Smoothed Bootstrap Methods for Right-Censored Data and Bivariate Data

AL-LUHAYB, ASAMH,SALEH,M

AL-LUHAYB, ASAMH,SALEH,M (2021) *Smoothed Bootstrap Methods for Right-Censored Data and Bivariate Data*, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/14096/

Smoothed Bootstrap Methods for Right-Censored Data and Bivariate Data

Asamh Saleh Muhammad Al-Luhayb

A Thesis presented for the degree of
Doctor of Philosophy

Statistics and Probability
Department of Mathematical Sciences
University of Durham
England

June 2021

Dedicated to

My parents
for all their unlimited supports and prayers

My wife Abeer and daughter Mohra
who make my life convivial

My siblings
for their believes and wishes

My friends
for encouraging and believing

Smoothed Bootstrap Methods for Right-Censored Data and Bivariate Data

Asamh Saleh Muhammad Al-Luhayb

Submitted for the degree of Doctor of Philosophy
June 2021

Abstract

This thesis introduces a smoothed bootstrap method for univariate right-censored data and investigates this bootstrap method for the coverage probability and survival function inferences through simulations. The bootstrap method relies on the right-censoring $A_{(n)}$ assumption, which was proposed by Coolen and Yan [21]. This assumption allows sampling from the whole data range and avoids the complication in computation that occurs due to ties and right-censored observations which often occur in the samples created by Efron's bootstrap method [31]. The performance of the proposed bootstrap method is studied on finite and infinite data ranges, and compared to the performance of Efron's bootstrap method through simulations. It is found that the smoothed bootstrap method mostly outperforms Efron's bootstrap method, in particular when the sample size is small. Also, the smoothed bootstrap method and Efron's bootstrap method are compared through simulations to compute the actual Type 1 error rates of quartiles tests and two sample medians test.

For bivariate data, three smoothed bootstrap methods are introduced. Two of them are based on the generalization of Nonparametric Predictive Inference for random quantities (X, Y) with copulas, proposed by Coolen-Maturi *et al.* [22] and Muhammad *et al.* [65]. The third one is by using uniform kernels. These smoothed bootstrap methods are compared to Efron's bootstrap method [33] through simulations. It is found that the smoothed bootstrap methods mostly outperform Efron's bootstrap method in terms of the coverage probabilities for Pearson correlation and the means of $T_1 = X + Y$ and $T_2 = XY^2$ when the data distribution is symmetric.

Also, these bootstrap methods are compared to compute the Type 1 error rates of the Pearson and Kendall correlation tests to provide insight into the methods' performances. For the Pearson correlation test, the smoothed bootstrap methods mostly perform better than Efron's method, but Efron's method provides better results for the Kendall correlation test.

Declaration

The work in this thesis is based on research carried out in the Department of Mathematical Sciences at Durham University. No part of this thesis has been submitted elsewhere for any degree or qualification, and it is all my own work unless referenced to the contrary in the text.

Acknowledgements

The one I should thank most is my God; Allah. I am really grateful for everything you have given me in general and letting me accomplish this thesis particularly.

Unlimited thanks go to my supervisors Prof. Frank Coolen and Dr. Tahani Coolen-Maturi for being a source of support in my academic life. I truly appreciate all your advices and the guidance during the research.

To my parents, brothers and sisters, thank you all for your support and encouragement. Special thanks go to my sister Asma for her enthusiasm, untiring support and calm advice.

Thanks to Durham University, Qassim University, and the Saudi government for the supports that have enabled me to study smoothly.

Contents

Chapter 1

Introduction

1.1 Overview

Measuring the uncertainty of a sample estimate is central to statistical inference. In a simple situation, it is possible to assume a probability model and make inferences, but with a complicated problem this may be difficult and the results can mislead if model assumptions are not suitable. To avoid this obstacle, Efron [30] used the power of computer to develop a bootstrap method, which relies on few assumptions, but with more computations. This method has been widely used due to its simplicity and potential to provide good approximate results to the sample estimates.

In 1979, Efron [30] proposed the standard bootstrap for real data to measure the accuracy of statistical estimates. The method relies on sampling with replacement from the original data set to create the bootstrap samples. In 1981, he generalized his method for right-censored data [31]. He assumed the empirical distribution on the original data set, which puts a probability mass $\frac{1}{n}$ at each data point, whether it is an event or censored observation. The bootstrap samples are created by sampling with replacement from the original sample, then he used the Kaplan-Meier estimator for inferences based on each bootstrap sample [31, 53].

For bivariate data, he described a bootstrap technique which is close to the standard bootstrap method [33]. He assumed the data distribution is empirical, so that each pair is sampled with probability $\frac{1}{n}$. To create the bootstrap samples, he sampled pairs with replacement from the original sample. This bootstrap method is used to estimate the standard error of the Pearson correlation in [33].

Banks [5] developed a smoothed bootstrap method for univariate real-valued data by linear interpolation between the observed points. The method requires that the underlying distribution is continuous with finite support and no ties. This bootstrap method starts with ordering the observed data points, then partitioning the sample space into $n+1$ intervals by the original observations, and putting mass $\frac{1}{n+1}$ at each interval. To create one bootstrap sample, n intervals are sampled with replacement, then from each chosen interval, one observation is drawn uniformly. This method allows sampling from the whole data range, and ties occur with probability zero in the bootstrap samples.

Dolker *et al.* [28] and Lee and Rodgers [58] showed that Efron's bootstrap method has problems when the sample size is small. The method restricts the process to sampling with replacement from the original data set, so bootstrap samples often include ties. In this situation, it may become harder to make inference based on those bootstrap samples; assumptions are needed to break the ties. Also, the method does not provide good results for small data sets as explained in [14, 28, 58]. These drawbacks motivate the development of smoothed bootstrap methods for right-censored data and bivariate data to avoid ties and obtain better results in terms of the coverage probability and testing.

This thesis presents smoothed bootstrap methods for right-censored data and bivariate data to quantify the uncertainty of sample estimates. The proposed bootstrap methods will be compared to Efron's methods in terms of the coverage probability through simulations. They are also compared in computing the Type 1 error rates of some statistical tests.

This chapter introduces some bootstrap methods from the literature. Section 1.2 presents Efron's bootstrap methods for real-valued data, right-censored data and bivariate data. In Section 1.3, Banks' bootstrap method is described with its generalization on the whole real line and the positive real line. Section 1.4 provides an introduction to nonparametric predictive inference and Section 1.5 provides the outline of this thesis.

1.2 Efron's bootstrap methods

The standard bootstrap method, presented by Efron [30], is a technique to quantify the uncertainty of sample estimates. It has been widely used in applied statistics as it relies on few mathematical assumptions and its application is straightforward with statistical software [34]. Suppose that there is a random sample from an unknown distribution. In the literature, traditional statistical methods are mostly based on an assumed model to make inferences reflecting the whole population from which the data sample was drawn. In real world applications, it may be difficult to choose a suitable model and hence powerful statistical methods for inference without an assumed underlying model are of great practical use [26]. The bootstrap method, which is based on the empirical distribution, has become one of the most used statistical methods [30]. In this section, we describe Efron's bootstrap methods for real-valued data, right-censored data and bivariate data, where the methods for right-censored data and bivariate data are particularly used in Chapters 2 and 3, respectively.

1.2.1 Efron's bootstrap method for real-valued data

Efron [30] presents the standard version of bootstrap method, which is a resampling technique from the original data set. It is used to specify the accuracy of a sample estimate, e.g. estimating the standard error or the percentile confidence interval of a sample mean. This bootstrap method used the empirical distribution to measure the uncertainty, but it requires a lot of computations. From the original data set, B bootstrap samples are created by resampling, then find the function of interest in each bootstrap sample. The empirical distribution for the B resulting values can be considered as a proxy distribution for the sample statistic. It is advised that the number of B should be large, e.g. $B = 1000$.

There are many references describing the bootstrap method along with examples and applications, e.g. Berrar [8], Davison and Hinkley [26] and Efron and Tibshirani [34], and the idea of bootstrap has been used for a variety of statistical inferences. For example, Rosenkranz [72] used the bootstrap to estimate the bias of treatment effect estimators. Mandel and Betensky [60] derive simultaneous confidence intervals for several parameters based on the percentile bootstrap approach. Davison and Hinkley [26] and Efron and Tibshirani [34] used it to provide a bootstrap estimate of standard error and a bootstrap confidence interval for a statistic. They also used the bootstrap method for testing and regression problems. Davison and Hinkley [26] provided R codes to put this method into practice.

Suppose that $x_1, x_2, \ldots, x_n$ are observations corresponding to independent and identically distributed random variables $X_1, X_2, \ldots, X_n$, and F is a continuous distribution defined on a sample space $\aleph$. Let $\theta(F)$ be a function of interest, e.g. the mean or median, which can be estimated by $\hat{\theta} = \theta(\hat{F})$, where $\hat{F}$ is the empirical distribution function putting mass $\frac{1}{n}$ at each value x_i,

$$\hat{F}(x) = \frac{1}{n}\sum_{i=1}^{n} I(x_i \leq x) \tag{1.1}$$

where $I(x_i \leq x)$ is the indicator function which is 1 if $x_i \leq x$ and 0 otherwise.

The bootstrap procedure uses the empirical distribution function $\hat{F}$ of the original data set to be a proxy for the unknown distribution F, so each observation is assigned probability $\frac{1}{n}$. Hence, the algorithm of Efron's bootstrap is as follows [30, 31]:

(i) Construct the empirical cumulative distribution function $\hat{F}$ using Equation (1.1).

(ii) Draw B random samples of size n with replacement from the original data set.

(iii) Compute the statistic of interest for each bootstrap sample. This leads to $\hat{\theta}^{*1}$, $\hat{\theta}^{*2}$, ..., $\hat{\theta}^{*B}$. Then the empirical distribution of $\hat{\theta}^{*1}, \hat{\theta}^{*2}, \ldots, \hat{\theta}^{*B}$ can be used to approximate the sampling distribution of $\theta(F)$.

To provide a bootstrap estimate of standard error $\hat{\sigma}_{boot}$ for the sample statistic $\hat{\theta}$, we compute the standard deviation of $\hat{\theta}^{*1}, \hat{\theta}^{*2}, \ldots, \hat{\theta}^{*B}$ by

$$\hat{\sigma}_{boot} = \sqrt{\frac{\sum_{j=1}^{B}(\hat{\theta}^{*j})^2 - (\sum_{j=1}^{B}\hat{\theta}^{*j})^2/B}{B-1}} \tag{1.2}$$

For estimating a standard error, Efron and Tibshirani [34] determined B to be in the range $25 - 200$.

1.2.2 Efron's bootstrap method for right-censored data

Right-censored data often occurs in reliability and medical applications because of many reasons. For example, if a study is ended while some individuals do not experience the event of interest. These observations will be considered as right-censored times. The only fact we know about right-censored observations is that these observations do not experience the event at given times [56, 69]. A medical example for right-censored data can be the results for a clinical trial of a drug 6-mercaptopurine versus a placebo in 42 children with acute leukemia reported by Freireich *et al.* [38]. Patients were followed until the disease returned. Some of them had not experienced the disease until the end of study, so these observations were considered as right-censored observations. Also, some patients were missed to follow for any reason during the study with known that they had not experienced the leukemia again, so those patients were considered right-censored observations as well.

In 1981, Efron [31] presented a bootstrap method for right-censored data by sampling with replacement from the original data set, similar to his original bootstrap method explained in Subsection 1.2.1. It assigns probability $\frac{1}{n}$ to each observation, regardless of whether it is an observed event time or a right-censored observation.

Suppose that (X_1, D_1), $(X_2, D_2), \ldots, (X_n, D_n)$ are independent and identically distributed random variables, with interest in the random quantities X_i, drawn from a distribution F. D_i represents the censoring status, where it takes 0 for censored time and 1 for event time. Note that we do not need to introduce notation for an assumed distribution for the censoring random quantities D_i as we restrict the attention explicitly to characteristics of the distribution F. Let (x_1, d_1), $(x_2, d_2), \ldots,$ (x_n, d_n) denote the corresponding observations. The pair (x_i, d_i), for $i = 1, 2, \ldots, n$, is the form for right-censored data where x_i is the i^{th} observation and $d_i = 1$ if x_i is uncensored, and $d_i = 0$ if x_i is censored. Let $\theta(F)$ be the functional of interest, which can be estimated through $\theta(\hat{F})$. Efron's bootstrap method for right-censored data is described as follows [31]:

(i) From the original sample, which has n observations, a simple random sample with size n is drawn with replacement. The bootstrap sample is denoted by $D^* = \{(x_1^*, d_1^*), (x_2^*, d_2^*), \ldots, (x_n^*, d_n^*)\}$.

(ii) Calculate the function of interest $\hat{\theta}^* = \hat{\theta}(D^*)$ based on the Kaplan-Meier estimator [53].

(iii) Repeat (i) and (ii) B times, and this leads to have $\hat{\theta}^{*1}, \hat{\theta}^{*2}, \ldots, \hat{\theta}^{*B}$.

It should be noted that this bootstrap method often leads to include ties and right-censored observations in the bootstrap samples. This can cause complications when the sample size is small and the censoring proportion is high.

1.2.3 Efron's bootstrap method for bivariate data

In 1986, Efron and Tibshirani [33] presented a bootstrap method for bivariate data with a description of the bootstrap estimate of the standard error for the Pearson correlation in an American law school data set. Efron's bootstrap method for bivariate data is quite similar to the standard Efron's bootstrap method explained in [30]. The empirical distribution is assumed for the original data set, so each observation is assigned probability $\frac{1}{n}$. This technique is used for regression models [32] and different measures of statistical accuracy [33]. For more explanations and applications, see the book by Efron and Tibshirani [34]. However, Lee and Rodgers [58] showed that the method performs poorly for Pearson correlation test in comparison to their univariate sampling bootstrap method, in particular when the sample size is small.

To describe the method, suppose (X_1, Y_1), $(X_2, Y_2), \ldots, (X_n, Y_n)$ are independent and identically distributed random variables following an unknown probability distribution F with a sample space $\aleph$. Let (x_1, y_1), (x_2, y_2), $\ldots$, (x_n, y_n) denote the observed data and $\theta(F)$ be the function of interest, which can be estimated through $\theta(\hat{F})$, where $\hat{F}$ is the empirical distribution. Efron's procedure for bivariate data is as follows [33]:

(i) From the original sample, which has n observations, a simple random sample with size n is drawn with replacement. The bootstrap sample is denoted by $D^* = \{(x_1^*, y_1^*), (x_2^*, y_2^*), \ldots, (x_n^*, y_n^*)\}$.

(ii) Calculate the function of interest $\hat{\theta}^* = \hat{\theta}(D^*)$.

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

In the literature, many references present tips and guidance for the number of bootstrap replications B depending on the function of interest. To estimate the standard error, it is mostly sufficient to set $B = 50$, but for confidence intervals and hypotheses tests, B should be at least 1000 for well results [34]. Some researchers use larger numbers, but this causes long computation time. In this thesis, we set $B = 1000$ for our use not only because we think it is suitable for our purposes, but also this number is widely used in the literature for confidence intervals and hypotheses tests.

1.3 Banks' bootstrap method for real-valued data

As described in Subsection 1.2.1, Efron's method [30] restricts the bootstrap sampling process to re-sampling with replacement from the original data set, so no observation can be outside the range of the original data set, and also ties can occur in the Efron's bootstrap samples, which can lead to a complication in computation and poor results in the sense of statistical accuracy when the sample size is small. Banks [5] presented an alternative smoothed bootstrap method for real-valued data to avoid such problems and improve the estimation results. The method allows to sample from the whole data range, but it requires the underlying distribution to be continuous with known limited support.

Suppose that $x_1, x_2, \ldots, x_n$ are observations corresponding to independent and identically distributed random variables $X_1, X_2, \ldots, X_n$, and F is a continuous distribution on $[a, b]$. Banks' bootstrap method [5] orders the observations and creates $n + 1$ intervals partitioning the sample space, where each interval is assigned probability $\frac{1}{n+1}$. Banks' procedure is described as follows:

(i) The observations $x_{(0)} < x_{(1)} < \ldots < x_{(n)} < x_{(n+1)}$ lead to $n + 1$ open intervals $(x_{(i)}, x_{(i+1)})$, where $i = 0, 1, 2, \ldots, n$, $x_{(0)} = a$ and $x_{(n+1)} = b$.

(ii) Sample one interval with probability $1/(n + 1)$.

(iii) Sample one observation uniformly from this chosen interval.

(iv) Repeat steps (ii) and (iii) n times to derive one Banks' bootstrap sample.

(v) Calculate the function of interest $\hat{\theta}^*$.

(vi) Perform steps (ii)-(v) B times in order to obtain B Banks' bootstrap samples along with their corresponding functions of interest.

Banks [5] used simulation studies to compare his method to other bootstrap competitors by the chi-squared goodness of fit test. He concluded that his approach outperforms Efron's method in terms of the coverage accuracy for the mean, median and variance, in particular when the sample size is small. However, the smoothed bootstrap method requires a limited compact support and this is not usually the case in real applications.

Suppose that the data is defined on the whole real line $(-\infty, \infty)$. Based on Banks' bootstrap technique, it is known that the first and last intervals, $(-\infty, x_{(1)})$ and $(x_{(n)}, \infty)$, are both with probability $\frac{1}{n+1}$, but it is difficult to sample uniformly from these intervals during the bootstrap procedure. For these unlimited intervals, Binhimd and Coolen [11] assumed the tails of a Normal distribution, where the parameters μ and σ are set equal to

$$\begin{aligned} \mu &= \frac{x_{(1)} + x_{(n)}}{2} \\ \sigma &= \frac{x_{(n)} - \mu}{\Phi^{-1}(\frac{n}{n+1})} \end{aligned} \tag{1.3}$$

where Φ is the standard Normal cumulative distribution function.

When either one of the unlimited intervals are chosen during the bootstrap procedure, an observation is sampled from $\mathrm{N}(\mu, \sigma^2)$ and this sampled observation is accepted for $(x_{(n)}, \infty)$ if it is greater than $x_{(n)}$. For the interval $(-\infty, x_{(1)})$, we sample an observation from $\mathrm{N}(\mu, \sigma^2)$ and the sampled observation is accepted if it is less than the observation $x_{(1)}$.

If the data is supported on $(0, \infty)$, Binhimd and Coolen [11] assumed the tail of an Exponential distribution for the last interval $(x_{(n)}, \infty)$ with the rate parameter λ set equal to

$$\lambda = \frac{\ln(n+1)}{x_{(n)}} \tag{1.4}$$

where the last interval $(x_{(n)}, \infty)$ is set with probability $\frac{1}{n+1}$.

1.4 Nonparametric Predictive Inference (NPI)

Nonparametric predictive inference (NPI) is a statistical method that makes inferences on a future observation based on a past sample of size n by using Hill's assumption $A_{(n)}$ [49, 50, 51]. NPI produces lower and upper probabilities for the next future observation X_{n+1}, denoted by $\underline{P}(X_{n+1})$ and $\overline{P}(X_{n+1})$, respectively, where $0 \leq \underline{P}(X_{n+1}) \leq \overline{P}(X_{n+1}) \leq 1$. It is a method using imprecise probability [20], and it has been used in many statistical applications for inference on a future observation based on past data observations. For more details see Section 7.6 in the book by Augustin *et al.* [4].

Hill [49] introduced the assumption $A_{(n)}$ to present predictive probabilities for one future observation X_{n+1} with no prior information about an underlying distribution. Suppose that $x_1, x_2, \ldots, x_n$ are the observed data corresponding to the continuous and exchangeable real-valued random quantities $X_1, X_2, \ldots, X_n$. Let $x_{(1)} < x_{(2)} < \ldots < x_{(n)}$ be the ordered observations and, for ease of notation, define $x_{(0)} = -\infty$ ($x_{(0)} = 0$ if we have lifetime data) and $x_{(n+1)} = +\infty$. For one future observation X_{n+1}, the assumption $A_{(n)}$ is

$$P\left(X_{n+1} \in I_i\right) = \frac{1}{n+1} \tag{1.5}$$

where $I_i = (x_{(i)}, x_{(i+1)})$, and $i = 0, 1, 2, \ldots, n$. If there are ties, a very small value can be added to one observed point to break the ties. This way to break the ties has been widely used in the literature [50].

Based on the assumption $A_{(n)}$, NPI is a frequentist statistical technique [3, 17, 18], and the inferences based on Hill's assumption $A_{(n)}$ are nonparametric and predictive. The assumption $A_{(n)}$ is considered as a suitable technique if there is no knowledge about the random quantity of interest. NPI is sufficient to provide imprecise probability for any event A, but not if we are interested in precise probabilities. NPI produces lower and upper probabilities, which are considered in interval probability theory [18, 84]. NPI based on the $A_{(n)}$ assumption provides a strong consistency property in the frequentist theory of statistics [20, 84], and the results based on NPI will be never contradicted to the ones based on the empirical probabilities [57].

In NPI, the lower and upper probabilities are the maximum lower bound and the minimum upper bound for the precise probability for A, respectively. The precise probability for A, $P(A)$, is a special case of imprecise probability, and this is when the lower and upper probabilities are equal, $\underline{P}(A) = \overline{P}(A)$. The NPI lower and upper probabilities for $X_{n+1} \in B$, where $B \subseteq \mathbb{R}$, can be calculated by [10]

$$\underline{P}(X_{n+1} \in B) = \frac{1}{n+1} \sum_{i=0}^{n} I(I_i \subseteq B) \tag{1.6}$$

$$\overline{P}(X_{n+1} \in B) = \frac{1}{n+1} \sum_{i=0}^{n} I(I_i \cap B \neq \emptyset) \tag{1.7}$$

where $I(.)$ is the indicator function. For the lower probability, we sum up all probability masses assigned to intervals which are completely within the set B, and the upper probability is equal to the summation of the probability masses assigned to intervals which intersect with the set B; so all I_i such that $I_i \cap B \neq \phi$.

For data containing right-censored observations, NPI is generalized based on the generalization of $A_{(n)}$ introduced by Coolen and Yan [21]. This generalization is known as the right-censoring $A_{(n)}$ assumption, rc-$A_{(n)}$. The assumption provides a probability distribution for next future observation based on n observed data points including u event times and v right-censored times. It creates $n + 1$ intervals dividing the sample space and each interval is assigned a certain probability mass defined by mass function value for next future observation within each interval. The rc-$A_{(n)}$ assumption can present predictive probabilities for next future observation based on a past sample with no knowledge about the underlying distribution. This assumption is used for survival function inferences [2] and for imprecise reliability [83]. It is also used to provide the age replacement of technical units by Coolen-Schrijner and Coolen [24].

Coolen-Maturi *et al.* [22] and Muhammad *et al.* [65] generalized the NPI technique for bivariate data with parametric and nonparametric copulas. These generalizations use the NPI method for the variables X and Y separately, and then take the dependence structure into account by parametric and nonparametric copulas. Muhammad [64] applied the generalized methods using data from the literature, and their performances were investigated through simulation studies [22, 65]. Muhammad *et al.* [66] used the parametric generalization of NPI for combining bivariate diagnostic test.

In many applications, NPI has been used due to its simplicity to implement and relying on few assumptions with good performance. NPI has been introduced for different data, see e.g. for real data [3], Bernoulli data [16], and ordinal data [35]. It is also used for right-censored data [21]. For statistical inference and decision support, NPI has been widely considered, see e.g. accuracy of diagnostic tests [23, 36], precedence testing for two groups [25], acceptance sampling [19] and sequential acceptance problems [37]. In Chapter 3, we will use the generalizations of NPI proposed by Muhammad [64] to introduce smoothed bootstrap methods for bivariate data.

1.5 Outline of thesis

In this thesis, we propose different versions of bootstrap for two different data types. For right-censored data, we introduce a generalization of Banks' bootstrap method by using the right-censoring $A_{(n)}$ assumption introduced by Coolen and Yan [21], where this assumption is presented in Section 2.2. For bivariate data, three smoothed bootstrap methods are proposed based on the semi-parametric and nonparametric predictive methods, introduced by Coolen-Maturi *et al.* [22] and Muhammad *et al.* [65], and a uniform kernel assigned to each observed point. The third proposed method uses a uniform kernel around each observation without any further restriction on the spread of the probability mass within that kernel. All assumptions that we used to provide smoothed bootstrap methods for bivariate data are introduced in Chapter 3. We use these proposed bootstrap methods for different statistical inferences in the perspective of estimation.

Chapter 2 introduces the smoothed bootstrap method for right-censored data. It is then compared to Efron's method through simulation studies in three different scenarios with different data supports. The smoothed bootstrap method is described with both finite and infinite intervals. In these comparisons, we use the technique Banks' proposed in [5]. We also provide a brief comparison of our approach and the well-known Kaplan-Meier estimator of the survival function through two examples from the literature. This is followed by a comparison between our smoothed bootstrap method and an alternative smoothed bootstrap technique through simulations in terms of the coverage probability for the survival function.

Some results of this chapter were presented in the papers "Generalizing Banks' smoothed bootstrap method for right-censored data" and "Smoothed bootstrap for survival function inference", which are published in the proceedings of the 29th European Safety and Reliability Conference (ESREL 2019) and the International Conference on Information and Digital Technologies (IDT 2019) [1, 2], respectively.

In Chapter 3, parametric and nonparametric copulas in combination with NPI are described and these combinations are used to provide smoothed bootstrap methods for bivariate data. These combinations were proposed in [22, 65] to generalize NPI for bivariate data. Another smoothed bootstrap method is introduced by the uniform kernels. These proposed methods are compared to Efron's method through simulation studies. In those comparisons, different settings of sample sizes and dependence levels between the variables are considered.

Chapter 4 shows how the bootstrap methods, presented in Chapters 2 and 3, can be used as alternative methods for testing hypotheses. For right-censored data, the smoothed bootstrap and Efron's method are used to compute the Type 1 error rates of quartiles hypothesis tests and two sample medians test. The smoothed bootstrap methods for bivariate data and Efron's bootstrap are used to compute the Type 1 error rates of Pearson and Kendall correlation tests. In Chapter 5, we point out some remarks and conclusions. In the appendix, there are extra simulation results and we have included the R commands that have been used for the smoothed bootstrap methods. The calculations in this thesis were performed using the statistical software R version 3.6.1.

Chapter 2

A smoothed bootstrap method for right-censored data

2.1 Introduction

In 1981, Efron [31] introduced the bootstrap method for right-censored data. This bootstrap method is simple and straightforward for calculating some measures of statistical accuracy. However, it performs poorly when the sample size is small and the censoring proportion is large in the original sample. It also causes some issues, e.g. ties and censored observation in the bootstrap samples.

This chapter introduces a smoothed bootstrap method based on the right-censoring $A_{(n)}$ assumption, which was proposed by Coolen and Yan [21]. Arguably, the advantages of the smoothed bootstrap method are threefold. First, this bootstrap method can avoid ties and right-censored observations in the bootstrap samples and this eases the calculations. Secondly, the proposed bootstrap method is easy to apply. Finally, simulation studies support the general superiority of the smoothed bootstrap method.

This chapter is organized as follows. Section 2.2 presents the right-censoring $A_{(n)}$ assumption. Section 2.3 presents the smoothed bootstrap method for right-censored data. In Section 2.4, the smoothed bootstrap approach is compared to Efron's bootstrap method for right-censored data through simulation studies. In Section 2.5, we fit an Exponential distribution to each interval, then sample from the fitted distributions. We provide a brief comparison of our approach and the well-known Kaplan-Meier estimator of the survival function in Section 2.6. We compare the smoothed bootstrap method to an alternative smoothed bootstrap method in Section 2.7. The final section provides some concluding remarks.

2.2 The right-censoring $A_{(n)}$ assumption

The assumption $A_{(n)}$, which was proposed by Hill [49, 50], is defined as the observable random quantities $X_1, X_2, \ldots, X_n$ are exchangeable and ties have probability zero. The probability that next future observation X_{n+1} falls in the open interval $(x_{(j)}, x_{(j+1)})$ is $1/(n+1)$, for all $j = 0, 1, 2, \ldots, n$, where $x_{(0)} = -\infty$ (or $x_{(0)} = 0$ if the random quantities are non-negative), and $x_{(n+1)} = +\infty$.

Berliner and Hill [7] used the $A_{(n)}$ assumption for right-censored data. They order the event time observations and denote them by $0 < t_{(1)} < t_{(2)} < \ldots < t_{(u)}$, where $0 \leq u \leq n$. Then, they assign a specific probability for the next future observation X_{n+1} to be between any two ordered event times $(t_{(u)}, t_{(u+1)})$ only. They only used the event observations to create the intervals; not using the precise right-censored observations. Coolen and Yan [21] generalized the $A_{(n)}$ assumption for right-censored data, denoted by rc-$A_{(n)}$. They use all event and censored observations to create $n+1$ intervals partitioning the sample space. They specify a probability for X_{n+1} to be between any two consecutive event times and between any censored observation and its following event time. This generalization leads to divide the support into $n+1$ intervals while Berliner and Hill [7] create only $u+1$ intervals, where $u \leq n$. The rc-$A_{(n)}$ assumption presented by Coolen and Yan [21] will be used to generalize Banks' bootstrap method for right-censored data.

The rc-$A_{(n)}$ assumption provides a partially specified predictive probability distribution for a future observation, and it is expressed through M-function values [21], which are basic probability assignments following the general theory introduced by Shafer [76]. A partial specification of a probability distribution for a real-valued random quantity X can be provided through probability masses assigned to intervals without any further restriction on the spread of the probability mass within each interval. A probability mass assigned to an interval (a, b) is denoted by $M_X(a, b)$, and referred to as M-function value for X on (a, b). These M-function values are within $[0, 1]$ and they sum up to one over all intervals considered [21].

In order to introduce the rc-$A_{(n)}$ assumption, further notation is needed. The information is assumed to be from exchangeable non-negative real-valued random quantities $X_1, X_2, \ldots, X_n$. Assume that the observations include u event times and v right-censoring times. The event times are denoted by $0 < t_{(1)} < t_{(2)} < \ldots < t_{(u)}$, where $0 \leq u \leq n$. The right-censoring times are denoted by $0 < c_{(1)} < c_{(2)} < \ldots < c_{(v)}$, where $v = n - u$. Let $I_i = (t_{(i)}, t_{(i+1)})$, for $0 \leq i \leq u$, and denote the ordered right-censoring times within I_i by $c_1^i < c_2^i < \ldots < c_{l_i}^i$, where l_i is the number of right-censored observations in I_i. The data range is divided into $n + 1$ intervals, which are $(t_{(i)}, t_{(i+1)})$ and $(c_k^i, t_{(i+1)})$, where $1 \leq k \leq l_i$, $t_{(0)} = 0$ and $t_{(u+1)} = +\infty$ (or $t_{(u+1)} = b$ if the support is known to be $[0, b]$).

The rc-$A_{(n)}$ assumption [21] can be defined as the probability distribution for a non-negative random quantity X_{n+1}, on the basis of data including u event times and v right-censoring times, is partially specified by the following M-function values:

$$M_{X_{n+1}}(t_{(i)}, t_{(i+1)}) = \frac{1}{n+1} \prod_{\{r: c_{(r)} < t_{(i)}\}} \frac{\tilde{n}_{c_{(r)}} + 1}{\tilde{n}_{c_{(r)}}} \tag{2.1}$$

$$M_{X_{n+1}}(c_k^i, t_{(i+1)}) = \frac{1}{(n+1)\tilde{n}_{c_k^i}} \prod_{\{r: c_{(r)} < c_k^i\}} \frac{\tilde{n}_{c_{(r)}} + 1}{\tilde{n}_{c_{(r)}}} \tag{2.2}$$

where $\tilde{n}_{c_{(r)}}$ is the number of individuals remaining at risk (still alive) just before time $c_{(r)}$ plus one, and $\tilde{n}_{c_k^i}$ is the number of individuals remaining at risk just before time c_k^i plus one.

2.3 A smoothed bootstrap method

In applying Banks' bootstrap method for a real-valued data set with a limited data range, as described in Section 1.3, the method randomly selects one of the intervals in the partition created by the data, then samples one observation uniformly from the chosen interval. For the whole real line or the positive real line, Binhimd and

Coolen [11] assumed distribution tails for the unbounded intervals and sampled observations from the tails for those intervals. The assumption Banks' used to smooth the bootstrap method is related to the assumption $A_{(n)}$, which was proposed by Hill [49, 50]. They both divide the sample space into $n+1$ intervals, and each interval is assigned probability $1/(n+1)$. Hence, it is possible to generalize Banks' bootstrap method for right-censored data based on the rc-$A_{(n)}$ assumption. However, before generalizing Banks' bootstrap method for right-censored data, it is important to recall that the aims of this smoothed bootstrap method are to learn about uncertainty in estimation of a population characteristic θ and to avoid the complication in computation that occurs due to ties and censored observations.

Suppose that there are n observations including u event times and v right-censored times. We first order the observations, then create $n+1$ intervals and derive the M-function values as given by Equations (2.1) and (2.2). We present the smoothed bootstrap (SB) algorithm for right-censored data in the following steps, the R codes are provided in Appendix B.1:

(i) Create $n+1$ intervals based on the n observed data points of the form $(t_{(i)}, t_{(i+1)})$ and $(c_k^i, t_{(i+1)})$ where $i = 0, 1, \ldots, u \leq n$ and $1 \leq k \leq l_i$.

(ii) Compute the M-function values for the $n+1$ intervals.

(iii) Sample with replacement n intervals with the assignment probabilities, then sample one observation uniformly from each chosen interval to obtain a smoothed bootstrap sample of size n.

(iv) Calculate the function of interest, $\hat{\theta}^*$.

(v) Perform steps (iii) to (iv) B times. This leads to B smoothed bootstrap samples with their corresponding functions of interest.

If the support of the random quantities is $(0, \infty)$, it is possible to have one or more intervals in the form of $(x_{(i)}, \infty)$ during the smoothed bootstrap technique, where $x_{(i)}$ is either an observed event or a right-censored time. For such cases, Binhimd and Coolen [11] suggested to fit Exponential distributions to the tail intervals, such that the probability masses for such intervals are the same as the assigned M-function values. Following this approach, we assume an Exponential distribution for any interval of the form $(x_{(i)}, \infty)$ with the rate parameter for the Exponential distribution set at

$$\lambda_{(i)} = \frac{-\ln\left(M_{X_{n+1}}(x_{(i)}, \infty)\right)}{x_{(i)}} \tag{2.3}$$

where $M_{X_{n+1}}(x_{(i)}, \infty)$ can be calculated by either Equation (2.1) or (2.2) depending on the observation $x_{(i)}$, which is either an observed event or a right-censored time.

The process of the smoothed bootstrap method has advantages over Efron's method. As described before, Efron's bootstrap restricts to sampling with replacement from the original data set and this often leads the bootstrap samples to contain ties and right-censored observations. When the censoring proportion is large in the original sample, Efron's bootstrap samples may include only censored observations and this causes an obstacle for inference. In contrast, the smoothed bootstrap method allows to sample from the whole data range, and ties occur in the bootstrap samples with probability zero. Also, the bootstrap samples consist only of event time observations, so they do not contain any right-censored observations. These advantages of the smoothed bootstrap method ease the computations for inference.

2.4 Comparison with Efron's method

In the literature, the primary requirement for confidence regions is that nominal coverage probability be close to the actual coverage probability [5]. This has motivated many simulations with accuracy at certain confidence levels, e.g. $0.90, 0.95$ and 0.99. However, Banks [5] investigated the global measure of coverage accuracy to show the best bootstrap method with starting point in coverage. He creates 20 confidence regions with nominal coverage probability 0.05 by

$$CRL_{(i)} = \left(q_{(\frac{\alpha_{i+1}}{2})}, q_{(\frac{\alpha_i}{2})}\right) \tag{2.4}$$

$$CRR_{(i)} = \left(q_{(1-\frac{\alpha_i}{2})}, q_{(1-\frac{\alpha_{i+1}}{2})}\right) \tag{2.5}$$

where $i = 1, 2, \ldots, 10$, $\alpha_{i+1} = \alpha_i - 0.10$, $\alpha_1 = 1$ and $q_{(z)}$ is the z^{th} quantile of functional values, so $CRL_{(i)}$ are the confidence regions presenting the left tail of the global measure of coverage accuracy, and $CRR_{(i)}$ are the confidence regions presenting the right tail of the global measure of coverage accuracy.

He also creates 10 confidence regions with nominal coverage probability 0.10 by

$$CR_{(i)} = \left(q_{(\frac{\alpha_{i+1}}{2})}, q_{(\frac{\alpha_i}{2})}\right) \cup \left(q_{(1-\frac{\alpha_i}{2})}, q_{(1-\frac{\alpha_{i+1}}{2})}\right) \tag{2.6}$$

where $CR_{(i)} = CRL_{(i)} \cup CRR_{(i)}$ for $i = 1, 2, \dots, 10$.

