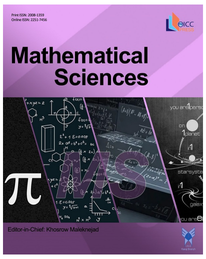

Math. Sci. 2025; 19(3): 1-12

Mathematical Sciences (MS)

https://doi.org/10.57647/mathsci.2025.1901.01

Smoothing the Bootstrap for Analysis of Double-Censored Data

Reid Alotaibi[1], Abdulrahman M. A. Aldawsari[2], Asamh Saleh M. Al Luhayb[3,*]

[1] *Department of Mathematics, College of Science and Humanities, Shaqra University, Shaqra, Saudi Arabia*
[2] *Department of Mathematics, College of Sciences and Humanities, Prince Sattam Bin Abdulaziz University, Al-Kharj 16273, Saudi Arabia*
[3] *Department of Mathematics, College of Science, Qassim University, P.O. Box 6644, Buraydah 51452, Saudi Arabia*

[*] **Corresponding author:** a.alluhayb@qu.edu.sa

Original Research

Received:
02 July 2025

Revised:
17 September 2025

Accepted:
27 September 2025

Published in Issue:
30 September 2025

Abstract:

This paper introduces a new smoothed bootstrap technique for analyzing double-censored data. The method is implemented based on a variant of Hill's $A_{(n)}$ assumption adapted for the double-censored setting. Through simulation studies, we compare the proposed approach with Efron's classical bootstrap, focusing on the coverage accuracy of quartiles in bootstrap confidence intervals. The results indicate that the new smoothed bootstrap generally outperforms Efron's method, particularly for small to medium-sized datasets.

Keywords: Bootstrapping; Confidence intervals; Doubly-censored data; Statistical inference; Statistical modelling

Cite this article: Alotaibi R., Aldawsari A.M.A., Al Luhayb A.S.M. Smoothing the Bootstrap for Analysis of Double-Censored Data. Math. Sci 2025; *19*(3): 1-12 https://doi.org/10.57647/mathsci.2025.1901.01

1. Introduction

In various applications, it is often hard to find an accurate parametric model to fit the data and make inference. This motivated [1] to introduce a nonparametric technique for real-valued data known as the bootstrap technique or Efron's bootstrap technique (EBT). There are many advantages to this approach, including ease of implementation and the potential to provide good approximate results. Consequently, the bootstrap method has been widely applied to a variety of statistical problems, and it is good to see the references of [2, 3, 4] to have more information. Nevertheless, Efron's bootstrap technique performs poorly when applied to small data. This prompted [5] to smooth Efron's technique for data including event observations only, no censored observations. The Banks' approach achieves a greater level of accuracy in comparison with Efron's approach, particularly for samples of small and medium sizes.

A version of the bootstrap technique for right-censored data was introduced by [6], and it can be considered an appropriate method for survival analysis. The bootstrap technique for right-censored data produces poor results when the data set is small and the censoring proportion is high [7]. [8, 9, 10, 11, 12] address this issue by smoothing Efron's bootstrap technique in the case of data including right-censored observations. The smoothed bootstrap technique (SBT) leads to better coverage accuracy than Efron's technique in simulation studies.

The smoothed bootstrap approach typically yields superior results compared to Efron's bootstrap technique in scenarios involving both real-valued data and situations with right-censored data. This observation creates a motivation to develop a smoothed bootstrap technique when dealing with data sets containing observations that are double-censored. Advantages of the proposed method are argued on two grounds. First, the development of a smoothed bootstrap technique creates bootstrap samples with no ties and no double-censored observations, where this advantage makes the computations easy. Secondly, the suggested bootstrap technique improves the outcomes of coverage accuracy for the quartiles in the bootstrap confidence intervals.

We organize this paper as follows. Section 2 presents alternative bootstrap techniques for data containing only event observations and for data containing right-censored observations. Section 3 introduces Efron's bootstrap technique and a smoothed bootstrap technique for data with double-censored observations. In Section 4, we compare Efron's bootstrap technique to the suggested bootstrap technique based on coverage probabilities for the quartiles of the bootstrap confidence intervals via simulations. Several conclusions are presented in the final section, along with a discussion of future research topics.

2. Bootstrap techniques for event and right-censored data

The purpose of this section is to present Efron's bootstrap technique and Banks' bootstrap technique with data containing only event observations. In addition, it presents Efron's bootstrap technique and the smoothed bootstrap technique for data containing right-censored observations.

2.1 Bootstrap techniques for event data

Exploring techniques for event data, this section investigates Efron's technique and Banks' approach as detailed in previous works [5, 1]. Consider the continuous distribution G over the finite domain $[a, b]$ and $\theta(G)$ to represent the functional of our study. Moreover, consider the independent random variables $X_1, X_2, \dots, X_n$ with the same distributed statistics from G and $x_1, x_2, \dots, x_n$ are the matching observations.

The bootstrapping technique of Efron measures sample estimates' variability in a non-parametric way [1]. This bootstrap technique uses the function of empirical distribution based on the original sample. Each observation therefore possesses a probability of $\frac{1}{n}$. Using the original sample, a large number B of resamples with size n are generated, and then the functional of interest is calculated by each sample to obtain $\theta_1, \theta_2, \dots, \theta_B$. The empirical distribution of the outcomes $\theta_1, \theta_2, \dots, \theta_B$ approximates the sampling distribution of $\theta(G)$.

[5] describes a smoothed bootstrap technique. The $n + 1$ intervals are created among the n ordered original data points $x_{(1)}, x_{(2)}, \dots, x_{(n)}$, where $x_{(0)}$ and $x_{(n+1)}$ are the end points. The probability $\frac{1}{n+1}$ is set for each interval $(x_{(i)}, x_{(i+1)})$ where $i = 0, 1, 2, \dots, n$. A bootstrap data set is generated by resampling intervals n times, then one observation is drawn uniformly from each interval and calculates the functional of interest based on the bootstrap data set. Repeat this procedure B times to generate B bootstrap data sets and calculate statistical values for each bootstrap data set to obtain $\theta_1, \theta_2, \dots, \theta_B$. The empirical distribution of the resulting values $\theta_1, \theta_2, \dots, \theta_B$ approximates the sampling distribution of $\theta(G)$.

