Contributions to Nonparametric Predictive Inference with Right-Censored Data

MAHNASHI, ALI,MOHAMMED,Y

Contributions to Nonparametric Predictive Inference with Right-Censored Data

Ali Mohammed Y. Mahnashi

A Thesis presented for the degree of
Doctor of Philosophy

Statistics and Probability
Department of Mathematical Sciences
University of Durham
England

May 2022

Dedicated to

My parents' soul.

My beloved wife and children.

My brothers and sisters.

Contributions to Nonparametric Predictive Inference with Right-Censored Data

Ali Mohammed Y. Mahnashi

Submitted for the degree of Doctor of Philosophy
May 2022

Abstract

A right-censored data set is most common in reliability and survival analyses. It occurs when a particular event of interest is not fully observed in an experiment and when there is no information provided about a random quantity except that it exceeds a certain value. Nonparametric Predictive Inference (NPI) is a frequentist statistical method based on only few assumptions. It focuses explicitly on future observations and uses imprecise probabilities, based on Hill's assumption $A_{(n)}$, to quantify uncertainty. NPI has been developed for several types of data, including right-censored data. However, NPI with right-censored data has only taken into consideration a single future observation.

This thesis presents three contributions to NPI with right-censored data. First, some statistical methods on extreme values assume that the endpoint of the support is equal to the largest observed value in a data set. However, a question that may be of interest is whether, for some right-censored observations in a data set, their actual value might exceed the largest observed value.

Secondly, the actuarial estimator provides information on the number of events and censorings at any given discrete point in time. The nature of this estimator is such that, at every time point (except if all people in the data set have died) there is right-censoring, the data themselves are not necessarily right-censored. A similar approach is followed here, but we aim to develop an alternative method to the actuarial estimator, based on NPI with right-censored data. The proposed method will be used to derive NPI lower and upper probabilities for a variety of events of

interest. As an example application, we apply the newly developed method to obtain NPI lower and upper survival probabilities for reliability of systems.

Thirdly, NPI has been developed for real-valued data that contain right-censored observations but only a single future observation was considered. There may be reasons to be interested in multiple future observations, and it is important that in the NPI approach such multiple future observations are not conditionally independent given the data. We extend NPI for right-censored data by considering two future observations. Particularly, we present NPI lower and upper probabilities for the event that both future observations are greater than time t. We apply the proposed method to system reliability.

The results in this thesis widen the applicability of NPI for several real-world scenarios, while also suggesting new related topics for research.

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.

Acknowledgements

My first and primary acknowledgement must go to Allah my God for giving me countless blessings in general, and especially for what it has been done towards this thesis, so Alhamdulillah. I give my endless gratitude to my parents, my father Mohammed and my mother Fatimah, who always believed in my ability to succeed in the academic arena. Although they are gone, their belief in me has made this journey possible.

Secondly, my special appreciation and sincere gratitude must go to my academic supervisors and mentors, Professor Frank Coolen and Dr. Tahani Coolen-Maturi for their unlimited support, professional advice and guidance in all the time of my PhD study and over the period of writing this thesis. I can honestly say that what I have accomplished would not have been possible without their guidance and support. In particular, their enthusiasm, untiring support, psychological encouragement and calm advice during the difficult period due to COVID-19 pandemic, would not be forgotten as long as I am alive. It has really been the greatest fortune in my education to be able to work with and learn from them. Along with my profound appreciation, they will always have my utmost admiration and respect.

Thirdly, a big special thanks go to my beloved family. My wife Fatimah Mahnashi, my daughter Hemss, and my son Wesam, all are a priceless gift, greater than anything I can imagine. Without their help, support, and encouragements, I am not sure what I would do. I do not have the words to tell them how truly fortunate I feel to have them in my life. Thank you very much for your patience on me throughout my academic journey abroad, especially over my years in the doctoral program. Also, a great gratitude is expressed to my brothers and sisters for their confidence, encouragement, and their prayers for me.

Fourthly, I would like to acknowledge the financial support from Jazan University in Saudi Arabia and the Saudi Arabian Cultural Bureau (SACB) in London that were instrumental in helping me complete my PhD program in Durham University successfully. Also, my appreciation to Durham University through the Department of Mathematical Sciences for offering the most important elements for successfully achieving the goal of my doctoral program such as funding for participating in conferences.

Finally, I would like to express a general note of thanks to everyone who has helped and provided important support during this process, especially my close friends and colleagues.

Contents

Chapter 1

Introduction

1.1 Motivation

In survival analysis, one of the primary characteristics is that some data may not be fully observable, but are instead censored. In many cases, event times are subject to right-censoring, which simply means that for a specific individual it is known that the event has not yet occurred at a particular time [46]. In other words, an observation for an individual is right-censored at c if its lifetime is known only to be greater than c. While there are several other common types of censoring, including left-censoring and interval-censoring, right-censoring occurs most frequently in applied statistics. This thesis considers data sets including right-censored observations.

Nonparametric Predictive Inference (NPI) is a frequentist statistical method that is based on Hill's assumption $A_{(n)}$ [39, 40], as introduced in Chapter 2, which uses imprecise probabilities [13, 45, 55, 61, 63] to quantify uncertainty. NPI gives lower and upper probabilities for a future, observable, random quantity, conditioned on observed values of related random quantities, based on the assumption $A_{(n)}$ [12]. NPI has been developed for a variety of data types, such as Bernoulli data [20, 27], real-valued data [29, 30, 52], data with right-censored observations [31, 32], bivariate data [34], multinomial data [14, 22], and circular data [21]. Moreover, NPI has been developed for a wide variety of statistical applications, such as reliability analysis [2], operational research [26] and medical survival data [43]. This thesis is mostly theoretical in nature, and uses data from the literature to illustrate how the developed methods are used. In this thesis, we present three new contributions to NPI for data with right-censored observations.

In Chapter 3, it is commonly assumed in the literature on extreme values that the endpoint of the support is equal to the largest observed value. Alves et al. [8], for example, assumed a fixed endpoint and may have set it to the highest observation. In this case, it further leads to a question that, if some of the values in a data set that have been right-censored exceed the largest observed value? We use the $A_{(n)}$ assumption and assume exchangeability for all random quantities known to be in the risk set just prior to a certain censoring time and for random quantities that are right-censored at that certain censoring time, to answer the question of interest. To illustrate the utility of the proposed method, it is applied to a real data set termed "Supercentenarian data" [8].

In Chapter 4, we assume that the time is discrete, as it was in respect of the actuarial model, with regard to the number of events and censorings that take place at each discrete time. In such a situation, Bernoulli data, that is, the data on how many future individuals, who were alive at a certain time, would survive the next time, is of interest. The actuarial estimator is used when there is right-censored data in order to determine the probability of survival. Considering discrete-time data, we present the NPI approach as a predictive alternative to the actuarial estimator with right-censored data. By using the NPI method for Bernoulli data [20], we construct the NPI alternative to the actuarial estimator. The NPI alternative to the actuarial estimation method will be used to derive NPI lower and upper probabilities with respect to the multiple future observations for some events of interest. This approach will be applied together with the survival signature method to some reliability systems.

Coolen and Yan [32] have developed NPI for right-censored data based on a generalization of $A_{(n)}$, called the right-censoring $A_{(n)}$ assumption, or rc-$A_{(n)}$, but it was only developed for a single future observation. In practice, however, there may be reasons to be interested in multiple future observations; it is important that in the NPI approach, such multiple future observations are not conditionally independent given the data. In Chapter 5, we develop NPI for two future observations based on the assumption rc-$A_{(n)}$ without further assumptions [49] and as an example application we consider reliability of series systems.

1.2 Censoring

When collecting data in an experiment or observational study, censoring occurs when the event of interest is not fully observed. A common form of censoring is when the only information about a random quantity is that is exceeds a particular value. This is called right-censoring, and typically occurs in reliability and survival analyses.

While this thesis considers right-censored data, we first give a brief general introduction to censoring. Lawless [47] identified three types of censoring: (i) right-censoring, (ii) left-censoring, and (iii) interval-censoring. As a first type of censoring, *right-censored* data set, is the most common in reliability and survival analyses [53]. A dataset is referred to as *right-censored* if an individual has been removed from the study or the study has been terminated and an individual still has not yet experienced the event of interest. We can use the example of light bulbs to illustrate *right-censoring* data. In such a study, if a light bulb had not failed before the study ends, while we know that this light bulb is still functional, we do not know when it will fail beyond the end of the study; therefore, its lifetime is considered *right-censored.* Two main types of right-censoring exist. *Type I censoring* occurs when we conduct an experiment, and a decision is made to stop the study at a fixed time. As a result, any individuals remaining functional beyond the termination of the study are said to be right-censored. In addition, *Type II* censoring arises when a study with a particular number of observations is conducted, however, a decision has been made to stop the study when a specific number of failure times has been reached. Then, any individuals remaining functional or surviving after termination are said to be right-censored. Moreover, right-censoring may also occur due to different reasons, not only *Types I* or *II*, e.g. if an individual leaves the study for a different reason.

The second type is called *left-censored* data. Here, failure appears before a specific time. For example, suppose a study was conducted to estimate light bulb performance and a decision was made to terminate the study at a particular time, for example, in five years. Regular monthly checks were assigned during the lifetime of the study. In such a study, if one light bulb had failed before the first monthly check, then the only information we know is that the light bulb's failure time is less than one month, but we do not know exactly by how much. Therefore, this situation illustrates left-censored data.

The third type of censoring is called *interval-censored* data. In the case where we do not know exactly the time of failure for an individual unit within a study's duration, but we know that the failure occurs within a particular finite interval, then this case is said to represent interval-censored data. Using the light bulb example, if one light bulb is found to have failed between the first- and second-month checks, then the only information that we have about this bulb is that it was found to have survived the first month while it failed in the second month. Therefore, this is known as an interval-censored item since its failure occurs between two failures times.

Some assumptions are assumed with regard to the mechanisms of censoring as well as their relationship to the event time. We assume that a censoring time can be either predetermined or random [36, 53]. One can do statistical inference in case of non-informative censoring, that the censoring arises due to reasons which are independent of the random quantity of interest for a unit. Other one can also do statistical inference in case of informative censoring, but then things change and one needs to model the relation between the censoring mechanism and the random quantities of interest. So we only need to assume that for a right-censored observation, the remaining time till the event of interest, at the moment of censoring, is exchangeable with such remaining times for all other units which are still in the study. This is the key assumption underlying rc-$A_{(n)}$ [32], which will be explained in Section 2.4. And right centering of *Type I* and *II*, all of those types can fit perfectly with the other assumption and for our interest, it is irrelevant as long as it is non-informative right-censoring.

There are many nonparametric statistical methods can easily deal with all kinds of censoring, such as the Kaplan-Meier estimator (KM) [44], which is the most commonly method used for dealing with right-censored data. In the following, the KM estimator is briefly presented. Then, we will briefly discuss the link between NPI for right-censored data and KM estimator [32] in Section 2.4.

The KM estimator (also known as the Product-Limit (PL) estimator) is a classical nonparametric method for estimating the survival function using lifetime data that contains right-censored data. The KM estimator, presented by Kaplan-Meier [44], has become one of the most used methods in applications where event times are considered, e.g. in medical statistics and in reliability.

Suppose that there are observations on n individuals, and there are $h (h \leq n)$ distinct event times $t_1 < t_2 < ... < t_h$. Let d_{t_i} be the number of events occurred simultaneously at time t_i. For $n - \sum_{i=1}^{h} d_{t_i}$ individuals, no event time is observed, but assume that these individuals are right-censored. Let c_{t_i} be the number of individuals censored at time t_i. Suppose that there are l different right-censoring times, $c_1 < c_2 < \cdots < c_l$. The survival function at time t is defined as $S(t) = P(X \geq t)$. The KM estimator of the survival function $S(t)$ is given by

$$\hat{S}(t) = \prod_{i:t_i \leq t} \frac{\tilde{n}_{t_i} - d_{t_i}}{\tilde{n}_{t_i}} = \prod_{i:t_i \leq t} \left(1 - \frac{d_{t_i}}{\tilde{n}_{t_i}}\right),$$

where $\tilde{n}_{t_i}$ is the number of individuals in the risk set (still functioning or alive and uncensored) just prior to t_i, then we have the relation $\tilde{n}_{t_i} = \tilde{n}_{t_{i-1}} - d_{t_{i-1}} - c_{t_{i-1}}$.

For $i = 0, 1, 2, ..., h - 1$, with $t_0 = 0$, the KM estimator is a step function which decreases at event time t_i by a factor $(\tilde{n}_{t_i} - d_{t_i})/\tilde{n}_{t_i}$. We notice that $\hat{S}(t) = 0$ when the largest observation is at the event time t_h. The KM estimator will be a positive constant on $[t_h, c_k)$, $k = 1, 2, \ldots, l$, if the largest observation is a right-censoring time at c_l, but it is often left undefined for interval $[c_l, \infty)$. In addition, $\hat{S}(t) = 1$ over the interval $[0, t_1)$. In general, the censoring times do not directly affect the KM estimator [44]. They have a direct effect only on the size of the later steps; i.e., every change in value happens at an event time, there is no change at censoring times. In case there is no censoring, then the KM estimator of the survival function is equal to the empirical estimate of the survival function [44].

1.3 Outline of the thesis

This thesis is structured as follows. Chapter 2 introduces the basic concepts needed in this thesis. We start by providing brief introductions of Hill's assumption $A_{(n)}$ [39] as well as the imprecise probability [12, 13]. Next, a brief introduction to NPI [12, 21] is provided. Furthermore, we briefly review NPI for Bernoulli data [20], which will be used in Chapter 4. NPI for right-censored data for a future observation based on the rc-$A_{(n)}$ assumption [32] is also presented, which will be used in Chapter 5.

Chapters 3 to 5 contain the main contributions of this thesis. In some statistical methods on extreme values, the endpoint of the support is considered to be equal to the largest observed value in a data set. There is, however, a question that may be of interest, namely whether the actual value of some right-censored observations in a data set would be larger than the largest observed value. In Chapter 3, we will present new results regarding our investigation of the question of interest. We illustrate our new results using a real data set about supercentenarian people.

In Chapter 4, we consider discrete time and an NPI-based alternative to the actuarial estimator with right-censored data is introduced, using NPI for Bernoulli data [20]. Based on this method, NPI lower and upper probabilities for several events of interest are derived. This method will then be compared to a nonparametric statistical method, called 'NPI for grouped data' [65], which is discussed in this chapter. Further insights are presented in terms of deriving the NPI lower and upper probabilities for the event that there are at least one trial succeeds in multiple future Bernoulli trials, which will be applied to reliability of systems using the concept of the survival signature [4, 21, 23, 25] and the NPI for Bernoulli data [20].

In Chapter 5, NPI for two future observations with right-censored data [49] is introduced. We generalise the methodology of NPI for right-censored data that has been presented for a single future observation [32] to two future observations, by considering the NPI lower and upper probabilities for the event that both future observations are greater than time t. The results presented in this chapter are applied to system reliability by considering a small series system [49].

It is worth mentioning that part of Chapter 4 was presented at the 12th International Conference of the ERCIM WG on Computational and Methodological

Statistics and at the 13th International Conference on Computational and Financial Econometrics in London on 14–16 December 2019. Part of Chapter 5 has been presented at the 29th European Safety and Reliability Conference (ESREL) conducted in Hannover on 22–26 September 2019 and a related short paper was published in the conference proceedings [49]. Also, the results of this chapter have been presented at the 1st UK Reliability Meeting at Durham University on 1–3 April 2019 as well as at the 12th Workshop on Principles and Methods of Statistical Inference with Interval Probabilities 2019 (WPMSIIP) at Durham University on 9–12 September 2019.

Chapter 2

Nonparametric Predictive Inference (NPI)

Over the last two decades, Nonparametric Predictive Inference (NPI) has been developed for a range of data types, and for a variety of applications and problems in statistics and related areas such as risk, reliability, operations research and finance. NPI is a statistical method that requires only a few assumptions, based on Hill's assumption $A_{(n)}$ [39], and uses imprecise probabilities to quantify uncertainty [12, 21]. The purpose of this chapter is to provide a brief overview of the background concepts from the literature that are explicitly relevant to the new statistical inferences proposed in this thesis.

This chapter is organized as follows. Section 2.1 presents a brief overview of the Hill's assumption $A_{(n)}$ and the imprecise probability. Section 2.2 provides an overview of NPI for real-valued data. NPI for Bernoulli quantities is provided in Section 2.3. Section 2.4 presents an overview of NPI for right-censored data, based on the assumption rc-$A_{(n)}$, for a single future observation.

2.1 $A_{(n)}$ assumption and imprecise probability

Due to the fact that the NPI method is a frequentist method based on Hill's assumption $A_{(n)}$ and utilizes the imprecise probability theory to quantify uncertainty, in this section, we will discuss the nature and properties of $A_{(n)}$ as well as some basic aspects of imprecise probability theory. Assume that $X_1, X_2, \ldots, X_n, X_{n+1}$ are real-valued absolutely continuous and exchangeable random quantities. Let the ordered observed values of $X_1, \ldots, X_n$ be denoted as $x_1 < x_2 < \cdots < x_n$. To simplify notation, let $x_0 = -\infty$ and $x_{n+1} = \infty$, or we assume $x_0 = 0$ in case of nonnegative random quantities [42]. It is assumed that there are no ties between the observations of the data. In the case of ties, we assume that the tied observations differ by a small amount, which is a common strategy in statistics to break ties [40]. These n observations divide up the real-line into $n+1$ intervals $I_j = (x_j, x_{j+1})$, where $j = 0, 1, \ldots, n$. Based on n observations, the assumption $A_{(n)}$ [41] is that the probability that the next future observation X_{n+1} is equally likely to fall in each open interval (x_j, x_{j+1}), for all $j = 0, 1, \ldots, n$, so

$$P_{X_{n+1}}(x_j, x_{j+1}) = \frac{1}{n+1} \qquad \text{for all } \; j = 0, 1, \ldots, n \tag{2.1}$$

The data carry information about the location but no information about the rank of the future observations, corresponding to the absence of prior knowledge, so $A_{(n)}$ is considered as a post-data assumption related to finite exchangeability, and assumes nothing else [35]. For a detailed presentation and discussion of $A_{(n)}$, see Hill [41].

The assumption $A_{(n)}$ alone is insufficient for constructing precise probabilities for many events of interest, but it is still useful to derive bounds for probability, effectively by applying De Finetti's Fundamental Theorem of Probability [35], or Walley's concept of natural extension [61], which provide lower and upper probabilities in interval probability theory. Weichselberger [64] also developed a formal foundation for interval probability, via lower and upper probabilities, by applying the principles of Kolmogorov's axioms. These lower and upper probabilities are also known as imprecise probabilities in accordance with the imprecise probability theory [12, 13].

Imprecise probabilities have been proposed and studied since at least the middle of the nineteenth century [18]. Recently, the topic of imprecise probabilities has become increasingly prominent, resulting in a series of conferences and a project website*. There are several interpretations of the lower and upper probabilities for event A, which are denoted by $\underline{P}(A)$ and $\overline{P}(A)$, respectively [21]. According to Walley [61], for instance, the lower and upper probabilities for event A can be interpreted as supremum buying price and infimum selling price, respectively, of a gamble on the event A, in which 1 is paid when the event occurs and 0 if the event does not occur. From a classical perspective, lower and upper probabilities can be interpreted as bounds on precise probabilities, because of the lack of information or the desire not to make further assumptions. The theory of imprecise probability clearly demonstrates that bounds provide valuable information regarding the uncertainty of events caused by a lack of information [12, 61, 62, 63, 64]. The precise classical probability of an event A is simply a special case of the imprecise probability, when $\underline{P}(A) = \overline{P}(A)$, whereas the total absence of information about the event A can be reflected by $\underline{P}(A) = 0$ and $\overline{P}(A) = 1$. Next, we outline several important aspects of imprecise probability theory relevant to $A_{(n)}$-based inference [12]. As a general rule, in imprecise probability theory, the lower and upper probabilities for the event A are $\underline{P}(A) = 1 - \overline{P}(A^c)$, which is the conjugacy property, where A^c represents the complementary event of A. In many cases, this conjugacy property can be utilised in order to simplify the calculation of imprecise probabilities for events of interest and their complementary events. For events A and B, such that $A \cap B = \emptyset$, the lower probability is superadditive and the upper probability is subadditive, so

$$\underline{P}(A \cup B) \geq \underline{P}(A) + \underline{P}(B)$$
$$\overline{P}(A \cup B) \leq \overline{P}(A) + \overline{P}(B)$$

In the following section, we will introduce the statistical method NPI which assigns lower and upper probabilities to events involving a future random observation X_{n+1}.

* *www.sipta.org*

2.2 NPI for real-valued data

Nonparametric Predictive Inference (NPI) is a statistical technique derived from Hill's assumption $A_{(n)}$ [39] to give direct probabilities for a future observable random quantity based on observed values of corresponding random quantities [12, 21]. In NPI, uncertainty about events of interest is quantified by lower and upper probabilities. On the basis of $A_{(n)}$, Augustin and Coolen [12] introduced predictive lower and upper probabilities as follows.

The NPI lower and upper probabilities for the event $X_{n+1} \in B$, where $B \subset \mathbb{R}$, given the intervals $I_j = (x_j, x_{j+1})$, $j = 0, 1, \ldots, n$, resulting from n real-valued non-tied observations and based on the assumption $A_{(n)}$, are

$$\underline{P}(X_{n+1} \in B) = \frac{1}{n+1} \; |\{j : I_j \subseteq B\}| \tag{2.2}$$

$$\overline{P}(X_{n+1} \in B) = \frac{1}{n+1} \; |\{j : I_j \cap B \neq \emptyset\}| \tag{2.3}$$

According to Equation (2.2), the NPI lower probability can be calculated by summing only the $A_{(n)}$-based probabilities assigned to intervals I_j which are fully contained in B. The NPI upper probability, Equation (2.3), can be calculated by summing all the $A_{(n)}$-based probabilities assigned to intervals I_j which have non-empty intersection with B. Augustin and Coolen [12] proved strong properties of NPI lower and upper probabilities which fit well in the theory of interval probability [61, 63, 64]. When assuming only $A_{(n)}$ corresponding logically to exchangeability over all the $n+1$ random quantities, the NPI lower and upper probabilities are the sharpest bounds on the probability for a given event of interest [12].

The nature of $A_{(n)}$ results in NPI being a frequentist statistical methodology [12, 39, 40], which can be interpreted in a way similar to that of Bayesian statistics [20, 42]. As in Bayesian statistics, NPI draws its inferences only from the actual data observed, but it is worth pointing out that one must accept the exchangeability assumption for the data and future observations, which may not always be straightforward, depending on experimental setup.

Furthermore, NPI has been adapted to suit different types of data and a multitude of applications. For instance, NPI has been applied to Bernoulli data [20, 27], which will be introduced in detail in Section 2.3, data with right-censored observations [31, 32], as stated in more detail in Section 2.4, bivariate data [34], multinomial data [14, 22] and circular data [21]. The NPI approach, in the way of dealing with right-censoring times, has also been applied to real medical survival data including right-censored observations by Janurová and Briš [43].

In NPI approach, when dealing with tied observations, it is quite straightforward to assume that the tied observations differ by small amounts [40, 51]. If there is a tie between the event time and the right-censoring time, then we assume that the right-censoring time is just beyond the event time, a common approach to deal with a such situation in statistics [44]. Throughout this thesis, we assume that there are no ties between data observations; however, if ties do exist, we use the same approach that is used in literature to deal with tied observations (see [32, 40, 44, 51] for more details).

2.3 NPI for Bernoulli quantities

This section summarizes NPI for Bernoulli random quantities, based on the $A_{(n)}$ assumption, as introduced by Coolen [20]. The special cases given in this section are all used in Chapter 4.

Assume that there are $n + m$ exchangeable Bernoulli trials with 'failure' and 'success' as possible outcomes for each trial, and data containing s successes in n trials. Let X_1^n denote the random number of successes in trials 1 to n, and X_{n+1}^{n+m} denote the random number of successes in trials $n+1$ to $n+m$. Let $R_t = \{r_1, \ldots, r_t\}$, with $0 \leq r_1 < r_2 < \ldots < r_t \leq m$ and $1 \leq t \leq m+1$, and to simplify the notation, we define $\binom{s+r_0}{s} = 0$. Coolen [20] gives the general formulas of NPI for Bernoulli data for the event of interest that $X_{n+1}^{n+m} \in R_t$, given $X_1^n = s$, where $s \in \{0, 1, \ldots, n\}$. The NPI upper probability for the event $X_{n+1}^{n+m} \in R_t \mid X_1^n = s$, is

$$\overline{P}(X_{n+1}^{n+m} \in R_t \mid X_1^n = s) = \binom{n+m}{n}^{-1} \times \sum_{j=1}^{t} \left[\binom{s+r_j}{s} - \binom{s+r_{j-1}}{s} \right] \binom{n-s+m-r_j}{n-s} \tag{2.4}$$

The corresponding NPI lower probability for the event $X_{n+1}^{n+m} \in R_t \mid X_1^n = s$ can be derived via the conjugacy property,

$$\underline{P}(X_{n+1}^{n+m} \in R_t \mid X_1^n = s) = 1 - \overline{P}(X_{n+1}^{n+m} \in R_t^c \mid X_1^n = s) \tag{2.5}$$

with R_t^c the complement of R_t, that is $R_t^c = \{0, 1, \ldots, m\} \backslash R_t$.

The above NPI lower and upper probabilities were derived by Coolen [20] using direct counting arguments. This method is developed based on the $A_{(n)}$ assumption for m future observations given n observed values, and on the $A_{(n+m-1)}$ assumption for a latent variable representation of Bernoulli quantities represented by values on the real line, with a threshold such that values on one side are successes and values on the other side are failures. Assuming $A_{(n)}, \ldots, A_{(n+m-1)}$ sequentially and all observations being exchangeable, that this leads to all orderings $\binom{n+m}{n}$ being equally likely.

For any specific ordering of the m future observations among the n data observations, we can find all possible combinations of s successes in the data observations and r successes in the future data, so if s can have $n+1$ values and r can have $m+1$ values, there are $(n+1) \times (m+1)$ possible combinations (s, r).

With some counting arguments, Coolen [20] gives the formulas of NPI for Bernoulli data for the event of interest that $X_{n+1}^{n+m} = r | X_1^n = s$, where $s \in \{0, 1, \ldots, n\}$. The NPI lower probability for the event $(X_{n+1}^{n+m} = r | X_1^n = s)$ is

$$\underline{P}(X_{n+1}^{n+m} = r \mid X_1^n = s) = \binom{n+m}{n}^{-1} \left[\binom{s-1+r}{s-1} \binom{n-s-1+m-r}{m-r} \right] \tag{2.6}$$

and in case m future Bernoulli trials are all successes, so $r = m$, Equation (2.6) can be reduced to the following

$$\underline{P}(X_{n+1}^{n+m} = m \mid X_1^n = s) = \binom{n+m}{n}^{-1} \binom{s-1+r}{s-1} \tag{2.7}$$

which can be simplified even further to a simple product as

$$\underline{P}(X_{n+1}^{n+m} = m \mid X_1^n = s) = \prod_{i=1}^{m} \frac{s+i-1}{n+i} \tag{2.8}$$

The associated NPI upper probability for the event $X_{n+1}^{n+m} = r | X_1^n = s$ is

$$\overline{P}(X_{n+1}^{n+m} = r \mid X_1^n = s) = \binom{n+m}{n}^{-1} \left[\binom{s+r}{s} \binom{n-s+m-r}{n-s} \right] \tag{2.9}$$

and for $r = m$, Equation (2.9) can be reduced to the follows

$$\overline{P}(X_{n+1}^{n+m} = m \mid X_1^n = s) = \binom{n+m}{n}^{-1} \binom{s+r}{s} \tag{2.10}$$

which can be simplified even further to a simple product as

$$\overline{P}(X_{n+1}^{n+m} = m \mid X_1^n = s) = \prod_{i=1}^{m} \frac{s+i}{n+i} \tag{2.11}$$

where Equations (2.8) and (2.11) are all used in Chapter 4.

Given data (n, s), with further counting arguments, Aboalkhair [1] also derives the formulas of NPI for Bernoulli data for r successes out of m future Bernoulli trials, where $r \in \{0, 1, \dots, m\}$ and $s \in \{0, 1, \dots, n\}$. The NPI lower and upper probabilities for the event of interest that $X_{n+1}^{n+m} \geq r \mid X_1^n = s$, are

$$\begin{aligned}\overline{P}(X_{n+1}^{n+m} \geq r \mid X_1^n = s) = \binom{n+m}{n}^{-1} \times \Bigg[\binom{s+r}{s} \binom{n-s+m-r}{n-s} \\ + \sum_{\ell=r+1}^{m} \binom{s+\ell-1}{s-1} \binom{n-s+m-\ell}{n-s} \Bigg]\end{aligned} \tag{2.12}$$

$$\begin{aligned}\underline{P}(X_{n+1}^{n+m} \geq r \mid X_1^n = s) = 1 - \overline{P}(X_{n+1}^{n+m} \leq r \mid X_1^n = s) = \\ 1 - \binom{n+m}{n}^{-1} \times \left[\sum_{\ell=0}^{r-1} \binom{s+\ell-1}{s-1} \binom{n-s+m-\ell}{n-s} \right]\end{aligned} \tag{2.13}$$

As a special cases, for example, in case $m = 1$ future observation, the NPI lower and upper probabilities for the conditional event that $X_{n+1}^{n+1} = 1$ given data $X_1^n = s$, for $s \in \{0, ..., n\}$ and $m = r = 1$, are as follows [19]

$$\underline{P}(X_{n+1}^{n+1} = 1 \mid X_1^n = s) = \frac{s}{n+1} \tag{2.14}$$

$$\overline{P}(X_{n+1}^{n+1} = 1 \mid X_1^n = s) = \frac{s+1}{n+1} \tag{2.15}$$

In case the observed data are all successes, so $s = n$, then the NPI lower and upper probabilities for all $r \in \{0, 1, \dots, m\}$, are as follows

$$\underline{P}(X_{n+1}^{n+m} \geq r \mid X_1^n = n) = 1 - \binom{n+m}{n}^{-1} \binom{n+r-1}{n} \tag{2.16}$$

$$\overline{P}(X_{n+1}^{n+m} \geq r \mid X_1^n = n) = 1 \tag{2.17}$$

In case the observed data are all failures, so $s = 0$, then the NPI lower and upper probabilities for all $r \in \{0, 1, \ldots, m\}$, are as follows

$$\underline{P}(X_{n+1}^{n+m} \geq r \mid X_1^n = 0) = 0 \tag{2.18}$$

$$\overline{P}(X_{n+1}^{n+m} \geq r \mid X_1^n = 0) = \binom{n+m}{n}^{-1} \binom{n+m-r}{n} \tag{2.19}$$

Using the counting arguments presented by Aboalkhair [1], we determine the probability relationships between the events $X_{n+1}^{n+m} \geq r$ and $X_{n+1}^{n+m} \geq r+1$ for $r = 0, 1, \ldots, m-1$, as follows.

$$\underline{P}(X_{n+1}^{n+m} \geq r) - \underline{P}(X_{n+1}^{n+m} \geq r+1) = \binom{n+m}{n}^{-1} \binom{s+r-1}{s-1} \binom{n-s+m-r}{n-s} \tag{2.20}$$

$$\overline{P}(X_{n+1}^{n+m} \geq r) - \overline{P}(X_{n+1}^{n+m} \geq r+1) = \binom{n+m}{n}^{-1} \binom{s+r}{s} \binom{n-s+m-r-1}{n-s} \tag{2.21}$$

In Chapter 4, the development of NPI for Bernoulli quantities, presented in this section, will be utilised to present new statistical inferences related to NPI for discrete-time data with right-censoring.

