International Journal of Contemporary Issues in Social Sciences
Volume 3, Issue 2, 2024 ISSN(P):2959-3808 | 2959-2461

A NOVEL CENSORING RANK TEST FOR SURVIVAL DATA

Syed Almas[1], Qamruz Zaman[2], Danish Wasim[*3], Sofia Mansoor[4]

[1,2,4]Department of Statistics, University of Peshawar, Pakistan;
[*3]Department of Management Sciences, Abasyn University Peshawar, Pakistan

[1]attabaqi11@gmail.com; [2]qamruzzaman@uop.edu.pk; [*3]danishwasim.std@icp.edu.pk ; [4]sofia@uop.edu.pk

Corresponding Author: *

Received: 10 April, 2024 **Revised:** 10 May, 2024 **Accepted:** 28 May, 2024 **Published:** 10 June, 2024

ABSTRACT

The objective of this research was to overcome the problem of moderate to heavy censoring in survival analysis for comparing the survival functions of two groups. In literature several tests are available, e.g. the unweighted log-rank test and the weighted Wilcoxon, tarone ware tests. These tests considered only the events and give no importance to the censoring. But in real life some times, the data set consists of more censored observations than the events. To solve this issue a novel censoring rank test is developed in this research. The test considered the three quartiles i.e. Q1, Q2 and Q3 respectively. Performance of the new method with the existing methods compared by the simulation study. Results of the study showed that the novel test perform much better than the tests available in the literature.The proposed methods along with the available methods are applied to two real data sets i.e. Leukemia and Thalassaemia data sets. Results of the applications are also support the novel test.

Keywords: Survival Analysis, Kaplan-Meier Survival Function, Novel Censoring Rank Test, Log-Rank Test, Leukemia data and Thalassemia data

INTRODUCTION

The term survival analysis is a subfield of biostatistics related to the empirical analysis of the lifetime data. In other words, the analysis of data in which the time until the occurrence of an event is the variable of interest; such as the time spent until occurring of death or recovery from disease etc. Survival analysis is considered the most widely applied area of biostatistics [1]. Survival analysis not only helps in analyzing the data related to the survival of humans in the medical field. But even also technological sciences like engineering are not free from its applications.

Over the last couple of decades, there has been a fast improvement in probabilistic models and analysis tools for modeling the survival time of technological equipment and humans. It has different names in different fields of research. The modeling of lifetime data is known as reliability analysis in engineering sciences, event history analysis in sociology, duration analysis in economics and survival analysis in bio-statistics [2]. The comparison of the survival time across the various groups or diseases is one of the basic approaches of the survival analysis. In many circumstances, the analyst need to compare multiple curve e.g. survival curves of male and female smokers receiving treatments. The graphical comparison can be done with the help of Kaplan Meier (KM) curve.

The log-rank test is based on a very critical assumption of proportional hazards in the two groups. In such circumstances the estimates of survival curves are very beneficial, resulting in a constant hazard ratio and reliable log-rank test results. On the other hand, when the proportional hazard assumption does not hold, then the log-rank statistic loses its power to detect the difference among the groups [3-4].

However, in real life, there are situations when the proportionality assumption of hazard functions is difficult to hold. In these cases, the medication may produce short-term health gain and constantly lose its impact as time progresses. Consequently, an alternative approach, known as the proportional odds model, results in converging hazard functions that usually fit well the survival dataset [5-6]. In general

the $G\rho,\gamma$ test can also be recast as weighted log rank test [7].

Besides these tests, one cannot declare one ideal test that is appropriate in all situations [8]. Lai and Ying [9] proposed a class of estimators in the presence of left-truncation and right-censoring on the responses. The asymptotic normality of the estimators is postulated considering some specific assumptions. This test is more powerful in case of large sample but fails in case of small sample as compared to the log rank test.

Moreau, et al, [10] introduces a generalized approach to semi-parametric alternatives; assigning different weights to the events as proposed by Schoenfeld [11]. For representation purposes, different situations were considered; including two Weibull distributions whose shape parameters were different. This model assigns different weights to first and last event in people.

The empirical comparison of the survival time of the two more groups can be done with the help of the log-rank test and commonly known as the Mental-Cox test. The log-rank test is also considered as the un-weighted test and is more powerful when the proportional hazard assumption is satisfied. In other words, when there is a significant difference between the two curves, the test is more suitable [12].

One of the drawbacks of the log-rank test is that it gives equal weight to every waiting time, as in medical sciences sometimes the initial treatment is more important than the one at later stages. This problem is detected and overcome by Wilcoxon Test [13]. The test is directly related to the number of persons at risk at that time point i.e. n_i. The test is very useful for small sample size but for large it underestimates the results at the lower tail. This problem was solved by Tarone and Ware [14] and assigned the square root of the weight given by Wilcoxon.

