	1.13971	1.00252
25	2.83110	2.29323	1.84565	1.57539	1.39781	1.22420	1.04688
30	2.79975	2.31781	1.94185	1.67732	1.45333	1.25434	1.13419

Table 3. Critical points of the proposed test, when the time to failure distribution is Weibull and time to censoring distribution is Exponential

n \ pcens	0.1	0.125	0.15	0.175	0.2	0.225	0.25
10	3.60395	3.63164	2.70654	2.03883	1.91145	1.50136	1.27092
15	3.18578	2.44709	2.09012	1.66678	1.42322	1.20934	1.06403
20	2.93699	2.41518	1.98969	1.61361	1.37435	1.20519	1.04461
25	2.93609	2.31926	1.88534	1.63482	1.45403	1.28143	1.08959
30	2.78932	2.35783	2.02662	1.74267	1.54213	1.35647	1.21875

5. ASYMPTOTIC RELATIVE EFFICIENCY

In this section, we find the asymptotic relative efficiency (ARE) of the proposed test statistic W relative to Gehan's [3] test statistic G and Ayushee et al. [5] test statistic V by assuming Exponential lifetime distribution. Let us suppose that the cumulative distribution function of time to failure for random variable X is given by

$$F_1(x) = 1 - e^{-x}, \qquad (x > 0) \tag{10}$$

and the cumulative distribution function of time to failure for random variable Y is given by

$$F_2(y) = 1 - e^{-\theta y}. \qquad (y > 0) \tag{11}$$

We have considered Lehmann-type alternatives (for $\mathrm{a} > 1$),

$$\overline{G_1}(x) = \left(\overline{F_1}(x)\right)^a = e^{-ax}, \tag{12}$$

and

$$\overline{G_2}(y) = \left(\overline{F_2}(y)\right)^a = e^{-a\theta y}. \tag{13}$$

where $G_1(x)$ and $G_2(y)$ are denoting the cumulative distribution functions of time to censoring for random variables X and Y respectively.

Further, the ratio of reverse hazard rates, calculated using the time to censoring distribution and time to failure distribution, remains constant (a) over time. Mathematically, for random variable X,

$$\text{(Reverse hazard ratio)}_X = \frac{r_{g_1}(x)}{r_{f_1}(x)} = a$$

Similarly, for random variable Y,

$$\text{(Reverse hazard ratio)}_Y = \frac{r_{g_2}(y)}{r_{f_2}(y)} = a$$

where, $r_{g_1}(.)$, $r_{f_1}(.)$ are consecutively denoting the reverse hazard rates for time to censoring and time to failure for random variable X and $r_{g_2}(.)$, $r_{f_2}(.)$ are denoting the reverse hazard rates for time to censoring and time to failure for random variable Y.

Our interest is to test the hypothesis

$$H: F_1(t) = F_2(\theta t) \qquad (t \le T).$$

Thus, our null hypothesis would be, $H_0: \theta = 1$. This type of test would be relevant in the situations, when we are interested in checking whether there is any constant proportion (θ) of failure times of the patients receiving two different treatments.

We want to find the ARE of the proposed test relative to Gehan's test, under the setup that is described above and when all the individuals enter study at time zero and experiment is stopped at time T. The ARE of proposed W test relative to G test is given by,

$$ARE_{WG} = \lim_{n\to\infty} \frac{\left(\frac{\partial E(n^{-2}W)}{\partial\theta}\Big|_{\theta=1}\right)^2}{\left(nvar(n^{-2}W|H_0)\right)} \times \frac{\left(nvar(n^{-2}G|H_0)\right)}{\left(\frac{\partial E(n^{-2}G)}{\partial\theta}\Big|_{\theta=1}\right)^2}. \tag{14}$$

