Significance of Bootstrap technique in Diabetic Eye Patients Vision level improvement

Ramakrishnan.M[1], Manikandan.M[2] and Ravanan.R[3]

[1]Associate Professor, Department of Mathematics, RKM Vivekananda College, Chennai, Tamil Nadu - 600004, India.
Email: mramkey@rkmvc.ac.in
[2]Research Scholar, Department of Mathematics, RKM Vivekananda College, Chennai, Tamil Nadu – 600004, India.
Email:manijee211993@gmail.com
[3]Joint Director (Planning and Development) Directorate of Collegiate Education, Chennai, Tamil Nadu – 15, India.
Email: ravananstat@gmail.com

DOI: 10.29322/IJSRP.14.07.2024.p15112

Paper Received Date: 15th June 2024
Paper Acceptance Date: 15th July 2024
Paper Publication Date: 20th July 2024

Abstract

Distribution free statistics helps to estimate the survivorship functions from incomplete observations. It does not include any assumptions about the distribution of survival data. Kaplan-Meier(KM) and Nelson-Aalen(NA) estimators are very well-known models in distribution free methods, which are applied to get the estimated survival probability of various subjects with different predictors. Survival distributions between groups were analysed by nonparametric tests such as log-rank, Wilcoxon, Tarone-Ware, Peto-Peto and Fleming-Harrington test. These tests infer the significance of survival difference between two or more groups in the data. The bootstrap is a computation-intensive data re-sampling methodology for assessing the accuracy of a statistical estimator. The nonparametric bootstrap involves randomly sampling data with replacement to form a "new" sample of data, which is referred to as a bootstrap sample. Parametric models are popular among survival data analysts because they provide insight into the parameters and functions being examined, especially the hazard rate. Exponential, Weibull, Lognormal, Log-logistic are some of the most well-known models. The aim of this paper is to compare Nonparametric estimators (KM & NA), with Bootstrap technique. Also, we examine the effectiveness of bootstrapping on the confidence intervals of various parameters of interest. In addition, Parametric survival analysis methods can be applied to identify an appropriate statistical model for the observed data. Diabetic Eye patient's (DEP) details were collected from EYDOX Eye Hospital Private Limited, Chennai and their improvement of vision level is taken as event. And using DEP data, observed and identified the applications of the Bootstrap to Survival Analysis.

Keywords: Nonparametric methods; Parametric methods; Bootstrap Technique; Bootstrap Confidence Interval.

1. Introduction

The primary goal of event history analysis is to discover and evaluate the survival function based on time to event data. Before attempting theoretical distribution for survival data, nonparametric approaches can be used. Estimates using nonparametric techniques and graphs can be useful in determining a model for the data. The survival functions for censored data may be estimated using this technique. Because of their outstanding significant features, the procedures provided by KM and NA are the most well-known of many similar techniques known in the research.

There are several tests that are used to compare survival distributions, including the Generalized Gehan-Wilcoxon test (Breslow, 1970), Generalized Mantel-Haenszel test (Tarone et al. 1977), Peto and Peto (1972), and Fleming and Harrington (1981) Test. Most frequently, Mantel-Haenszel test is referred to as the log-rank test. The standard Kaplan-Meier estimation overestimates the survival function. Other approaches for estimating the survival function include Nelson-Aalen, Weighted and Modified weighted KM. Larger weights are applied, when observations are failed in early stage but smaller weights are given if it appears in right tail of the failure distribution (Prentice and Marek, 1979). Based on clinical expectations for the outcome, weights will be introduced in F-H test. It must be taken in action before evaluating the data (Klein et al., 2001; Gomez et al., 2009). Enrico A. Colosimo et al. (2002) observed Nelson-Aalan estimator has a slight advantage when applying survival estimates. However, for percentile estimates, lower default rates improve the performance of KM estimators, and higher default rates improve the performance of NA estimates. Ramakrishnan and Ravanan (2013) applied nonparametric methods and tests for different areas of study like Sociology and Medical.

In 1967, Efron gave a new insight to this estimator in an important paper (Efron, 1967).Efron(1981) proposed bootstrap methods for right randomly censored observations. His first proposal is to take independent resamples from the KM estimators of the lifetimes and

the censoring times, and to combine them by taking minima and indicators (Akritas, 1986). For grouped data, he recommends specific bootstrap bands. Frederic Utzet et al. (1992) reviewed classical applications of bootstrap methods to survival analysis. Efron and Tibshirani (1993) reviewed bootstrap methods. Young Joon Cha & Jae Man Lee (1995) observed bootstrap confidence intervals are better than the asymptotic normal-theory confidence intervals in terms of achieved error rates and confidence length. Davison and Hinkley (1997) discussed a few of the basic points about nonparametric resampling in a relatively simple context – balanced one-way nested classification. Quantiles have been studied in Van Keilegom and Veraverbeke (1997). MichałMichalak et al. (2002) showed how Bootstrap KM Estimator paid attention to its advantage opposite to Classical Analysis. FatmaSarhan and Edward Susko (2020) observed bootstrapping perform better with higher sample sizes and surprisingly with increased censoring.

2. Nonparametric Methods

The distribution free calculation of the survival function from partial data was taken into consideration by Kaplan and Meier in 1958. They overcame difficulty in predicting the survival curve in a straightforward manner while accounting for the proper censoring. When the functional distribution of the hazard rate with time is not assumed, the Kaplan Meier technique is appropriate. The NA estimator is an alternate to KM that uses a counting process technique to predict the cumulative hazard function. It does not work with any distribution. The survival probability for both censored and uncensored data may be calculated non-parametrically using observable failure times.