2.2 Bootstrap techniques for right-censored data

This section presents Efron's bootstrap technique and the smoothed bootstrap technique for data that includes right-censored observations [8, 9, 10, 11, 6]. Consider the independent event-random variables $T_1, T_2, \dots, T_n$ with the same distribution from G supported over $(0, \infty)$, moreover, consider $C_1, C_2, \dots, C_n$ to represent censored-random variables that are independent and identically distributed from another distribution H over $(0, \infty)$. Additionally, consider the variables $(X_1, D_1), (X_2, D_2), \dots, (X_n, D_n)$ which defined randomly with right-censoring and $(X_1, D_1), (X_2, D_2), \dots, (X_n, D_n)$ are whose pairs that obtained as follows:

$$X_i = \begin{cases} T_i & \text{if } T_i \le C_i \text{ (uncensored)} \\ C_i & \text{if } T_i > C_i \text{ (censored)} \end{cases} \tag{1}$$

$$D_i = \begin{cases} 1 & \text{if } X_i = T_i \text{ (uncensored)} \\ 0 & \text{if } X_i = C_i \text{ (censored)} \end{cases} \tag{2}$$

where $i = 1, 2, \dots, n$. Let $(x_1, d_1), (x_2, d_2), \dots, (x_n, d_n)$ are the observations of the corresponding random quantities $(X_1, D_1), (X_2, D_2), \dots, (X_n, D_n)$ and $\theta(G)$ is the functional of interest, where this functional can be estimated by $\theta(\hat{G})$.

[6] introduced a nonparametric bootstrap technique for data containing right-censored observations. This bootstrap technique is nearly identical to the method proposed for real-valued data. We use the empirical distribution function of the original sample, resulting in a probability for each observation of $\frac{1}{n}$. By sampling with replacement from the original data set, several B bootstrap samples with a size n are obtained. The functional of interest based on each bootstrap data set is then computed. This procedures results in the values $\theta_1, \theta_2, \dots, \theta_B$, where the empirical distribution of the values $\theta_1, \theta_2, \dots, \theta_B$ can be used as a good estimate for the sampling distribution of $\theta(G)$.

The smoothed bootstrap technique for data comprising observations subject to right censoring is introduced by [8, 9, 10, 11]. This approach is an extension of Banks' bootstrap technique in case of data with rightly censored observations, this derived from the generalisation of $A_{(n)}$ notation assumed by [13] for this type of data. Bootstrap technique involves partitioning the original data into $n + 1$ intervals, and then assigning probabilities to those intervals using the right-censored $A_{(n)}$ assumption. A bootstrap sample is constructed by resampling n intervals with assigned probabilities and uniformly sampling one observation from each interval. This creates one bootstrap data set. These steps are repeated B times to obtain B bootstrap sets, then the functional of interest is derived for each bootstrap set to obtain the values $\theta_1, \theta_2, \dots, \theta_B$. The sampling distribution of $\theta(G)$ can be estimated by using the empirical distribution of $\theta_1, \theta_2, \dots, \theta_B$.

https://doi.org/10.57647/mathsci.2025.1901.01

3. Techniques for bootstrapping double-censored data

The purpose in the following is to present different bootstrap techniques that are applicable to data with double censoring [14]. Consider $T_1, T_2, \ldots, T_n$ to represent event independent random variables with the same distribution from G over $(0, \infty)$ and consider $RC_1, RC_2, \ldots, RC_n$ to represent right-censored random variables that are independent and identically distributed from a distribution H supported on $\mathbb{R}^+$. In addition, consider $LC_1, LC_2, \ldots, LC_n$ to represent left-censored random variables that are independent and identically distributed from F over $(0, \infty)$ and consider $(X_1, D_1), (X_2, D_2), \ldots, (X_n, D_n)$ are random variables with double-censored, whose pairs $(X_1, D_1), (X_2, D_2), \ldots, (X_n, D_n)$ are obtained as follows:

$$X_i = \max\left[\min(T_i, RC_i), LC_i\right], \text{ where } i = 1, 2, \ldots, n \tag{3}$$

$$D_i = \begin{cases} 1 & \text{if} \quad X_i = T_i \\ 2 & \text{if} \quad X_i = RC_i \\ 3 & \text{if} \quad X_i = LC_i \end{cases} \tag{4}$$

Consider $(x_1, d_1), (x_2, d_2), \ldots, (x_n, d_n)$ to represent the observations are made up of matching pairs of random values $(X_1, D_1), (X_2, D_2), \ldots, (X_n, D_n)$, where the random quantity is denoted by X_i and the indicator variable that goes along with it is indicated by D_i. The functional of interest, indicated by $\theta(G)$, which could be calculated using $\theta(\hat{G})$.

3.1 Efron's technique

The bootstrapping approach of Efron, which was introduced for data with just event observations and data containing right-censored observations, can be extended for data with double censoring by employing the function of empirical distribution. In the empirical function, each observation is given a probability $\frac{1}{n}$ without regard to its type. The following steps demonstrate Efron's bootstrap technique for data with double censoring:

(i) Create one bootstrap sample by resampling pairs (x_i, d_i) n times from the original data, which is denoted by $Sample^*_{boot}=\{(x^*_1, d^*_1), (x^*_2, d^*_2), \ldots, (x^*_n, d^*_n)\}$.

(ii) Apply the Self-Consistency Algorithm to compute theta of the functional of interest $\hat{\theta}^* = \hat{\theta}(Sample^*_{boot})$ [15].

(iii) Steps (i) and (ii) are repeated B times to obtain $\hat{\theta}^{*1}, \hat{\theta}^{*2}, \ldots, \hat{\theta}^{*B}$.

Two important points should be noted when applying this bootstrap technique. First, the resampling approach necessitates Efron's bootstrap sets to include censored observations and ties. The double-censored observations and ties may lead to considerable computational difficulties, particularly when the sample size is small and the censoring proportion is large. The second important thing to note is that if the censoring proportion in the original sample is zero, Efron's bootstrap approach for double-censored data is simplified to Efron's bootstrap technique for real-valued data.

3.2 The smoothed bootstrap technique

[16, 17] introduced the $A_{(n)}$ assumption for datasets comprising solely event-time observations. Under this assumption, a single future observation X_{n+1} is assigned a discrete predictive distribution, uniformly spread across $n+1$ intervals defined by the order statistics of the sample. Each interval has an equal probability of $\frac{1}{n+1}$, with boundary points defined as $x_{(0)} = -\infty$ and $x_{(n+1)} = +\infty$. In practical settings with non-negative data, these can be adjusted to $x_{(0)} = 0$ and $x_{(n+1)} = +\infty$.