In both divisions, the confidence regions are used to see whether they have equal coverage probabilities for a specific function of interest. Then Banks [5] used the chi-square goodness of fit test to measure the discrepancy in coverage probability. He compared his bootstrap method to other bootstrap techniques, e.g. Efron's method [30], Rubin's Bayesian bootstrap [73] and smoothed Rubin's bootstrap [5], and considered the best method is the one having the lowest χ^2 value among all values. In other words, the lowest discrepancy between the nominal and actual coverage probability for a true specific statistic is provided by the method whose χ^2 value is lowest among all methods. We here intend to generalize the comparison with right-censored data.

To establish a right-censored data set, we first generate n observations from one distribution, which are referred to by $\{t_1, t_2, \dots, t_n\}$, and n observations from another distribution, which are denoted by $\{c_1, c_2, \dots, c_n\}$. Then, we define the right-censored data set by

$$x_i = \min(t_i, c_i), \text{ for } \; i = 1, 2, \dots, n \tag{2.7}$$

$$d_i = \begin{cases} 1 & \text{if } \; x_i = t_i \; \text{ (uncensored)} \\ 0 & \text{if } \; x_i = c_i \; \text{ (censored)} \end{cases} \tag{2.8}$$

where x_i is the time and d_i is the censored indicator.

Three different scenarios are considered to compare the smoothed bootstrap method to Efron's method by using Banks' comparison strategy. For the first scenario, we use the Beta distribution with two shape parameters α and β to generate event times and we use the uniform distribution with two parameters a and b to generate right-censored observations, where their density functions are as follows

$$f_1(t) = \frac{t^{\alpha-1}(1-t)^{\beta-1}}{\beta(\alpha, \beta)}; \quad t \in [0, 1] \tag{2.9}$$

$$g_1(c) = \begin{cases} \dfrac{1}{b-a} & ; \quad c \in [a,b] \\ 0 & ; \quad \text{otherwise} \end{cases} \tag{2.10}$$

The data range of uniform distribution will be determined based on a predefined censoring proportion and this will be discussed in Subsection 2.4.1. Note that the data range for the first scenario is finite.

For the second scenario, we use the Weibull distribution with shape parameter α and scale parameter β to create event times, and the Exponential distribution with rate parameter λ to create right-censored observations with the following density functions

$$f_2(t) = \begin{cases} \frac{\alpha}{\beta}(\frac{t}{\beta})^{\alpha-1}\exp(-(\frac{t}{\beta})^{\alpha}) & ; \quad t \geq 0 \\ 0 & ; \quad t < 0 \end{cases} \tag{2.11}$$

$$g_2(c) = \lambda \exp(-\lambda c); \quad c \in [0, \infty) \tag{2.12}$$

For the third scenario, we will use the standard log-normal distribution to generate event times and Weibull(α, β) to generate right-censored observations with the following density functions

$$f_3(t) = \frac{1}{t\sqrt{2\pi}} \exp(-\frac{(\ln(t))^2}{2}); \quad t \in (0, \infty) \tag{2.13}$$

$$g_3(c) = \begin{cases} \frac{\alpha}{\beta}(\frac{c}{\beta})^{\alpha-1}\exp(-(\frac{c}{\beta})^{\alpha}) & ; \quad c \geq 0 \\ 0 & ; \quad c < 0 \end{cases} \tag{2.14}$$

It is important to note that the data range in the second and third scenarios is $(0, \infty)$. Therefore, we will use different tails' assumptions for the last intervals when we apply the smoothed bootstrap method.

2.4.1 First scenario: finite support

To compare the smoothed bootstrap method to Efron's method for the case of finite support, we create right-censored data sets from the distributions used in the first scenario. These generated data sets can be set with a fixed censoring proportion p by determining the two parameters of the uniform distribution [85], where the censoring proportion p is important to be fixed to eliminate its impact on the performances of the bootstrap methods. We first fix the parameters of the Beta distribution and the parameter a of the uniform distribution, then use the following formula to get the parameter b of the uniform distribution. In the following simulations, we set the Beta distribution parameters $\alpha = 1.2$ and $\beta = 3.2$. We also define the censoring proportion $p = 0.15$ and the parameter a of the uniform distribution is equal to 0.

$$\begin{aligned} p &= \int_0^1 (\frac{x-a}{b-a}) \left(\frac{x^{\alpha-1}(1-x)^{\beta-1}}{Beta(\alpha,\beta)} \right) dx; \; Beta(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \\ p &= \frac{\alpha}{(b-a)\times(\alpha+\beta)} - \frac{a}{(b-a)} \end{aligned} \tag{2.15}$$

$$\implies \quad b = \frac{\left(\frac{\alpha}{\alpha+\beta} + pa - a\right)}{p} = \frac{(\frac{1.2}{1.2+3.2})}{0.15} = 1.82 \tag{2.16}$$

We generate $N = 1000$ data sets with sample size $n = 6$. Then, the smoothed bootstrap method and Efron's method are applied to each data set $B = 1000$ times. We compute the medians of the bootstrap samples, then we compute the 10 and 20 confidence regions by Equations (2.4), (2.5) and (2.6). We then discover which confidence regions include the true median, the median of Beta distribution. We repeat this procedure in all $N = 1000$ generated data sets in order to see the actual coverage probabilities for the true median in the 10 and 20 confidence regions. The actual coverage probabilities for the true median in the 10 and 20 confidence regions are outlined in Tables 2.1 and 2.2, respectively.

With Efron's bootstrap procedure, the samples often include some right-censored observations, so we use the Kaplan-Meier estimator in order to compute their corresponding medians by finding a time t, so that $\hat{S}(t) = 0.50$. It is wished to get 1000 bootstrap samples' medians, but unfortunately we do not find the corresponding medians for some bootstrap samples because the censoring rates in those bootstrap samples are large. In this case, we have considered three options. The first one is just neglecting all the not applicable medians, so the 10 and 20 confidence regions are based on a number of medians which can be less than 1000. This option is referred to by $E_{(1)}$. The second option is applying Efron's suggestion at each bootstrap sample whose median is not found by the Kaplan-Meier estimator. He suggested that the median is assumed to be the maximum event time of that bootstrap sample [29]. This option is referred to by $E_{(2)}$. Finally, we apply an Exponential tail with rate parameter $\hat{\lambda}^* = -\ln(\hat{S}(t_{max}))/t_{max}$, where t_{max} is the maximum event time of

$CR_{(i)}$	SB	$\text{E}_{(1)}$	$\text{E}_{(2)}$	$\text{E}_{(3)}$
1	0.121	0.024	0.025	0.025
2	0.129	0.131	0.129	0.131
3	0.109	0.107	0.115	0.110
4	0.120	0.166	0.155	0.163
5	0.088	0.109	0.119	0.110
6	0.106	0.095	0.090	0.095
7	0.106	0.012	0.014	0.012
8	0.083	0.118	0.118	0.117
9	0.070	0.103	0.105	0.107
10	0.068	0.135	0.130	0.130

Table 2.1: The actual coverage probabilities for the median $Q_2 = 0.236$, in the 10 confidence regions, where $n = 6$.

method	SB		$\text{E}_{(1)}$		$\text{E}_{(2)}$		$\text{E}_{(3)}$	
i	$CRL_{(i)}$	$CRR_{(i)}$	$CRL_{(i)}$	$CRR_{(i)}$	$CRL_{(i)}$	$CRR_{(i)}$	$CRL_{(i)}$	$CRR_{(i)}$
1	0.056	0.065	0.012	0.012	0.009	0.016	0.009	0.016
2	0.068	0.061	0.056	0.075	0.059	0.070	0.054	0.077
3	0.056	0.053	0.062	0.045	0.065	0.050	0.065	0.045
4	0.061	0.059	0.076	0.090	0.068	0.087	0.073	0.090
5	0.049	0.039	0.059	0.050	0.065	0.054	0.060	0.050
6	0.049	0.057	0.064	0.031	0.059	0.031	0.064	0.031
7	0.054	0.052	0.002	0.010	0.004	0.010	0.003	0.009
8	0.048	0.035	0.048	0.070	0.048	0.070	0.047	0.070
9	0.039	0.031	0.053	0.050	0.056	0.049	0.055	0.052
10	0.046	0.022	0.070	0.065	0.064	0.066	0.067	0.063

Table 2.2: The actual coverage probabilities for the median $Q_2 = 0.236$, in the 20 confidence regions, where $n = 6$.

	10 CR				20 CR			
Trial	SB	$E_{(1)}$	$E_{(2)}$	$E_{(3)}$	SB	$E_{(1)}$	$E_{(2)}$	$E_{(3)}$
1	41.92	205.50	188.22	197.62	53.12	231.48	210.56	228.08
2	42.50	202.80	184.96	193.66	56.32	228.56	204.56	221.20
3	35.00	205.64	187.32	197.04	50.52	233.76	212.20	228.92
4	30.40	187.40	168.88	176.14	46.36	216.68	191.92	209.68
5	33.64	178.98	166.46	171.96	53.12	208.16	189.68	202.52
6	32.68	188.20	174.96	179.68	53.08	211.12	194.08	202.56
7	36.92	192.74	177.18	181.52	54.84	215.56	198.44	205.20
8	46.08	188.32	174.14	176.50	57.16	210.56	194.76	198.04
9	43.98	192.62	180.00	179.58	55.12	217.20	200.24	204.16
10	42.04	195.98	183.28	183.78	58.12	215.36	199.56	203.00

Table 2.3: The chi-squared values obtained from coverage probabilities for $Q_2 = 0.236$ where $n = 6$ and $B = 1000$.

the bootstrap sample. Thereby the corresponding median, X_{med}, can be found by, $X_{med} = -\ln(0.50)/\hat{\lambda}^*$. This suggestion is presented in [13], we denote it by $E_{(3)}$. In the last two cases, we can assure that the 10 and 20 confidence regions are based on 1000 bootstrap samples' medians.

Tables 2.1 and 2.2 present the results of these simulations. They show the superiority of the smoothed bootstrap method in making the coverage probabilities in each of the 10 and 20 confidence regions close to 0.100 and 0.050, respectively. In contrast, Efron's bootstrap method, with the three options for the method $E_{(1)}$, $E_{(2)}$ and $E_{(3)}$, leads to the coverage probabilities far from the nominal sizes 0.100 and 0.050, respectively, in most confidence regions. As a result, the discrepancy between the estimated and nominal coverage probabilities is very high. To get the observed statistics of the chi-square test of goodness of fit, we multiply the coverage probabilities in Tables 2.1 and 2.2 by 1000 due to $N = 1000$. Hence, the discrepancy between the actual and nominal coverage probabilities at distinct confidence levels based on the two bootstrap methods can be assessed via the χ^2-test. The resulting χ^2 values are outlined in the first row of Table 2.3.

		10 CR				20 CR			
n	measures	SB	$E_{(1)}$	$E_{(2)}$	$E_{(3)}$	SB	$E_{(1)}$	$E_{(2)}$	$E_{(3)}$
6	χ^2	45.34	207.84	186.38	196.36	55.24	231.48	206.48	223.04
	NA	—	8628	0	0	—	8628	0	0
	ABS	—	493	493	493	—	493	493	493
	P-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
10	χ^2	22.08	167.86	162.66	158.46	25.76	182.08	179.80	175.08
	NA	—	2332	0	0	—	2332	0	0
	ABS	—	1	1	1	—	1	1	1
	P-value	0.009	0.000	0.000	0.000	0.137	0.000	0.000	0.000
20	χ^2	8.32	58.46	59.44	59.44	20.20	73.64	75.04	75.04
	NA	—	98	0	0	—	98	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.502	0.000	0.000	0.000	0.383	0.000	0.000	0.000
40	χ^2	8.96	6.44	6.44	6.44	12.48	18.12	18.12	18.12
	NA	—	0	0	0	—	0	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.441	0.695	0.695	0.695	0.864	0.514	0.514	0.514
100	χ^2	6.94	3.66	3.66	3.66	16.52	10.80	10.80	10.80
	NA	—	0	0	0	—	0	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.643	0.932	0.932	0.932	0.622	0.930	0.930	0.930

Table 2.4: The chi-squared values for $Q_2 = 0.236$, and their *P-values* with the corresponding NA and ABS numbers.

Now, we consider the other two quartiles (Q_1 and Q_3) to evaluate the bootstrap methods. We compute the chi-squared goodness of fit values gained from the coverage probabilities of both bootstrap methods for the quartiles at different sample sizes n. They will be presented with increasing sample sizes and $N = 1000$ simulations as we did before to show the performance of the methods. The number of simulations is set equal to $N = 1000$ because we think it is suitable for our purpose. We repeat simulations for the case of $N = 1000$ several times with different seeds and they give nearly identical outcomes as illustrated in Table 2.3. We consider a variety of sample sizes to explore whether there is an influence of sample size on the chi-squared values or not. This helps us to observe how the bootstrap methods perform as the sample size increases.

In Table 2.4, the chi-squared values obtained from the coverage probabilities for the true median based on the two bootstrap procedures are shown. When $n = 6, 10$ and 20, the generalization of Banks' bootstrap method performs a better coverage accuracy. It makes the discrepancy between the nominal and estimated coverage probabilities lower than Efron's method does and this is apparent from having lower chi-squared values. Both bootstrap methods distribute the actual coverage probabilities good over the 10 and 20 confidence regions when n increases to 40 and 100, but Efron's method is mostly better because it provides lower chi-squared values. With large sample sizes, the empirical distribution can be a good proxy model for the underlying distribution, so that Efron's method performs well. By observing the patterns of chi-squared values in both confidence region divisions when the sample size increases, the corresponding χ^2 values to Efron's bootstrap method go down. For the corresponding χ^2 values to the smoothed bootstrap method, there is no pattern. The three options, $\text{E}_{(1)}$, $\text{E}_{(2)}$ and $\text{E}_{(3)}$, give the same χ^2 values when $n = 40, 100$ because we find the corresponding median to each bootstrap sample by the Kaplan-Meier estimator.

The number NA, in Table 2.4, indicates the number of Efron bootstrap samples whose medians are not found because there is no time satisfying $\hat{S}(t) = 0.50$ in these bootstrap samples, so we use methods $\text{E}_{(2)}$ and $\text{E}_{(3)}$. Also, The number ABS means the number of Efron bootstrap samples that are including only right-censored observations, no events, so we replace those samples by other Efron bootstrap samples at least including one event time in each. The numbers NA and ABS are out of $1,000,000$; they go down when the sample size increases as shown in Table 2.4. These measures with the smoothed bootstrap method are not applicable because the method generates bootstrap samples including only event time observations.

Table 2.5 shows the chi-squared values obtained from the coverage probabilities for the first quartile. The SB method performs better at all different sample sizes with only one exception when $n = 100$ and the confidence level divided into 20 confidence regions. The differences between the corresponding chi-squared values of the smoothed procedure and Efron's method decrease as the sample size increases.

		10 CR				20 CR			
n	measures	SB	$\mathrm{E}_{(1)}$	$\mathrm{E}_{(2)}$	$\mathrm{E}_{(3)}$	SB	$\mathrm{E}_{(1)}$	$\mathrm{E}_{(2)}$	$\mathrm{E}_{(3)}$
6	χ^2	29.92	934.14	933.96	921.52	97.00	2366.84	2373.40	2355.72
	NA	—	1998	0	0	—	1998	0	0
	ABS	—	493	493	493	—	493	493	493
	P-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
10	χ^2	22.20	876.66	875.62	875.62	57.40	1857.84	1855.64	1855.64
	NA	—	104	0	0	—	104	0	0
	ABS	—	1	1	1	—	1	1	1
	P-value	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000
20	χ^2	11.70	45.98	45.98	45.98	27.44	286.68	286.68	286.68
	NA	—	1	0	0	—	1	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.231	0.000	0.000	0.000	0.095	0.000	0.000	0.000
40	χ^2	11.60	78.76	78.76	78.76	22.48	162.32	162.32	162.32
	NA	—	0	0	0	—	0	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.237	0.000	0.000	0.000	0.261	0.000	0.000	0.000
100	χ^2	9.24	14.98	14.98	14.98	28.48	27.52	27.52	27.52
	NA	—	0	0	0	—	0	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.415	0.091	0.091	0.091	0.075	0.093	0.093	0.093

Table 2.5: The chi-squared values for $Q_1 = 0.117$, and their *P-values* with the corresponding NA and ABS numbers.

When $n = 6$, the smoothed bootstrap's χ^2 values are 29.92 for 10 confidence regions, and 97.00 for 20 confidence regions, while the values corresponding to Efron's method with the three options are equal to 934.14, 933.96 and 921.52 for 10 confidence regions, and 2366.84, 2373.40 and 2355.72 for 20 confidence regions, respectively. When $n = 100$, the corresponding χ^2 values for the SB method are 9.24 for 10 confidence regions, and 28.48 for 20 confidence regions, and for Efron's bootstrap with its three options the values are 14.98 for 10 confidence regions, and 27.52 for 20 confidence regions. This means that Efron's procedure works well when the sample size gets large. The three options, $\mathrm{E}_{(1)}$, $\mathrm{E}_{(2)}$ and $\mathrm{E}_{(3)}$, applied for calculating the first quartile lead to the same χ^2 values when the sample sizes are $n = 20, 40, 100$. This is because we get a time t, so that $\hat{S}(t) = 0.75$ in most bootstrap samples.

		10 CR				20 CR			
n	measures	SB	$E_{(1)}$	$E_{(2)}$	$E_{(3)}$	SB	$E_{(1)}$	$E_{(2)}$	$E_{(3)}$
6	χ^2	157.04	751.88	678.06	557.32	211.40	1640.24	1628.72	1335.80
	NA	—	52306	0	0	—	52306	0	0
	ABS	—	493	493	493	—	493	493	493
	P-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
10	χ^2	40.58	108.58	99.38	76.08	62.80	277.60	262.68	211.36
	NA	—	26739	0	0	—	26739	0	0
	ABS	—	1	1	1	—	1	1	1
	P-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
20	χ^2	3.78	13.08	12.24	11.54	13.88	58.52	57.2	56.44
	NA	—	5141	0	0	—	5141	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.925	0.159	0.200	0.241	0.791	0.000	0.000	0.000
40	χ^2	6.78	17.20	17.20	17.20	29.20	32.92	32.92	32.92
	NA	—	113	0	0	—	113	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.660	0.046	0.046	0.046	0.063	0.025	0.025	0.025
100	χ^2	12.48	11.74	11.74	11.74	45.56	44.16	44.16	44.16
	NA	—	0	0	0	—	0	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.188	0.228	0.228	0.228	0.001	0.001	0.001	0.001

Table 2.6: The chi-squared values for $Q_3 = 0.396$, and their *P-values* with the corresponding NA and ABS numbers.

Table 2.6 presents the outcomes of chi-squared values obtained from the actual coverage probabilities for the true third quartile. The smoothed bootstrap approach provides a better accuracy when $n = 6, 10, 20, 40$ than Efron's method does in both divisions. In other words, the SB technique leads to coverage probabilities closer to the desired nominal sizes than Efron's method does. When $n = 100$, Efron's method provides a better accuracy for the 10 and 20 confidence regions, but differences are small because the corresponding chi-squared values for both bootstrap methods are very close. When $n = 40, 100$, the corresponding χ^2 values to the three options of Efron's method are equal because we get a time t met the condition $\hat{S}(t) = 0.25$ in most bootstrap samples. The NA numbers in the Efron bootstrap samples decrease as the sample size increases. When $n = 6, 10, 20, 40, 100$, the NA number is $52306, 26739, 5141, 113, 0$, respectively.

To illustrate the performances of the two bootstrap methods for the true quartiles when the censoring proportion is increased, we set $p = 0.30$ while keeping the Beta distribution parameters $\alpha = 1.2$ and $\beta = 3.2$ as before. We also fixed the parameter a of the uniform distribution as 0. By using Equation (2.16), the parameter b is equal to 0.91. These values are used in the simulation studies to investigate how the two bootstrap methods perform in terms of the coverage probabilities for the true quartiles. The simulation results are provided in Appendix A.1, and they show that the smoothed bootstrap method mostly outperforms Efron's method for all quartiles, in particular when the sample size is small or medium. Efron's method with a large censoring proportion performs poorly except when the sample size is large.

2.4.2 Second scenario: infinite support

In Subsection 2.4.1, we compared the two bootstrap methods for the case of finite support. When applying the SB method, we sample one observation uniformly from each chosen interval. Those intervals are limited, so it is easy to sample one observation uniformly. If the support is infinite, $[0, \infty)$, we will have one or more intervals of the form $(x_{(i)}, \infty)$, so it will be impossible to draw one observation uniformly from such intervals. To overcome this issue, we need a further assumption for such intervals. BinHimd [10] assumed Exponential tail(s) for the last interval(s), and this assumption will be applied in the following simulations.

Let us first create a right-censored data set from infinite supported distributions. We generate n observations from each distribution of the second scenario, where T_i ∼Weibull(α,β) and C_i ∼Exponential(λ). Then, we can define the right-censored data set through Equations (2.7) and (2.8). We can predefine the censoring proportion, p, in the data sets in order to determine the parameter of the Exponential distribution [85]. We first fix the parameters α and β, then we use the following formula to get the parameter λ.

$$p = \int_0^\infty (1 - \exp(-\lambda x)) \, \frac{\alpha}{\beta} \, (\frac{x}{\beta})^{\alpha-1} \, \exp(-(x/\beta)^\alpha) \, dx \qquad (2.17)$$

It is hard to get a closed form expression for this integral, so we solve it numerically using the R software. Let us set $\alpha = 1.5$, $\beta = 1$, and $p = 0.15$, then λ will be approximately equal to 0.187. These fixed parameters will be used in the following simulations to generate data sets including right-censored observations with different sample sizes to compare the two bootstrap methods. We use the same comparison method as used in Subsection 2.4.1.

Tables 2.7 and 2.8 present the estimated coverage probabilities for the median of Weibull($\alpha = 1.5$,$\beta = 1$), $Q_2 = 0.7832$, in the 10 and 20 confidence regions based on the smoothed bootstrap method, SB, and Efron's bootstrap with its three options $\mathrm{E}_{(1)}$, $\mathrm{E}_{(2)}$ and $\mathrm{E}_{(3)}$ when $n = 6$. The SB method illustrates its superiority in assigning the coverage probabilities in each of the 10 confidence regions close to 0.100 and 0.050 in each of the 20 confidence regions. In contrast, Efron's bootstrap method with the three options, $\mathrm{E}_{(1)}$, $\mathrm{E}_{(2)}$ and $\mathrm{E}_{(3)}$, leads the coverage probabilities far from the nominal level 0.100 in most of the 10 confidence regions and far from 0.050 in most of the 20 confidence regions. As a result, the discrepancy between the estimated and nominal coverage probabilities for Efron's bootstrap method will be very high. The resulting chi-squared values are given in the first row of Table 2.9, and it is obvious that the SB method shows its superiority to Efron's method in achieving the smaller discrepancy between the nominal and actual coverage probabilities at all confidence levels.

Table 2.9 shows the chi-squared values obtained from the coverage probabilities for the median. When $n = 6, 10, 20, 40$ and the number of confidence regions is 10, the generalization of Banks' bootstrap has a better coverage accuracy, and it makes the discrepancy between the nominal and estimated coverage probabilities lower than Efron's method does. The SB technique is still performing well when $n = 100$, but Efron's method is better. This could be due to the influence of the Exponential tail(s) assumed for the end interval(s) when the SB method is applied. When $n = 100$ and the number of confidence regions increases to 20, the SB method provides a better coverage accuracy. From this result, the SB method has a better accuracy when the confidence level is divided into more cells. By observing the χ^2 values as the sample size increases, the ones of Efron's bootstrap decrease while there

$CR_{(i)}$	SB	$\text{E}_{(1)}$	$\text{E}_{(2)}$	$\text{E}_{(3)}$
1	0.115	0.026	0.027	0.028
2	0.106	0.119	0.118	0.120
3	0.118	0.125	0.125	0.123
4	0.100	0.130	0.130	0.128
5	0.087	0.090	0.093	0.095
6	0.101	0.100	0.098	0.099
7	0.102	0.012	0.015	0.014
8	0.091	0.109	0.110	0.109
9	0.093	0.145	0.143	0.143
10	0.087	0.144	0.141	0.141

Table 2.7: The actual coverage probabilities for $Q_2 = 0.7832$, in the 10 confidence regions.

method	SB		$\text{E}_{(1)}$		$\text{E}_{(2)}$		$\text{E}_{(3)}$	
i	$CRL_{(i)}$	$CRR_{(i)}$	$CRL_{(i)}$	$CRR_{(i)}$	$CRL_{(i)}$	$CRR_{(i)}$	$CRL_{(i)}$	$CRR_{(i)}$
1	0.059	0.056	0.009	0.017	0.010	0.017	0.012	0.016
2	0.049	0.057	0.035	0.084	0.035	0.083	0.034	0.086
3	0.057	0.061	0.063	0.062	0.062	0.063	0.061	0.062
4	0.058	0.042	0.066	0.064	0.064	0.066	0.065	0.063
5	0.044	0.043	0.054	0.036	0.059	0.034	0.057	0.038
6	0.065	0.036	0.072	0.028	0.069	0.029	0.072	0.027
7	0.068	0.034	0.004	0.008	0.006	0.009	0.006	0.008
8	0.056	0.035	0.054	0.055	0.055	0.055	0.053	0.056
9	0.054	0.039	0.086	0.059	0.084	0.059	0.083	0.060
10	0.048	0.039	0.086	0.058	0.083	0.058	0.086	0.055

Table 2.8: The actual coverage probabilities for $Q_2 = 0.7832$, in the 20 confidence regions.

is no pattern with the smoothed bootstrap method's chi-squared values. The reason of missing a pattern with the SB method is the sampling process, where we sample uniformly from the limited intervals, and we use the Exponential tail(s) assumed for

		10 CR				20 CR			
n	measures	SB	$E_{(1)}$	$E_{(2)}$	$E_{(3)}$	SB	$E_{(1)}$	$E_{(2)}$	$E_{(3)}$
6	χ^2	10.58	192.48	180.86	179.30	41.48	255.08	239.28	245.44
	NA	—	5968	0	0	—	5968	0	0
	ABS	—	699	699	699	—	699	699	699
	P-value	0.306	0.000	0.000	0.000	0.002	0.000	0.000	0.000
10	χ^2	6.74	115.10	113.94	117.18	14.12	165.36	161.68	165.20
	NA	—	981	0	0	—	981	0	0
	ABS	—	14	14	14	—	14	14	14
	P-value	0.664	0.000	0.000	0.000	0.777	0.000	0.000	0.000
20	χ^2	13.36	33.40	33.40	33.40	34.52	88.72	88.72	88.12
	NA	—	106	0	0	—	106	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.147	0.000	0.000	0.000	0.016	0.000	0.000	0.000
40	χ^2	16.10	17.76	17.76	17.76	31.48	33.24	33.24	33.24
	NA	—	0	0	0	—	0	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.065	0.038	0.038	0.038	0.036	0.023	0.023	0.023
100	χ^2	12.40	6.56	6.56	6.56	24.48	26.16	26.16	26.16
	NA	—	0	0	0	—	0	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.192	0.683	0.683	0.683	0.178	0.126	0.126	0.126

Table 2.9: The chi-squared values for $Q_2 = 0.7832$, and their *P-values* with the NA and ABS numbers.

the end interval(s) to sample from. The three options, $E_{(1)}$, $E_{(2)}$ and $E_{(3)}$, give the same χ^2 values when the sample sizes are 40 and 100 because the median exists for each bootstrap sample.

For the first quartile, the SB method has a better accuracy of all different sample sizes considered as shown in Table 2.10. The differences between the corresponding chi-squared values of the smoothed procedure and Efron's method decrease as the sample size increases in both divisions of confidence level. For example when $n = 6$ and the number of confidence regions is 10, the SB method's chi-squared value is 36.28 while Efron's method with the three options are equal to 959.10, 955.56, and 955.56, respectively. With the largest sample size considered, $n = 100$, the χ^2 value

		10 CR				20 CR			
n	measures	SB	$E_{(1)}$	$E_{(2)}$	$E_{(3)}$	SB	$E_{(1)}$	$E_{(2)}$	$E_{(3)}$
6	χ^2	36.28	959.10	955.56	955.56	96.04	2171.16	2159.92	2159.72
	NA	—	2051	0	0	—	2051	0	0
	ABS	—	699	699	699	—	699	699	699
	P-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
10	χ^2	14.48	756.78	752.94	756.78	60.48	1406.24	1404.48	1406.24
	NA	—	124	0	0	—	124	0	0
	ABS	—	14	14	14	—	14	14	14
	P-value	0.106	0.000	0.000	0.000	0.000	0.000	0.000	0.000
20	χ^2	15.58	73.26	73.26	73.26	43.08	343.96	343.96	343.96
	NA	—	0	0	0	—	0	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.076	0.000	0.000	0.000	0.001	0.000	0.000	0.000
40	χ^2	7.48	98.82	98.82	98.82	22.08	147.52	147.52	147.52
	NA	—	0	0	0	—	0	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.587	0.000	0.000	0.000	0.280	0.000	0.000	0.000
100	χ^2	7.30	13.04	13.04	13.04	23.52	30.88	30.88	30.88
	NA	—	0	0	0	—	0	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.606	0.161	0.161	0.161	0.215	0.042	0.042	0.042

Table 2.10: The chi-squared values for $Q_1 = 0.4358$, and their *P-values* with NA and ABS numbers.

of the SB method is 7.30 while those of Efron's method with its three options are equal to 13.04. This illustrates that Efron's procedure works better when the sample size gets large, as was also shown in Table 2.5 for finite support. The three options, $E_{(1)}$, $E_{(2)}$ and $E_{(3)}$, applied for calculating the first quartiles lead to the same χ^2 values when the sample sizes are $n = 20, 40$, and 100. This is because we get a time t, so that $\hat{S}(t) = 0.75$ in each bootstrap sample. The NA number as shown in this table decreases as the sample size increases. The NA number is 2051 and 124 for $n = 6$ and 10, respectively, while there is no NA values for $n = 20, 40$, and 100. Each Efron's bootstrap sample has a higher chance to have at least one event time observation when the sample size gets larger. The ABS number is 699 and 14 for $n = 6$ and 10, respectively, and it is 0 for $n = 20, 40$ and 100.

		10 CR				20 CR			
n	measures	SB	$\mathrm{E}_{(1)}$	$\mathrm{E}_{(2)}$	$\mathrm{E}_{(3)}$	SB	$\mathrm{E}_{(1)}$	$\mathrm{E}_{(2)}$	$\mathrm{E}_{(3)}$
6	χ^2	23.70	689.00	618.70	549.36	54.96	1323.96	1271.80	1114.52
	NA	—	32959	0	0	—	32959	0	0
	ABS	—	699	699	699	—	699	699	699
	P-value	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000
10	χ^2	25.50	110.32	109.88	112.62	50.92	181.80	194.68	183.24
	NA	—	15032	0	0	—	15032	0	0
	ABS	—	14	14	14	—	14	14	14
	P-value	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000
20	χ^2	12.20	10.82	10.24	10.56	23.76	25.80	27.36	27.56
	NA	—	3019	0	0	—	3019	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.202	0.288	0.331	0.307	0.206	0.136	0.097	0.092
40	χ^2	11.12	11.26	11.26	11.26	22.64	21.16	21.16	21.16
	NA	—	96	0	0	—	96	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.268	0.258	0.258	0.258	0.254	0.328	0.328	0.328
100	χ^2	12.68	8.76	8.76	8.76	26.24	25.64	25.64	25.64
	NA	—	0	0	0	—	0	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.178	0.460	0.460	0.460	0.124	0.141	0.141	0.141

Table 2.11: The chi-squared values for $Q_3 = 1.2433$, and their *P-values* with the NA and ABS numbers.

Table 2.11 outlines the chi-squared values conducted from the coverage probabilities for the third quartile. The SB method provides a better accuracy when $n = 6, 10$ and 40 than Efron's method does in the division of 10 confidence regions. This means that the discrepancy between nominal and estimated coverage probabilities are lower with the SB method. When $n = 20$ and 100, it distributes the actual coverage probability equally over the 10 confidence regions, but Efron's method is better. Also Efron's method is better when $n = 40$ and 100 and the global measure of coverage accuracy is divided into 20 segments. At these sample sizes, the χ^2 values corresponding to the three options $\mathrm{E}_{(1)}$, $\mathrm{E}_{(2)}$, $\mathrm{E}_{(3)}$ of Efron's method are equal because we find the corresponding third quartiles for most bootstrap samples. As the sample size increases, the chi-squared values of Efron's method generally decline, which was also the case of the first and second quartiles in Tables 2.9 and 2.10.

One may want to investigate how the two bootstrap methods perform for different quantiles. Therefore, we repeat the same procedure of comparisons as we did before, but for quantiles 0.10 and 0.90. We use the same seeds and the same parameter values of the generating data sets used for the second scenario, where the censoring proportion is equal to 0.15. The summary of the chi-squared values obtained from the estimated coverage probabilities for $q_{0.10}$ and $q_{0.90}$ is presented in Appendix A.2.

During simulations, some practical notes have been observed, so it is good to present them in order to give a close view for the two bootstrap methods' performances. The first note is that Efron's bootstrap method samples from the original data set, so it is possible for the bootstrap samples to include right-censored observations and ties. Also in some cases, this method generates some samples that include only right-censored observations, so they must be replaced by other bootstrap samples that include at least one event time. In contrast, the generalization of Banks' bootstrap generates samples that contain only events and ties are with probability zero because the method allows to sample from the whole data range. Secondly, the Kaplan-Meier estimator is used in Efron's bootstrap samples to compute the functions of interest, but we sometimes cannot find those functions due to the large proportion of censoring, so we have used the modifying assumptions $\mathrm{E}_{(1)}$, $\mathrm{E}_{(2)}$ and $\mathrm{E}_{(3)}$. Thirdly, due to the process of sampling for the SB method, there is no pattern for the corresponding chi-squared values as the sample size increases. Finally, when the sample size is large, Efron's method mostly performs better than the SB method because the empirical distribution fits the data well.

In the following subsection, we shall use Weibull tails for the last intervals and compare this to the Exponential tails to investigate if this may lead to lower chi-squared values.

2.4.3 Third scenario: infinite support

In Subsection 2.4.2, we assumed Exponential tails for the infinite intervals of the SB method, and we conducted comparisons with Efron's bootstrap method through simulations to study their performances in terms of the coverage probabilities for the quartiles. We found that the smoothed bootstrap achieved better coverage probabilities, in particular when the sample size is small or medium, but in some cases when n is large, Efron's method performs better. We next assume Weibull tails instead to investigate whether this may lead to better coverage probabilities for the same functions of interest.

In the following simulation studies, we assume a standard log-normal distribution for event times and the Weibull($\alpha = 3$, $\beta = 3.7$) distribution for right-censored observations, where the censoring proportion in the generating data sets is set equal to $p = 0.15$. To fit Weibull tails for the last intervals of the smoothed bootstrap method, we assume that each generating data set is from a Weibull distribution with shape parameter α^* and scale parameter β^*, and then we get the maximum likelihood estimates for α^* and β^* through Equation (2.18), which is presented in [56, 70].

$$L = \prod_{i=1}^{n} \left(f(x_i; \alpha^*, \beta^*)\right)^{d_i} \left(S(x_i; \alpha^*, \beta^*)\right)^{1-d_i} \tag{2.18}$$

where $f(.)$ is the probability density function of a Weibull distribution, and $S(.)$ is the survival function of a Weibull distribution. $d_i = 0$ if x_i is a censored observation, and $d_i = 1$ if x_i is an event time.

After estimating the parameters α^* and β^*, we only keep the estimated shape parameter, $\hat{\alpha}^*$, and use it for the last intervals, and neglect the estimated scale parameter, $\hat{\beta}^*$. We do that to have the same shape for all tails assumed for the last intervals, but with different scales depending on the corresponding $M_{X_{n+1}}(x_{(i)}, \infty)$ values assigned to the end intervals. We assume the tail of Weibull($\hat{\alpha}^*$, $\beta_{(i)}$) for any interval in the form of ($x_{(i)}$,∞), with scale parameter $\beta_{(i)}$ obtained by the following formula

$$\beta_{(i)} = \frac{x_{(i)}}{\left(-\ln(M_{X_{n+1}}(x_{(i)}, \infty))\right)^{1/\hat{\alpha}^*}} \tag{2.19}$$

Hence, we use the fitted Weibull tail for the interval ($x_{(i)}$,∞) if it is chosen during the smoothed bootstrap. This new assumption for the tail intervals is referred to by SBw, and we refer to the method with Exponential tails by SB as before.