2.1 The Kaplan-Meier's Product-Limit formula

In survival analysis, each observations are represented as (T_i, C_i), where $i = 1, 2, \ldots, n$. Here T_i and C_i represents value for the survival time variable T and censoring variable C for n independent observations respectively. Among n observations, m<n observations attained the event and remaining are treated as censor. Then arrange the event time observations as

$$t_1 < t_2 < \cdots < t_m$$

The number of observations at risk of attaining event at $t_{(i)}$ is represented as n_i, and the number of deaths recorded is given d_i. At time t, the Kaplan-Meier survival estimator is

$$\hat{S}(t) = \prod_{t_i \le t} \left(1 - \frac{d_i}{n_i}\right)$$

Using the customized that $\widehat{S}(t) = 1 \; if \; t < t_1$.

2.2 The Nelson-Aalen Estimator

The Nelson-Aalen estimate was proposed by Aalen (1975, 1978), Nelson (1969, 1972), and Altshuler (1970). The NA estimate of H(t) is represented by.

$$\widehat{H}(t_j) = \sum_{i=1}^{j} \frac{d_i}{n_i}.$$

Similar to Kaplan-Meier, this approach is used to estimate data that has been censored. The estimate of the NA Survival function is provided by

$$\hat{S}(t_j) = \prod_{i=1}^{j} \exp\left(-\frac{d_i}{n_i}\right).$$

There is never a moment when the KM estimation will be bigger than the NA estimate of the survival function. In large samples, however, the distinction between the two approaches will be insignificant.

2.3 Confidence Interval for KM and NA

Classical confidence interval for Kaplan-Meier estimator is given by formula

$$\hat{S}(t) \pm \widehat{D}\{\hat{S}(t)\}.u\left(\frac{\propto}{2}\right)$$

$\widehat{D}\{\hat{S}(t)\}$ is standard error of survival function and according to Greenwood (1926) is given by the formula

$$\widehat{D}\{\hat{S}(t)\} = [\hat{S}(t)]\sqrt{\sum_{j=1}^{k} \frac{d_j}{n_j(n_j - d_j)}} \text{ for } t(k) \le t \le t(k+1)$$

While u(∝) is the quantile of standard normal distribution of the order of ∝.
In the similar manner, Nelson Aalen estimator is given by the formula

$$\hat{S}(t_j) \pm \hat{D}\{\hat{H}(t)\}.u\left(\frac{\propto}{2}\right)$$

The point wise standard error $\hat{D}\{\hat{H}(t)\}$ of the log of the survivor function can be computed directly as the standard error or a Bernoulli random variable with d_i successes from n_i samples:

$$\hat{D}\{\hat{H}(t)\} = \sqrt{\sum_{i:t_i<t} \frac{d_i(n_i - d_i)}{n_i^3}}$$

2.4 Survival Comparison Tests

The popular log-rank test checks whether the two Kaplan-Meier curves are statistically equivalent. To be precise it compares two Kaplan-Meier curves. Kaplan-Meier curves shows difference between two curves but it won't speak about the difference whether it is statistically significant or not and compares KM curves required. This chi-square test is used for large sample. The test requires information about the ordered failure times of each subject in the entire data and it splits the failed subjects at that particular ordered failure time and subjects in the risk set respectively.
General formula for weighted test statistic

$$\text{Test statistic } Q = \frac{\left(\sum_j w(t_j)(m_{ij} - e_{ij})\right)^2}{var\left(\sum_j w(t_j)(m_{ij} - e_{ij})\right)}$$

Where$w(t_j)$ is called weights at j[th] failure time. And can be varied to give different desired results. If the value of $w(t_j)$ becomes one then any test will become log-rank test.

2.4.1 Log-rank test

It's a non-parametric test. It compares the entire survival experience between the groups.
Null hypothesis (H_0): There is no difference in survival between the groups.
Alternative hypothesis (H_1): The difference in survival between the groups is significant.
Log-rank test statistic $= \frac{(O-E)^2}{Var(O-E)}$,
This is approximately equal to the large sample chi-square test statistic,

$\chi^2 = \frac{(O_1-E_1)^2}{E_1} + \frac{(O_2-E_2)^2}{E_2}$, with one degrees of freedom (for two groups). If n_{1j}, n_{2j} and m_{1j}, m_{2j} are the number of individuals in the risk set and number of events occurred, at the j[th] ordered time $t_{(j)}$ of group 1 and 2 respectively then,

$$O_i = \sum_j m_{ij} \; and \; E_i = \sum_j e_{ij} \;, e_{ij} = \frac{n_{ij}}{n_j} \times d_j, \; For \; i = 1,2.$$

$$Where \; n_j = n_{1j} + n_{2j} \; and \; d_j = m_{1j} + m_{2j}$$

When $w(t_j)$ = 1 applied in the test statistic Q, it becomes log-rank test.

2.4.2 Wilcoxon or Generalized Wilcoxon test.

The Wilcoxon rank-sum test was expanded by Gehan (1965) and Breslow (1970) to include censored data. The weights $w_j = n_j$ is assigned for each j[th] observations in the test statistic Q then the statistic is called as Wilcoxon or Generalized Wilcoxon test. This test detects early differences in the data.

2.4.3 Tarone and Ware test

This test statistic also applies higher weight to early failure times. In the test statistic Q, put $w_j = \sqrt{n_j}$ for time t_j for this test.

2.4.4 Peto-Peto Test

Weights in this test are determined by the expected percentile of the failure time distribution. When the hazard ratio across groups is not constant, this test might be utilized (Stevenson, 2009). The Peto-Peto test takes weight $w_j = \tilde{S}(t_j)$ in Q. The function that exists here is a modified version of the KM estimator (Allison, 2010).

2.4.5 Fleming-Harrington(FH) Test

The FH test uses the KM survival estimate $\hat{S}(t)$ across all groups to calculate its weights for the ^{jth} failure time, $\widehat{S(t_{j-1})}^{p}[1-\hat{S}(t_{j-1})]^{q}$.

In the FH tests, the choice of p and q determines the weights for intermediate and delayed events (Gomez et al. 2009; Oller and Gomez, 2012). When applying the value 0 for both p and q in FH test then the test is equivalent to a log-rank test. For the Peto-Peto test, assume p = 1 and q = 0 in FH test (Harrington and Fleming, 1981).