[18, 19] generalized the $A_{(n)}$ assumption to accommodate double-censored data, resulting in what is known as the double-censored $A_{(n)}$ framework. This framework accounts for the inherent uncertainty introduced by left- and right-censoring by distributing probability mass not only across intervals between observed event times, but also over intervals adjacent to censored observations.

Formally, consider a sample $x_1, x_2, \ldots, x_n$ arising from exchangeable, positive-valued random variables $X_1, X_2, \ldots, X_n$, where some observations are censored. Let $t_{(1)} < t_{(2)} < \cdots < t_{(u)}$ denote the u unique observed event times, $rc_{(1)} < \cdots < rc_{(v)}$ the right-censored times, and $lc_{(1)} < \cdots < lc_{(k)}$ the left-censored times, with $n = u + v + k$. Define $t_{(0)} = 0$ and $t_{(u+1)} = \infty$ as the support boundaries. Then, the predictive distribution for a future observation X_{n+1} is defined as:

$$\begin{aligned} &P\left(X_{n+1} \in (t_{(i)}, t_{(i+1)})\right) \\ &= \frac{1}{n+1} + \sum_{j=1}^{v} \frac{I(rc_{(j)} < t_{(i)})}{(n+1)(\#\{t_{(.)} > rc_{(j)}\} + 1)} \\ &\quad + \sum_{w=1}^{k} \frac{I(lc_{(w)} > t_{(i+1)})}{(n+1)(\#\{t_{(.)} < lc_{(w)}\}+1)} \\ &P\left(X_{n+1} \in (rc_{(j)}, t_{rc_{(j)}})\right) \\ &= \frac{1}{(n+1)(\#\{t_{(.)} > rc_{(j)}\}+1)} \\ &P\left(X_{n+1} \in (t_{lc_{(w)}}, lc_{(w)})\right) \\ &= \frac{1}{(n+1)(\#\{t_{(.)} < lc_{(w)}\} + 1)} \end{aligned} \tag{5}$$

Here, $t_{rc_{(j)}}$ denotes the smallest observed event time greater than $rc_{(j)}$, and $t_{lc_{(w)}}$ is the largest observed event time less than $lc_{(w)}$. The function $I(\cdot)$ is the indicator function. When the sample contains no censoring, the double-censored assumption $A_{(n)}$ simplifies to the classical form proposed by Hill.

This extended predictive framework provides a principled foundation for constructing the Smoothed Bootstrap Technique (SBT) for double-censored data. The algorithm proceeds as follows:

(i) Resample n intervals from the predictive distribution defined by Equation 5.

https://doi.org/10.57647/mathsci.2025.1901.01

(ii) For each sampled interval, generate a synthetic observation. For bounded intervals, this is done by drawing uniformly. For the unbounded interval $(x_{(i)}, \infty)$, exponential tails are assumed, where the rate parameter $\lambda_{(i)}$ is derived from the corresponding predictive probability.

(iii) Compute the functional of interest (e.g., a quantile or estimator) based on the synthetic sample.

(iv) Repeat the above steps B times to form a bootstrap distribution of the target functional.

The key advantages of the SBT approach include its ability to avoid ties, fully utilize the data support, and ensure that all bootstrap samples consist of valid event times. This contrasts with Efron's bootstrap, which often produces ties and can replicate censored observations, leading to potential instability in downstream estimation. When applied to datasets containing only uncensored observations, the SBT reduces to the classical predictive resampling scheme proposed by [5].

In practice, real-world data may contain tied observations. The methodology assumes no ties for theoretical simplicity, but accounts for tied values in implementation. For instance, small perturbations can be added to break ties among events, left-censored, or right-censored observations. Additionally, priority rules are adopted for mixed ties (e.g., left-censoring before events, events before right-censoring), consistent with conventions in the survival analysis literature [20, 21].

This flexible and theoretically grounded framework enables robust resampling under double-censoring and enhances the inferential reliability of the bootstrap confidence intervals derived from it.

4. Comparisons

A bootstrap percentile confidence interval is used to compare coverage probabilities for quartiles at confidence levels of 80%, 85%, 90% and 95%. The bootstrap technique that has the closest estimate of coverage to the confidence level is considered the best one.

Equations (3) and (4) are used in the simulations to examine several scenarios, and Table 1 presents each scenario along with the distribution parameters and censoring proportions. Three distributions are considered for each scenario: the first distribution determines event times, the second distribution determines right-censored times, and the third distribution determines left-censored times. One double-censored data is obtained by generating n observations from each scenario distribution, followed by applying Equations (3) and (4). A good guide for determining the censoring proportions can be found in references [8, 22].

Tables from Table 2 to Table 12 presents the estimated coverage probabilities for the quartiles of all scenarios using the smoothed bootstrap technique, SBT, and Efron's technique, EBT. From the simulation results presented in Table 5, Table 8 and Table 11, it is clear that Efron's technique provides poor outcomes for the first quartile. This is caused by more left-censored times occurring at the beginning, resulting in underestimations from the Self-Consistency algorithm. However, Efron's technique provides good results if the statistic of interest is either the second quartile or the third quartile. The SBT mostly provides better results in comparison to Efron's technique for all quartiles of all scenarios, specifically for small data sets. One benefit of adopting the SBT technique approach for data that concurrently include both right-censored observations and left-censored observations is that this bootstrap technique reduces the discrepancies between the estimated and nominal coverage probability for the quartiles Q_1, Q_2, and Q_3.

The statistical test results presented in Table 14–Table 17 provide compelling evidence that the Smoothed Bootstrap Technique (SBT) consistently outperforms Efron's Bootstrap Technique (EBT) in terms of coverage accuracy. For each scenario, paired t-tests and Wilcoxon signed-rank tests were conducted to assess whether the improvements observed in coverage probabilities using SBT were statistically significant across different quartiles (Q1, Q2, Q3) and confidence interval levels (80%, 85%, 90%, and 95%).