Table 2.12 presents the chi-squared values obtained from the actual coverage probabilities for $Q_1 = 0.509$ on the base of using the smoothed bootstrap with its

		10 CR					20 CR				
n	measures	SB	SBw	$E_{(1)}$	$E_{(2)}$	$E_{(3)}$	SB	SBw	$E_{(1)}$	$E_{(2)}$	$E_{(3)}$
6	χ^2	24.06	24.38	962.34	946.68	974.80	109.68	110.20	2509.36	2460.76	2539.16
	NA	—	—	5841	0	0	—	—	5841	0	0
	ABS	—	—	1207	1207	1207	—	—	1207	1207	1207
	P-value	0.004	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
10	χ^2	9.04	9.04	917.08	917.62	912.78	67.20	67.20	2286.88	2270.72	2287.48
	NA	—	—	2437	0	0	—	—	2437	0	0
	ABS	—	—	145	145	145	—	—	145	145	145
	P-value	0.434	0.434	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
20	χ^2	14.80	14.80	64.08	64.08	64.08	45.84	45.84	154.84	154.84	154.84
	NA	—	—	11	0	0	—	—	11	0	0
	ABS	—	—	0	0	0	—	—	0	0	0
	P-value	0.097	0.097	0.000	0.000	0.000	0.001	0.001	0.000	0.000	0.000
40	χ^2	19.04	19.04	25.98	25.98	25.98	39.96	39.96	59.12	59.12	59.12
	NA	—	—	0	0	0	—	—	0	0	0
	ABS	—	—	0	0	0	—	—	0	0	0
	P-value	0.025	0.025	0.002	0.002	0.002	0.003	0.003	0.000	0.000	0.000
100	χ^2	9.06	9.06	5.84	5.84	5.84	37.60	37.60	27.56	27.56	27.56
	NA	—	—	0	0	0	—	—	0	0	0
	ABS	—	—	0	0	0	—	—	0	0	0
	P-value	0.432	0.432	0.756	0.756	0.756	0.007	0.007	0.092	0.092	0.092

Table 2.12: The chi-squared values for $Q_1 = 0.509$, and their *P-values* with the NA and ABS numbers.

two assumptions for the tail intervals, SB and SBw, and Efron's method with its three modifications, $E_{(1)}$, $E_{(2)}$ and $E_{(3)}$. A very slight difference occurs between the χ^2 values corresponding to SB and SBw when the sample size is 6 in both divisions of the confidence regions. Both SB and SBw provide the same χ^2 values when $n = 10$, 20, 40 and 100 in both divisions of the global measure of coverage accuracy. This is because the first quartile is located far from the tail interval(s), so the estimated coverage probabilities are not very much affected by the two distributions' tails. Efron's bootstrap approach performs better coverage probabilities only when the sample size is 100 in both divisions of the confidence regions.

From the estimated coverage probabilities for $Q_2 = 1$, Table 2.13 presents the χ^2

		10 CR					20 CR				
n	measures	SB	SBw	$E_{(1)}$	$E_{(2)}$	$E_{(3)}$	SB	SBw	$E_{(1)}$	$E_{(2)}$	$E_{(3)}$
6	χ^2	26.10	26.84	791.26	746.94	936.34	46.28	47.72	879.16	844.68	991.00
	NA	—	—	50287	0	0	—	—	50287	0	0
	ABS	—	—	1207	1207	1207	—	—	1207	1207	1207
	P-value	0.002	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
10	χ^2	8.88	9.36	116.74	108.60	113.76	26.68	27.48	139.08	140.28	146.12
	NA	—	—	19858	0	0	—	—	19858	0	0
	ABS	—	—	145	145	145	—	—	145	145	145
	P-value	0.448	0.405	0.000	0.000	0.000	0.112	0.094	0.000	0.000	0.000
20	χ^2	6.80	6.80	104.44	103.54	107.36	13.32	13.32	135.44	138.48	144.28
	NA	—	—	2854	0	0	—	—	2854	0	0
	ABS	—	—	0	0	0	—	—	0	0	0
	P-value	0.658	0.658	0.000	0.000	0.000	0.822	0.822	0.000	0.000	0.000
40	χ^2	4.62	4.62	42.76	42.76	42.76	12.04	12.04	57.84	57.84	57.84
	NA	—	—	18	0	0	—	—	18	0	0
	ABS	—	—	0	0	0	—	—	0	0	0
	P-value	0.866	0.866	0.000	0.000	0.000	0.884	0.884	0.000	0.000	0.000
100	χ^2	5.96	5.96	12.02	12.02	12.02	12.88	12.88	17.08	17.08	17.08
	NA	—	—	0	0	0	—	—	0	0	0
	ABS	—	—	0	0	0	—	—	0	0	0
	P-value	0.744	0.744	0.212	0.212	0.212	0.845	0.845	0.584	0.584	0.584

Table 2.13: The chi-squared values for $Q_2 = 1$, and their *P-values* with the NA and ABS numbers.

values of the smoothed bootstrap and Efron's method. When the sample sizes are 6 and 10, SB and SBw provide about the same χ^2 values, and they provide identical χ^2 values when $n = 20$, 40 and 100 in both divisions of the confidence regions. These results occur due to two reasons. First, increasing the sample size increases the number of intervals, $n + 1$, partitioning the sample space and the number of infinite intervals is few in comparison to the number of finite intervals. Secondly, the median location is not close to the tail region. Therefore, the coverage probabilities are very slightly influenced by the two tails assumptions, and this is obvious by looking up the corresponding χ^2 values. The smoothed bootstrap with both assumptions for the tail intervals are much better than Efron's method with its three modifications

		10 CR					20 CR				
n	measures	SB	SBw	E$_{(1)}$	E$_{(2)}$	E$_{(3)}$	SB	SBw	E$_{(1)}$	E$_{(2)}$	E$_{(3)}$
6	χ^2	22.44	52.52	1670.96	1516.50	352.86	83.92	185.60	3478.60	3684.96	1100.52
	NA	—	—	218705	0	0	—	—	218705	0	0
	ABS	—	—	1207	1207	1207	—	—	1207	1207	1207
	P-value	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
10	χ^2	24.04	26.68	532.76	590.16	181.30	71.28	77.84	1380.88	1706.72	647.64
	NA	—	—	153425	0	0	—	—	153425	0	0
	ABS	—	—	145	145	145	—	—	145	145	145
	P-value	0.004	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
20	χ^2	10.84	8.62	29.80	39.32	28.86	26.40	23.72	140.28	177.36	99.56
	NA	—	—	79850	0	0	—	—	79850	0	0
	ABS	—	—	0	0	0	—	—	0	0	0
	P-value	0.287	0.473	0.000	0.000	0.001	0.119	0.207	0.000	0.000	0.000
40	χ^2	10.02	9.98	10.08	17.74	19.74	33.28	33.40	56.60	50.72	47.48
	NA	—	—	29450	0	0	—	—	29450	0	0
	ABS	—	—	0	0	0	—	—	0	0	0
	P-value	0.349	0.352	0.344	0.038	0.020	0.022	0.022	0.000	0.000	0.000
100	χ^2	11.94	11.94	5.10	6.24	6.24	27.68	27.68	23.92	24.36	24.36
	NA	—	—	2356	0	0	—	—	2356	0	0
	ABS	—	—	0	0	0	—	—	0	0	0
	P-value	0.217	0.217	0.826	0.716	0.716	0.090	0.090	0.199	0.183	0.183

Table 2.14: The chi-squared values for $Q_3 = 1.963$, and their *P-values* with the NA and ABS numbers.

in terms of making the discrepancy between the estimated and nominal coverage probabilities for the true median small, particularly when $n = 6$, 10, 20 and 40. This is true for both divisions of the confidence regions.

In Table 2.14, we compute the χ^2 values for the third quartile, $Q_3 = 1.963$. When $n = 6$, the smoothed bootstrap method provides very distinct χ^2 values depending on whether we use Exponential or Weibull tails for the end intervals in both divisions of the confidence regions. When the number of confidence regions is 10, the χ^2 value with the Exponential assumption is 22.44 while it is 52.52 when we assume Weibull tails. The corresponding χ^2 value of the SB method is 83.92, and the one of SBw is 185.60 for the case of 20 confidence regions. These huge differences between the χ^2 values occur because Weibull tails are pulled in, so they do not cover the true third quartile. In other words, the tails of Weibull distribution are more likely to be as normal tails because the shape parameters, $\hat{\alpha}^*$, are greater than 1. As the sample size increases to 10, 20 and 40, the differences between the χ^2 values get smaller in both divisions of the confidence regions. When $n = 100$, the χ^2 values of the SB and SBw methods are identical in both divisions of the confidence regions. Efron's bootstrap technique with its three modifications provides better coverage probabilities than the smoothed bootstrap at this sample size, but the smoothed bootstrap method with the two tails assumptions performs better at the other different sample sizes.

The Exponential and Weibull assumptions for the end intervals generally lead to identical results for the three quartiles, in particular when n is large. For a variety of sample sizes, the Weibull assumption provides about the same χ^2 values as Exponential tails, which is obvious from Tables 2.12, 2.13 and 2.14. However, when the sample size is small and the function of interest is Q_3, which is located in or close to the tail part, it is very much better to assume Exponential tails because they are most likely to cover the function of interest. One example for this exception can be presented from Table 2.14. When $n = 6$, the corresponding χ^2 values to SB are 22.44 and 83.92 while the SBw's χ^2 values are 52.52 and 185.60 with the 10 and 20 confidence regions, respectively. Assuming Exponential tails consumes less calculation time; the SBw method needs approximately 47% more computational time than the SB procedure. Therefore, the overall conclusion is that the smoothed bootstrap method with Exponential tails assumption is better than with the Weibull assumption.

2.5 Smoothing using Exponential distributions

In the previous section, when applying the SB method, one observation is sampled uniformly from each interval. We now assume an Exponential tail for each interval regardless whether it is finite or not, then we sample observations from these tails. This assumptions is investigated to show whether it helps to provide lower chi-squared values or not. For each interval, we assume an Exponential distribution with rate parameter $\lambda_{(i)}$ or λ_k^i depending on the form of interval, whether it is $(t_{(i)}, t_{(i+1)})$

or $(c_k^i, t_{(i+1)})$, where $i = 0, 1, \dots, u \leq n$, $1 \leq k \leq l_i$ and the rate parameters can be computed by Equations (2.20) and (2.21).

$$X_{n+1} \sim \text{Exp}(\lambda_{(i)}) : M_{X_{n+1}}(t_{(i)}, t_{(i+1)}) = \exp(-\lambda_{(i)} t_{(i)}) - \exp(-\lambda_{(i)} t_{(i+1)}) \tag{2.20}$$

$$X_{n+1} \sim \text{Exp}(\lambda_k^i) : M_{X_{n+1}}(c_k^i, t_{(i+1)}) = \exp(-\lambda_k^i c_k^i) - \exp(-\lambda_k^i t_{(i+1)}) \tag{2.21}$$

Then we sample from the assumed tails within the bounds of the corresponding intervals. Equations (2.20) and (2.21) are hard to be in simple forms for $\lambda_{(i)}$ and λ_k^i, but it is possible to calculate the rates numerically by R program with `nleqslv` package [45].

For the following simulations, we repeat the same comparisons as in Subsection 2.4.1, but we perform the simulation using this new assumption, which is referred to by SB$_{exp}$. Recall that we assume the Beta distribution with $\alpha = 1.2$ and $\beta = 3.2$ for event time observations, and the uniform distribution with $a = 0$ and $b = 1.82$, for right-censored observations. The censoring proportion in the generating data sets is set equal to 0.15.

Table 2.15 reports the chi-squared values obtained from the coverage probabilities for the first quartile Q_1 based on the two assumptions of sampling for the smoothed bootstrap method and Efron's method with its three options. When the sample size is small, $n = 6, 10$, a clear difference is observed between the corresponding chi-squared values of the two ways of sampling for the smoothed bootstrap method, SB$_{exp}$ and SB. As a consequence of increasing the sample size, the difference between the chi-squared values decreases, and they are nearly identical when n gets large. This is of course because the intervals' lengths become shorter as the sample size increases, so the two assumptions of sampling lead to the same results. The smoothed bootstrap method with its two modifications provides better accuracy than Efron's method in all cases except when $n = 100$ and the division includes 20 confidence regions.

Table 2.16 lists the chi-squared values for the second quartile Q_2 based on the two bootstrap methods with their related modifications. For the corresponding options to the smoothed bootstrap method, a small difference occurs between their corresponding chi-squared values when the sample size is small, then the difference

		10 CR					20 CR				
n	measures	SB$_{exp}$	SB	E$_{(1)}$	E$_{(2)}$	E$_{(3)}$	SB$_{exp}$	SB	E$_{(1)}$	E$_{(2)}$	E$_{(3)}$
6	χ^2	34.90	27.44	936.00	931.78	919.72	96.24	98.12	2370.80	2375.96	2358.80
	NA	—	—	2324	0	0	—	—	2324	0	0
	ABS	—	—	623	623	623	—	—	623	623	623
	P-value	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
10	χ^2	13.72	20.82	898.34	897.22	897.22	52.80	62.60	1919.40	1917.04	1917.04
	NA	—	—	127	0	0	—	—	127	0	0
	ABS	—	—	1	1	1	—	—	1	1	1
	P-value	0.133	0.013	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
20	χ^2	14.16	14.22	46.26	46.26	46.26	32.24	32.40	287.36	287.36	287.36
	NA	—	—	0	0	0	—	—	0	0	0
	ABS	—	—	0	0	0	—	—	0	0	0
	P-value	0.117	0.115	0.000	0.000	0.000	0.029	0.028	0.000	0.000	0.000
40	χ^2	8.14	8.30	70.12	70.12	70.12	18.52	18.24	158.44	158.44	158.44
	NA	—	—	0	0	0	—	—	0	0	0
	ABS	—	—	0	0	0	—	—	0	0	0
	P-value	0.520	0.504	0.000	0.000	0.000	0.488	0.506	0.000	0.000	0.000
100	χ^2	11.80	11.36	13.58	13.58	13.58	29.68	30.32	26.28	26.28	26.28
	NA	—	—	0	0	0	—	—	0	0	0
	ABS	—	—	0	0	0	—	—	0	0	0
	P-value	0.225	0.252	0.138	0.138	0.138	0.056	0.048	0.123	0.123	0.123

Table 2.15: The chi-squared values for $Q_1 = 0.117$, and their *P-values* with the corresponding the NA and ABS numbers.

almost disappears for large n. In other words, the two ways of sampling applied on the smoothed bootstrap method provide about the same coverage probabilities when the sample size is large, so as a result they have similar chi-squared values. In this table, it is clear that the SB and SB$_{exp}$ methods are better in terms of defining the coverage probabilities for the median when $n = 6, 10$ and 20, and they provide good coverage probabilities when $n = 40$ and 100. However, Efron's method mostly provides better outcomes for large n.

Table 2.17 presents the chi-squared values obtained from the coverage probabilities for the third quartile Q_3 based on the smoothed bootstrap method with the two sampling techniques and Efron's method with its three options. The SB and

		10 CR					20 CR				
n	measures	SB$_{exp}$	SB	E$_{(1)}$	E$_{(2)}$	E$_{(3)}$	SB$_{exp}$	SB	E$_{(1)}$	E$_{(2)}$	E$_{(3)}$
6	χ^2	45.48	42.80	204.42	186.96	196.40	55.60	54.08	230.56	209.76	226.96
	NA	—	—	9407	0	0	—	—	9407	0	0
	ABS	—	—	623	623	623	—	—	623	623	623
	P-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
10	χ^2	20.82	24.02	159.46	152.62	151.80	26.84	29.60	174.72	168.84	167.28
	NA	—	—	2490	0	0	—	—	2490	0	0
	ABS	—	—	1	1	1	—	—	1	1	1
	P-value	0.013	0.004	0.000	0.000	0.000	0.108	0.057	0.000	0.000	0.000
20	χ^2	6.08	6.12	56.86	57.88	57.88	17.76	15.36	72.36	73.72	73.72
	NA	—	—	96	0	0	—	—	96	0	0
	ABS	—	—	0	0	0	—	—	0	0	0
	P-value	0.732	0.728	0.000	0.000	0.000	”0.539 ”	0.699	0.000	0.000	0.000
40	χ^2	10.74	10.84	6.26	6.26	6.26	14.68	14.92	16.28	16.28	16.28
	NA	—	—	0	0	0	—	—	0	0	0
	ABS	—	—	0	0	0	—	—	0	0	0
	P-value	0.294	0.287	0.714	0.714	0.714	0.743	0.728	0.639	0.639	0.639
100	χ^2	9.98	9.86	6.48	6.48	6.48	19.56	19.88	13.40	13.40	13.40
	NA	—	—	0	0	0	—	—	0	0	0
	ABS	—	—	0	0	0	—	—	0	0	0
	P-value	0.352	0.362	0.691	0.691	0.691	0.421	0.402	0.817	0.817	0.817

Table 2.16: The chi-squared values for $Q_2 = 0.236$, and their *P-values* with the corresponding the NA and ABS numbers.

SB$_{exp}$ methods provide about the same chi-squared results in all cases while the options of Efron's method have different outcomes when the sample size is small, and they provide identical results when n gets large due to having corresponding third quartile to almost all bootstrap samples, the NA number is zero or close to zero. Both smoothed bootstrap methods are better in defining the coverage probabilities in the 10 and 20 confidence regions when $n = 6, 10, 20$ and 40, but Efron's method is better at the sample size 100 in both divisions of the confidence regions. This is because the empirical distribution fits the data well, and as a consequence, we obtain smaller chi-squared values than those of the SB and SB$_{exp}$ methods. From Tables 2.15, 2.16 and 2.17, it seems that the SB and SB$_{exp}$ methods provide about

		10 CR					20 CR				
n	measures	SB$_{exp}$	SB	E$_{(1)}$	E$_{(2)}$	E$_{(3)}$	SB$_{exp}$	SB	E$_{(1)}$	E$_{(2)}$	E$_{(3)}$
6	χ^2	150.04	148.04	712.70	648.10	533.68	207.28	199.20	1583.12	1586.80	1295.72
	NA	—	—	53422	0	0	—	—	53422	0	0
	ABS	—	—	623	623	623	—	—	623	623	623
	P-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
10	χ^2	40.90	41.34	101.66	91.66	67.40	69	63.88	268.44	251.96	199.84
	NA	—	—	26855	0	0	—	—	26855	0	0
	ABS	—	—	1	1	1	—	—	1	1	1
	P-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
20	χ^2	5.26	4.70	14.46	14.34	13.50	17.60	16.72	54.00	52.20	51.24
	NA	—	—	5027	0	0	—	—	5027	0	0
	ABS	—	—	0	0	0	—	—	0	0	0
	P-value	0.811	0.860	0.107	0.111	0.141	0.549	0.609	0.000	0.000	0.000
40	χ^2	10.82	10.44	19.14	19.14	19.14	32.28	30.88	39.04	39.04	39.04
	NA	—	—	111	0	0	—	—	111	0	0
	ABS	—	—	0	0	0	—	—	0	0	0
	P-value	0.288	0.316	0.024	0.024	0.024	0.029	0.042	0.004	0.004	0.004
100	χ^2	12.58	12.58	9.74	9.74	9.74	41.32	41.24	37.60	37.60	37.60
	NA	—	—	0	0	0	—	—	0	0	0
	ABS	—	—	0	0	0	—	—	0	0	0
	P-value	0.183	0.183	0.372	0.372	0.372	0.002	0.002	0.007	0.007	0.007

Table 2.17: The chi-squared values for $Q_3 = 0.396$, and their *P-values* with the corresponding the NA and ABS numbers.

the same results.

2.6 Estimates for the survival function

The interest of analyzing survival data arises in many important and interesting applications related to medicine, engineering, economics and biology. The Kaplan-Meier (KM) estimator [53] has been widely used for inferences on the survival function. In this section, the proposed smoothed bootstrap method is used to estimate the survival function, and the results are compared to the KM estimates. Pointwise bootstrap confidence intervals for the survival function are also derived. The presentation is based on two data sets from the literature [6, 81].

Interval	M-function
(0,2)	0.0909
(2,4)	0.0909
(4,14)	0.0909
(14,24)	0.0909
(21+,24)	0.0130
(24,27)	0.1039
(27,33)	0.1039
(33,51)	0.1039
(51,∞)	0.1039
(60+,∞)	0.0519
(72+,∞)	0.1558

Table 2.18: The M-function values for the 11 intervals created by Bartholomew's data set.

Example 2.6.1 A lifetime data set for 10 pieces of equipment was provided by Bartholomew [6]; the collected data points are $2, 4, 14, 21+, 24, 27, 33, 51, 60+, 72+$, where the + sign indicates a right-censored observation. We want to estimate the lifetime distribution of the equipment by the smoothed bootstrap method, as presented in Section 2.3. Table 2.18 shows the 11 intervals and corresponding M-function values, according to Equations (2.1) and (2.2), for the SB method.

We create $B = 1000$ bootstrap samples of size $n = 10$, and derive the empirical survival function of each sample. This leads to 1000 empirical survival functions, so there are 1000 survival function estimates at each time t. We take the averages of these estimates at times $t = 0, 1, 2, 3, \ldots, 100$ to show the estimated survival curve, and consider these resulting values as the bootstrap estimate $\hat{S}^*_{boot}(t)$ at those times t. Figure 2.1 presented the estimated survival function.

Figure 2.1 shows the estimated survival function based on the SB method (orange line) and the Kaplan-Meier estimator (black step function). It is obvious that the survival function estimate by our proposed bootstrap method is smoother than the Kaplan-Meier estimate because the latter only decreases at the 7 observed event

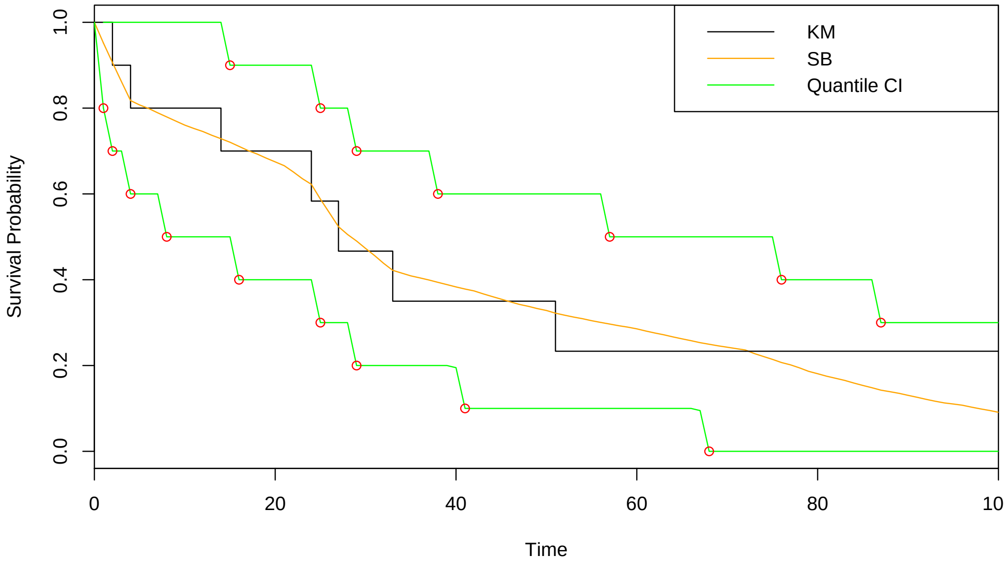

Figure 2.1: The survival curves for Bartholomew's data set based on the smoothed bootstrap method with its 90% bootstrap quantile confidence intervals, and the KM estimate.

times. At each observed event time, the smoothed bootstrap estimate is greater than the Kaplan-Meier estimate, and the smoothed bootstrap estimate crosses the Kaplan-Meier estimate between each two observed event times. The two green lines present the 90% confidence intervals for the survival function based on the SB method at times $t = 0, 1, 2, 3, \ldots, 100$. These are derived by taking the 50th and 950th ordered values of the 1000 empirical survival functions resulting from the application of the SB method. At each value of t, the method provides $\hat{S}^*_{(50)}(t)$ and $\hat{S}^*_{(950)}(t)$, where the former is the lower bound and the latter is the upper bound for the confidence interval. These bounds of the bootstrap confidence intervals are actually step functions because each empirical survival function based on a SB sample, at any time t, can be equal to only one of the 11 values $0, 0.1, \ldots, 0.9, 1$.

Example 2.6.2 To study the performances of the KM estimator and the smoothed bootstrap method with large sample size, we use the lung cancer data, which was provided by Therneau [81] and it is available in the `survival` package in R. This data contains the survival times for 138 patients in days since the cancer was detected. In

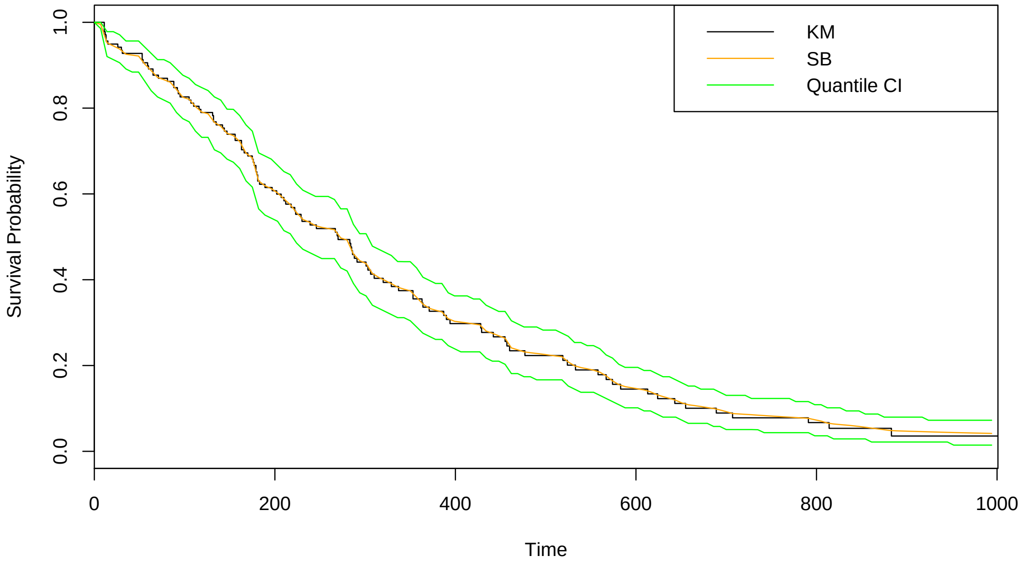

Figure 2.2: The survival curves for lung cancer data based on the smoothed bootstrap method with its 90% quantile confidence intervals, and the KM estimator.

this data set, the censoring proportion is 0.188, and there are 17 tied observations. These ties are broken by adding a small value to some of the tied observations in order to apply our smoothed bootstrap method. It is possible to deal with the ties in several ways [41], but this is not discussed further in this thesis because we think the resulting inferences are almost identical. The influence of the ways that ties are dealt with is left as a future work. Effectively, if one has a tie between two or more observed event times, this time value has a positive probability to be chosen for the bootstrap samples. If there is a tie between an observed event time and a right-censored observation, then we assume that the censored time has occurred fractionally later than the event time. If we get ties between multiple right-censored times, this does not really affect the method.

Figure 2.2 presents the estimated survival function for the lung cancer data based on the smoothed bootstrap method and the Kaplan-Meier estimate. Because the samples size is large and we have many event times, these two estimates are nearly identical. The two green lines indicate the 90% quantile confidence intervals for the survival function at each time t, which are narrower in this example than the confidence intervals in Example 2.6.1, due to the larger sample size. It should be noted that these bounds are still step functions, but of course the steps are quite small, so the functions are quite smooth.

It is apparent that the SB method provides a smoother survival function estimate than the Kaplan-Meier estimate as the KM curve only decreases at the observed event times, and this is obvious when the sample size is small. The bounds of the bootstrap confidence intervals are actually step functions, because each empirical survival function at any time t can be equal to only one of $n+1$ values. When the sample size is large with many observed event times, the estimates based on the two methods are nearly identical as appeared in Example 2.6.2.

To obtain smoothed pointwise confidence intervals for the survival function, we can use the linear and log-transformed confidence intervals as presented by Borgan and Liestøl [12] and Klein and Moeschberger [54], which for confidence level $100(1-2\alpha)\%$, have the property that the true survival function at time t falls in the interval with approximate confidence $100(1-2\alpha)\%$. We will study the performance of such confidence bounds based on our smoothed bootstrap method through simulations in Section 2.7. The $100(1-2\alpha)\%$ linear pointwise confidence interval can be calculated by

$$\left(\hat{S}(t) \mp Z_{(1-2\alpha)} \times SD(\hat{S}(t))\right) \tag{2.22}$$

and the $100(1-2\alpha)\%$ log-transformed pointwise confidence interval can be calculated by

$$\left((\hat{S}(t))^{\frac{1}{\theta}}, (\hat{S}(t))^{\theta}\right) \tag{2.23}$$

where $\theta = \exp(\frac{Z_{(1-2\alpha)} \times SD(\hat{S}(t))}{\hat{S}(t) \times \ln(\hat{S}(t))})$ and $Z_{(1-2\alpha)}$ is the $(1-2\alpha)$ percentile of the standard normal distribution. $SD(\hat{S}(t))$ is the standard deviation of the estimated survival function, where the estimate can be either based on the smoothed bootstrap method or the Kaplan-Meier estimate. If specified lower bound becomes negative or the upper bound exceeds 1, we restrict these intervals to be in $[0,1]$. Note that the linear confidence interval is symmetric about the survival function estimate as long as the specified bounds are inside the interval $[0,1]$, but the log-transformed confidence interval is not symmetric about the survival function estimate.

To construct either one of the smoothed pointwise confidence intervals for a

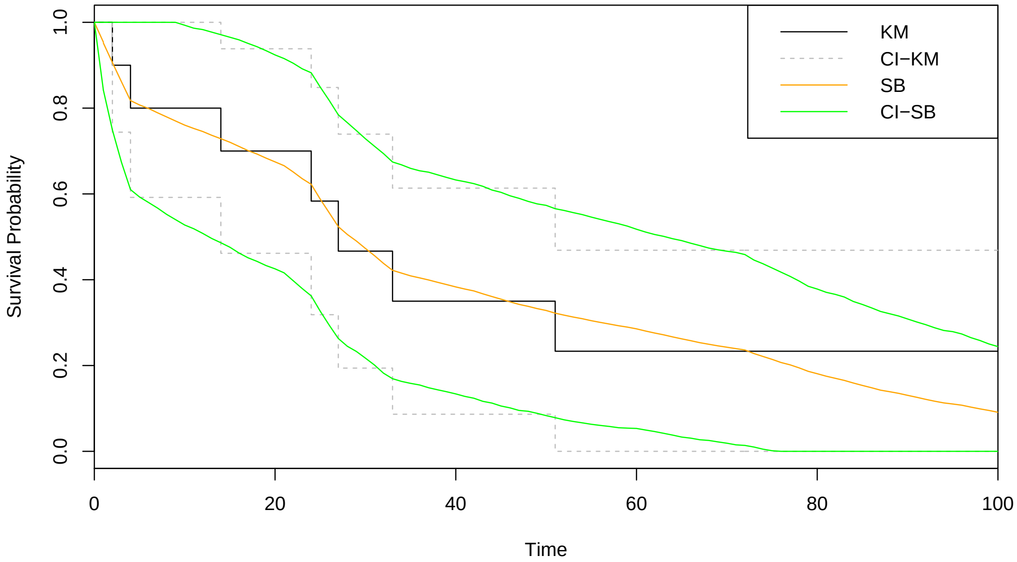

Figure 2.3: The estimated survival functions for the equipment data based on the smoothed bootstrap method and the KM estimator with their 90% linear confidence intervals.

survival function at time t, the standard deviation of the survival function estimate at time t should be computed first. For the survival function estimate based on the smoothed bootstrap at time t, $\hat{S}^*_{boot}(t)$, the standard deviation can be calculated by Equation (1.2), with $\hat{\theta}^{*j}$ replaced by $\hat{S}^{*j}(t)$. For the Kaplan-Meier estimate, this is provided by the well-known Greenwood's formula [42, 53]. Then Equations (2.22) and (2.23) can be used to derive the linear and the log-transformed confidence intervals for both estimation methods discussed here. Next we revisit Examples 2.6.1 and 2.6.2 to show the smoothed confidence intervals.

For the data of Example 2.6.1, Figure 2.3 presents the smoothed bootstrap and the Kaplan-Meier estimates of the survival function, together with their corresponding 90% linear confidence intervals. The smoothed bootstrap method provides smooth confidence interval bounds for the survival function, and this is contrary to the confidence intervals' bounds based on the KM estimate. This is due to the fact that the KM estimate drops only at the event times, and this leads to the confidence intervals' bounds being step functions. The confidence intervals are mostly of simi-

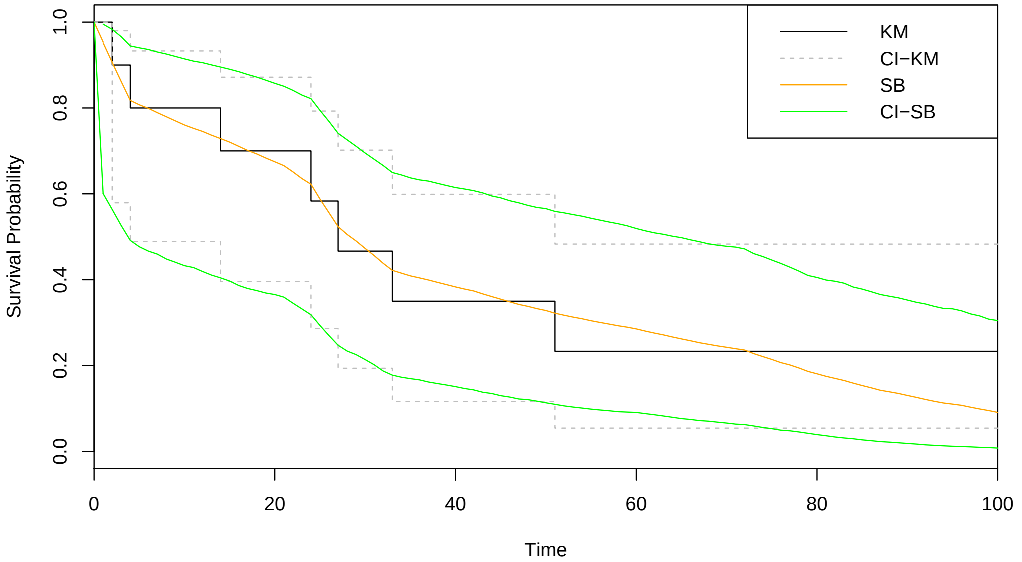

Figure 2.4: The estimated survival functions for the equipment data based on the smoothed bootstrap method and the KM estimator with their 90% log-transformed confidence intervals.

lar width, however for values of t beyond the largest observation, the Kaplan-Meier estimate and its confidence interval remain constant while the smoothed bootstrap estimate decreases and its confidence interval becomes more narrow.

The log-transformed confidence interval is better than the linear confidence interval in the sense of coverage accuracy [12], in particular when the confidence level is high; for example 0.90 or 0.95. Figure 2.4 presents the 90% log-transformed confidence interval for the survival function based on the same estimates used to present the 90% linear confidence interval in Figure 2.3. It is clear that the confidence interval bounds based on the Kaplan-Meier estimate are step functions with steps only at the observed event times, and this is contrary to the bounds based on the smoothed bootstrap estimate, which are more smooth. Both confidence intervals, whether they are based on the SB method or the KM estimator, have similar pattern, and this can be seen in Figures 2.3 and 2.4. The main difference is that the symmetry is not observed with the log-transformed confidence interval around the survival function while the linear confidence interval is symmetric as presented in Figure 2.3 (except

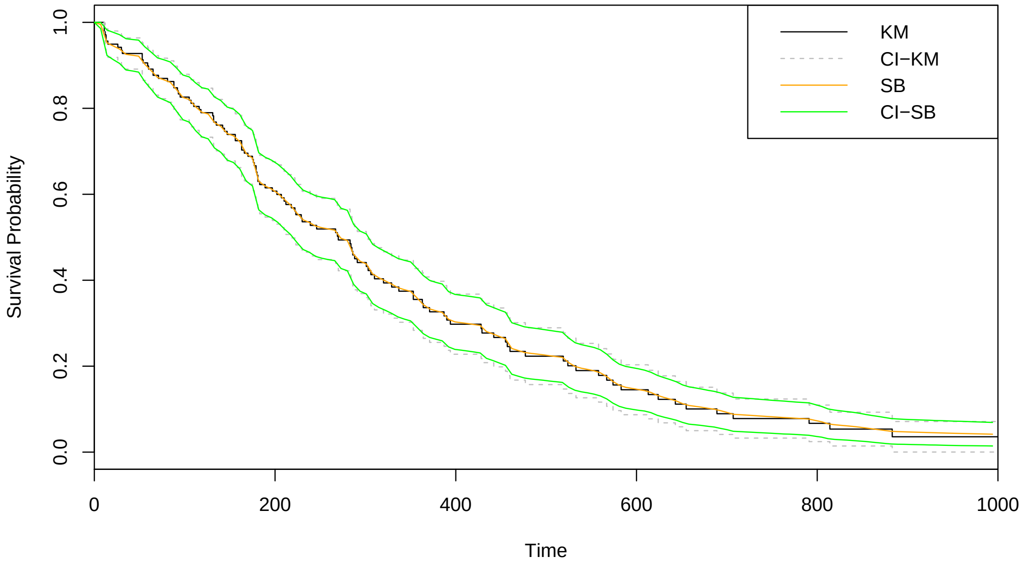

Figure 2.5: The survival curves for the lung cancer data based on the smoothed bootstrap method and the KM estimator with their 90% linear confidence intervals.