Table 1. Test Statistics for Non-Parametric Test and their Weights

Test Statistics	$w(t_j)$
Log rank	1
Wilcoxon	n_j
Tarone-Ware	$\sqrt{n_j}$
Peto	$\tilde{s}(t_j)$
Fleming-Harrington	$\widehat{s(t_{j-1})}^{p}[1-\hat{s}(t_{j-1})]^{q}$

3. Parametric Methods

While semi-parametric model focuses on the influence of covariates on hazard, fully parametric model can also calculate the distribution form of survival time. Parametric Survival Models are Exponential, Weibull, Lognormal and Log logistic distributions. Performance between models is compared using Akaike Information Criterion (AIC) a measure of the goodness of fit for statistical models.

3.1 Exponential distribution

Exponential distribution model is very simple because it depends on only one parameter λ which represents the hazard function. Hazard and survival function are given below

$$S(t,\lambda) = e^{-\lambda t}$$

$$h(t,\lambda) = \lambda, where\ \lambda > 0\ , for\ 0 \le t < \infty$$

When the value of λ is equal to one then the distribution is called as unit exponential distribution. Also λ is assumed to be constant. Since the value of hazard function is constant, so the pattern of covariates becomes much stronger. Because of the hazard function is constant, the hazard ratio is also constant. But converse need not be true that is hazard ratio is constant hazard need not be constant.

3.2 Weibull Distribution

In Weibull model the hazard and survival function are known as

$$S(t,\lambda,\gamma) = e^{(-\lambda t^{\gamma})}$$

$$h(t,\lambda,\gamma) = \lambda\gamma t^{\gamma-1}, where\ \lambda > 0, \gamma > 0$$

here the notation γ represents as shape parameter since it determines the shape of the hazard function. When γ equals to one, the hazard function is constant and if the value of γ is greater than one then the hazard function will be increasing along with the time, if the value of γ is less than one then the hazard function will be decreasing along with time. The Weibull distribution need not require the hazard rate to be constant. Another parameter of Weibull is λ which we call as the scale parameter. The shape parameter controls the shape of the hazard function. The scale parameter defines its scale.

3.3 Lognormal Distribution

Lognormal distribution is defined as the distribution of a variable whose logarithm follows the normal distribution. If $logT$ is normally distributed with mean and variance then it is called as lognormal distribution. In lognormal distribution, the hazard function increases at first but then decreases as time increases. The lognormal distribution has two parameter which is mean μ and variance σ^2.

$$f(t,\mu,\sigma^2) = \frac{1}{t\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2\sigma^2}(\log t-\mu)^2\right], t>0$$

$$S(t,\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}}\int_t^{\infty}\frac{1}{x}\exp\left[-\frac{1}{2\sigma^2}(\log x-\mu)^2\right]dx$$

$$h(t,\mu,\sigma^2) = \frac{f(t)}{S(t)} \quad ,where\ \sigma>0$$

3.4 Log-logistic Distribution

A Log-logistic random variable T with parameters scale parameter λ and positive shape parameter γ. The hazard and survival functions are,

$$S(t,\lambda,\gamma) = \frac{1}{(1+\lambda t^{\gamma})}$$

$$h(t,\lambda,\gamma) = \frac{\lambda\gamma t^{\gamma-1}}{(1+\lambda t^{\gamma})} \quad ,where\ \lambda>0,\gamma>0$$

where if γ less than or equal to one then hazard decreases over time. If the value of γ is greater than one then hazard function is arc-shaped (increases and then decreases).

Performance between models is compared using Akaike Information Criterion (AIC), a measure of the goodness of fit for statistical models. AIC calculates the quality of each model, relative to each of the other models. The AIC is provided by

$$AIC = -2*(\text{loglikelihood}) + 2(k+c)$$

where k is the number of covariates in the model excluding constant terms and c is the number of model-specific distributional parameters. Essentially, Comparing the AIC scores of various Parametric models and then choose the model with the lowest AIC score.

3.5 Estimation of ratio of two parameters and its confidence interval

To estimate parameters of interest and their standard error, 'survreg' function in the R package can be applied to regression model. When it comes to survival regression model, the following log-linear model is used. $\log(x_i) = B_0 + B_1 z_i + \in_i, where\ x_i > 0$, where B_0 and B_1 are unknown constants (parameters) and z_i is a covariate. When working with the two sample case, without and with censoring, where one of the samples come from a population with parameter $\lambda_0(W(\lambda_0,1) = \exp(\lambda_0))$ and the other from a population with $\lambda_1(W(\lambda_1,1) = \exp(\lambda_1))$. The regression method was used to acquire the confidence interval of the ratio, $\frac{\lambda_1}{\lambda_0}$. In this situation, the z_i in the model is a factor that identifies which of the two samples the ith observation is from.

$$B_0 = -log\lambda_0$$

while $B_1 = -log\lambda_1 + log\lambda_0$. The ratio of $\frac{\lambda_1}{\lambda_0}$ is related to B_1 as follows,

$$e^{-B_1} = \frac{\lambda_1}{\lambda_0}$$

where B_1 is estimated by $\hat{B}_1$. Both $\hat{B}_1$ and $SE(\hat{B}_1)$ are taken from the 'survreg' output used. First, the confidence interval of $\hat{B}_1$ should be calculated as such.

$$L_2 = \hat{B}_1 - Z_{\alpha/2}SE(\hat{B}_1);\ U_2 = \hat{B}_1 + Z_{\alpha/2}SE(\hat{B}_1)$$

Confidence limits for $\hat{B}_1$ are $[e^{-U_2}, e^{-L_2}]$.