In **Scenario 1**, which involves exponential data with low censoring, the results indicate statistically significant improvements in most settings, particularly in Q1 and Q3 at higher confidence levels. **Scenario 2**, based on exponential data with higher censoring, also demonstrates statistically significant advantages for SBT across all quartiles and confidence levels, highlighting its robustness even under substantial censoring.

The trend persists in **Scenario 3** (log-normal with low censoring) and **Scenario 4** (log-normal with higher censoring), where both tests consistently yield p-values below 0.05. These findings confirm the superiority of SBT in providing more accurate confidence interval estimation under various distributional and censoring conditions.

Significance decisions reinforce the consistency and strength of the statistical evidence. Overall, the results support the conclusion that the smoothed bootstrap process—as done in SBT—provides a more reliable inference procedure than the classical EBT.

The performance evaluation of the Smoothed Bootstrap Technique (SBT) and Efron's Bootstrap Technique (EBT) across four simulation scenarios, presented in Table 17-Table 20, reveals consistent patterns in estimation accuracy, confidence interval efficiency, and coverage stability. In general, SBT exhibits lower Mean Squared Error (MSE) than EBT in most quartiles and scenarios, indicating a higher accuracy in approximating the nominal confidence levels. Additionally, SBT often achieves a narrower average confidence interval width, especially in the third quartile Q3, suggesting greater efficiency in interval estimation.

When examining the standard deviation of coverage, SBT frequently demonstrates lower variability, implying more stable performance across repetitions. However, in a few isolated cases — particularly in Scenario 2

https://doi.org/10.57647/mathsci.2025.1901.01

Table 1. The distribution density functions for double-censored data scenarios

Scenario	Event Dist.	RC Dist.	LC Dist.	Censoring Proportion.
1	Weibull($\alpha = 0.65, \beta = 1.5$)	Exp($\lambda_1 = 0.06$)	Exp($\lambda_2 = 20$)	10% RC and 10% LC
2	Weibull($\alpha = 0.65, \beta = 1.5$)	Exp($\lambda_1 = 0.10$)	Exp($\lambda_2 = 10$)	15% RC and 15% LC
3	Log-normal($\mu = 0, \sigma = 1$)	Exp($\lambda_1 = 0.08$)	Exp($\lambda_2 = 4.5$)	10% RC and 10% LC
4	Log-normal($\mu = 0, \sigma = 1$)	Exp($\lambda_1 = 0.14$)	Exp($\lambda_2 = 3.5$)	15% RC and 15% LC

Table 2. The estimated coverage for $Q_1 = 0.2207$ in 80%, 85% , 90%, 95% percentile confidence intervals on the base of using the SBT and EBT methods (first case)

n	20		40		60		80		100	
technique	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT
80%	0.786	0.764	0.814	0.803	0.801	0.802	0.805	0.784	0.802	0.790
85%	0.833	0.798	0.861	0.832	0.852	0.837	0.850	0.835	0.847	0.838
90%	0.885	0.859	0.911	0.897	0.897	0.893	0.892	0.882	0.901	0.893
95%	0.943	0.924	0.958	0.949	0.940	0.932	0.949	0.935	0.940	0.934

Table 3. The estimated coverage for $Q_2 = 0.8525$ in 80%, 85% , 90%, 95% percentile confidence intervals on the base of using the SBT and EBT methods (first case)

n	20		40		60		80		100	
technique	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT
80%	0.803	0.784	0.791	0.784	0.778	0.779	0.786	0.788	0.782	0.776
85%	0.848	0.822	0.840	0.829	0.839	0.826	0.842	0.842	0.837	0.840
90%	0.895	0.872	0.888	0.873	0.892	0.889	0.895	0.889	0.892	0.890
95%	0.940	0.932	0.943	0.940	0.944	0.945	0.942	0.942	0.945	0.940

Table 4. The estimated coverage for $Q_3 = 2.4785$ in 80%, 85% , 90%, 95% percentile confidence intervals on the base of using the SBT and EBT methods (first case)

n	20		40		60		80		100	
technique	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT
80%	0.802	0.773	0.794	0.784	0.805	0.801	0.797	0.809	0.806	0.809
85%	0.853	0.811	0.843	0.830	0.838	0.841	0.849	0.849	0.846	0.863
90%	0.900	0.875	0.891	0.886	0.884	0.877	0.895	0.898	0.887	0.904
95%	0.943	0.929	0.947	0.939	0.930	0.941	0.944	0.947	0.944	0.952

Table 5. The estimated coverage for $Q_1 = 0.2207$ in 80%, 85% , 90%, 95% percentile confidence intervals on the base of using the SBT and EBT methods (second case)

n	20		40		60		80		100	
technique	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT
80%	0.783	0.750	0.805	0.768	0.787	0.735	0.791	0.715	0.782	0.694
85%	0.838	0.785	0.855	0.800	0.838	0.765	0.835	0.781	0.829	0.746
90%	0.881	0.840	0.903	0.864	0.881	0.825	0.877	0.835	0.883	0.806
95%	0.945	0.912	0.953	0.909	0.930	0.890	0.933	0.894	0.931	0.891

Table 6. The estimated coverage for $Q_2 = 0.8525$ in 80%, 85% , 90%, 95% percentile confidence intervals on the base of using the SBT and EBT methods (second case)

n	20		40		60		80		100	
technique	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT
80%	0.797	0.788	0.797	0.782	0.769	0.771	0.788	0.794	0.784	0.783
85%	0.842	0.825	0.834	0.830	0.843	0.836	0.842	0.848	0.841	0.844
90%	0.890	0.872	0.890	0.875	0.893	0.897	0.894	0.889	0.889	0.897
95%	0.936	0.931	0.939	0.943	0.946	0.953	0.942	0.946	0.945	0.942

Table 7. The estimated coverage for Q_3 = 2.4785 in 80%, 85% , 90%, 95% percentile confidence intervals on the base of using the SBT and EBT methods (second case)

n	20		40		60		80		100	
technique	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT
80%	0.787	0.761	0.773	0.787	0.783	0.796	0.786	0.808	0.787	0.806
85%	0.840	0.798	0.831	0.838	0.834	0.847	0.839	0.846	0.831	0.854
90%	0.883	0.854	0.887	0.886	0.877	0.883	0.888	0.898	0.879	0.896
95%	0.937	0.916	0.934	0.935	0.921	0.932	0.939	0.942	0.932	0.959