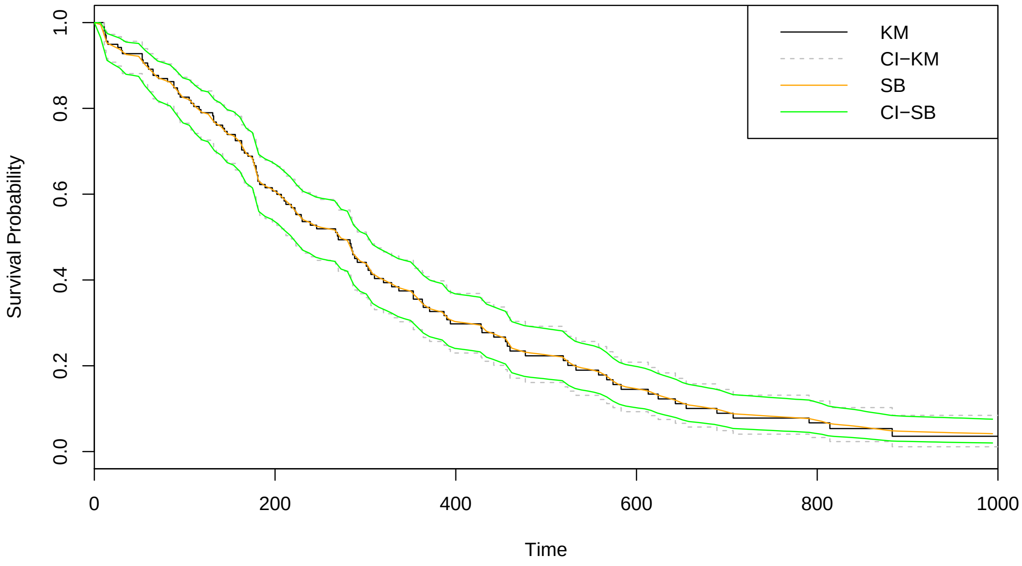

Figure 2.6: The survival curves for the lung cancer data based on the smoothed bootstrap method and the KM estimator with their 90% log-transformed confidence intervals.

in the tails due to the logical restriction to the interval $[0, 1]$). We can conclude that these two pointwise confidence intervals can have smooth bounds when using the smoothed bootstrap estimate of the survival function, which is intuitively attractive. Of course, due to the small sample size in this data set, the Kaplan-Meier estimate and the smoothed bootstrap estimate for the survival function are quite different, and this leads to have quite different confidence intervals. Detailed investigation of the performance of these bounds in terms of coverage of the real underlying survival function is studied in the next section.

For the data in Example 2.6.2, Figures 2.5 and 2.6 present the linear and log-transformed confidence intervals, respectively, for the survival function estimated by the smoothed bootstrap method and the Kaplan-Meier estimator. In Figure 2.2, we observed that the survival function estimates by the two methods are quite similar, and Figures 2.5 and 2.6 show that the corresponding confidence intervals are very similar as well. The confidence intervals based on the smoothed bootstrap method and the KM estimator are nearly identical. From this situation, it is possible to conclude that for large sample size with many observed event times, both methods are likely to provide about the same outcomes, and this is particularly useful because of the well-known excellent large sample properties of the Kaplan-Meier estimator [54], which therefore also seem to hold for the SB method.

2.7 A smoothed Kaplan-Meier bootstrap method

The bootstrap method for right-censored data, presented by Efron [31], has been widely used in the literature for survival inferences, see e.g. [43, 63]. Bilker and Wang [9] used the bootstrap method to obtain confidence bands for the bootstrap estimate of the survival curve, and Heller and Venkatraman [47] used it for testing. This bootstrap method is easy to implement, but it does not perform well if the sample size is small due to the ties and censored observations occurring in the bootstrap samples [63]. Therefore, one can consider a new alternative smoothed bootstrap method which basically relies on the Kaplan-Meier estimator. In this section, we present the alternative bootstrap method and compare it to our smoothed bootstrap method through simulation studies.

The alternative method is based on the Kaplan-Meier estimate, and the method is referred to by SBa. Suppose that there are n observations including u events and $v = n - u$ right-censored observations with no ties. The n observations are ordered, then based on the event data points, $u + 1$ intervals of the form $(t_{(i)},t_{(i+1)})$ are created, where $0 \leq i \leq u$ and $t_{(0)} = 0$ and $t_{(u+1)} = \infty$. Each interval of the form $(t_{(i)},t_{(i+1)})$ is specified with probability $\mathrm{P}_{(i)}$, where the probability $\mathrm{P}_{(i)}$ can be calculated by

$$\mathrm{P}_{(i)} = \mathrm{P}(t_{(i)} < t \leq t_{(i+1)}) = \hat{S}(t_{(i)}) - \hat{S}(t_{(i+1)}) \tag{2.24}$$

where $\hat{S}(t_{(i)})$ and $\hat{S}(t_{(i+1)})$ are the KM estimates for the survival functions at fixed times $t_{(i)}$ and $t_{(i+1)}$, respectively.

For clarity, we present this alternative bootstrap algorithm for right-censored data in the following steps:

(i) Order the data set, and add the end points of the possible data range, $t_{(0)} = 0$ and $t_{(u+1)} = +\infty$ (or $t_{(u+1)} = b$ in case of finite support is $[0, b]$), so we have $t_{(0)} < t_{(1)} < \ldots < t_{(u)} < t_{(u+1)}$.

(ii) Create $u + 1$ intervals among the n observations of the data set on the form of $(t_{(i)}, t_{(i+1)})$, $i = 0, \ldots, u$.

(iii) Compute the $\mathrm{P}_{(i)}$ probabilities by Equation (2.24) for the $u + 1$ intervals.

(iv) Sample with replacement n intervals with the assignment probabilities, then draw one observation from each chosen interval to obtain one bootstrap sample of size n. If the interval is finite, we sample uniformly; otherwise, we use a fitted Exponential tail for sampling as describe is Section 2.3.

(v) Use the bootstrap sample to calculate the empirical survival function at a fixed time t, $\hat{S}^*(t)$.

(vi) Repeat steps (iv) and (v) B times in order to have B survival function values at a fixed time t.

It is important to note that if the last observation of the original data set is an event time, then $\hat{S}(t_{(u)}) = 0$. This causes that the last interval $(t_{(u)}, t_{(u+1)})$ will never be selected during the alternative bootstrap procedure because $\mathrm{P}_{(u)} = 0$. If this case occurs, the SBa method performs poorly for the survival function at time t being greater than $t_{(u)}$ because the estimated survival function at that time t is zero in all bootstrap samples. If we have right-censored observation(s) after $t_{(u)}$, then we fit an Exponential tail with rate parameter λ^* for the end interval [13], with λ^* set of

$$\hat{\lambda}^* = \frac{-\ln(\hat{S}(t_{(u)}))}{t_{(u)}} \tag{2.25}$$

Then we sample an observation greater than $t_{(u)}$ for the last interval. Note that ties and censored observations can occur with probability zero in the bootstrap samples.

In the following simulation studies, we use Weibull($\alpha = 1.5$, $\beta = 1$) to generate the event time observations, and we use Exponential distribution with rate parameter $\lambda = 0.187$ to generate the right-censored observations, where the censoring proportion in the generating data sets is 0.15. We investigate the performances of the quantile, linear and log-transformed pointwise confidence intervals with 90% confidence level in combination with the smoothed bootstrap method and the smoothed Kaplan-Meier bootstrap method to find the estimated coverage proportions for the true survival function at 19 distinct fixed times, as listed in the first column in Tables 2.19, 2.20 and 2.21. These 19 values of t are equally spaced quantiles of the underlying distribution, so this allows to discover which bootstrap method performs better in terms of the coverage proportions for the survival function at different locations.

Next, $N = 1000$ data sets with sample size n are generated. Then, each bootstrap method is applied $B = 1000$ times on each generated data set. We then derive the empirical survival function corresponding to each bootstrap sample, and this leads to 1000 empirical survival functions. We compute the average and the standard deviation of the 1000 survival function estimates at each value of t to conduct the 90% linear and log-transformed confidence intervals by Equations (2.22) and (2.23). Also, we compute the 90% quantile confidence interval $[\hat{S}^*_{(50)}(t), \hat{S}^*_{(950)}(t)]$. We count which confidence intervals include the true survival function at each value of t. We repeat this procedure for all 1000 generated data sets in order to observe the actual

$n =$		6		10		20		40		100	
t	$S(t)$	SB	SBa	SB	SBa	SB	SBa	SB	SBa	SB	SBa
0.138	0.95	0.998	0.994	0.985	0.978	0.993	0.991	0.988	0.986	0.896	0.899
0.223	0.90	0.979	0.954	0.995	0.988	0.918	0.895	0.948	0.944	0.884	0.895
0.298	0.85	0.938	0.877	0.967	0.942	0.894	0.883	0.943	0.942	0.879	0.882
0.368	0.80	0.972	0.956	0.923	0.872	0.890	0.871	0.947	0.936	0.911	0.916
0.436	0.75	0.965	0.913	0.908	0.885	0.886	0.858	0.937	0.932	0.879	0.883
0.503	0.70	0.936	0.855	0.947	0.928	0.930	0.918	0.888	0.890	0.869	0.865
0.570	0.65	0.886	0.881	0.899	0.856	0.872	0.842	0.916	0.909	0.876	0.879
0.639	0.60	0.909	0.857	0.943	0.902	0.925	0.906	0.876	0.873	0.868	0.869
0.710	0.55	0.913	0.819	0.896	0.838	0.870	0.831	0.906	0.897	0.887	0.890
0.783	0.50	0.956	0.897	0.954	0.892	0.879	0.830	0.873	0.862	0.874	0.881
0.861	0.45	0.900	0.821	0.904	0.826	0.880	0.825	0.905	0.885	0.902	0.897
0.943	0.40	0.923	0.772	0.938	0.876	0.878	0.832	0.874	0.844	0.878	0.873
1.033	0.35	0.897	0.691	0.901	0.800	0.918	0.892	0.908	0.899	0.900	0.903
1.132	0.30	0.902	0.765	0.945	0.854	0.880	0.801	0.879	0.859	0.904	0.898
1.243	0.25	0.930	0.694	0.899	0.762	0.868	0.774	0.902	0.883	0.902	0.880
1.373	0.20	0.935	0.601	0.894	0.686	0.877	0.729	0.905	0.865	0.899	0.891
1.533	0.15	0.905	0.552	0.916	0.665	0.937	0.815	0.913	0.835	0.908	0.885
1.744	0.10	0.943	0.422	0.930	0.503	0.896	0.619	0.926	0.818	0.909	0.868
2.078	0.05	0.967	0.230	0.955	0.357	0.925	0.459	0.900	0.617	0.863	0.720

Table 2.19: The actual coverage proportions for the true survival function at each value of t in the 90% quantile confidence intervals based on the two smoothed bootstrap methods.

coverage proportions for the true survival function at each value of t in the three pointwise confidence intervals based on the two smoothed bootstrap methods.

Table 2.19 shows the simulation results for the true survival function at each value of t in the 90% quantile confidence intervals based on the smoothed bootstrap method and the smoothed Kaplan-Meier bootstrap method, SBa. The SB method performs better than the SBa method for the true survival functions at almost all values of t, in particular when the sample size is small. At most values

$n =$		6		10		20		40		100	
t	$S(t)$	SB	SBa	SB	SBa	SB	SBa	SB	SBa	SB	SBa
0.138	0.95	0.998	0.989	0.992	0.985	0.985	0.981	0.954	0.952	0.908	0.903
0.223	0.90	0.987	0.970	0.988	0.973	0.938	0.943	0.905	0.902	0.894	0.897
0.298	0.85	0.981	0.945	0.962	0.951	0.905	0.895	0.898	0.897	0.881	0.881
0.368	0.80	0.969	0.930	0.922	0.922	0.889	0.892	0.898	0.894	0.890	0.886
0.436	0.75	0.955	0.908	0.898	0.872	0.886	0.876	0.909	0.901	0.884	0.888
0.503	0.70	0.935	0.878	0.895	0.862	0.878	0.860	0.891	0.881	0.872	0.863
0.570	0.65	0.911	0.848	0.888	0.852	0.881	0.858	0.882	0.870	0.879	0.877
0.639	0.60	0.895	0.837	0.880	0.846	0.869	0.855	0.880	0.867	0.867	0.871
0.710	0.55	0.913	0.827	0.890	0.827	0.865	0.849	0.873	0.860	0.891	0.892
0.783	0.50	0.918	0.803	0.882	0.837	0.881	0.852	0.866	0.851	0.881	0.875
0.861	0.45	0.910	0.789	0.894	0.812	0.862	0.844	0.861	0.841	0.873	0.871
0.943	0.40	0.899	0.762	0.901	0.796	0.872	0.841	0.876	0.851	0.872	0.864
1.033	0.35	0.906	0.733	0.894	0.791	0.878	0.827	0.882	0.863	0.889	0.868
1.132	0.30	0.919	0.719	0.897	0.771	0.875	0.827	0.881	0.842	0.889	0.870
1.243	0.25	0.926	0.670	0.889	0.740	0.869	0.790	0.870	0.828	0.883	0.863
1.373	0.20	0.936	0.615	0.895	0.716	0.882	0.758	0.860	0.808	0.880	0.855
1.533	0.15	0.950	0.534	0.913	0.654	0.874	0.712	0.868	0.788	0.882	0.860
1.744	0.10	0.955	0.413	0.935	0.544	0.900	0.671	0.882	0.753	0.881	0.830
2.078	0.05	0.974	0.242	0.965	0.361	0.950	0.490	0.914	0.657	0.871	0.743

Table 2.20: The actual coverage proportions for the true survival function at each value of t in the 90% linear confidence intervals based on the two smoothed bootstrap methods.

of t, the SB method provides over-coverage. As the sample size increases, the discrepancy between the simulation results and the nominal size 0.90 at all values of t decreases. For the true survival functions at the values of t located in the tail region, $2.078, 1.744, 1.533, 1.373, 1.243$, the SBa method performs poorly while the SB method shows its advantage in making the discrepancy between the simulation results and the nominal size 0.90 small, particularly for sample sizes $6, 10$ and 20. The results based on the SBa method are improved when the sample size gets large, and this is apparent when $n = 40$ and 100.

Table 2.20 reports the actual coverage proportions for the true survival function at each value of t in the 90% linear confidence intervals based on the two smoothed bootstrap methods. The SB method provides better actual coverage proportions for the true survival functions at almost all values of t and for different sample sizes. For sample sizes 6 and 10, it provides over-coverage at most values of t, particularly at small and large values of t. When the sample size gets larger, the discrepancies between the actual coverage proportions and the nominal size 0.90 decrease at all values of t. The SBa method performs poorly when the sample size is small, it gets better when n increases. It does not provide a good performance to define the actual coverage proportions for the true survival functions at values of t located in the tail region. For example, when $t = 2.078$, the actual coverage proportions based on the SBa method are $0.242, 0.361, 0.490, 0.657, 0.743$ when $n = 6, 10, 20, 40, 100$, respectively. This is because this method provides only one Exponential tail if we have censored observation(s) after $t_{(u)}$. Also, if the greatest observation of a generating data set is event, then $(t_{(u)}, t_{(u+1)})$ will never be chosen during the alternative procedure because $\mathrm{P}_{(u)} = 0$. Therefore, the estimated survival function at any time $t \geq t_{(u)}$ is zero.

Table 2.21 presents the actual coverage proportions for the true survival function at each value of t in the 90% log-transformed confidence intervals based on the smoothed bootstrap method and the smoothed Kaplan-Meier bootstrap method. The SB method performs better in defining the actual coverage proportions for the true survival functions at all different sample sizes n. In the 90% log-transformed confidence intervals based on the SB method, over-coverage occurs at most values of t for $n = 6, 10, 20$, but there is under-coverage at most values of t for $n = 40, 100$. The log-transformed confidence intervals based on the SBa method provide good results for the true survival functions, contrary to the quantile and linear confidence intervals, in particular when the sample size is small. At small values of t, for example $t = 0.138, 0.223$, the alternative bootstrap method does not provide good coverage proportions when $n = 6, 10$. As the sample size increases, both smoothed bootstrap methods decrease the discrepancies between the actual coverage

$n =$		6		10		20		40		100	
t	$S(t)$	SB	SBa	SB	SBa	SB	SBa	SB	SBa	SB	SBa
0.138	0.95	0.919	0.836	0.891	0.863	0.913	0.899	0.924	0.909	0.921	0.917
0.223	0.90	0.922	0.854	0.910	0.877	0.906	0.876	0.925	0.915	0.910	0.907
0.298	0.85	0.936	0.867	0.931	0.877	0.912	0.889	0.914	0.897	0.907	0.900
0.368	0.80	0.941	0.868	0.931	0.876	0.918	0.894	0.918	0.906	0.891	0.884
0.436	0.75	0.945	0.858	0.930	0.876	0.914	0.873	0.909	0.898	0.904	0.897
0.503	0.70	0.941	0.860	0.937	0.869	0.907	0.874	0.904	0.887	0.872	0.869
0.570	0.65	0.948	0.850	0.932	0.869	0.905	0.866	0.899	0.887	0.883	0.876
0.639	0.60	0.952	0.859	0.931	0.863	0.909	0.875	0.891	0.873	0.885	0.881
0.710	0.55	0.968	0.880	0.933	0.861	0.895	0.872	0.884	0.868	0.893	0.888
0.783	0.50	0.973	0.880	0.933	0.868	0.904	0.867	0.881	0.867	0.887	0.883
0.861	0.45	0.972	0.889	0.943	0.855	0.903	0.868	0.882	0.860	0.882	0.879
0.943	0.40	0.969	0.908	0.933	0.860	0.902	0.870	0.892	0.870	0.878	0.871
1.033	0.35	0.973	0.937	0.933	0.870	0.894	0.865	0.890	0.878	0.891	0.879
1.132	0.30	0.985	0.988	0.934	0.882	0.900	0.859	0.899	0.857	0.892	0.881
1.243	0.25	0.982	0.991	0.940	0.921	0.896	0.844	0.876	0.854	0.887	0.867
1.373	0.20	0.974	0.987	0.943	0.970	0.899	0.839	0.868	0.836	0.886	0.871
1.533	0.15	0.972	0.983	0.933	0.965	0.901	0.900	0.879	0.834	0.890	0.870
1.744	0.10	0.954	0.973	0.931	0.962	0.914	0.964	0.888	0.879	0.884	0.858
2.078	0.05	0.936	0.969	0.917	0.950	0.906	0.951	0.895	0.960	0.876	0.847

Table 2.21: The actual coverage proportions for the true survival function at each value of t in the 90% log-transformed confidence intervals based on the two smoothed bootstrap methods.

proportions and the nominal size 0.90 for the true survival functions.

In this section, we introduced the SBa method to be as an alternative smoothed bootstrap method, and to compare its performance to the SB method through simulations. We found that the smoothed bootstrap method mostly performs better than the alternative bootstrap method in terms of the actual coverage proportions for the true survival functions at different values of t due to two reasons. The first one is that the SB method divides the support into $n + 1$ intervals while the alternative method divides the support into $u+1$ intervals, where $u \leq n$. Secondly, with the smoothed bootstrap method, we apply Exponential tail(s) for the last interval(s). In contrast to the alternative smoothed bootstrap method, we assume one Exponential tail for ($t_{(u)}$,∞) in case we have censored observation(s) after $t_{(u)}$. The last interval, ($t_{(u)}$,∞), will never be chosen during the alternative procedure if the largest observation is event because $\mathrm{P}_{(u)} = 0$, and this causes that the estimated survival function at any time $t \geq t_{(u)}$ is zero.

2.8 Concluding remarks

This chapter introduced a generalization of Banks' bootstrap method for right-censored data [1, 2] based on the rc-$A_{(n)}$ assumption [21]. This method can be used for inferences with high accuracy as illustrated through simulations. We compared it to Efron's method in terms of the coverage accuracy for the quartiles through simulations to investigate its performance.

Efron's bootstrap method for right-censored data restricts the process to sampling with replacement from the original data set. This method often leads to some issues such as right-censored observations and ties in bootstrap samples, especially when the sample size is small. The smoothed bootstrap method overcomes these issues by sampling from the whole data range using the rc-$A_{(n)}$ assumption.

Based on the simulations, the smoothed bootstrap method performs well with a low discrepancy between the nominal and actual coverage probabilities for the quartiles, in particular when the sample size is small. This is because the method has more variation in sampling, so that the percentile bootstrap confidence intervals are wider than the ones based on Efron's method. Efron's bootstrap method provides good accuracy when the sample size is large, but collecting a large data set is not always feasible in real applications, e.g. in early stage clinical trials, the sample size is mostly small, but important decisions must be made as explained in [38, 67, 75]. For such situations, it may be beneficial to use the smoothed bootstrap method rather than Efron's bootstrap method.

In running R codes, applying the smoothed bootstrap method takes approximately 15% longer than Efron's method. This may be due to computing the M-function values for each generated data set, sampling uniformly from the bounded intervals and getting observations from fitted Exponential tails to be accepted for the end interval(s).

When the support is unlimited, we assumed Exponential tail(s) for the final interval(s) in the smoothed bootstrap method, and this assumption affects the χ^2 value, in particular when the sample size is large. Therefore, we investigated another way to fit the tail based on the whole data. We used the Weibull distribution, but assuming the Weibull tail(s) for the end interval(s) has a small impact. In many cases, the smoothed bootstrap method with Weibull tail(s) for the end interval(s) led to about the same χ^2 values as the SB method with Exponential tail(s). When the sample size is small and the function of interest is located in or close to the tail region, it is better to assume Exponential tail(s) because they are able to cover the last quantiles, so that we are likely to obtain a smaller χ^2 value. The smoothed bootstrap method with Weibull tail(s) takes approximately 47% longer time than the SB procedure due to estimation of the Weibull shape parameter.

To investigate the smoothed bootstrap method in defining the actual coverage proportions for the true survival functions in the quantile, linear and log-transformed confidence intervals, it was compared to the smoothed Kaplan-Meier bootstrap method, SBa. The smoothed bootstrap method mostly performed better simulation results in all three pointwise confidence intervals especially when the sample size is small. Also, the smoothed bootstrap method is much better in the estimated coverage proportions for the true survival functions located in the tail region because we assumed an Exponential tail for each infinite interval. In applying the smoothed Kaplan-Meier bootstrap method, the last interval will not be selected unless we have right-censored observation(s) greater than the largest event time. Moreover, one Exponential tail would be assumed for the last interval if and only if we get right-censored observation(s) being greater than the largest event time.

Chapter 3

Smoothed bootstrap methods for bivariate data

3.1 Introduction

In 1986, Efron and Tibshirani [33] presented the bootstrap method for bivariate data. It is similar to the standard bootstrap method for univariate data, which creates the bootstrap samples by sampling with replacement from the original data set. It is simple and straightforward to measure the uncertainty of sample estimate. However, Efron's bootstrap provides poor results when the sample size is small. This motivates to introduce new smoothed bootstrap methods for bivariate data.

For univariate data, there are some smoothed bootstrap methods proposed based on kernels and linear interpolation, histospline smoothing, to obtain more accuracy in estimations; see Banks [5], De Angelis and Young [27], Hall [44], Silverman and Young [78], Young [87] for more details. Those methods have not been yet generalized for multivariate situation due to their complexity in application perhaps. The only method widely used in the literature for multivariate data is Efron's method, which uses the empirical distribution that assigns probability $\frac{1}{n}$ to each data point [33].

This chapter introduces three smoothed bootstrap methods for bivariate data. First, two bootstrap methods are based on the methodologies presented by Coolen-Maturi *et al.* [22] and Muhammad *et al.* [65]. They use both a parametric and a nonparametric copula model in combination with NPI to provide a partial distribution for one future bivariate observation. Based on each distribution, we present a smoothed bootstrap method. The two smoothed bootstrap methods rely on dividing the sample space into $(n+1)^2$ blocks based on the original data set, where n is the sample size. From the blocks, n blocks are sampled with replacement and with the assignment probabilities, which are computed based on the two predictive methods introduced by Coolen-Maturi *et al.* [22] and Muhammad *et al.* [65], then one bivariate observation is sampled uniformly from each selected block. A third smoothed bootstrap method is proposed based on assigning individual box kernels to the original data set. To study the proposed methods' performances, we compare them to Efron's bootstrap for bivariate data through simulations. We follow the similar procedure for comparison as in Chapter 2.

This chapter is organized as follows: Section 3.2 presents an overview of the combination of NPI with parametric and nonparametric copulas for bivariate data [64]. Section 3.3 introduces two smoothed bootstrap methods based on these methods. We present the third smoothed bootstrap method in Section 3.4, which is based on box kernels assigned to the original observations, and it will be referred to by the smoothed Efron's bootstrap. One example from the literature is presented in Section 3.5 to illustrate applications of the proposed smoothed bootstrap methods and Efron's method. Comparisons between the proposed smoothed bootstrap methods and Efron's bootstrap method for bivariate data are presented in Section 3.6. Section 3.7 presents concluding remarks of this chapter.

3.2 NPI combined with copulas

Copula is a well-known concept introduced by Sklar [79] to model dependence between random variables and construct multivariate distributions. This concept has been widely used in many applications because it allows to formalize the dependence between random variables separately from their marginal distributions. In the literature, many copula models have been presented and most of them are symmetric, see [15, 52, 68, 86] for more details. A copula model is symmetric if the variables

X and Y are exchangeable in a copula, $C(F_X(x), F_Y(y)) = C(F_Y(y), F_X(x))$, where $F_X(x)$ and $F_Y(y)$ are the cumulative distribution functions of X and Y, respectively, and $C(.,.)$ is the cumulative distribution function of the copula.

Muhammad [64] introduced two methods to provide a partially specified predictive distribution for one future bivariate observation. Those two methods are based on parametric and nonparametric copulas in combination with nonparametric predictive inference (NPI). Both methods consist of two parts. They apply NPI for the marginals in the first part, then a copula is assumed in the second part. For the semi-parametric predictive method, a parametric copula is assumed in the second part, but for the nonparametric predictive method, the bivariate nonparametric kernel-based copula is used in the second part.

To present the two predictive methods, we follow notation and definitions from Muhammad [64]. Suppose that there are n bivariate real-valued observations (x_i, y_i), for $i = 1, 2, \ldots, n$, corresponding to n exchangeable bivariate random quantities with no ties. The observations of the marginals are ordered and denoted by x_i and y_j for simplicity. Therefore, we have $x_1 < x_2 < \ldots < x_i < \ldots < x_n$ and $y_1 < y_2 < \ldots < y_j < \ldots < y_n$. The assumption $A_{(n)}$ is used for the marginals, so we have

$$P(X_{n+1} \in (x_{i-1}, x_i)) = \frac{1}{n+1} \quad \text{and} \quad P(Y_{n+1} \in (y_{j-1}, y_j)) = \frac{1}{n+1} \tag{3.1}$$

for $i, j = 1, 2, \ldots, n+1$, where $x_0 = a_x$, $y_0 = a_y$, $x_{n+1} = b_x$ and $y_{n+1} = b_y$ if the support is limited; and $x_0 = -\infty$, $y_0 = -\infty$, $x_{n+1} = +\infty$ and $y_{n+1} = +\infty$ if the sample space is $\mathbb{R}^2$.

To link this part to the second part of the proposed methods, where the dependence structure in the data is taken into account, Muhammad [64] introduced a natural transformation of the random variables individually by letting $\tilde{X}_{n+1}$ and $\tilde{Y}_{n+1}$ denote transformed versions of the random quantities X_{n+1} and Y_{n+1}, respectively, such that

$$\begin{gathered}(X_{n+1} \in (x_{i-1}, x_i), Y_{n+1} \in (y_{j-1}, y_j)) \iff \\ \left(\tilde{X}_{n+1} \in (\frac{i-1}{n+1}, \frac{i}{n+1}), \tilde{Y}_{n+1} \in (\frac{j-1}{n+1}, \frac{j}{n+1})\right)\end{gathered} \tag{3.2}$$

for $i, j = 1, 2, \ldots, n+1$.

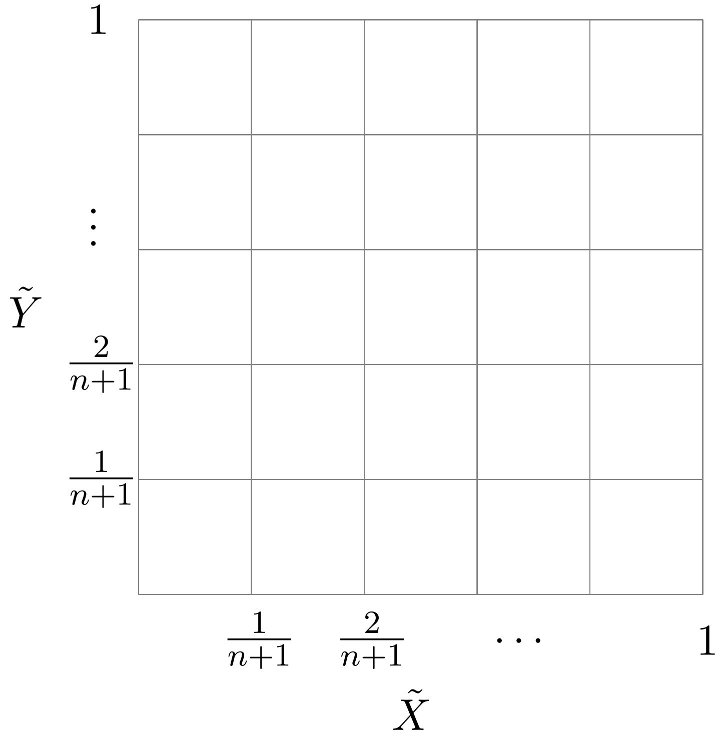

Figure 3.1: Presentation of the transformed space.

The $A_{(n)}$ assumptions for the marginals after the transformation are

$$P\left(\tilde{X}_{n+1} \in (\frac{i-1}{n+1}, \frac{i}{n+1})\right) = P\left(X_{n+1} \in (x_{i-1}, x_i)\right) = \frac{1}{n+1} \tag{3.3}$$

$$P\left(\tilde{Y}_{n+1} \in (\frac{j-1}{n+1}, \frac{j}{n+1})\right) = P\left(Y_{n+1} \in (y_{j-1}, y_j)\right) = \frac{1}{n+1} \tag{3.4}$$

This transformation leads us from the real space $\mathbb{R}^2$ to $[0,1]^2$, where $[0,1]^2$ is divided into $(n+1)^2$ equal-sized blocks by the n observed bivariate observations as shown in Figure 3.1. It should be noted that uniform marginal distributions have been discretized on $[0,1]^2$.

3.2.1 The semi-parametric predictive method

After using the NPI approach for the marginals in the first step, a parametric copula is assumed [64] and the copula parameter is estimated. It is possible to estimate the copula parameter by using the transformed data, where the observed pairs are replaced by $(\frac{r_i^x}{n+1}, \frac{r_i^y}{n+1})$, where r_i^x is the rank of the observation x_i among the x-observations, and r_i^y is the rank of the observation y_i among the y-observations.

It is obvious that the NPI approach assumed on the marginals is now combined with the estimated copula to provide a partially specified predictive distribution for one future bivariate observation. Each of the $(n+1)^2$ squares dividing the sample space is assigned a specific probability by the following formula

$$h_{ij}(\hat{\theta}) = P\left(\tilde{X}_{n+1} \in (\frac{i-1}{n+1}, \frac{i}{n+1}), \tilde{Y}_{n+1} \in (\frac{j-1}{n+1}, \frac{j}{n+1}) | \hat{\theta}\right) \tag{3.5}$$

for $i, j = 1, 2, \dots, n+1$, where $P(.|\hat{\theta})$ represents the assumed copula-based probability for the transformed data based on the original data and the assumed copula with estimated parameter $\hat{\theta}$. Each value of h_{ij} is between 0 and 1, and $\sum_{i=1}^{n} h_{ij} = \frac{1}{n+1}$ for all $j \in \{1, 2, \dots, n+1\}$ and $\sum_{j=1}^{n} h_{ij} = \frac{1}{n+1}$ for all $i \in \{1, 2, \dots, n+1\}$. Of course, $\sum_{i,j} h_{ij} = 1$.

To estimate the parameter of a copula model, many procedures have been presented in the literature, see [39, 40, 55] for more details. For the semi-parametric predictive method, we consider two estimation methods, which are widely used in the literature. The first one is pseudo maximum likelihood estimation (MLE), where the log pseudo likelihood function is [39]

$$\ell^*(\theta) = \sum_{i=1}^{n} \ln\left(c_\theta(\frac{r_i^x}{n+1}, \frac{r_i^y}{n+1})\right) \tag{3.6}$$

where $c_\theta(u, v) = \frac{\partial^2}{\partial u \partial v} C_\theta(u, v)$, and $C_\theta(u, v)$ is the cumulative distribution function of a parametric copula. The pseudo maximum likelihood estimator is the value $\hat{\theta}$ that maximizes ℓ^*.

The second estimation technique is the inversion of Kendall's tau (Itau), where the Kendall's tau formula and its population version in terms of the copula are [39]

$$\tau_n = \frac{4}{n(n-1)} P_n - 1 \quad \text{and} \quad \tau(C_\theta) = -1 + 4 \int_0^1 \int_0^1 C_\theta(u, v)\, dC_\theta(u, v) \tag{3.7}$$

where P_n is the number of concordant pairs in the sample, and τ_n is the sample Kendall's correlation. The Itau estimator is the value $\hat{\theta}$ resulting from solving the equation $\tau_n = \tau(C_\theta)$. The two estimation methods can be easily implemented by the R package `VineCopula` [74].

The studies by Genest *et al.* [40] and Kojadinovic and Yan [55] suggest that the pseudo maximum likelihood estimation is the best choice in terms of mean square error to estimate the copula parameter when the sample size is greater than 100 or $\tau_n \geq 0.4$, where τ_n is the Kendall's correlation. Otherwise, the inversion of Kendall's tau procedure provides more precise estimates.

3.2.2 The nonparametric predictive method

For this method, we first use the NPI approach for the marginals, and then do the transformation as described before. This is the first step of this nonparametric predictive method, which is as the first step in Subsection 3.2.1. In the second step for the nonparametric predictive method, a kernel smoothing copula is assumed and an estimated probability density function $\hat{c}$ can be defined as [64]:

$$\hat{c}(x,y) = \frac{1}{nb_X b_Y} \sum_{i=1}^{n} K\left(\frac{x - F_X(\tilde{X}_i)}{b_X}, \frac{y - F_Y(\tilde{Y}_i)}{b_Y}\right) \tag{3.8}$$

where $K : \mathbb{R}^2 \to \mathbb{R}$ is a bivariate kernel function, $b_X, b_Y > 0$ are the bandwidths or the smoothing parameter, $F_X(\tilde{X}_i) = \frac{r_x^i}{n+1}$ and $F_Y(\tilde{Y}_i) = \frac{r_y^i}{n+1}$.

Now, the NPI approach, which is assumed for the marginals, can easily be combined with this nonparametric kernel-based copula to take the dependence into account. $K(.,.)$ in Equation (3.8) can be any kernel function. Some kernel functions, e.g. Gaussian, Epanechnikov or Uniform, are available in the R package `np` [46]. Now we can find the h_{ij} values assigned to the $(n+1)^2$ blocks by the following equation

$$h_{ij}(\hat{c}) = P\left(\tilde{X}_{n+1} \in (\frac{i-1}{n+1}, \frac{i}{n+1}), \tilde{Y}_{n+1} \in (\frac{j-1}{n+1}, \frac{j}{n+1}) | \hat{c}\right) \tag{3.9}$$

where $i, j = 1, 2, ..., n+1$ and $P(.|\hat{c})$ represents the nonparametric kernel-based copula probability with estimated kernel density function, $\hat{c}$. Note that the h_{ij} values must satisfy the three conditions presented after Equation (3.5) in Subsection 3.2.1.