Some researchers used different forms of survival function as the weight [15-16]. Similarly, for comparing the survival of imprecise data, more modifications of the rank tests are given by different researchers [17-18].

1.1 NEED OF THE STUDY

The most frequently applied test for comparing the survival times of two treatments are the log-rank, Wilcoxon, Tarone-Ware tests etc. Out of these, the log-rank test assigns equal weight to all observations. While other tests are influenced by early differences [19] and neither gives any importance to the censoring observations directly.

When these tests are applied to the heavily censored data, they are unable to detect the differences or may have small power. To handle this problem, we suggest a new weighted testing procedure that has better power than log-rank, Wilcoxon, and Tarone-Ware tests. Heavy censoring is very common in lifelong diseases e.g. thalassemia. Several works has been done so far the development of rank test from different point of view. Some of them are described below.

An asymptotic test was developed by Peto [15] for comparing the two survival curves. The procedure can also be generalized to more than two groups. Harrington & Fleming[16] determined a test for comparing the survival probabilities of two or more than two groups. In the presence of left and right censoring in the survival data set, Lai and Ying[21] suggested the procedure for regression analysis when the information about dependent variable is not completed. Tsiatis [22] determined the consistent and efficient estimators for the parametric linear regression models in the presence of right censoring. For comparing the groups with censored data a test was produced by Prentice [23] based on linear rank statistics. Similarly, for right censoring data Multiple comparison procedures for comparing the censored data was developed by Koziol & Reid[24] . The base of the test was pairwise ranking scheme. The comparison of different durations of hazard ratios was modeled by Yang and Prentice[25]. The model is more powerful in a situation where proportional hazard assumption is not satisfied.

A nonparametric extension of Wilcoxon test for right censoring was proposed by Gehan[26]. The test gives more satisfactory results than the log-rank test. The comparison and the procedure for finding the significant factors were proposed by Cox [27] through semi parametric cox regression model. Lee [28] recently suggested the improves log-rank test by considering the crossing hazards and the presence of large amount censoring in the survival data. Lee proved the results with the help of detailed simulation study. For the noncontact hazard ratio Wilcoxon test is more powerful than log-rank test. Liu and Yin [29] proposed a modified version of log-rank test based the partitioning of the test to overcome the problem of non-constant hazard ratio. The power of test revealed the results infavor of new procedure. Log-rank test is more suitable in a

situation where the hazard ratio is constant between the two groups. This assumption can be checked by different graphical procedures [30].

Xie and Lie [31] proposed a survival function based on the distribution of confounders to overcome the problem of biasness in Kaplan-Meier survival function by applying the Inverse Probability of Treatment Weighting. A new log-rank test was also introduced for the comparison of two or more groups under the same conditions.

2. METHODOLOGY

Two basic components of survival analysis are event or the occurrence of anything under observation and the concept of censoring. Survival analysis is used for the analysis of univariate, bivariate and multivariate cases. The most popular technique of univariate case is the Kaplan-Meier. The technique is used for calculating the survival probabilities. Asymptotically the function is normally distributed with the Greenwood's variance. For comparing the life expectancy of two groups, different survival tests are used. Out of these tests, the most popular test is the log-rank test. If the data set consists of two or more than two covariates then the concept of Cox Regression is used.

The main objective the present research is based on the log-rank and other weighted tests.

Unweight log-rank test is powerful when the hazard rate is constant between the two groups. For example, consider the death due to headache and due to stage-IV cancer patients. Obviously the death rate of cancer patients is very high as compared to death due to headache, which is almost negligible. But if we compare the cancer patients with the hepatitis A, B patients, then the death rate due to cancer patient is very close to the death rate of hepatitis A, B patients. In such cases weighted tests are more powerful than the log-rank test. The popular weighted tests are Gehan-Wilcoxon test, Tarone-Ware test, Peto test etc. All these weighted and unweighted tests consider the event concept only and giving no importance to the censoring. This can be illustrated below:

Consider a failure time study and suppose k distinct ordered failure times for two groups (labelled I and II) as $t_{(1)} < t_{(2)} < \cdots < t_{(k)}$. At time $t_{(i)}$, d_{1i} and d_{2i} be the number of observed failures in the two groups respectively. Similarly, at time $t_{(i)}$, c_{1i} and c_{2i} be the number of observed censoring in the two groups respectively. Similarly, n_{1i} and n_{2i} denote the number of person's at risk prior to time $t_{(i)}$, respectively, for $i = 1,2,\cdots k$. Consequently, the total number of failure and number of person's at and prior to time $t_{(i)}$ are $d_i = d_{1i} + d_{2i}$, $n_i = n_{1i} + n_{2i}$ respectively. Consider the following two groups consisting of time, number of censored and number of events are described in table 3.1 as:

Table 1:. Layout of the two survival data groups

Group-I							
Time	t_1	t_2	t_3	t_4	t_5	…	t_g
censored	c_1	c_2	c_3	c_4	c_5	…	c_g
Event	d_1	d_2	d_3	d_4	d_5	…	d_g
Group-II							
Time	t_1	t_2	t_3	t_4	t_5	…	t_h
censored	c_1	c_2	c_3	c_4	c_5	…	c_h
Event	d_1	d_2	d_3	d_4	d_5	…	d_h

Furthermore, all information's can be summarized in a contingency Table 2.