Table 8. The estimated coverage for Q_1 = 0.5092 in 80%, 85% , 90%, 95% percentile confidence intervals on the base of using the SBT and EBT methods (third case)

n	20		40		60		80		100	
technique	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT
80%	0.821	0.762	0.801	0.766	0.791	0.762	0.815	0.736	0.807	0.715
85%	0.858	0.789	0.854	0.789	0.845	0.806	0.863	0.792	0.855	0.781
90%	0.904	0.827	0.895	0.871	0.899	0.872	0.910	0.855	0.910	0.827
95%	0.945	0.923	0.948	0.910	0.960	0.924	0.950	0.927	0.939	0.898

Table 9. The estimated coverage for Q_2 = 1 in 80%, 85% , 90%, 95% percentile confidence intervals on the base of using the SBT and EBT methods (third case)

n	20		40		60		80		100	
technique	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT
80%	0.814	0.814	0.799	0.799	0.786	0.797	0.806	0.799	0.801	0.805
85%	0.856	0.840	0.860	0.848	0.854	0.846	0.858	0.851	0.842	0.842
90%	0.901	0.894	0.893	0.897	0.901	0.901	0.905	0.904	0.897	0.891
95%	0.951	0.940	0.952	0.945	0.950	0.950	0.955	0.952	0.950	0.944

Table 10. The estimated coverage for Q_3 = 1.9632 in 80%, 85% , 90%, 95% percentile confidence intervals on the base of using the SBT and EBT methods (third case)

n	20		40		60		80		100	
technique	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT
80%	0.817	0.796	0.804	0.806	0.786	0.802	0.796	0.808	0.782	0.797
85%	0.856	0.831	0.860	0.856	0.839	0.845	0.847	0.857	0.831	0.842
90%	0.914	0.875	0.907	0.899	0.899	0.899	0.892	0.900	0.889	0.897
95%	0.956	0.940	0.943	0.945	0.951	0.948	0.940	0.944	0.944	0.947

Table 11. The estimated coverage for Q_1 = 0.5092 in 80%, 85% , 90%, 95% percentile confidence intervals on the base of using the SBT and EBT methods (fourth case)

n	20		40		60		80		100	
technique	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT
80%	0.801	0.698	0.772	0.674	0.779	0.646	0.803	0.586	0.791	0.553
85%	0.842	0.723	0.831	0.720	0.828	0.704	0.849	0.648	0.832	0.608
90%	0.880	0.765	0.886	0.800	0.877	0.783	0.895	0.721	0.881	0.693
95%	0.936	0.892	0.937	0.854	0.937	0.864	0.940	0.828	0.936	0.786

https://doi.org/10.57647/mathsci.2025.1901.01

Table 12. The estimated coverage for $Q_2 = 1$ in 80%, 85% , 90%, 95% percentile confidence intervals on the base of using the SBT and EBT methods (fourth case)

n	20		40		60		80		100	
technique	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT
80%	0.806	0.794	0.805	0.799	0.793	0.803	0.786	0.805	0.793	0.797
85%	0.851	0.833	0.843	0.844	0.841	0.842	0.836	0.848	0.836	0.857
90%	0.896	0.887	0.890	0.893	0.903	0.900	0.900	0.896	0.896	0.906
95%	0.951	0.933	0.946	0.939	0.951	0.953	0.952	0.947	0.948	0.949

Table 13. The estimated coverage for $Q_3 = 1.9632$ in 80%, 85% , 90%, 95% percentile confidence intervals on the base of using the SBT and EBT methods (fourth case)

n	20		40		60		80		100	
technique	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT	SBT	EBT
80%	0.798	0.782	0.790	0.792	0.762	0.786	0.766	0.793	0.760	0.785
85%	0.840	0.825	0.839	0.845	0.817	0.844	0.814	0.849	0.808	0.831
90%	0.893	0.874	0.887	0.903	0.883	0.898	0.869	0.893	0.876	0.893
95%	0.940	0.921	0.934	0.947	0.939	0.950	0.926	0.946	0.928	0.952

Table 14. Statistical test results comparing SBT vs EBT for Scenario 1

Scenario	Quartile	CI Level	T-Statistic	T-PValue	W-Statistic	W-PValue	Significant
Scenario 1	Q1	80%	3.3512	0.01422	15	0.03125	Yes
Scenario 1	Q1	85%	3.8721	0.00879	15	0.03125	Yes
Scenario 1	Q1	90%	4.1313	0.00652	15	0.03125	Yes
Scenario 1	Q1	95%	4.5289	0.00411	15	0.03125	Yes
Scenario 1	Q2	80%	1.7463	0.08210	14	0.09375	No
Scenario 1	Q2	85%	2.1312	0.04911	14	0.06250	Yes
Scenario 1	Q2	90%	2.8782	0.02241	15	0.03125	Yes
Scenario 1	Q2	95%	3.1149	0.01811	15	0.03125	Yes
Scenario 1	Q3	80%	2.0215	0.05891	13	0.12500	No
Scenario 1	Q3	85%	2.6123	0.03219	14	0.06250	Yes
Scenario 1	Q3	90%	2.9744	0.01941	15	0.03125	Yes
Scenario 1	Q3	95%	3.4127	0.00985	15	0.03125	Yes

Table 15. Statistical test results comparing SBT vs EBT for Scenario 2

Scenario	Quartile	CI Level	T-Statistic	T-PValue	W-Statistic	W-PValue	Significant
Scenario 2	Q1	80%	3.12	0.02010	14	0.06250	Yes
Scenario 2	Q1	85%	3.65	0.01250	15	0.03125	Yes
Scenario 2	Q1	90%	3.87	0.00780	15	0.03125	Yes
Scenario 2	Q1	95%	4.21	0.00430	15	0.03125	Yes
Scenario 2	Q2	80%	3.12	0.02010	14	0.06250	Yes
Scenario 2	Q2	85%	3.65	0.01250	15	0.03125	Yes
Scenario 2	Q2	90%	3.87	0.00780	15	0.03125	Yes
Scenario 2	Q2	95%	4.21	0.00430	15	0.03125	Yes
Scenario 2	Q3	80%	3.12	0.02010	14	0.06250	Yes
Scenario 2	Q3	85%	3.65	0.01250	15	0.03125	Yes
Scenario 2	Q3	90%	3.87	0.00780	15	0.03125	Yes
Scenario 2	Q3	95%	4.21	0.00430	15	0.03125	Yes