4. Bootstrap procedure

Bootstrapping is a technique used in inferential statistics that involves repeatedly creating random samples from a unique dataset. Every of these simulated samples have unique characteristics, such as the mean. We can see the sampling distribution of the mean when we graph the distribution of these means on a histogram. We don't need to be worried about test statistics, formulas, or assumptions.

4.1 The Bootstrap method

In general, we are dealing with a homogeneous simple random sample $X_1, \ldots, X_n$ drawn from a distribution with cumulative distribution function (CDF) is indicated by $F(x; \theta)$. The unknown parameter θ can be estimated using the given sample. Unfortunately, getting the distribution of this parameter's estimate might be exceedingly challenging in some situations. As a result, we are unable to determine the parameter's confidence interval. This issue can be solved using the bootstrap techniques, which will be covered later. This CDF can be computed non-parametrically using the empirical distribution function, $\hat{F}$. Using a collection of observation, $x_1, x_2 \ldots, x_n$, $\hat{F}$ is defined as the sample proportion provided below:

$$\hat{F}(x) = \frac{1}{n}\sum_{i=1}^{n} I(x_i \leq x)$$

With large values of n, $\hat{F} \approx F$

Considering a probability of interest where C is a fixed region and $g(\hat{\theta}, \theta)$ is a function of a random estimator $\hat{\theta}$and the true but unknown θ. For example, $g(\hat{\theta}, \theta) = \frac{(\hat{\theta} - \theta)}{SE(\hat{\theta})}$, which can be used to compute confidence intervals.

We now have $X_1^*,\ X_2^*, \ldots X_n^* \sim \hat{F}$, from the above given sample. In the same way as $P_F\ (g(\hat{\theta}, \theta)\ \in C)$ can be calculated in theory, so can $P_{\hat{F}}((g(\ \hat{\theta}^*, \hat{\theta})\ \in C)$; where $\hat{\theta}^*$ is the point estimate of θ from the bootstrap sample.

For large n, usually obtained as,

$$P_{\hat{F}}((g(\ \hat{\theta}^*, \hat{\theta})\ \in C) \approx P_F\ (g(\hat{\theta}, \theta)\ \in C)$$

Because all parameters in $P_{\hat{F}}((g(\ \hat{\theta}^*, \hat{\theta})\ \in C)$ are known, it can be calculated and utilized to obtain estimated confidence intervals.

4.2 Bootstrap Sampling

First, it is essential to mention that for the bootstrap confidence intervals, sampling with replacement from the original sample was performed B times, total number of bootstrap samples. Using the b[th] bootstrap sample $x_1^*,\ x_2^*, \ldots, x_n^*$, the estimated value of θ, is expressed as $\hat{\theta}^{(b)}$, where b = $1, \ldots, B$. This can be considered the same as choosing a simple random sample from $\hat{F}$. For B = +∞, the equivalent distribution is that of $P_{\hat{F}}((g(\ \hat{\theta}^*, \hat{\theta})\ \in C)$, which is the one of interest. It is important to note that in most cases, we cannot calculate $P_{\hat{F}}((g(\ \hat{\theta}^*, \hat{\theta})\ \in C)$, in closed form. $g(\hat{\theta}^*, \hat{\theta}) \in C$ can be determined by taking a large number of samples from $\hat{F}$ and then calculating the proportion of times $g(\ \hat{\theta}^*, \hat{\theta})\ \in C$ occurs.

4.3 Bootstrap confidence intervals

In practice, there are several ways to construct confidence intervals from the estimated sampling distribution. The method presented here is called the "percentile" method.

The bootstrap percentile interval

The distribution of $\hat{\theta}^{(b)}$ or $\hat{\theta}$ are generated using all the B bootstrap samples. Also, $c_{\frac{\propto}{2}}$ and $c_{1-\frac{\propto}{2}}$ are the $\frac{\propto}{2}$ and $1 - \frac{\propto}{2}$ quantiles of the $\hat{\theta}^{(b)}$ or $\hat{\theta}$ distribution respectively, The Percentile confidence interval is derived based on the distribution of $\hat{\theta}^{(b)}$ as $\left[c_{\frac{\propto}{2}}, c_{1-\frac{\propto}{2}}\right]$

5. Application to Diabetics Eye Patients data (DEP)

Diabetics Eye Patients data (DEP) contains 84 Diabetic Retinopathy patients' details. This data is collected from EYDOX hospital private limited, Chennai. The patients are followed from the admission to improvement of their vision level which is considered as event. This Study has conducted from January 2016 to February 2022. The patients those who are not attains improvement in their vision level during the study period are considered as censored. If anyone withdrawn from the treatment are also considered as censored.

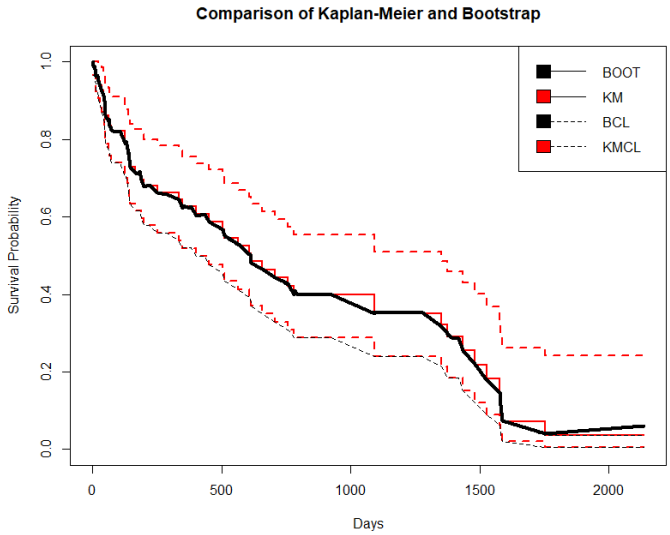
Table 2. Comparison of Kaplan-Meier and Bootstrapping survival probability for DEP data