The choice of bandwidths b_X and b_Y is crucial for implementation. In the univariate data situation, the normal reference rule-of-thumb has been widely used in the literature. It works well if the underlying distribution is close to a normal distribution [77]. This bandwidth [77] is given by

$$b_Z = 1.06\ \hat{\sigma}_Z\ n^{-\frac{1}{5}} \tag{3.10}$$

where $\hat{\sigma}_Z$ is the standard deviation of the variable Z, and n is the sample size.

Silverman [77] stated that Equation (3.10) can be improved for long-tailed, skew and bi-modal distributions by

$$b_Z = 0.90\ \hat{A}_Z\ n^{-\frac{1}{5}} \tag{3.11}$$

where $\hat{A}_Z = \min(\hat{\sigma}_Z, \frac{IQR(Z)}{1.349})$. $IQR(Z)$ is the interquartile of Z, and the 1.349 value is the interquartile of the standard normal distribution.

Li and Racine [59] generalized the normal reference rule-of-thumb for multivariate data as follows

$$b_i = 1.06\ \hat{\sigma}_i\ n^{-\frac{1}{4+q}} \tag{3.12}$$

where $i = 1, 2, \ldots, q$, q is the number of variables, $\hat{\sigma}_i$ is the estimated standard deviation of i^{th} variable, and n is the sample size.

Henderson and Parmeter [48] derived a general formula of the normal reference rule-of thumb bandwidths for q variables, v^{th}-order kernel density and r^{th} derivative estimator by using an approximation form of the mean integrated squared error as follows

$$b_i = CS\ \hat{\sigma}_i\ n^{-\frac{1}{2v+q}} \tag{3.13}$$

where $CS = (\frac{4}{2+q})^{(\frac{1}{4+q})}$, $i = 1, 2, \ldots, q$, q is the number of variables, v is the v^{th}-order kernel density, which is defined as the order of the first non-zero moment. $\hat{\sigma}_i$ is the standard deviation of i^{th} variable, and n is the sample size.

For the nonparametric predictive method, Muhammad [64] used the following equation

$$b_Z = 1.06\ \hat{A}_Z\ n^{-\frac{1}{4}} \tag{3.14}$$

where $\hat{A}_Z = \min\left(\hat{\sigma}_Z, \frac{IQR(Z)}{1.349}\right)$, and n is the sample size.

Muhammad [64] used Equation (3.14) as the normal reference rule-of-thumb bandwidths, but for our use, we replace $n^{-\frac{1}{4}}$ by n^{-1} to have smaller bandwidths, and this change leads to better results in terms of the coverage probability.

3.3 Smoothed bootstrap based on NPI with copulas

In this section, we present the smoothed bootstrap methods for bivariate data by using the semi-parametric and nonparametric predictive methods. These smoothed bootstrap methods are proposed to learn about uncertainty in estimation of a population characteristic θ and to avoid the issue of ties, which often occurs with Efron's bootstrap method for bivariate data. We refer to the smoothed bootstrap methods by SBSP and SBNP, respectively. Suppose that (X_1, Y_1), $(X_2, Y_2), \ldots, (X_n, Y_n)$ are independent and identically distributed random variables, where $X \sim F$, $Y \sim G$ and F and G represent unknown probability distributions defined on a sample space $\aleph \times \xi$. Let (x_1, y_1), (x_2, y_2), $\ldots$, (x_n, y_n) denote the observed data and let $\theta(F)$ be the functional of interest. Based on the n data points, the predictive methods construct $(n+1)^2$ squares dividing the sample space, and their probabilities h_{ij} are specified via either Equation (3.5) or (3.9) [64]. For clarity, we describe the smoothed bootstrap algorithms for bivariate data in the following steps, the R codes are provided in Appendixes B.2 and B.3:

(i) Apply one of the introduced predictive methods to the observed data to construct $(n+1)^2$ squares with their estimated probabilities $\hat{h}_{ij}$.

(ii) Sample with replacement n squares with the assignment probabilities, then sample one bivariate observation uniformly from each selected square to obtain a smoothed bootstrap sample of size n.

(iii) Calculate the function of interest, $\hat{\theta}^*$.

(iv) Perform steps (ii) and (iii) B times in order to have B resulting values of the function of interest.

When the proposed bootstrap methods are applied for the case of finite support, all blocks partitioning the sample space are finite, so that it is easy to sample one bivariate observation uniformly from those blocks. However in an infinite support situation, we will have blocks unlimited from either one or two sides, so it is impossible to draw one observation uniformly. To resolve this complicated issue, we will use the idea introduced by BinHimd [10] for the end intervals when the data is univariate, but here it will be generalized for bivariate data situation. If a block that is not limited from the left or right side, which the block's x-axis is $(-\infty, x_1)$ or $(x_n, +\infty)$, is sampled during the smoothed bootstrap, we will use normal distribution tails, where the normal distribution's parameters μ and σ can be estimated by Equations (1.3).

Then we get a value less than x_1 from the left tail for $(-\infty, x_1)$, and greater than x_n from the right tail for $(x_n, +\infty)$. The x observation sampled from either tail of the normal distribution will be considered as the x value of the bivariate future observation. In case a block is not limited from the lower or upper bound, which means the block's y-axis is $(-\infty, y_1)$ or $(y_n, +\infty)$, we apply the same assumption, but in terms of Y, then the y observation sampled from either tail of the normal distribution will be consider as the y value of the bivariate future observation.

3.4 Smoothed Efron's bootstrap

For Efron's bootstrap method, the empirical distribution is used, so the bootstrap method samples are created by sampling observations with replacement from the original data set. In this section, we want to relax this assumption and introduce an alternative smoothed bootstrap method for bivariate data. This can be done by creating a block centered each observation. After creating blocks around the observed data points, the alternative bootstrap method samples blocks with replacement, then one bivariate observation is sampled uniformly from each selected block. For each point in the original data set, we assign a block with size $b_X \times b_Y$, where the observed point is located in the center of its corresponding block. Those blocks could be overlapping due to their size and the distances among their centers. This alternative bootstrap method provides some smoothness to avoid the issue of ties, and it can be considered as applying a uniform kernel, where each point inside any individual kernel has the same probability regardless of the distance from

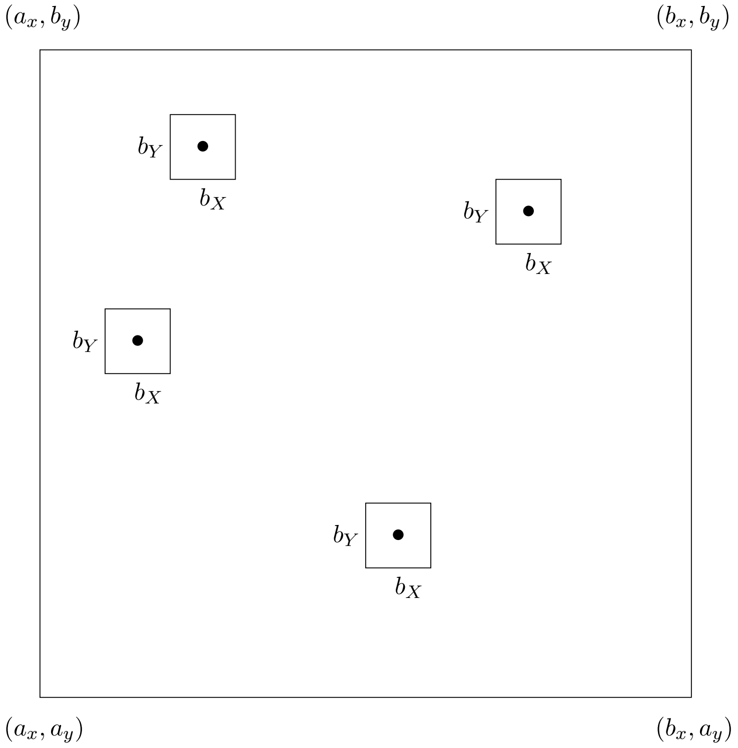

Figure 3.2: The kernels assigned to the observed values.

the center. Suppose that there are n bivariate real-valued observations (x_i, y_i), for $i = 1, 2, \ldots, n$, corresponding to n bivariate random quantities defined on a data space $(a_x, b_x) \times (a_y, b_y)$. Thus, n equal-sized blocks are made and assigned to the observations, as shown in Figure 3.2 for $n = 4$, where b_X and b_Y are computed by Equation (3.14). This smoothed version of Efron's bootstrap, SEB, can be described in the following steps, the R codes are provided in Appendix B.4:

(i) Create the blocks around the observed points.

(ii) Sample n blocks with replacement, where each block is set with probability $\frac{1}{n}$.

(iii) From each chosen block, sample one bivariate observation uniformly.

(iv) Calculate the function of interest, $\hat{\theta}^*$.

(v) Perform steps (ii), (iii) and (iv) B times in order to have B bootstrap estimates of the function θ^*.

In the following section, one example is presented using a data set from the literature to illustrate the proposed smoothed bootstrap methods along with Efron's bootstrap method for bivariate data.

3.5 Body Mass Index example

A data set for 30 eleven-year-old girls attending Heaton Middle School in Bradford is collected [64] and it is presented in Table 3.1 and Figure 3.3. This data set presents the heights (m) and weights (kg) of these girls. The observed Pearson, Kendall and Spearman correlations for these 30 observations are $0.742, 0.631, 0.807$, respectively. Suppose that one is interested in estimating the average of the body mass index $\overline{\mathrm{BMI}}$ and the standard error of $\overline{\mathrm{BMI}}$ along with the 90% quantile confidence interval for $\overline{\mathrm{BMI}}$. In this section, we illustrate how to use the proposed smoothed bootstrap methods and Efron's method for bivariate data to estimate $\overline{\mathrm{BMI}}$, SE($\overline{\mathrm{BMI}}$) and the 90% quantile confidence interval. We expect that the methods will provide nearly identical results because the sample size is not small and it seems there is a roughly linear relationship between the variables as shown in Figure 3.3. The sample mean of the body mass index $\overline{\mathrm{BMI}}$ is equal to 17.11, where BMI is calculated by

$$\mathrm{BMI} = \frac{\text{Weight (kg)}}{[\text{Height (m)}]^2} \tag{3.15}$$

We create 1000 bootstrap samples of size $n = 30$ based on each bootstrap method, and derive the mean of BMI of each bootstrap sample. This leads to 1000 resulting values of $\overline{\mathrm{BMI}}$. The bootstrap estimate for $\overline{\mathrm{BMI}}$ can be computed by taking the average of the 1000 resulting values. To estimate the standard error of $\overline{\mathrm{BMI}}$, we use Equation (1.2). For the 90% quantile confidence interval, we take the 50th and 950th ordered values of the 1000 resulting values. The estimated results based on each bootstrap method are presented in Table 3.2.

Table 3.2 shows the estimated results for $\overline{\mathrm{BMI}}$, SE($\overline{\mathrm{BMI}}$) and the 90% confidence interval for $\overline{\mathrm{BMI}}$ based on the proposed smoothed bootstrap methods and Efron's method for bivariate data. All bootstrap methods provide nearly identical results for the mean of BMI. However, the smoothed bootstrap method based on the nonparametric predictive method, SBNP, has larger estimate for the standard error of

Height (m)	Weight (kg)	BMI	Height (m)	Weight (kg)	BMI
1.35	26	14.27	1.33	31	17.53
1.46	33	15.48	1.49	34	15.31
1.53	55	23.50	1.41	32	16.10
1.54	50	21.08	1.64	47	17.47
1.39	32	16.56	1.46	37	17.36
1.31	25	14.57	1.49	46	20.72
1.49	44	19.82	1.47	36	16.66
1.37	31	16.52	1.52	47	20.34
1.43	36	17.60	1.40	33	16.84
1.46	35	16.42	1.43	42	20.54
1.41	28	14.08	1.48	32	14.61
1.36	28	15.14	1.49	32	14.41
1.54	36	15.18	1.41	29	14.59
1.51	48	21.05	1.37	34	18.11
1.55	36	14.98	1.35	30	16.46

Table 3.1: The heights (m), weights (kg) and BMI of 30 eleven-year-old girls.

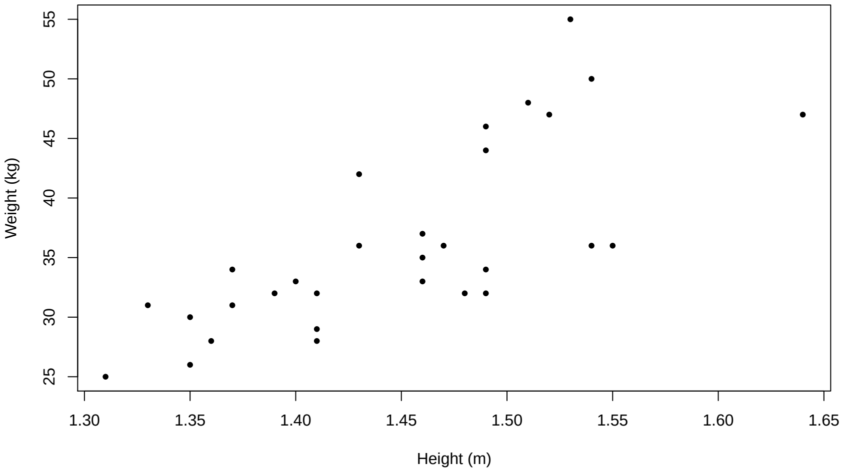

Figure 3.3: Heights (m) and corresponding weights (kg) values of 30 eleven-year-old girls.

Method	$\overline{\text{BMI}}$	SE($\overline{\text{BMI}}$)	90% confidence interval
SBSP	17.10	0.479	(16.35, 17.87)
SBNP	17.23	0.591	(16.31, 18.25)
Efron	17.11	0.455	(16.41, 17.90)
SEB	17.08	0.446	(16.36, 17.83)

Table 3.2: The bootstrap estimates for $\overline{\text{BMI}}$, SE($\overline{\text{BMI}}$) and the 90% confidence interval for $\overline{\text{BMI}}$ based on each bootstrap method.

$\overline{\text{BMI}}$ and wider 90% confidence interval. This could be because the method provides more variation in distributing the probabilities h_{ij} over the $(n+1)^2$ blocks. In contrast, Efron's method and the SEB method provide smaller estimates for SE($\overline{\text{BMI}}$) and shorter confidence intervals. These results are expected because we have less variation in sampling when applying these bootstrap methods in comparison to the SBSP and SBNP methods. Overall, the methods provided about the same estimates, as we had expected, due to the large sample size and there is a roughly linear relationship between the variables.

3.6 Comparison with Efron's method

To compare the smoothed bootstrap methods, proposed in Sections 3.3 and 3.4, and Efron's bootstrap method for bivariate data, we use the same comparison technique as in Chapter 2. The null hypothesis proposes that the coverage probabilities are equally distributed over the 10 confidence regions; each confidence region or cell is with nominal size 0.10. The test is set as follows

$$\begin{aligned} &H_0 : \text{The coverage probabilities are equally distributed over the 10 cells.} \\ &H_1 : \text{At least one coverage probability is not equal.} \end{aligned} \tag{3.16}$$

We use the chi-squared goodness of fit test with significance level 0.05, the 95th percentile of the chi-squared distribution with 9 degrees of freedom is equal to 16.92. If we obtain a chi-squared value less than 16.92, we fail to reject the null hypothesis and conclude that the actual coverage probabilities could be equally distributed over the 10 confidence regions. We use the over (under) lines for larger chi-squared values to indicate the over (under) coverage proportions in the first confidence regions.

In this section we consider two different scenarios to compare the bootstrap methods for bivariate data. Each scenario is created from a specific copula family, where the variables are both following the Uniform distribution with $a_x = a_y = 0$ and $b_x = b_y = 1$. For the first scenario, we generate data sets from the Normal copula model with different levels of dependence modelled by Kendall's tau, τ, and different sample sizes. We expect that the smoothed bootstrap based on the semi-parametric predictive method, SBSP, will provide better results than those of the other methods in this scenario because we generate data sets from the same copula family we assume for the analysis.

For the second scenario, data sets are generated from Gumbel copula model with different values of τ and n as well. In this scenario, we assume Normal copula when we apply the semi-parametric predictive method. It will be a nice comparison to show the performance of SBSP method when the model used to generate the data sets is different from the model we assume for the analysis. For this smoothed bootstrap method, we will use the pseudo MLE method [39] when $\tau \geq 0.4$ to estimate the copula parameter θ; otherwise we use the inversion of Kendall's tau. For the smoothed bootstrap method based on the nonparametric predictive technique, we use the Normal kernel as a kernel type and the bandwidths are computed by Equation (3.14), but we replace $n^{-\frac{1}{4}}$ by n^{-1} for better results. This is implemented in R by the package `np` [46].

3.6.1 First scenario: Normal copula model

To compare the proposed bootstrap methods for bivariate data with a finite support, we generate $N = 1000$ data sets from the Normal copula family with uniform margins. Then for each generated data set, we apply each bootstrap method $B = 1000$ times. We compute the estimate for the function of interest based on each bootstrap sample, then we define the 10 confidence regions by Equation (2.6). We then compute the proportions of confidence regions which include the true function of interest. We repeat this procedure for all $N = 1000$ generated data sets in order to observe the actual coverage proportions for the true function of interest in the 10 confidence regions. Based on the results, we compute the corresponding chi-squared value for each bootstrap method. We consider sample sizes $n = 10, 50, 100$, and different functions of interest. We investigate the actual coverage proportions for the Pearson (r), Kendall (τ) and Spearman (rs) correlations. For the two variables X and Y, we investigate the actual coverage proportions for the mean of $T_1 = X + Y$, denoted by $\overline{T_1}$ and the mean of $T_2 = XY^2$, denoted by $\overline{T_2}$. Tables 3.3 to 3.7 outline the chi-squared values obtained from the coverage probabilities for r, τ, rs, $\overline{T_1}$ and $\overline{T_2}$, respectively, based on each bootstrap method.

For the Pearson correlation r, Table 3.3 shows the chi-squared values conducted from the coverage proportions based on the bootstrap methods. It is important to recall that the over (under) lines are used for larger chi-squared values to indicate the over (under) coverage proportions in the first confidence regions. When $n = 10$, the SBSP, SBNP and SEB methods provide under-coverage proportions when $\tau = -0.25, 0$ and 0.25, and over-coverage proportions if $\tau = -0.75$ and 0.75. However, Efron's bootstrap method always provides under-coverage proportions regardless of the dependence level. At this sample size, the SBNP method provides the smallest chi-squared value at most values of τ. For large sample sizes, all methods perform well, but the SEB method is the best. The results are shown in Figure 3.4. From this plot, where it is easy to observe that the best method not rejecting H_0 is the SBNP method, followed by the SEB method. These two methods lead to not reject H_0 in 16 and 15 cases, respectively, out of 21 cases.

Table 3.4 presents the chi-squared values for Kendall's correlation τ. This function of interest relies on the concordance of data, and the proposed bootstrap methods affect the rank not only through the probabilities assigned for blocks, but also because we sample uniformly from the chosen blocks during the bootstrap procedures. This can be seen when the dependence level is high, in particular when we use the SEB method. At large sample sizes, the SBSP and SBNP methods provide small chi-squared values in most dependence level cases. The SBSP method performs well because we assume the Normal copula in the semi-parametric method and the same parametric family is used to generate the data sets. These results are expected and rational as we mentioned before in the introduction of this section.

$n =$			10				50				100			
τ	θ	r	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB
-0.75	-0.9239	-0.92	$\overline{29.98}$	$\overline{54.34}$	$\underline{91.22}$	$\overline{79.36}$	12.32	7.32	14.96	13.06	16.82	12.08	9.82	4.16
-0.50	-0.7071	-0.69	15.04	11.06	$\underline{70.16}$	14.58	9.28	15.70	7.66	6.18	7.32	7.34	6.62	10.32
-0.25	-0.3827	-0.37	$\underline{27.92}$	$\underline{24.16}$	$\underline{48.90}$	$\underline{30.98}$	7.46	12.02	9.48	4.40	$\overline{20.20}$	4.82	10.24	3.46
0	0	0	$\underline{24.52}$	$\underline{35.54}$	$\underline{56.84}$	$\underline{34.62}$	8.18	6.50	14.48	6.86	11.72	11.44	11.92	7.82
0.25	0.3827	0.37	$\underline{43.18}$	$\underline{20.44}$	$\underline{37.86}$	$\underline{24.90}$	8.96	8.04	9.94	5.96	$\underline{17.56}$	10.62	$\underline{26.24}$	15.98
0.50	0.7071	0.69	8.00	4.82	$\underline{80.84}$	$\underline{24.94}$	7.28	9.64	13.54	5.26	$\overline{21.22}$	11.04	4.78	11.14
0.75	0.9239	0.92	$\overline{126.20}$	$\overline{65.24}$	$\underline{116.46}$	$\overline{64.38}$	$\overline{17.32}$	7.72	$\underline{25.62}$	14.86	$\overline{24.88}$	11.92	5.76	11.64

Table 3.3: The chi-squared values obtained from the coverage proportions for the true Pearson correlation r.

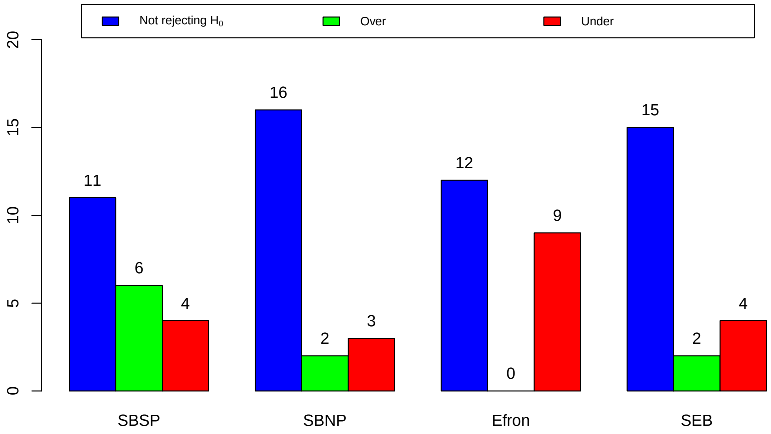

Figure 3.4: The number of times that H_0 is not rejected, and the number of times that H_0 is rejected due to the over- and under- coverage proportions for Pearson correlation r based on each bootstrap.

$n =$		10				50				100			
τ	θ	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB
-0.75	-0.9239	$\overline{261.68}$	$\overline{65.06}$	$\overline{70.50}$	$\overline{63.92}$	5.96	14.08	$\overline{19.08}$	$\underline{69.00}$	6.60	13.50	16.26	$\underline{86.06}$
-0.50	-0.7071	$\underline{53.06}$	$\overline{52.00}$	$\overline{18.70}$	$\overline{104.02}$	8.14	7.56	3.60	9.20	3.10	10.26	10.08	6.66
-0.25	-0.3827	8.78	$\overline{21.00}$	12.70	$\overline{32.22}$	7.66	5.08	5.76	3.98	14.90	$\overline{17.18}$	$\overline{18.30}$	11.60
0	0	$\underline{17.98}$	14.54	$\overline{33.54}$	$\overline{21.74}$	8.34	3.78	12.28	6.14	8.58	15.82	9.62	14.90
0.25	0.3827	$\underline{18.60}$	16.38	6.86	$\overline{28.68}$	11.12	8.86	9.58	9.92	$\underline{19.92}$	13.98	16.72	13.18
0.50	0.7071	5.44	$\overline{71.66}$	$\overline{24.74}$	$\overline{103.40}$	16.94	6.94	8.40	8.98	11.28	8.32	10.44	10.12
0.75	0.9239	$\overline{259.48}$	$\overline{92.50}$	$\overline{149.00}$	$\overline{73.92}$	$\overline{23.72}$	$\overline{24.94}$	11.38	$\underline{20.02}$	$\overline{19.76}$	10.20	12.18	$\underline{46.66}$

Table 3.4: The chi-squared values obtained from the coverage proportions for the true Kendall correlation τ.

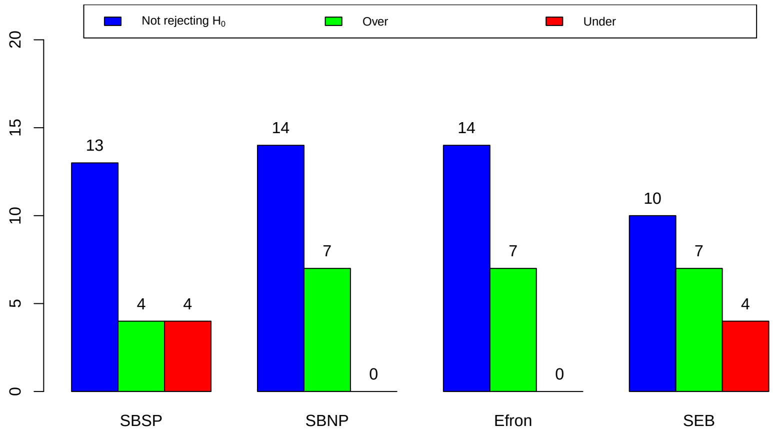

Figure 3.5: The number of times that H_0 is not rejected, and the number of times that H_0 is rejected due to the over- and under- coverage proportions for Kendall correlation τ based on each bootstrap.

From Figure 3.5, it is obvious that the SBNP method and Efron's bootstrap method are the best methods providing small chi-squared values, where both methods lead to not reject H_0 14 times, with H_0 rejected 7 times due to over coverage proportions occurring in the first confidence regions.

In Table 3.5, we list the chi-squared values obtained from the coverage proportions for Spearman's correlation rs. For this kind of correlation, the difference between the ranks of the x and y observations is the base to compute the correlation between variables, and this can be influenced by the blocks' probabilities and the uniform sampling from blocks in the smoothed bootstrap methods. When $\tau = -0.75$ and 0.75, the SEB method is always the worst method amongst all applied methods due to having the largest chi-squared values, but it provides good results if τ is small. The SBSP and SBNP methods along with Efron's method mostly provide small chi-squared values, but Efron's method is better, in particular when the sample size is large. To have a better view of the methods' performances, we visualize the results in Figure 3.6. From this figure, where it is clear that the best method not rejecting H_0 is Efron's method, as out of 21 different cases, it leads to not reject H_0 17 times, then the second best method is SBNP; it leads to not reject H_0 16 times.

Tables 3.6 and 3.7 present the chi-squared values obtained from the coverage proportions for the means of T_1 and T_2, respectively. All smoothed bootstrap methods distribute the coverage proportions for both functions $\overline{T_1}$ and $\overline{T_2}$ over the 10 cells better than Efron's method does at most dependence levels when $n = 10$. At this sample size, Efron's method provides under-coverage proportions for all values of τ. When the sample size increases to 50 and 100, Efron's method accomplishes good coverage proportions for both functions, regardless of the value of τ. This mostly leads to not reject H_0, but the smoothed bootstrap methods perform better at these sample sizes. Figures 3.7 and 3.8 present the number of times that H_0 is not rejected, and the number of times that H_0 is rejected due to the over- and under-coverage proportions. It is clear that the SBSP and SBNP methods are the best in the sense of providing small chi-squared values for $\overline{T_1}$; they both lead to not reject H_0 17 times in the 21 settings. For $\overline{T_2}$, the SBSP and SBNP methods accomplish

$n =$			10				50				100			
τ	θ	rs	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB
-0.75	-0.9239	-0.92	$\overline{125.18}$	$\underline{140.32}$	$\overline{137.38}$	$\underline{410.24}$	4.46	$\underline{22.26}$	10.02	$\underline{142.30}$	14.38	8.38	7.86	$\underline{94.86}$
-0.50	-0.7071	-0.69	$\underline{39.38}$	$\overline{34.02}$	$\overline{38.24}$	$\overline{69.18}$	7.12	7.72	6.70	9.40	11.66	8.38	9.18	9.80
-0.25	-0.3827	-0.37	10.82	10.52	5.80	$\overline{27.80}$	6.84	6.88	8.12	6.34	15.06	8.18	10.66	7.00
0	0	0	$\underline{17.58}$	15.46	14.72	$\overline{18.64}$	11.74	3.12	7.34	11.62	7.48	6.98	6.98	15.96
0.25	0.3827	0.37	$\underline{20.04}$	8.64	6.26	$\overline{19.84}$	5.50	15.04	15.92	$\overline{20.32}$	13.70	12.38	13.36	12.00
0.50	0.7071	0.69	$\overline{35.92}$	$\overline{51.54}$	$\overline{39.00}$	$\overline{70.34}$	13.40	4.18	9.34	14.08	16.90	12.48	3.58	9.70
0.75	0.9239	0.92	$\overline{89.14}$	$\underline{116.40}$	$\overline{147.30}$	$\underline{445.88}$	$\overline{30.48}$	11.70	3.60	$\underline{82.30}$	$\overline{33.88}$	12.52	4.58	$\underline{63.20}$

Table 3.5: The chi-squared values obtained from the coverage proportions for the true Spearman correlation rs.

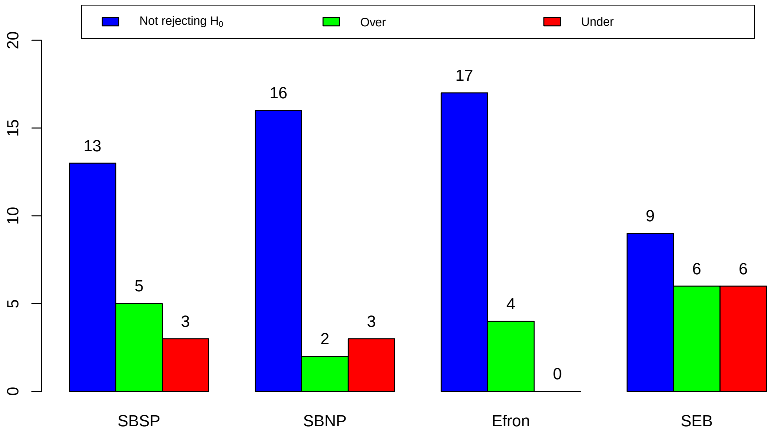

Figure 3.6: The number of times that H_0 is not rejected, and the number of times that H_0 is rejected due to the over- and under- coverage proportions for Spearman correlation rs based on each bootstrap.

$n =$			10				50				100			
τ	θ	$\overline{T_1}$	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB
-0.75	-0.9239	1	$\overline{85.86}$	$\overline{131.72}$	$\underline{47.60}$	8.16	9.08	9.68	11.30	4.76	12.40	15.44	13.92	7.64
-0.50	-0.7071	1	$\overline{24.44}$	$\overline{23.84}$	$\underline{47.12}$	$\underline{17.10}$	8.54	5.88	10.14	5.26	10.80	12.48	$\underline{18.02}$	11.46
-0.25	-0.3827	1	13.90	$\overline{18.16}$	$\underline{39.96}$	$\underline{28.14}$	7.32	7.04	12.88	7.54	7.94	16.00	$\underline{28.02}$	$\underline{20.72}$
0	0	1	15.76	11.24	$\underline{40.04}$	$\underline{28.26}$	10.92	4.08	10.82	11.20	12.46	$\underline{17.32}$	$\underline{19.04}$	16.56
0.25	0.3827	1	8.90	8.56	$\underline{41.94}$	$\underline{33.04}$	$\underline{20.10}$	7.10	$\underline{17.98}$	15.20	14.22	4.60	10.44	$\underline{17.68}$
0.50	0.7071	1	$\overline{18.10}$	10.76	$\underline{37.50}$	$\underline{29.56}$	10.12	5.38	6.90	9.72	7.42	12.28	9.80	10.42
0.75	0.9239	1	9.72	9.74	$\underline{36.48}$	$\underline{31.62}$	13.28	6.60	7.08	6.58	13.38	9.12	13.48	13.00

Table 3.6: The chi-squared values obtained from the coverage proportions for the true $\overline{T_1}$.

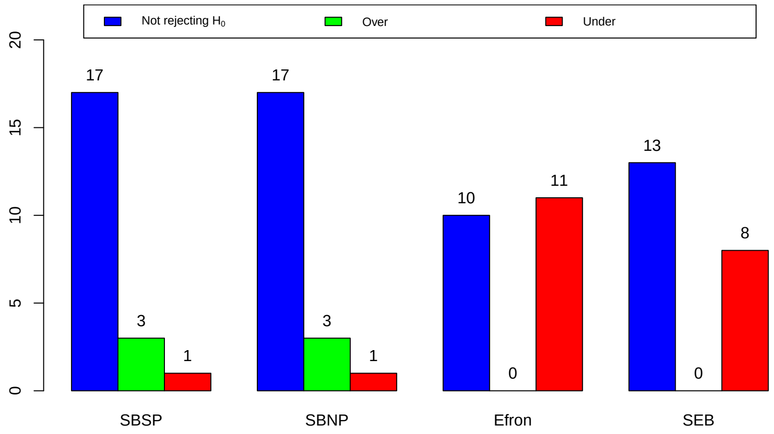

Figure 3.7: The number of times that H_0 is not rejected, and the number of times that H_0 is rejected due to the over- and under- coverage proportions for $\overline{T_1}$ based on each bootstrap.

$n =$			10				50				100			
τ	θ	$\overline{T_2}$	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB
-0.75	-0.9239	0.090	$\overline{49.10}$	$\overline{49.72}$	$\underline{46.28}$	5.42	$\overline{26.02}$	8.08	$\overline{17.86}$	14.64	7.70	8.66	9.60	3.52
-0.50	-0.7071	0.109	13.28	8.52	$\underline{98.94}$	$\underline{44.70}$	7.12	11.88	10.44	9.04	6.12	8.56	7.70	3.66
-0.25	-0.3827	0.136	$\overline{18.46}$	$\underline{17.26}$	$\underline{91.22}$	$\underline{75.34}$	12.46	8.64	10.64	$\underline{19.38}$	4.26	15.18	16.08	6.14
0	0	0.167	10.04	$\underline{17.02}$	$\underline{110.96}$	$\underline{87.78}$	1.80	4.94	7.56	12.14	4.54	15.62	$\underline{21.80}$	13.76
0.25	0.3827	0.197	8.12	$\overline{19.06}$	$\underline{87.32}$	$\underline{73.42}$	12.22	10.30	5.14	14.94	11.62	$\underline{21.78}$	$\underline{17.76}$	15.14
0.50	0.7071	0.224	$\overline{19.04}$	11.40	$\underline{71.12}$	$\underline{61.26}$	14.68	5.94	8.80	11.34	2.90	$\underline{17.94}$	4.40	16.56
0.75	0.9239	0.243	15.30	15.74	$\underline{42.60}$	$\underline{36.12}$	16.54	9.76	8.98	$\underline{18.94}$	6.04	9.50	5.92	12.28

Table 3.7: The chi-squared values obtained from the coverage proportions for the true $\overline{T_2}$.

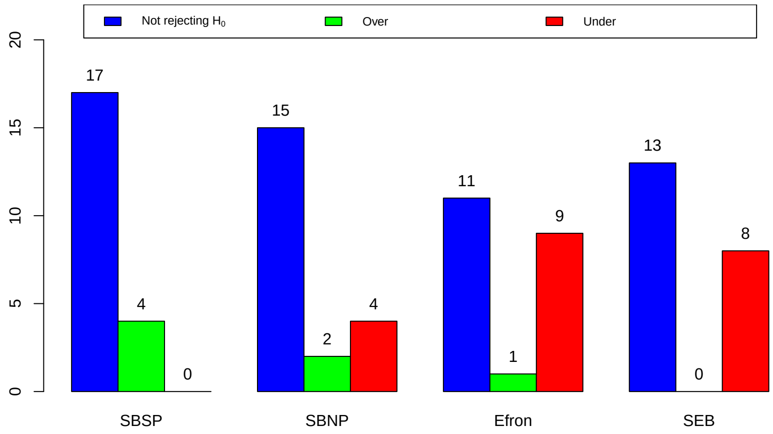

Figure 3.8: The number of times that H_0 is not rejected, and the number of times that H_0 is rejected due to the over- and under- coverage proportions for $\overline{T_2}$ based on each bootstrap.

Function	SBSP	SBNP	Efron	SEB
r	2 (0,1,1)	8 (6,0,2)	3 (3,0,0)	8 (7,1,0)
τ	6 (6,0,0)	6 (6,0,0)	4 (3,1,0)	5 (3,2,0)
rs	5 (2,3,0)	5 (4,1,0)	9 (9,0,0)	3 (3,0,0)
$\overline{T_1}$	6 (6,0,0)	9 (9,0,0)	0 (0,0,0)	6 (5,0,1)
$\overline{T_2}$	9 (9,0,0)	6 (5,0,1)	3 (3,0,0)	3 (3,0,0)

Table 3.8: The summary of scores for the bootstrap methods in the five functions of interest along with the number of chi-squared values of each score that are based on equal, over- and under- coverage proportions, (#Equal, #Over, #Under).

good accuracy in most cases, but the SBSP method is better. The latest one does not reject H_0 in 17 cases while the SBNP method does not reject H_0 in 15 cases only.