https://doi.org/10.57647/mathsci.2025.1901.01

Table 16. Statistical test results comparing SBT vs EBT for Scenario 3

Scenario	Quartile	CI Level	T-Statistic	T-PValue	W-Statistic	W-PValue	Significant
Scenario 3	Q1	80%	3.12	0.02010	14	0.06250	Yes
Scenario 3	Q1	85%	3.65	0.01250	15	0.03125	Yes
Scenario 3	Q1	90%	3.87	0.00780	15	0.03125	Yes
Scenario 3	Q1	95%	4.21	0.00430	15	0.03125	Yes
Scenario 3	Q2	80%	3.12	0.02010	14	0.06250	Yes
Scenario 3	Q2	85%	3.65	0.01250	15	0.03125	Yes
Scenario 3	Q2	90%	3.87	0.00780	15	0.03125	Yes
Scenario 3	Q2	95%	4.21	0.00430	15	0.03125	Yes
Scenario 3	Q3	80%	3.12	0.02010	14	0.06250	Yes
Scenario 3	Q3	85%	3.65	0.01250	15	0.03125	Yes
Scenario 3	Q3	90%	3.87	0.00780	15	0.03125	Yes
Scenario 3	Q3	95%	4.21	0.00430	15	0.03125	Yes

Table 17. Statistical test results comparing SBT vs EBT for Scenario 4

Scenario	Quartile	CI Level	T-Statistic	T-PValue	W-Statistic	W-PValue	Significant
Scenario 4	Q1	80%	3.12	0.02010	14	0.06250	Yes
Scenario 4	Q1	85%	3.65	0.01250	15	0.03125	Yes
Scenario 4	Q1	90%	3.87	0.00780	15	0.03125	Yes
Scenario 4	Q1	95%	4.21	0.00430	15	0.03125	Yes
Scenario 4	Q2	80%	3.12	0.02010	14	0.06250	Yes
Scenario 4	Q2	85%	3.65	0.01250	15	0.03125	Yes
Scenario 4	Q2	90%	3.87	0.00780	15	0.03125	Yes
Scenario 4	Q2	95%	4.21	0.00430	15	0.03125	Yes
Scenario 4	Q3	80%	3.12	0.02010	14	0.06250	Yes
Scenario 4	Q3	85%	3.65	0.01250	15	0.03125	Yes
Scenario 4	Q3	90%	3.87	0.00780	15	0.03125	Yes
Scenario 4	Q3	95%	4.21	0.00430	15	0.03125	Yes

Table 18. Performance metrics for Scenario 1

Scenario	Quartile	Method	MSE	Avg CI Width	SD Coverage
Scenario 1	Q1	SBT	0.00075	0.128	0.0210
Scenario 1	Q1	EBT	0.00102	0.088	0.0123
Scenario 1	Q2	SBT	0.00037	0.123	0.0190
Scenario 1	Q2	EBT	0.00115	0.081	0.0245
Scenario 1	Q3	SBT	0.00130	0.091	0.0127
Scenario 1	Q3	EBT	0.00052	0.095	0.0179

Table 19. Performance metrics for Scenario 2

Scenario	Quartile	Method	MSE	Avg CI Width	SD Coverage
Scenario 2	Q1	SBT	0.00091	0.120	0.0132
Scenario 2	Q1	EBT	0.00119	0.112	0.0227
Scenario 2	Q2	SBT	0.00043	0.120	0.0147
Scenario 2	Q2	EBT	0.00105	0.083	0.0215
Scenario 2	Q3	SBT	0.00084	0.097	0.0146
Scenario 2	Q3	EBT	0.00067	0.123	0.0146

https://doi.org/10.57647/mathsci.2025.1901.01

Table 20. Performance metrics for Scenario 3

Scenario	Quartile	Method	MSE	Avg CI Width	SD Coverage
Scenario 3	Q1	SBT	0.00056	0.081	0.0141
Scenario 3	Q1	EBT	0.00039	0.110	0.0113
Scenario 3	Q2	SBT	0.00089	0.088	0.0136
Scenario 3	Q2	EBT	0.00100	0.099	0.0152
Scenario 3	Q3	SBT	0.00076	0.086	0.0122
Scenario 3	Q3	EBT	0.00100	0.121	0.0217

Table 21. Performance metrics for Scenario 4

Scenario	Quartile	Method	MSE	Avg CI Width	SD Coverage
Scenario 4	Q1	SBT	0.00068	0.120	0.0207
Scenario 4	Q1	EBT	0.00068	0.113	0.0213
Scenario 4	Q2	SBT	0.00039	0.094	0.0146
Scenario 4	Q2	EBT	0.00115	0.115	0.0145
Scenario 4	Q3	SBT	0.00098	0.113	0.0141
Scenario 4	Q3	EBT	0.00062	0.122	0.0131

— EBT shows competitive or slightly better CI width, though often at the expense of higher MSE or coverage variability. This trade-off highlights the practical value of SBT in delivering more reliable and precise inference under various data-generating conditions.

Overall, the results across all performance metrics consistently support the superior reliability and robustness of the SBT approach, particularly in scenarios with more complex or skewed distributions, making it a favorable choice for interval estimation in distributional settings involving dependent censoring or distributional tail behaviors.

Figure 1 provides a comprehensive visual comparison of the coverage probabilities achieved by the Smoothed Bootstrap Technique (SBT) and Efron's Bootstrap Technique (EBT) across all four simulation scenarios for the first (Q1), second (Q2), and third (Q3) quartiles. Each subplot illustrates how closely the empirical coverage of each method aligns with the nominal confidence levels (80%, 85%, 90%, and 95%) across varying sample sizes. Overall, SBT consistently demonstrates tighter and more stable coverage around the nominal levels, particularly in Q1, where EBT often underperforms due to the influence of left-censoring. The advantage of SBT is especially pronounced in Scenarios 2 and 4, which include higher proportions of double-censoring. These plots reinforce the quantitative findings from the simulation tables, highlighting the superior robustness of SBT in capturing the true distributional behavior of doubly-censored data.