2.4 NPI for right-censored data

The Hill's assumption $A_{(n)}$ [16], presented in Section 2.1, by itself is not suitable for right-censored data, so Coolen and Yan [32] presented a generalization of $A_{(n)}$, called the right-censoring $A_{(n)}$ assumption, abbreviated as rc-$A_{(n)}$, for right-censored data. They added a new assumption to $A_{(n)}$ to makes it more suitable for dealing with right-censored data. It is assumed that, at the moment of censoring, the residual lifetime of a right-censored observation is exchangeable with the residual lifetimes of all other observations that are not yet failed or censored [32].

According to the $A_{(n)}$ assumption [16], the probability distribution for a real-valued random quantity X_{n+1} is partially specified by probability mass assigned to open intervals, without any further restriction on the spread of the probability mass within each interval [32, 33]. A probability mass assigned in such a way to an interval (a, b) is denoted by $M_X(a, b)$, and referred to as a M-function value for $X \in (a, b)$. The M-function value should satisfy $0 \leq M_X(a, b) \leq 1$ and the M-function values specified on all intervals should sum up to one [32]. These M-functions are in the theory presented by Shafer [59].

In this section, we follow the notation and definitions presented by Maturi [51]. Consider the following data when determining the predictive probabilities for a future observation. Assume $X_1, \ldots, X_n, X_{n+1}$ are non-negative, exchangeable and continuous random quantities representing lifetimes. Suppose that there are in total n observations containing u failure times observations, $x_1 < x_2 < \cdots < x_u$, and $\nu = n - u$ right-censoring times, $c_1 < c_2 < \cdots < c_\nu$. For ease of notation, $x_0 = 0$ and $x_{u+1} = \infty$. Suppose further that there are s_i right-censored observations in the interval $I^i = (x_i, x_{i+1})$, denoted by $c_1^1 < c_2^i < \cdots < c_{s_i}^i$, so $\sum_{i=}^{u} s_i = \nu$, such that $c_{i^*}^i \in (x_i, x_{i+1})$, where $i = 0, 1, \ldots, u$ and $i^* = 1, 2, \ldots, s_i$. Again we assume that no ties occur between the data observations as discussed in Section 2.2.

On the basis of n given event times, the assumption $A_{(n)}$ offers a partially specified probability distribution for X_{n+1} in terms of M-function values. To deal with right-censored observations being present in the data, a generalization of $A_{(n)}$ was considered, that is the assumption $\tilde{A}_{(n)}$ [32].

Definition 2.4.1 ($\tilde{A}_{(n)}$ assumption)
On the basis of data including u event times and $\nu = n-u$ right-censoring times, the assumption $\tilde{A}_{(n)}$ partially specifies the probability distribution for the next observation X_{n+1} assigning probability masses to two types of open intervals, one formed by consecutive event times, (x_i, x_{i+1}), and the other is formed by a censoring time and infinity, $(c_{i^*}^i, \infty)$, expressed via the following M-function values:

$$\tilde{M}_{X_{n+1}}(x_i, x_{i+1}) = \frac{1}{n+1} \tag{2.22}$$

$$\tilde{M}_{X_{n+1}}(c_{i^*}^i, \infty) = \frac{1}{n+1} \tag{2.23}$$

where $i = 0, 1, \ldots, u$ and $i^* = 1, 2, \ldots, s_i$.

Note that the notation $\tilde{M}$ used in Equations (2.22) and (2.23) indicates that these M-function values are based on the assumption $\tilde{A}_{(n)}$. According to Equation (2.22), the probability masses for the intervals (x_i, x_{i+1}) created by the u event times are equal to $\frac{1}{n+1}$. Furthermore, probability mass $\frac{1}{n+1}$ is assigned to the interval $(c^i_{i^*}, \infty)$, without making any other assumptions, so the lifetime of this observation will occur at any point beyond $c^i_{i^*}$, as in Equation (2.23). Next, they [32] split the probability mass assigned to the interval $(c^i_{i^*}, \infty)$ into masses on sub-intervals.

Let $X_{c^i_{i^*}}$ denote the random quantity corresponding to the right-censoring at time $c^i_{i^*}$. According to [32], the probability masses assigned to intervals $(c^i_{i^*}, \infty)$ may have caused wide bounds on probabilities, so it would be helpful if these probability masses can be split into probability masses on sub-intervals. For this reason, Coolen and Yan [32, 65] proposed the assumption Shifted-$\tilde{A}_{(n)}$ for $X_{c^i_{i^*}}$, for which all we know is that the random quantity $X_{c^i_{i^*}}$ exceeds $c^i_{i^*}$.

Definition 2.4.2 (Shifted-$\tilde{A}_{(n)}$ assumption)
The assumption shifted-$\tilde{A}_{(n)}$ partially specifies the probability distribution for $X_{c^i_{i^*}}$, given that $X_{c^i_{i^*}} > c^i_{i^*}$, expressed via the following M-function values:

$$M_{X_{c^i_{i^*}}}(x_k, x_{k+1}) = \frac{1}{\tilde{n}_{c^i_{i^*}} + 1} \qquad \text{for } k = i+1, \ldots, u, \tag{2.24}$$

$$M_{X_{c^i_{i^*}}}(c^i_{i^*}, x_{k+1}) = \frac{1}{\tilde{n}_{c^i_{i^*}} + 1}, \tag{2.25}$$

$$M_{X_{c^i_{i^*}}}(c^i_l, \infty) = \frac{1}{\tilde{n}_{c^i_{i^*}} + 1} \qquad \text{for } l = i^*+1, \ldots, \nu. \tag{2.26}$$

where $\tilde{n}_{c^i_{i^*}}$ represents the number of observations in the risk set at time $c^i_{i^*}$, for $c^i_{i^*} \in (x_i, x_{i+1})$, $i^* = 1, 2, \ldots, s_i$.

This assumption is related to the fact that if the random quantities $X_1, X_2, \ldots, X_r$ are exchangeable, then the random quantities in any subset of $X_1, X_2, \ldots, X_r$ are also exchangeable [32, 65]. It also follows that as long as the random quantities $X_1, X_2, \ldots, X_r$ are exchangeable, then all are also exchangeable when they exceed a given value c [32, 65]. In this sense, the exchangeability assumption of all random quantities known to be in the risk set just prior to c_i is an appropriate assumption to handle random quantities that are right-censored at time $c^i_{i^*}$, and in fact implies the assumption of non-informative censoring [32, 65].

Based on the assumption of non-informative censoring, reviewed in Section 1.2, the assumption shifted-$\tilde{A}_{(n)}$ allows us to apply $A_{(n)}$ but with the starting point shifted from the value 0 to the observed right-censoring time $c^i_{i^*}$ [32, 65]. Clearly, the sum of the M-function values for $X_{c^i_{i^*}}$ over these sub-intervals, as in Equations (2.24), (2.25) and (2.26), is equal to one [32, 65].

Taking into account the two previously proposed assumptions '$\tilde{A}_{(n)}$' for X_{n+1} and 'shifted-$\tilde{A}_{(n)}$' for $X_{c^i_{i^*}}$, Coolen and Yan [32, 65] proposed the right-censoring $\tilde{A}_{(n)}$ assumption, denoted by rc-$\tilde{A}_{(n)}$, which allows splitting the total M-function values for X_{n+1} assigned to interval $(c^i_{i^*}, \infty)$ into separate M-function values for X_{n+1} assigned to sub-intervals of $(c^i_{i^*}, \infty)$.

Definition 2.4.3 (rc-$\tilde{A}_{(n)}$ assumption)
Let $\mathcal{P}_{c^i_{i^*}} = M_{X_{n+1}}(c^i_{i^*}, \infty)$ be the M-function value for X_{n+1} on the interval $(c^i_{i^*}, \infty)$, taking into account the effects of all previous right-censorings and $A_{(n)}$. The assumption rc-$\tilde{A}_{(n)}$ splits the probability mass of $M_{X_{n+1}}(c^i_{i^*}, \infty)$ as

$$M^{c^i_{i^*}}_{X_{n+1}}(x_k, x_{k+1}) = \frac{\mathcal{P}_{c^i_{i^*}}}{\tilde{n}_{c^i_{i^*}} + 1} \qquad \text{for } \; k = i+1, \ldots, u, \tag{2.27}$$

$$M^{c^i_{i^*}}_{X_{n+1}}(c^i_{i^*}, x_{k+1}) = \frac{\mathcal{P}_{c^i_{i^*}}}{\tilde{n}_{c^i_{i^*}} + 1}, \tag{2.28}$$

$$M^{c^i_{i^*}}_{X_{n+1}}(c^i_l, \infty) = \frac{\mathcal{P}_{c^i_{i^*}}}{\tilde{n}_{c^i_{i^*}} + 1} \qquad \text{for } \; l = i^*+1, \ldots, \nu. \tag{2.29}$$

where $\tilde{n}_{c^i_{i^*}}$ represents the number of observations in the risk set at time $c^i_{i^*}$, for $c^i_{i^*} \in (x_i, x_{i+1})$, where $i = 0, 1, \ldots, u$ and $i^* = 1, 2, \ldots, s_i$.

With the combined assumptions $\tilde{A}_{(n)}$ and rc-$\tilde{A}_{(n)}$ for $r = 1, 2, \ldots, i^*-1$, $i^* = 1, 2, \ldots, s_i$, and for any right-censoring time $c^i_{i^*}$, the $\mathcal{P}_{c^i_{i^*}}$ can be computed by

$$\mathcal{P}_{c^i_{i^*}} = M_{X_{n+1}}(c^i_{i^*}, \infty) = \frac{1}{n+1} \prod_{\{r: r<i^*\}} \frac{\tilde{n}_{c_r} + 1}{\tilde{n}_{c_r}}. \tag{2.30}$$

where $\tilde{n}_{c_r}$ is the number of individuals in the risk set just prior to time c_r [32, 65]. Note that throughout this thesis, a product over an empty set is defined to be equal to 1.

Consequently, based on the assumptions $\tilde{A}_{(n)}$, given by Definition 2.4.1, and rc-$\tilde{A}_{(n)}$, given by the Definition 2.4.3, the M-function values for X_{n+1} are finally all assigned to intervals (x_i, x_{i+1}) or $(c^i_{i^*}, x_{i+1})$ for $i = 0, 1, \ldots, u$ and $i^* = 1, 2, \ldots, s_i$, via considering an assumption called right-censoring $A_{(n)}$, which is also denoted as rc-$A_{(n)}$ [32, 65].

Definition 2.4.4 (rc-$A_{(n)}$ assumption)
The assumption rc-$A_{(n)}$ partially specifies the NPI-based probability distribution for the observable and non-negative random quantity X_{n+1}, via the following M-function values [33],

$$M_{X_{n+1}}(x_i, x_{i+1}) = \frac{1}{n+1} \prod_{\{r: c_r < x_i\}} \frac{\tilde{n}_{c_r} + 1}{\tilde{n}_{c_r}} \tag{2.31}$$

$$M_{X_{n+1}}(c^i_{i^*}, x_{i+1}) = \frac{1}{(n+1)\tilde{n}_{c^i_{i^*}}} \prod_{\{r: c_r < c^i_{i^*}\}} \frac{\tilde{n}_{c_r} + 1}{\tilde{n}_{c_r}} \tag{2.32}$$

where $i = 0, 1, \ldots, u$, $i^* = 1, 2, \ldots, s_i$ and $\tilde{n}_{c_r}$ represents the number of observations in the risk set just before time c_r.

Following Maturi [51] and based on the assumption rc-$A_{(n)}$, all M-function values that are assigned for X_{n+1} to be in one interval created by two consecutive observed event times, (x_i, x_{i+1}), lead to the following probability for the event $X_{n+1} \in (x_i, x_{i+1})$,

$$\begin{aligned} P_{X_{n+1}}(x_i, x_{i+1}) &= M_{X_{n+1}}(x_i, x_{i+1}) + \sum_{i^*=1}^{s_i} M_{X_{n+1}}(c^i_{i^*}, x_{i+1}) \\ &= \frac{1}{n+1} \prod_{r: c_r < x_{i+1}} \frac{\tilde{n}_{c_r} + 1}{\tilde{n}_{c_r}} \end{aligned} \tag{2.33}$$

Based on the rc-$A_{(n)}$ assumption, Maturi [51] derived simple closed-form expressions for the NPI lower and upper survival functions, $\underline{S}_{X_{n+1}}(t)$ and $\overline{S}_{X_{n+1}}(t)$. For

$t \in [t_a^i, x_{i+1})$ with $i = 1, 2, ..., u$ and $a = 0, 1, ..., s_i$, the NPI lower survival function is [51]

$$\underline{S}_{X_{n+1}}(t) = \frac{1}{n+1}\tilde{n}_{t_a^i} \prod_{r: c_r \le t_a^i} \frac{\tilde{n}_{c_r}+1}{\tilde{n}_{c_r}} \tag{2.34}$$

and, for $t \in [x_i, x_{i+1})$ with $i = 1, 2, ..., u$, the NPI upper survival function is [51]

$$\overline{S}_{X_{n+1}}(t) = \frac{1}{n+1}\tilde{n}_{x_i} \prod_{r: c_r \le x_i} \frac{\tilde{n}_{c_r}+1}{\tilde{n}_{c_r}} \tag{2.35}$$

where $\tilde{n}_{t^i_{i*}}$ and $\tilde{n}_{x_i}$ represent the number of observations in the risk set just prior to times t^i_{i*} and x_i, respectively, and $\tilde{n}_{c_r}$ represents the number of observations in the risk set just before time c_r. In Chapter 5, the rc-$A_{(n)}$ method for X_{n+1}, as presented in this section, will be extended to two future observations.

Coolen and Yan [32] compared the NPI lower and upper survival functions based on the rc-$A_{(n)}$ assumption with the Kaplan–Meier estimator (KM), which was reviewed in Section 1.3. They showed that the lower survival function for X_{n+1}, based on the assumption rc-$A_{(n)}$, becomes zero after the largest observation, also the KM estimator will behave this way if that observation is an event time. The upper survival function always remains positive, unless the range of possible values for X_{n+1} is restricted by choosing a finite upper bound [32, 65]. The KM estimate is always equal to one for the first interval $(0, x_1)$. This is also the case for the NPI upper survival function for the event $t \in (0, x_1)$. Remember that the KM estimate only decreases at observed event times. The NPI lower survival function decreases at every observation but the NPI upper survival function decreases only at event times, like the KM. Coolen and Yan [32] claimed that the rc-$A_{(n)}$-based lower and upper survival functions for X_{n+1} are more suitable for graphical presentation rather than the KM-based lower and upper survival functions, as they show the data in full, including right-censored observations, and can be interpreted in a predictive manner [32, 65].

The following example briefly illustrates the NPI lower and upper survival functions and the KM estimator.

Example 2.4.1 (Cervical cancer survival data)

Treatment A data			
90	291	> 890	1153
142	> 468	1037	1297
150	680	> 1090	1429
269	837	> 1113	> 1577

Table 2.1: Cervical cancer survival data (Treatment A) ($> t$ represents right-censoring at time t).

$t_i \in (.,.)$	$\tilde{n}_{t_i}$	d_i	c_i	$(1 - \frac{d_i}{\tilde{n}_{t_i}})$	$\underline{S}_{X_{n+1}}(t)$	$\overline{S}_{X_{n+1}}(t)$	$\hat{S}(t)$
(0,90)	16	0	0	1.000	0.941	1	1
(90,142)	16	1	0	0.938	0.882	0.941	0.938
(142,150)	15	1	0	0.933	0.824	0.882	0.875
(150,269)	14	1	0	0.929	0.765	0.824	0.813
(269,291)	13	1	0	0.923	0.706	0.765	0.750
(291,680)	12	1	1	0.917	0.642	0.706	0.688
(680,837)	10	1	0	0.900	0.578	0.642	0.619
(837,1037)	9	1	1	0.889	0.505	0.578	0.551
(1037,1153)	7	1	2	0.857	0.404	0.505	0.472
(1153,1297)	4	1	0	0.750	0.303	0.404	0.354
(1297,1429)	3	1	0	0.667	0.202	0.303	0.236
(1429,∞)	2	1	1	0.500	0	0.202	0.118

Table 2.2: NPI lower and upper survival functions $\underline{S}_{X_{n+1}}(t)$ and $\overline{S}_{X_{n+1}}(t)$, respectively, and KM estimator $\hat{S}(t)$, for data in Table 2.1.

In this example, we use a subset of data obtained from 183 patients entered into a randomised Phase III trial conducted by the Medical Research Council Working Party on Advanced Carcinoma of the Cervix [48]. Parmar and Machin [48] used 30 observations to illustrate nonparametric methods for survival data, and 14 of these patients received a new therapy known as a radiosensitiser, which was added to their radiotherapy treatment for their cancer and is referred to as Treatment B. The remaining 16 patients received only radiotherapy as a control, referred to as Treatment A. The death of patients due to cancer is the event of interest and the data observations are in days. We only use the 16 patients with Treatment A to illustrate the NPI lower and upper survival functions and compare them to the KM estimator.

The data are presented in Table 2.1, there are eleven event times and five right-censoring times. Table 2.2 presents the NPI lower and upper survival function on the intervals created by the data, together with the KM estimator. These are also plotted in Figure 2.1. The R package 'Survival' is used to derive the KM estimator, see [56].

As shown in Figure 2.1, the NPI lower survival function for X_{17} based on the rc-$A_{(n)}$ method is zero beyond the largest observation, 1577, which is a right-censored observation, while the KM estimate is 0.202 beyond 1577. Also, it is clear from Figure 2.1 that the KM estimator for the first interval $(0, 90)$ is equal to 1. The upper survival function is equal to 1 for interval $(0, 90)$. Figure 2.1 shows that the NPI lower survival function decreases at each observation while the NPI upper survival function and the KM estimate only decrease at observed event times. The results also indicate that KM estimator lies between the NPI lower and upper.

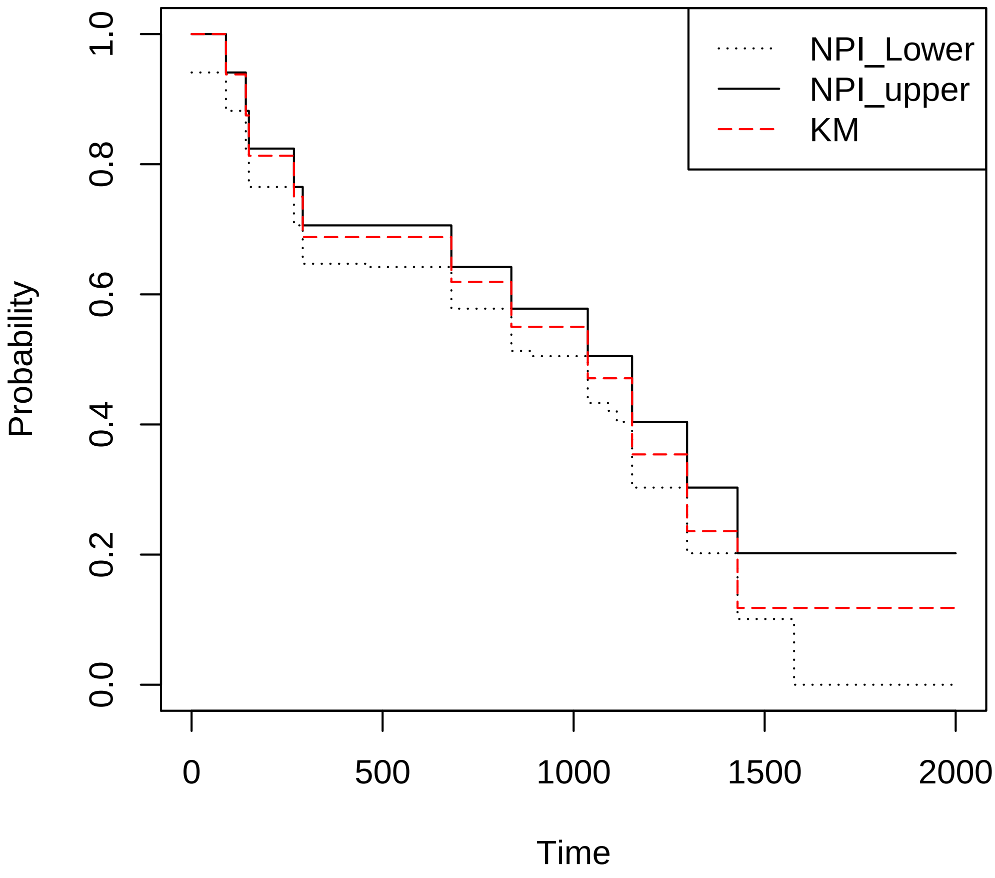

Figure 2.1: NPI lower and upper survival functions and the KM estimate for the data in Table 2.1.

Chapter 3

On Exceedance of the Largest Observed Value

The first contribution to Nonparametric Predictive Inference (NPI) with right-censored data in this thesis has been inspired by the literature on extreme value theory, where in some papers (see e.g. Alves et al [7] and [8]) it is assumed that the support of the random quantities of interest has a finite maximum which is set equal to the largest observation in the available data. This is a somewhat questionable assumption, in particular if the data set contains some right-censored observations, and the question we asked is how likely it is that one or more of the right-censored observations would actually have resulted in a data observation exceeding the largest data value.

This chapter is presented with specific focus on the Supercentenarian data set, which was also used by Alves et al [8]. This data set contains ages at death of people who lived beyond the age of 110, where right-censoring of their death time occurs for those who were still alive at the time the data were collected.

We start this chapter by a brief introduction of Extreme value theory from literature in Section 3.1. In Section 3.2, we consider the exceedance of the largest observed value from NPI perspective. New additions to the study have been presented in Section 3.3 taking into account future observations. The proposed methods, presented in Sections 3.2 and 3.3, are applied to the Supercentenarian data set in Section 3.4. In Section 3.5, we consider the exceedance of the j_{th} largest observations and the results will be applied to the Supercentenarian data set. Finally, this chapter ends up with concluding remarks in Section 3.6.

3.1 Extreme value theory

Extreme value theory (EV) is a statistical method which can be used to draw inferences about rare events from large (or small) samples [7]. Fischer and Tippet [37] obtained three asymptotic limits that describe the distributions of extremes when independent and identically distributed random variables are assumed. We consider the order statistics $X_{n,n} \geq X_{n-1,n} \geq \cdots \geq X_{1,n}$ of i.i.d. random variables $X_1, \ldots, X_n$ with distribution function F. We assume that distribution function F has a finite right endpoint, denoted as x^F, such that $x^F := \sup\{x : F(x) < 1\} \in \mathbb{R}$. For $a_n > 0$ and $b_n \in \mathbb{R}$ such that $\lim_{n\to\infty} F_n(a_n\ x + b_n) = G(x)$, there are three classes of distributions that can occur for the right endpoint, x^F, denoted by Φ_α, Ψ_α and Λ [38], thus

$$\begin{aligned}
&\Phi_\alpha(x) = \exp\{-x^{-\alpha}\},\ x > 0, \quad \alpha > 0,\\
&\Psi_\alpha(x) = \exp\{-(-x)^{\alpha}\},\ x < 0, \quad \alpha > 0,\\
&\Lambda(x) = \exp\{-\exp(-x)\}, \quad x \in \mathbb{R}.
\end{aligned}$$

These distributions can be nested within the Generalized Extreme Value (GEV) distribution with distribution function [7].

$$G_\gamma(x) := \exp\{-(1+\gamma x)^{-1/\gamma}\}, \quad 1 + \gamma x > 0,\ \gamma \in \mathbb{R}. \tag{3.1}$$

To estimate the right endpoint x^F of a distribution function, we tend to use a semi-parametric approach, i.e., the Gumbel extremal domain of attraction [7], rather than using the limiting GEV distribution in Equation (3.1), thus

$$\hat{x}^F := X_{n,n} + \sum_{i=0}^{k-1} a_{i,k}\left(X_{n-k,n} - X_{n-k-i,n}\right)$$

where $a_{i,k} := -\left(\log 2\right)^{-1}\left(\log(k+i) - \log(k+i+1)\right) > 0$, with $\sum_{i=0}^{k-1} a_{i,k} = 1$, and then the $\hat{x}^F$ is defined as a functional of the top observations of the original sample based on an intermediate sequence $k = k_n$, with $k_n \to \infty$, $k_n = O(n)$, as $n \to \infty$. For more information about the Extreme value theory (EV), see [7, 8, 37, 38].

3.2 Exceedance of the largest observed value

Based on the $A_{(n)}$ assumption and non-informative right censoring, described in Chapter 2, this section presents a new method on the exceedance of the largest observed value in a data set containing some right-censored observations that aims to answer the question of whether one or more lifetimes of right-censored observations would exceed the largest observed value.

Let $X_1, X_2, \ldots, X_n$ be non-negative, exchangeable and continuous random quantities. We have observations for these random quantities, such that for u we observed the actual value, these are denoted by $x_1 < x_2 < \cdots < x_u$, while for $v = n-u$ we only observed a right-censoring time, these are denoted by $c_1 < c_2 < \cdots < c_v$. For ease of notation we define $x_0 = 0$ and $x_{u+1} = \infty$. Further, let $X_{c_1}, X_{c_2}, \ldots, X_{c_{v-1}}, X_{c_v}$ be the random quantities corresponding to the individuals whose lifetimes have been right-censored at censoring time c_r, where $r = 1, 2, \ldots, v$. Let $\tilde{n}_{c_r}$ represent the number of observations in the risk set just before time c_r. Let $\mathcal{R} = x_u$, so $\mathcal{R}$ denotes the largest observed event time in the data set. We assume that no ties occur among all observations as reviewed in Section 2.2.

In addition to the assumed exchangeability of all $X_1, X_2, \ldots, X_n$, we assume that, at any right-censoring time, the remaining time to observing the event for a right-censored observation is exchangeable with the remaining times until the event for all other random quantities in the risk set at that time [32, 65]. Based on the assumption of non-informative right censoring [32, 65], discussed in Section 1.2, the assumption shifted-$\tilde{A}_{(n)}$, stated in Definition 2.4.2, is used. The assumption shifted-$\tilde{A}_{(n)}$ allows us to apply $A_{(n)}$ but with the starting point shifted from the value 0 to the observed right-censoring time c_r [32, 65]. Thus, the assumption shifted-$\tilde{A}_{(n)}$ partially specifies the probability distribution for X_{c_r} via the following M-function values:

$$M_{X_{c_r}}(x_i, x_{i+1}) = \frac{1}{\tilde{n}_{c_r} + 1} \qquad \text{for } i = 0, \ldots, u \tag{3.2}$$

where c_r is between (x_i, x_{i+1}) and $\tilde{n}_{c_r}$ is the number of observations in the risk set just prior to time c_r, with $r = 1, 2, \ldots, v$.

Now, consider the event of interest that for at least one of the individuals whose lifetimes have been right-censored, the actual value of the lifetime would be larger than the largest observed value $\mathcal{R}$. For ease of notation, let $G_{\mathcal{R}}(0)$ denote this event of interest. Note that, here we refer to the notation $G_{\mathcal{R}}(0)$ since we do not take any future observations into account, that is $m = 0$, so only the data set that contains n observations is taken into account.

Theorem 3.2.1 The probability for the event of interest $G_{\mathcal{R}}(0)$, denoted by $P(G_{\mathcal{R}}(0))$, is

$$P(G_{\mathcal{R}}(0)) = 1 - \prod_{r=1}^{v} \frac{\tilde{n}_{c_r}}{\tilde{n}_{c_r} + 1} \tag{3.3}$$

where $\tilde{n}_{c_r}$ represents the number of observations in the risk set (still functioning or alive and uncensored) just before time c_r.

Proof: We first consider the individual X_{c_v}, who is the last one who was censored at censoring time c_v, such that there are no further censorings beyond it. Then for X_{c_v}, we can just apply the shifted-$\tilde{A}_{(n)}$, in Equation (3.2), which allows us to apply $A_{(n)}$ but with the starting point shifted from the value 0 to the highest right-censoring time c_v. Then the lifetime of this individual X_{c_v} will either survive the value $\mathcal{R}$ or not. If the lifetime of X_{c_v} would actually be beyond $\mathcal{R}$, then on the basis of the shifted-$\tilde{A}_{(n)}$ as in Equation (3.2), the probability for $X_{c_v} > \mathcal{R}$ is

$$P(X_{c_v} > \mathcal{R}) = \frac{1}{\tilde{n}_{c_v} + 1} \tag{3.4}$$

If the lifetime of X_{c_v} is not going to be beyond $\mathcal{R}$, then the probability for the event of interest $X_{c_v} < \mathcal{R}$ with knowing the value of $\tilde{n}_{c_v}$, is

$$P(X_{c_v} < \mathcal{R}) = 1 - \frac{1}{\tilde{n}_{c_v} + 1} = \frac{\tilde{n}_{c_v}}{\tilde{n}_{c_v} + 1} \tag{3.5}$$

where $\tilde{n}_{c_v}$ is the number of observations in the risk set just prior to time c_v.

We now consider the previous individual with the second censoring time c_{v-1}, namely $X_{c_{v-1}}$, conditional on $X_{c_v} < \mathcal{R}$. It is critical to recognize that for $X_{c_{v-1}}$, it does not matter where exactly the final individual's fail time or lifetime, X_{c_v}, is, as long as it is earlier than $\mathcal{R}$. To be precise, we do not need to take censoring into account for X_{c_v}, because it is conditioning on what is happening before the value $\mathcal{R}$ and thus it must be involved in between X_{c_v} and $\mathcal{R}$, but it does not matter what the exact value of X_{c_v} is within the interval $(X_{c_{v-1}}, \mathcal{R})$. Therefore, the probability for the event that $X_{c_{v-1}}$ exceeds $\mathcal{R}$ given that $X_{c_v} < \mathcal{R}$, on the basis of the shifted-$\tilde{A}_{(n)}$ as in Equation (3.2), is

$$P(X_{c_{v-1}} > \mathcal{R} | X_{c_v} < \mathcal{R}) = \frac{1}{\tilde{n}_{c_{v-1}} + 1} \tag{3.6}$$

and then the probability for the event of interest that $X_{c_{v-1}} < \mathcal{R}$ given $X_{c_v} < \mathcal{R}$, with knowing the value of $\tilde{n}_{c_{v-1}}$, is

$$P(X_{c_{v-1}} < \mathcal{R} | X_{c_v} < \mathcal{R}) = 1 - \frac{1}{\tilde{n}_{c_{v-1}} + 1} = \frac{\tilde{n}_{c_{v-1}}}{\tilde{n}_{c_{v-1}} + 1} \tag{3.7}$$

where $\tilde{n}_{c_{v-1}}$ is the number of observations in the risk set just prior to time c_{v-1}.