To compare the methods in each setting, we give 1 for the method providing the lowest chi-squared value among all values, regardless whether that value is based on equal, over or under coverage probabilities, and we give 0 to the other methods. For example in Table 3.3, the corresponding chi-squared values to the SBSP, SBNP and SEB methods are $\overline{29.98}, \overline{54.34}$ and $\overline{79.36}$, respectively, while Efron's value is $\underline{91.22}$ when $n = 10$ and $\tau = -0.75$. In this case, we count 1 for the SBSP method because its chi-squared value is the lowest value among all. We count the times that each method gives lowest chi-squared values in comparison to those of the other methods on each function of interest, and the summary of Tables 3.3 to 3.7 is presented in Table 3.8. As indicated in this table, each score is assigned with three numbers written within a parenthesis. These numbers indicate how many chi-squared values of that score are based on equal, over- and under- coverage proportions, (#Equal, #Over, #Under), respectively. For the Pearson's correlation r, the SBNP and SEB methods get the best scores; they both score 8, but the SEB method is better than the SBNP method because it provides small chi-squared values 7 times while the SBNP method provides only 6. For the Kendall's correlation τ, the SBSP and SBNP methods provide the best score, 6, and the SEB method is the second best with score 5. For the Spearman's correlation rs, Efron's method has the highest score, 9, and all of them are based on equal coverage proportions. The SBNP method is the second best with score 5, and 4 out of 5 are not rejecting H_0. For the mean of T_1, the SBNP method performs best, followed by the SBSP method. For the mean of T_2, the SBSP method is the winner with score 9, and the second best method is the SBNP method with score 6, and 5 out of 6 are not rejecting H_0.

3.6.2 Second scenario: Gumbel copula model

In this subsection, we generate data sets from the Gumbel copula model, where the variables are both following Uniform(0,1). For the SBSP method, we assume the Normal copula in order to compute the probabilities $\hat{h}_{ij}$ by Equation (3.5). We want to investigate the performance of this proposed bootstrap method when the copula model used to generate the data is different to the one we assume for the semi-parametric predictive method. Note that both copula models used here are symmetric, and their density functions at three different dependence levels τ are presented in Figure 3.9.

We compare the proposed methods to Efron's bootstrap method through simulations. Tables 3.9 to 3.13 present the chi-squared values for r, τ, rs and the means of T_1 and T_2, and the results are visualized in Figures 3.10 to 3.14, respectively. The best method not rejecting H_0 when the Pearson correlation r is considered is the SBNP method; it provides 10 small chi-squared values out of 12 settings as shown in Figure 3.10. For the correlation of τ, Figure 3.11 shows that the SBNP method and Efron's bootstrap both do not reject H_0 at 8 cases, and they also both reject H_0 4 times due to the over-coverage proportions in the first confidence regions. For the Spearman correlation rs, Efron's method continues to provide the best results as shown in Figure 3.12. In 9 cases, we obtain small chi-squared values, and the results corresponding to the 3 other cases are large due to over-coverage. Figures 3.13 and 3.14 indicate that the SBSP method is the best one to not reject H_0. For each of $\overline{T_1}$ and $\overline{T_2}$, it provides small chi-squared values 10 times. For these two statistics, the actual coverage rates based on the SEB and Efron's methods are always under-coverage when $n = 10$, then as n increases to 50 and 100, they both accomplish good values in most settings.

(a) $\tau = 0.25$, Gumbel copula (b) $\tau = 0.25$, Normal copula

(c) $\tau = 0.50$, Gumbel copula (d) $\tau = 0.50$, Normal copula

(e) $\tau = 0.75$, Gumbel copula (f) $\tau = 0.75$, Normal copula

Figure 3.9: Density functions of Gumbel and Normal copulas with uniform margins when $\tau = 0.25, 0.50$ and 0.75.

$n =$			10				50				100			
τ	θ	r	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB
0	1	0	$\underline{28.50}$	9.84	$\underline{31.42}$	$\underline{17.36}$	6.00	7.98	13.40	9.08	7.80	8.80	7.10	5.68
0.25	1.3333	0.36	$\underline{23.72}$	$\underline{19.10}$	$\underline{33.02}$	$\underline{18.96}$	16.86	14.52	13.08	7.88	14.12	6.28	9.86	5.32
0.50	2	0.68	12.26	8.12	$\underline{68.14}$	$\underline{25.66}$	$\underline{25.24}$	10.28	$\underline{21.92}$	7.58	$\underline{19.40}$	6.96	7.64	7.86
0.75	4	0.92	16.58	$\overline{41.54}$	$\underline{116.02}$	$\overline{60.10}$	$\underline{33.00}$	16.82	12.90	$\underline{35.94}$	$\underline{22.66}$	15.32	9.60	$\underline{59.52}$

Table 3.9: The chi-squared values obtained from the coverage proportions for the true Pearson correlation r.

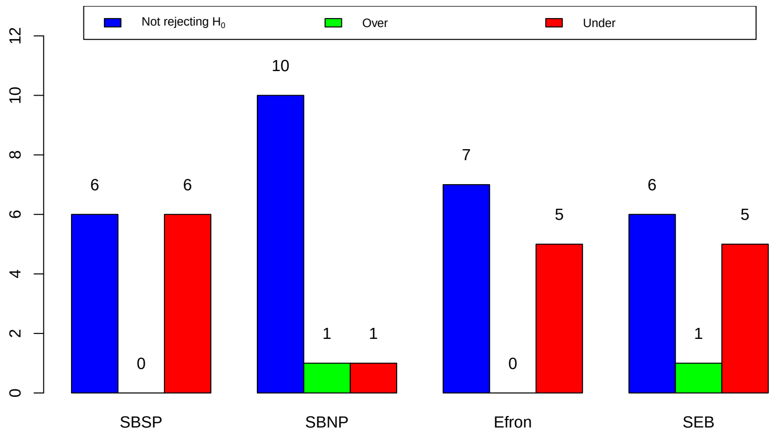

Figure 3.10: The number of times that H_0 is not rejected, and the number of times that H_0 is rejected due to the over- and under- coverage proportions for Pearson correlation r based on each bootstrap.

$n=$		10				50				100			
τ	θ	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB
0	1	14.40	$\overline{30.58}$	$\overline{41.98}$	$\overline{26.98}$	6.84	4.78	10.32	8.10	12.84	12.28	5.66	11.46
0.25	1.3333	$\overline{31.68}$	$\overline{43.54}$	$\overline{30.44}$	$\overline{62.58}$	$\underline{18.74}$	13.46	12.88	12.88	$\underline{17.52}$	4.14	3.18	3.40
0.50	2	$\underline{26.34}$	$\overline{52.60}$	$\overline{36.94}$	$\overline{98.36}$	$\underline{29.18}$	8.06	9.74	8.92	$\underline{17.02}$	11.22	3.14	4.98
0.75	4	$\overline{303.20}$	$\overline{58.60}$	$\overline{85.00}$	$\overline{49.04}$	$\underline{28.76}$	15.72	13.16	$\underline{33.10}$	$\underline{18.90}$	10.16	9.86	$\underline{49.56}$

Table 3.10: The chi-squared values obtained from the coverage proportions for the true Kendall correlation τ.

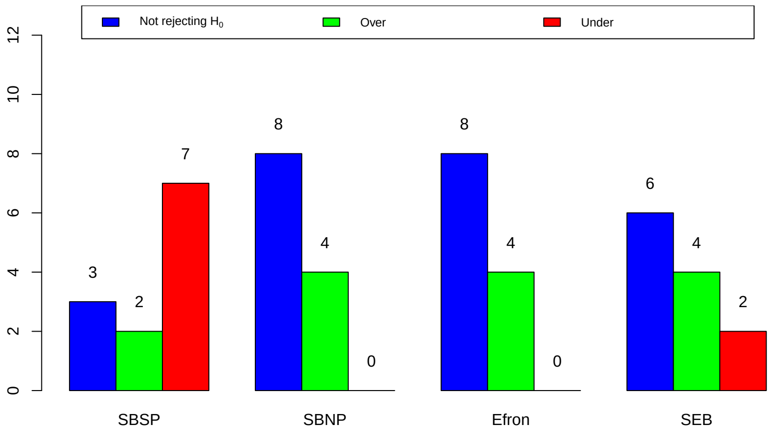

Figure 3.11: The number of times that H_0 is not rejected, and the number of times that H_0 is rejected due to the over- and under- coverage proportions for Kendall correlation τ based on each bootstrap.

$n =$			10				50				100			
τ	θ	rs	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB
0	1	0.00	$\underline{27.70}$	13.38	15.38	$\overline{21.02}$	10.30	13.10	11.70	7.78	9.20	15.48	16.20	6.86
0.25	1.3333	0.36	$\underline{23.76}$	$\overline{23.48}$	$\overline{20.86}$	$\overline{36.52}$	$\underline{19.74}$	5.74	11.26	9.58	13.42	7.56	12.16	15.94
0.50	2	0.68	$\overline{42.74}$	$\overline{44.46}$	$\overline{29.50}$	$\overline{69.06}$	$\underline{18.92}$	5.18	7.04	6.82	16.28	5.70	5.72	7.94
0.75	4	0.92	$\overline{123.14}$	$\underline{112.48}$	$\overline{105.14}$	$\underline{398.76}$	$\underline{18.00}$	$\underline{25.68}$	10.70	$\underline{266.32}$	12.02	$\underline{22.94}$	15.02	$\underline{248.46}$

Table 3.11: The chi-squared values obtained from the coverage proportions for the true Spearman correlation rs.

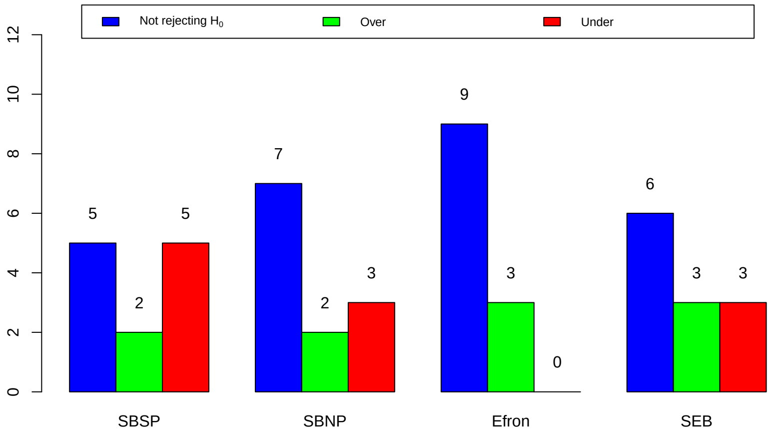

Figure 3.12: The number of times that H_0 is not rejected, and the number of times that H_0 is rejected due to the over- and under- coverage proportions for Spearman correlation rs based on each bootstrap.

$n =$			10				50				100			
τ	θ	$\overline{T_1}$	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB
0	1	1	13.46	8.68	$\underline{48.98}$	$\underline{42.68}$	12.86	10.42	5.44	2.58	5.16	12.96	9.34	10.50
0.25	1.3333	1	13.46	11.88	$\underline{33.98}$	$\underline{33.18}$	13.28	$\overline{17.16}$	14.62	14.88	$\underline{19.56}$	7.90	$\underline{16.96}$	13.88
0.50	2	1	7.94	13.38	$\underline{35.64}$	$\underline{28.92}$	6.26	3.72	6.54	6.56	12.64	8.50	14.64	$\underline{17.16}$
0.75	4	1	$\overline{20.86}$	$\overline{22.52}$	$\underline{24.82}$	$\underline{27.12}$	7.60	8.52	8.70	12.14	13.56	$\underline{21.94}$	$\underline{17.02}$	12.22

Table 3.12: The chi-squared values obtained from the coverage proportions for the function $\overline{T_1}$.

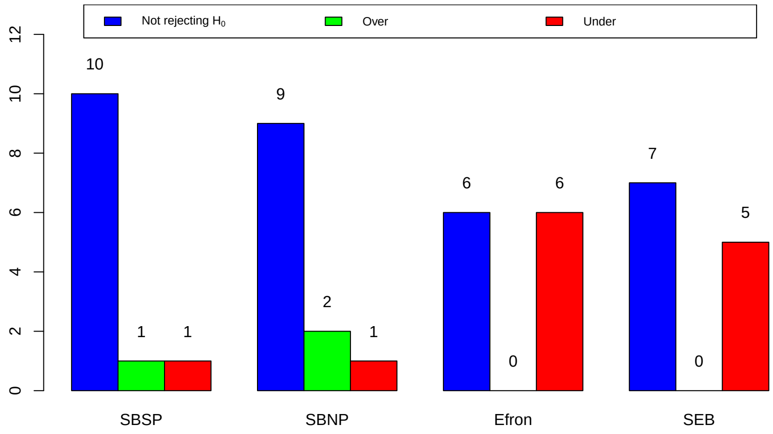

Figure 3.13: The number of times that H_0 is not rejected, and the number of times that H_0 is rejected due to the over- and under- coverage proportions for $\overline{T_1}$ based on each bootstrap.

$n =$			10				50				100			
τ	θ	$\overline{T_2}$	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB
0	1	0.167	8.46	$\overline{18.20}$	$\underline{110.82}$	$\underline{104.00}$	9.36	9.54	7.40	4.96	8.54	6.80	9.30	10.00
0.25	1.3333	0.199	5.24	$\overline{26.44}$	$\underline{109.44}$	$\underline{99.06}$	12.04	12.00	12.14	13.18	$\underline{34.80}$	14.78	16.64	$\underline{17.72}$
0.50	2	0.225	11.70	$\overline{24.08}$	$\underline{83.70}$	$\underline{68.78}$	5.76	6.12	8.30	11.80	$\underline{23.40}$	$\underline{17.14}$	$\underline{20.18}$	$\underline{21.70}$
0.75	4	0.243	15.02	$\overline{18.26}$	$\underline{78.00}$	$\underline{54.00}$	5.66	10.48	10.90	6.86	10.18	12.76	12.34	9.56

Table 3.13: The chi-squared values obtained from the coverage proportions for the function $\overline{T_2}$.

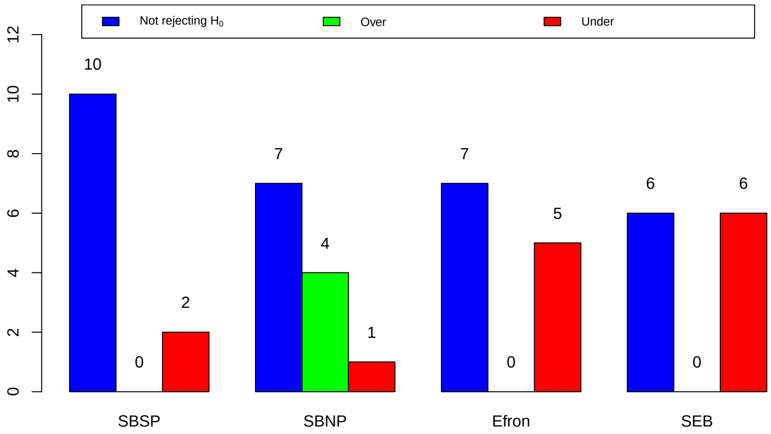

Figure 3.14: The number of times that H_0 is not rejected, and the number of times that H_0 is rejected due to the over- and under- coverage proportions for $\overline{T_2}$ based on each bootstrap.

Function	SBSP	SBNP	Efron	SEB
r	2 (2,0,0)	3 (3,0,0)	2 (2,0,0)	5 (4,0,1)
τ	2 (1,0,1)	2 (2,0,0)	7 (6,1,0)	2 (1,1,0)
rs	1 (1,0,0)	5 (5,0,0)	4 (1,3,0)	2 (2,0,0)
$\overline{T_1}$	5 (4,1,0)	5 (5,0,0)	0 (0,0,0)	2 (2,0,0)
$\overline{T_2}$	6 (6,0,0)	4 (3,0,1)	0 (0,0,0)	2 (2,0,0)

Table 3.14: The summary of scores for the bootstrap methods on the five functions of interest along with the number of chi-squared values of each score that are based on equal, over- and under- coverage proportions, (#Equal, #Over, #Under).

The summary of scores for each bootstrap method providing the lowest chi-squared values on each function of interest is presented in Table 3.14. For the Pearson's correlation r, the SEB method provides a lower chi-squared value than those of the other methods 5 times. Efron's method accomplishes lowest values in 7 cases for the Kendall's correlation τ, and in 6 cases, H_0 is not rejected. The SBNP method is the winer for the Spearman's correlation rs and $\overline{T_1}$; it scores 5 for these functions, and all values in these cases do not support rejection H_0. The SBSP method is a distant second when we consider $\overline{T_1}$, and it is a distant first for $\overline{T_2}$.

These simulation studies have led to some important insights. First, for r, τ and rs, the SEB method mostly provides large chi-squared values when there is high dependence between variables. These large values occur because we observed under-coverage proportions in the first cells usually. Based on the simulation results, it is advised to use small bandwidths for the SEB method when $\tau > 0.50$ or $\tau < -0.50$ in order to improve the results. For clarity, we repeat the simulations for the second scenario, but with different bandwidth sizes. We use Equation (3.14) for large bandwidths, and replace $n^{-\frac{1}{4}}$ by n^{-1} to make small bandwidths. This change in the bandwidths leads to less variability in sampling during the bootstrap procedure. The simulation results are presented in Table 3.15, and it is clear that the small bandwidths mostly lead to better results.

Secondly, for the semi-parametric predictive method, it is better to use the Itau estimation technique to estimate the copula parameter θ, in particular when $\tau \leq 0.4$.

$n =$	10		50		100	
Bandwidth	small (n^{-1})	large ($n^{-\frac{1}{4}}$)	small (n^{-1})	large ($n^{-\frac{1}{4}}$)	small (n^{-1})	large ($n^{-\frac{1}{4}}$)
$r = 0.92$	$\underline{85.10}$	$\overline{60.10}$	$\underline{19.84}$	$\underline{35.94}$	11.30	$\underline{59.52}$
$\tau = 0.75$	$\overline{91.18}$	$\overline{49.04}$	12.52	$\underline{33.10}$	9.20	$\underline{49.56}$
$rs = 0.92$	15.22	$\underline{398.76}$	$\underline{25.76}$	$\underline{266.32}$	$\underline{18.20}$	$\underline{248.46}$

Table 3.15: The chi-squared values obtained from the coverage proportions for r, τ and rs using the SEB method with two sizes of bandwidths, where the copula parameter $\theta = 4$.

$n =$	10				50				100			
Function	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB
$r = 0.89$	$\underline{37.54}$	$\underline{4341.90}$	$\underline{249.22}$	15.70	$\underline{425.88}$	$\underline{154.50}$	$\underline{33.40}$	13.48	$\underline{1488.24}$	$\underline{34.22}$	$\underline{19.36}$	12.44
$\tau = 0.75$	$\overline{210.52}$	$\underline{5701.30}$	$\overline{34.54}$	$\overline{64.18}$	$\underline{70.46}$	$\underline{297.40}$	8.54	6.34	$\underline{42.84}$	11.58	9.80	$\underline{18.66}$
$rs = 0.91$	$\overline{185.64}$	$\underline{7369.40}$	$\overline{87.42}$	$\overline{19.40}$	$\underline{37.02}$	$\underline{142.04}$	7.04	11.44	$\underline{31.12}$	7.42	8.94	10.34
$\overline{T_1} = 6.00$	$\underline{27.14}$	$\underline{47.86}$	$\underline{50.04}$	$\underline{44.28}$	12.56	5.54	6.80	5.60	7.68	10.62	12.62	13.94
$\overline{T_2} = 58.63$	$\underline{89.92}$	$\underline{236.22}$	$\underline{183.74}$	$\underline{174.48}$	$\overline{18.52}$	$\overline{17.40}$	12.68	10.28	10.20	3.40	7.60	$\overline{20.28}$

Table 3.16: The chi-squared values obtained from the coverage proportions for the true functions of interest, where the copula parameter $\theta = 6$.

The Itau estimation technique helps the SBSP method to provide better accuracy in terms of defining the actual coverage proportions for some functions of interest. Thirdly, the purpose of assuming Normal copula in the second step of the semi-parametric predictive method is that this copula can model negative and positive correlations. Lastly, the SBSP and SBNP methods do not work well for the correlations r, τ and rs when the data distribution is asymmetric, and this will be shown in the next simulations.

We generate $N = 1000$ data sets from the Clayton copula, where X follows Normal($\mu = 1, \sigma = 1$) and Y follows Normal($\mu = 5, \sigma = 3$). The data distribution is asymmetric as can be seen from Figure 3.15. We apply the bootstrap methods $B = 1000$ times for each generated data set and the simulation results are presented in Table 3.16. For the Pearson's, Kendall's and Spearman's correlations, the SBSP and SBNP methods provide poor outcomes in terms of the coverage proportions. As n increases, the SBSP method performs poorly, but the results of the SBNP

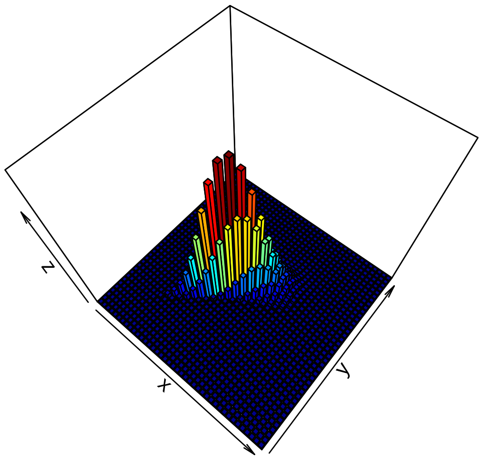

Figure 3.15: The empirical function of Clayton copula, where X∼Normal(μ=1,σ=1) and Y∼Normal(μ=5,σ=3), $\tau = 0.75$ and $\theta = 6$.

$n =$	10				50				100			
Function	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB
$\overline{T_1} = 0.58$	$\overline{48.33}$	$\overline{32.10}$	$\underline{52.56}$	$\underline{42.72}$	12.42	13.42	7.62	8.74	3.80	11.52	11.40	6.70
$\overline{T_2} = 0.01$	$\overline{68.81}$	$\overline{61.10}$	$\underline{436.24}$	$\underline{442.42}$	12.96	10.98	10.64	16.66	$\underline{85.84}$	12.50	9.60	5.26

Table 3.17: The chi-squared values for the averages of T_1 and T_2, where $X \sim \text{Uniform}(0,1)$ and $Y = (X - 0.5)^2$.

method are improved. For the functions $\overline{T_1}$ and $\overline{T_2}$ when $n = 10$, they provide large chi-squared values due to the under-coverage proportions in the first cells; for larger n, results are better. However, all bootstrap methods provide good results when $n = 50$ and 100.

To investigate the performances of the bootstrap methods in different scenarios to the ones before, we consider two relationships between X and Y. For a uniformly distributed random variable X, we set Y first as $Y = (X - 0.5)^2$ and second as $Y = (X - 0.5)^2 + \epsilon$, where $\epsilon \sim \text{Normal}(\mu = 0, \sigma = 0.02)$. The shape of quadratic relationship is as shown in Figure 3.16. In this investigation, we take the means of

$n =$	10				50				100			
Function	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB
$\overline{T_1} = 0.5833$	5.02	7.56	73.36	75.72	6.58	7.84	7.20	17.78	7.12	4.60	8.28	9.30
$\overline{T_2} = 0.0064$	410.28	286.08	526.46	524.34	186.16	20.42	40.74	44.58	182.60	16.10	11.88	15.36

Table 3.18: The chi-squared values for the averages of T_1 and T_2; $X \sim \text{Uniform}(0,1)$ and $Y = (X-0.5)^2 + \epsilon$, where $\epsilon \sim \text{Normal}(0, 0.02)$.

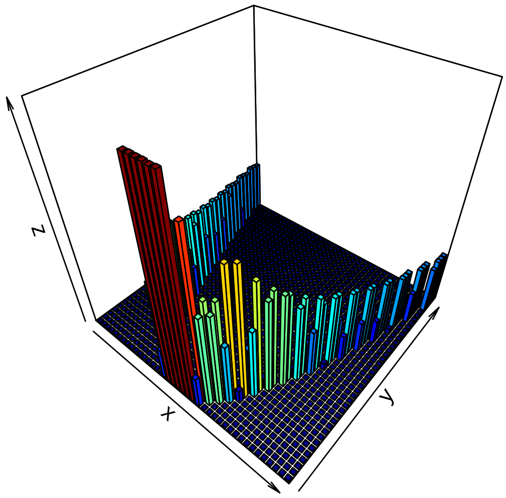

Figure 3.16: The empirical function of quadratic distribution, where X ~Uniform(0,1) and $Y = (X-0.5)^2$.

$T_1 = X + Y$ and $T_2 = XY^2$ into account only, and their corresponding chi-squared values are presented in Tables 3.17 and 3.18. For $\overline{T_1}$ in the first case without the epsilon random term, the SBSP and SBNP methods lead to over-coverage proportions in the first cells when $n = 10$, contrary to Efron's bootstrap and the SEB method; they both lead to under-coverage. As n increases to 50 and 100, all methods fail to reject H_0. In the second case where $Y = (X-0.5)^2 + \epsilon$, the SBSP and SBNP methods provide small chi-squared values for the same function at all different sample sizes. However, Efron's bootstrap and the SEB method provide large values when $n = 10$, and the results are improved as n increases. For $\overline{T_2}$ in both simulation studies, all methods provide large chi-squared values when the sample size is small, and as n increases to 50 and 100, the results are improved except the ones of SBSP. The SBSP method continues providing large values as n gets large due to the under-coverage proportions occur in the first cells.

3.7 Concluding remarks

In this chapter, we introduced three smoothed bootstrap methods for bivariate data based on the semi-parametric and nonparametric predictive methods and uniform kernels. We compared these to Efron's method through simulation studies, computing the coverage proportions for the Pearson, Kendall and Spearman correlations along with the means of T_1 and T_2, where $T_1 = X + Y$ and $T_2 = XY^2$. The comparisons were conducted with different dependence levels τ and $n = 10, 50$ and 100.

Efron's bootstrap method for bivariate data is sampling with replacement from the original data set, so this method often leads to include ties to the bootstrap samples especially when the sample size is small. The smoothed bootstrap methods using the semi-parametric and nonparametric predictive methods overcome this issue as they lead to more variation in sampling. Also, the bootstrap method based on uniform kernels leads to more variation in sampling, so the complication of ties is avoided.

Through the simulation studies when the data distribution is symmetric, the SBNP method mostly outperforms the other methods when we consider the Pearson's correlation r, and Efron's method mostly performs better for Kendall τ and Spearman rs correlations. For $\overline{T_1}$ and $\overline{T_2}$, the smoothed bootstrap methods based on the semi-parametric and nonparametric predictive methods mostly provide the lowest discrepancies between the nominal and observed coverage proportions, in particular when the sample size is small. When the data distribution is not symmetric, it is advised to use either the smoothed Efron's bootstrap or the original Efron's method because they lead to perform well. In case Kendall τ is very large, it is advised to use either Efron's method or the smoothed Efron's method with small bandwidths to make inferences for r, τ and rs. From the simulations, it seems that the smoothed bootstrap methods mostly perform better than Efron's method for any function of interest in case there is no relationship between the variables and the sample size is small. Also, if the variables show such a linear relationship with randomness, the smoothed bootstrap methods can lead to better results in comparison to those of Efron's method.

In running R codes, applying the SBSP method takes approximately four times as long as Efron's procedure. This is mainly because of fitting the Normal copula on each generated data set to compute the probabilities h_{ij}, sampling uniformly from the bounded blocks and getting observations from the fitted normal tails to be accepted for the end blocks. For the SBNP and SEB methods, their calculation times are approximately equal to the time corresponding to Efron's method.

Chapter 4

Hypothesis tests using smoothed bootstrap methods

4.1 Introduction

For testing statistical hypotheses, the classical statistical methods have been widely used in the literature, and they are considered as the standard methods although their assumptions are often not met if the observed data set is complicated. To avoid the mathematical assumptions, Efron and Tibshirani [33], and Rasmussen [71] and Strube [80] used the standard bootstrap method to test statistical hypotheses. This bootstrap method is easy to implement with good approximation results, but it requires a lot of computations. This motivates to use the proposed smoothed bootstrap methods for testing of statistical hypotheses and to compare the results to those of Efron's method. Tests based on both right-censored data and bivariate data will be considered in this chapter.

In Chapter 2, the smoothed bootstrap method for right-censored data was compared to Efron's method in terms of the coverage probability, and it was also used to estimate the survival function. In this chapter, we want to use the bootstrap methods to compute the Type 1 error rate for a proposed null hypothesis with significance level 2α. This can be done by the $100(1-2\alpha)\%$ bootstrap confidence intervals. Based on the bootstrap methods, the achieved significance level is used to test whether the two samples have equal median or not, and compute the Type 1 error rate. The evaluations of power and Type 2 error rate are useful and important when comparing the bootstrap methods and this is left as a topic for future research.

For bivariate data, Dolker *et al.* [28] stated that Efron's bootstrap method leads to bad approximation results for Pearson correlation when the sample size is small due to the ties occurred in the bootstrap samples. Lee and Rodgers [58] proposed an univariate sampling procedure for bootstrapping correlation coefficients, and compared it to Efron's method through simulation studies. They performed hypothesis testing and compute their corresponding Type 1 error rates via bootstrap confidence intervals. They found that the proposed bootstrap method performs better. In this chapter, we use the proposed smoothed bootstrap methods presented in Chapter 3 to test hypotheses and compute their Type 1 error rates. The results will be compared to those of Efron's method. For future research, we will investigate the power and Type 2 error rate, which are helpful and crucial.

In Section 4.2, the Type 1 error rates for quartiles' tests are computed based on the smoothed bootstrap method and Efron's bootstrap method. In Section 4.3, the Type 1 error rate for two sample medians' tests is computed using the achieved significance level based on the bootstrap methods. Note that the simulations considered in Sections 4.2 and 4.3 include right-censored observations. Section 4.4 compares the proposed bootstrap methods, introduced in Chapter 3, and Efron's bootstrap in computing the Type 1 error rate for Pearson correlation test. Section 4.5 compares the proposed bootstrap methods, introduced in Chapter 3, and Efron's bootstrap in computing the Type 1 error rate for Kendall correlation test. In the last section, we point out some concluding remarks.

4.2 Hypothesis tests for quartiles

In this section, we compute the Type 1 error rates of quartiles' hypothesis tests based on the bootstrap methods, presented in Chapter 2, where data contains right-censored observations. To do investigations based on the bootstrap methods, we simulate data sets including right-censored observations from two different scenarios. For the first scenario, we set that the event time T follows Beta distribution with parameters $\alpha = 1.2$ and $\beta = 3.2$, and the censored time C follows Uniform distribution with parameters $a = 0$ and $b = 1.82$. The second scenario is defined as $T \sim$ Log-Normal$(\mu = 0, \sigma = 1)$ and $C \sim$ Weibull$(\alpha = 3, \beta = 3.7)$. The censoring proportion p in the generated data sets for both scenarios is 15%.

To conduct comparisons between the two bootstrap methods, we generate $N = 1000$ data sets from either one of the scenarios we proposed above. Then for each generated data set, we apply the methods $B = 1000$ times. This leads to have 1000 bootstrap samples based on each method. We then compute the quartile of interest at each bootstrap sample, and from the resulting values, we can define the $100(1 - 2\alpha)\%$ bootstrap confidence interval for the quartile. We count one if the value of the quartile specified in the null hypothesis is not included in the confidence interval; otherwise we count zero. We repeat this procedure for all $N = 1000$ generated data sets, then count the number of times the null hypothesis was rejected over the 1000 trials. This ratio will be the Type 1 error rate of the quartile's hypothesis test with significance level 2α.

It should be noted that the bootstrap samples obtained by Efron's method often include some censored observations, so we use the Kaplan-Meier estimator to find their corresponding quartiles. Suppose we are interested in the median, we should find a time t such that $\hat{S}(t) = 0.50$ in each bootstrap sample. Unfortunately in some samples, we cannot find that time t. To overcome this issue, we use the three options, denoted by $\mathrm{E}_{(1)}$, $\mathrm{E}_{(2)}$ and $\mathrm{E}_{(3)}$, which are described in Subsection 2.4.1. In the following simulations, the NA number represents the number of Efron's bootstrap samples whose quartiles are not found, and this number is out of 1000000. In case there is a bootstrap sample including only right-censored observations, that sample will be replaced by another sample that includes at least one event time. This action is counted and denoted by ABS, which is out of 1000000.

For the smoothed bootstrap method, we consider three different strategies in sampling observations from the $n + 1$ intervals partitioning the sample space. The first one is sampling uniformly from all intervals, and this is denoted by SB. Secondly, we assume exponential tail for each interval and sample from the tails to create the bootstrap samples, which is denoted by SB_{exp}. The third one is sampling uniformly

H_0 :		$Q_1 = 0.117$						$Q_2 = 0.236$						$Q_3 = 0.396$					
n	measures	SB	SB$_{exp}$	SB$_{Lexp}$	E$_{(1)}$	E$_{(2)}$	E$_{(3)}$	SB	SB$_{exp}$	SB$_{Lexp}$	E$_{(1)}$	E$_{(2)}$	E$_{(3)}$	SB	SB$_{exp}$	SB$_{Lexp}$	E$_{(1)}$	E$_{(2)}$	E$_{(3)}$
10	Type 1	0.103	0.103	0.105	0.107	0.107	0.107	0.090	0.096	0.097	0.151	0.151	0.149	0.068	0.110	0.111	0.200	0.202	0.172
	NA	—	—	—	228	0	0	—	—	—	3736	0	0	—	—	—	32821	0	0
	ABS	—	—	—	12	12	12	—	—	—	12	12	12	—	—	—	12	12	12
50	Type 1	0.098	0.098	0.101	0.108	0.108	0.108	0.126	0.126	0.114	0.117	0.117	0.117	0.121	0.121	0.126	0.108	0.107	0.107
	NA	—	—	—	0	0	0	—	—	—	0	0	0	—	—	—	56	0	0
	ABS	—	—	—	0	0	0	—	—	—	0	0	0	—	—	—	0	0	0
100	Type 1	0.100	0.100	0.098	0.100	0.100	0.100	0.120	0.120	0.117	0.104	0.104	0.104	0.133	0.133	0.134	0.114	0.114	0.114
	NA	—	—	—	0	0	0	—	—	—	0	0	0	—	—	—	0	0	0
	ABS	—	—	—	0	0	0	—	—	—	0	0	0	—	—	—	0	0	0
500	Type 1	0.104	0.104	0.104	0.100	0.100	0.100	0.126	0.126	0.126	0.110	0.110	0.110	0.121	0.121	0.121	0.094	0.094	0.094
	NA	—	—	—	0	0	0	—	—	—	0	0	0	—	—	—	0	0	0
	ABS	—	—	—	0	0	0	—	—	—	0	0	0	—	—	—	0	0	0

Table 4.1: Type 1 error rates with significance level $2\alpha = 0.10$, $T \sim \text{Beta}(\alpha = 1.2, \beta = 3.2)$, $C \sim \text{Unif}(a = 0, b = 1.82)$ and $p = 0.15$.

H_0 :		$Q_1 = 0.117$						$Q_2 = 0.236$						$Q_3 = 0.396$					
n	measures	SB	SB$_{exp}$	SB$_{Lexp}$	E$_{(1)}$	E$_{(2)}$	E$_{(3)}$	SB	SB$_{exp}$	SB$_{Lexp}$	E$_{(1)}$	E$_{(2)}$	E$_{(3)}$	SB	SB$_{exp}$	SB$_{Lexp}$	E$_{(1)}$	E$_{(2)}$	E$_{(3)}$
10	Type 1	0.051	0.049	0.050	0.088	0.088	0.088	0.046	0.048	0.050	0.070	0.070	0.068	0.020	0.072	0.065	0.183	0.181	0.146
	NA	—	—	—	228	0	0	—	—	—	3736	0	0	—	—	—	32821	0	0
	ABS	—	—	—	12	12	12	—	—	—	12	12	12	—	—	—	12	12	12
50	Type 1	0.054	0.054	0.045	0.059	0.059	0.059	0.066	0.066	0.070	0.069	0.069	0.069	0.067	0.067	0.067	0.059	0.059	0.059
	NA	—	—	—	0	0	0	—	—	—	0	0	0	—	—	—	56	0	0
	ABS	—	—	—	0	0	0	—	—	—	0	0	0	—	—	—	0	0	0
100	Type 1	0.047	0.047	0.052	0.045	0.045	0.045	0.057	0.057	0.061	0.057	0.057	0.057	0.078	0.078	0.083	0.061	0.061	0.061
	NA	—	—	—	0	0	0	—	—	—	0	0	0	—	—	—	0	0	0
	ABS	—	—	—	0	0	0	—	—	—	0	0	0	—	—	—	0	0	0
500	Type 1	0.058	0.058	0.058	0.054	0.054	0.054	0.062	0.062	0.062	0.054	0.054	0.054	0.072	0.072	0.072	0.049	0.049	0.049
	NA	—	—	—	0	0	0	—	—	—	0	0	0	—	—	—	0	0	0
	ABS	—	—	—	0	0	0	—	—	—	0	0	0	—	—	—	0	0	0

Table 4.2: Type 1 error rates with significance level $2\alpha = 0.05$, $T \sim \text{Beta}(\alpha = 1.2, \beta = 3.2)$, $C \sim \text{Unif}(a = 0, b = 1.82)$ and $p = 0.15$.

from all intervals except the last intervals, for which we sample from the exponential tails. For this strategy, we indicate it by SB$_{Lexp}$. We consider these three ways in sampling to improve the results for the last quartile when n gets large.