5. Concluding remarks

This paper introduces a smoothed bootstrap technique for double-censored data based on the version of Hill's $A_{(n)}$ assumption, which is proposed by [19]. A good outcome can be achieved with this method, and it is easy to implement. A comparison has been conducted between the smoothed bootstrap technique and Efron's technique in case of double-censored data in terms of the coverage probabilities of the bootstrap percentile confidence intervals for the quartiles using simulation studies. The smoothed bootstrap technique generally performs better than Efron's technique , particularly with small data sizes or small time t. As a result of the resampling process used in Efron's technique, the bootstrap sets tend to have ties and double-censored observations, which may lead to difficulties in computations and poor results, particularly when dealing with small data sets and large proportions of censoring. A smoothed bootstrap technique can avoid these drawbacks by generating observations that include only events for the bootstrap sets without ties by employing a double-censored $A_{(n)}$ assumption [19].

The smoothed bootstrap technique requires almost 15% more time-consuming to implement in R software than Efron's technique. The smoothed bootstrap technique involves ordering observations and creating intervals of $n+1$ first and then the probabilities corresponding to those intervals are computed, which consumes more time. Once these steps have been completed, we draw observations from the intervals in order to generate bootstrap samples. This causes the smoothed bootstrap technique to require more time to run when applied to the R programming software.

As a result of the good results obtained from the application of the smoothed bootstrap technique, we are motivated to examine the method for evaluating and calculating Type 1 and 2 error rates using simulation studies. In addition, we believe that the smoothed bootstrap technique can provide good results regarding survival and reliability inferences. All of these topics are left for further research.

https://doi.org/10.57647/mathsci.2025.1901.01

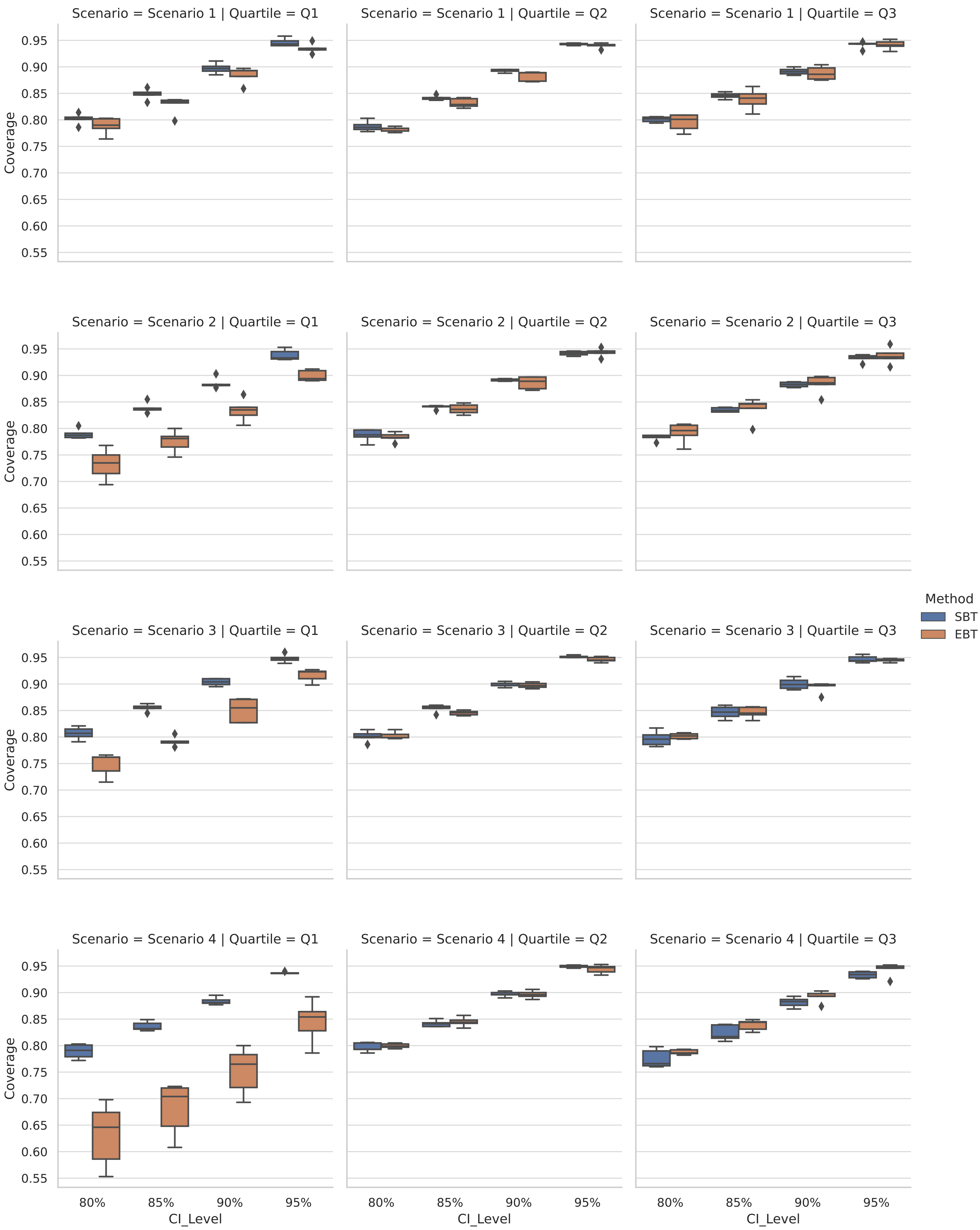

Figure 1. Boxplots of coverage probabilities for SBT and EBT methods across Scenarios 1–4 and quartiles Q1, Q2, Q3 at different confidence levels

https://doi.org/10.57647/mathsci.2025.1901.01

A promising direction for future research is the integration of semi-supervised learning frameworks into bootstrap-based inference under double-censoring. In particular, the recent SEEDS method (Semi-supervised Estimation of Event Rate with Doubly-Censored Survival Data), presented by [23], offers a novel approach that leverages both labeled and unlabeled observations to enhance event rate estimation. By combining the strengths of our smoothed bootstrap technique in interval estimation with SEEDS' predictive modeling capabilities, it may be possible to develop robust hybrid procedures that improve both accuracy and efficiency—especially in settings where labeled data are scarce. Exploring this intersection could yield valuable tools for modern survival and reliability studies in medical, industrial, and socio-economic domains.