The same procedures are repeated for all other individuals whose lifetimes have been right-censored at censoring time c_r, where $r = 1, 2, \ldots, v-3, v-2$. If the lifetime of an individual, X_{c_r}, is not going to be beyond $\mathcal{R}$, then we are going to check the previous individuals at those censoring times c_r. The important thing to note is that for these individuals, it does not matter precisely where their failure of lifetimes occur as long as they have already died before $\mathcal{R}$. Generally, for the lifetime of those later individuals, censoring does not need to be taken into consideration since it is based on what is happening before $\mathcal{R}$. Accordingly, for an individual X_{c_r} at time c_r, we only know the number of individuals in between X_{c_r} and $\mathcal{R}$ and we also know that all of them failed before $\mathcal{R}$. Therefore, the probability for an event that $X_{c_r} > \mathcal{R}$, with $r = 1, 2, \ldots, v-3, v-2$, given that all of the individuals failed before $\mathcal{R}$, on the basis of the shifted-$\tilde{A}_{(n)}$ as in Equation (3.2), are

$$P(X_{c_r} > \mathcal{R} | X_{c_{r+1}} < \mathcal{R}, \ldots, X_{c_{v-1}} < \mathcal{R}, X_{c_v} < \mathcal{R}) = \frac{1}{\tilde{n}_{c_r} + 1} \tag{3.8}$$

and then the probabilities for the event of interest that nobody survives the value $\mathcal{R}$, with knowing the values of $\tilde{n}_{c_r}$, $r = 1, 2, \ldots, v-3, v-2$, are

$$P(X_{c_r} < \mathcal{R} | X_{c_{r+1}} < \mathcal{R}, \ldots, X_{c_{v-1}} < \mathcal{R}, X_{c_v} < \mathcal{R}) = 1 - \frac{1}{\tilde{n}_{c_r} + 1} = \frac{\tilde{n}_{c_r}}{\tilde{n}_{c_r} + 1} \tag{3.9}$$

It is critical to emphasize that for the event of interest above we do not need to apply $A_{(n)}$ with censoring, since it is written as a conditional event that all individuals are less than the value $\mathcal{R}$. Of course, if an individual's lifetime is greater than $\mathcal{R}$, then we know that the event that they are all less than R is not true.

Consequently, the probability for the event of interest $G_{\mathcal{R}}(0)$, denoted by $P(G_{\mathcal{R}}(0))$, is derived in terms of a product of Equation (3.9), the probability for that all events are less than $\mathcal{R}$, thus

$$P(G_{\mathcal{R}}(0)) = 1 - \prod_{r=1}^{v} \frac{\tilde{n}_{c_r}}{\tilde{n}_{c_r} + 1}$$

Thus, the proof is complete.

□

The following example illustrates the probabilities presented in this section.

Example 3.2.1 Suppose we have a data set consisting of $n = 10$ observations. Of these ten individuals, seven died at ages 111, 113, 115, 116, 119, 120 and 122, while three observations were still alive at the time the data were collected, their lifetimes were right-censored at ages 112, 114 and 117. Note that the largest observation that was recorded is 122, so $\mathcal{R} = 122$. Let X_{c_1}, X_{c_2} and X_{c_3} denote the random quantities corresponding to the right-censorings at times 112, 114 and 117, respectively.

We first consider the individual X_{c_3}, who is the last one who was censored at age 117, such that there are no further censorings beyond it. Then the lifetime of X_{c_3} will either survive the value $\mathcal{R}$ or not. If $X_{c_3} > 122$, then on the basis of the shifted-$\tilde{A}_{(3)}$, using Equation (3.4), with $\tilde{n}_{c_3} = 3$ observations in the risk set just prior to time c_3, the probability for $X_{c_3} > \mathcal{R}$ is

$$P(X_{c_3} > 122) = \frac{1}{4}$$

If $X_{c_3} < 122$, then by using Equation (3.5) with $\tilde{n}_{c_3} = 3$, the probability for the event $X_{c_3} < 122$ is $1 - \frac{1}{4} = \frac{3}{4}$.

We now consider X_{c_2}, who was censored at age 114, conditional on $X_{c_3} < 122$. For X_{c_2}, we do not need to take censoring into account for X_{c_3}, since $X_{c_3} < 122$, so it does not matter what the exact value of X_{c3} is within the interval $(117, 122)$, and we only know that there are 3 times of death between 116 and 122. Thus, the probability for the event that $X_{c_2} > 122$ given that $X_{c_3} < 122$, on the basis of the shifted-$\tilde{A}_{(6)}$ using Equation (3.6), with $\tilde{n}_{c_2} = 6$, is

$$P(X_{c_2} > 122 | X_{c_3} < 122) = \frac{1}{7}$$

and then the probability for the event of interest that $X_{c_2} < 122$ given $X_{c_3} < 122$, using Equation (3.7) with $\tilde{n}_{c_2} = 6$, is $1 - \frac{1}{7} = \frac{6}{7}$.

Next, we consider X_{c_1}, who was censored at age 112, conditional on that $X_{c_2} < 122$ and $X_{c_3} < 122$. So, for X_{c_1}, we again do not need to take censoring into account for X_{c_2} and X_{c_3}, since both died before 122, so it does not matter what the exact values of X_{c_2} and X_{c_3} are within the interval $(114, 122)$, and we only know that there are 6 times of death between 113 and 122. Thus, the probability for the event that $X_{c_1} > 122$ given that $X_{c_2} < 122$ and $X_{c_3} < 122$, on the basis of the shifted-$\tilde{A}_{(8)}$ using Equation (3.8), with $\tilde{n}_{c_1} = 8$, is

$$P(X_{c_1} > 122 | X_{c_2} < 122, X_{c_3} < 122) = \frac{1}{9}$$

and then the probability for the event of interest that $X_{c_1} < 122$ given $X_{c_2} < 122$ and $X_{c_3} < 122$, using Equation (3.9) with $\tilde{n}_{c_1} = 8$, is $1 - \frac{1}{9} = \frac{8}{9}$.

Consequently, to calculate the probability for the event that at least one of the three individuals, X_{c_1}, X_{c_2} and X_{c_3}, with lifetimes right-censored at ages 112, 114, and 117, would be larger than the value $\mathcal{R} = 122$, we use Equation (3.3) of Theorem 3.2.1, thus

$$P(G_{122}(0)) = 1 - \prod_{r=1}^{3} \frac{\tilde{n}_{c_r}}{\tilde{n}_{c_r} + 1} = 1 - \left[\frac{3}{4} \times \frac{6}{7} \times \frac{8}{9}\right] = 1 - \frac{4}{7} = 0.4286$$

With this illustrative example of deriving the probability for the event of interest $G_{122}(0)$, we do not have to deal with any censoring in the $A_{(n)}$ setting because we are conditioning on that individuals are all less than the value $\mathcal{R} = 122$.

3.3 New additions to the study by including future items

In this section, we present new additions to the study when taking future items into account. In Section 3.2, we ask what is the probability that anyone, who was censored in the data set, would actually have a real random quantity would exceed the largest value and then in this section, we will include new future individuals. It is important because the main interest in these inferences is not only on the people in the data set, but also on the future individuals, and even when there is no censoring, future ones are still interesting.

In addition to the notation provided in Section 3.2, let $X_{n+1}, X_{n+2}, \ldots, X_{n+m}$ be non-negative, exchangeable and continuous random quantities for the future lifetimes who are including to the given $n = u + v$ data set. Considering $x_0 = 0$, let $\tilde{n}_{x_0} = n$ represent the number of all observations in the risk set at time x_0. Remember that $\mathcal{R}$ denotes the largest observed event time in the data set and we assume that no ties occur among all observations.

Now, consider the event of interest that for at least either one of the individuals whose lifetimes have been right-censored, or one of the $m \geq 1$ future individuals, added to the study, the actual value of the lifetime would be larger than the largest observed value $\mathcal{R}$. For ease of notation, let $G_{\mathcal{R}}(m)$ denote this event of interest. It is important to note that we use the notation $G_{\mathcal{R}}(m)$ for the event of interest that takes into account both future observations as well as the data set consisting of n observations while using the notation $G_{\mathcal{R}}(0)$ for the event of interest, introduced in Section 3.2, which only considers the data set that contains n observations without considering any future observations.

Theorem 3.3.1 Consequently, the probability for the event of interest $G_{\mathcal{R}}(m)$, denoted by $P_{\mathcal{R}}(G(m))$, is

$$P(G_{\mathcal{R}}(m)) = 1 - \left[\prod_{i=1}^{m} \frac{n+i-1}{n+i} \quad \prod_{r=1}^{v} \frac{\tilde{n}_{c_r}}{\tilde{n}_{c_r}+1}\right] = 1 - \left[\frac{n}{n+m} \prod_{r=1}^{v} \frac{\tilde{n}_{c_r}}{\tilde{n}_{c_r}+1}\right] \quad (3.10)$$

where $\tilde{n}_{c_r}$ represents the number of observations in the risk set (still functioning or alive and uncensored) just before time c_r.

Proof: For $m = 1$, we consider the lifetime of the first future individual X_{n+1}, conditional on that all individuals whose lifetimes have been right-censored at censoring time c_r, where $r = 1, 2, \ldots, v$ have been failed before the value $\mathcal{R}$. It is crucial that for all right-censored individuals, it does not matter where exactly their lifetimes are, as long as they are earlier than $\mathcal{R}$. So the only thing that we need to know is the number of individuals in the risk set at time x_0, that is $\tilde{n}_{x_0} = n$. Then the probability for the event that $X_{n+1} > \mathcal{R}$ given that all $X_{c_r} < \mathcal{R}$, $r = 1, 2, \ldots, v$, on the basis of the shifted-$\tilde{A}_{(n)}$ as in Equation (3.2), with $\tilde{n}_{x_0} = n$, is

$$P(X_{n+1} > \mathcal{R}|X_{c_1} < \mathcal{R}, X_{c_2} < \mathcal{R}, \ldots, X_{c_v} < \mathcal{R}) = \frac{1}{\tilde{n}_{x_0} + 1} = \frac{1}{n+1} \tag{3.11}$$

and then the probability for the event of interest that $X_{n+1} < \mathcal{R}$ given all $X_{c_r} < \mathcal{R}$, $r = 1, 2, \ldots, v$, with $\tilde{n}_{x_0} = n$, is

$$P(X_{n+1} < \mathcal{R}|X_{c_1} < \mathcal{R}, X_{c_2} < \mathcal{R}, \ldots, X_{c_v} < \mathcal{R}) = 1 - \frac{1}{n+1} = \frac{n}{n+1} \tag{3.12}$$

In case, $m = 2$, we consider the lifetime of the second future individual X_{n+2}, conditional on that the lifetime of the first future individual X_{n+1} and all individuals whose lifetimes have been right-censored at censoring time c_r, where $r = 1, 2, \ldots, v$, have been failed before the value $\mathcal{R}$. Then again the probability for the event that $X_{n+2} > \mathcal{R}$ given that $X_{n+1} < \mathcal{R}$ and all $X_{c_r} < \mathcal{R}$, $r = 1, 2, \ldots, v$, on the basis of the shifted-$\tilde{A}_{(n)}$ as in Equation (3.2), with $\tilde{n}_{x_0} + 1$, as X_{n+1} is added, so $\tilde{n}_{x_0} + 1 = n + 1$, thus

$$P(X_{n+2} > \mathcal{R}|X_{n+1} < \mathcal{R}, X_{c_1} < \mathcal{R}, \ldots, X_{c_v} < \mathcal{R}) = \frac{1}{(\tilde{n}_{x_0} + 1) + 1} = \frac{1}{n+2}$$

and then the probability for the event of interest that $X_{n+2} < \mathcal{R}$ given $X_{n+1} < \mathcal{R}$ and all $X_{c_r} < \mathcal{R}$, $r = 1, 2, \ldots, v$, with $\tilde{n}_{x_0} + 1 = n + 1$, is

$$P(X_{n+2} < \mathcal{R}|X_{n+1} < \mathcal{R}, X_{c_1} < \mathcal{R}, \ldots, X_{c_v} < \mathcal{R}) = 1 - \frac{1}{n+2} = \frac{n+1}{n+2}$$

In general, we derive the probability for an event $X_{n+i} > \mathcal{R}$, when $i = 2, 3, \ldots, m$, conditional on that the previous future individuals and all individuals whose lifetimes have been right-censored at censoring time c_r, where $r = 1, 2, \dots, v$, have been failed before the value $\mathcal{R}$, on the basis of the shifted-$\tilde{A}_{(n)}$ as in Equation (3.2), as follows.

$$P(X_{n+i} > \mathcal{R} | X_{n+1} < \mathcal{R}, \dots, X_{n+i-1} < \mathcal{R}, X_{c_1} < \mathcal{R}, \dots, X_{c_v} < \mathcal{R}) = \frac{1}{(\tilde{n}_{x_0} + i - 1) + 1} = \frac{1}{n+i} \tag{3.13}$$

and then the probability for an event $X_{n+i} < \mathcal{R}$ given $X_{n+1} < \mathcal{R}, \dots, X_{n+i-1} < \mathcal{R}$, when $i = 1, 2, \dots, m$, and all $X_{c_r} < \mathcal{R}$, $r = 1, 2, \dots, v$, with $\tilde{n}_{x_0} + i = n + i$, is

$$P(X_{n+i} < \mathcal{R} | X_{n+1} < \mathcal{R}, \dots, X_{n+i-1} < \mathcal{R}, X_{c_1} < \mathcal{R}, \dots, X_{c_v} < \mathcal{R}) = 1 - \frac{1}{n+i} = \frac{n+i-1}{n+i} \tag{3.14}$$

Since the event of interest $G_{\mathcal{R}}(m)$ takes into account both future observations and the data set that contains the n observations, Equation (3.9) that is related to the data set that contains only the n observations without considering any future observations and Equation (3.14) that is related to future observations, are required to compute the probability that all right-censored times exceed $\mathcal{R}$, stated for the event of interest $G_{\mathcal{R}}(m)$. Consequently, the probability for the event of interest $G_{\mathcal{R}}(m)$, denoted by $P_{\mathcal{R}}(G(m))$, in terms of products of Equations (3.9) of Theorem 3.2.1 and (3.14), will be

$$P(G_{\mathcal{R}}(m)) = 1 - \left[\prod_{i=1}^{m} \frac{n+i-1}{n+i} \quad \prod_{r=1}^{v} \frac{\tilde{n}_{c_r}}{\tilde{n}_{c_r} + 1} \right] = 1 - \left[\frac{n}{n+m} \prod_{r=1}^{v} \frac{\tilde{n}_{c_r}}{\tilde{n}_{c_r} + 1} \right]$$

Thus, the proof is complete.

□

With respect to Equation (3.10) of Theorem 3.3.1, $P(G_{\mathcal{R}}(m))$ value increases as m increases. In addition, if m tends to infinity, the second term in Equation (3.10) tends to zero, so that the value of $P(G_{\mathcal{R}}(m))$ tends to 1.

The following example illustrates the probabilities presented in this section.

Example 3.3.1 We again use the same data on $n = 10$ observations (as in Example 3.2.1). We consider that X_{n+1} and X_{n+2} are lifetimes of the first and second future ones to be included to the study and then we ask that what is the probability for the event that for at least either one of the three individuals, X_{c_1}, X_{c_2} and X_{c_3}, with lifetimes right-censored at ages 112, 114, and 117, or one of the future individuals, X_{n+1} and X_{n+2}, the actual value of the lifetime would be larger than the largest observed value $\mathcal{R}$.

We first consider the lifetime of X_{n+1}, conditional on that X_{c_1}, X_{c_2} and X_{c_3}, with lifetimes right-censored at ages 112, 114, and 117, have been failed before the value $\mathcal{R} = 122$. Then the probability for the event that $X_{n+1} > 122$ given that $X_{c_1} < 122, X_{c_2} < 122$ and $X_{c_3} < 122$, on the basis of the shifted-$\tilde{A}_{(10)}$, with the $\tilde{n}_{x_0} = 10$ individuals in the risk set at time x_0, is calculated by using Equation (3.11) as

$$P(X_{n+1} > 122 | X_{c_1} < 122, X_{c_2} < 122, X_{c_3} < 122) = \frac{1}{11}$$

and then the probability for the event $X_{n+1} < 122$ given that $X_{c_1} < 122, X_{c_2} < 122$ and $X_{c_3} < 122$, with $\tilde{n}_{x_0} = 10$, is calculated by using Equation (3.12) as

$$P(X_{n+1} < 122 | X_{c_1} < 122, X_{c_2} < 122, X_{c_3} < 122) = 1 - \frac{1}{11} = \frac{10}{11}$$

We now consider the lifetime of X_{n+2}, conditional on that the lifetime of X_{n+1} and X_{c_1}, X_{c_2} and X_{c_3}, with lifetimes right-censored at ages 112, 114, and 117, have been failed before the value $\mathcal{R} = 122$. Then the probability for the event $X_{n+2} > 122$ given that $X_{n+1} < 122$, $X_{c_1} < 122, X_{c_2} < 122$ and $X_{c_3} < 122$, on the basis of the shifted-$\tilde{A}_{(11)}$, with the $\tilde{n}_{x_0} + 1 = 11$, is calculated by using Equation (3.11) as

$$P(X_{n+2} > 122 | X_{n+1} < 122, X_{c_1} < 122, X_{c_2} < 122, X_{c_3} < 122) = \frac{1}{12}$$

with respect to that for $X_{n+1} < 122$, $X_{c_1} < 122, X_{c_2} < 122$ and $X_{c_3} < 122$, it does not matter where exactly their lifetimes are, as long as they died before 122. Then the probability for the event $X_{n+2} < 122$ given $X_{n+1} < 122$, $X_{c_1} < 122, X_{c_2} < 122$ and $X_{c_3} < 122$, with $\tilde{n}_{x_0} + 1 = 11$, is

$$P(X_{n+2} < 122 | X_{n+1} < 122, X_{c_1} < 122, X_{c_2} < 122, X_{c_3} < 122) = 1 - \frac{1}{12} = \frac{11}{12}$$

Consequently, by using Theorem 3.3.1, the probability for the event $G_{122}(2)$, denoted by $P(G_{122}(2))$, in terms of products of Equations (3.9) and (3.14), is derived as

$$P(G_{122}(2)) = 1 - \left[\prod_{i=1}^{2} \frac{n+i-1}{n+i} \prod_{r=1}^{3} \frac{\tilde{n}_{c_r}-1}{\tilde{n}_{c_r}}\right]$$
$$= 1 - \left[(\frac{10}{11} \times \frac{11}{12}) \times (\frac{3}{4} \times \frac{6}{7} \times \frac{8}{9})\right] = 1 - \frac{40}{84} = 0.5238$$

In the following section, the proposed method presented in Sections 3.2 and 3.3 will be applied to the full supercentenarian data set, but separately for the women and the men.

3.4 Application to the Supercentenarian data

The data set considered in this chapter was used by Alves et al. [8], it contains the ages at death of 1740 people who had lived past the age of 110, together with the ages of such people who were still alive when the data were collected. These data were taken from Tables B and C of the dataset that had been collected by the Gerontology Research Group (GRG), on 22 April 2018.* The lifespan of individuals indicated in this data set contains the number of years a person may be able to live. One interesting thing is an inference on the maximum number of years that is possible for a person to live. For the sake of computation, the ages in the data set are in days but in here, the number of years as well as days are needed and we assume that there are no ties occur between these ages. It should be noted that all years were interpreted as 365 days, and that leap years were not taken into account. According to the GRG, Jeanne Calment, a supercentenarian woman from France, has the oldest recorded age of a person, which is 122.5 years old, whereas Jiroemon Kimura from Japan has the oldest recorded age of a supercentenarian man, which is 116.2 years old.

The data set consists of 1580 lifetimes of supercentenarian females and 160 lifetimes of supercentenarian males, who exceeded the age of 110. These numbers also

*The data set is available at *http://www.grg.org/Adams/Tables.htm*, see also [8].

reflect that supercentenarian women tend to live longer than supercentenarian men. Of these 1580 supercentenarian women, 72 were still alive on 22 April 2018, the date of collecting the data, so their lifetimes are considered to be right-censored. In contrast, there were only two supercentenarian males still alive on the date of collecting the data, out of the 160 supercentenarian males in the study. In this study, the objective is to determine the probability for the event that at least one of the right-censored supercentenarians women would live beyond Jeanne Calment's age at death, and the event that at least one of the right-censored supercentenarian men would live beyond Jiroemon Kimura's age at death. In the following examples, the methods presented in Sections 3.2 and 3.3 will be applied to the supercentenarian men and women separately.

Example 3.4.1 (Supercentenarian women data) In this example, we consider the supercentenarian data for women. There are $n = 1580$ supercentenarian women in the data set, 72 were still alive at the time of the study and hence their lifetimes are right-censored. As Jeanne Calment's age was the largest age recorded in the data set with the age of 122.5, let $\mathcal{R} = 122.5$. Then the interest will be on the question of what is the probability for the event, $G_{\mathcal{R}}(0)$, that at least one of the 72 supercentenarian women whose lifetimes have been right-censored, the actual value of the lifetime would be larger than the largest observed value $\mathcal{R} = 122.5$. This probability is obtained by using Equation (3.3) of Theorem 3.2.1, as follows

$$P(G_{122.5}(0)) = 1 - \prod_{r=1}^{72} \frac{\tilde{n}_{c_r}}{\tilde{n}_{c_r} + 1} = 1 - 0.6567 = 0.3433$$

As a result of what we have assumed in our model, which is based on the $A_{(n)}$ assumption and non-informative right censoring, described in Section 3.2, the probability that at least a lifetime of one of the 72 supercentenarian women, who were still alive at the time of data set, would live beyond Jeanne Calment's age ($\mathcal{R} = 122.5$), is 0.3433.

Now, let us consider $m = 1$ future supercentenarian women, X_{n+1}, added to the study, given the $n = 1580$ supercentenarian women. The lifetime of X_{n+1} is considered, conditional on that all the 72 supercentenarian women, whose lifetimes have been right-censored, have been failed before the value $\mathcal{R} = 122.5$. Then the probability for the event, $G_{122.5}(1)$, that at least one of the 72 right-censored supercentenarian women or the lifetime of X_{n+1}, would live longer than $\mathcal{R} = 122.5$, using Equation (3.10) of Theorem 3.3.1, is

$$P(G_{122.5}(1)) = 1 - \left[\frac{1580}{1580+1} \prod_{r=1}^{72} \frac{\tilde{n}_{c_r}}{\tilde{n}_{c_r}+1}\right] = 1 - 0.6563 = 0.3437$$

Consider $m = 2$ future supercentenarian women, X_{n+2}, added to the study, given the $n = 1580$ supercentenarian women and the first future supercentenarian women, X_{n+1}. The lifetime of X_{n+2} is considered, conditional on that all the 72 supercentenarian women, whose lifetimes have been right-censored, and the lifetime of X_{n+1}, all have been failed before the value $\mathcal{R} = 122.5$. Then the probability for the event, $G_{122.5}(2)$, that at least one of the 72 right-censored supercentenarian women or one of the lifetimes of X_{n+1} and X_{n+2}, would live longer than $\mathcal{R} = 122.5$, using Equation (3.10) of Theorem 3.3.1, is

$$P(G_{122.5}(2)) = 1 - \left[\frac{1580}{1580+2} \prod_{r=1}^{72} \frac{\tilde{n}_{c_r}}{\tilde{n}_{c_r}+1}\right] = 1 - 0.6559 = 0.3441$$

Considering $m \geq 2$ future supercentenarian women to be added to the study, then the probability for the event, $G_{122.5}(m)$, that at least one of the 72 right-censored supercentenarian women or one of the lifetimes of $m \geq 2$ future supercentenarian women, would live longer than $\mathcal{R} = 122.5$, is calculated by using Equation (3.10) of Theorem 3.3.1, and shown in Figure 3.1, thus

$$P(G_{122.5}(m)) = 1 - \left[\frac{1580}{1580+m} \prod_{r=1}^{72} \frac{\tilde{n}_{c_r}}{\tilde{n}_{c_r}+1}\right] = 1 - \left[\frac{1580}{1580+m} \times 0.6567\right]$$

Remark: As $m \to \infty$, the probability $P(G_{122.5}(m)) \to 1$.

Example 3.4.2 (Supercentenarian men data) In this example, we consider the supercentenarian data for men. The data set consists of 160 supercentenarian men, two of them were still alive at the time of the study and hence their lifetimes are right-censored. As Jiroemon Kimura's age was the largest age recorded in the data set with the age of 116.2, let $\mathcal{R} = 116.2$. Then the interest will be on the question of what is the probability for the event, $G_{116.2}(0)$, that at least one of the two supercentenarian men whose lifetimes have been right-censored, the actual value of the

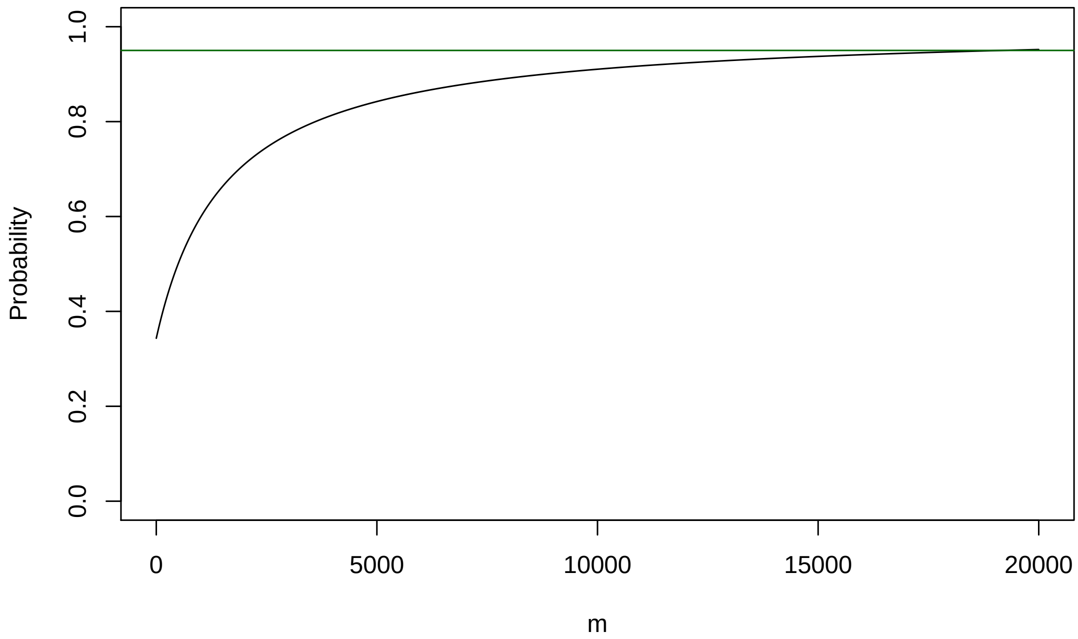

Figure 3.1: $P(G_{122.5}(m))$ for the supercentenarian women data, according to Example 3.4.1.

lifetime would be larger than the largest observed value $\mathcal{R} = 116.2$. This probability is obtained by using Equation (3.3) of Theorem 3.2.1, as follows

$$P(G_{116.2}(0)) = 1 - \prod_{r=1}^{2} \frac{\tilde{n}_{c_r}}{\tilde{n}_{c_r} + 1} = 1 - 0.9444 = 0.0556$$

so with probability of 0.0556, there is at least one lifetime of the two supercentenarian men, who were still alive at the time of data set, would live beyond Jiroemon Kimura's age ($\mathcal{R} = 116.2$).

Take $m = 1$ future supercentenarian men, X_{n+1}, added to the study, into considration, given the $n = 160$ supercentenarian men. The lifetime of X_{n+1} is considered, conditional on that all the two supercentenarian men, whose lifetimes have been right-censored, have been failed before the value $\mathcal{R} = 116.2$. Then the probability for the event, $G_{116.2}(1)$, that at least one of the two right-censored supercentenarian men or the lifetime of X_{n+1}, would live longer than $\mathcal{R} = 116.2$, using Equation

(3.10) of Theorem 3.3.1, is

$$P(G_{116.2}(1)) = 1 - \left[\frac{160}{160+1} \prod_{r=1}^{2} \frac{\tilde{n}_{c_r}}{\tilde{n}_{c_r}+1}\right] = 1 - 0.9385 = 0.0615$$

Then we consider $m = 2$ future supercentenarian men, X_{n+2}, added to the study, given the $n = 160$ supercentenarian men and the first future supercentenarian men, X_{n+1}. The lifetime of X_{n+2} is considered, conditional on that all the two supercentenarian men, whose lifetimes have been right-censored, and the lifetime of X_{n+1}, all have been failed before the value $\mathcal{R} = 116.2$. Then the probability for the event, $G_{116.2}(2)$, that at least one of the two right-censored supercentenarian men or one of the lifetimes of X_{n+1} and X_{n+2}, would live longer than $\mathcal{R} = 116.2$, using Equation (3.10) of Theorem 3.3.1, is

$$P(G_{116.2}(2)) = 1 - \left[\frac{160}{160+2} \prod_{r=1}^{2} \frac{\tilde{n}_{c_r}}{\tilde{n}_{c_r}+1}\right] = 1 - 0.9327 = 0.0673$$

Considering $m \geq 2$ future supercentenarian men to be added to the study, then the probability for the event, $G_{116.2}(m)$, that at least one of the two right-censored supercentenarian men or one of the lifetimes of $m \geq 2$ future supercentenarian men, would live longer than $\mathcal{R} = 116.2$, is calculated by using Equation (3.10) of Theorem 3.3.1, and shown in Figure 3.2, thus

$$P(G_{116.2}(m)) = 1 - \left[\frac{160}{160+m} \prod_{r=1}^{2} \frac{\tilde{n}_{c_r}}{\tilde{n}_{c_r}+1}\right] = 1 - \left[\frac{160}{160+m} \times 0.9444\right]$$

One interesting argument is that one would get the smallest m for which value of the probability $P(G_{\mathcal{R}}(m))$ in Example 3.4.1 or in Example 3.4.2, is greater than a specific P value, where $P \in [0, 1]$, i.e. $P = 0.95$. For example, from Figure 3.1, for the supercentenarian women data, according to Example 3.4.1, such that

$$P(G_{122.5}(m)) = 1 - \left[\frac{1580}{1580+m} \times 0.6567\right] > P$$

where the $P(G_{122.5}(m))$ is greater than $P = 0.95$ for $m \geq 19200$ future supercentenarian women.