Using proposed kernel, we can find mean and variance of statistic W, using

$$\begin{aligned}\mathrm{E}(W) = n^2\{&Pr\left(M_d(X_{i_1}, X_{i_2}, X_{i_3}) > M_d(Y_{j_1}, Y_{j_2}, Y_{j_3}); \delta_{i_1} = \delta_{i_2} = \delta_{i_3} = \varepsilon_{j_1} = \varepsilon_{j_2} = \varepsilon_{j_3} = 1\right) \\ &+ Pr\left(Med(X_{i_1}, X_{i_2}, X_{i_3}) > Med(Y_{j_1}, Y_{j_2}, Y_{j_3}), \delta_{i_1} = \delta_{i_2} = \varepsilon_{j_1} = \varepsilon_{j_2} = \varepsilon_{j_3} = 1, \delta_{i_3} = 0\right) \\ &- Pr\left(Med(X_{i_1}, X_{i_2}, X_{i_3}) < Med(Y_{j_1}, Y_{j_2}, Y_{j_3}), \delta_{i_1} = \delta_{i_2} = \delta_{i_3} = \varepsilon_{j_1} = \varepsilon_{j_2} = 1, \varepsilon_{j_3} = 0\right) \\ &- Pr\left(M_d(X_{i_1}, X_{i_2}, X_{i_3}) < M_d(Y_{j_1}, Y_{j_2}, Y_{j_3}), \delta_{i_1} = \delta_{i_2} = \delta_{i_3} = \varepsilon_{j_1} = \varepsilon_{j_2} = \varepsilon_{j_3} = 1\right)\}\end{aligned} \tag{15}$$

when the sub-sample observations $(X_{i_1} < X_{i_2} < X_{i_3})$, $(Y_{j_1} < Y_{j_2} < Y_{j_3})$ are in increasing order of their magnitude and

$$\mathrm{v}ar(n^{-2}W|H_0) = n^{-4}E\left\{\left(\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} U_{ij} - E\left(\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} U_{ij}\middle| H_0\right)\right)\right\}^2$$

$$\Rightarrow \qquad var(n^{-2}W|H_0) = n^{-4}E\left\{\left(\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} U_{ij}\middle| H_0\right)\right\}^2. \tag{16}$$

Since, $E\left(\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}U_{ij}\right)=0$ and $E\left(\sum_{i\neq i'=1}^{n_1}\sum_{j}^{n_2}U_{ij}U_{i'j}\right)=E\left(\sum_{i=1}^{n_1}\sum_{j\neq j'=1}^{n_2}U_{ij}U_{ij'}\right)$. Also, $E\left(\sum_{i\neq i'=1}^{n_1}\sum_{j\neq j'=1}^{n_2}U_{ij}U_{i'j'}\right)=0$, since $\mathrm{U_{ij}}$ and $\mathrm{U_{i'j'}}$ are independent of each other and has expectation zero. Similarly, we can find the mean and variance of Gehan's test statistic G.

On evaluating each term of eq. (14), the final expression for ARE of proposed statistic W to Gehan's statistic G as,

$$ARE_{WG}=\frac{9(1-e^{-T})\times(A_1)\times(A_2)^2}{(-1+e^{-T})^6\times(D_1)^2\times(D_2)}. \tag{17}$$

where,

$$A_1=(-1+e^{-T})^2-\frac{6(1-e^{-T})(1-e^{-aT})}{2+3a+a^2}+\frac{3(-1+e^{-aT})^2}{1+2a}+6(1-e^{-T})(1-e^{-aT})\left(\frac{1}{2+a}+\frac{a\Gamma(a)}{\Gamma(3+a)}\right), \tag{18}$$