Table 2: At observed time $t_{(i)}$, the number of failures in two groups

Group	Number of failures at $t_{(i)}$	Number of censored at $t_{(i)}$	Number surviving	Number at risk just before $t_{(i)}$
I	d_{1i}	c_{1i}	$n_{1i} - d_{1i}$	n_{1i}
II	d_{2i}	c_{2i}	$n_{2i} - d_{2i}$	n_{2i}
Total	d_i	c_i	$n_i - d_i$	n_i

In survival analysis, Log-rank and other weighted tests are used for comparing the survival functions of two or more groups testing:

$$H_0 \equiv S_1(t) = S_2(t)$$

Vs

$$H_1 \equiv S_1(t) \neq S_2(t)$$

Where $S_1(t)$ and $S_2(t)$ are the survival functions of group I and II respectively.

Based under the assumption of null hypothesis that survival is independent of group membership, both the number of failures for two groups and the number of persons surviving can be determined from the value of d_{1i} alone. Once the marginal sums of Table 2 remain fixed, d_{1i} follows a hyper-geometric distribution with mean $E(d_{1i}) = n_{1i}d_i / n_i$ and variance of d_{1i} is

$$Var(d_{1i}) = \frac{n_{1i}n_{2i}d_i(n_i - d_i)}{n_i^2(n_i - 1)}$$

For the overall measure of deviation between the observed and expected failure, sum their differences over the total number of death times to get the statistic.

$$U = \sum_{i=1}^{k}(d_{1i} - E(d_{1i}))$$

and

$$Var(U) = Var\left(\sum_{i=1}^{k}(d_{1i} - E(d_{1i}))\right)$$

$$Var(U) = \sum_{i=1}^{k}\frac{n_{1i}n_{2i}d_i(n_i - d_i)}{n_i^2(n_i - 1)}$$

Furthermore, according to Collett [32]

$$\frac{U}{\sqrt{Var(U)}} \sim N(0,1)$$

$$\frac{U^2}{Var(U)} \sim \chi^2_{(1)}$$

Therefore, test statistic for the log-rank test is:

$$\frac{\left(\sum_{i=1}^{k}(d_{1i} - E(d_{1i}))\right)^2}{Var(U)} \sim \chi^2_{(1)} \quad (1)$$

The log-rank test is more suitable, powerful, and consistent when compared to the test statistics used in circumstances where the survival curves remains parallel. In the case of a crossing curve, we use weighted tests. The most frequently used weighted test is Wilcoxon test, which assigns more weight to initial failure. The test statistic is:

$$\frac{\left(\sum_{i=1}^{k}n_i(d_{1i} - E(d_{1i}))\right)^2}{Var_w(U)} \sim \chi^2_{(1)} \quad (2)$$

Where n_i is the total number of person's prior to time $t_{(i)}$. Where

$$Var_w(U) = \sum_{i=1}^{k}\frac{n_{1i}n_{2i}d_i(n_i - d_i)}{(n_i - 1)}$$

The other commonly used weighted test is the Tarone-Ware, where weight suggested is the square root of the total number of person's prior to time $t_{(i)}$ i.e. $tw = \sqrt{n_i}$.

The Tarone-Ware test is:

$$\frac{\left(\sum_{i=1}^{k}\sqrt{n_i}(d_{1i} - E(d_{1i}))\right)^2}{Var_{tw}(U)} \sim \chi^2_{(1)} \quad (3)$$

$$Var_{tw}(U) = \sum_{i=1}^{k}\frac{n_{1i}n_{2i}d_i(n_i - d_i)}{n_i(n_i - 1)}$$

All these tests do not give any importance to the censoring in a data set and fewer events are observed. The existing methods give us underestimate results, e.g. the thalassemia is life long and difficult process due to which most of the patients leave the study or move to some other places which generates the censoring concept. In such cases we observe more censored observations than the events. So the current tests give us underestimates results. To overcome

this problem, new weighted is developed in this study. The procedure is as follows:

3. PROPOSED METHOD

A comparative weight is assigned to the rank test. The following equation described the weight as.

$$wc = \begin{cases} Q_1 \text{ of the censored, if the } 25\% \text{ censiored observations are present in the data set} \\ Q_2 \text{ of the censored , if the } 50\% \text{ of censored observations are present in the data set} \\ Q_3 \text{of the censored observations, if equal or more than } 75\% \text{ of the data consists of censored observations} \end{cases}$$

The method can also be illustrated through Figure 1.