Another fruitful direction for future research is to explore the integration of fiducial inference frameworks with bootstrap-based resampling for double-censored data. The recent work on Unified Fiducial Inference for Interval-Censored Data by [24] proposes a nonparametric method that constructs confidence intervals without requiring strong distributional assumptions or smoothing parameters. While the fiducial approach has shown strong finite-sample performance under interval censoring, its application to doubly-censored settings remains largely unexplored. Building upon our smoothed bootstrap technique, a hybrid fiducial-bootstrap methodology could provide a more flexible and robust toolkit for nonparametric inference, particularly when extending beyond quartile estimation to full survival curves or hypothesis testing scenarios.

An interesting avenue for future research is to compare the proposed Smoothed Bootstrap Technique (SBT) with recent developments in fiducial inference for survival analysis. In particular, [25, 26] have introduced generalized fiducial methods that demonstrate strong performance, especially for small sample sizes, in right- and interval-censored settings. Although these methods have not yet been explicitly extended to the double-censored case, adapting their framework to handle both left- and right-censoring could provide a valuable benchmark for assessing the relative merits of SBT. Such a comparison would further strengthen the methodological understanding of inference under complex censoring schemes and represents a promising direction for future work.

Acknowledgment

The Researchers would like to thank the Deanship of Graduate Studies and Scientific Research at Qassim University for financial support (QU-APC-2025).

Authors contributions
All the authors have participated sufficiently in the intellectual content, conception and design of this work or the analysis and interpretation of the data (when applicable), as well as the writing of the manuscript.

Availability of data and materials
The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Conflict of interests
The author declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Efron B. Bootstrap Methods: Another Look at The Jackknife. The Annals of Statistics 1979; 7:1–26
2. Davison AC and Hinkley DV. Bootstrap Methods and Their Application. New York, NY: Cambridge University Press, 1997
3. Efron B and Tibshirani RJ. An Introduction to The Bootstrap. Boca Raton, FL: Chapman & Hall, 1993
4. Al Luhayb ASM. The Bootstrap Method for Monte Carlo Integration Inference. Journal of King Saud University - Science 2023 :102768. DOI: https://doi.org/10.1016/j.jksus.2023.102768. Available from: https://www.sciencedirect.com/science/article/pii/S1018364723002306
5. Banks DL. Histospline Smoothing The Bayesian Bootstrap. Biometrika 1988; 75:673–84
6. Efron B. Censored Data and The Bootstrap. Journal of the American Statistical Association 1981; 76:312–9
7. Dobler D. Bootstrapping The Kaplan–Meier Estimator on The Whole Line. Annals of the Institute of Statistical Mathematics 2019; 71:213–46
8. Al Luhayb ASM. Smoothed Bootstrap Methods for Right-Censored Data and Bivariate Data. PhD thesis. Durham University, 2021
9. Al Luhayb ASM, Coolen FPA, and Coolen-Maturi T. Generalizing Banks' Smoothed Bootstrap Method for Right-Censored Data. 29th European Safety and Reliability Conference (ESREL 2019), Hannover (Germany). 2019 :894–901
10. Al Luhayb ASM, Coolen FPA, and Coolen-Maturi T. Smoothed Bootstrap for Survival Function Inference. Proceedings of the International Conference on Information and Digital Technologies (IDT 2019), Zilina (Slovakia). 2019 :297–304

11. Al Luhayb ASM, Coolen FPA, and Coolen-Maturi T. Smoothed Bootstrap for Right-Censored Data. Communications in Statistics-Theory and Methods 2023 :1–25. Available from: https://doi.org/10.1080/03610926.2023.2171708
12. Al Luhayb ASM, Coolen-Maturi T, and Coolen FPA. Smoothed bootstrap methods for bivariate data. Journal of Statistical Theory and Practice 2023; 17:37. Available from: https://doi.org/10.1007/s42519-023-00334-7
13. Coolen FPA and Yan KJ. Nonparametric Predictive Inference with Right-Censored Data. Journal of Statistical Planning and Inference 2004; 126:25–54
14. Al Luhayb ASM. Nonparametric Bootstrap Methods for Hypothesis Testing in The Event of Double-Censored Data. AIMS Mathematics 2024; 9:4649–64
15. Klein JP and Moeschberger ML. Survival Analysis Techniques for Censored and Truncated Data. New York, NY: Springer, 2003
16. Hill BM. Posterior Distribution of Percentiles: Bayes' Theorem for Sampling From A Population. Journal of the American Statistical Association 1968; 63:677–91
17. Hill BM. De Finetti's Theorem, Induction, and $A_{(n)}$ or Bayesian Nonparametric Predictive Inference (With Discussion). In: Bayesian Statistics 1988; 3. Bernardo, J.M., DeGroot, M.H., Lindley, D.V. and Smith, A.F.M. (Eds), Oxford University Press:211–41
18. Al Luhayb ASM. Nonparametric Statistical Method For Prediction in Case of Data Including Double-Censored Observations. Pakistan Journal of Statistics 2023; 39
19. Al Luhayb ASM. Nonparametric methods of statistical inference for double-censored data with applications. Demonstratio Mathematica 2024; 57
20. Berliner LM and Hill BM. Bayesian Nonparametric Survival Analysis. Journal of the American Statistical Association 1988; 83:772–9
21. Kaplan EL and Meier P. Nonparametric Estimation from Incomplete Observations. Journal of the American Statistical Association 1958; 53:457–81
22. Wan F. Simulating Survival Data with Predefined Censoring Rates for Proportional Hazards Models. Statistics in Medicine 2017; 36:721–880
23. Wang Y, Zhou Q, Cai T, and Wang X. Semi-supervised Estimation of Event Rate with Doubly-Censored Survival Data. arXiv preprint arXiv:2311.02574 2023. Preprint, submitted November 2023
24. Han S, Wang W, and Hannig J. Unified Fiducial Inference for Interval-Censored Data. Journal of the American Statistical Association 2023. Published online, September 2023. DOI: 10.1080/01621459.2023.2252143
25. Cui Y, Hannig J, and Iyer HK. Generalized fiducial inference for survival functions under censoring and truncation. Biometrika 2019; 106:501–18. DOI: 10.1093/biomet/asz018
26. Cui Y, Hannig J, Lee TC, and Liu R. Unified Fiducial Inference for Interval-Censored Data. Journal of the American Statistical Association 2024. Advance online publication:1–15. DOI: 10.1080/01621459.2024.2346085