$$\begin{aligned}A_2=&-\frac{2089}{225}-4e^{-T}T-2e^{-2T}(1+4T)+\frac{8}{9}e^{-3T}(-4+15T)+e^{-4T}(1-4T)\\&-\frac{4}{25}e^{-5T}(-24+5T)-2e^{-6T}+4ae^{-aT}T-2e^{-2aT}-2e^{-2(2+a)T}\\&+\frac{4}{(1+a)^2}\Big(2a(1-a)-e^{(-1-a)T}(4a-3T-5aT-5a^2T-3a^3T)\\&+e^{-2(1+a)T}(6a-2T+3a^2-2aT+2)\Big)\\&+\frac{4}{(2+a)^2}\Big(a^2+2(2-a)-2e^{(-2-a)T}(4-8aT-6a^2T-a^3T)\Big)\\&-\frac{4}{(3+a)^2}\Big(3(a^2-4a+3)-2e^{(-3-a)T}(-12a+15T-13aT-3a^2T+a^3T-4)\Big)\\&+\frac{4}{(4+a)^2}\Big(3a^2+8(1-a)+e^{(-4-a)T}(32a-32T+24aT-4a^2T-3a^3T+40)\Big)\\&-\frac{4}{(5+a)^2}\Big((1-a)^2+e^{(-5-a)T}(12a+5T+9aT-3a^2T-a^3T-24)\Big)\\&+\frac{16(1+a)(2a+5)}{(2+a)^2(3+a)^2}+\frac{8}{(1+2a)^2}\Big(1+e^{(-1-2a)T}(4a-T+4a^2-2aT)\Big)\\&+\frac{8(a^2+5a+7)}{(2+a)(3+a)}+\frac{8}{(3+2a)^2}\Big(1+e^{(-3-2a)T}(12a-3T+4a^2-2aT+8)\Big),\end{aligned} \tag{19}$$

and

$$D_1=\frac{1}{2}e^{-2T}(-1+e^{2T})+\frac{a\left(1-(1+T-aT)e^{(-1-a)T}\right)}{1+a}-\frac{a(a-1)\left(1-e^{(-1-a)T}\right)}{(1+a)^2}, \tag{20}$$

$$D_2 = (1-e^{-T})^3 + (1-e^{-aT})^3 + 3e^{(-2-a)T}(-1+e^{T})^2(-1+e^{aT})$$
$$+24(1-e^{-T})(-1+e^{-aT})^2\left(\frac{1}{8} - \frac{4}{1+a} - \frac{66}{2+a} + \frac{12}{3+a} - \frac{102}{4+a} + \frac{114}{5+a} - \frac{60}{6+a}\right.$$
$$\left. + \frac{12}{7+a} + \frac{60}{3+2a} + \frac{120}{5+2a} + \frac{12}{7+2a} - \frac{72\Gamma(2+a)}{\Gamma(8+a)} + \frac{1440\Gamma(2+2a)}{\Gamma(8+2a)}\right). \quad (21)$$

In the same way, we can find the ARE of proposed test W and Statistic V proposed by Ayushee et al. [5] as,

$$ARE_{WV} = \frac{9 \times (A_3) \times (A_4)^2}{(-1+e^{-T})^2 \times (D_3)^2 \times D_4} \quad (22)$$

where,

$$A_3 = (-1+e^{-T})^2 + 6\left(\frac{11}{12} - \frac{5}{2(1+a)} + \frac{2}{3+a} + \frac{1}{1+2a}\right)(-1+e^{-aT})^2$$
$$+ 3(1-e^{-t})(1-e^{-at})\left(\frac{7}{15} - \frac{144}{(1+a)(2+a)(3+a)(4+a)(5+a)}\right.$$
$$\left. + \frac{8(a^2+9a-16)\,\Gamma(1+a)}{\Gamma(6+a)}\right), \quad (23)$$

and

$$A_4 = -\frac{289}{225} - 4e^{-T}T - 2e^{-2T}(1+4T) + \frac{8}{9}e^{-3T}(-4+15T) + e^{-4T}(1-4T)$$
$$- \frac{4}{25}e^{-5T}(-24+5T) - 2e^{-6T} + e^{-aT}T - 2e^{-2aT} - 2e^{-2(2+a)T} - \frac{4a}{2+a}$$
$$+ \frac{4(7+3a)}{3+a} - \frac{12a}{4+a} + \frac{4a}{5+a} + \frac{16(1+a)(5+2a)}{(2+a)^2(3+a)^2}$$
$$+ \frac{4}{(1+a)^2}\Big(a(a-2) + e^{(-1-a)T}(4a - 3T - 5aT - 5a^2T - 3a^3T)$$
$$- e^{-2(1+a)T}(6a - 2T + 3a^2 - 2aT + 2)\Big)$$
$$+ \frac{4}{(2+a)^2}\Big((a-2)^2 + 2e^{(-2-a)T}(8aT + 6a^2T + a^3T - 4)\Big)$$
$$- \frac{4}{(3+a)^2}\Big(3(a^2-4a+3)$$
$$- 2e^{(-3-a)T}(12a + 15T + 13aT - 3a^2T + a^3T - 4)\Big)$$
$$+ \frac{1}{(4+a)^2}\Big(4(3a^2 - 8a + 8)$$
$$+ 4e^{(-4-a)T}(32a - 32T + 24aT - 4a^2T - 3a^3T + 40)\Big)$$
$$- \frac{4}{(5+a)^2}\Big((a^2 - 2a + 2) + e^{(-5-a)T}(12a - 5T + 9aT - 3a^2T - a^3T + 24)\Big)$$
$$+ \frac{8}{(1+2a)^2}\Big(1 + e^{(-1-2a)T}(4a + T + 4a^2 - 2aT)\Big)$$
$$- \frac{8}{(3+2a)^2}\Big(1 + e^{(-3-2a)T}(12a - 3T + 4a^2 - 2aT + 8)\Big), \quad (24)$$