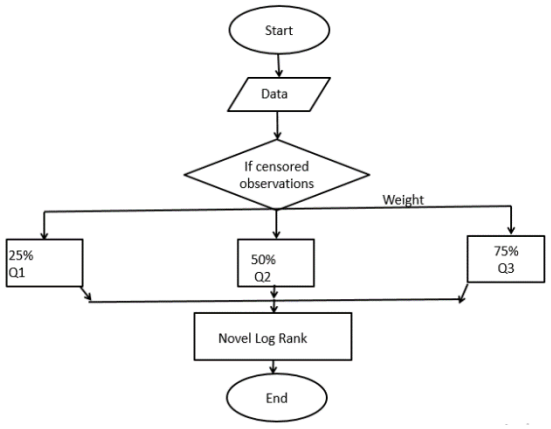

Figure 1:. Illustration of the Novel Log-Rank Test through flowchart.

So the procedure can be illustrated below:

3.1 METHOD FOR UP TO 25% CENSORED OBSERVATIONS

If the data set consists of less than equal to 25% observation the procedure for testing the hypothesis is illustrated with the modification of the table 2 as:

Table 3:. At observed time $t_{(i)}$, the number of failures and censored in two groups

Group	Number of failures at $t_{(i)}$	Number of censored at $t_{(i)}$	Number surviving	Number at risk just before $t_{(i)}$
I	d_{1i}	c_{1i}	$n_{1i} - d_{1i}$	n_{1i}
II	d_{2i}	c_{2i}	$n_{2i} - d_{2i}$	n_{2i}
Total	d_i	c_i	$n_i - d_i$	n_i
	If c_i is <= 25% censored observations then calculate Q_1 If c_i is <= 50% censored observations then calculate Q_2 If c_i is <= 75% censored observations then calculate Q_3			

In survival analysis, null and alternative hypothesis for comparing the two survival groups are given below:

$$H_0 \equiv S_1(t) = S_2(t)$$

Vs

$$H_1 \equiv S_1(t) \neq S_2(t)$$

Where $S_1(t)$ and $S_2(t)$ are the survival functions of group I and II respectively.

Based under the assumption of null hypothesis that survival is independent of group membership, both the number of failures for two groups and the number of persons surviving can be determined from the value of d_{1i} alone. Once the marginal sums of Table 3.3 remain fixed, d_{1i} follows a hyper-geometric distribution with mean $E(d_{1i}) = n_{1i}d_i / n_i$ and variance of d_{1i} is

$$Var(d_{1i}) = \frac{n_{1i}n_{2i}d_i(n_i - d_i)}{n_i^2(n_i - 1)}$$

For the overall measure of deviation between the observed and expected failure, sum their differences over the total number of death times to get the statistic.

$$U = \sum_{i=1}^{k}(d_{1i} - E(d_{1i}))$$

and

$$Var(U) = Var\left(\sum_{i=1}^{k}(d_{1i} - E(d_{1i}))\right)$$

$$Var(U) = \sum_{i=1}^{k}\frac{n_{1i}n_{2i}d_i(n_i - d_i)}{n_i^2(n_i - 1)}$$

Furthermore, according to Collett [20]

$$\frac{U}{\sqrt{Var(U)}} \sim N(0,1)$$

$$\frac{U^2}{Var(U)} \sim \chi^2_{(1)}$$

Therefore, test statistic for the log-rank test is:

$$\frac{\left(\sum_{i=1}^{k}(d_{1i} - E(d_{1i}))\right)^2}{Var(U)} \sim \chi^2_{(1)}$$

Since in log-rank test statistic no weight is assigned to the censored observations so if the data consists of < =25% censored observations. Calculate Q_1 for the censored observations in the entire data set. So the test statistics of the Novel Log-Rank test for <=25% censored observations is:

$$\frac{[\sum_{i=1}^{k} Q_1(d_{1i} - E(d_{1i}))]^2}{Var_{Q_1}(U)} \sim \chi^2_{(1)} \qquad (4)$$

Where

$$Var_{Q_1}(U) = \sum_{i=1}^{k}\frac{Q_1\ n_{1i}n_{2i}d_i(n_i - d_i)}{n_i^2(n_i - 1)} \qquad (5)$$

3.2 METHOD FOR UP TO 50% CENSORED OBSERVATIONS

If the data set composed of 50% of the censored observations, then the median or Q_2 is assigned as the weight to the censored observations. So the Novel-Rank test for the 50% censored observations is obtained by replacing Q_1 by Q_2 in equations 4 and 5.