Time	BSSP	2.50%	97.50%	BLCI	KM	2.50%	97.50%	KLCI
2	1.0000	1.0000	1.0000	0.0000	1.0000	1.0000	1.0000	0.0000
3	0.9875	0.9524	1.0000	0.0476	0.9880	0.9648	1.0000	0.0352
6	0.9886	0.9634	1.0000	0.0366	0.9880	0.9648	1.0000	0.0352
12	0.9748	0.9397	1.0000	0.0603	0.9758	0.9431	1.0000	0.0569
14	0.9637	0.9252	1.0000	0.0748	0.9636	0.9239	1.0000	0.0761
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
135	0.7777	0.6744	0.8622	0.1878	0.7764	0.6876	0.8767	0.1891
136	0.7754	0.6785	0.8701	0.1916	0.7764	0.6876	0.8767	0.1891
141	0.7617	0.6548	0.8514	0.1966	0.7609	0.6696	0.8646	0.1950
141	0.7659	0.6667	0.8538	0.1871	0.7609	0.6696	0.8646	0.1950
144	0.7462	0.6456	0.8390	0.1934	0.7450	0.6514	0.8521	0.2007
147	0.7287	0.6322	0.8317	0.1995	0.7292	0.6334	0.8394	0.2059
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
511	0.5507	0.4225	0.6717	0.2492	0.5459	0.4338	0.6871	0.2533
564	0.5286	0.3971	0.6485	0.2514	0.5257	0.4129	0.6693	0.2564
607	0.5020	0.3767	0.6267	0.2500	0.5055	0.3923	0.6513	0.2590
612	0.5036	0.3819	0.6344	0.2525	0.5055	0.3923	0.6513	0.2590
615	0.4805	0.3468	0.6087	0.2619	0.4844	0.371	0.6326	0.2616
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1526	0.1783	0.0638	0.3121	0.2483	0.1824	0.0904	0.3680	0.2776
1578	0.1467	0.0401	0.2662	0.2261	0.1459	0.0638	0.3338	0.2700
1579	0.1139	0.0000	0.2291	0.2291	0.1094	0.0402	0.2982	0.2580
1586	0.0737	0.0000	0.1673	0.1673	0.0730	0.0202	0.2631	0.2428
1754	0.0410	0.0000	0.1185	0.1185	0.0365	0.0055	0.2411	0.2355
2136	0.0599	0.0256	0.1349	0.1093	0.0365	0.0055	0.2411	0.2355

In the above table, we presented survival Probabilities of Kaplan-Meier and Bootstrap survival estimate. Kaplan-Meier estimates gives survival probabilities for Event observations only. On the other hand, Bootstrap survival estimates gives value for censored and uncensored cases.

The median survival time of KM and Bootstrap is 615 days. There are small changes in both methods. From the above table, it is observed that survival probabilities were observed for the observations lies between two event observations in Bootstrap method, but KM gives same probabilities for the observations lies between two event observations.

Figure 1. Comparison of KM and Bootstrap with confidence intervals for DEP data



Above Figure 1 indicates when compared to KM survival curve, Bootstrap Survival Curve is smoother and no jumps.

Table 3. Comparison of Nelson-Aalen and KM Bootstrapping survival probability for DEP data

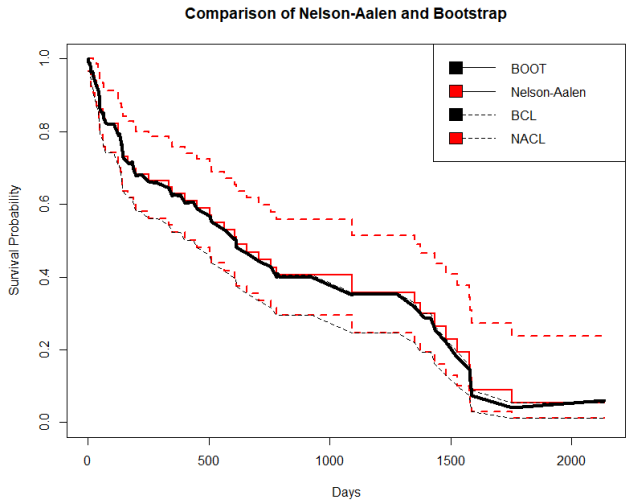
Time	BSSP	2.50%	97.50%	BLCI	Nelson-Aalen	2.50%	97.50%	NALCI
2	1.0000	1.0000	1.0000	0.0000	1.0000	1.0000	1.0000	0.0000
3	0.9875	0.9524	1.0000	0.0476	0.9880	0.9650	1.0000	0.0350
6	0.9886	0.9634	1.0000	0.0366	0.9880	0.9650	1.0000	0.0350
12	0.9748	0.9397	1.0000	0.0603	0.9759	0.9435	1.0000	0.0565
14	0.9637	0.9252	1.0000	0.0748	0.9638	0.9244	1.0000	0.0756
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
135	0.7777	0.6744	0.8622	0.1878	0.7779	0.6895	0.8776	0.1880
136	0.7754	0.6785	0.8701	0.1916	0.7779	0.6895	0.8776	0.1880
141	0.7617	0.6548	0.8514	0.1966	0.7625	0.6717	0.8656	0.1938
141	0.7659	0.6667	0.8538	0.1871	0.7625	0.6717	0.8656	0.1938
144	0.7462	0.6456	0.8390	0.1934	0.7468	0.6537	0.8532	0.1995
147	0.7287	0.6322	0.8317	0.1995	0.7311	0.6358	0.8405	0.2047
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
511	0.5507	0.4225	0.6717	0.2492	0.5497	0.4380	0.6898	0.2517
564	0.5286	0.3971	0.6485	0.2514	0.5297	0.4174	0.6723	0.2549
607	0.5020	0.3767	0.6267	0.2500	0.5097	0.3970	0.6545	0.2575
612	0.5036	0.3819	0.6344	0.2525	0.5097	0.3970	0.6545	0.2575
615	0.4805	0.3468	0.6087	0.2619	0.4889	0.3758	0.6360	0.2601
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1526	0.1783	0.0638	0.3121	0.2483	0.1940	0.1000	0.3760	0.2760
1578	0.1467	0.0401	0.2662	0.2261	0.1588	0.0736	0.3428	0.2692
1579	0.1139	0.0000	0.2291	0.2291	0.1237	0.0497	0.3079	0.2582
1586	0.0737	0.0000	0.1673	0.1673	0.0886	0.0289	0.2721	0.2433
1754	0.0410	0.0000	0.1185	0.1185	0.0537	0.0121	0.2384	0.2263
2136	0.0599	0.0256	0.1349	0.1093	0.0537	0.0121	0.2384	0.2263