$$D_3 = \frac{23}{18} - 3e^{-2t} - \frac{4}{9}e^{-3T}(-5+3T) - \frac{e^{-4T}}{2} + \frac{2}{1+a} - \frac{4}{2+a} + \frac{2}{3+a}$$
$$- \frac{1}{(1+a)^2}\left(2(1-a) - e^{(-1-a)T}(2T + 4aT + 2a^2T)\right)$$
$$+ \frac{4}{(2+a)^2}\left((1-a) + 2e^{(-2-a)T}(1 + a - at - 2T)\right)$$
$$- \frac{1}{(3+a)^2}\left(2(1-a) - e^{(-3-a)T}(-2a + 6T + 8aT + 2a^2T - 3)\right), \qquad (25)$$

and

$$D_4 = (1-e^{-T})^3 + (1-e^{-aT})^3 + 3e^{(-2-a)T}(-1+e^T)^2(-1+e^{aT})$$
$$+ 12(1-e^{-t})(-1+e^{-at})^2\left(\frac{1}{4} - \frac{12}{1+a} - \frac{132}{2+a} + \frac{24}{3+a} - \frac{204}{4+a} + \frac{228}{5+a}\right.$$
$$- \frac{120}{6+a} + \frac{24}{7+a} + \frac{120}{3+2a} + \frac{240}{5+2a} + \frac{24}{7+2a} - \frac{72(a^2+13a+2)\Gamma(2+a)}{\Gamma(8+a)}$$
$$\left. + \frac{2880\,\Gamma(2+2a)}{\Gamma(8+2a)}\right). \qquad (26)$$

Using eq's (14) and (26), the $\mathrm{ARE}'s$ of the proposed test W with respect to statistic G and statistic V for different values of a and T are given in Table 4.

Table 4. ARE of the W test w.r.t. *G* (Gehan) and *V* tests for different values of a and T

a	Tests	T					
		1	**2**	**3**	**4**	**5**	$\to \infty$
1	G	43.65	53.80	47.84	43.15	40.91	39.43
	V	10.39	22.22	24.81	24.48	24.02	23.62
2	G	11.88	16.06	18.93	21.44	23.09	24.86
	V	6.20	14.58	16.19	19.68	21.79	23.84
3	G	5.89	12.75	18.13	21.39	23.21	25.01
	V	3.95	10.99	18.31	23.08	25.69	28.11
4	G	4.41	13.16	19.18	22.56	24.43	26.28
	V	3.25	12.27	20.74	26.05	28.93	31.57
5	G	4.15	13.94	20.16	23.62	25.54	27.42
	V	3.21	13.53	22.98	28.95	31.18	33.97

The ARE for our test W with respect to both G test and V test decreases with a and increases with T, except for $a = 1$. The value of ARE with respect to both the tests is always greater than 1, i.e., our test is performing better than both of the tests for all considered values of a and T.

6. REAL-LIFE EXAMPLE

A real-life example taken from Kirk et al. [17], which is based on the results of Royal Free Hospital proposed controlled trial of patients suffering from chronic active hepatitis, who have given prednisolone therapy in hepatitis B surface antigen negative chronic active hepatitis. These data give the survival times (in months) for a control group and a group treated with prednisolone, of 22 patients suffering from chronic active hepatitis.