$$\frac{[\sum_{i=1}^{k} Q_2(d_{1i} - E(d_{1i}))]^2}{Var_{Q_2}(U)} \sim \chi^2_{(1)} \qquad (6)$$

Where

$$Var_{Q_2}(U) = \sum_{i=1}^{k}\frac{Q_2\ n_{1i}n_{2i}d_i(n_i - d_i)}{n_i^2(n_i - 1)} \qquad (7)$$

3.3 METHOD FOR UP TO 75% CENSORED OBSERVATIONS

If the data set composed of 75% of the censored observations, then Q_3 is assigned as the weight to the censored observations. So the Novel-Rank test for the 75% censored observations is obtained by replacing Q_1 by Q_3 in equations 4 and 5.

$$\frac{[\sum_{i=1}^{k} Q_3(d_{1i}-E(d_{1i}))]^2}{Var_{Q_3}(U)} \sim \chi^2_{(1)} \quad (8)$$

Where

$$Var_{Q_3}(U) = \sum_{i=1}^{k} \frac{Q_3\ n_{1i}n_{2i}d_i(n_i-d_i)}{n_i^2(n_i-1)} \quad (9)$$

3.4 NOVEL RANK TEST FORMULA

The following test statistic is used for the Novel Rank Test:

$$\frac{[\sum_{i=1}^{k} W_c(d_{1i}-E(d_{1i}))]^2}{Var_{W_c}(U)} \sim \chi^2_{(1)} \quad (10)$$

Where

$$Var_{W_c}(U) = \sum_{i=1}^{k} \frac{W_c\ n_{1i}n_{2i}d_i(n_i-d_i)}{n_i^2(n_i-1)} \quad (11)$$

SIMULATION STUDY

To compare the performance of novel test with the Log-rank, Wilcoxon and Tarone-Ware tests simulation study is performed. For generating the different percentages of censoring Uniform different distribution is used. Two survival distributions Weibull and Exponential distributions are used for drawing the failure times. R package is used for analysis purpose. For generating the data comparatively large sample sizes i.e. 100, 500 and 1000 are used for each group. Simulation study is performed for 5000 times. Under different conditions, the results of simulation study are summarized below.

4. FOR OBTAINING THE 25% CENSORING

For obtaining the required number of censored observations, the procedure is adopted as:

Survival times for Group I and Group II were generated from Exponential distribution (E(1)). The censoring distribution for Group I was Uniform (U(0, 6.5)). Similarly, U(0, 3.2) was used to obtain censoring for Group II. The censoring percentage is 25%. The pattern is shown in Figure 4.1.

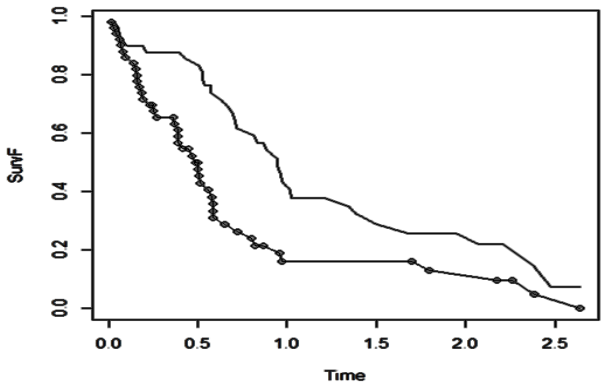

Figure 4. 1: Survival Curves at 25% censoring

Table 4. 1: Comparison of different tests for 25% censoring

Information	Results	Information	Results
Group1		Group2	
n1=100		**n2=100**	
Survival Distribution	E(1)	Survival Distribution	E(1)
Censoring Distribution	U(0, 6.5)	Censoring Distribution	U(0, 3.2)
Overall Censoring Percentage			25%
Novel Rank Test		**22.26886 24.37898**	
P-Value		**3.245677e-07**	
LOGRANKTEST		7.12345	
P-Value		0.0024567831	
TARONEWARETEST		13.1691	
P-Value		0.0004456789	
WILCOXONTEST		13.57687	
P-Value		0.0003404508	

Table4.1shows that the novel rank test has the smallest P-value as compared to the other tests. This indicates the significant difference between the survivals of two groups. The log-rank test which is considered as the most powerful test in proportional hazard case has even smaller P-value than Tarone Ware and Wilcoxon Tests. The simulation study of sample sizes of 500 and 1000 also give the results infavor of novel rank test

4.2 FOR OBTAINING THE 50% CENSORING

Survival times for Group I and Group II were generated from Weibull distribution (Weibull(0.5, 50)). The censoring distribution for Group I was Uniform (U(2, 25)). Similarly, U(2, 25) was used for Group II. Overall nearly 50% censored observations are obtained from the two groups.