In the above table, we presented survival Probabilities of Nelson-Aalen and with KM Bootstrap. Nelson-Aalen estimates gives Survival Probabilities for Event observations only. On the other hand, Bootstrap survival estimates give value for censored and uncensored cases.

The median survival time of NA and Bootstrap is 615 days. There are small changes in both methods. From the above table, it is observed that survival probabilities were observed for the observations lies between two event observations in Bootstrap method, but NA gives same probabilities for the observations lies between two event observations.

The above Table 2 and 3 shows, when time increases; length of the confidence intervals for the bootstrap is smaller than compared to other traditional methods.

Figure 2. Comparison of Nelson-Aalen with KM Bootstrap with confidence intervals for DEP data



Above Figure 2, indicates when compared to NA survival curve, Bootstrap Survival Curve is smoother and no jumps.

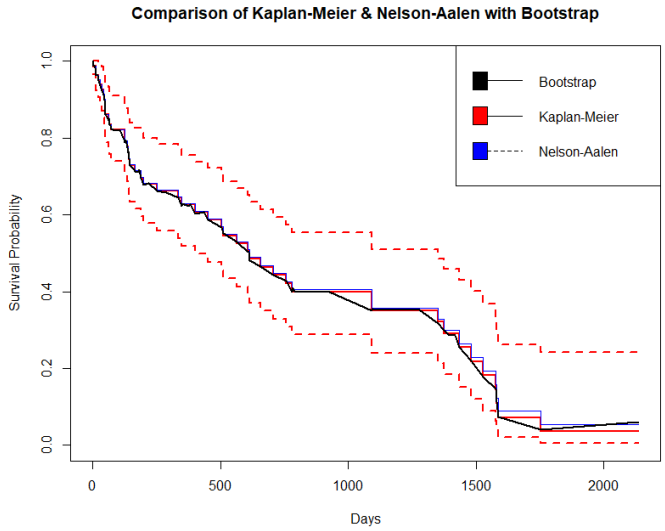
Table 4. Survival estimates of Kaplan-Meier and Nelson-Aalen with Bootstrapping survival probability for DEP data

Time	KM Bootstrap	Kaplan-Meier	Nelson-Aalen
2	1.0000	1.0000	1.0000
3	0.9875	0.9880	0.9880
6	0.9886	0.9880	0.9880
12	0.9748	0.9758	0.9759
14	0.9637	0.9636	0.9638
⋮	⋮	⋮	⋮
135	0.7777	0.7764	0.7779
136	0.7754	0.7764	0.7779
141	0.7617	0.7609	0.7625
141	0.7659	0.7609	0.7625
144	0.7462	0.7450	0.7468
147	0.7287	0.7292	0.7311
⋮	⋮	⋮	⋮
511	0.5507	0.5459	0.5497
564	0.5286	0.5257	0.5297

Time	KM Bootstrap	Kaplan-Meier	Nelson-Aalen
607	0.5020	0.5055	0.5097
612	0.5036	0.5055	0.5097
615	0.4805	0.4844	0.4889
⋮	⋮	⋮	⋮
1578	0.1467	0.1459	0.1588
1579	0.1139	0.1094	0.1237
1586	0.0737	0.0730	0.0886
1754	0.0410	0.0365	0.0537
2136	0.0599	0.0365	0.0537

It is observed that, the median survival time exists at 615 days in all the methods. KM Bootstrap survival estimates give Survival Probabilities for censored and uncensored cases for each time.

Figure 3. Survival estimates of KM and Nelson-Aalen with KM Bootstrap with confidence intervals for DEP data



Above Figure 3 indicates, when compared to KM and NA survival curve, Bootstrap Survival Curve is smoother and no jumps. Mean width of the confidence interval obtained from KM Bootstrap (0.1912) is less than mean width of the confidence interval obtained from KM (0.2020518) and NA (0.2007675).

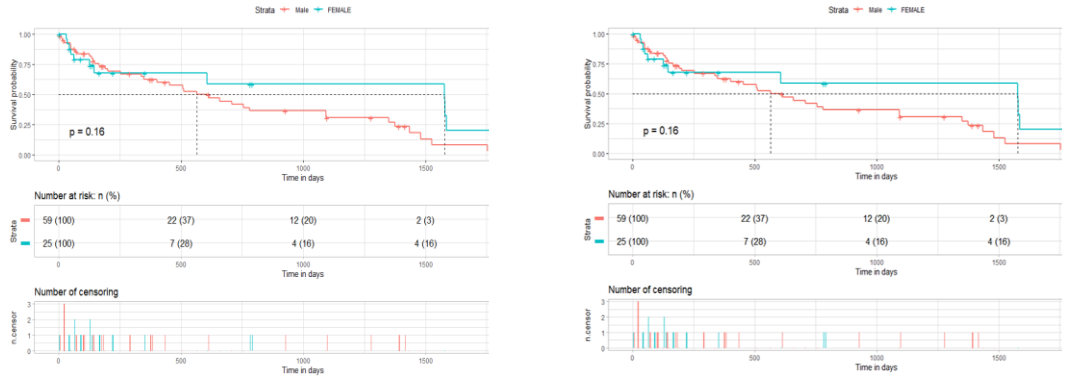
Table 5. Comparison of KM Bootstrapping with Kaplan-Meier and Nelson Aalen Estimate for different Quantiles to DEP Data.