While from Figure 3.2, for the supercentenarian men data, according to Example 3.4.2, such that

$$P(G_{116.2}(m)) = 1 - \left[\frac{160}{160+m} \times 0.9444\right] > P$$

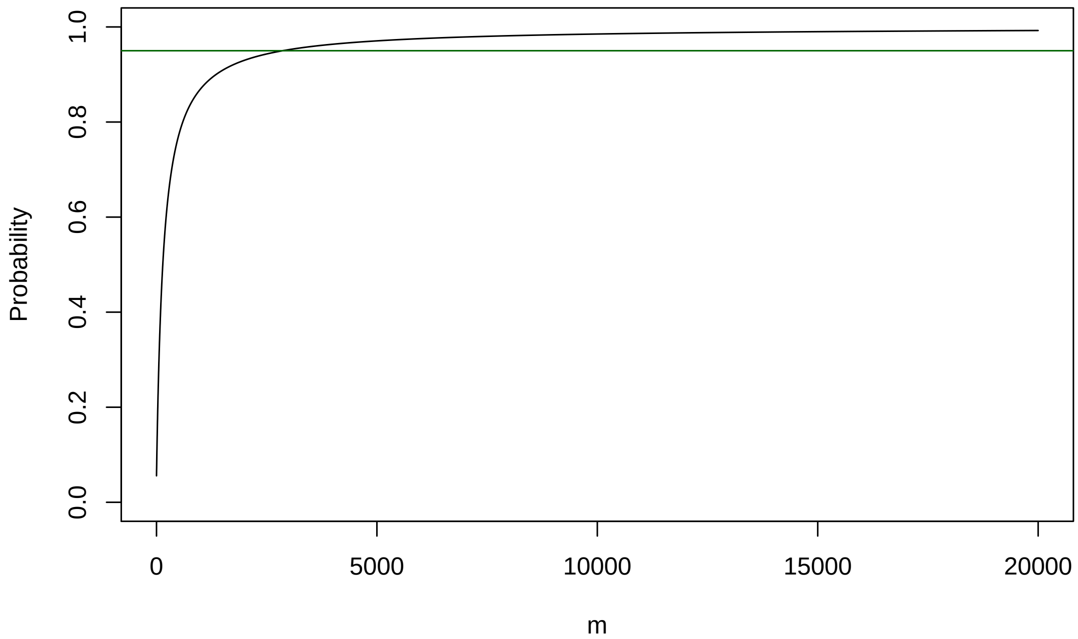

Figure 3.2: $P(G_{116.2}(m))$ for the supercentenarian men data, according to Example 3.4.2.

where the $P(G_{116.2}(m))$ is greater than $P = 0.95$ for $m \geq 2900$ future supercentenarian men.

Note that for the supercentenarian men data in Example 3.4.2, the standard of study centered on whether they survived or not the age of the oldest male 116.2, while for the supercentenarian women data in Example 3.4.1, it is on whether they survived or not the age of the oldest female 122.5, because there was no supercentenarian man has lived to be older than 116.2.

3.5 Exceedance of the j_{th} largest observations

In Sections 3.2 and 3.3, we took into account the exceedance of the largest observation, $\mathcal{R}$, with right-censored data and in this section we take into account the exceedance of the second, third, fourth, ... j_{th} largest observations, as long as they are greater than the largest censored observation, X_{c_v}. Then, we look at time t in between any two of those largest observations and we find lower and upper probabilities for the exceedance of this time t.

We use the same notation as in Sections 3.2 and 3.3, in addition to a few additional notation. In order to simplify notation, the notation $\mathcal{R}$, provided in Sections 3.2 and 3.3, which is $\mathcal{R} = x_u$, should now be written as $\mathcal{R}_1 = x_u$ to represent the first largest event time in the data set. Also, let $\mathcal{R}_2 = x_{u-1}$, so $\mathcal{R}_2$ denotes the second largest observed event time in the data set, and let $\mathcal{R}_3 = x_{u-2}$, so $\mathcal{R}_3$ denotes the third largest observed event time in the data set, up to the largest observed event time just beyond the time c_v, such that $\mathcal{R}_1 > \mathcal{R}_2 > \mathcal{R}_3 > \ldots > \mathcal{R}_j$ corresponding to $x_u > x_{u-1} > x_{u-2} > \ldots > x_{u-i}$, as long as $x_{u-i} > c_v$, where $i = 0, 1, \ldots, u$, and $j = 1, \ldots, u$. Remember that $\tilde{n}_{c_r}$, $r = 1, 2, \ldots, v$, represent the number of observations in the risk set just before time c_r. We assume that no ties occur among all observations as stated in Section 2.2. In addition, the method introduced in this section will be based on the shifted-$\tilde{A}_{(n)}$ as in Equation (3.2), under the exchangeability assumption discussed in Section 3.2, and the assumption of non-informative right censoring discussed in Section 1.2 [32, 65].

Now, we take into account the second largest observed value in the data set, $\mathcal{R}_2 = x_{u-1}$, as long as there are no censorings past it, so that $x_{u-1} > c_v$. We consider the event of interest that for at least one of the individuals whose lifetimes have been right-censored, the actual value of the lifetime would be larger than the second largest observed value $\mathcal{R}_2$. For ease of notation, let $G_{\mathcal{R}_2}(0)$ denote this event of interest. Then we find the probability for the event $G_{\mathcal{R}_2}(0)$ as we did for the event $G_{\mathcal{R}}(0)$ for exceeding the first largest observation, in Section 3.2.

For those individuals whose lifetimes have been right-censored at censoring time c_r, where $r = 1, 2, \ldots, v$, censoring does not need to be taken into consideration as long as they all failed before $\mathcal{R}_2$, and we only know the number of individuals in between X_{c_r} and $\mathcal{R}_2$. On the basis of the shifted-$\tilde{A}_{(n)}$ as in Equation (3.2), then $P_{X_{c_r}}(\mathcal{R}_2, \mathcal{R}) = P_{X_{c_r}}(\mathcal{R}, \infty) = \frac{1}{\tilde{n}_{c_r}+1}$, where $\tilde{n}_{c_r}$ is the number of observations in the risk set just prior to time c_r, with $r = 1, 2, \ldots, v$. Therefore, the probability for the event $X_{c_r} > \mathcal{R}_2$, with $r = 1, 2, \ldots, v$, given that all of the individuals failed before

$\mathcal{R}_2$, are

$$P(X_{c_r} > \mathcal{R}_2 | X_{c_{r+1}} < \mathcal{R}_2, \ldots, X_{c_{v-1}} < \mathcal{R}_2, X_{c_v} < \mathcal{R}_2) = \frac{2}{\tilde{n}_{c_r} + 1} \tag{3.15}$$

and then the probability for the event of interest that nobody survives the value $\mathcal{R}_2$, with knowing the values of $\tilde{n}_{c_r}$, $r = 1, 2, \ldots, v$, are

$$P(X_{c_r} < \mathcal{R}_2 | X_{c_{r+1}} < \mathcal{R}_2, \ldots, X_{c_{v-1}} < \mathcal{R}_2, X_{c_v} < \mathcal{R}_2) = 1 - \frac{2}{\tilde{n}_{c_r} + 1} = \frac{\tilde{n}_{c_r} - 1}{\tilde{n}_{c_r} + 1} \tag{3.16}$$

Consequently, the probability for the event of interest $G_{\mathcal{R}_2}(0)$, denoted by $P(G_{\mathcal{R}_2}(0))$, in terms of a product of Equation (3.16), is derived as

$$P(G_{\mathcal{R}_2}(0)) = 1 - \prod_{r=1}^{v} \frac{\tilde{n}_{c_r} - 1}{\tilde{n}_{c_r} + 1} \tag{3.17}$$

Using the same logic as above, we obtain the probability for the event of interest $G_{\mathcal{R}_3}(0)$, that is for at least one of the individuals whose lifetimes have been right-censored, the actual value of the lifetime would be larger than the third largest observed value $\mathcal{R}_3 = x_{u-2}$, where $x_{u-2} > c_v$, thus

$$P(G_{\mathcal{R}_3}(0)) = 1 - \prod_{r=1}^{v} \frac{\tilde{n}_{c_r} - 2}{\tilde{n}_{c_r} + 1} \tag{3.18}$$

Similar to the above explanation, one could straightforwardly obtain any probability for the event that at least one of the individuals, whose lifetimes have been right-censored, the actual value of the lifetime would be larger than any other largest observed value, as long as it is greater than the largest censored observation at c_v.

We are now considering including future items to the study, as we did for Section 3.3. Then we consider the event of interest that for at least either one of the individuals whose lifetimes have been right-censored, or one of the $m \geq 1$ future individuals, added to the study, the actual value of the lifetime would be larger than the second largest observed value $\mathcal{R}_2 = x_{u-1}$, where $x_{u-1} > c_v$. For ease of notation, let $G_{\mathcal{R}_2}(m)$ denote this event of interest.

The probability for the event $X_{n+i} > \mathcal{R}_2$, when $i = 1, 2, \ldots, m$, conditional on that the previous future individuals and all individuals whose lifetimes have been right-censored at censoring time c_r, where $r = 1, 2, \ldots, v$, have been failed before the value $\mathcal{R}_2 = x_{u-1}$, $x_{u-1} > c_v$, is derived on the basis of the shifted-$\tilde{A}_{(n)}$ in Equation (3.2), as follows.

$$P(X_{n+i} > \mathcal{R}_2 | X_{n+1} < \mathcal{R}_2, \ldots, X_{n+i-1} < \mathcal{R}_2, X_{c_1} < \mathcal{R}_2, \ldots, X_{c_v} < \mathcal{R}_2)$$
$$= \frac{1}{n+i} + \frac{1}{n+i} = \frac{2}{n+i} \tag{3.19}$$

Then the probability for an event $X_{n+i} < \mathcal{R}_2$ given $X_{n+1} < \mathcal{R}_2, \ldots, X_{n+i-1} < \mathcal{R}_2$, when $i = 1, 2, \ldots, m$, and all $X_{c_r} < \mathcal{R}_2$, $r = 1, 2, \ldots, v$, with $n+i$, is

$$P(X_{n+i} < \mathcal{R}_2 | X_{n+1} < \mathcal{R}_2, \ldots, X_{n+i-1} < \mathcal{R}_2, X_{c_1} < \mathcal{R}_2, \ldots, X_{c_v} < \mathcal{R}_2)$$
$$= 1 - \frac{2}{n+i} = \frac{n+i-2}{n+i} \tag{3.20}$$

Consequently, the probability for the event of interest $G_{\mathcal{R}_2}(m)$, denoted by $P(G_{\mathcal{R}_2}(m))$, in terms of products of Equations (3.16) and (3.20), is derived as

$$\begin{aligned} P(G_{\mathcal{R}_2}(m)) &= 1 - \left[\prod_{i=1}^{m} \frac{n+i-2}{n+i} \prod_{r=1}^{v} \frac{\tilde{n}_{c_r} - 1}{\tilde{n}_{c_r} + 1} \right] \\ &= 1 - \left[\frac{n(n-1)}{(n+m)(n+m-1)} \prod_{r=1}^{v} \frac{\tilde{n}_{c_r} - 1}{\tilde{n}_{c_r} + 1} \right] \end{aligned} \tag{3.21}$$

Using the same logic as above, we obtain the probability for the event of interest $G_{\mathcal{R}_3}(m)$, that is for at least either one of the individuals whose lifetimes have been right-censored, or one of the $m \geq 1$ future individuals, added to the study, the actual value of the lifetime would be larger than the third largest observed value $\mathcal{R}_3 = x_{u-2}$, where $x_{u-2} > c_v$, as follows

$$\begin{aligned} P(G_{\mathcal{R}_3}(m)) &= 1 - \left[\prod_{i=1}^{m} \frac{n+i-3}{n+i} \prod_{r=1}^{v} \frac{\tilde{n}_{c_r} - 2}{\tilde{n}_{c_r} + 1} \right] \\ &= 1 - \left[\frac{n(n-1)(n-2)}{(n+m)(n+m-1)(n+m-2)} \prod_{r=1}^{v} \frac{\tilde{n}_{c_r} - 2}{\tilde{n}_{c_r} + 1} \right] \end{aligned} \tag{3.22}$$

Similar to the above explanation, one could straightforwardly obtain any probability for the event that for at least either one of the individuals whose lifetimes have been right-censored, or one of the $m \geq 1$ future individuals, added to the study, the actual value of the lifetime would be larger than any other largest observed value, as long as it is greater than the largest censored observation at c_v. Consequently, the probability that somebody would survive any largest observed value recorded in a data set, when it exceeds the largest censored observation, increases when it is calculated backwards from the largest recorded value to the j_{th} largest observed value as long as it past the largest censored observation.

Now, consider the event of interest that for at least either one of the individuals whose lifetimes have been right-censored, or one of the $m \geq 1$ future individuals, added to the study, the actual value of the lifetime would be larger than time t, where t is in between any two consecutive largest observed values, i.e. x_i and x_{i+1}, were $i = 0, 1, \ldots, u$, as long as the x_i is greater than the largest censored observation c_v. For ease of notation, let $G_{t\in(x_i,x_{i+1})}(m)$ denote this event of interest. Then for $t \in (x_i, x_{i+1})$, the upper survival of such a t is the survival of the previous one x_i and the lower survival of such a t is the survival of the next one x_{i+1}. So, the lower probability for the event $G_{t\in(x_i,x_{i+1})}(m)$ is the probability for the event $G_{x_{i+1}}(m)$, that is $\underline{P}(G_{t\in(x_i,x_{i+1})}(m)) = P(G_{x_{i+1}}(m))$, where x_{i+1} is the next largest observation. And the corresponding upper probability for the event $G_{t\in(x_i,x_{i+1})}(m)$ is the probability for the event $G_{x_i}(m)$, that is $\overline{P}(G_{t\in(x_i,x_{i+1})}(m)) = P(G_{x_i}(m))$, where x_i is the previous largest observation. For example, if $t \in (x_{u-1}, x_u)$, where $\mathcal{R}_1 = x_u$ and $\mathcal{R}_2 = x_{u-1}$ representing to the first and second largest event value in the data set and $x_{u-1} > c_v$, then the lower survival of t is that $\underline{P}(G_{t\in(\mathcal{R}_2,\mathcal{R}_1)}(m)) = P(G_{\mathcal{R}_1}(m))$, which is obtained by applying to Equation (3.10). And the corresponding upper survival of t is that $\overline{P}(G_{t\in(\mathcal{R}_2,\mathcal{R}_1)}(m)) = P(G_{\mathcal{R}_2}(m))$, which is obtained by applying to Equation (3.21).

The following examples illustrate the proposed method presented in this section, which will be applied to the full supercentenarian data set, but separately for the women and the men.

Example 3.5.1 (Supercentenarian women data) In this example, we again use the same data on $n = 1580$ supercentenarian women (as in Example 3.4.1). There were 72 supercentenarian women still alive at the time of the study and hence their lifetimes are right-censored. Also, there are eight supercentenarian women whose ages exceed the largest censored supercentenarian women at age 117.1. Previously, in Example 3.4.1, the first largest age recorded, $\mathcal{R}_1 = 122.5$, were considered, and now we are considering the second and the third largest ages recorded, $\mathcal{R}_2 = 119.3$ and $\mathcal{R}_3 = 117.8$, respectively.

The interest now will be on the question of what is the probability for the event, $G_{\mathcal{R}_2}(0)$, that at least one of the 72 supercentenarian women whose lifetimes have been right-censored, the actual value of the lifetime would be larger than the second largest observed value $\mathcal{R}_2 = 119.3$. This probability is obtained by using Equation (3.17), as follows

$$P(G_{119.3}(0)) = 1 - \prod_{r=1}^{72} \frac{\tilde{n}_{c_r-1}}{\tilde{n}_{c_r}+1} = 1 - 0.4228 = 0.5772$$

The probability for the event, $G_{\mathcal{R}_3}(0)$, that at least one of the 72 supercentenarian women whose lifetimes have been right-censored, the actual value of the lifetime would be larger than the third largest observed value $\mathcal{R}_3 = 117.8$. This probability is obtained by using Equation (3.18), as follows

$$P(G_{117.8}(0)) = 1 - \prod_{r=1}^{72} \frac{\tilde{n}_{c_r-1}}{\tilde{n}_{c_r}+1} = 1 - 0.2655 = 0.7345$$

As a result of what we have assumed in our model, which is based on the $A_{(n)}$ assumption and non-informative right censoring, described in this section, the probability that at least a lifetime of one of the 72 supercentenarian women, who were still alive at the time of data set, would live beyond the second largest observed age ($\mathcal{R}_2 = 119.3$), is 0.5772, and this probability increased to reach 0.7345 in case of surviving the third largest observed age ($\mathcal{R}_3 = 117.8$). Moreover, it is more likely that somebody would survive any one of the eighth supercentenarian women, as long as it is calculated backwardly from the first largest age to the eighth largest age that exceeds the largest censored supercentenarian women at age 117.1.

Now, let consider $m = 2$ future supercentenarian women, X_{n+2}, added to the study, given the $n = 1580$ supercentenarian women and the first future supercentenarian women, X_{n+1}. The lifetime of X_{n+2} is considered, conditional on that all the 72 supercentenarian women, whose lifetimes have been right-censored, and the lifetime of X_{n+1}, all have been failed before the value $\mathcal{R}_2 = 119.3$. Then the probability for the event, $G_{\mathcal{R}_2}(2)$, that at least one of the 72 right-censored supercentenarian women or one of the lifetimes of X_{n+1} and X_{n+2}, would live longer than the age

$\mathcal{R}_2 = 119.3$, using Equation (3.21), is

$$P(G_{119.3}(2)) = 1 - \left[\frac{1580(1579)}{(1580+2)(1580+1)} \prod_{r=1}^{72} \frac{\tilde{n}_{c_r} - 1}{\tilde{n}_{c_r} + 1} \right] = 1 - 0.4217 = 0.5783$$

Taking into account the survival of the third largest age $\mathcal{R}_3 = 117.8$, with considering $m = 2$ future supercentenarian women, then the lifetime of X_{n+2} is considered, conditional on that all the 72 supercentenarian women, whose lifetimes have been right-censored, and the lifetime of X_{n+1}, all have been failed before the value $\mathcal{R}_3 = 117.8$. Then the probability for the event, $G_{\mathcal{R}_3}(2)$, that at least one of the 72 right-censored supercentenarian women or one of the lifetimes of X_{n+1} and X_{n+2}, would live longer than the age $\mathcal{R}_3 = 117.8$, using Equation (3.22), is

$$P(G_{117.8}(2)) = 1 - \left[\frac{1580(1579)(1578)}{(1580+2)(1580+1)(1580)} \prod_{r=1}^{72} \frac{\tilde{n}_{c_r} - 2}{\tilde{n}_{c_r} + 1} \right] = 1 - 0.2645 = 0.7355$$

If we look at t to be in between $\mathcal{R}_2 = 119.3$ and $\mathcal{R}_3 = 117.8$, in case of $m = 2$, then the upper survival of t is the survival of $\mathcal{R}_3 = 117.8$, that is 0.7355, and the lower survival of t is the survival of $\mathcal{R}_2 = 119.3$, that is 0.5783.

Considering $m \geq 2$ future supercentenarian women to be added to the study, then the probabilities for the events, $G_{\mathcal{R}_1}(m)$, $G_{\mathcal{R}_2}(m)$ and $G_{\mathcal{R}_3}(m)$, respectively, that at least one of the 72 right-censored supercentenarian women or one of the lifetimes of $m \geq 2$ future supercentenarian women, would live longer than $\mathcal{R}_1 = 122.5$, $\mathcal{R}_2 = 119.3$ and $\mathcal{R}_3 = 117.8$, respectively, are shown in Figure 3.3.

From Figure 3.3, if we consider a specific P value, say $P = 0.95$, and then we look at the smallest m for which value of the probabilities $G_{122.5}(m)$, $G_{119.3}(m)$ and $G_{117.8}(m)$, respectively, are greater than $P = 0.95$. We can conclude that the smallest value of m future supercentenarian women at $P = 0.95$ will decrease, as long as the probability for each event that somebody would survive the largest recorded age, ordered backwardly, increases. So, it can be seen from Figure 3.3 that the $P(G_{122.5}(m))$ is greater than $P = 0.95$ for $m \geq 19200$ future supercentenarian women, the probability $P(G_{119.3}(m))$ is greater than $P = 0.95$ for $m \geq 3050$ future supercentenarian women, and the probability $P(G_{117.8}(m))$ is greater than $P = 0.95$ for $m \geq 1180$ future supercentenarian women.

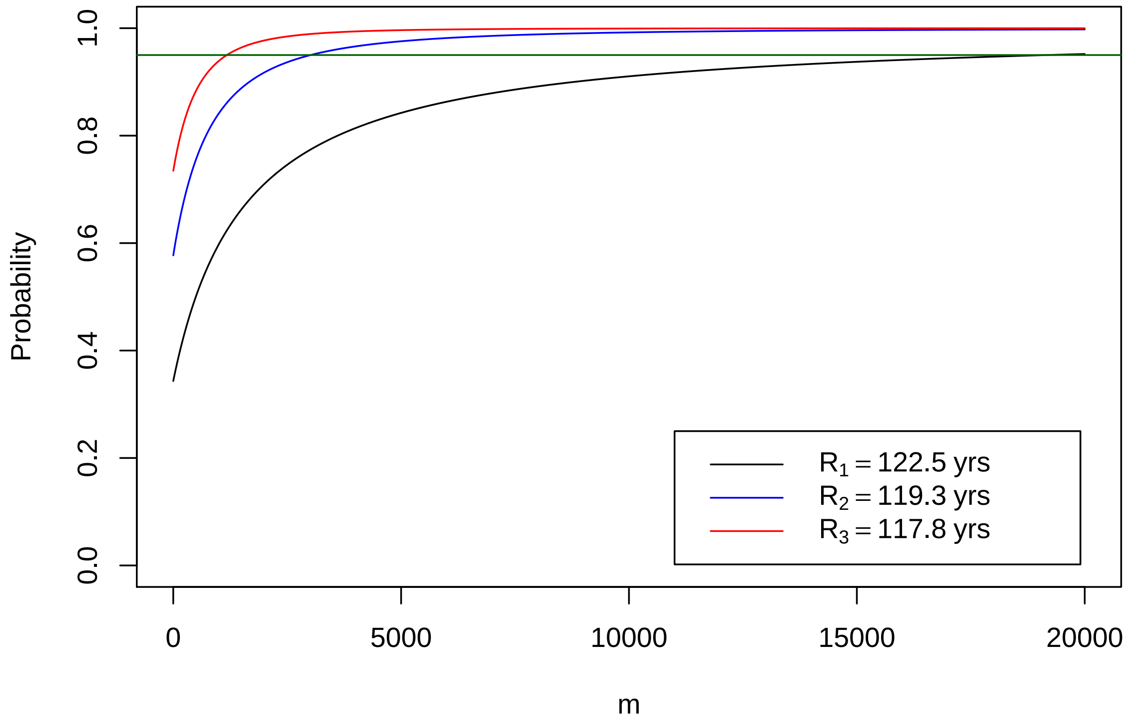

Figure 3.3: $P(G_{\mathcal{R}_1}(m))$, $P(G_{\mathcal{R}_2}(m))$ and $P(G_{\mathcal{R}_3}(m))$, for the supercentenarian women data, according to Example 3.5.1.

Example 3.5.2 (Supercentenarian men data) In this example, we again use the same data on $n = 160$ supercentenarian men (as in Example 3.4.2). There were two supercentenarian men still alive at the time of the study and hence their lifetimes are right-censored. Also, there are 33 supercentenarian men whose ages exceed the largest censored supercentenarian men at age 111.9. Previously, in Example 3.4.2, the first largest age recorded, $\mathcal{R}_1 = 116.2$, were considered, and now we are considering the second and the third largest ages recorded, $\mathcal{R}_2 = 115.7$ and $\mathcal{R}_3 = 115.5$, respectively.

The interest now will be on the question of what is the probability for the event, $G_{\mathcal{R}_2}(0)$, that at least one of the two supercentenarian men whose lifetimes have been right-censored, the actual value of the lifetime would be larger than the second largest observed value $\mathcal{R}_2 = 115.7$. This probability is obtained by using Equation

(3.17), as follows

$$P(G_{115.7}(0)) = 1 - \prod_{r=1}^{2} \frac{\tilde{n}_{c_r-1}}{\tilde{n}_{c_r}+1} = 1 - 0.8903 = 0.1097$$

The probability for the event, $G_{\mathcal{R}_3}(0)$, that at least one of the two supercentenarian men whose lifetimes have been right-censored, the actual value of the lifetime would be larger than the third largest observed value $\mathcal{R}_3 = 115.5$. This probability is obtained by using Equation (3.18), as follows

$$P(G_{115.5}(0)) = 1 - \prod_{r=1}^{2} \frac{\tilde{n}_{c_r-1}}{\tilde{n}_{c_r}+1} = 1 - 0.8378 = 0.1622$$

As a result of what we have assumed in our model, which is based on the $A_{(n)}$ assumption and non-informative right censoring, described in this section, the probability that at least a lifetime of one of the two supercentenarian men, who were still alive at the time of data set, would live beyond the second largest observed age ($\mathcal{R}_2 = 115.7$), is 0.1097, and this probability increased to reach 0.1622 in case of surviving the third largest observed age ($\mathcal{R}_3 = 115.5$). Moreover, it is more likely that somebody would survive any one of the 33 supercentenarian men, as long as it is calculated backwardly from the first largest age to the third largest age that exceeds the largest censored supercentenarian men at age 111.9.

Now, let consider $m = 2$ future supercentenarian men, X_{n+2}, added to the study, given the $n = 160$ supercentenarian men and the first future supercentenarian men, X_{n+1}. The lifetime of X_{n+2} is considered, conditional on that all the two supercentenarian men, whose lifetimes have been right-censored, and the lifetime of X_{n+1}, all have been failed before the value $\mathcal{R}_2 = 115.7$. Then the probability for the event, $G_{\mathcal{R}_2}(2)$, that at least one of the two right-censored supercentenarian men or one of the lifetimes of X_{n+1} and X_{n+2}, would live longer than the age $\mathcal{R}_2 = 115.7$, using Equation (3.21), is

$$P(G_{115.7}(2)) = 1 - \left[\frac{160(159)}{(160+2)(160+1)} \prod_{r=1}^{2} \frac{\tilde{n}_{c_r}-1}{\tilde{n}_{c_r}+1}\right] = 1 - 0.8684 = 0.1316 \tag{3.23}$$

Taking into account the survival of the third largest age $\mathcal{R}_3 = 115.5$, with considering $m = 2$ future supercentenarian men, then the lifetime of X_{n+2} is considered, conditional on that all the two supercentenarian men, whose lifetimes have been right-censored, and the lifetime of X_{n+1}, all have been failed before the value $\mathcal{R}_3 = 115.5$. Then the probability for the event, $G_{\mathcal{R}_3}(2)$, that at least one of the two right-censored supercentenarian men or one of the lifetimes of X_{n+1} and X_{n+2}, would live longer than the age $\mathcal{R}_3 = 115.5$, using Equation (3.22), is

$$P(G_{115.5}(2)) = 1 - \left[\frac{160(159)(158)}{(160+2)(160+1)(160)} \prod_{r=1}^{2} \frac{\tilde{n}_{c_r} - 2}{\tilde{n}_{c_r} + 1} \right] = 1 - 0.8070 = 0.1930 \tag{3.24}$$

Now, let consider the event $G_{t\in(\mathcal{R}_3,\mathcal{R}_2)}(2)$, where t is in between $\mathcal{R}_2 = 115.7$ and $\mathcal{R}_3 = 115.5$, in case of $m = 2$ future supercentenarian men. Then the lower probability for the event $G_{t\in(115.5,115.7)}(2)$ is obtained by deriving the probability for the event $P(G_{115.7}(2))$, using Equation (3.21), so $\underline{P}(G_{t\in(115.5,115.7)}(2)) = P(G_{115.7}(2)) = 0.1316$ (see Equation (3.23)). The corresponding upper probability for the event $G_{t\in(115.5,115.7)}(2)$ is obtained by deriving the probability for the event $P(G_{115.5}(2))$, using Equation (3.22), so $\overline{P}(G_{t\in(115.5,115.7)}(2)) = P(G_{115.5}(2)) = 0.1930$ (see Equation (3.24)).

Considering $m \geq 2$ future supercentenarian men to be added to the study, then the probabilities for the events, $G_{\mathcal{R}_1}(m)$, $G_{\mathcal{R}_2}(m)$ and $G_{\mathcal{R}_3}(m)$, respectively, that at least one of the two right-censored supercentenarian men or one of the lifetimes of $m \geq 2$ future supercentenarian men, would live longer than $\mathcal{R}_1 = 116.2$, $\mathcal{R}_2 = 115.7$ and $\mathcal{R}_3 = 115.5$, respectively, are shown in Figure 3.4.

From Figure 3.4, if we consider a specific P value, say $P = 0.95$, and then we look at the smallest m for which value of the probabilities $G_{116.2}(m)$, $G_{115.7}(m)$ and $G_{115.5}(m)$, respectively, are greater than $P = 0.95$. We can conclude that the smallest value of m future supercentenarian men at $P = 0.95$ will decrease, as long as the probability for each event that somebody would survive the largest recorded age, ordered backwardly, increases. So, it can be seen from Figure 3.4 that the $P(G_{116.2}(m))$ is greater than $P = 0.95$ for $m \geq 2900$ future supercentenarian men, the probability $P(G_{115.7}(m))$ is greater than $P = 0.95$ for $m \geq 515$ future supercentenarian men, and the probability $P(G_{115.5}(m))$ is greater than $P = 0.95$ for $m \geq 250$ future supercentenarian men.

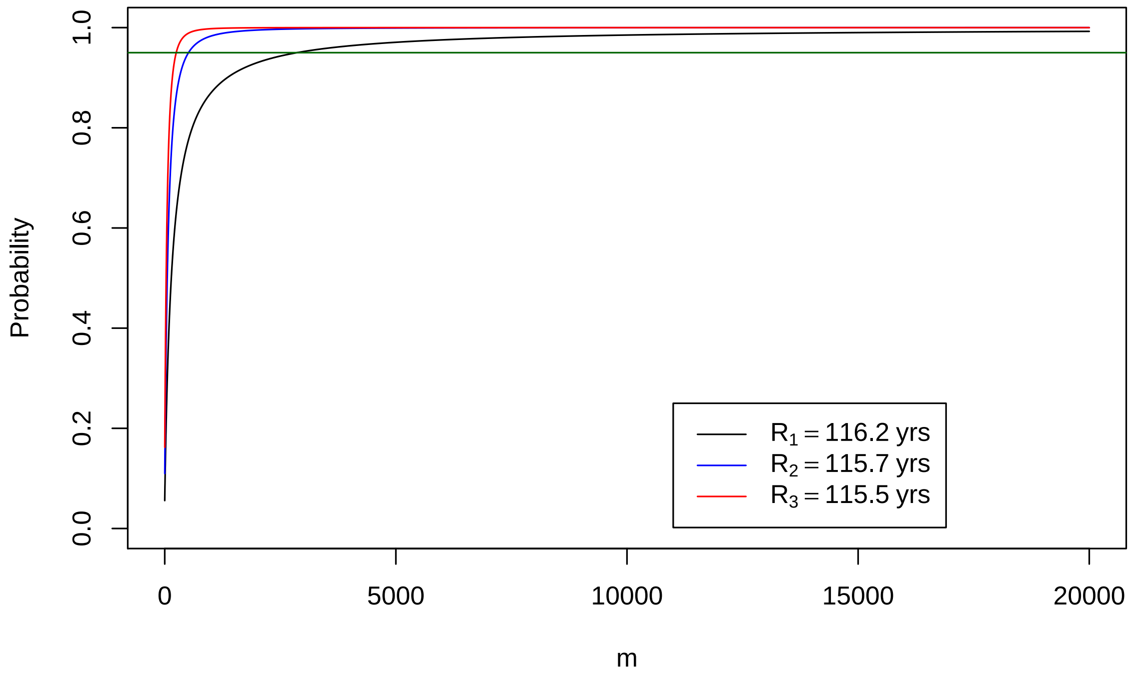

Figure 3.4: $P(G_{\mathcal{R}_1}(m))$, $P(G_{\mathcal{R}_2}(m))$ and $P(G_{\mathcal{R}_3}(m))$, for the supercentenarian men data, according to Example 3.5.2.