Figure 4.2: Survival Curves at 50% censoring

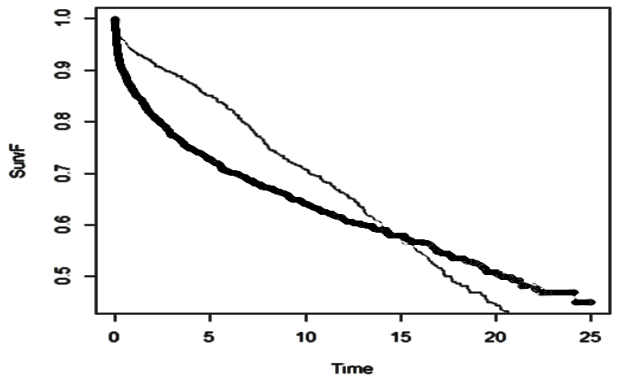

Table 4. 2: Comparison of different tests for 50% censoring

Information	Results	Information	Results
Group1		Group2	
n1=500		**n2=500**	
Survival Distribution	Weibull(0.5, 50)	Survival Distribution	Weibull(0.5, 50)
Censoring Distribution	U(2, 25)	Censoring Distribution	U(2, 25)
Overall Censoring Percentage			49.5%
Novel Rank Test		**5.345456**	
P-Value		**0.01231345**	
LOGRANKTEST			
PValue		0.4312568	
TARONEWARETEST		0.00987456	
PValue		0.9768909	
WILCOXONTEST		0.376858985	
PValue		0.64564356	

Figure 4. 2: Survival Curves at 50% censoring

The novel rank test has the largest value resulting the smallest P-value indicating the significant difference between the two groups. The conventional log-rank test, Wilcoxon and tarone test gave the largest P-values as compared to the proposed test.

4.3 FOR OBTAINING THE 75% CENSORING

For obtaining the largest percentage of censoring i.e 75%, Weibull (0.5, 50) and U(2, 25) were used for generating the failure time and censoring observations from both the group. The pattern of the curves is illustrated in Figure 4.3 and the simulation results are summarized in table 4.3.

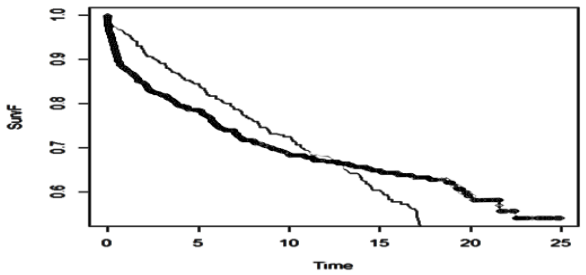

Figure 4. 3: Survival Curves at 64.3% censoring

Test Statistics and other information's are summarized in the Table 4.6:

Table 4. 3: Comparison of different tests for 75% censoring

Information	Results	Information	Results
Group1		Group2	
n1=1000		**n2=1000**	
Survival Distribution	Weibull(0.5, 50)	Survival Distribution	Weibull(0.5, 50)
Censoring Distribution	U(2, 25)	Censoring Distribution	U(2, 25)
Overall Censoring Percentage			75%
Novel Rank Test		**14.356467**	
P-Value		**0.0008765656**	
LOGRANKTEST		3.45634	
P-Value		0.0776466	
TARONEWARETEST		0.03695264	
P-Value		0.799876	
WILCOXONTEST		1.876567	
P-Value		0.1631335	

Table4.3 also shows the results infavour of new test. The test statistic value of the test is 14.356467with P-Value 0.0008765656 indicates the significant difference between the two groups. P-values for the rest of test statistics greater than 0.05, this indicates no difference between two groups. So the test based on novel rank test, performs much better than the other tests.

4.4 SMALL PERCENTAGE OF CENSORING (15%) FOR UNEQUAL SAMPLE SIZES

For obtaining the small percentage of censoring, U(0, 6.5) was used for both Group I and Group II. Similarly, the failure distribution was E(1) for both the groups. The sample size for group I was 500 and for group II was 1000. The pattern is shown in Figure 4.4 the summary of results is given in Table 4.4.