Tables 4.1 and 4.2 present the Type 1 error rates for the quartiles' hypothesis tests with significance levels 0.10 and 0.05, respectively, where the simulated data sets are from the first scenario. The smoothed bootstrap with its three assumptions, SB, SB$_{exp}$ and SB$_{Lexp}$, provides lower discrepancies between the actual and nominal error rates for all quartiles' tests when $n = 10$. This superiority occurs not only due to the event observations obtained for the smoothed bootstrap samples, but also the KM estimator used in Efron's bootstrap samples is often not able to find the

H_0 :		$Q_1 = 0.509$					$Q_2 = 1$					$Q_3 = 1.963$				
n	measures	SB$_{exp}$	SB$_{Lexp}$	E$_{(1)}$	E$_{(2)}$	E$_{(3)}$	SB$_{exp}$	SB$_{Lexp}$	E$_{(1)}$	E$_{(2)}$	E$_{(3)}$	SB$_{exp}$	SB$_{Lexp}$	E$_{(1)}$	E$_{(2)}$	E$_{(3)}$
10	Type 1	0.092	0.096	0.103	0.103	0.103	0.098	0.093	0.128	0.119	0.126	0.104	0.108	0.287	0.304	0.172
	NA	—	—	1813	0	0	—	—	23589	0	0	—	—	167582	0	0
	ABS	—	—	61	61	61	—	—	61	61	61	—	—	61	61	61
50	Type 1	0.089	0.092	0.121	0.121	0.121	0.095	0.092	0.106	0.106	0.106	0.118	0.115	0.118	0.119	0.119
	NA	—	—	0	0	0	—	—	2	0	0	—	—	18178	0	0
	ABS	—	—	0	0	0	—	—	0	0	0	—	—	0	0	0
100	Type 1	0.084	0.090	0.100	0.100	0.100	0.097	0.102	0.101	0.101	0.101	0.119	0.117	0.116	0.117	0.117
	NA	—	—	0	0	0	—	—	0	0	0	—	—	1421	0	0
	ABS	—	—	0	0	0	—	—	0	0	0	—	—	0	0	0
500	Type 1	0.103	0.106	0.098	0.098	0.098	0.106	0.103	0.104	0.104	0.104	0.120	0.120	0.112	0.112	0.112
	NA	—	—	0	0	0	—	—	0	0	0	—	—	0	0	0
	ABS	—	—	0	0	0	—	—	0	0	0	—	—	0	0	0

Table 4.3: Type 1 error rates with significance level $2\alpha = 0.10$, $T \sim$ Log-Normal($\mu = 0, \sigma = 1$), $C \sim$ Weibull($\alpha = 3, \beta = 3.7$) and $p = 0.15$.

H_0 :		$Q_1 = 0.509$					$Q_2 = 1$					$Q_3 = 1.963$				
n	measures	SB$_{exp}$	SB$_{Lexp}$	E$_{(1)}$	E$_{(2)}$	E$_{(3)}$	SB$_{exp}$	SB$_{Lexp}$	E$_{(1)}$	E$_{(2)}$	E$_{(3)}$	SB$_{exp}$	SB$_{Lexp}$	E$_{(1)}$	E$_{(2)}$	E$_{(3)}$
10	Type 1	0.040	0.045	0.084	0.084	0.084	0.047	0.050	0.070	0.069	0.070	0.065	0.069	0.250	0.268	0.138
	NA	—	—	1813	0	0	—	—	23589	0	0	—	—	167582	0	0
	ABS	—	—	61	61	61	—	—	61	61	61	—	—	61	61	61
50	Type 1	0.047	0.049	0.066	0.066	0.066	0.054	0.054	0.056	0.056	0.056	0.066	0.059	0.062	0.062	0.062
	NA	—	—	0	0	0	—	—	2	0	0	—	—	18178	0	0
	ABS	—	—	0	0	0	—	—	0	0	0	—	—	0	0	0
100	Type 1	0.042	0.046	0.047	0.047	0.047	0.050	0.047	0.049	0.049	0.049	0.061	0.065	0.065	0.065	0.065
	NA	—	—	0	0	0	—	—	0	0	0	—	—	1421	0	0
	ABS	—	—	0	0	0	—	—	0	0	0	—	—	0	0	0
500	Type 1	0.054	0.057	0.054	0.054	0.054	0.047	0.043	0.050	0.050	0.050	0.070	0.066	0.062	0.062	0.062
	NA	—	—	0	0	0	—	—	0	0	0	—	—	0	0	0
	ABS	—	—	0	0	0	—	—	0	0	0	—	—	0	0	0

Table 4.4: Type 1 error rates with significance level $2\alpha = 0.05$, $T \sim$ Log-Normal($\mu = 0, \sigma = 1$), $C \sim$ Weibull($\alpha = 3, \beta = 3.7$) and $p = 0.15$.

H_0 :	$Q_1 = 0.117$		$Q_2 = 0.236$		$Q_3 = 0.396$	
n	Banks	Efron	Banks	Efron	Banks	Efron
10	0.102	0.099	0.080	0.136	0.081	0.096
50	0.089	0.113	0.099	0.112	0.099	0.111
100	0.099	0.103	0.113	0.109	0.095	0.103
500	0.097	0.103	0.101	0.102	0.087	0.091

Table 4.5: Type 1 error rates with significance level $2\alpha = 0.10$, Beta($\alpha = 1.2, \beta = 3.2$) and $p = 0$.

quartiles, in particular the second and third ones. In $228, 3736$ and 32821 bootstrap samples, we cannot find the first, second and third quartiles, respectively. As the sample size increases to 50, 100 and 500, both methods provide good results, but Efron's method is better, and the NA and ABS numbers decrease toward zero. These decreases lead to equal results when $\text{E}_{(1)}$, $\text{E}_{(2)}$ and $\text{E}_{(3)}$ are used. Also, SB, SB_{exp} and SB_{Lexp} provide approximately equal outcomes at these large sample sizes.

In the second scenario, we should note that the data space is $(0, \infty)$, so that the last intervals for the smoothed method are not bounded. In this case, we can only use the SB_{exp} and SB_{Lexp} assumptions, not SB. The results of Type 1 error rates for the quartiles' hypothesis tests with significance levels 0.10 and 0.05 are listed in Tables 4.3 and 4.4. The SB_{exp} and SB_{Lexp} methods again outperform Efron's method in defining the Type 1 error rates when the sample size is small. As n gets large, both methods perform well as we observed in Tables 4.1 and 4.2.

As a special case when data includes only failures, no censored observations, we will use the original Banks' bootstrap and the standard Efron's bootstrap methods to compute the Type 1 error rates for the quartiles' hypothesis tests. In the simulations, we use Beta($\alpha = 1.2, \beta = 3.2$) to create data sets and repeat the same comparison procedure as in the previous simulations. Tables 4.5 and 4.6 present the Type 1 error rates for the quartiles' hypothesis tests based on Banks' and Efron's methods with significance levels 0.10 and 0.05, respectively. It is again that the Banks' bootstrap method performs better, in particular when $n = 10$ and $2\alpha = 0.05$. As n gets large, both methods perform well.

$H_0:$	$Q_1 = 0.117$		$Q_2 = 0.236$		$Q_3 = 0.396$	
n	Banks	Efron	Banks	Efron	Banks	Efron
10	0.052	0.089	0.046	0.064	0.014	0.086
50	0.046	0.059	0.058	0.060	0.055	0.069
100	0.043	0.042	0.054	0.060	0.054	0.058
500	0.052	0.058	0.057	0.056	0.040	0.042

Table 4.6: Type 1 error rates with significance level $2\alpha = 0.05$, Beta($\alpha = 1.2, \beta = 3.2$) and $p = 0$.

4.3 The two sample problem

In the literature, the achieved significance level ASL has been used to draw a conclusion for a hypothesis test $H_0 : \theta_1 = \theta_2$, where θ_1 and θ_2 represent the function of interest in the first and second samples, respectively. Having observed $\hat{\theta} = \hat{\theta}_1 - \hat{\theta}_2$, the achieved significance level is defined as the probability of observing at least that large a value when the null hypothesis is true,

$$ASL = \text{Prob}_{H_0}\{\hat{\theta}^* \geq \hat{\theta}\} \tag{4.1}$$

The smaller the value of ASL, the stronger the evidence against H_0. The variable $\hat{\theta}$ is fixed at its observed value and the random quantity $\hat{\theta}^*$ has the null hypothesis distribution, the distribution of $\hat{\theta}$ if H_0 is true [34].

Efron and Tibshirani [34] used the achieved significance level to test whether the two samples have equal mean or not. Suppose we have two samples $\mathbf{z} = z_1, z_2, \dots, z_n$ and $\mathbf{y} = y_1, y_2, \dots, y_m$ from possibly different probability distributions, and we wish to test the null hypothesis $H_0 : \mu_1 = \mu_2$. Efron and Tibshirani [34] use Efron's bootstrap method to approximate the ASL value, then H_0 is rejected when $\widehat{ASL} < 2\alpha$. The computation of the bootstrap test statistic for testing the null hypothesis is as follows

(i) Combine $\mathbf{z}$ and $\mathbf{y}$ samples together, so we get a sample $\mathbf{x}$ of size $n + m$. $\mathbf{x} = z_1, z_2, \dots, z_n, y_1, y_2, \dots, y_m$

(ii) Draw B bootstrap samples of size $n + m$ with replacement from $\mathbf{x}$, and call the first n observations $\mathbf{z}^{*b}$ and the remaining m observations $\mathbf{y}^{*b}$ for $b = 1, 2, \ldots, B$.

(iii) On each bootstrap sample, we compute the means of $\mathbf{z}^{*b}$ and $\mathbf{y}^{*b}$, then find $A^{*b} = \overline{\mathbf{z}}^{*b} - \overline{\mathbf{y}}^{*b}$, $b = 1, 2, \ldots, B$.

(iv) The achieved significance level ASL can be approximated by

$$\widehat{ASL} = \frac{\sum_{b=1}^{B}\{A^{*b} \geq A_{obs}\}}{B} \tag{4.2}$$

where $A_{obs} = \overline{\mathbf{z}} - \overline{\mathbf{y}}$, and $\overline{\mathbf{z}}$ and $\overline{\mathbf{y}}$ are the sample means of the two original samples.

In this section, we want to use the proposed strategy to test whether two samples have equal median or not based on the bootstrap methods, presented in Chapter 2, and conduct comparisons through simulations. We will compute the Type 1 error rate for the following hypothesis test

$$H_0 : Q_2^1 = Q_2^2 \quad \text{VS.} \quad H_1 : Q_2^1 \neq Q_2^2 \tag{4.3}$$

To conduct comparisons between the bootstrap methods through simulations, we generate two data sets, both of size n, from the second scenario we proposed in Section 4.2. We compute their medians $\hat{Q}_2^1$ and $\hat{Q}_2^2$ and calculate $A_{obs} = \hat{Q}_2^1 - \hat{Q}_2^2$. The two generated data sets will be combined together, so that the new sample will be of size $2n$. Based on each bootstrap method, we draw 1000 samples of size $2n$, and call the first n observations $\mathbf{z}^{*b}$ and the remaining n observations $\mathbf{y}^{*b}$ for $b = 1, 2, \ldots, B$. We then compute $A^{*b} = \hat{Q}_2(\mathbf{z}^{*b}) - \hat{Q}_2(\mathbf{y}^{*b})$ on each bootstrap sample. This leads to have 1000 A^* values. Finally, we compute the approximation of ASL, and reject H_0 if $\widehat{ASL} < 2\alpha$. We repeat this procedure 1000 times, then count the number of times that we reject the null hypothesis. We take the ratio out of 1000 due to the 1000 trials, and consider the best method is the one that its corresponding ratio is closer to 2α. The final results of the simulations are presented in Tables 4.7 and 4.8 with two different significance levels.

As the sample space of the underlying distribution is $(0, \infty)$, we only consider SB$_{exp}$ and SB$_{Lexp}$ for the smoothed bootstrap method, and with Efron' method, we

n	SB$_{exp}$	SB$_{Lexp}$	E$_{(2)}$	E$_{(3)}$
10	0.078	0.075	0.091	0.089
50	0.079	0.079	0.090	0.090
100	0.100	0.101	0.107	0.107
500	0.105	0.101	0.104	0.104

Table 4.7: Type 1 error rates with significance level $2\alpha = 0.10$, and all samples created by $(T \sim \text{Log-Normal}(\mu = 0, \sigma = 1), C \sim \text{Weibull}(\alpha = 3, \beta = 3.7), \text{where} \;\; p = 0.15)$.

n	SB$_{exp}$	SB$_{Lexp}$	E$_{(2)}$	E$_{(3)}$
10	0.025	0.025	0.031	0.031
50	0.039	0.041	0.039	0.039
100	0.047	0.046	0.049	0.049
500	0.043	0.042	0.043	0.043

Table 4.8: Type 1 error rates with significance level $2\alpha = 0.05$, and all samples created by $(T \sim \text{Log-Normal}(\mu = 0, \sigma = 1), C \sim \text{Weibull}(\alpha = 3, \beta = 3.7), \text{where} \;\; p = 0.15)$.

consider E$_{(2)}$ and E$_{(3)}$ because they can obtain a median for each set of each bootstrap sample. Tables 4.7 and 4.8 present the Type 1 error rates of the hypothesis test (4.3) with significance levels 0.10 and 0.05, respectively. The SB$_{exp}$ and SB$_{Lexp}$ methods mostly provide lower actual Type 1 error rates than the ones based on E$_{(2)}$ and E$_{(3)}$ at different sample sizes. However, E$_{(2)}$ and E$_{(3)}$ make the discrepancies between the actual and nominal Type 1 error levels lower, in particular when the sample size is small. When $n = 500$, all methods provide nearly identical results.

In the previous simulations, we created both samples in each trial from one scenario, but now we desire to have the samples from two different scenarios. In each trial, we set that the first sample is created from $T \sim \text{Log-Normal}(\mu = 0, \sigma = 1), C \sim \text{Weibull}(\alpha = 3, \beta = 3.7)$ and the second sample is from $T \sim \text{Weibull}(\alpha = 1, \beta = 1.443), C \sim \text{Exponential}(\lambda = 0.12), \text{where}\ p = 0.15$ in both scenarios. We want to investigate how the bootstrap methods perform when the two samples have

n	SB_{exp}	SB_{Lexp}	$\mathrm{E}_{(2)}$	$\mathrm{E}_{(3)}$
10	0.082	0.079	0.083	0.083
50	0.103	0.105	0.095	0.095
100	0.101	0.097	0.093	0.093
500	0.089	0.092	0.084	0.084

Table 4.9: Type 1 error rates with significance level $2\alpha = 0.10$, the first samples from $(T \sim \text{Log-Normal}(\mu = 0, \sigma = 1), C \sim \text{Weibull}(\alpha = 3, \beta = 3.7), \text{where} \ \ p = 0.15)$ and the second samples from $(T \sim \text{Weibull}(\alpha = 1, \beta = 1.443), C \sim \text{Exponential}(\lambda = 0.12), \text{where} \ \ p = 0.15)$.

n	SB_{exp}	SB_{Lexp}	$\mathrm{E}_{(2)}$	$\mathrm{E}_{(3)}$
10	0.030	0.027	0.038	0.038
50	0.046	0.047	0.046	0.046
100	0.041	0.043	0.034	0.034
500	0.045	0.047	0.043	0.043

Table 4.10: Type 1 error rates with significance level $2\alpha = 0.05$, the first samples from $(T \sim \text{Log-Normal}(\mu = 0, \sigma = 1), C \sim \text{Weibull}(\alpha = 3, \beta = 3.7), \text{where} \ \ p = 0.15)$ and the second samples from $(T \sim \text{Weibull}(\alpha = 1, \beta = 1.443), C \sim \text{Exponential}(\lambda = 0.12), \text{where} \ \ p = 0.15)$.

two different distributions, but they have the same median, which is equal to 1. Tables 4.9 and 4.10 outline the Type 1 error rates with significance levels 0.10 and 0.05, respectively. All methods perform well at all different sample sizes, and the results are nearly close to the nominal size 2α, in particular when the sample size is large.

4.4 Pearson correlation test

To compare the proposed bootstrap methods, presented in Chapter 3, to Efron's method, we compute the corresponding Type 1 error rate, where the method is considered superior if its corresponding Type 1 error rate is closer to the significance level 2α. In this section, we consider two different distributions to simulate the data sets. For the first scenario, we generate data sets from Gumbel copula, where the marginals X and Y are both following the standard uniform distribution. The second scenario is Clayton copula, where X follows normal distribution with mean 1 and standard deviation 1 and Y follows normal distribution as well, but with mean 5 and standard deviation 3. In both scenarios, we consider three dependence levels of ρ and three sample sizes with two significance levels. With the simulation results in the tables, we present the dependence parameters of copulas and their concordance measure Kendall's τ as there is a one-to-one relationship between them.

To compute Type 1 error rate for the null hypothesis of $\rho = \rho^{\star}$ based on a bootstrap method, we create $N = 1000$ data sets with sample size n and dependence level $\rho = \rho^{\star}$ from one scenario we presented above. Then for each generated data set, we apply each bootstrap method $B = 1000$ times, and compute the Pearson correlation of each bootstrap sample. We order the 1000 Pearson correlation values from lowest to highest and obtain the $100(1-2\alpha)\%$ bootstrap confidence interval. If the null hypothesis value is not included in the confidence interval, we reject H_0 and count 1; otherwise, we do not reject H_0 and count 0. The number of times that the null hypothesis was rejected over the 1000 trials will be the Type 1 error rate.

Table 4.11 presents the Type 1 error rates based on the bootstrap methods, where the significance level is 0.10. For small sample size, $n = 10$, the SBSP and SBNP methods, presented in Chapter 3, provide error rates closer to the nominal rate, = 0.10, compared to Efron's and the smoothed Efron's methods. However, the SBNP method is the best one when $\rho = 0.4$ and 0.8. When n increases to 50 and 100, all methods decrease the discrepancies between the actual and nominal error rates, but the SBNP method is the superior one in most cases.

With significance level 0.05, the actual Type 1 error rates based on the bootstrap methods are listed in Table 4.12. The SBSP and SBNP methods again provide lower discrepancies between the nominal and actual Type 1 error rates compared to Efron's and the smoothed Efron's methods, in particular when $n = 10$. When the sample size increases to 50 and 100, all methods perform better, but the SBNP method is the best one in most settings.

$n =$			10				50				100			
τ	θ	$H_0:$	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB
0	1	ρ=0	0.114	0.120	0.139	0.142	0.105	0.113	0.106	0.105	0.106	0.102	0.107	0.106
0.275	1.3793	ρ=0.4	0.137	0.129	0.147	0.149	0.136	0.122	0.128	0.127	0.129	0.105	0.109	0.106
0.610	2.5641	ρ=0.8	0.133	0.075	0.189	0.184	0.129	0.123	0.121	0.126	0.126	0.103	0.111	0.107

Table 4.11: Type 1 error rates with significance level 0.10, Gumbel copula, $X \sim \text{Unif}(0,1)$ and $Y \sim \text{Unif}(0,1)$.

$n =$			10				50				100			
τ	θ	$H_0:$	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB
0	1	ρ=0	0.064	0.072	0.085	0.081	0.046	0.051	0.053	0.058	0.056	0.052	0.055	0.057
0.275	1.3793	ρ=0.4	0.075	0.070	0.100	0.098	0.067	0.079	0.080	0.075	0.066	0.061	0.061	0.058
0.610	2.5641	ρ=0.8	0.079	0.034	0.131	0.127	0.074	0.070	0.078	0.076	0.080	0.066	0.071	0.071

Table 4.12: Type 1 error rates with significance level 0.05, Gumbel copula, $X \sim \text{Unif}(0,1)$ and $Y \sim \text{Unif}(0,1)$.

$n =$			10				50				100			
τ	θ	$H_0:$	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB
0	0	ρ=0	0.119	0.026	0.144	0.147	0.119	0.097	0.117	0.115	0.116	0.097	0.102	0.102
0.259	0.6990	ρ=0.4	0.142	0.039	0.167	0.165	0.150	0.102	0.122	0.125	0.135	0.114	0.116	0.119
0.630	3.4054	ρ=0.8	0.144	0.175	0.189	0.196	0.218	0.110	0.141	0.132	0.277	0.104	0.111	0.118

Table 4.13: Type 1 error rates with significance level 0.10, Clayton copula, $X \sim \text{Normal}(\mu = 1, \sigma = 1)$ and $Y \sim \text{Normal}(\mu = 5, \sigma = 3)$.

$n =$			10				50				100			
τ	θ	$H_0:$	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB
0	0	ρ=0	0.065	0.009	0.086	0.086	0.053	0.048	0.060	0.064	0.063	0.056	0.054	0.058
0.259	0.6990	ρ=0.4	0.088	0.012	0.108	0.105	0.076	0.048	0.066	0.068	0.083	0.066	0.063	0.070
0.630	3.4054	ρ=0.8	0.080	0.039	0.130	0.132	0.147	0.051	0.079	0.079	0.200	0.054	0.062	0.063

Table 4.14: Type 1 error rates with significance level 0.05, Clayton copula, $X \sim \text{Normal}(\mu = 1, \sigma = 1)$ and $Y \sim \text{Normal}(\mu = 5, \sigma = 3)$.

From the second scenario, we simulate $N = 1000$ data sets with dependence level $\rho = \rho^{\star}$, and we compute Type 1 error rates by using the bootstrap methods as shown in Tables 4.13 and 4.14. When $n = 10$, the SBSP method provides the closer results to the nominal error rates at most levels of ρ. Then as n increases to 50 and 100, its performance worsens for $H_0 : \rho = 0.8$ because the underlying distribution is not symmetric. At these large sample sizes, the results of SBNP, Efron and SEB methods are better than those of the SBSP method, in particular the results based on the SBNP method. The SBNP method mostly provides the lowest discrepancies between the nominal and actual error rates in comparison to those of the other methods in both significance levels 0.10 and 0.05. However, when $n = 10$ and $\rho = 0, 0.4$, the SBNP method provides very small error rates.

4.5 Kendall correlation test

In Section 4.4, we compute the Type 1 error rate for Pearson correlation test at different sample sizes and different dependence levels. For this section, we want to repeat the same comparisons, but we will consider the Kendall correlation test instead. We use the same scenarios to create data sets and conduct the simulations, where $n = 10, 50$ and 100, and the dependence level $\tau = 0, 0.4$ and 0.8 with significance levels 0.10 and 0.05.

We use Gumbel copula, where both marginals follow Uniform(0,1), to generate data sets and create Tables 4.15 and 4.16. From these tables, it is clear that the SBSP method performs well when $\tau = 0$ at all different sample sizes. However, it performs poorly as the sample size increases when $\tau = 0.4$ and 0.8. This is contrary to the results based on SBNP, Efron's and smoothed Efron's methods. These methods provide lower error rates than the nominal levels when the sample size is small at all different dependence levels. As n increases to 50 and 100, the error rates become closer to the nominal level 2α.

Tables 4.17 and 4.18 present the Type 1 error rates for Kendal correlation test at different dependence levels with significance levels 0.10 and 0.05, respectively. When $\tau = 0$ and $n = 10$, the error rate based on the SBNP method is extremely

$n =$			10				50				100			
τ	θ	$H_0:$	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB
0	1	$\tau = 0$	0.103	0.076	0.073	0.076	0.103	0.094	0.087	0.092	0.105	0.096	0.102	0.099
0.4	1.667	$\tau = 0.4$	0.120	0.059	0.078	0.065	0.133	0.107	0.110	0.100	0.128	0.100	0.100	0.098
0.8	5	$\tau = 0.8$	0.047	0.062	0.094	0.046	0.132	0.076	0.076	0.077	0.130	0.077	0.081	0.070

Table 4.15: Type 1 error rates of Kendall correlation test with significance level 0.10, Gumbel copula, $X \sim \text{Unif}(0,1)$ and $Y \sim \text{Unif}(0,1)$.

$n =$			10				50				100			
τ	θ	$H_0:$	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB
0	1	$\tau = 0$	0.057	0.035	0.037	0.040	0.048	0.041	0.041	0.045	0.055	0.047	0.049	0.052
0.4	1.667	$\tau = 0.4$	0.063	0.032	0.038	0.035	0.071	0.053	0.055	0.055	0.079	0.047	0.052	0.049
0.8	5	$\tau = 0.8$	0.021	0.025	0.021	0.025	0.072	0.039	0.032	0.037	0.068	0.043	0.038	0.042

Table 4.16: Type 1 error rates of Kendall correlation test with significance level 0.05, Gumbel copula, $X \sim \text{Unif}(0,1)$ and $Y \sim \text{Unif}(0,1)$.

$n =$			10				50				100			
τ	θ	$H_0:$	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB
0	0	0	0.103	0.020	0.073	0.076	0.103	0.087	0.087	0.092	0.105	0.095	0.102	0.099
0.4	1.333	0.4	0.125	0.037	0.089	0.074	0.140	0.094	0.101	0.099	0.121	0.090	0.098	0.089
0.8	8	0.8	0.049	0.918	0.110	0.046	0.165	0.456	0.078	0.080	0.160	0.169	0.088	0.094

Table 4.17: Type 1 error rates of Kendall correlation test with significance level 0.10, Clayton copula, $X \sim \text{Normal}(\mu = 1, \sigma = 1)$ and $Y \sim \text{Normal}(\mu = 5, \sigma = 3)$.

$n =$			10				50				100			
τ	θ	$H_0:$	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB	SBSP	SBNP	Efron	SEB
0	0	0	0.057	0.006	0.037	0.040	0.048	0.033	0.041	0.045	0.055	0.055	0.049	0.052
0.4	1.333	0.4	0.065	0.013	0.041	0.032	0.076	0.039	0.050	0.046	0.067	0.045	0.054	0.048
0.8	8	0.8	0.020	0.749	0.027	0.019	0.096	0.307	0.044	0.028	0.107	0.103	0.043	0.040

Table 4.18: Type 1 error rates of Kendall correlation test with significance level 0.05, Clayton copula, $X \sim \text{Normal}(\mu = 1, \sigma = 1)$ and $Y \sim \text{Normal}(\mu = 5, \sigma = 3)$.

lower than the nominal level 2α, while the results of other methods are close to the nominal levels. As the sample size increases to 50 and 100, all methods provide well outcomes. If there is a strong relation between the variables, it is advised to use either Efron's bootstrap method or the SEB method. They are both able to accomplish well results because these methods have much less effects than the SBSP and SBNP methods on the observations rank which is the base to compute Kendall correlation.

4.6 Concluding remarks

This chapter explored how the smoothed bootstrap methods, proposed for right-censored data and bivariate data, can be used to compute the Type 1 error rates for some hypothesis tests, and compare their results to those of Efron's bootstrap methods for right-censored data and bivariate data through simulations. For right-censored data, the null hypothesis tests are that the quartiles are equal to those of the underlying distributions. We also test whether two sample medians are equal regardless whether the two samples are from the same underlying distribution or not. For bivariate data, we compute the Type 1 error rates for Pearson and Kendall correlation tests.

We found, in the right-censored situation, the smoothed bootstrap method performs better when the sample size is small. It provides lower discrepancies between the actual and nominal error rates. As the sample size gets large, both bootstrap methods provide well results, but Efron's method mostly performs better for the third quartile. For two sample medians, we use the achieved significance level ASL to test whether the two samples have equal median or not. Through the simulations, it is clear that all bootstrap methods perform well, and the Type 1 error rates are close to the nominal levels.

For Pearson correlation test, the SBSP and SBNP methods lead to lower discrepancies between the actual and nominal Type 1 error rates compared to Efron's and the smoothed Efron's methods when the sample size is small. For large n, all methods provide good results, but the SBNP method performs better in most dependence levels. If the data distribution is asymmetric, the SBSP method does not perform that well, in particular when $\tau \neq 0$, which results from the Normal copula assumption.

For Kendall correlation test, it is advised to use either Efron's bootstrap method or the SEB method, in particular when the underlying distribution is asymmetric and has a strong Kendall correlation. Their influences on the observations rank are much less than those of the SBSP and SBNP methods. When $\tau = 0$ and the sample size is small, all bootstrap methods perform good, and as n gets large, their performances are improved and the Type 1 error rates become closer to the nominal level 2α.

In this chapter, we used the bootstrap methods for right-censored data and bivariate data to compute the Type 1 error rates for different hypothesis tests. It will be of interest to use these bootstrap methods to compute the power or Type 2 error rates for some hypothesis tests, this is left as an interesting topic for future research.

Chapter 5

Conclusions

This chapter summarizes the main results in this thesis, and some topics are discussed for future research. For right-censored data, we introduced a new smoothed bootstrap method, which has been used for inferences in terms of coverage probabilities, survival functions and Type 1 error rates of hypothesis tests. We also presented three new smoothed bootstrap methods for bivariate data and studied the coverage proportions for some functions of interest. Moreover, we used these methods to derive the Type 1 error rates of Pearson and Kendall correlation tests.

In Chapter 2, we presented the smoothed bootstrap method for right-censored data on a finite support and on a positive real line. This bootstrap method allows us to sample from the whole data range to create the bootstrap samples and avoids the issues of ties and censored observations. This method has more variation in sampling than the one of Efron, and this mostly leads to better accuracy in the coverage probabilities, in particular when the sample size is small or medium. For survival function inferences, the smoothed bootstrap method performs better than the alternative smoothed bootstrap method in terms of the coverage probability for a survival function at a certain time t. For future research, it will be interesting to use the smoothed bootstrap method to derive the Type 2 error rates of hypothesis tests. The asymptotic behaviour of smoothed bootstrap method is also of interest, the study of this is left as a topic for future research. We expect that, for large sample sizes, the behaviour of our method will be similar to that of Efron's method because the difference between sampling from intervals between observations or sampling from observations becomes less when the number of observations increases.

Chapter 3 introduced three smoothed bootstrap methods for bivariate data, and these methods were compared to Efron's bootstrap method through simulation studies. The first two methods are based on the semi-parametric and nonparametric predictive methods, proposed by Coolen-Maturi *et al.* [22] and Muhammad *et al.* [65], and the third one is based on the uniform kernels. These smoothed methods avoid the issue of ties that often occurred in the bootstrap samples conducted by Efron's method, and this leads to better accuracy in terms of the coverage probability for some functions of interest. If the data distribution is symmetric or approximately symmetric, the smoothed bootstrap methods perform better than Efron's bootstrap method when we are interested in the Pearson correlation and the means of T_1 and T_2. However, Efron's bootstrap performs better if we consider the Kendall and Spearman correlations. In case of no symmetry is observed for the data distribution, it is advised to use either Efron's bootstrap or the smoothed Efron's method because they are more likely to provide better results. For further research, it will be interesting to derive a smoothed bootstrap method for bivariate data, where one or both marginals include right-censored observations. This could be based on a combination of the right-censoring $A_{(n)}$ assumption and a parametric or nonparametric copula model. As a starting point to derive the method, we can attempt to the right-censoring $A_{(n)}$ assumption on the marginals first, then to take the dependence structure into account, we use copulas. This will provide a partial distribution for one future bivariate observation and this partial distribution will be used to derive a smoothed bootstrap method for bivariate right-censored data. We expect that the $(n+1)^2$ blocks partitioned the sample space will be overlapping due to the censored observations and challenges may be experienced, so we reckon some heuristic method will be needed to link copulas to the NPI marginals in this case. Also as a topic for future research, it is of interest to study the asymptotic behaviour of the smoothed bootstrap methods as the sample size increases.

In Chapter 4, the proposed bootstrap methods were used as alternative methods for some hypothesis tests. Simulations were used to study their performance in deriving the Type 1 error rates. When the data include right-censored observations, we found that the smoothed bootstrap provides smaller discrepancies between the actual and nominal error rates of the quartiles' hypothesis tests compared to Efron's method, in particular when the sample size is small. For the two sample median test, the smoothed bootstrap method and Efron's method both provide good results. For Pearson correlation test at any dependence level, the smoothed bootstrap based on the nonparametric predictive method mostly provides lower discrepancies between the actual and nominal Type 1 error rates. For Kendall's correlation test, it is advised to use either Efron's bootstrap or the smoothed Efron's method. For future research, the proposed bootstrap methods will be used to compute either the Type 2 error rate or the power of such hypothesis tests and compared the results to the ones of Efron's methods. The evaluation of these statistics is important and useful.

The proposed methods can be used or introduced to the field of system reliability, where for example for a coherent system, the standard bootstrap has been used to provide the lifetime distribution. Marks *et al.* [61] presented the resampling bootstrap method to find the empirical system lifetime distributions of parallel and series systems. In reliability applications, our approaches can be used to find the empirical system lifetime distributions of any system, and inferences can be made with high accuracy. This is left as a future research topic in system reliability.

Bootstrap methods have been widely used in many statistical fields for precise inferences due to their simplicity and efficiency to give good estimates. The bootstrap, in general, can be used in many statistical situations. For future research, we want to use it for imprecise inferences, and this could be of interest in practical use to provide a range of estimates. Instead of providing one single value as an estimation for a function of interest, the bootstrap methods could lead to an imprecise bootstrap estimate. In other words, the imprecise bootstrap for any function of interest can replace a single estimate with an interval of estimations including lower and upper bounds. As this will be the first research into developing bootstrap methods for imprecise inferences, there are many related research opportunities and challenges.

Tukey [82] introduced a nonparametric estimation technique referred to as statistically equivalent blocks and tolerance regions. This technique uses the original data points to divide the sample space into $n+1$ blocks, where each block is assigned with probability $1/(n+1)$. Based on these blocks, it could be possible to introduce a new smoothed bootstrap method for any data type; univariate, bivariate or multivariate data. This is left as a future research topic. Also as a topic for future research, it is of interest to explore the asymptotic behaviour of the suggested bootstrap methods as both the sample size n and the number of variables increase. It is expected that smoothing will be advantageous as the number of variables increases, but there may be a need for a very large sample size n in order to be able to do meaningful nonparametric inference in higher dimensional spaces.

Appendix A

Comparing the smoothed bootstrap to Efron's bootstrap for right-censored data

A.1 Simulations for the quartiles

Table A.1 presents the chi-squared values obtained from the actual coverage probabilities based on the bootstrap methods for the true median where the censoring proportion p in the generating data sets is 0.30. The smoothed bootstrap method performs better than Efron's method when $n = 6, 10, 20, 40$, but Efron's method with its three modifications accomplishes better results when $n = 100$. The three modifications have an identical chi-squared value when $n = 100$ because we found a median at each bootstrap sample; the number NA is equal to 0. Due to increasing the censoring proportion to 0.30, The NA and ABS numbers increase as well, and they are greater than the results in Table 2.4 where the censoring proportion p is set equal to 0.15.

Table A.2 presents the chi-squared values obtained from the actual coverage probabilities for Q_1, where the censoring proportion p is 0.30. The smoothed bootstrap method mostly provides better results than those of Efron's method, in particular when the sample size is small or medium. When $n = 6, 10, 20$ and 40, the smoothed bootstrap method achieves better coverage probabilities than Efron's method in

		10 CR				20 CR			
n	measures	SB	$E_{(1)}$	$E_{(2)}$	$E_{(3)}$	SB	$E_{(1)}$	$E_{(2)}$	$E_{(3)}$
6	χ^2	43.38	163.72	127.52	132.68	63.84	179.16	152.92	146.04
	NA	—	41629	0	0	—	41629	0	0
	ABS	—	4582	4582	4582	—	4582	4582	4582
	P-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
10	χ^2	9.34	58.60	52.40	62.48	23.12	69.64	65.56	73.60
	NA	—	20044	0	0	—	20044	0	0
	ABS	—	175	175	175	—	175	175	175
	P-value	0.406	0.000	0.000	0.000	0.232	0.000	0.000	0.000
20	χ^2	6.22	7.80	7.46	6.48	13.32	17.88	17.56	16.72
	NA	—	3096	0	0	—	3096	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.718	0.554	0.589	0.691	0.822	0.530	0.552	0.609
40	χ^2	11.32	11.38	11.10	11.10	22.84	30.60	30.64	30.64
	NA	—	37	0	0	—	37	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.254	0.251	0.269	0.269	0.244	0.045	0.044	0.044
100	χ^2	21.40	5.86	5.86	5.86	26.40	18.52	18.52	18.52
	NA	—	0	0	0	—	0	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.011	0.754	0.754	0.754	0.119	0.488	0.488	0.488

Table A.1: The chi-squared values for $Q_2 = 0.236$, and their *P-values* with the corresponding NA and ABS numbers.

both divisions, but Efron's method performs better when $n = 100$. The smoothed bootstrap method is still good to provide reasonable coverage probabilities at this sample size, but the three options of Efron's bootstrap method are better.