3.6 Concluding remarks

This chapter presented a method on which taking the largest observation in a data set, including right-censored observations, as end point of support. The new method finds the probability for the event of interest that the actual lifetime corresponding to a right-censored observation would exceed the largest observed value. Taking into account new future items added to the study, we find the probability for the event of interest that for at least either one of the observations whose lifetimes have been right-censored, or one of multiple future items, added to the study, the actual value of the lifetime would be larger than the largest observed value. In order to derive these results, we assumed that the remaining times to the event of interest for all individuals reaching a certain age, are exchangeable. So, these new results are presented based on the exchangeability assumption, using $A_{(n)}$ assumption [39] and non-informative right censoring [32]. In particular, the assumption of exchangeability seems reasonable in the absence of additional information about the people, for instance, we have no details on their health or other factors that would influence their remaining lifetime.

The proposed method is extended to derive the probability for the event of interest that the actual lifetime corresponding to a right-censored observation would exceed a largest observed value, as long as it past the largest censored observation in the data set. Also, it is extended to derive the probability for the event of interest that for at least either one of the observations whose lifetimes have been right-censored, or one of the multiple future items, added to the study, the actual value of the lifetime would be exceed the largest observed value. These new methods, presented in this chapter, applied to the full Supercentenarian data set, but separately for the women and the men. [8].

On the basis of our investigation of the Supercentenarian data, where the probabilities that somebody woud survive the largest observed age were quite high, we think that it is not appropriate for analysis of extreme values to assume that the largest value is the end-point of support. Due to the weak assumptions underlying NPI, it cannot be used for more detailed prediction of observations beyond the largest observation, this would require additional distributional assumptions.

Chapter 4

NPI Alternative to the Actuarial Estimator

4.1 Introduction

In the real world, time or time till occurence of an event is usually considered to be a continuous variable. One can, however, argue that recorded observations have a discrete character, because all recordings could be interpreted as discrete values. But if very many different values are possible, any difference between the continuous or discrete nature of a variable tends to be neglectable. However, in some application areas time tends to be modelled as a discrete variable, with relatively few possible values, this has particularly been the case for actuarial models, e.g. the actuarial estimator of the survival function.

For such models, typically a cohort of people, either real or just as a concept, is followed through time and events are recorded per a year. The main event of interest is typically death of a person, but right-censoring tends to occur if the person exits the cohort for another reason. In such cases, time is usually recorded as the age of the person at the time of the event, hence time is considered as a discrete variable.

At any given point in time, we consider how many people are alive; that is, how many people have survived that time, so this is effectively Bernoulli data; then we look ahead and assess how many people will be alive in the future. The actuarial estimator, i.e. a nonparametric method used to estimate the survival function, explicitly restricts attention to discrete time, when using right-censored data [6, 50, 60]. In this chapter we take a similar approach, but we do it from the perspective of NPI. In particular, we propose to use the NPI method as an alternative predictive approach to the actuarial estimator with right-censored data.

Since NPI for Bernoulli data [20], as discussed in Section 2.3, was introduced for multiple future observations, it can be used to develop the NPI alternative to the actuarial estimator. The proposed method is developed based on the assumption of non-informative right censoring [32, 65], as discussed in Section 1.2.

The discrete time approach, considered in this chapter, can simply visualize as a table with t_j discrete-time points such that $t_1 < t_2 < \ldots < t_k$, where $j = 1, 2, \ldots, k$. We only have the number of events and the number of right-censored individuals at each given discrete time point, with respect to that no events or right-censoring have occurred at time t_0. Consequently, the proposed method, i.e. NPI alternative to the actuarial estimator, is appropriate to use when data consists of numbers of event times and right-censoring times at specific discrete-time points t_j.

This chapter is organised as follows. A brief introduction to the actuarial estimator of the survival function is presented in Section 4.2. Section 4.3 presents NPI as an alternative to the actuarial estimator with right-censored data, which allows us to derive the lower and upper probabilities for the event that all future observations are greater than a specific discrete time t_j. In the proposed method there are no intervals, time is discrete. But NPI for grouped data [65] is somewhat similar as we only know the number of events and the number of right-censorings in each interval, hence it is of interest to compare it with our proposed method. This comparison is presented in Section 4.4. In Section 4.5, the proposed method will be applied to system reliability using survival signatures [4, 25] combined with NPI for Bernoulli data [20]. This chapter ends with some concluding remarks in Section 4.6.

4.2 Actuarial estimator of the survival function

In the discrete-time approach, one common nonparametric method for estimating the survival function is the actuarial estimator. To introduce the actuarial estimator, we first consider n individuals alive at time t_0. Let $X_1, X_2, \dots, X_n$ be positive, exchangeable and discrete random variables, of which their discrete lifetimes are assumed to be independent and identically distributed, that takes values at discrete-time points t_j, where $j = 1, \dots, k$, with $t_1 < t_2 < \cdots < t_k$. Consider the event of interest as 'death'. The discrete time hazard function at a specific time t_j is defined as the conditional probability that a randomly selected individual, X_i, $i = 0, 1, \dots, n$, will experience the event of interest at time t_j, given that this individual did not experience the event prior to t_j, so that

$$h_{t_j} = P(X_i = t_j | X_i \geq t_j) \tag{4.1}$$

We introduce the following notation. Let d_{t_j} denotes the number of deaths (n individuals who died) at time t_j and let c_{t_j} denotes the number of individuals whose lifetimes are right-censored at time t_j. Let $\grave{n}_{t_j}$ denotes the number of individuals known to be at risk (still functioning or alive and uncensored) at time t_j, that is $\grave{n}_{t_j} = \grave{n}_{t_{j-1}} - d_{t_{j-1}} - c_{t_{j-1}}$. Let $n_{t_0} = 0$, so $\grave{n}_{t_0} = n$ (all individuals are at risk). Then the discrete time hazard function, h_{t_j}, at a discrete time t_j, can be estimated by the actuarial estimator [6, 50, 60], as

$$\widehat{h}_{t_j} = \frac{d_{t_j}}{\grave{n}_{t_j}} \tag{4.2}$$

The survival function at time t_j is defined as $S_{t_j} = P(X \geq t_j)$, note that $S_{t_0} = P(X \geq 0) = 1$. The survival function S_{t_j} can be expressed in terms of the hazard h_{t_j} at all earlier times $t_1, t_2, \dots, t_{j-1}$ as

$$\widehat{S}_{t_j} = \prod_{l=1}^{j-1} \left(\frac{\grave{n}_{t_l} - d_{t_l}}{\grave{n}_{t_l}} \right) = \prod_{l=1}^{j-1} (1 - \widehat{h}_{t_l}) \tag{4.3}$$

The following example illustrates how to estimate the survival function using the actuarial estimate approach.

Example 4.2.1 (Actuarial estimator for the survival function) To illustrate the actuarial estimator for the survival function, we consider a simple example involving $n = 9$ observations, available at discrete times t_j, for $j = 1, 2, 3, 4$, (the data are displayed in Table 4.1).

t_j	d_{t_j}	c_{t_j}	$\grave{n}_{t_j}$	$1-\widehat{h}_{t_j}$	$\widehat{S}_{t_j}$
t_1	1	0	9	0.8889	0.8889
t_2	2	1	8	0.7500	0.6667
t_3	2	1	5	0.6000	0.4000
t_4	0	1	2	1.0000	0.4000

Table 4.1: Actuarial estimator for the survival function (Example 4.2.1).

The probability of hazard function, h_{t_j}, at a discrete time t_j can be estimated by the actuarial estimator using Equation (4.2) and then the corresponding survival probability at t_j is obtained by $\grave{n}_{t_j} - (d_{t_j}/\grave{n}_{t_j})$. Then, cumulatively, the estimated probability of surviving t_j, for $j = 1, 2, 3, 4$, is derived by using Equation (4.3). These results are presented in the Table 4.1.

Next, a similar approach to the actuarial estimator will be considered under NPI methodology as an alternative to the actuarial estimator, using NPI for Bernoulli data [20], as was reviewed in Section 2.3.

4.3 NPI alternative to the actuarial estimator

In this section, we introduce an NPI based alternative to the actuarial estimator. We use the same notation as introduced in Section 4.2. Let $X_1, X_2, \ldots, X_n$ be positive, exchangeable and discrete random variables, with $i = 0, 1, \ldots, n$. Assume that the population only has n individuals, consists of event and censoring times given k distinct discrete-time points, with $t_1 < t_2 < \cdots < t_k$, defining $\grave{n}_{t_0} = n$ for the start of the study where all individuals survived and $t_{k+1} = \infty$. Let d_{t_j} represent the number of observed events at time t_j. Let c_{t_j} represent the number of censored at times t_j.

As considered in this approach that the censored observations are assumed to occur at discrete times t_j, for $j = 1, 2, \ldots, k$, so the number of individuals at risk at time t_j, denoted by $\hat{n}_{t_j}$, is computed by $\hat{n}_{t_j} = n_{t_{j-1}} - c_{t_j}$. Thus, the number of individuals at risk at time t_j, $\hat{n}_{t_j}$, will be decreasing at next discrete times. Further, let $X_l^{t_j} > t_j$, for $l = 1, \ldots, \hat{n}_{t_j}$ be the event times for the individuals in the risk set at time t_j.

Even though NPI is predictive inference, not estimation, it is still interesting to compare it with the actuarial estimator, as NPI is well suited for dealing with right-censored data based on the rc-$A_{(n)}$ assumption as was reviewed in Section 2.4, and it is developed for multiple future observations via the NPI for Bernoulli data as was reviewed in Section 2.3.

Now, let us consider X_{n+i} for the time of event of the i^{th} future individual, for $i = 1, 2, \ldots, m$. We consider the event of interest that all m multiple future observations X_{n+i} survive a specific discrete time t_j given that they survived the earlier discrete time t_{j-1}, so $\bigcap_{i=1}^{m}\{X_{n+i} > t_j | X_{n+i} > t_{j-1}\}$. For convenience, we indicate this event as $E_j(m)$. Then, we aim to derive NPI lower and upper probabilities for the event $E_j(m)$, based on NPI for Bernoulli data [20], which has been presented in Section 2.3.

4.3.1 Lower and upper probabilities for the event $E_j(m)$

In this section, we derive the NPI alternative to the actuarial estimator in terms of lower and upper probabilities for the event $E_j(m)$, by utilising NPI for Bernoulli data [20], presented in Section 2.3. We denote the conditional lower and upper probabilities for the event $\bigcap_{i=1}^{m}\{X_{n+i} > t_j | X_{n+i} > t_{j-1}\}$ by $\underline{P}\left(E_j(m)\right)$ and $\overline{P}\left(E_j(m)\right)$, respectively.

We consider the survival of all m future observations at time t_j as exchangeable with the survival of the individuals in the risk set $\hat{n}_{t_j}$. So, we assume that the random quantities $X_{n+1}, X_{n+2}, \ldots, X_{n+i}$, with respect to the event $X_{n+i} > t_j$, $i = 1, \ldots, m$, are exchangeable with $X_1^{t_j}, X_2^{t_j}, \ldots, X_l^{t_j}$, with respect to the event $X_l^{t_j} > t_j$, for $l = 1, \ldots, \hat{n}_{t_j}$, where $X_l^{t_j}$ are the event times for the individuals in the risk set at time t_j.

The NPI alternative to the actuarial estimator, consists of conditional lower and upper probabilities, $\underline{P}\left(E_j(m)\right)$ and $\overline{P}\left(E_j(m)\right)$, respectively, at a specific discrete time t_j. These conditional lower and upper probabilities can be derived by utilising the NPI for Bernoulli data [20] via applying the Equations (2.8) and (2.11), respectively, regarding the number of individuals that known to be alive at time t_j out of the number of individuals at risk at time t_j, $\hat{n}_{t_j}$. Thus, the NPI conditional lower probability for the event $E_j(m)$ is

$$\underline{P}\left(E_j(m)\right) = \underline{P}\left(\bigcap_{i=1}^{m}\{X_{n+i} > t_j | X_{n+i} > t_{j-1}\}\right) = \prod_{i=1}^{m} \frac{n_{t_j} + i - 1}{\hat{n}_{t_j} + i} \tag{4.4}$$

The coreresponding NPI conditional upper probability for the event $E_j(m)$ is

$$\overline{P}\left(E_j(m)\right) = \overline{P}\left(\bigcap_{i=1}^{m}\{X_{n+i} > t_j | X_{n+i} > t_{j-1}\}\right) = \prod_{i=1}^{m} \frac{n_{t_j} + i}{\hat{n}_{t_j} + i} \tag{4.5}$$

4.3.2 NPI probabilities for the event $\bigcap_{i=1}^{m}\{X_{n+i} > t_j\}$

We now consider the event that the m future observations will all exceed t_j, that is $\bigcap_{i=1}^{m}\{X_{n+i} > t_j\}$. The NPI lower and upper probabilities for this event, denoted by $\underline{P}\left(\bigcap_{i=1}^{m}\{X_{n+i} > t_j\}\right)$ and $\overline{P}\left(\bigcap_{i=1}^{m}\{X_{n+i} > t_j\}\right)$, can be expressed in terms of the NPI conditional lower and upper probabilities, $\underline{P}\left(E_j(m)\right)$ and $\overline{P}\left(E_j(m)\right)$, provided in Equations (4.4) and (4.5), respectively, at all earlier times $t_1, t_2, \dots, t_j$, thus

$$\underline{P}\left(\bigcap_{i=1}^{m}\{X_{n+i} > t_j\}\right) = \prod_{\ell=1}^{j} \underline{P}\left(E_\ell(m)\right) = \prod_{\ell=1}^{j}\left(\prod_{i=1}^{m} \frac{n_{t_\ell} + i - 1}{\hat{n}_{t_\ell} + i}\right) \tag{4.6}$$

$$\overline{P}\left(\bigcap_{i=1}^{m}\{X_{n+i} > t_j\}\right) = \prod_{\ell=1}^{j} \overline{P}\left(E_\ell(m)\right) = \prod_{\ell=1}^{j}\left(\prod_{i=1}^{m} \frac{n_{t_\ell} + i}{\hat{n}_{t_\ell} + i}\right) \tag{4.7}$$

We consider the case $m = 1$, considering the first future observation, X_{n+1}, for comparisons later on. So, the NPI lower and upper probabilities for the event $X_{n+1} > t_j$ are directly resulting from Equations (4.6) and (4.7), respectively, thus

$$\underline{P}(X_{n+1} > t_j) = \prod_{l=1}^{j} \frac{n_{t_l}}{\hat{n}_{t_l} + 1} \tag{4.8}$$

$$\overline{P}(X_{n+1} > t_j) = \prod_{l=1}^{j} \frac{n_{t_l} + 1}{\hat{n}_{t_l} + 1} \tag{4.9}$$

NPI lower and upper probabilities for the event $\bigcap_{i=1}^{m}\{X_{n+i} > t_j\}$, as presented in Equations (4.6) and (4.7), takes into account the dependence among all these future observations when there is limited information in the form of n observations in the data set. It is of interest to see the effect of taking this dependence carefully into account. For this reason, we will compare our method with the results one would get if one, mistakenly, when interested in m future observations, would use the NPI lower and upper probabilities for the event $X_{n+1} > t_j$, presented in Equations (4.8) and (4.9), raised to the power of m, i.e. $\left[\underline{P}, \overline{P}\right]^m (X_{n+1} > t_j)$. This will be illustrated in Example 4.3.2.

The following examples illustrate the method presented above and compare the method with with the results when ignoring the dependency between the multiple future observations as well as with the actuarial estimator reviewed in Section 4.2.

4.3.3 Examples

In this subsection, two examples are presented to illustrate the NPI alternative to the actuarial estimator.

Example 4.3.1 We consider the data set used in Example 4.2.1, with $n = 9$ observations, (the data are presented in Table 4.1).

The first step in applying the proposed method is to draw a table that will assist in the calculation of the NPI lower and upper probabilities for the event $\bigcap_{i=1}^{m}\{X_{9+i} > t_j\}$ for $m \in \{1, 2, 3, 10, 15\}$ future observations. As shown in Table 4.2, there are four discrete-time points at t_1, t_2, t_3 and t_4. At each discrete time point, this table provides the number of observed events, d_{t_j}, the number of censored individuals, c_{t_j}, the number of individuals known to be alive at time t_j, n_{t_j}, the number of individuals at risk at time t_j, $\hat{n}_{t_j}$, where $\hat{n}_{t_j}$ is computed differently than that in Example 4.2.1, $\grave{n}_{t_j}$, i.e. $\hat{n}_{t_2} = 7$ but $\grave{n}_{t_2} = 8$, (see Sections 4.2 and 4.3), and the NPI lower and upper probabilities of surviving time t_j.

It is noteworthy that, at the start of the study at time t_0, no events or censorings have been recorded, so $\underline{P}(\bigcap_{i=1}^{m}\{X_{9+i} > t_0\}) = \overline{P}(\bigcap_{i=1}^{m}\{X_{9+i} > t_0\}) = 1$. Then we apply the NPI alternative to the actuarial estimator e.g. leading to the conditional lower and upper probabilities, as given by Equations (4.4) and (4.5), respectively, for the discrete-time points t_1, t_2, t_3 and t_4.

Following the results in Table 4.2, the difference between the NPI upper probabil-

					$m=1$		$m=2$		$m=3$		$m=10$		$m=15$	
t_j	d_{t_j}	c_{t_j}	$\hat{n}_{t_j}$	n_{t_j}	$\underline{P}$	$\overline{P}$	$\underline{P}$	$\overline{P}$	$\underline{P}$	$\overline{P}$	$\underline{P}$	$\overline{P}$	$\underline{P}$	$\overline{P}$
t_1	1	0	9	8	0.8000	0.9000	0.7273	0.9091	0.6667	0.9167	0.4211	0.9474	0.3333	0.9583
t_2	2	1	7	5	0.5000	0.6750	0.4040	0.7071	0.3333	0.7333	0.1238	0.8359	0.0758	0.8712
t_3	2	1	4	2	0.2000	0.4050	0.1347	0.4714	0.0952	0.5238	0.0177	0.7165	0.0080	0.7795
t_4	0	1	1	1	0.1000	0.4050	0.0449	0.4714	0.0238	0.5238	0.0016	0.7165	0.0005	0.7795

Table 4.2: NPI lower and upper probabilities for $\bigcap_{i=1}^{m}\{X_{9+i} > t_j\}$, $m \in \{1, 2, 3, 10, 15\}$ (Example 4.3.1).

ity and the NPI lower probability is quite small at time t_1 for all considered numbers of future observations, and becomes larger later on. So, there are two effects that cause the difference to increase, due to fewer individuals in the risk set, $\hat{n}_{t_j}$, later on, at times t_2, t_3 and t_4, and due to taking products of lower probabilities and of upper probabilities, so each term (so each time point) adds to the imprecision.

Comparison of the results based on our proposed method for $m = 1$ future observation, given in Table 4.2, with those resulting from estimating the survival function based on the actuarial estimator, given in Table 4.1, which has been considered for $m = 1$, shows that the $\widehat{S}_{t_j}$ values, based on using the actuarial estimator, fall between our NPI lower and upper probabilities for $X_{10} > t_j$, but more closer to the upper probability values.

Example 4.3.2 The data set used in this example has been used by Berkson and Gage [15] to interpret the survival experience of a set of patients who had operations for a certain type of cancer; the dataset has also been used by Lawless [47] and Yan [65]. The data consists of 374 observations, 95 of which are right-censored observations and the remaining observations are event times considered at 10 discrete times, measured in years. The dataset is summarized in the first three columns of Table 4.3.

By applying the proposed method given by Equations (4.6) and (4.7), respectively, we derive the NPI lower and upper probabilities for the event $\bigcap_{i=1}^{m}\{X_{n+i} > t_j\}$ for $m \in \{1, 2, 3, 10\}$ future observations at discrete-time points t_1, t_2, t_3, t_4. The results are presented in Table 4.3.

					$m = 1$		$m = 2$		$m = 5$		$m = 10$		$[\underline{P}, \overline{P}]^5(X_{375} > t_j)$	
t_j	d_{t_j}	c_{t_j}	$\hat{n}_{t_j}$	n_{t_j}	$\underline{P}$	$\overline{P}$	$\underline{P}$	$\overline{P}$	$\underline{P}$	$\overline{P}$	$\underline{P}$	$\overline{P}$	$[\underline{P}]^5$	$[\overline{P}]^5$
t_1	90	0	374	284	0.757	0.760	0.574	0.578	0.2513	0.2557	0.0644540	0.0667235	0.2486	0.2536
t_2	76	0	284	208	0.553	0.557	0.306	0.311	0.0527	0.0549	0.0029253	0.0031739	0.0517	0.0536
t_3	51	0	208	157	0.415	0.421	0.173	0.178	0.0128	0.0138	0.0001793	0.0002070	0.0123	0.0132
t_4	25	12	145	120	0.341	0.349	0.117	0.123	0.0049	0.0055	0.0000269	0.0000336	0.0046	0.0051
t_5	20	5	115	95	0.280	0.289	0.079	0.084	0.0018	0.0022	0.0000040	0.0000055	0.0017	0.0020
t_6	7	9	86	79	0.254	0.266	0.065	0.071	0.0011	0.0014	0.0000016	0.0000025	0.0011	0.0013
t_7	4	9	70	66	0.236	0.251	0.056	0.063	0.0008	0.0011	0.0000008	0.0000014	0.0007	0.0010
t_8	1	3	63	62	0.229	0.247	0.053	0.061	0.0007	0.0010	0.0000006	0.0000012	0.0006	0.0009
t_9	3	5	57	54	0.213	0.234	0.046	0.055	0.0005	0.0008	0.0000003	0.0000008	0.0004	0.0007
t_{10}	2	5	49	47	0.200	0.225	0.040	0.051	0.0004	0.0006	0.0000002	0.0000005	0.0003	0.0006

Table 4.3: NPI lower and upper probabilities for $\bigcap_{i=1}^{m}\{X_{374+i} > t_j\}$, $m \in \{1, 2, 5, 10\}$ and $[\underline{P}, \overline{P}]^5(X_{375} > t_j)$ (Example 4.3.2).

Now, in order to see the effect of taking the dependence of all future observations carefully into account, we compare our results $[\underline{P}, \overline{P}](\bigcap_{i=1}^{5}\{X_{374+i} > t_j\})$, for $m = 5$, with those if we mistakenly take the NPI lower and upper for only the first future observation $(X_{375} > t_j)$ raised to the power of $m = 5$, i.e. $[\underline{P}, \overline{P}]^5(X_{375} > t_j)$. Due to the positive dependence among X_{375}, X_{376}, X_{377}, X_{378} and X_{379}, our correct NPI lower and upper probabilities for the event $\bigcap_{i=1}^{5}\{X_{374+i} > t_j\}$, are greater than the values resulting from mistakenly taking the lower and upper probabilities for $X_{375} > t_j$ raised to the power of 5. Even though the differences (NPI upper probability - NPI lower probability) are quite small, they will become larger for $m > 5$ future observations due to the positive dependence of all future observations.

In the following section, we compare our results, which are given in Section 4.2, with an alternative nonparametric predictive approach, which could also be applied in cases of discrete time.

4.4 Comparison with NPI for grouped data

From a reliability and survival analysis perspective, lifetime data can be recorded in groups, represented by a finite number of intervals, by recording or using only the numbers of event times and the numbers of censoring times in these intervals, rather than the exact observed times. In this case, time is continuous but there either is no further information about the exact observations within the intervals, or such information is neglected.

Using grouped data, if we can ignore its baseline continuous time characteristics and this method keeps time continuous but only uses the information per interval that the numbers of event times and the numbers of censoring times can only occur at specific discrete recorded time points that occur at the beginning of each given grouped-timed period, we compare this approach with our proposed methodology, presented in Section 4.3. So, the proposed method, presented in Section 4.3, will be compared with an alternative methodology developed for grouped data with right-censored observations, called "NPI for grouped data" [65].

Coolen and Yan [65] presented NPI for grouped data, which has been applied to grouped data to derive NPI lower and upper survival functions for the event that the first future observation X_{n+1} is greater than time t.

In Subsection 4.4.1, we provide a brief overview of NPI for grouped data [65] and this will be compared, via an example, to our proposed method in Subsection 4.4.2.

4.4.1 NPI for grouped data

Coolen and Yan [65] have developed NPI for grouped data, with real-valued observations including right-censored data, based on the rc-$A_{(n)}$ assumption as presented in Section 2.1. Suppose that the time-axis is partitioned into $k+1$ intervals. Let $\mathcal{I}_j = [t_j, t_{j+1})$ represent the intervals in which the time-axis is separated into t_j points, where $j = 0, 1, 2, \ldots, k$, such that $t_1 < t_2 < \cdots < t_k$, with defining $t_0 = 0$, and $t_{k+1} = \infty$. For each interval $\mathcal{I}_j$, the number of events and the number of right-censorings in the interval are known, but not the exact times. Let d_{t_j} be the number of event times in $\mathcal{I}_j$ and the c_{t_j} be the number of right-censorings in $\mathcal{I}_j$. Because of this, the order of events and censoring times for each interval must be taken into account, to deal with such grouped data. Coolen and Yan [65] used 'optimal configurations', to derive NPI for grouped data [65], which has been developed for only $m = 1$ future observation. The optimal configurations indicate that orderings of event and censoring times within the intervals are required to derive the M-function values for X_{n+1} to be within a particular interval, which ones lead to the lower and upper probabilities for the event $X_{n+1} > t$, for $t \in \mathcal{I}_j$ (see [65] for the discussion of the optimal configurations). Then, the NPI upper probability, denoted by $\overline{P}^{\mathcal{I}}(X_{n+1} > t)$, for $t \in \mathcal{I}_j = [t_j, t_{j+1})$, with $j = 0, 1, 2, \ldots, k$, is derived as follows [65].

$$\overline{P}^{\mathcal{I}}(X_{n+1} > t) = \overline{P}^{\mathcal{I}_{j-1}}(X_{n+1} > t) - d_{t_{j-1}} p^{j-1}, \quad \text{for } j \geq 1 \tag{4.10}$$

where,

$$p^{j-1} = p^0 = \frac{1}{n+1}, \qquad \text{for } j = 1, \quad \text{and}$$

$$p^{j-1} = p^{j-2} \times \frac{n - \sum_{i=0}^{j-3} n_{t_i} - d_{t_{j-2}} + 1}{n - \sum_{i=0}^{j-2} n_{t_i} + 1}, \qquad \text{for } j \geq 2$$

Note that $\overline{P}^{\mathcal{I}_{j-1}}(X_{n+1} > t)$ in Equation (4.10) refers to the NPI upper probability for the event $X_{n+1} > t$ where $t \in \mathcal{I}_{j-1} = [t_{j-1}, t_j)$, with $j = 0, 1, 2, \ldots, k$.

The corresponding NPI lower probability, denoted by $\underline{P}^{\mathcal{I}}(X_{n+1} > t)$, for $t \in \mathcal{I}_j = (t_j, t_{j+1}]$, with $j = 0, 1, 2, \ldots, k$, is derived as follows [65].

$$\underline{P}^{\mathcal{I}}(X_{n+1} > t) = \underline{P}^{\mathcal{I}_{j-1}}(X_{n+1} > t) + q^{j-1} - (d_{t_j} + 1) q^j, \quad \text{for } j \geq 1 \tag{4.11}$$

where,

$$q^0 = \frac{1}{n - c_{t_0} + 1}, \qquad \text{for } j = 1, \quad \text{and}$$

$$q^j = q^{j-1} \times \frac{n - \sum_{i=0}^{j-1} n_{t_i} + 1}{n - \sum_{i=0}^{j-1} n_{t_i} - c_{t_j} + 1}, \qquad \text{for } j \geq 2$$

Based on grouped data analysis approach, as in [65], we only know the numbers of event times and right-censoring times in $\mathcal{I}_j = [t_j, t_{j+1})$; these values could be anywhere in this interval without any additional assumptions. Using the assumption rc-$A_{(n)}$ [32] given in Section 2.4, the optimal configurations via M-function values are used to derive the maximum probability mass to the right of t_j taking into account that all observations in the interval $\mathcal{I}_j$ are assumed to be greater than t, which produced the NPI upper probability for the grouped data $\overline{P}^{\mathcal{I}}(X_{n+1} > t_j)$ [65], so $\overline{P}^{\mathcal{I}}(X_{n+1} > t)$ in Equation (4.10) is equal $\overline{P}^{\mathcal{I}}(X_{n+1} > t_j)$ for $t \in \mathcal{I}_j = [t_j, t_{j+1})$, with $j = 0, 1, 2, \ldots, k$. Also, by placing minimum probability mass to the right of t_j taking into account that all observations in the interval $\mathcal{I}_j$ are assumed to be less than t, we produced the NPI lower probability for grouped data $\underline{P}^{\mathcal{I}}(X_{n+1} > t)$, so $\underline{P}^{\mathcal{I}}(X_{n+1} > t)$ in Equation (4.11) is equal to $\underline{P}^{\mathcal{I}}(X_{n+1} > t_{j+1})$ for $t \in \mathcal{I}_j = (t_j, t_{j+1}]$, with $j = 0, 1, 2, \ldots, k$.

As it is assumed that no censorings or events occurred at time t_0, $\overline{P}^{\mathcal{I}}(X_{n+1} > t_0) = 1$, but $\underline{P}^{\mathcal{I}}(X_{n+1} > t_0) = \frac{n-n_{t_0}}{n-c_{t_0}+1}$, where d_{t_0} and c_{t_0} are the the number of events and the number of right-censorings, respectively, in the first interval $\mathcal{I}_0 = (0, t_1]$.

In the following subsection, the theory of NPI for grouped data [65], presented in Subsection 4.4.1, is compared to our method, presented in Section 4.3. In addition, an example is given to show these comparisons.