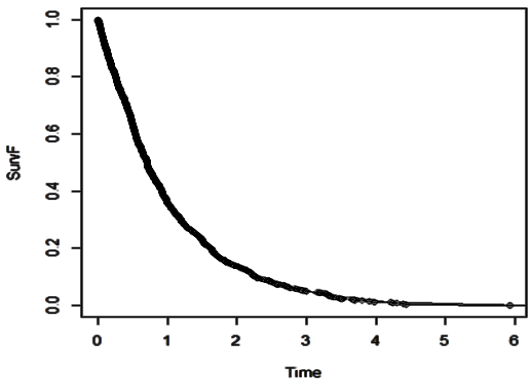

Figure 4. 4: Survival Curves of 15.7% censorin

Table 4. 4: Comparison of different tests for 15.7% censoring for unequal sample sizes

Information	Results	Information	Results
Group1		Group2	
n1=500		**n2=1000**	
Survival Distribution	E(1)	Survival Distribution	E(1)
Censoring Distribution	U(0, 6.5)	Censoring Distribution	U(0, 6.5)
Overall Censoring Percentage			15.7%
Novel Rank Test		0.0035678	
P-Value		0.95050705	
LOGRANKTEST		0.0072973	
P-Value		0.9319241	
TARONEWARETEST		0.02042892	
P-Value		0.8863456	
WILCOXONTEST		0.05253605	
P-Value		0.8187077	

The figure 4.4 indicates the no difference between the two groups. Table 4.4 shows that the test based on novel test gives the high P-Value. Rest of the tests gives less P-Value. In this situation, novel test gives better results. This indicates no difference between the two survival curves.

5. APPLICATION TO REAL DATA SETS

Novel censoring Rank Test is applied to two real data sets namely, Leukemia and Thalassemia.

5.1 APPLICATION TO LEUKEMIA DATA SET

The data of leukemia patients of two groups is taken from the Kleinbaum and Klein (1996) book. The remission times in weeks for two groups of 21 patients in each group i.e. 21 in treatment group and 21 in placebo group. The hypothesis is based on survival experience of two groups. Details of the groups are given below:

Group 1: 6, 6, 6, 7, 10,13, 16, 22, 23,6+, 9+, 10+, 11+,17+, 19+, 20+,25+, 32+, 32+,34+, 35+

Group1 consists of 9 failure times and 12 censored observations. Where + denotes the censored observation.

Group2: 1, 1, 2, 2, 3,4, 4, 5, 5,8, 8, 8, 8,11, 11, 12, 12,15, 17, 22, 23

Group2 consists of 21 failure times and 12no censored observation.

The pattern of the two survival curves is described in Figure 5.1 and the result of analysis is summarized in table 5.1.

Figure 5.1 Survival curves of Treatment vs Placebo Groups

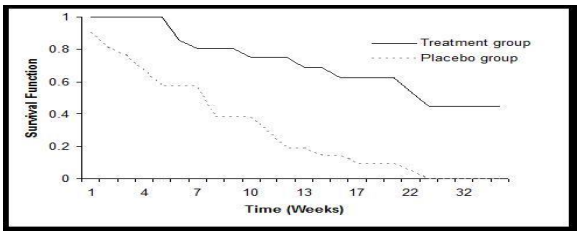

Table 5. 1: Comparison of Treatment vs Placebo groups

Information	Results	Information	Results
Group1: Treatment Group		Group2: Placebo Group	
n1=21		**n2=21**	
Overall Censoring Percentage			23.5%
Novel Rank Test		**27.12565**	
P-Value		**0.0000000**	

LOGRANKTEST	16.79
P-Value	0.000045
TARONEWARETEST	15.12
P-Value	0.0001
WILCOXONTEST	13.46
P-Value	0.0002

Table5.1shows that the test statistic value of novel censoring rank test is greater than the rest of the test statistics values indicates the significant difference between two groups. Similarly P- value of the test is much smaller than the rest of the test statistics. This indicates the significant difference between the survivals of two groups. The logrank test which is considered as the most powerful test in proportional hazard case has even smaller P-value than the proposed method.

Conclusion and Discussion

The main theme of this research was to develop a method to overcome the problem of censoring in rank tests. In literature several tests are available, some based on weighted and some on unweighted procedures. Out of these tests, the most commonly used tests are the unweighted log-rank test and the weighted Wilcoxon, tarone ware tests. These and all the tests are considered only the events and give no importance to the censoring. Although it is censoring, which differentiate the survival analysis procedure from the traditional regression method. In real life some times, the data set consists of more censored observations than the events. To solve this issue a novel test is developed in this research. The novel test gives weight on the basis of quartiles. First quartile weight is assigned to the less than 25% censored observation data. Second quartile or median weight is assigned to the 50% censored data and third quartile weight is assigned to the more than 75% censored data.

To compare the performance of new tests with the available tests, simulation study based on 5000 iterations was performed. Two famous failure distributions namely Exponential and Weibull were used for drawing the failure times and Uniform distribution was used for drawing the censoring time. Sample sizes 100, 500 and 1000 were considered for the simulation study. Results of the study showed that the novel test perform much better than the tests available in the literature.

The proposed methods along with the available methods are applied to two real data sets i.e. Leukemia data set. Results of the applications are also support the novel test.

REFERENCES

Zaman, Q.; Pfeiffer, K. P. Survival Analysis in Medical Research. *Inter stat.* 2011, 17(4): 1-36.