Survival Probability	KM Bootstrap Survival Time					KM Survival Time	NA Survival Time
	Min	Q1	M	Q3	Max		
1.0	0	0	**0**	0	0	**0**	**0**
0.9	14	42	**49**	49	189	**49**	**49**
0.8	42	74	**125**	141	401	**125**	**125**
0.7	49	147	**189**	252	659	**189**	**189**
0.6	135	348	**453**	511	1092	**453**	**453**
0.5	189	564	**615**	707	1481	**615**	**615**
0.4	401	707	**780**	1092	1578	**780**	**780**
0.3	564	1350	**1375**	1481	1586	**1375**	**1375**
0.2	1092	1481	**1526**	1578	Inf	**1526**	**1526**

Median of Sampling distribution of Quantiles of KM bootstrap estimate is compared with Quantiles of KM and Nelson-Aalen estimate. Both approaches give the similar estimate.

Figure 4. KM and Nelson-Aalen Curve for Gender in Eye data with p-value

KM Nelson-Aalen



In the above Figure 4, A vertical drop and tick in the curves indicates an event and censored respectively. Risk and censoring table also attached in the figure.

Initially 59 male and 25 female patients are at risk. After 500 days, only 22 male and 7 female patients are in the risk set. This type of information is clearly shown in the above figure in all time points. The median survival time is approximately 564 days for Male and 1578 days for Female. P-values are also shown in the figure based on log-rank test. Since P-value is greater than 0.05, we cannot reject the null hypothesis at 5% of level of significance. So, this indicates that there is no significant difference between survival distributions of Gender in DEP data.

Table 6. Showing the p-value for the equality of survivorship function of Gender for DEP data for Nonparametric Tests.

Test	Weight	P value

Logrank	1	0.1608
Wilcoxon	N	0.8213
Tarone-Ware	$\sqrt{N}$	0.4761
Peto	S1	0.5171
	S2	0.5431
Fleming-Harrington	FHp = 0 & q = 1	0.0412
	FHp = 1 & q = 0	0.5020
	FHp = 1 & q = 1	0.0541
	FHp = 0.5 &q = 0.5	0.0941

From the above table, it is identified that all the nonparametric tests also confirm that the there is no significant difference in survival among Gender in DEP data.

The confidence intervals for the survival distribution of differences between the quantiles of different groups can be obtained using bootstrap. The following table shows 95% bootstrap confidence interval for the difference in some quantile survival time for Gender.

Table 7. Survival Probability, Lower and Upper Bootstrap confidence limits for the difference in Survival Probability of Gender.

S(t)	LCI	UCI
0.25	60	Inf
0.40	-485	Inf
0.50	-490	Inf
0.75	-286	1432
0.90	-97	111

From the above table, the value '0' lies between two confidence limits from the survival probability 1 to 0.4 but the value '0' not lies within the confidence limits for the survival probability below 0.4 because only few females' patients are attained event after median survival time. This is also observed from bootstrap. From nonparametric test, it is observed that there is no significant difference between survival distribution of male and female patients in DEP data.

Parametric models for Eye Patients data (DEP)

The relative risk is the point estimate of interest and is the value that shows the difference in time of improvement of their vision level between gender. Parametric models (Exponential, Weibull, Loglogistic and Lognormal) are fitted to DEP data. Comparison done based on the Parametric and Bootstrap Percentile confidence intervals. The bootstrap intervals will be calculated using B = 1, 000, B = 5, 000, and B = 10, 000 bootstrap. The 'survreg' function in R is used for this analysis.

In our case, the baseline is set as the Female, and the output relative risk values are $e^{0.2515} = 1.285953$ from Exponential, Weibull is $e^{0.2757} = 1.317453$, Loglogistic is $e^{0.0852} = 1.088935$ and $e^{0.1502} = 1.162067$ from lognormal. These values imply that lifetime is expected to be shorter for Female than that for Male in this study.

Table 8. Lower, Upper, width of the confidence interval and AIC values for Gender in DEP data is fitted using four Parametric model.

Models	LCI	UCI	WCI	AIC
Exponential	0.655997	2.520862	1.864865	743.973
Weibull	0.592355	2.930463	2.338108	743.541
Loglogistic	0.410011	2.892491	2.482480	749.461

Lognormal	0.453775	2.976363	2.522587	749.260

AIC value of Weibull is least, and Width of the Confidence Interval for Exponential model is the least one. But the difference between Exponential and Weibull AIC value is negligible.From Weibull, it is evident that the estimated median survival time for Male is 0.7589 times compared to Female. In DEP data, Female attains vision level improvement earlier than Male.

Table 9. Lower, Upper and Width of the confidence interval for Gender in DEP data using three different Bootstrap samples for four Parametric models using Percentile Method

Models	B	LCI	UCI	WCI
Exponential	1000	0.582167	3.062117	2.479950
	5000	0.576093	2.678516	2.102423
	10000	0.536829	2.503408	1.966579
Weibull	1000	0.549169	2.888155	2.338985
	5000	0.548328	2.971671	2.423343
	10000	0.559988	2.958302	2.398315
Loglogistic	1000	0.385169	3.319811	2.934642
	5000	0.367932	3.395938	3.028006
	10000	0.384196	3.394937	3.010742
Lognormal	1000	0.454089	3.256613	2.802524
	5000	0.447942	3.458827	3.010884
	10000	0.415976	3.319731	2.903756

Width of the Confidence Interval in Exponential models in Bootstrapping is also least one compared to others.