We applied Kolmogorov-Smirnov test to check that the data follows which distribution, and found that this data set follows Exponential distribution. Now, we wish to check whether there is significant difference in the survival times of the patients of control group and the grouped treated with prednisolone therapy. The critical value of the proposed test statistic for sample size of 22 each, in case of Exponential distribution is found out to be 0.83188 using the procedure as described in Section 4. The standardized test statistic (z) for this data is 2.40679. Since the calculated test statistic value found to be greater than its critical value, thus the null hypothesis of no significant difference is rejected and it is concluded that there is a significant difference in the survival times of the patients of control group and the patients that are treated with prednisolone therapy.

7. POWER OF THE PROPOSED TEST

Statistical power of a test is defined as the probability that the test rejects the null hypothesis when it is true. Using the critical values given in Section 4, Statistical power of the proposed test has been found through Monte-Carlo simulation study. Data is simulated from three lifetime distributions viz., Exponential, Lindley and Weibull for 10,000 times, at equal sample sizes $\mathrm{n} = 10, 15, 20, 25 \text{ and } 30$ from both the samples and scale parameter of the second sample as $\theta = 2(1)4$. The statistical power of the proposed test W is given in the Tables $5 - 7$ at same sample sizes and censoring percentages that we have considered for calculating the critical points.

Table 5. Statistical power of the proposed test W, when both time to failure and time to censoring distributions are Exponential

n	θ	$pcens$						
		0.1	0.125	0.15	0.175	0.2	0.225	0.25
10	2	0.226	0.201	0.172	0.158	0.146	0.139	0.134
	3	0.301	0.239	0.226	0.213	0.199	0.194	0.185
	4	0.352	0.271	0.258	0.248	0.240	0.223	0.221
15	2	0.251	0.206	0.196	0.191	0.189	0.182	0.169
	3	0.330	0.322	0.317	0.306	0.304	0.294	0.290
	4	0.403	0.397	0.389	0.380	0.369	0.361	0.358
20	2	0.286	0.273	0.267	0.251	0.248	0.243	0.225
	3	0.458	0.442	0.440	0.438	0.437	0.412	0.403
	4	0.570	0.563	0.551	0.543	0.531	0.525	0.511
25	2	0.351	0.328	0.323	0.314	0.292	0.276	0.261
	3	0.601	0.597	0.591	0.587	0.573	0.548	0.532
	4	0.738	0.730	0.716	0.701	0.668	0.665	0.658
30	2	0.443	0.441	0.429	0.420	0.415	0.377	0.359
	3	0.743	0.732	0.731	0.724	0.705	0.685	0.657
	4	0.876	0.858	0.856	0.851	0.828	0.809	0.797

Table 6. Statistical power of the proposed test W, when the time to failure distribution is Lindley and time to censoring distribution is Exponential

n	θ	pcens						
		0.1	0.125	0.15	0.175	0.2	0.225	0.25
10	2	0.281	0.253	0.226	0.199	0.187	0.172	0.167
	3	0.334	0.308	0.279	0.254	0.248	0.239	0.231
	4	0.352	0.335	0.313	0.285	0.272	0.263	0.254
15	2	0.313	0.242	0.229	0.217	0.205	0.193	0.182
	3	0.394	0.365	0.323	0.311	0.298	0.276	0.268
	4	0.489	0.437	0.404	0.387	0.364	0.351	0.338
20	2	0.350	0.329	0.318	0.301	0.285	0.269	0.254
	3	0.482	0.469	0.451	0.438	0.429	0.413	0.396
	4	0.623	0.591	0.574	0.558	0.539	0.522	0.513
25	2	0.397	0.376	0.363	0.347	0.338	0.311	0.281
	3	0.625	0.612	0.598	0.576	0.562	0.547	0.529
	4	0.764	0.752	0.738	0.724	0.713	0.702	0.690
30	2	0.582	0.561	0.523	0.498	0.476	0.452	0.429
	3	0.824	0.810	0.792	0.775	0.743	0.712	0.689
	4	0.923	0.906	0.885	0.869	0.831	0.809	0.783