4.4.2 Comparing our method with NPI for grouped data

In this section, we compare our approach presented in Section 4.3, for the discrete-time case, with NPI for grouped data as presented in Subsection 4.4.1 [65]. The theory of NPI for grouped data has been developed for predicting only a single future observation for the event $X_{n+1} > t_j$, while our method, NPI alternative to the actuarial estimator as presented in Section 4.3, allows the NPI methodology using right-censored data to predict multiple future observations for the event that all these future observations survived discrete times t_j. In this section, we present only the comparison of NPI lower and upper probabilities for one future observation to be greater than the time t_j from both perspectives.

Considering the event $X_{n+1} > t_j$, where $j = 0, 1, 2, \ldots, k$, we compare the NPI lower and upper probabilities for grouped data, which we denoted by $\underline{P}^{\mathcal{I}}(X_{n+1} > t_j)$ and $\overline{P}^{\mathcal{I}}(X_{n+1} > t_j)$, given by Equations (4.10) and (4.11), respectively, with the NPI lower and upper probabilities following our method stated in Section 4.3, which we denoted by $\underline{P}(X_{n+1} > t_j)$ and $\overline{P}(X_{n+1} > t_j)$, given by Equations (4.8) and (4.9), respectively.

As discussed in the Subsection 4.4.1 that, $\underline{P}^{\mathcal{I}}(X_{n+1} > t) = \underline{P}^{\mathcal{I}}(X_{n+1} > t_{j+1})$ for $t \in \tilde{\mathcal{I}}_j = (t_j, t_{j+1}]$, and $\overline{P}^{\mathcal{I}}(X_{n+1} > t) = \overline{P}^{\mathcal{I}}(X_{n+1} > t_j)$ for $t \in \mathcal{I}_j = [t_j, t_{j+1})$, with $j = 0, 1, 2, \ldots, k$. So, the upper probability based on NPI for grouped data at t_0 is

t_j	d_{t_j}	c_{t_j}	$\underline{P}(X_{375} > t_j)$	$\overline{P}(X_{375} > t_j)$	$\underline{P}^{\mathcal{I}}(X_{375} > t_j)$	$\overline{P}^{\mathcal{I}}(X_{375} > t_j)$	$\overline{P} - \underline{P}$	$\overline{P}^{\mathcal{I}} - \underline{P}^{\mathcal{I}}$
t_0	90	0	1	1	0.757	1	0	0.243
t_1	76	0	0.757	0.760	0.555	0.760	0.003	0.205
t_2	76	0	0.553	0.557	0.419	0.557	0.004	0.138
t_3	51	0	0.415	0.421	0.346	0.421	0.006	0.075
t_4	25	12	0.341	0.349	0.286	0.355	0.008	0.069
t_5	20	5	0.280	0.289	0.262	0.296	0.009	0.034
t_6	7	9	0.254	0.266	0.247	0.275	0.012	0.028
t_7	4	9	0.236	0.251	0.243	0.261	0.015	0.018
t_8	1	3	0.229	0.247	0.230	0.257	0.018	0.027
t_9	3	5	0.213	0.234	0.220	0.245	0.021	0.025
t_{10}	2	5	0.200	0.225	0	0.236	0.025	0.236

Table 4.4: NPI lower and upper probabilities for discrete time and grouped data.

identical to the lower and upper probabilities based on our proposed method at t_0, as all of them are equal to 1.

Next, a data set from the literature will be used to compare these methods.

Example 4.4.1 Using the same dataset as in Example 4.3.2, we compare the proposed method presented in Section 4.3 for discrete time with NPI for grouped data [65].

Table 4.4 presents the NPI lower and upper probabilities for $X_{375} > t_j$ for discrete time, together with the NPI lower and upper probabilities for the grouped data for $X_{375} > t_j$.

As discussed above, at time t_0, the NPI lower and upper probabilities, based on our proposed method, and the NPI upper probability, based on NPI for grouped data, are equal to 1, whereas the NPI lower probability, based on NPI for grouped data, is equal to 0.757. At time t_{10}, the NPI lower probability, based on NPI for grouped data, is equal to 0, whereas the NPI upper probability, based on NPI for grouped data, and the NPI lower and upper probabilities, based on our proposed method, all remain positive without any further assumptions added.

In this example, we see that, $\overline{P}^{\mathcal{I}}(X_{n+1} > t_j)$ and $\overline{P}(X_{n+1} > t_j)$ are equal when there are no censorings in $\mathcal{I}_j = [t_j, t_{j+1})$. Also, we see that $\underline{P}(X_{n+1} > t_j) >$

$\underline{P}^{\mathcal{I}}(X_{n+1} > t_j)$ and $\overline{P}(X_{n+1} > t_j) \leq \overline{P}^{\mathcal{I}}(X_{n+1} > t_j)$. The proof of these inequalities is difficult, as the calculations of these two methods are actually very different when applied.

From Table 4.4, we see that the difference of the upper and lower probabilities resulting from our proposed method is quite small at time t_1 for the first future observation, and becomes larger later on, due to fewer individuals in the risk set later on. In contrast, the difference of the upper and lower probabilities resulting from NPI for grouped data method is high at time t_1 and becomes smaller later on (with the obvious exception when the lower probability beyond t_{10} is equal to 0).

Next, we use NPI for Bernoulli data [20], presented in Section 2.3, to derive NPI lower and upper probabilities for the event that at least one future observation out of multiple future observations will survive for all t_j. Then, these results will be applied to systems reliability using the concept of survival signatures [23, 25].

4.5 Application to system reliability using survival signatures

On the basis of the NPI for Bernoulli data [20], presented in Section 2.3, we derive NPI lower and upper probabilities for the event that at least x out of m future observations will survive a discrete time t_j, for all future observations and for all t_j. For a specific t_j, we apply these lower and upper probabilities to system reliability, with one or more than one types of components, using survival signatures. Throughout this section, the values of the survival signatures are given.

In Subsection 4.5.1, we derive NPI lower and upper probabilities for the event that at least x out of m future observations will survive a discrete time t_j, using NPI for Bernoulli data. A brief introduction to the survival signatures is provided in Subsection 4.5.2. In Subsection 4.5.3, the results presented in Subsection 4.5.1 will be used to derive NPI lower and upper probabilities for discrete time system reliability, using the survival signature combined with NPI for Bernoulli data. Finally, in Subsection 4.5.4, the proposed methods will be applied to some discrete time systems reliability with a single type of components and multiple types of components.

4.5.1 NPI lower and upper for the event $N_{t_j} \geq x$

This section follows the same notation used in Section 4.3, with additional notation needed. It should be noted that the results presented in this section will be derived on the basis of utilising the results of NPI lower and upper probabilities for Bernoulli data, presented in Section 2.3, explicitly by using Equations (2.12) and (2.13), respectively. Let N_{t_j} denote the number out of m future observations that survive a discrete time t_j. Consider the event $N_{t_j} \geq x$ that is at least x out of m future observations will survive a discrete time t_j, $x \in \{0, 1, \ldots, m\}$. Given $\hat{n}_{t_j}$ Bernoulli trials, and out of these we have $\hat{n}_{t_j} - d_{t_j}$ survived at time t_j, we aim to derive the NPI lower and upper probabilities for the event $N_{t_j} \geq x$ for all $x \in \{0, 1, \ldots, m\}$ and for all t_j. These lower and upper probabilities are denoted by $\underline{P}(N_{t_j} \geq x)$ and $\overline{P}(N_{t_j} \geq x)$, respectively.

The NPI upper probability for the event $N_{t_j} \geq x$, for $x \in \{0, 1, \ldots, m\}$ and for all t_j, is derived by utilising Equation (2.12), as

$$\overline{P}(N_{t_j} \geq x) = \sum_{y=x}^{m} \overline{P}(N_{t_j} \geq x | N_{t_{j-1}} = y) \left[\overline{P}(N_{t_{j-1}} \geq y) - \overline{P}(N_{t_{j-1}} \geq y+1)\right] \quad (4.12)$$

The first term of Equation (4.12) is derived by applying to the Equation (2.12) presented in Section 2.3. The second term of Equation (4.12) refers to the maximum value of the probability that $N_{t_{j-1}}$ is y; there are some orderings for which $N_{t_{j-1}}$ can be y and also can be $y+1$. The argument $\left[\overline{P}(N_{t_{j-1}} \geq y) - \overline{P}(N_{t_{j-1}} \geq y+1)\right]$ is achieved by applying to Equation (2.12) first for the event $N_{t_{j-1}} \geq y$ and then for the event $N_{t_{j-1}} \geq y+1$, for $y \in \{0, 1, \ldots, m\}$ future observations, and we calculate the difference between them. The second term of Equation (4.12) is also obtained by deriving the difference of the NPI upper probibility for the event $N_{t_{j-1}} \geq y | \left(\hat{n}_{t_j}, \hat{n}_{t_j} - d_{t_j}\right)$ and the NPI upper probibility for the event $N_{t_{j-1}} \geq$

$y+1|\left(\hat{n}_{t_j},\hat{n}_{t_j}-d_{t_j}\right)$, using the Expression (2.21), so

$$\left[\overline{P}(N_{t_{j-1}} \geq y|\left(\hat{n}_{t_j},\hat{n}_{t_j}-d_{t_j}\right)) - \overline{P}(N_{t_{j-1}} \geq y+1)|\left(\hat{n}_{t_j},\hat{n}_{t_j}-d_{t_j}\right)\right] = \binom{\hat{n}_{t_{j-1}}+m}{\hat{n}_{t_{j-1}}}^{-1}\binom{(\hat{n}_{t_{j-1}}-d_{t_{j-1}})+y}{(\hat{n}_{t_{j-1}}-d_{t_{j-1}})}\binom{\hat{n}_{t_{j-1}}-(\hat{n}_{t_{j-1}}-d_{t_{j-1}})+m-y-1}{\hat{n}_{t_{j-1}}-(\hat{n}_{t_{j-1}}-d_{t_{j-1}})} \tag{4.13}$$

where $y \in \{0,1,\ldots,m\}$ future observations. It is important to point out that for the case $m+1$, the NPI upper probability for the event $(N_{t_{j-1}} \geq y+1|\left(\hat{n}_{t_j},\hat{n}_{t_j}-d_{t_j}\right))$ is equal to 0.

The NPI lower probability for the event $N_{t_j} \geq x$, for $x \in \{0,1,\ldots,m\}$ and for all t_j, is derived by utilising Equation (2.13), as

$$\underline{P}(N_{t_j} \geq x) = \sum_{y=x}^{m} \underline{P}(N_{t_j} \geq x|N_{t_{j-1}} = y)\left[\overline{P}(N_{t_{j-1}} \leq y) - \overline{P}(N_{t_{j-1}} \leq y-1)\right] \tag{4.14}$$

The first term of Equation (4.14) is derived by applying to the Equation (2.13) presented in Section 2.3. The second term of Equation (4.14) refers to the least possible value for the large values of y. This term is achieved by giving maximum value to smaller y, so for $y=0$, we can get the maximum value to $y=0$ by deriving the upper probability for $y=0$, $\overline{P}(y=0)$. For example, we achieve the maximum value to smaller $y=1$ by subtracting the upper probability for $y=0$ from the upper probability for $y \leq 1$, thus $\overline{P}(y \leq 1) - \overline{P}(y=0)$ leads to maximum value to smaller $y=1$. The argument $\left[\overline{P}(N_{t_{j-1}} \leq y) - \overline{P}(N_{t_{j-1}} \leq y-1)\right]$ is computed as follows.

$$\begin{aligned}\left[\overline{P}(N_{t_{j-1}} \leq y) - \overline{P}(N_{t_{j-1}} \leq y-1)\right] &= 1 - \underline{P}(N_{t_{j-1}} \geq y+1) - \left[1 - \underline{P}(N_{t_{j-1}} \geq y+1)\right] \\ &= \underline{P}(N_{t_{j-1}} \geq y) - \underline{P}(N_{t_{j-1}} \geq y+1)\end{aligned} \tag{4.15}$$

So, the NPI lower probability $\underline{P}(N_{t_j} \geq x)$ stated in Equation (4.14), becomes

$$\underline{P}(N_{t_j} \geq x) = \sum_{y=x}^{m} \underline{P}(N_{t_j} \geq x|N_{t_{j-1}} = y)\left[\underline{P}(N_{t_{j-1}} \geq y) - \underline{P}(N_{t_{j-1}} \geq y+1)\right] \tag{4.16}$$

The argument $\left[\underline{P}(N_{t_{j-1}} \geq y) - \underline{P}(N_{t_{j-1}} \geq y+1)\right]$ stated in Equation (4.16) is achieved by applying to Equation (2.13) first for the event $N_{t_{j-1}} \geq y$ and then for the event $N_{t_{j-1}} \geq y+1$, for $y \in \{0,1,\ldots,m\}$ future observations. The argument $\left[\underline{P}(N_{t_{j-1}} \geq y) - \underline{P}(N_{t_{j-1}} \geq y+1)\right]$ is also achieved by using the Expression (2.20),

t_j	d_{t_j}	c_{t_j}	n_{t_j}	$\hat{n}_{t_j}$	$\hat{n}_{t_j}-d_{t_j}$		$x=0$		$x=1$		$x=2$		$x=3$	
							$\underline{P}$	$\overline{P}$	$\underline{P}$	$\overline{P}$	$\underline{P}$	$\overline{P}$	$\underline{P}$	$\overline{P}$
t_1	1	0	8	9	8	$y=0$	1	1						
						$y=1$	1	1	0.8000	0.9000				
						$y=2$	1	1	0.9455	0.9818	0.6545	0.8182		
						$y=3$	1	1	0.9818	0.9955	0.8727	0.9545	0.5455	0.7500
t_2	2	1	5	7	5	$y=0$	1	1						
						$y=1$	1	1	0.6250	0.7500				
						$y=2$	1	1	0.8333	0.9167	0.4167	0.5833		
						$y=3$	1	1	0.9167	0.9667	0.6667	0.8167	0.2917	0.4667
t_3	2	1	2	4	2	$y=0$	1	1						
						$y=1$	1	1	0.4000	0.6000				
						$y=2$	1	1	0.6000	0.8000	0.2000	0.4000		
						$y=3$	1	1	0.7143	0.8857	0.3714	0.6286	0.1143	0.2857
t_4	1	1	0	1	0	$y=0$	1	1						
						$y=1$	1	1	0	0.5000				
						$y=2$	1	1	0	0.6667	0	0.3333		
						$y=3$	1	1	0	0.7500	0	0.5000	0	0.2500

Table 4.5: NPI lower and upper for the event $N_{t_j} \geq x | N_{t_{j-1}} = y, (\hat{n}_{t_j}, \hat{n}_{t_j} - d_{t_j}))$ for $x \in \{0, 1, 2, 3\}$ and $y \in \{0, 1, 2, 3\}$, with $x \leq y$.

so

$$\left[\underline{P}(N_{t_{j-1}} \geq y | \left(\hat{n}_{t_j}, \hat{n}_{t_j} - d_{t_j}\right)) - \underline{P}(N_{t_{j-1}} \geq y+1 | \left(\hat{n}_{t_j}, \hat{n}_{t_j} - d_{t_j}\right))\right] = \binom{\hat{n}_{t_{j-1}} + m}{\hat{n}_{t_{j-1}}}^{-1} \binom{(\hat{n}_{t_{j-1}} - d_{t_{j-1}}) + y - 1}{(\hat{n}_{t_{j-1}} - d_{t_{j-1}}) - 1} \binom{\hat{n}_{t_{j-1}} - (\hat{n}_{t_{j-1}} - d_{t_{j-1}}) + m - y}{\hat{n}_{t_{j-1}} - (\hat{n}_{t_{j-1}} - d_{t_{j-1}})} \tag{4.17}$$

where $y \in \{0, 1, \ldots, m\}$ future observations. It should be remarked that the NPI lower probability for the event $(N_{t_{j-1}} \geq y + 1 | \left(\hat{n}_{t_j}, \hat{n}_{t_j} - d_{t_j}\right))$ for the case $m + 1$ is equal to 0.

Next, we present an example to illustrate the results presented in this section.

Example 4.5.1 To illustrate the NPI lower and upper for the event $N_{t_j} \geq x$ for $x \in \{0, 1, 2, 3\}$, presented above, we consider a simple example involving $n = 9$ observations, available at discrete times t_j, for $j = 1, 2, 3, 4$, (the data are summarised in Table 4.5).

						$x=0$		$x=1$		$x=2$		$x=3$	
t_j	d_{t_j}	c_{t_j}	n_{t_j}	$\hat{n}_{t_j}$	$\hat{n}_{t_j}-d_{t_j}$	$\underline{P}$	$\overline{P}$	$\underline{P}$	$\overline{P}$	$\underline{P}$	$\overline{P}$	$\underline{P}$	$\overline{P}$
t_1	1	0	8	9	8	1	1	0.9625	0.9955	0.7883	0.8832	0.4091	0.7500
t_2	2	1	5	7	5	1	1	0.8409	0.9432	0.5000	0.7318	0.1591	0.3500
t_3	2	1	2	4	2	1	1	0.5334	0.7834	0.1833	0.4334	0.0333	0.1333
t_4	1	1	0	1	0	1	1	0	0.5714	0	0.2571	0	0.0714

Table 4.6: NPI lower and upper for the event $N_{t_j} \geq x$, for $x \in \{0,1,2,3\}$.

Table 4.5 shows the NPI lower and upper probabilities for the event $N_{t_j} \geq x|N_{t_{j-1}} = y$ where $x \in \{0,1,2,3\}$ and $y \in \{0,1,2,3\}$, with $y \geq x$. Table 4.6 presents the NPI lower and upper for the event $(N_{t_j} \geq x)$, for $x \in \{0,1,2,3\}$ future observations. Note that some of the cells in Table 4.5 are empty due to the calculation of probabilities for the event that is at least x out of y future observations will survive a discrete time t_j, where $x \in \{0,1,2,3\}$ and $y \in \{0,1,2,3\}$.

From Table 4.5, we see that, at a specific discrete time t_j, the NPI lower and upper probabilities are decreasing in x when keeping y, $\hat{n}_{t_j}$ and $\hat{n}_{t_j} - d_{t_j}$ constant and increasing in y when keeping x, $\hat{n}_{t_j}$ and $\hat{n}_{t_j} - d_{t_j}$ constant, where x and y are varying from 0 to 3 with respect to that $y \geq x$. For $x = 0$, the NPI lower and upper probabilities for the event $N_{t_j} \geq x|N_{t_{j-1}} = y$, are equal 1, for $y \in \{0,1,2,3\}$ and for all t_j, due to the fact that no future observation out of y, $y \in \{0,1,2,3\}$, future observations will survive a discrete time t_j.

Based on the results provided in Table 4.5, the NPI lower and upper probabilities for the event $N_{t_j} \geq x$ are derived using to Equations (4.16) and (4.12), and shown in Table 4.6. From Table 4.6, we see that, the difference of the NPI lower and upper probabilities are decreasing in x when keeping m, $\hat{n}_{t_j}$ and $\hat{n}_{t_j} - d_{t_j}$, at each discrete time t_j, constant. Without any further assumptions added, the values of the NPI lower probability at t_4 are 0 for $x \in \{1,2,3\}$, whereas the NPI upper probabilities are positive.

The results presented in this subsection will be used in Subsection 4.5.3 to derive NPI lower and upper probabilities for discrete time system reliability for the event $T_S > t$, where T_S represents the random failure time of the system, using the concept of the survival signature [23, 25] combined with NPI for Bernoulli data [21]. As such, we will first provide a brief overview of the survival signatures in the following section.

4.5.2 The survival signature

The signature has been introduced to evaluate the reliability for systems consisting of only one type of components and is used to model the structure of a system, separating this from the random failure times of the components [4]. The NPI method is used in order to learn about the components within the system, based on data consisting of failure times for components that are exchangeable with those within the system. We therefore assume that such data are available, such as those obtained from testing or previous use of the components [4, 25]. Following the literature, the assumption of exchangeability is often replaced by the stronger assumption of independent and identically distributed (*iid*) component failure times [58]. Taking into account a system consisting of m components with exchangeable failure times, Samaniego [54, 57] introduced the system signature as a tool for reliability assessment for systems consisting of components of a single type. However, the use of signatures becomes very complicated in the case of quantifying reliability of systems with multiple types of components. Coolen and Coolen-Maturi [23] have introduced an alternative concept called the 'survival signature'. The idea of the survival signature is to generalise the signature to systems with multiple types of components. When quantifying the reliability of systems with only one type of components, the survival signature is closely related to the signature [23, 25]. The NPI methodology has been introduced for system reliability using the survival signature via lower and upper survival functions for the failure time T_S of a system consisting of multiple types of components [25], combined with NPI for Bernoulli data [20]. Aslett [11] created a package in the statistical software R to compute the survival signature, given a graphical representation of the system structure.

For a system with m exchangeable components, we need to consider the state vector $\underline{x} = (x_1, x_2, \ldots, x_m) \in \{0, 1\}^m$ taking into account that for each i, if the i^{th} component functions, then $x_i = 1$, otherwise $x_i = 0$ when the i^{th} component does not function. For all possible state vectors $\underline{x}$, the following structure function is defined as $\phi : \{0,1\}^m \rightarrow \{0,1\}$, so that $\phi(\underline{x}) = 1$ if the system functions and $\phi(\underline{x}) = 0$ if the system does not function. Throughout this section, the system is assumed to be coherent, which means that the structure function $\phi(\underline{x})$ must not be decreasing in any of the components of $\underline{x}$, and this leads to the fact that the functioning of the system cannot be improved by worse performance of one or more of its components. Furthermore, we assume that the system functions if all its components function, so $\phi(1) = 1$, and the system fails if all its components fail, so $\phi(0) = 0$.

For a system consisting only of m exchangeable components, the survival signature, denoted by $\Phi(l)$, for $l = 1, \ldots, m$, is defined as the probability that the system functions given that precisely l of its components function [23]. For coherent systems, $\Phi(l)$ is an increasing function of l, and assume that $\Phi(0) = 0$ and $\Phi(m) = 1$. There are $\binom{m}{l}$ state vectors $\underline{x}$ with precisely l components $x_i = 1$, so with $\sum_{i=1}^{m} x_i = l$; the set of these state vectors is denoted by S_l. Inspired by the *iid* assumption which has been considered for the failure times of the m components, all these state vectors are equally likely to occur [23]. Thus, the survival signature $\Phi(l)$ can be achieved as follows [23]

$$\Phi(l) = \binom{m}{l}^{-1} \sum_{\underline{x} \in S_l} \phi(\underline{x}) \tag{4.18}$$

Let $C(t) \in \{0, 1, \ldots, m\}$ represent the number of components in the system with a single type that function at time $t > 0$. The following equation holds for $l \in \{0, 1, \ldots, m\}$, in case the probability distribution of the *iid* component failure times is known and has CDF $F(t)$.

$$P(C(t) = l) = \binom{m}{l} [F(t)]^{m-l} [1 - F(t)]^l \tag{4.19}$$

Also, we can derive the probability for the event $T_S > t$ as follows.

$$P(T_S > t) = \sum_{l=0}^{m} \Phi(l) P(C(t) = l) \tag{4.20}$$

Then, the survival signature $\Phi(l)$ combined with NPI for Bernoulli data [4, 5] is used to present NPI lower and upper survival functions for T_S; the random failure time of a system, which consists of a single type of components. For a sinle type, we consider n to represent the number of components for which test failure data are available. These are not the components that are in the system but their failure times are assumed to be exchangeable with those in the system. Considering m to represent the number of components for which test failure data are available, as well as considering $s(t)$ to represent the number of components that are still functioning at time t, the NPI lower survival function for T_S is obtained as follows [4, 5].

$$P(T_S > t) \geq \sum_{l=0}^{m} \Phi(l)\overline{D}(C(t) = l) \tag{4.21}$$

where $\overline{D}(C(t) = l)$ can be computed by the following expression, in relationship to Expression (4.17),

$$\begin{aligned}\overline{D}(C(t) = l) &= \overline{P}(C(t) \leq l) - \overline{P}(C(t) \leq l - 1)\\ &= \binom{n+m}{n}^{-1}\binom{s(t)+l-1}{s(t)-1} \times \binom{n-s(t)+m-l}{n-s(t)}\end{aligned}$$

The notation $\overline{P}$ in this expression, indicates the NPI upper probability for Bernoulli data [20], given by Equation (2.12). The strict function $\overline{D}$ gives maximum possible probability mass to the small values through the event $C(t) = 0$, thus $D(C(t) = 0) = P(C(t) = 0)$ and then gives the maximum possible remaining probability mass, denoted by $\overline{D}(C(t) = 1)$, from the total probability mass available for the event $C(t) \leq 1$, to the event $C(t) = 1$. Thus, $\overline{D}(C(t) = 1)$ can be computed as follows, $\overline{D}(C(t) = 1) = \overline{P}(C(t) \leq l) - \overline{P}(C(t) = 0)$.

It is clear that the right-hand side of the inequality (4.21) is considered as a maximum possible lower bound due to the assumption that the survival signature $\Phi(l)$, is an increasing function of l for coherent systems, as well as the function, $\overline{D}$, is a probability distribution. As a result, the NPI lower probability for the event $T_S > t$, giving the NPI lower survival function for the system failure time for the event $t > 0$, can be achieved by the following equation [4, 25].

$$\underline{S}_{T_S}(t) = \underline{P}(T_S > t) = \sum_{l=0}^{m} \Phi(l)\overline{D}(C(t) = l) \tag{4.22}$$

Similarly, the corresponding NPI upper survival function for T_S can be obtained as follows [4].

$$P(T_S > t) \leq \sum_{l=0}^{m} \Phi(l)\underline{D}(C(t) = l) \tag{4.23}$$

where $\underline{D}(C(t) = l)$ can be computed by following expression.

$$\begin{aligned}\underline{D}(C(t) = l) &= \underline{P}(C(t) \leq l) - \underline{P}(C(t) \leq l-1)\\ &= \binom{n+m}{n}^{-1}\binom{s(t)+l}{s(t)} \times \binom{n-s(t)+m-l-1}{n-s(t)}\end{aligned}$$

The notation $\underline{P}$, appeared in this expression, indicates the NPI lower probability for Bernoulli data [20], given by Equation (2.13). The function $\underline{D}$ gives minimum possible weight to the small values of $C(t)$. Thus, the NPI upper probability for the event $T_S > t$, giving the NPI upper survival function for the system failure time for the event $t > 0$, can be achieved by the following equation.

$$\overline{S}_{T_S}(t) = \overline{P}(T_S > t) = \sum_{l=0}^{m} \Phi(l)\underline{D}(C(t) = l) \tag{4.24}$$

For a system consisting of $K \geq 2$ types of components, the survival signature, denoted by $\Phi(l_1, \ldots, l_K)$, for $l_k = 0, \ldots, m_k$, is defined as the probability that a system functions given that precisely l_k of its components of type k function, for each $k \in \{1, 2, \ldots, K\}$ [23]. There are $\binom{m_k}{l_k}$ state vectors $\underline{x}^k$ with precisely l_k of its m_k components $x_i^k = 1$, so with $\sum_{i=1}^{m_k} x_i^k = l_k$; we denote the set of these state vectors for components of type k by S_l^k. In addition, let $S_{l_1,\ldots,l_k}$ denote the set of these state vectors for the whole system for which $\sum_{i=1}^{m_k} x_i^k = l_k$, $k \in \{1, 2, \ldots, K\}$. Inspired by the *iid* assumption which has been considered for the failure times of the m_k components of type k, all these state vectors $\underline{x}^k \in$ are equally likely to occur [23]. Thus, the survival signature $\Phi(l_1, \ldots, l_K)$ can be achieved as follows [23].

$$\Phi(l_1, \ldots, l_K) = \left[\prod_{k=1}^{K} \binom{m_k}{l_k}^{-1}\right] \times \sum_{\underline{x} \in S_{l_1,\ldots,l_K}} \phi(\underline{x}) \tag{4.25}$$

Let $C_k(t) \in \{0, 1, \ldots, m_k\}$ represents the number of components of type k in the system which function at time $t > 0$. So, the probability that the system functions at time $t > 0$ is

$$P(T_S > t) = \sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(l_1, \ldots, l_K) P\left(\bigcap_{k=1}^{K}\{C_k(t) = l_k\}\right) \tag{4.26}$$

For each different type k, we consider n_k to represent the number of components of type k for which test failure data are available, as well as considering $s_k(t)$ to represent the number of components of type k still functioning at time t [4, 25]. Assuming that the failure times of components of different types are independent, while the exchangeability is assumed for the failure times of components of the same type [25]. Then the NPI lower survival function for T_S can be obtained as follows [25].

$$\underline{S}_{T_S}(t) = \underline{P}(T_S > t) = \sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(l_1 \ldots l_K) \prod_{k=1}^{K} \overline{D}(C_k(t) = l_k) \tag{4.27}$$

where $C_k(t) \in \{0, 1, \ldots, m_k\}$ represents the number of components of type k in the system functioning at time t, $k = 1, 2, \ldots, K$, while we can compute the $\overline{D}(C_k(t) = l_k)$ by using the following expression [25].

$$\begin{aligned}\overline{D}(C_k(t) = l_k) &= \overline{P}(C_k(t) \leq l_k) - \overline{P}(C_k(t) \leq l_k - 1) \\ &= \binom{n_k + m_k}{n_k}^{-1} \binom{s_k(t) + l - 1}{s_k(t) - 1} \times \binom{n_k - s_k(t) + m_k - l_k}{n_k - s_k(t)}\end{aligned}$$

where the notation $\overline{P}$ appearing in this expression refers to the NPI upper probability for Bernoulli data [20]. The function $\overline{D}$ gives maximum possible probability mass to the small values through the event $C_k(t) = 0$, thus $D(C_k(t) = 0) = P(C_k(t) = 0)$ and then gives the maximum possible remaining probability mass, denoted by $\overline{D}(C_k(t) = 1)$, from the total probability mass available for the event $C_k(t) \leq 1$, to the event $C_k(t) = 1$. Thus, $\overline{D}(C_k(t) = 1)$ can be computed as follows, $\overline{D}(CK(t) = 1) = \overline{P}(C_k(t) \leq l) - \overline{P}(C_k(t) = 0)$ [25].

Similarly, the corresponding NPI upper survival function for T_S is obtained as follows [25].

$$\overline{S}_{T_S}(t) = \overline{P}(T_S > t) = \sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(l_1 \ldots l_K) \prod_{k=1}^{K} \underline{D}(C_k(t) = l_k) \tag{4.28}$$

where $\underline{D}(C_k(t) = l)$ can be computed by following expression

$$\begin{aligned}\underline{D}(C_k(t) = l) &= \underline{P}(C_k(t) \leq l_k) - \underline{P}(C_k(t) \leq l_k - 1) \\ &= \binom{n_k + m_k}{n_k}^{-1} \binom{s_k(t) + l}{s_k(t)} \times \binom{n_k - s_k(t) + m_k - l_k - 1}{n_k - s_k(t)}\end{aligned}$$

where the notation $\underline{P}$ appearing in this expression refers to the NPI lower probability for Bernoulli data [20]. The strict function $\underline{D}$ gives minimum possible value to the small values of $C_k(t)$ [25].