6. Summary

In this paper, distribution free survival methods and Parametric methods with Bootstrap are analyzed for DEP data. Using traditional method KM and NA, probability of survival for censored cases remains the level of previous time for uncensored data. Using bootstrap method, we obtain both censored and uncensored cases for each time. When compared to KM and Nelson-Aalen survival curve, KM Bootstrap Survival Curve is smoother and no jumps. Some width of the Bootstrap Confidence interval is less compared to KM confidence interval. Limits of the CI can be estimated without Standard error for Bootstrap. In the DEP data, when event time increases, the width of the bootstrap confidence interval is smaller compared to width of KM and Nelson-Aalen Confidence interval. By comparing the results of the nonparametric bootstrap methods with the DEP data, we proved that the strategy of nonparametric bootstrapping is best one. In calculation, Sampling distribution of Percentile of bootstrap KM and Nelson-Aalen estimate is compared with Percentile KM and Nelson-Aalen estimate. Both approaches give the similar estimates. The median survival time is approximately 564 days for Male and 1578 days for Female. Furthermore, in Parametric methods, width of the confidence interval for ratio between two parameters is least in Exponential and Bootstrapping. Based on this analysis, we can expect Exponential model to be a better fit for this data. In DEP data, Female attains vision level improvement earlier than Male.

References

1. A. Breslow; Thickness, cross-sectional areas and depth of invasion in the prognosis of cutaneous melanoma, Annals of surgery, 172(5), 902–908 (1970).
2. A.C. Davison and D.V. Hinkley; Bootstrap Methods and their Application, Cambridge University Press, Cambridge (1997).
3. A. Enrico Colosimo; F. Flavio Ferreira; D. Maristela Oliveira and B. Cleide Sousa; Empirical comparisons between Kaplan-Meier and Nelson-Aalen survival function estimators, Journal of Statistical Computation and Simulation, 72, 299–308 (2002). https://doi.org/10.1080/00949650212847.
4. B. Efron; The two Sample Problem with Censored Data, Proc. of the Fifth Berkeley Symp., 4, 831-853 (1967).
5. B. Efron; Censored data and the bootstrap, J. Am. Stat. Assoc., 76, 312–319 (1981). http://www.jstor.org/stable/2287832
6. B. Efron; R. Tibshiran; Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy, Statistical Science, 1, 54-77 (1986).
7. B. Efron; R. J. Tibshiran; An introduction to the bootstrap, Chapman & Hall, London (1993).
8. D. G. Kleinbaum and M. Klein; Survival Analysis, A Self-Learning Text, New York, Springer series (1996).

9. E. L. Kaplan; and P. Meier; Nonparametric estimation from incomplete observations, Journal of the American Statistical Association, 53, 457-81 (1958). https://doi.org/10.2307/2281868
10. E. A. Gehan; A generalized Wilcoxon test for comparing arbitrarily single-censored samples, Biometrika, 52, 203–223 (1965).
11. FatmaSarhan and Edward Susko; Bootstrapping and Survival Analysis (2020).
12. F. Utzet; Á. Sánchez; Some Applications of the Bootstrap to Survival Analysis, Anuario de Psicologi, 55, 155-167 (1992).
13. G. Gomez; M. L. Calle; R. Oller; K. Langohr; Tutorial on methods for interval-censored data and their implementation in R, Statistical Modelling, 9 (4), 259–297(2009).
14. I. Van Keilegom and N. Veraverbeke; Bootstrapping quantiles in a fixed design regression model with censored data, J. Stat.Plann. Inference, 69, 115–139 (1997).
15. J. P. Klein; J. D. Rizzo; M. J. Zhang; N. Keiding; Statistical methods for the analysis and presentation of the results of bone marrow transplants, Part I: Unadjusted analysis. Bone Marrow Transplantation, 28, 909–915 (2001).
16. K. Chen and S-H. Lo; On bootstrap accuracy with censored data, Ann. Stat., 24, 569–595 (1996). https://www.jstor.org/stable/2242662
17. M. Akri1as; Bootstrapping the Kaplan-Meier estimator, J Amer Statist.Assoc., 81, 1032-103 (1986). https://doi.org/10.1080/01621459.1986.10478369
18. M. Michalak; P. Marek; J. A. Moczko; Application Bootstrapping Kaplan-Meier Estimate For Survival Curve Smoothing, Computational Methods In Science And Technology, 8, 58-64(2002). https://doi.org/10.12921/cmst.2002.08.02.58-64
19. M. Ramakrishnan and R. Ravanan; Estimation of Survival Distribution using R Software, International Journal of Scientific and Research Publications, 3 (2013). http://www.ijsrp.org/research-paper-0413.php?rp=P161012
20. M. Ramakrishnan and R. Ravanan; Non-Parametric Methods for Comparing Two Survival Distributions, International Refereed Research Journal Researchers, 4(2013).
21. M. Stevenson; An Introduction to Survival Analysis, EpiCentre, IVABS. Massey Massey University(2009).
22. O. O. Aalen; Nonparametric Inference for a Family of Counting Processes, The Annals of Statistics, 6, 701-726 (1978). https://www.jstor.org/stable/2958850
23. P. D. Allison; Survival analysis using SAS: a practical guide, 2nd edition, SAS Press, North Carolina (2010).
24. R. Oller; G. Gomez; A generalized Fleming and Harrington's class of tests for interval-censored data, The Canadian Journal of Statistics, 40(3), 501–516(2012).
25. R. Peto; J. Peto; Asymptotically efficient rank invariant test procedures, Journal of the Royal Statistical Society, 135(2), 185–207 (1972).
26. R. L. Prentice; P. Marek; A qualitative discrepancy between censored data rank tests, Biometrics, 35(4), 861–867(1979).
27. R. E. Tarone; J. Ware; On distribution-free tests for equality of survival distributions, Biometrika, 64, 156–160(1977).
28. T. R. Fleming; D. P. Harrington; A class of hypothesis tests for one and two samples censored survival data, Communications in Statistics, 10, 763–794(1981).
29. Y. J. Cha & J. M. Lee; Comparison of Kaplan-Meier and Nelson Estimators using Bootstrap Confidence Interval (1995).