Kleinbaum, D. G.; Klein, M. Survival Analysis A Self-Learning Text. *Springer:* New York. 1996.

Martinez M. Diagnostics for choosing between Log-Rank and Wilcoxon test (Ph.D). *Western Michgan University*. 2007.

Yang S.; Prentics R. Improved Log rank-Type Tests for Survival Data Using Adaptive Weights. *Biometrics.* 2010, 66(1): 30-38.

Bennett, S. Analysis of survival data by the proportional odds model. *Statistics in Medicine.* 1983, 2: 273–277.

Zucker DM. The efficiency of a weighted log-rank test under a percent error miss specification model for the log hazard ratio. *Biometrics.* 1992, 48: 893–899.

Fleming T. R.; Harrington D. P. Counting Processes and Survival Analysis. *New York: Wiley.* 1991.

Klein, J. P.; Moeschberger, M. Survival Analysis, Techniques for censored and Truncated Data. *Springer*. 1998.

Lai, T. Z.; Ying, Z. Rank regression methods for left-truncated and right-censored data. *Annals of Statistics*, 1991, 19: 531–556.

Moreau, T.; Maccario, J.; Lellouch. J.; Huber, C. Weighted log rank statistics for comparing two distributions. *Biometrika.* 1992, 79: 195-198.

Schoenfel, D. The asymptotic properties of nonparametric tests for comparing survival distributions. *Biometrika*. 1981, 68: 316-9.

Yang, S.; Prentics, R. Improved Log rank-Type Tests for Survival Data Using Adaptive Weights. *Biometrics.* 2010, 66(1): 30-38.

Genhan, E. A. A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. *Biometrika*. 1965, 52: 203-223.

Tarone, R. E.; Ware J. On distribution-free tests for equality of survival distributions. *Biometrika.* 1977, 64: 156-160.

Peto, R.; Peto J. Asymptotically Efficient Rank Invariant Test Procedures, *Journal of the Royal Statistical Society, Series A.* 1972, 135: 185-206.

Fleming, T. R.; O'Fallon, J. R.; O'Brien P. D.; Harrington D. P. Modified Kolmogorov Smirnov test procedures with applications to arbitrarily right-censored data, *Biometrics.* 1980, 36: 607-625.

Augustin, T.; Coolen, F. P. A.; De Cooman, G.; Troffaes. M. C. M. Introduction to imprecise probabilities. *Chichester: Wiley.* 2014.

Coolen, F. P. A.; Ahmadini, A. A. H.; Coolen-Maturi, T. Imprecise inference based on the log-rank test for accelerated life testing. *Metrika*. 2021.

Spoto, R.; Stablein, D.; Carter-Camobell, S. A partially grouped log rank test, *Statistics in medicine.* 1997, 16: 695-704.

Lai, T.L.; Ying Z. Rang regression methods for left-truncated and right censored data. The Annals of statistics. 1991, 531-556.

Tsiatis, A. A. Estimating regression parameters using linear rank tests for censored data. *The Annals of Statistics.* 1990, 354-372.

Prentice, R. L. Linear rank tests with right censored data. *Biometrika*, 1978, *65*(1), 167-179.

Harrington,D. P.; Fleming, T. R. A class of rank test procedures for censored survival data.Biometrika. 1982, 553-556.

Koziol, J.A.; Reid, N. On multiple comparisons among K samples subject to unequal patterns of censorship. Communications in statistics theory and methods. 1997, 1149-1164.

Yang, S.; Prentice, R. Semi parametric analysis of short-term and long-term hazard ratios with two- sample survival data. Biometrica. 2005, 1-17.

Gehan, A. A. A generalized Wilcoxon test for comparing arbitrarily singly-censored samples. *Biometrika*,1965, 203-224.

Cox, D. R. Regression Models and life-tables. Journal of royal statistical society, 1972, 187-220.

Lee, S. H. Weighted Log-Rank Statistics for Accelerated Failure Time Model. *Stats*, 2021, 4(2): 348-358.

Lee, E.T .; Desu, M. M.; Gehan, E. A. A Monte Carlo Study of the Power of Some Two-Sample Tests. *Biometrika*. 1975, 62(2): 425-432.

Liu, Y.; Yin, G. Partitioned log-rank tests for the overall homogeneity of hazard rate functions. *Lifetime Data Anal.* 2017, 23: 400–425.

Hess, K. Graphical Methods for Assessing Violations of the Proportional Hazards Assumption in Cox Regression. Statistics in Medicine, 1995, 14: 1707-1723.

Xie, J.; Liu, C. Adjusted Kaplan–Meier estimator and log-rank test with inverse probability of treatment weighting for survival data. *Statistics in medicine*, 2005, *24*(20): 3089-3110.

Collett, D. Modelling Survival Data in Medical Research. *Chapman and Hall.* 1994.