For the third quartile Q_3, Table A.3 presents the chi-squared values obtained from the actual coverage probabilities. The smoothed bootstrap method has a better accuracy in terms of defining the actual coverage probabilities when $n = 6, 10$ and 20 in both divisions. However, it does not achieve the same level of performance when the sample size increases to 40 and 100. When $n = 40, 100$, Efron's method with its three options shows its superiority providing lower chi-squared values, in particular when the third option is assumed. The third suggested option for Efron's

		10 CR				20 CR			
n	measures	SB	$\mathrm{E}_{(1)}$	$\mathrm{E}_{(2)}$	$\mathrm{E}_{(3)}$	SB	$\mathrm{E}_{(1)}$	$\mathrm{E}_{(2)}$	$\mathrm{E}_{(3)}$
6	χ^2	24.80	1049.84	989.94	979.14	96.32	2139.52	1985.40	2015.48
	NA	—	13616	0	0	—	13616	0	0
	ABS	—	4582	4582	4582	—	4582	4582	4582
	P-value	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000
10	χ^2	17.54	775.76	746.30	759.66	64.88	1318.44	1291.76	1308.04
	NA	—	2279	0	0	—	2279	0	0
	ABS	—	175	175	175	—	175	175	175
	P-value	0.041	0.000	0.000	0.000	0.000	0.000	0.000	0.000
20	χ^2	13.08	122.50	122.50	122.50	39.76	430.88	430.88	430.88
	NA	—	38	0	0	—	38	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.159	0.000	0.000	0.000	0.004	0.000	0.000	0.000
40	χ^2	10.58	91.82	91.82	91.82	24.40	117.16	117.16	117.16
	NA	—	0	0	0	—	0	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.306	0.000	0.000	0.000	0.181	0.000	0.000	0.000
100	χ^2	14.54	6.10	6.10	6.10	31.52	19.32	19.32	19.32
	NA	—	0	0	0	—	0	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.104	0.730	0.730	0.730	0.035	0.436	0.436	0.436

Table A.2: The chi-squared values for $Q_1 = 0.117$, and their *P-values* with the corresponding NA and ABS numbers.

method, $\mathrm{E}_{(3)}$, clearly has better coverage probabilities in comparison to the other two options, $\mathrm{E}_{(1)}$ and $\mathrm{E}_{(2)}$, in all different sample sizes and in both divisions. More variations are occurred in the chi-squared values corresponding to the three options of Efron's method if the NA number is large, and this can be shown from Table A.3. When $n = 6$, the NA number is 140875 and the chi-square values of $\mathrm{E}_{(1)}$, $\mathrm{E}_{(2)}$ and $\mathrm{E}_{(3)}$ are $1318.90, 1290.20$ and 500.02, respectively. The three options have about the same chi-squared value when $n = 100$ because the NA number is only 163 out of 1000000, which is small.

		10 CR				20 CR			
n	measures	SB	$E_{(1)}$	$E_{(2)}$	$E_{(3)}$	SB	$E_{(1)}$	$E_{(2)}$	$E_{(3)}$
6	χ^2	147.28	1318.90	1290.20	500.02	198.72	2710.36	2991.12	1221.84
	NA	—	140875	0	0	—	140875	0	0
	ABS	—	4582	4582	4582	—	4582	4582	4582
	P-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
10	χ^2	41.88	481.80	461.30	141.02	57.28	1052.12	1123.68	415.52
	NA	—	101223	0	0	—	101223	0	0
	ABS	—	175	175	175	—	175	175	175
	P-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
20	χ^2	7.28	74.88	64.74	22.20	16.52	184.68	184.40	74.88
	NA	—	51700	0	0	—	51700	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.608	0.000	0.000	0.008	0.622	0.000	0.000	0.000
40	χ^2	38.70	12.80	12.44	10.44	60.80	40.44	40.68	33.80
	NA	—	11572	0	0	—	11572	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.000	0.172	0.190	0.316	0.000	0.003	0.003	0.019
100	χ^2	49.18	12.54	12.42	12.42	71.20	28.64	28.20	28.20
	NA	—	163	0	0	—	163	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.000	0.185	0.191	0.191	0.000	0.072	0.080	0.080

Table A.3: The chi-squared values for $Q_3 = 0.396$, and their *P-values* with the corresponding NA and ABS numbers.

A.2 Simulations for $q_{0.10}$ and $q_{0.90}$

		10 CR				20 CR			
n	measures	SB	$E_{(1)}$	$E_{(2)}$	$E_{(3)}$	SB	$E_{(1)}$	$E_{(2)}$	$E_{(3)}$
6	χ^2	59.86	3092.98	3092.98	3092.98	374.96	6832.68	6832.68	6832.68
	NA	—	580	0	0	—	580	0	0
	ABS	—	699	699	699	—	699	699	699
	P-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
10	χ^2	90.20	1146.48	1146.48	1146.48	235.56	2680.56	2680.56	2680.56
	NA	—	15	0	0	—	15	0	0
	ABS	—	14	14	14	—	14	14	14
	P-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
20	χ^2	19.20	149.52	149.52	149.52	107.56	462.60	462.60	462.60
	NA	—	0	0	0	—	0	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.024	0.000	0.000	0.000	0.000	0.000	0.000	0.000
40	χ^2	22.08	176.42	176.42	176.42	73.44	441.44	441.44	441.44
	NA	—	0	0	0	—	0	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.009	0.000	0.000	0.000	0.000	0.000	0.000	0.000
100	χ^2	5.40	68.24	68.24	68.24	35.96	400.24	400.24	400.24
	NA	—	0	0	0	—	0	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.798	0.000	0.000	0.000	0.011	0.000	0.000	0.000

Table A.4: The chi-squared values for $q_{0.10} = 0.2231$, and their *P-values* with the corresponding NA and ABS numbers.

Table A.4 outlines the chi-squared values obtained from the coverage probabilities for $q_{0.10} = 0.2231$. The smoothed bootstrap (SB) approach apparently achieves a better accuracy than Efron's method with its three options at all different sample sizes in both divisions. In other words, the SB method provides lower discrepancy between the nominal and estimated coverage probabilities. The three options of Efron's method are equal because we get a time t met the condition $\hat{S}(t) = 0.90$ in most of the bootstrap samples. The NA number goes down as the sample size increases. The NA numbers are 580 and 15 when $n = 6$ and 10, respectively, then it gets 0 with large sample sizes. As the sample size increases, Efron' chi-squared values generally decrease as they do for the quartiles in Tables 2.9, 2.10 and 2.11.

		10 CR				20 CR			
n	measures	SB	$E_{(1)}$	$E_{(2)}$	$E_{(3)}$	SB	$E_{(1)}$	$E_{(2)}$	$E_{(3)}$
6	χ^2	50.32	3379.08	3474.22	2854.48	70.96	7282.16	7615.24	6288.72
	NA	—	58470	0	0	—	58470	0	0
	ABS	—	699	699	699	—	699	699	699
	P-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
10	χ^2	9.92	1652.02	1661.78	1205.94	30.52	3591.60	3727.68	2750.52
	NA	—	46573	0	0	—	46573	0	0
	ABS	—	14	14	14	—	14	14	14
	P-value	0.357	0.000	0.000	0.000	0.046	0.000	0.000	0.000
20	χ^2	7.68	432.74	419.80	333.92	48.64	1223.04	1196.24	968.08
	NA	—	32670	0	0	—	32670	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.567	0.000	0.000	0.000	0.000	0.000	0.000	0.000
40	χ^2	13.68	71.78	72.48	71.94	42.16	244.04	238.52	235.88
	NA	—	6160	0	0	—	6160	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.134	0.000	0.000	0.000	0.002	0.000	0.000	0.000
100	χ^2	19.10	8.24	8.16	8.16	34.96	26.40	26.24	26.24
	NA	—	100	0	0	—	100	0	0
	ABS	—	0	0	0	—	0	0	0
	P-value	0.024	0.510	0.518	0.518	0.014	0.119	0.124	0.124

Table A.5: The chi-squared values for $q_{0.90} = 1.7437$, and their *P-values* with the corresponding NA and ABS numbers.

In Table A.5, the chi-squared values obtained from the coverage probabilities for $q_{0.90} = 1.7437$ are outlined. When $n = 6, 10, 20$ and 40 and regardless of the number of confidence regions, the SB method distinctly performs a better coverage accuracy, and it makes the discrepancy between nominal and estimated coverage probabilities lower than Efron's method does. The SB method is still performing well when $n = 100$, but Efron's method provides better results. This could be from the influence of Exponential tail(s) assumed for the end interval(s) when the SB method is applied. By observing the pattern of chi-squared values as the sample size increases, the chi-squared values of Efron's bootstrap method become smaller while there is no pattern with those of the smoothed bootstrap method, and this is another influence from the Exponential tail(s) assumed for the last interval(s).

Appendix B

R codes

B.1 The smoothed bootstrap method (SB)

The R codes to compute M-function values are written by Maturi [62].

```
# The functions to compute the m-function values

X.c <- function(X) {
# to get the censored data
ifelse(length(X[X[, 2] == 0,]) > 2, x1 <-
X[X[, 2] == 0,][, 1], x1 <- X[X[, 2] == 0,][1])
return(x1)
}




X.u <- function(X) {
# to get the failure data
ifelse(sum(X[, 2] == 1) == 1, x1 <- X[X[, 2] == 1,][1],
x1 <- X[X[, 2] == 1,][, 1])
return(x1)
}
```

```
Xu1 <- function(X) {
# all censored, no failure occurs
ifelse(sum(X[, 2] == 1) == 0, Y <- 1, Y <- c(X.u(X),
1))
return(Y)
}



Xt0 <- function(X) {
Y <- c(0, X[, 1])
return(Y)
}



cond <- function(X, y) {
P1 <- NULL
n <- nrow(X)
Xc <- X.c(X)
ncc <-
function(X, cr) {
# calculate the term in the product term
(sum(X[, 1] >= cr) + 1) / sum(X[, 1] >= cr)
}
cr.obs <- Xc[Xc < y]
n.cr.obs <-
length(cr.obs) # calculate the condition under the product term
if (n.cr.obs == 0 | sum(X[, 2] == 0) == 0)
{
P1 <- 1
}
else{
for (j in 1:n.cr.obs) {
P1[j] <- ncc(X, cr.obs[j])
}
}
P3 <- prod(P1) / (n + 1)
return(P3)
```

```
}


# calculate Mfun and prob

Mfun <- function(X) {
Y <- rbind(c(0, 1), X)
ny <- nrow(Y)
if (sum(Y[, 2]) == ny) {
Mu = rep(1 / ny, ny)
} else{
Mu <- NULL
for (i in 1:ny) {
Mu[i] <-
(sum(X[, 1] >= Y[, 1][i])) ^ (Y[, 2][i] - 1) * cond(X, Y[, 1][i])
}
}
return(Mu)
}


# Create a data set with right-censored observations

n = 10

T = rbeta(n, 1.2, 3.2)
C = runif(n, 0, 1.82)

time = ifelse(T <= C, T, C)
cens = ifelse(T <= C, 1, 0)

dataa1 = data.frame(time, cens)
dataa = as.matrix(dataa1[order(dataa1$time), ])

mfun = Mfun(dataa)
U = Xu1(dataa)
L = Xt0(dataa)
```

```
lower = NULL
upper = NULL

for (r in 1:(length(T) + 1)) {
lower[r] = L[r]
upper[r] = U[U > L[r]][1]
}

mdaataa = data.frame(lower, upper, mfun)
daataa = as.matrix(mdaataa)

median = NULL

# Create 1000 bootstrap samples and compute their corresponding
   medians

for (a in 1:1000) {
mysample = daataa[sample(length(T) + 1, length(T) , replace = TRUE,
    prob =
mfun), ]

obs = NULL


for (i in 1:length(T)) {
obs[i] = runif(1, min = mysample[i, 1], max = mysample[i, 2])

}

median[a] = quantile(obs, prob = 0.5)

}
```

B.2 The smoothed bootstrap method based on the semi-parametric predictive method (SBSP)

Muhammad [64] provided the R codes to compute the probabilities h_{ij}.

```r
library(copula)
library(CDVine)
library(VineCopula)


fun1 = function(i, j) {
pMvdc(c(i / (n + 1), j / (n + 1)), mycopula)
}

ff3 <- function(XY) {
rXY <- apply(XY, 2, rank)
gg <- NULL
ss <- NULL
for (i in 1:(n + 1)) {
for (j in 1:(n + 1)) {
a <- rXY[i, 1]
b <- rXY[j, 2]
ff <- NULL
if (a > 1 &
b > 1)
ff <- fun1(a, b) - fun1(a - 1, b) - fun1(a, b - 1) + fun1(a - 1, b
   - 1)
if (a > 1 & b == 1)
ff <- fun1(a, b) - fun1(a - 1, b)
if (a == 1 & b > 1)
ff <- fun1(a, b) - fun1(a, b - 1)
if (a == 1 & b == 1)
ff <- fun1(a, b)
gg <- rbind(gg, c(XY[i, 1], XY[j, 2], a, b, ff))
}
}
return(gg)
}
```

```
#######################################################
# To set the dependence level
tau = 0

theta = 1 / (1 - tau)
theta

tau = 1 - (1 / theta)
tau

#######################################################

gumbel.cop = gumbelCopula(theta)
xyz = mvdc(gumbel.cop, c("unif", "unif"),
list(list(min = 0, max = 1), list(min = 0, max = 1)))
n = 10

data = rMvdc(n, xyz)

x = data[, c(1)]
y = data[, c(2)]
data = data.frame(x, y)

#######################################################
#To estimate parametric copula, using the data

eu = cbind((rank(x) / (n + 1)), (rank(y) / (n + 1)))
u = eu[, 1]
v = eu[, 2]

# The inversion of Kendall's tau estimate
fit2 = BiCopEst(u, v, family = 1, method = "itau")$par

#######################################################
#Using estimated theta
```

```
mycopula = mvdc(normalCopula(fit2), c("unif", "unif"),
list(list(min = 0, max = 1), list(min = 0, max = 1)))

#######################################################

xyU = rbind(data, c(1, 1)) #to add the Upper Bound
xyL = rbind(c(0, 0), data)  #to add the lower Bound
xyL = xyL[order(xyL[, 1], xyL[, 2]),]
xyU = xyU[order(xyU[, 1], xyU[, 2]),]

RR = ff3(xyU)
order_hij = RR[order(RR[, 3], RR[, 4]), ]

limitx = c(0, x[order(x)])
orderLX = rep(limitx, n + 1)[order(rep(limitx, n + 1))]

limity = c(0, y[order(y)])
orderLY = rep(limity, n + 1)

order_data = data.frame(orderLX,
order_hij[, 1],
orderLY,
order_hij[, 2],
order_hij[, 3],
order_hij[, 4],
order_hij[, 5])

colnames(order_data) = c("LX", "UX", "LY", "UY", "rX", "rY", "hij")
Or_data = order_data

#######################################################
#######################################################
# Applying the SBSP method 1000 times

Pearson = NULL
```

```
probability = Or_data[, 7]

for (aa in 1:1000) {
mysample = Or_data[sample(nrow(Or_data),
length(x) ,
replace = TRUE,
prob = probability), ]

for_X = NULL
for_Y = NULL

for (wre in 1:length(x)) {
divint1 = runif(1, min = mysample[wre, 1], max = mysample[wre, 2])
divint2 = runif(1, min = mysample[wre, 3], max = mysample[wre, 4])

for_X[wre] = divint1
for_Y[wre] = divint2

}

Boot_data = data.frame(for_X, for_Y)
Pearson[aa] = cor(Boot_data, method = c("pearson"))[1, 2]

}
```

B.3 The smoothed bootstrap method based on the nonparametric predictive method (SBNP)

Muhammad [64] provided the R codes to compute the probabilities h_{ij}.

```
library(gtools)
library(np)

tau = 0

theta = 1 / (1 - tau)
```

```
theta

tau = 1 - (1 / theta)
tau

#######################################################

gumbel.cop = gumbelCopula(theta)
xyz = mvdc(gumbel.cop, c("unif", "unif"),
list(list(min = 0, max = 1), list(min = 0, max = 1)))
n = 3

data = rMvdc(n, xyz)

x = data[, c(1)]
y = data[, c(2)]

data = data.frame(x, y)

h_x = 1.06 * min(sd(x), (IQR(x) / 1.349)) * (n ^ (-1))
h_y = 1.06 * min(sd(y), (IQR(y) / 1.349)) * (n ^ (-1))
xy = cbind(x, y)
n = nrow(xy)

#######################################################
#make a data frame (matrix)
mydat2 = data.frame(x = xy[, 1], y = xy[, 2])
mydat = apply(mydat2, 2, rank) / (n + 1)
mydat = data.frame(x = mydat[, 1], y = mydat[, 2])

#include upper and lower limit
mydat1 = rbind(c(0, 0), mydat, c(1, 1))
n1 = nrow(mydat1)

grid.dat = mydat1

#######################################################
```

```
#######################################################
## Estimate the copula

bw = npudistbw(
~ x + y,
data = mydat,
bws = c(h_x, h_y),
bandwidth.compute = FALSE,
ckertype = "gaussian"
)
copula = npcopula(bws = bw, data = mydat, u = grid.dat)

#######################################################
copulaa = copula[order(copula[, 2]), ]

HIJ = matrix(copulaa[, 1], n1, n1, byrow = T)
nn = nrow(HIJ)
HIJorder = order(HIJ[nn, ])
HIJO = HIJ[, HIJorder]

#######################################################
# For Hij

BB = HIJO[-1, -1]
n2 = nrow(BB)

#######################################################
# For hij

D0 = permutations(n2, 2, 1:n2, repeats = TRUE, set = FALSE)
D1 = D0 / n2
B0 = rep(0, n2)
BB2 = cbind(B0, BB[, -n2], deparse.level = 0)
BB3 = rbind(B0, BB[-n2, ], deparse.level = 0)
BB4 = cbind(B0, rbind(B0[-1], BB[-n2, -n2]), deparse.level = 0)
H1 = c(BB + BB4 - BB2 - BB3)
h_ij = H1
```

```
########################################
xyU = rbind(data, c(1, 1)) #to add the Upper Bound
xyL = rbind(c(0, 0), data)  #to add the lower Bound

xyL = xyL[order(xyL[, 2], xyL[, 1]),]
xyU = xyU[order(xyU[, 2], xyU[, 1]),]
limity = c(0, y[order(y)])
orderLY = rep(limity, n + 1)[order(rep(limity, n + 1))]
limityU = c(y[order(y)], 1)
orderLYU = rep(limityU, n + 1)[order(rep(limityU, n + 1))]
limitx = c(0, x[order(x)])
orderLX = rep(limitx, n + 1)
limitxU = c(x[order(x)], 1)
orderLXU = rep(limitxU, n + 1)

order_data = data.frame(orderLX, orderLXU, orderLY, orderLYU, h_ij)
colnames(order_data) = c("LX", "UX", "LY", "UY", "hij")
Or_data = order_data

#######################################################
#######################################################
# Applying the generalizing Banks' bootstrap for bivariate data:
Pearson = NULL

for (aa in 1:1000) {
mysample = Or_data[sample(nrow(Or_data),
length(x)
,
replace = TRUE,
prob = h_ij), ]

for_X = NULL
for_Y = NULL

for (wre in 1:length(x)) {
divint1 = runif(1, min = mysample[wre, 1], max = mysample[wre, 2])
```

```
divint2 = runif(1, min = mysample[wre, 3], max = mysample[wre, 4])

for_X[wre] = divint1
for_Y[wre] = divint2
}
Boot_data = data.frame(for_X, for_Y)
Pearson[aa] = cor(Boot_data, method = c("pearson"))[1, 2]
}
```

B.4 The smoothed Efron's bootstrap (SEB)

```
x = c(1, 2, 3)
y = c(3, 1, 2)
n = length(x)
data = data.frame(x, y)

h_x = 1.06 * min(sd(x), (IQR(x) / 1.349)) * (n ^ (-4))
h_y = 1.06 * min(sd(y), (IQR(y) / 1.349)) * (n ^ (-4))

LX = x - (h_x / 2)
UX = x + (h_x / 2)
LY = y - (h_y / 2)
UY = y + (h_y / 2)
Or_data = data.frame(LX, UX, LY, UY)
#######################################################
# Applying the SEB method 1000 times:

Pearson = NULL
for (aa in 1:1000) {
mysample = Or_data[sample(nrow(Or_data), length(x) , replace = TRUE
   ), ]

for_X = NULL
for_Y = NULL
for (wre in 1:length(x)) {
divint1 = runif(1, min = mysample[wre, 1], max = mysample[wre, 2])
divint2 = runif(1, min = mysample[wre, 3], max = mysample[wre, 4])
```

```
for_X[wre] = divint1
for_Y[wre] = divint2
}
Boot_data = data.frame(for_X, for_Y)
Pearson[aa] = cor(Boot_data, method = c("pearson"))[1, 2]
}
```

Bibliography

[1] Al Luhayb A.S.M., Coolen F.P.A. and Coolen-Maturi T. (2019). Generalizing Banks smoothed bootstrap method for right-censored data. pp. 894–901. 29th European Safety and Reliability Conference (ESREL 2019), Hannover (Germany).

[2] Al Luhayb A.S.M., Coolen F.P.A. and Coolen-Maturi T. (2019). Smoothed bootstrap for survival function inference. pp. 297–304. Proceedings of the International Conference on Information and Digital Technologies (IDT 2019), Zilina (Slovakia).

[3] Augustin T. and Coolen F.P.A. (2004). Nonparametric predictive inference and interval probability. *Journal of Statistical Planning and Inference*, 124(2), 251–272.

[4] Augustin T., Coolen F.P.A., De Cooman G. and Troffaes M.C.M. (2014). *Introduction to imprecise probabilities.* John Wiley & Sons.

[5] Banks D.L. (1988). Histospline smoothing the Bayesian bootstrap. *Biometrika*, 75, 673–684.

[6] Bartholomew D.J. (1957). A problem in life testing. *Journal of the American Statistical Association*, 52, 350–355.

[7] Berliner L.M. and Hill B.M. (1988). Bayesian nonparametric survival analysis. *Journal of the American Statistical Association*, 83, 772–779.

[8] Berrar D. (2019). Introduction to the non-parametric bootstrap. In *Encyclope-*

dia of bioinformatics and computational biology, pp. 766–773. Academic Press Oxford.

[9] Bilker W.B. and Wang M.C. (1997). Bootstrapping left truncated and right censored data. *Communications in Statistics-Simulation and Computation*, 26(1), 141–171.

[10] BinHimd S. (2014). *Nonparametric Predictive Methods for Bootstrap and Test Reproducibility.* Ph.D. thesis, Durham University, UK. Https://npi-statistics.com/pdfs/theses/SB14.pdf.

[11] Binhimd S. and Coolen F.P.A. (2012). On boostrapping using nonparametric predictive inference. pp. 465–468. The 1st ISM International Statistical Conference (ISM 2012), Malaysia.

[12] Borgan Ø. and Liestøl K. (1990). A note on confidence intervals and bands for the survival function based on transformations. *Scandinavian Journal of Statistics*, 17, 35–41.

[13] Brown B.W., Hollander M. and Korwar R.M. (1974). Nonparametric tests of independence for censored data with applications to heart transplant studies. *Reliability and Biometry*, pp. 327–354. Proschan, F. and Serfling, R. J. (Eds), SIAM, Philadelphia.

[14] Chen D. and Fritz M.S. (2021). Comparing Alternative Corrections for Bias in the Bias-Corrected Bootstrap Test of Mediation. *Evaluation & the Health Professions*, pp. 1–12. Https://doi.org/10.1177

[15] Cherubini U., Luciano E. and Vecchiato W. (2004). *Copula methods in finance.* New York: John Wiley & Sons.

[16] Coolen F.P.A. (1998). Low structure imprecise predictive inference for Bayes' problem. *Statistics & Probability Letters*, 36(4), 349–357.

[17] Coolen F.P.A. (2006). On nonparametric predictive inference and objective Bayesianism. *Journal of Logic, Language and Information*, 15(1-2), 21–47.

[18] Coolen F.P.A. (2011). Nonparametric predictive inference. In *International Encyclopedia of Statistical Sciences*, (Editor) M. Lovric, pp. 968–970. Berlin: Springer.

[19] Coolen F.P.A. and Elsaeiti M.A. (2009). Nonparametric predictive methods for acceptance sampling. *Journal of Statistical Theory and Practice*, 3(4), 907–921.

[20] Coolen F.P.A., Troffaes M.C.M. and Augustin T. (2011). Imprecise Probabilities. In *International Encyclopedia of Statistical Sciences*, (Editor) M. Lovric, pp. 645–648. Berlin: Springer.

[21] Coolen F.P.A. and Yan K.J. (2004). Nonparametric predictive inference with right-censored data. *Journal of Statistical Planning and Inference*, 126, 25–54.

[22] Coolen-Maturi T., Coolen F.P.A. and Muhammad N. (2016). Predictive inference for bivariate data: Combining nonparametric predictive inference for marginals with an estimated copula. *Journal of Statistical Theory and Practice*, 10(3), 515–538.

[23] Coolen-Maturi T., Coolen-Schrijner P. and Coolen F.P.A. (2012). Nonparametric predictive inference for binary diagnostic tests. *Journal of Statistical Theory and Practice*, 6(4), 665–680.

[24] Coolen-Schrijner P. and Coolen F.P.A. (2004). Adaptive age replacement strategies based on nonparametric predictive inference. *Journal of the Operational Research Society*, 55(12), 1281–1297.

[25] Coolen-Schrijner P., Maturi T.A. and Coolen F.P.A. (2009). Nonparametric predictive precedence testing for two groups. *Journal of Statistical Theory and Practice*, 3(1), 273–287.

[26] Davison A.C. and Hinkley D.V. (1997). *Bootstrap Methods and Their Application.* Cambridge University Press.

[27] De Angelis D. and Young G.A. (1992). Smoothing the bootstrap. *International Statistical Review/Revue Internationale de Statistique*, 60, 45–56.

[28] Dolker M., Halperin S. and Divgi D.R. (1982). Problems with bootstrapping Pearson correlations in very small bivariate samples. *Psychometrika*, 47(4), 529–530.

[29] Efron B. (1967). The two sample problem with censored data. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 4, pp. 831–853.

[30] Efron B. (1979). Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, 7, 1–26.

[31] Efron B. (1981). Censored data and the bootstrap. *Journal of the American Statistical Association*, 76, 312–319.

[32] Efron B. (1982). *The jackknife, the bootstrap, and other resampling plans*, volume 38. Society for Industrial and Applied Mathematics, Philadelphia.

[33] Efron B. and Tibshirani R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical science*, 1, 54–77.

[34] Efron B. and Tibshirani R.J. (1993). *An Introduction to the Bootstrap*. Chapman & Hall.

[35] Elkhafifi F.F. (2012). *Nonparametric Predictive Inference for Ordinal Data and Accuracy of Diagnostic Tests*. Ph.D. thesis, Durham University, UK. Https://npi-statistics.com/pdfs/theses/FE12.pdf.

[36] Elkhafifi F.F. and Coolen F.P.A. (2012). Nonparametric predictive inference for accuracy of ordinal diagnostic tests. *Journal of Statistical Theory and Practice*, 6(4), 681–697.

[37] Elsaeiti M.A. and Coolen F.P.A. (2012). A nonparametric predictive approach to sequential acceptance problems. *Journal of Statistical Theory and Practice*, 6(3), 383–401.

[38] Freireich E.J., Gehan E., Frei E., Schroeder L.R., Wolman I.J., Anbari R., Burgert E.O., Mills S.D., Pinkel D. *et al.* (1963). The effect of 6-mercaptopurine on the duration of steroid-induced remissions in acute leukemia: A model for evaluation of other potentially useful therapy. *Blood*, 21(6), 699–716.

[39] Genest C., Carabarín-Aguirre A. and Harvey F. (2013). Copula parameter estimation using Blomqvists beta. *Journal de la Société Française de Statistique*, 154(1), 5–24.

[40] Genest C., Ghoudi K. and Rivest L.P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, 82(3), 543–552.

[41] Gibbons J.D. and Chakraborti S. (2020). *Nonparametric statistical inference.* Chemical Rubber Company (CRC) press.

[42] Greenwood M. (1926). A report on the natural duration of cancer. *Reports on Public Health and Medical Subjects. Ministry of Health*, 33, 1–26.

[43] Gross S.T. and Lai T.L. (1996). Bootstrap methods for truncated and censored data. *Statistica Sinica*, 6, 509–530.

[44] Hall P. (1990). Using the bootstrap to estimate mean squared error and select smoothing parameter in nonparametric problems. *Journal of Multivariate Analysis*, 32, 177–203.

[45] Hasselman B. (2018). *nleqslv: Solve Systems of Nonlinear Equations.* R package version 3.3.2.

[46] Hayfield T. and Racine J.S. (2008). Nonparametric Econometrics: The np Package. *Journal of Statistical Software*, 27(5).

[47] Heller G. and Venkatraman E. (1996). Resampling procedures to compare two survival distributions in the presence of right-censored data. *Biometrics*, 52, 1204–1213.

[48] Henderson D.J. and Parmeter C.F. (2011). Normal reference bandwidths for the general order, multivariate kernel density derivative estimator. *Statistics & Probability Letters*, 82(12), 2198–2205.

[49] Hill B.M. (1968). Posterior distribution of percentiles: Bayes' theorem for sampling from a population. *Journal of the American Statistical Association*, 63, 677–691.

[50] Hill B.M. (1988). De Finettis theorem, induction, and $A_{(n)}$ or Bayesian nonparametric predictive inference (with discussion). *In: Bayesian Statistics*, 3, 211–241. Bernardo, J.M., DeGroot, M.H., Lindley, D.V. and Smith, A.F.M. (Eds), Oxford University Press.

[51] Hill B.M. (1993). Parametric models for A-(n): Splitting processes and mixtures. *Journal of the Royal Statistical Society. Series B (Methodological)*, 55, 423–433.

[52] Joe H. (1997). *Multivariate models and multivariate dependence concepts*, volume 73. New Jersey: Chapman & Hall.

[53] Kaplan E.L. and Meier P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53, 457–481.

[54] Klein J.P. and Moeschberger M.L. (2006). *Survival Analysis Techniques for Censored and Truncated Data.* Springer Science & Business Media.

[55] Kojadinovic I. and Yan J. (2010). Comparison of three semiparametric methods for estimating dependence parameters in copula models. *Insurance: Mathematics and Economics*, 47(1), 52–63.

[56] Lawless J.F. (1982). *Statistical Models and Methods for Lifetime Data.* John Wiley & Sons, New York.

[57] Lawless J.F. and Fredette M. (2005). Frequentist prediction intervals and predictive distributions. *Biometrika*, 92(3), 529–542.

[58] Lee W.C. and Rodgers J.L. (1998). Bootstrapping correlation coefficients using univariate and bivariate sampling. *Psychological Methods*, 3(1), 91–103.

[59] Li Q. and Racine J.S. (2006). *Nonparametric econometrics: theory and practice.* Princeton University Press.

[60] Mandel M. and Betensky R.A. (2008). Simultaneous confidence intervals based on the percentile bootstrap approach. *Computational statistics & data analysis*, 52(4), 2158–2165.

[61] Marks C.E., Glen A.G., Robinson M.W. and Leemis L.M. (2014). Applying bootstrap methods to system reliability. *The American Statistician*, 68, 174–182.

[62] Maturi T.A. (2010). *Nonparametric predictive inference for multiple comparisons.* Ph.D. thesis, Durham University, UK. Https://npi-statistics.com/pdfs/theses/TM10.pdf.

[63] Mazucheli J., Barros E.A.C. and Achcar J.A. (2005). Bootstrap confidence intervals for the mode of the hazard function. *Computer methods and programs in biomedicine*, 79(1), 39–47.

[64] Muhammad N. (2016). *Predictive inference with copulas for bivariate data.* Ph.D. thesis, Durham University, UK. Https://npi-statistics.com/pdfs/theses/NM16.pdf.

[65] Muhammad N., Coolen F.P.A. and Coolen-Maturi T. (2016). Predictive inference for bivariate data with nonparametric copula. *The American Institute of Physics (AIP) Conference Proceedings*, 1750(1), 0600041–0600048. Https://aip.scitation.org/doi/abs/10.1063/1.4954609.

[66] Muhammad N., Coolen-Maturi T. and Coolen F.P.A. (2018). Nonparametric predictive inference with parametric copulas for combining bivariate diagnostic tests. *Statistics, Optimization & Information Computing*, 6(3), 398–408.

[67] Nahman N.S., Middendorf D.F., Bay W.H., McElligott R., Powell S. and Anderson J. (1992). Modification of the percutaneous approach to peritoneal dialysis

catheter placement under peritoneoscopic visualization: clinical results in 78 patients. *Journal of the American Society of Nephrology*, 3(1), 103–107.

[68] Nelsen R.B. (2007). *An introduction to copulas.* New York: Springer Science & Business Media.

[69] Owen A.B. (2001). *Empirical likelihood.* Chapman and Hall/CRC.

[70] Patti S., Biganzoli E. and Boracchi P. (2007). Review of the Maximum Likelihood Functions for Right Censored Data. A New Elementary Derivation. *COBRA Preprint Series*, p. 21.

[71] Rasmussen J.L. (1987). Estimating correlation coefficients: Bootstrap and parametric approaches. *Psychological Bulletin*, 101(1), 136–139.

[72] Rosenkranz G.K. (2014). Bootstrap corrections of treatment effect estimates following selection. *Computational Statistics & Data Analysis*, 69, 220–227.

[73] Rubin D.B. (1981). The bayesian bootstrap. *The annals of statistics*, pp. 130–134.

[74] Schepsmeier U., Stoeber J., Brechmann E.C., Graeler B., Nagler T., Erhardt T., Almeida C., Min A., Czado C., Hofmann M. *et al.* (2015). Package VineCopula. *R package version*, 2(5).

[75] Sedmak D.D., Meineke T.A., Knechtges D.S. and Anderson J. (1989). Prognostic significance of cytokeratin-positive breast cancer metastases. *Modern Pathology*, 2(5), 516–520.

[76] Shafer G. (1976). *A mathematical theory of evidence*, volume 42. Princeton university press.

[77] Silverman B.W. (1986). *Density estimation for statistics and data analysis.* London: Chapman & Hall.

[78] Silverman B.W. and Young G.A. (1987). The bootstrap: to smooth or not to smooth? *Biometrika*, 74, 469–479.

[79] Sklar A.W. (1959). Fonctions de repartitiona‘ n-dimensions et leurs marges. *Publications de lInstitut de Statistique de lUniversite de Paris*, 8, 229–231.

[80] Strube M.J. (1988). Bootstrap Type I error rates for the correlation coefficient: An examination of alternate procedures. *Psychological Bulletin*, 104(2), 290–292.

[81] Therneau T.M. (2020). *A Package for Survival Analysis in R.* R package version 3.2-3.

[82] Tukey J.W. (1947). Non-parametric estimation II. statistically equivalent blocks and tolerance regions–the continuous case. *Annals of Mathematical Statistics*, 18(4), 529–539.

[83] Utkin L.V. and Coolen F.P.A. (2007). Imprecise reliability: an introductory overview. *Computational intelligence in reliability engineering*, pp. 261–306.

[84] Walley P. (1991). Statistical reasoning with imprecise probabilities. Chapman and Hall, London.

[85] Wan F. (2017). Simulating survival data with predefined censoring rates for proportional hazards models. *Statistics in Medicine*, 36, 721–880.

[86] Wu S. (2014). Construction of asymmetric copulas and its application in two-dimensional reliability modelling. *European Journal of Operational Research*, 238(2), 476–485.

[87] Young G.A. (1990). Alternative smoothed bootstraps. *Journal of the Royal Statistical Society: Series B (Methodological)*, 52, 477–484.