4.5.3 Discrete time system reliability

In the case of discrete time, we use the survival signature $\Phi(l)$ combined with NPI for Bernoulli data [25], presented in Subsection 4.5.2, to present NPI lower and upper survival functions for the event $T_S > t_j$ of a system reliability consisting of a single type and multiple types of components, by implementing our proposed method as presented in Subsection 4.5.1. These survival functions are expressed as functions at the assumed discrete-time points t_j.

For a specific discrete time t_j, we apply the method presented in Subsection 4.5.1, stated explicitly in Equations (4.12) and (4.16) to system reliability using survival signatures combined with NPI for Bernoulli data. We begin by considering a system consisting of only one type of component, so $k = 1$.

Consider the reliability data, which consist of the numbers of components that failed at time t_j and the number of components that were right-censored at time t_j. It is assumed that these right-censored components are data from earlier tests, so not from the actual system. We consider the notation presented in Subsection 4.5.1 along with the notation presented in Subsection 4.5.2. Let $N_{t_j} \in \{0, 1, \ldots, m\}$ denote the number of components in the system, out of m, that are still functioning at a discrete time t_j.

We obtain the NPI lower and upper probabilities for the event that $T_S > t_j$ for a system consisting of a single type of components, using the survival signature $\Phi(l)$ combined with NPI for Bernoulli data [25], as follows

$$\underline{P}(T_S > t_j) = \sum_{\ell_1=0}^{m} \Phi(\ell_1)\overline{D}(N_{t_j} = \ell_1) \tag{4.29}$$

and

$$\overline{P}(T_S > t_j) = \sum_{\ell_1=0}^{m} \Phi(\ell_1)\underline{D}(N_{t_j} = \ell_1) \tag{4.30}$$

where $\overline{D}(N_{t_j} = \ell_1)$ and $\underline{D}(N_{t_j} = \ell_1)$ are derived from our results presented in

Equations (4.13) and (4.17), respectively, so

$$\begin{aligned}\overline{D}(N_{t_j} = \ell_1) &= \overline{P}(N_{t_j} \leq \ell_1) - \overline{P}(N_{t_j} \leq \ell_1 - 1)\\ &= 1 - \underline{P}(N_{t_j} \geq \ell_1 + 1) - \left[1 - \underline{P}(N_{t_j} \geq \ell_1)\right]\\ &= \underline{P}(N_{t_j} \geq \ell_1) - \underline{P}(N_{t_j} \geq \ell_1 + 1)\\ &= \binom{\hat{n}_{t_j} + m}{\hat{n}_{t_j}}^{-1} \binom{(\hat{n}_{t_j} - d_{t_j}) + \ell_1 - 1}{(\hat{n}_{t_j} - d_{t_j}) - 1} \times \binom{\hat{n}_{t_j} - (\hat{n}_{t_j} - d_{t_j}) + m - \ell_1}{\hat{n}_{t_j} - (\hat{n}_{t_j} - d_{t_j})}\end{aligned} \tag{4.31}$$

and

$$\begin{aligned}\underline{D}(N_{t_j} = \ell_1) &= \overline{P}(N_{t_j} \geq \ell_1) - \overline{P}(N_{t_j} \geq \ell_1 - 1)\\ &= \binom{\hat{n}_{t_j} + m}{\hat{n}_{t_j}}^{-1} \binom{(\hat{n}_{t_j} - d_{t_j}) + \ell_1}{(\hat{n}_{t_j} - d_{t_j})} \times \binom{\hat{n}_{t_j} - (\hat{n}_{t_j} - d_{t_j}) + m - \ell_1 - 1}{\hat{n}_{t_j} - (\hat{n}_{t_j} - d_{t_j}) - 1}\end{aligned} \tag{4.32}$$

We now consider a system consisting of $K \geq 2$ types of components with m_k components of $k \in \{1, 2, \ldots, K\}$, with $\sum_{k=1}^{K} m_k = m$. For a specific time t_j, let $\hat{n}_{t_j}^k$ denote the number of components of type k for which test failure data are available, and $d_{t_j}^k$ denote the numbers of components that failed at time t_j, and therefore, $\hat{n}_{t_j}^k - d_{t_j}^k$ is the number of components of type k that are still functioning at time t_j [4, 25]. As assumed that the failure times of components of different types are assumed to be independent, while failure times of components of the same type are assumed to be exchangeable [25]. Let $N_{t_j}^k \in \{0, 1, \ldots, m_k\}$ denote the number of components of type k in the system, out of m_k, that are still functioning at a discrete time t_j, $k = 1, 2, \ldots, K$.

The NPI lower and upper probabilities for the event $T_S > t_j$ of a system consisting of multiple types of components, using the survival signature $\Phi(l_1, \ldots, l_k)$ combining with NPI for Bernoulli data [25], are obtained as follows

$$\underline{P}(T_S > t_j) = \sum_{\ell_1=0}^{m_1} \cdots \sum_{\ell_K=0}^{m_K} \Phi(\ell_1 \ldots \ell_K) \prod_{k=1}^{K} \overline{D}(N_{t_j}^k = \ell_k) \tag{4.33}$$

and

$$\overline{P}(T_S > t_j) = \sum_{\ell_1=0}^{m_1} \cdots \sum_{\ell_K=0}^{m_K} \Phi(\ell_1 \ldots \ell_K) \prod_{k=1}^{K} \underline{D}(N_{t_j}^k = \ell_k) \tag{4.34}$$

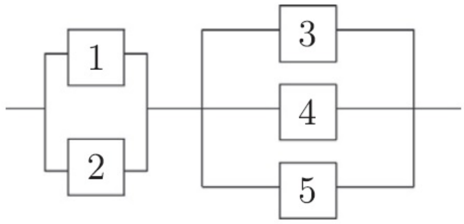

Figure 4.1: System with a single type of $m = 5$ components for Example 4.5.2.

where the values of $\overline{D}(N^k_{t_j} = \ell_k)$ and $\underline{D}(N^k_{t_j} = \ell_k)$, for $\ell_k \in \{0, 1, \ldots, m_k\}$, respectively, are derived from our results presented in Equations (4.13) and (4.17), respectively, thus

$$\begin{aligned}\overline{D}(N^k_{t_j} = \ell_k) &= \overline{P}(N^k_{t_j} \leq \ell_k) - \overline{P}(N^k_{t_j} \leq \ell_k - 1)\\ &= 1 - \underline{P}(N^k_{t_j} \geq \ell_k + 1) - \left[1 - \underline{P}(N^k_{t_j} \geq \ell_k)\right]\\ &= \underline{P}(N^k_{t_j} \geq \ell_k) - \underline{P}(N^k_{t_j} \geq \ell_k + 1)\\ &= \binom{\hat{n}^k_{t_j} + m_k}{\hat{n}^k_{t_j}}^{-1} \binom{(\hat{n}^k_{t_j} - d^k_{t_j}) + \ell_k - 1}{(\hat{n}^k_{t_j} - d^k_{t_j}) - 1} \times \binom{\hat{n}^k_{t_j} - (\hat{n}^k_{t_j} - d^k_{t_j}) + m_k - \ell_k}{\hat{n}^k_{t_j} - (\hat{n}^k_{t_j} - d^k_{t_j})}\end{aligned} \tag{4.35}$$

and

$$\begin{aligned}\underline{D}(N^k_{t_j} = \ell_k) &= \overline{P}(N^k_{t_j} \geq \ell_k) - \overline{P}(N^k_{t_j} \geq \ell_k - 1)\\ &= \binom{\hat{n}^k_{t_j} + m_k}{\hat{n}^k_{t_j}}^{-1} \binom{(\hat{n}^k_{t_j} - d^k_{t_j}) + \ell_k}{(\hat{n}^k_{t_j} - d^k_{t_j})} \times \binom{\hat{n}^k_{t_j} - (\hat{n}^k_{t_j} - d^k_{t_j}) + m_k - \ell_k - 1}{\hat{n}^k_{t_j} - (\hat{n}^k_{t_j} - d^k_{t_j}) - 1}\end{aligned} \tag{4.36}$$

Next, we apply our results, presented in this subsection, to some discrete time systems reliability consisting of a single type of components and multiple types of components.

4.5.4 Examples

This section presents two examples to illustrate the methodology presented in Subsection 4.5.3.

t_j	d_{t_j}	c_{t_j}	$\hat{n}_{t_j}$	$\hat{n}_{t_j} - d_{t_j}$	$\underline{P}(T_S > t_j)$	$\overline{P}(T_S > t_j)$
t_1	2	0	10	8	0.8811	0.9426
t_2	2	1	7	5	0.7765	0.8909
t_3	2	0	5	3	0.6190	0.8095
t_4	1	1	2	1	0.3857	0.7810
t_5	1	0	1	0	0	0.5833

Table 4.7: NPI lower and upper probabilities for $T_S > t_j$, for system in Figure 4.1, with $n = 10$ (Example 4.5.2).

t_j	d_{t_j}	c_{t_j}	$\hat{n}_{t_j}$	$\hat{n}_{t_j} - d_{t_j}$	$\underline{P}(T_S > t_j)$	$\overline{P}(T_S > t_j)$
t_1	4	0	20	16	0.9177	0.9465
t_2	4	2	14	10	0.8344	0.8921
t_3	3	1	9	6	0.7552	0.8559
t_4	2	2	4	2	0.4810	0.7333
t_5	2	0	2	0	0	0.3857

Table 4.8: NPI lower and upper probabilities for $T_S > t_j$, for system in Figure 4.1, with $n = 20$ (Example 4.5.2).

Example 4.5.2 The system in Figure 4.1 is used in this example was also used by Coolen and Coolen-Maturi [24]. We consider a discrete time system reliability with $m = 5$ exchangeable components, which has a survival signature with the following values: $\Phi(0) = 0$, $\Phi(1) = 0$, $\Phi(2) = 0.6$, $\Phi(3) = 0.9$, $\Phi(4) = 1$, and $\Phi(5) = 1$. Suppose that we have data consisting of $n = 10$ observations, including failure events and right-censored observations, for discrete times $t_1, \ldots, t_5$. We consider the data given in Table 4.7 for this example, which contains the numbers of components that failed at times $t_1, \ldots, t_5$, with $d_{t_1} = 2$, $d_{t_2} = 2$, $d_{t_3} = 2$, $d_{t_4} = 1$, and $d_{t_5} = 1$, along with the number of right-censored observations at times t_2 and t_4 with $c_{t_2} = 1$ and $c_{t_4} = 1$, respectively. Table 4.7 presents NPI lower and upper probabilities for $T_S > t_j$ at discrete times $t_1, \ldots, t_5$, with $n = 10$, based on using the survival signature values as given previously and the NPI for Bernoulli data.

We now increase the size of the data set for the same system in Figure 4.1, such

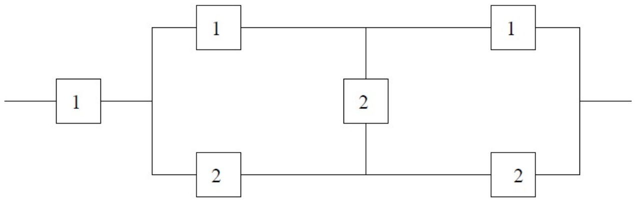

Figure 4.2: System with 2 types of components for Example 4.5.3.

	Type 1				Type 2			
t_j	$d^1_{t_j}$	$c^1_{t_j}$	$\hat{n}^1_{t_j}$	$(\hat{n}^1_{t_j} - d^1_{t_j})$	$d^2_{t_j}$	$c^2_{t_j}$	$\hat{n}^2_{t_j}$	$(\hat{n}^2_{t_j} - d^2_{t_j})$
t_1	2	1	9	7	3	0	10	7
t_2	3	2	5	2	3	1	6	3
t_3	2	0	2	0	2	1	2	0

Table 4.9: Data of a system with 2 types of components with $m_1 = m_2 = 3$, in Figure 4.2 (Example 4.5.3).

that the data set consists of $n = 20$ observations, including failure events and right-censored observations, for discrete times $t_1, \ldots, t_5$. Table 4.8 presents NPI lower and upper probabilities for $T_S > t_j$ at discrete times $t_1, \ldots, t_5$, with $n = 20$, based on using the survival signature values as given previously and the NPI for Bernoulli data.

In comparing the results in Tables 4.7 and 4.8, we observe that the imprecision in both tables is quite small at time t_1 and becomes larger later on, due to fewer observations in the risk set later on. Further, the differences between the lower and upper probabilities for $T_S > t_j$ for all t_j, with $n = 20$, according to Table 4.8, are quite smaller than those in Table 4.7 with $n = 10$ observations. So, the difference between the lower and upper probabilities for $T_S > t_j$ for all t_j decreases as the data set size increases.

Example 4.5.3 In this example, we consider the system with $K = 2$ types of components, types 1 and 2, as presented in Figure 4.2. This system was used by Coolen et al [23] to illustrate NPI for the system survival time. The survival

(ℓ_1, ℓ_2)	$\Phi(\ell_1, \ell_2)$	(ℓ_1, ℓ_2)	$\Phi(\ell_1, \ell_2)$
$(0, 0)$	0	$(2, 0)$	0
$(0, 1)$	0	$(2, 1)$	0
$(0, 2)$	0	$(2, 2)$	4/9
$(0, 3)$	0	$(2, 3)$	6/9
$(1, 0)$	0	$(3, 0)$	1
$(1, 1)$	0	$(3, 1)$	1
$(1, 2)$	1/9	$(3, 2)$	1
$(1, 3)$	3/9	$(3, 3)$	1

Table 4.10: Survival signature of the system in Figure 4.2 (Example 4.5.3).

signature for this system is presented in Table 4.10. We consider the data given in Table 4.9 for the two types with $m_1 = m_2 = 3$ components, and each type has 10 observations, i.e. $n_1 = n_2 = 10$, including failure events and right-censored observations, for discrete times t_1, t_2 and t_3. Table 4.11 presents the NPI lower and upper probabilities for $T_S > t_j$ at discrete times t_1, t_2 and t_5, based on using the given survival signature values and the NPI for Bernoulli data.

According to the NPI approach for real-valued data, it is natural for the lower probability value for $X_{n+1} > t$ when t is in a specific interval, to be less than or equal the upper probability value for $X_{n+1} > t$ when t is in the next interval, e.g. see the results in Table 5.7. In contrast, the NPI for discrete-time approach indicates that this does not always hold true, e.g. see the results of Table 4.11 as $\underline{P}(T_S > t_1) > \overline{P}(T_S > t_2)$ and see also the results of Table 4.8 as $\underline{P}(T_S > t_3) > \overline{P}(T_S > t_4)$. Most of the results presented in this chapter suggest that for discrete time cases, this issue may occur due to multiple failures that could occur in between the discrete-time points.

t_j	$\underline{P}(T_S > t_j)$	$\overline{P}(T_S > t_j)$
t_1	0.5500	0.7118
t_2	0.1412	0.3189
t_3	0	0.1478

Table 4.11: NPI lower and upper probabilities for $T_S > t_j$, for the system, with two types of components, in Figure 4.2 (Example 4.5.3).

4.6 Concluding remarks

Considering discrete-time data in this chapter, NPI method in terms of utilising NPI for Bernoulli data [20] are developed as an alternative predictive approach to the actuarial estimator dealing with right-censored data. This development allows us to present the NPI lower and upper probabilities for the event that all future observations survive a discrete time t_j. Further, one of the primary objectives that have been discussed in this chapter is the comparison between the proposed method and the NPI method for grouped data with right-censored data [65].

Taking into account that n Bernoulli trials are exchangeable with m future Bernoulli trials, the proposed method in this chapter, based on NPI for Bernoulli data [65], has been developed for deriving the NPI lower and upper probabilities for the event that there are x out of m future Bernoulli trials. Together with the survival signature method, this development has been applied to systems reliability for the event $T_S > t_j$ with single and multiple types of components at discrete times t_j.

Chapter 5

NPI for Two Future Observations with Right-Censored Data

5.1 Introduction

Coolen and Yan [32] presented NPI for right-censored data for a single future observation. There is a challenge to generalize the approach of NPI for right-censored data to multiple future observations. NPI has been developed to multiple future observations for uncensored real-valued data [9, 10] and for Bernoulli data [20], however, this is complicated for right-censored data. In this thesis, further theory is developed on NPI for two future observations with attention to right-censored data. This will be achieved by applying the rc-$A_{(n)}$ assumption [32], without further assumptions, for X_{n+1} and, conditionally on X_{n+1}, applying the rc-$A_{(n+1)}$ assumption for X_{n+2}. The focus is on NPI lower and upper probabilities for the event that both future observations X_{n+1} and X_{n+2} are greater than time t. We also illustrate how the proposed method can be applied to system reliability.

To achieve the aim of the chapter, which is to derive the joint lower and upper probabilities for the event $X_{n+1} > t$ and $X_{n+2} > t$, more notation needs to be introduced in addition to those introduced in Section 2.4.

In this chapter, the notation introduced in Section 2.4 is followed, together with new notations required. Let $X_1, X_2, \ldots, X_n, X_{n+1}$ be positive, continuous and exchangeable random quantities representing lifetimes. Suppose that there are n observations, including u event times, $x_1 < x_2 < \cdots < x_u$, and $\nu = n - u$ right-censored observations, $c_1 < c_2 < \ldots < c_\nu$. Let $x_0 = 0$ and $x_{u+1} = \infty$ for ease of notation. Suppose further that there are s_i right-censored observations in the interval $I^i = (x_i, x_{i+1})$, $i = 0, 1, \ldots, u$, denoted by $c_1^i < c_2^i < \ldots < c_{s_i}^i$, so $\sum_{i=0}^{u} s_i = \nu$. With regard to X_{n+1}, there are $n+1$ intervals created by the data set n including the right-censored observations. In order to simplify our presentation, we assume that no ties are present in our data, so no two observations (events or right-censoring) happen at the same time. The assumption rc-$A_{(n)}$ [32] partially specifies the NPI-based probability distribution for X_{n+1} by the M-function values, as given in Equations (2.31) and (2.32). The probability for the event that $X_{n+1} \in (x_i, x_{i+1})$, $i = 0, 1, \ldots, u$, where x_i and x_{i+1} are two consecutive failure times, is obtained by summing up all M-function values assigned to the interval (x_i, x_{i+1}), as given by the Equation (2.33).

This chapter is organised as follows. Using a new approach, we reformulate the NPI lower and upper probabilities for the event $X_{n+1} > t$ in Section 5.2. In Section 5.3, we present NPI for the event $X_{n+2} > t$ given $X_{n+1} > t$. NPI for the joint event $X_{n+1} > t$ and $X_{n+2} > t$ is presented in Section 5.4. Section 5.5 illustrates how these inferences can be applied to quantify the reliability of a small series system. Finally, this chapter ends with concluding remarks in Section 5.6.

5.2 Reformulating NPI for the first future observation

The overall goal of this chapter is to develop NPI for two future observations, X_{n+1} and X_{n+2}, with data including right-censored observations. Particularly, we present NPI lower and upper probabilities for the event $X_{n+1} > t$ and $X_{n+2} > t$. According to the rc-$A_{(n)}$ [32] assumption, the probability distribution for X_{n+1} is partially specified by probability mass assigned to open nested intervals via M-function values, without further restrictions on where it is in each interval. We consider X_{n+1} and X_{n+2} such that X_{n+2} is conditioned on X_{n+1} and the data set that contains n observations with right-censored observations. Without making any further assumptions, we aim to apply the rc-$A_{(n)}$ [32] assumption for X_{n+1}, and then, conditionally on X_{n+1}, we will apply the rc-$A_{(n+1)}$ assumption for X_{n+2}. However, where to put the probability masses for specific events of interest in order to get the NPI lower and upper probabilities is challenging. So, we must consider where the probability mass is for X_{n+1} within an interval (x_i, x_{i+1}), in order to apply rc-$A_{(n+1)}$ for X_{n+2}. In this case, this interval (x_i, x_{i+1}), which contains right-censored observations, must be specified into sub-intervals $(c^i_{i^*}, x_{i+1})$, $i^* = 1, 2, \ldots, s_i$, with respect to that the probability mass for X_{n+1} according to its M-function value assigned to the interval (x_i, x_{i+1}), will be distributed over these sub-intervals $(c^i_{i^*}, x_{i+1})$. To do this, we introduce probabilities denoted by $\underline{\alpha}^i$ and $\underline{\alpha}^{c^i_{i^*}}$, $i = 0, 1, \ldots, u$ and $i^* = 1, 2, \ldots, s_i$, to enable us to determine where to put the probability mass per interval over its sub-intervals. In this way, we can minimise or maximise the probability for any event of interest involving the one or two future observations with regard to the $\underline{\alpha}^i$ and $\underline{\alpha}^{c^i_{i^*}}$ values. Overall, this allows deriving the NPI lower and upper probabilities for the event $X_{n+1} > t$.

To achieve the overall goal, we start with deriving the lower and upper probabilities for the event $X_{n+1} > t$, which has been done by Coolen and Yan [32], in different way. For an interval $I^i = (x_i, x_{i+1})$, $i = 0, 1, 2, \ldots, u$, there are s_i right-censored observations in this interval, and

$$\underline{\alpha}^i = (\alpha^i_1, \alpha^i_2, \ldots, \alpha^i_{s_i+1}), \text{where } 0 \leq \alpha^i_{i^*} \leq 1 \text{ and } \sum_{i^*=1}^{s_i+1} \alpha^i_{i^*} = 1$$

If there are no censored observations in the interval (x_i, x_{i+1}), that is $s_i = 0$, then $\underline{\alpha}^i = \alpha^i_1 = 1$. Also, for each censored observation $c^i_{i^*}$, $i^* = 1, 2, \ldots, s_i$, in the interval (x_i, x_{i+1}),

$$\underline{\alpha}^{c^i_{i^*}} = (\alpha_1^{c^i_{i^*}}, \alpha_2^{c^i_{i^*}}, \ldots, \alpha_{s_i-i^*+1}^{c^i_{i^*}}), \text{where } 0 \leq \alpha_l^{c^i_{i^*}} \leq 1 \text{ and } \sum_{l=1}^{s_i-i^*+1} \alpha_l^{c^i_{i^*}} = 1.$$

If there is only one censored observation in the interval (x_i, x_{i+1}) then $\underline{\alpha}^{c^i_{i^*}} = \alpha_1^{c^i_{i^*}} = 1$.

The notation $\underline{\alpha}^i$ and $\underline{\alpha}^{c^i_{i^*}}$ are the proportion of (a specific) probability mass assigned to the intervals (x_i, x_{i+1}) and (c^i_l, x_{i+1}), respectively, that are distributed over sub-intervals. It is just a way to write how the probability mass is divided over sub-intervals, so that we can then find the NPI lower and upper probabilities for any event of interest involving X_{n+1}. The $\alpha^i_{i^*}$ are introduced to determine where to place the probability mass per interval (x_i, x_{i+1}) over its sub-intervals, whereas the $\alpha_l^{c^i_{i^*}}$ are introduced to determine where to place the probability mass per interval (c^i_l, x_{i+1}) over its sub-intervals.

Consequently, we reformulate the original M-function masses shown in Definition 2.4.4, having the notation $\underline{\alpha}^i$ introduced to them, to specify how much of each M-function value is in sub-intervals.

Definition 5.2.1 (rc-$A_{(n)}$- revisited)

Let $I^i_{i^*} = (t^i_{i^*}, t^i_{i^*+1})$ represent an interval created by the n data observations, where $i = 0, 1, 2, \ldots, u$, and

$$\begin{cases} i^* = 0 & \text{if} \quad t^i_0 = x_i \quad \text{(failure time or time 0)} \\ i^* = 1, 2, \ldots s_i, & \text{if} \quad t^i_{i^*} = c^i_{i^*} \quad \text{(right-censoring time)} \end{cases}$$

and for simplicity of notation let $t^i_{s_i+1} = t^{i+1}_0 = x_{i+1}$. Using α approach, the assumption rc-$A_{(n)}$ partially specifies the NPI-based probability distribution for the observable, nonnegative and real-valued random quantity X_{n+1}, via the following M-function values.

$$M_{X_{n+1}}(t^i_{i^*}, t^i_{i^*+1}) = \alpha^i_{i^*+1} M_{X_{n+1}}(x_i, x_{i+1}) + \sum_{l=1}^{i^*} \alpha^{c^i_l}_{i^*-l+1} M_{X_{n+1}}(c^i_l, x_{i+1}) \qquad (5.1)$$

In Equation (5.1), the M-function values $M_{X_{n+1}}(x_i, x_{i+1})$ and $M_{X_{n+1}}(c^i_{i^*}, x_{i+1})$ are derived from Equations (2.31) and (2.32), respectively. The $M_{X_{n+1}}(t^i_{i^*}, t^i_{i^*+1})$, stated in Equation (5.1), could be also written as $M_{X_{n+1} \in I^i_{i^*}}$. We do this for convenience in order to be used later in Section 5.4.

With respect to that for all $\alpha^i_{i^*} \in [0,1]$, $\alpha_l^{c^i_{i^*}} \in [0,1]$, $\sum_{i^*=1}^{s_i+1} \alpha^i_{i^*} = 1$ and $\sum_{l=1}^{s_i-i^*+1} \alpha_l^{c^i_{i^*}} = 1$, the M-function values as specified by rc-$A_{(n)}$ in Definition 5.2.1 lead to the probability for the event that $X_{n+1} \in (x_i, x_{i+1})$, $i = 0, 1, \ldots, u$, denoted by $P_{X_{n+1}}(x_i, x_{i+1})$, which can be calculated by summing up all M-function values assigned to the interval $I^i = (x_i, x_{i+1})$ along with all M-function values assigned to them sub-intervals $(c^i_{i^*}, x_{i+1})$ for X_{n+1}, so that

$$\begin{aligned}
P_{X_{n+1}}(x_i, x_{i+1}) &= \sum_{i^*=0}^{s_i} M_{X_{n+1}}(t^i_{i^*}, t^i_{i^*+1}) \\
&= \sum_{i^*=0}^{s_i} \alpha^i_{i^*+1} M_{X_{n+1}}(x_i, x_{i+1}) + \sum_{i^*=1}^{s_i}\sum_{l=1}^{i^*} \alpha^{c^i_l}_{i^*-l+1} M_{X_{n+1}}(c^i_l, x_{i+1}) \\
&= M_{X_{n+1}}(x_i, x_{i+1}) + \sum_{l=1}^{s_i}\sum_{i^*=1}^{s_i-l+1} \alpha^{c^i_l}_{i^*} M_{X_{n+1}}(c^i_l, x_{i+1}) \\
&= M_{X_{n+1}}(x_i, x_{i+1}) + \sum_{l=1}^{s_i} M_{X_{n+1}}(c^i_l, x_{i+1}) \qquad (5.2)
\end{aligned}$$

for $i = 0, 1, \dots, u$. The result in Equation (5.2) is the same result stated in Equation (2.33). For convenience, $P_{X_{n+1}}(x_i, x_{i+1})$, stated in Equation (5.2), will be also denoted by $P_{X_{n+1} \in I^i}$ We do this for convenience in order to be used later in Section 5.4. The first term after the second equality in Equation (5.2) is the sum of all M-function values assigned to the interval (x_i, x_{i+1}), and as $\sum_{i^*=1}^{s_i+1} \alpha^i_{i^*} = 1$, this first term is equal to $M_{X_{n+1}}(x_i, x_{i+1})$. The second term after the second equality in Equation (5.2) is the sum of all M-function values assigned to the sub-intervals (c^i_l, x_{i+1}) of (x_i, x_{i+1}), and as $\sum_{l=1}^{s_i-i^*+1} \alpha^{c^i_{i^*}}_l = 1$, for $i = 0, 1, \dots, u$ and $i^* = 1, 2, \dots, s_i$, this second term is equal to $\sum_{l=1}^{s_i} M_{X_{n+1}}(c^i_l, x_{i+1})$. Let us define the following equation

$$Q_{X_{n+1}}(t^i_a, x_{i+1}) = \sum_{i^*=a}^{s_i} \alpha^i_{i^*+1} M_{X_{n+1}}(x_i, x_{i+1}) + \sum_{i^*=a}^{s_i}\sum_{l=1}^{i^*} \alpha^{c^i_l}_{i^*-l+1} M_{X_{n+1}}(c^i_l, x_{i+1}) \qquad (5.3)$$

where for $a = 0$, Equation (5.2) and (5.3) are equivalent. This Equation (5.3) can be minimised or maximised in order to derive the NPI lower and upper probabilities for the event $X_{n+1} > t$. We sometimes denote the probability in Equation (5.3) by $Q_{X_{n+1} \in I^i_a}$ for convenience in order to be used later in Section 5.4. Now, let us consider the second term of Equation (5.3), and by rearranging the summations, we have

$$\begin{aligned}
&\sum_{i^*=a}^{s_i}\sum_{l=1}^{i^*} \alpha^{c^i_l}_{i^*-l+1} M_{X_{n+1}}(c^i_l, x_{i+1}) = \\
&\qquad \sum_{l=1}^{a-1}\sum_{i^*=a}^{s_i-l+1} \alpha^{c^i_l}_{i^*} M_{X_{n+1}}(c^i_l, x_{i+1}) + \sum_{l=a}^{s_i}\sum_{i^*=1}^{s_i-l+1} \alpha^{c^i_l}_{i^*} M_{X_{n+1}}(c^i_l, x_{i+1}) \qquad (5.4)
\end{aligned}$$

The first term on the right-hand side of Equation (5.4) is related to the probability masses to the right of t^i_a, corresponding to all $c^i_l < t^i_a$. The second term in

Equation (5.4) is related to the probability masses corresponding to all $c_l^i \geq t_a^i$, and as $\sum_{i^*=1}^{s_i-l+1} \alpha_{i^*}^{c_l^i} = 1$, this second term is equal to $\sum_{l=a}^{s_i} M_{X_{n+1}}(c_l^i, x_{i+1})$. So Equation (5.3) can be rewritten as

$$Q_{X_{n+1}}(t_a^i, x_{i+1}) = \sum_{i^*=a}^{s_i} \alpha_{i^*+1}^i M_{X_{n+1}}(x_i, x_{i+1}) + \sum_{l=1}^{a-1} \sum_{i^*=a}^{s_i-l+1} \alpha_{i^*}^{c_l^i} M_{X_{n+1}}(c_l^i, x_{i+1}) + \sum_{l=a}^{s_i} M_{X_{n+1}}(c_l^i, x_{i+1}) \tag{5.5}$$

If we want to find the values of $\underline{\alpha}^i$'s and $\underline{\alpha}^{c_l^i}$'s that minimize $Q_{X_{n+1}}(t_a^i, x_{i+1})$, stated in Equation (5.5), then this can be achieved by assigning all probability masses in (x_i, x_{i+1}) which can be assigned to the left of the t_a^i, that is

$$\sum_{i^*=a}^{s_i} \alpha_{i^*+1}^i = 0 \ , \qquad \sum_{i^*=0}^{a-1} \alpha_{i^*+1}^i = 1$$

and

$$\sum_{i^*=1}^{a-1} \alpha_{i^*}^{c_l^i} = 1 \ , \qquad \sum_{i^*=a}^{s_i-l+1} \alpha_{i^*}^{c_l^i} = 0$$

thus, the minimum value of $Q_{X_{n+1}}(t_a^i, x_{i+1})$ in Equation (5.5) is

$$Q_{X_{n+1}}^{\min}(t_a^i, x_{i+1}) = \sum_{l=a}^{s_i} M_{X_{n+1}}(c_l^i, x_{i+1}) \tag{5.6}$$

If we want to find the values of $\underline{\alpha}^i$'s and $\underline{\alpha}^{c_l^i}$'s that maximise $Q_{X_{n+1}}(t_a^i, x_{i+1})$, given in Equation (5.5), then this can be achieved by assigning all probability masses in the interval (x_i, x_{i+1}) to the right of t_a^i, that is

$$\sum_{i^*=a}^{s_i} \alpha_{i^*+1}^i = 1 \ , \qquad \sum_{i^*=0}^{a-1} \alpha_{i^*+1}^i = 0$$

and

$$\sum_{i^*=1}^{a-1} \alpha_{i^*}^{c_l^i} = 0 \ , \qquad \sum_{i^*=a}^{s_i-l+1} \alpha_{i^*}^{c_l^i} = 1$$

thus, the maximum value of $Q_{X_{n+1}}(t_a^i, x_{i+1})$ in Equation (5.5) is

$$\begin{aligned} Q_{X_{n+1}}^{\max}(t_a^i, x_{i+1}) &= M_{X_{n+1}}(x_i, x_{i+1}) + \sum_{l=1}^{a-1} M_{X_{n+1}}(c_l^i, x_{i+1}) + \sum_{l=a}^{s_i} M_{X_{n+1}}(c_l^i, x_{i+1}) \\ &= M_{X_{n+1}}(x_i, x_{i+1}) + \sum_{l=1}^{s_i} M_{X_{n+1}}(c_l^i, x_{i+1}) \\ &= P_{X_{n+1}}(x_i, x_{i+1}) \end{aligned} \tag{5.7}$$

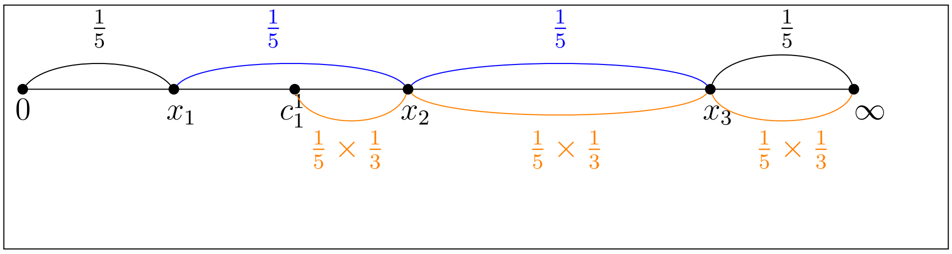

Figure 5.1: The original M-functions based on rc-$A_{(n)}$ assumption for X_5.