Table 7. Statistical power of the proposed test W, when the time to failure distribution is Weibull and time to censoring distribution is Exponential

n	θ	pcens						
		0.1	0.125	0.15	0.175	0.2	0.225	0.25
10	2	0.331	0.273	0.254	0.248	0.237	0.235	0.228
	3	0.408	0.340	0.334	0.332	0.325	0.313	0.293
	4	0.453	0.385	0.383	0.364	0.352	0.348	0.341
15	2	0.386	0.351	0.327	0.308	0.279	0.248	0.234
	3	0.542	0.505	0.489	0.422	0.385	0.351	0.323
	4	0.647	0.592	0.542	0.508	0.479	0.465	0.432
20	2	0.493	0.383	0.371	0.352	0.329	0.291	0.260
	3	0.683	0.596	0.541	0.497	0.463	0.437	0.389
	4	0.713	0.683	0.637	0.595	0.562	0.539	0.510
25	2	0.503	0.479	0.435	0.409	0.368	0.327	0.297
	3	0.741	0.699	0.658	0.615	0.587	0.539	0.424
	4	0.847	0.804	0.775	0.741	0.699	0.675	0.642
30	2	0.519	0.484	0.467	0.445	0.419	0.391	0.378
	3	0.810	0.768	0.755	0.721	0.705	0.691	0.650
	4	0.901	0.883	0.864	0.837	0.811	0.776	0.752

From the Tables $5-7$, we observe the following about Statistical power of the tests:

i. Statistical power of the proposed test increases with increase in sample size (n) and scale parameter (θ) but, decreases with increase in censoring percentage (pcens).

ii. In case of all the considered distributions, from Tables $5-7$, we can see that the proposed test attains its maximum power when the sample is ≥ 20, for all scale parameters considered here.

REFERENCES

[1] Wilcoxon F. Individual comparisons by ranking methods. Biometrics. 1945; 1:80-83.

[2] Mann HB., Whitney DR. On a test of whether one of two random variables is stochastically larger than the other. Annals of Mathematical Statistics. 1947;18:50–60.

[3] Gehan EA. A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. Biometrika. 1965;52:202-223.

[4] Efron B. The two-sample problem with censored data. Berkeley Symposium on Mathematical Statistics and Probability. 1967;5(4):831-853.

[5] Ayushee, Kumar N., Goyal M. Two-sample test for randomly censored data. Malaysian Journal of Science. 2022;*(In press)*.

[6] Mood AM. Introduction to the Theory of Statistics. New York: McGraw-Hill; 1950.

[7] Mathisen HC. A method of testing the hypothesis that two samples are from the same population. Annals of Mathematical Statistics. 1943;14(2):188-194.

[8] Chatterjee SK., Sen PK. Non-parametric tests for the bivariate two-sample location problem. Calcutta Statistical Association Bulletin. 1964;13:18-58.

[9] Hettmansperger TP. Statistical inference based on ranks. John Wiley & Sons, Inc; 1984.

[10] Kumar N. A class of two-sample tests for location based on sub-sample medians. Communications in Statistics - Theory and Methods. 1997;26(4):943-951.

[11] Kumar N., Singh RS., Öztürk Ö. A new class of distribution-free tests for location parameters. Sequential Analysis. 2003;22(1-2):107-128.

[12] Brookmeyer R., Crowley J. A k-sample median test for right censored data. Journal of American Statistical Association. 1982;77:433–440.

[13] Prentice RL. Linear Rank Tests with Right Censored Data. Biometrika. 1978;65: 167-179.

[14] Gastwirth J., Wang JL. Control percentile test procedures for censored data. Journal of Statistical Planning and Inference. 1988;18:267–276.

[15] Park HI., Desu MM. Multivariate control test for right censored data. Communications in Statistics- Theory and Methods. 1998;27:1923-1935.

[16] Park HI. Median test for multivariate and right censored data. Journal of Nonparametric Statistics. 2004;16(5):753-760.

[17] Kirk AP., Jain S., Pocock S., Thomas HC., Sherlock S. Late results of the Royal Free Hospital prospective controlled trial of prednisolone therapy in hepatitis B surface antigen negative chronic active hepatitis. Gut. 1980;21:78-83.