The probabilities, $Q^{\min}_{X_{n+1}}(t^i_a, x_{i+1})$ and $Q^{\max}_{X_{n+1}}(t^i_a, x_{i+1})$, given in Equations (5.6) and (5.7), are denoted by $Q^{\min}_{X_{n+1}\in I^i_a}$ and $Q^{\max}_{X_{n+1}\in I^i_a}$, respectively, for convenience in order to be used later in Section 5.4.

Consequently, the NPI lower probability for the event $X_{n+1} > t$, for $t \in [t^i_a, t^i_{a+1})$ with $i = 0, 1, \ldots, u$ and $a = 0, 1, \ldots, s_i$, given in Equation (2.34), can be written as follows.

$$\begin{aligned}\underline{P}(X_{n+1} > t) &= Q^{\min}_{X_{n+1}}(t^i_{a+1}, x_{i+1}) + \sum_{j=i+1}^{u} P_{X_{n+1}}(x_j, x_{j+1}) \\ &= \sum_{l=a+1}^{s_i} M_{X_{n+1}}(c^i_l, x_{i+1}) + \sum_{j=i+1}^{u} P_{X_{n+1}}(x_j, x_{j+1}) \qquad (5.8)\end{aligned}$$

The corresponding NPI upper probability for the event $X_{n+1} > t$, for $t \in [x_i, x_{i+1})$ with $i = 1, 2, \ldots, u$ and $a = 0, 1, \ldots, s_i$, given in Equation (2.35), can be written as follows.

$$\begin{aligned}\overline{P}(X_{n+1} > t) &= Q^{\max}_{X_{n+1}}(t^i_a, x_{i+1}) + \sum_{j=i+1}^{u} P_{X_{n+1}}(x_j, x_{j+1}) \\ &= \sum_{j=i}^{u} P_{X_{n+1}}(x_j, x_{j+1}) \qquad (5.9)\end{aligned}$$

The following example illustrates the above method.

Example 5.2.1 Suppose that a data set consists of three failure observations at times x_1, x_2, x_3 and one right-censored observation at time c^1_1, as shown in Figure 5.1. Let $X_{c^1_1}$ denote the random quantity corresponding to the right-censoring at time c^1_1, where $c^1_1 \in (x_1, x_2)$.

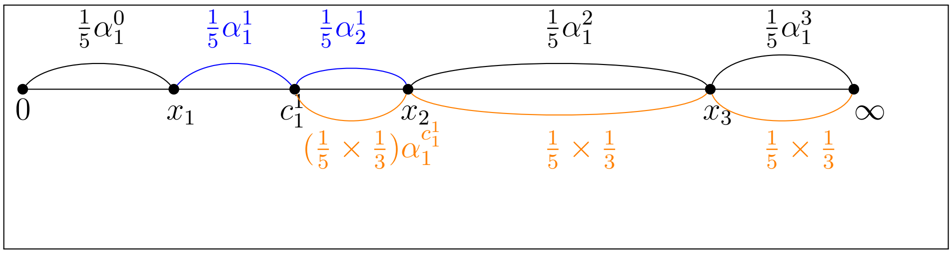

Figure 5.2: Reformulating the original M-functions for X_5.

Let us begin by briefly illustrating the assumption rc-$A_{(n)}$ [32]. According to the $\tilde{A}_{(4)}$ assumption, given by Definition 2.4.1, the M-function values for X_5 are $\tilde{M}_{X_5}(0, x_1) = \tilde{M}_{X_5}(x_1, x_2) = \tilde{M}_{X_5}(x_2, x_3) = \tilde{M}_{X_5}(x_3, \infty) = \frac{1}{5}$, and a further probability mass 1/5 is distributed over the interval (c_1^1, ∞), i.e. $\tilde{M}_{X_5}(c_1^1, \infty) = \frac{1}{5}$, since it is known without making any further assumptions that X_5 will be at any point beyond c_1^1.

According to the non-informative censoring assumption, the residual lifetime of the censored observation is independent of the censoring process, therefore, the assumption shifted-$\tilde{A}_{(2)}$, given by Definition 2.4.2, allows us to apply $A_{(2)}$ with the starting point shifted from 0 to the censoring time c_1^1. Based on the assumption shifted-$\tilde{A}_{(2)}$, the probability distribution for $X_{c_1^1}$, given $X_{c_1^1} > c_1^1$, is partially specified via M-function values for $X_{c_1^1}$ assigned to sub-intervals as $M_{X_{c_1^1}}(c_1^1, x_2) = M_{X_{c_1^1}}(x_2, x_3) = M_{X_{c_1^1}}(x_3, \infty) = \frac{1}{3}$. Moreover, the assumption rc-$\tilde{A}_{(4)}$, given by Definition 2.4.3, splits the probability mass of $\tilde{M}_{X_5}(c_1^1, \infty) = \frac{1}{5}$ to M-function values for X_5 assigned to sub-intervals as $M_{X_5}^{c_1^1}(c_1^1, x_2) = M_{X_5}^{c_1^1}(x_2, x_3) = M_{X_5}^{c_1^1}(x_3, \infty) = \frac{1}{5} \times \frac{1}{3} = \frac{1}{15}$ (see Figure 5.1).

The M-function values for X_5 based on the assumption $\tilde{A}_{(4)}$, given by Definition 2.4.1, are then combined with the M-function values for X_5 based on the assumption rc-$\tilde{A}_{(4)}$, given by Definition 2.4.3, leading to the M-function values for X_5 based on the rc-$A_{(4)}$ assumption, as given by the Definition 2.4.4 [32, 65]. For example, the M-function value for the event $X_5 \in (x_2, x_3)$ based on the assumption rc-$A_{(4)}$ is derived as $M_{X_5}(x_2, x_3) = \tilde{M}_{X_5}(x_2, x_3) + M_{X_5}^{c_1^1}(x_2, x_3) = \frac{1}{5} + \frac{1}{15} = \frac{4}{15}$.

Therefore, the original M-function values for the first future observation X_5, based on the assumption rc-$A_{(n)}$ [32], according to the Definition 2.4.4, are derived as follows (see Figure 5.1)

$$M_{X_5}(0, x_1) = \frac{1}{5} = \frac{3}{15}$$
$$M_{X_5}(x_1, x_2) = \frac{1}{5} = \frac{3}{15}$$
$$M_{X_5}(c_1^1, x_2) = \frac{1}{5} \times \frac{1}{3} = \frac{1}{15}$$
$$M_{X_5}(x_2, x_3) = \frac{1}{5} + \frac{1}{15} = \frac{4}{15}$$
$$M_{X_5}(x_3, \infty) = \frac{1}{5} + \frac{1}{15} = \frac{4}{15}$$

With the new technique presented in Section 5.2 on the basis of Definition 5.2.1, we have the opportunity to specify the original M-function values for X_5, shown in Figure 5.1, to probability mass values assigned to their sub-intervals (as shown in Figure 5.2).

From Figure 5.2, as the data set presented in this example does not include any censored observations in the intervals $I^0 = (0, x_1), I^2 = (x_2, x_3)$ and $I^3 = (x_3, \infty)$, we have $\alpha_1^0 = \alpha_1^2 = \alpha_1^3 = 1$. The interval $I^1 = (x_1, x_2)$ contains a single censored observation c_1^1, so we split this interval into two sub-intervals; $I_0^1 = (x_1, c_1^1)$ and $I_1^1 = (c_1^1, x_2)$ and we introduce α_1^1 and α_2^1 for these intervals, respectively, such that the sum of them is one. Using these α_1^1 and α_2^1 values, we can determine the distribution of a probability per interval over its sub-intervals in order to minimise or maximise the probability for the event $X_5 > t$.

As for $c_1^1 \in (x_1, x_2)$, it is necessary to determine where to put the probability mass for X_5, that is, $M_{X_5}(x_1, x_2) = \frac{1}{5}$, in this interval. Since there is only one right-censored observation in (x_1, x_2), the probability mass $M_{X_5}(x_1, x_2) = \frac{1}{5}$, given by Equation (2.27), is now assigned into two sub-intervals, with regard to α_1^1 and α_2^1 introduced respectively to the two sub-intervals, as

$$M_{X_5}(x_1, c_1^1) = \alpha_1^1 M_{X_5}(x_1, x_2) = \frac{1}{5}\alpha_1^1 \tag{5.10}$$
$$M_{X_5}(c_1^1, x_2) = \alpha_2^1 M_{X_5}(x_1, x_2) = \frac{1}{5}\alpha_2^1 \tag{5.11}$$

Taking into consideration the probability mass $M_{X_5}^{c_1^1}(c_1^1, x_2) = \frac{1}{15}$, given by Definition 2.4.3, we consider the following probability mass, using Definition 5.2.1, to be assigned to the sub-interval (c_1^1, x_2) for $c_1^1 \in (x_1, x_2)$

$$M_{X_5}(c_1^1, x_2) = \alpha_1^{c_1^1} M_{X_5}^{c_1^1}(c_1^1, x_2) = \frac{1}{15}\alpha_1^{c_1^1}, \qquad \text{where } \alpha_1^{c_1^1} = 1 \tag{5.12}$$

Therefore, the original M-function values for X_5 [32], given by the Definition 2.4.4 and shown in Figure 5.1, are now re-distributed based on the Definition 5.2.1, as follow (see Figure 5.2),

$$\begin{aligned}
M_{X_5}(0, x_1) =& \frac{1}{5} \\
M_{X_5}(x_1, c_1^1) =& \frac{1}{5}\alpha_1^1 \\
M_{X_5}(c_1^1, x_2) =& \frac{1}{5}\alpha_2^1 + \frac{1}{15}\alpha_1^{c_1^1} \\
M_{X_5}(x_2, x_3) =& \frac{1}{5} + \frac{1}{15} \\
M_{X_5}(x_3, \infty) =& \frac{1}{5} + \frac{1}{15}
\end{aligned}$$

Then, for the interval (x_1, x_2) which contains the only right-censored observation c_1^1, we consider $Q_{X_5}(c_1^1, x_2)$ as representing a probability that can either be maximised or minimised depending on how much the probability mass value is distributed over the sub-intervals of the interval (x_1, x_2). Using Equation (5.3), the function $Q_{X_5}(c_1^1, x_2)$ is defined by combining Equations (5.11) and (5.12), as

$$\begin{aligned}
Q_{X_5}(c_1^1, x_2) &= \alpha_2^1 M_{X_5}(x_1, x_2) + \alpha_1^{c_1^1} M_{X_5}^{c_1^1}(c_1^1, x_2) \\
&= \frac{1}{5}\alpha_2^1 + \frac{1}{15}\alpha_1^{c_1^1}
\end{aligned}$$

but $\alpha_1^{c_1^1} = 1$ since there is only one right-censored observation in the interval (x_1, x_2), so $Q_{X_5}(c_1^1, x_2) = \frac{1}{5}\alpha_2^1 + \frac{1}{15}$.

The function $Q_{X_5}(c_1^1, x_2)$ can be minimised and maximised in order to obtain the NPI lower and upper probabilities for the event $X_5 \in (c_1^1, x_2)$, using Equations (5.6) and (5.7). The minimum value of the function $Q_{X_5}(c_1^1, x_2)$ is obtained by assigning all probability masses within the interval (x_1, x_2) to the left of c_1^1, that is $\alpha_2^1 = 0$, so $\alpha_1^1 = 1$ and $Q_{X_5}^{\min}(c_1^1, x_2) = \frac{1}{15}$. The maximum value of the function $Q_{X_5}(c_1^1, x_2)$ is obtained by assigning all probability masses within the interval (x_1, x_2) to the right of c_1^1, that is $\alpha_2^1 = 1$, so $\alpha_1^1 = 0$ and $Q_{X_5}^{\max}(c_1^1, x_2) = \frac{1}{5} + \frac{1}{15} = \frac{4}{15}$.

The NPI lower and upper probabilities for the event $X_5 > t$, based on the Definition 5.1, are derived using Equations (5.8) and (5.9) respectively. The lower

$t \in (.)$	$\underline{P}(X_5 > t)$	$\overline{P}(X_5 > t)$
$(0, x_1)$	$\frac{4}{5}$	1
(x_1, c_1^1)	$\frac{3}{5}$	$\frac{4}{5}$
(c_1^1, x_2)	$\frac{8}{15}$	$\frac{4}{5}$
(x_2, x_3)	$\frac{4}{15}$	$\frac{8}{15}$
(x_3, ∞)	0	$\frac{4}{15}$

Table 5.1: $\underline{P}(X_5 > t)$ and $\overline{P}(X_5 > t)$ according to Example 5.2.1

probability $\underline{P}(X_5 > t)$ is obtained by considering only the probability mass that necessarily lies in (t, ∞). The corresponding upper probability $\overline{P}(X_5 > t)$ is obtained by considering the probability mass that could possibly lie within (t, ∞).

Taking the case $t \in (x_1, c_1^1)$ as an example, the lower probability for the event $X_5 > t$ is obtained by considering only probability masses that necessarily lie within (t, ∞), using Equation (5.8), i.e., $\underline{P}_{X_5}(x_1, c_1^1) = Q_{X_5}^{\min}(c_1^1, x_2) + P_{X_5}(x_2, x_3) + P_{X_5}(x_3, \infty) = 1/15 + 4/15 + 4/15 = 3/5$. For the case $t \in (c_1^1, x_2)$, the upper probability for the event $X_5 > t$ is obtained by summing up all probability masses that can be in (t, ∞), using Equation (5.9), i.e., $\overline{P}_{X_5}(c_1^1, x_2) = Q_{X_5}^{\max}(c_1^1, x_2) + P_{X_5}(x_2, x_3) + P_{X_5}(x_3, \infty) = 4/15 + 4/15 + 4/15 = 4/5$. For t in an interval which does not contain right-censored observations, the NPI lower and upper probabilities for the event $X_5 > t$ can be derived directly from the Equations (2.34) and (2.35), respectively. Consequently, the NPI lower and upper probabilities for the event $X_5 > t$, based on the data in this example, are given in Table 5.1.

Note that we can straightforwardly apply rc-$A_{(4)}$ for X_5, using Definition 2.4.4, where there are no assumptions on where the probability mass is within each interval. But, as we aim to apply rc-$A_{(5)}$ for X_6, based on rc-$A_{(4)}$ for X_5, later on, we had to consider where the probability mass is for X_5 in this example using the new techniques presented in this section. Also, this example will be used again later on in this chapter.

We next consider the second future observation, X_{n+2}, based on the first future observation, X_{n+1}, as well as the data set that includes n observations with right-censored observations. Section 5.3 derives the NPI lower and upper conditional probabilities for $X_{n+2} > t$ given $X_{n+1} > t$, which will enable us to derive the NPI lower and upper probabilities for the event that both future observations are greater than t, in Section 5.4.

5.3 Lower and upper probabilities for $X_{n+2} > t$ given $X_{n+1} > t$

This section presents the NPI conditional lower and upper probabilities for the event $X_{n+2} > t$ given $X_{n+1} > t$. We aim to apply the rc-$A_{(n+1)}$ assumption for X_{n+2}, conditionally on X_{n+1}, for which we apply the rc-$A_{(n)}$ assumption, which was presented in Section 5.2, without further assumptions.

Based on Definition 5.2.1, there are $n+1$ cases of which X_{n+1} falls in the intervals created by the data set that contains n observations including right-censored observations, denoted as $I^i_{i*} = (t^i_{i^*}, t^i_{i^*+1})$, where $i = 0, 1, \dots, u$, $i^* = 1, 2, \dots, s_i$, as provided in Section 5.2. For $X_{n+1} \in (t^i_{i^*}, t^i_{i^*+1})$, when considering X_{n+2}, there will be $n+1$ observations of which we have $u+1$ event times, $x_1 < x_2 < \cdots < x_u < x_{u+1}$, and $\nu = (n+1) - (u+1) = n - u$ right-censored observations, $c_1 < c_2 < \ldots < c_\nu$. Note that $u+1$ refer to the failure observations in the data set including X_{n+1}. So, there are $n+2$ intervals created by the data set that contains $n+1$ observations, included X_{n+1}, and the right-censored observations, denoted by $I^j_{j*} = (t^j_{j^*}, t^j_{j^*+1})$, where $j = 0, 1, \dots, u+1$, $j^* = 1, 2, \dots, s_j$. Let $x_0 = 0$ and $x_{u+2} = \infty$ for ease of notation. We assume, in order to simplify our presentation, that no ties exist in the data set, so no two observations (events or right-censoring) are at the same time value. In case there are ties, we refer to the discussion in Section 2.2.

In order to derive the NPI conditional lower and upper probabilities for the event $X_{n+2} > t$ given $X_{n+1} > t$, we are going to present the rc-$A_{(n+1)}$ assumption for X_{n+2} given $X_{n+1} \in I^i_{i*} = (t^i_{i^*}, t^i_{i^*+1})$, following what is presented in Section 5.2 regarding the rc-$A_{(n)}$ assumption for X_{n+1}, as we have considered where the probability mass is for X_{n+1} within an interval (x_i, x_{i+1}). Here, in case of the event $X_{n+2} > t$ given $X_{n+1} > t$, we use the same notation that used for the event $X_{n+1} > t$ in Section 5.2, with replacing the notation $\underline{\alpha}^i$ and $\underline{\alpha}^{c^i_{i*}}$ by $\underline{\beta}^i$ and $\underline{\beta}^{c^i_{i*}}$.

Given that $X_{n+1} \in I^i_{i*} = (t^i_{i^*}, t^i_{i^*+1})$ and for an interval $I^j = (x_j, x_{j+1})$, $j = 0, 1, 2, \ldots, u+1$, there are s_j right-censored observations in this interval, and

$$\underline{\beta}^j = (\beta^j_1, \beta^j_2, \ldots, \beta^j_{s_j+1}), \text{where } 0 \leq \beta^j_{j^*} \leq 1 \text{ and } \sum_{j^*=1}^{s_j+1} \beta^j_{j^*} = 1$$

If there are no censored observations in the interval (x_j, x_{j+1}), that is $s_j = 0$, then $\underline{\beta}^j = \beta^j_1 = 1$. Also, for each censored observation $c^j_{j^*}$, $j^* = 1, 2, \ldots, s_j$, in the interval (x_j, x_{j+1}),

$$\underline{\beta}^{c^j_{j^*}} = (\beta_1^{c^j_{j^*}}, \beta_2^{c^j_{j^*}}, \ldots, \beta_{s_j-j^*+1}^{c^j_{j^*}}), \text{where } 0 \leq \beta_l^{c^j_{j^*}} \leq 1 \text{ and } \sum_{l=1}^{s_j-j^*+1} \beta_l^{c^j_{j^*}} = 1.$$

if there is only one censored observation in the interval (x_j, x_{j+1}) then $\underline{\beta}^{c^j_{j^*}} = \beta_1^{c^j_{j^*}} = 1$.

The notation $\underline{\beta}^j$ and $\underline{\beta}^{c^j_{j^*}}$ are the proportion of (a specific) conditional probability mass assigned to the interval (x_j, x_{j+1}) that is distributed over sub-intervals, given $X_{n+1} \in (x_j, x_{j+1})$. It is just a way to write how the probability mass is divided over sub-intervals, so that we can then find the NPI conditional lower and upper probabilities for any event of interest involving X_{n+2} given X_{n+1}.

Given that $X_{n+1} \in I^i_{i*} = (t^i_{i^*}, t^i_{i^*+1})$, the rc-$A_{(n+1)}$ assumption partially specifies the probability distribution for the second future observation X_{n+2} by the conditional M-functions denoted as $M_{X_{n+2}|X_{n+1}}$. We present the conditional M-functions for X_{n+2} to be in the interval $I^j_{j*} = (t^j_{j^*}, t^j_{j^*+1})$, $j = 0, 1, \ldots, u+1$, $j^* = 1, 2, \ldots, s_j$, given that X_{n+1} is in the interval $I^i_{i*} = (t^i_{i^*}, t^i_{i^*+1})$, by the following definition.

Definition 5.3.1 (conditional M-functions)
The conditional M-function partially specifies the probability distribution for the second future observation X_{n+2} given $X_{n+1} \in I^i_{i^*}$, i.e., $x_{n+1} \in (t^i_{i^*}, t^i_{i^*+1})$, for $i =$

$0, 1, \ldots, u$, $i^* = 1, 2, \ldots, s_i$, as follows

$$\begin{aligned} M_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(t^j_{j^*}, t^j_{j^*+1}) &= \beta^j_{j^*+1} M_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(x_j, x_{j+1}) \\ &+ \sum_{k=1}^{j^*} \beta^{c^j_k}_{j^*-k+1} M_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(c^j_k, x_{j+1}) \end{aligned} \quad (5.13)$$

where

$$\begin{cases} j^* = 0 & \text{if} \quad t^j_0 = x_j \quad \text{(failure time or time 0)} \\ j^* = 1, 2, \ldots s_j & \text{if} \quad t^j_{j^*} = c^j_{j^*} \quad \text{(right-censoring time)} \end{cases}$$

for $j = 0, 1, \ldots, u+1$ and $j^* = 1, 2, \ldots, s_j$, and for simplicity of notation let $t^j_{s_j+1} = t^{j+1}_0 = x_{j+1}$. For convenience, $M_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(t^j_{j^*}, t^j_{j^*+1})$ can be denoted by $M_{X_{n+2}\in I^j_{j^*}|X_{n+1}\in I^i_{i^*}}$, in order to be used later in Section 5.4.

Using the same logic as in Definition 2.4.4 "rc-$A_{(n)}$ assumption" for X_{n+1}, the $M_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(x_j, x_{j+1})$ and $M_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(c^j_{j^*}, x_{j+1})$ values, given in Equation (5.13), are derived by applying $A_{(n+1)}$ for X_{n+2}, given $X_{n+1} \in I^i_{i^*}$ based on the assumption rc-$A_{(n+1)}$, as

$$M_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(x_j, x_{j+1}) = \frac{1}{n+2} \prod_{\{r: c_r < x_j\}} \frac{\tilde{n}_{c_r}+1}{\tilde{n}_{c_r}} \quad (5.14)$$

$$M_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(c^j_{j^*}, x_{j+1}) = \frac{1}{(n+2)\tilde{n}_{c^j_{j^*}}} \prod_{\{r: c_r < c^j_{j^*}\}} \frac{\tilde{n}_{c_r}+1}{\tilde{n}_{c_r}} \quad (5.15)$$

where $\tilde{n}_{c_r}$ represents the number of observations in the risk set (still functioning or alive and uncensored) just before time c_r. The product terms in Equations (5.14) and (5.15) are assumed to be equal to one if the product is taken over an empty set [32].

Using M-functions for X_{n+1}, based on Definition 5.2.1, leads to $P_{X_{n+1}}(x_i, x_{i+1})$ stated in Equation (5.2). In the same way, when using conditional M-functions for $X_{n+2}|X_{n+1}$, based on Definition 5.3.1, this will lead to $P_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(x_j, x_{j+1})$ in Equation (5.16). With $\beta^j_{j^*} \in [0,1]$, $\beta^{c^j_{j^*}}_l \in [0,1]$, $\sum_{j^*=1}^{s_j+1} \beta^j_{j^*} = 1$ and $\sum_{l=1}^{s_j-j^*+1} \beta^{c^j_{j^*}}_l = 1$, the conditional M-function values as specified by rc-$A_{(n+1)}$ in Definition 5.13 lead to the conditional probability for the event that $X_{n+2} \in I^j_{j^*}$, where $j = 0, 1, \ldots, u+1$, given $X_{n+1} \in I^i_{i^*}$, where $i = 0, 1, \ldots, u$, denoted by $P_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(x_j, x_{j+1})$. The

$P_{X_{n+2}|X_{n+1}\in I^i_{i*}}(x_j, x_{j+1})$ is calculated by summing up all conditional M-function values assigned to the interval $I^j = (x_j, x_{j+1})$ given $X_{n+1} \in I^i_{i^*}$, along with all conditional M-function values assigned to the sub-intervals $(c^j_{j^*}, x_{j+1})$ for X_{n+1}, given $X_{n+1} \in I^i_{i^*}$ so that

$$
\begin{aligned}
P_{X_{n+2}|X_{n+1}\in I^i_{i*}}(x_j, x_{j+1}) &= \sum_{j^*=0}^{s_j} M_{X_{n+2}|X_{n+1}\in I^i_{i*}}(t^j_{j^*}, t^j_{j^*+1}) \\
&= \sum_{j^*=0}^{s_j} \beta^j_{j^*+1} M_{X_{n+2}|X_{n+1}\in I^i_{i*}}(x_j, x_{j+1}) + \sum_{j^*=1}^{s_j}\sum_{l=1}^{j^*} \beta^{c^j_l}_{j^*-l+1} M_{X_{n+2}|X_{n+1}\in I^i_{i*}}(c^j_l, x_{j+1}) \\
&= M_{X_{n+2}|X_{n+1}\in I^i_{i*}}(x_j, x_{j+1}) + \sum_{l=1}^{s_j}\sum_{j^*=1}^{s_j-l+1} \beta^{c^j_l}_{j^*} M_{X_{n+2}|X_{n+1}\in I^i_{i*}}(c^j_l, x_{j+1}) \\
&= M_{X_{n+2}|X_{n+1}\in I^i_{i*}}(x_j, x_{j+1}) + \sum_{l=1}^{s_j} M_{X_{n+2}|X_{n+1}\in I^i_{i*}}(c^j_l, x_{j+1})
\end{aligned}
\tag{5.16}
$$

for $i = 0, 1, \ldots, u$ and $j = 0, 1, \ldots, u+1$. For convenience, $P_{X_{n+2}|X_{n+1}\in I^i}(x_j, x_{j+1})$, stated in Equation (5.16), would be denoted by $P_{X_{n+2}\in I^i|X_{n+1}\in I^i}$ in order to be used later in Section 5.4.

The first term after the second equality in Equation (5.16) is the sum of all conditional M-function values assigned to the interval (x_j, x_{j+1}), given $X_{n+1} \in I^i_{i^*}$, and as $\sum_{j^*=1}^{s_j+1} \beta^j_{j^*} = 1$, this first term is equal to $M_{X_{n+2}|X_{n+1}\in I^i_{i*}}(x_j, x_{j+1})$. The second term after the third equality in Equation (5.16) is the sum of all conditional M-function values assigned to the sub-intervals (c^j_l, x_{j+1}) of (x_j, x_{j+1}), given $X_{n+1} \in I^i_{i^*}$, and as $\sum_{l=1}^{s_j-j^*+1} \beta^{c^j_{j^*}}_l = 1$, for $j = 0, 1, \ldots, u+1$ and $j^* = 1, 2, \ldots, s_j$, this second term is equal to $\sum_{l=1}^{s_j} M_{X_{n+2}|X_{n+1}\in I^i_{i*}}(c^j_l, x_{j+1})$. And let us define the following

$$
\begin{aligned}
Q_{X_{n+2}|X_{n+1}\in I^i_{i*}}(t^j_a, x_{j+1}) &= \sum_{j^*=a}^{s_j} \beta^j_{j^*+1} M_{X_{n+2}|X_{n+1}\in I^i_{i*}}(x_j, x_{j+1}) \\
&\quad + \sum_{j^*=a}^{s_j}\sum_{l=1}^{j^*} \beta^{c^j_l}_{j^*-l+1} M_{X_{n+2}|X_{n+1}\in I^i_{i*}}(c^j_l, x_{j+1})
\end{aligned}
\tag{5.17}
$$

where for $a = 0$, Equation (5.16) and (5.17) are equivalent.

The $Q_{X_{n+2}|X_{n+1}\in I^i_{i*}}(t^j_a, x_{j+1})$, given by Equation (5.17), can be minimised or maximised in order to derive the NPI conditional lower and upper probabilities for the event $X_{n+2} > t$ given $X_{n+1} > t$. We sometimes denote the conditional probability in Equation (5.17) by $Q_{X_{n+2}\in I^j_a|X_{n+1}\in I^i_{i*}}$ for convenience. Now, let us consider the second term of Equation (5.17), and by rearranging the summations, we have

$$\sum_{j^*=a}^{s_j}\sum_{l=1}^{j^*}\beta_{j^*-l+1}^{c_l^j} M_{X_{n+2}|X_{n+1}\in I_{i^*}^i}(c_l^j, x_{j+1}) = \sum_{l=1}^{a-1}\sum_{j^*=a}^{s_j-l+1}\beta_{j^*}^{c_l^j} M_{X_{n+2}|X_{n+1}\in I_{i^*}^i}(c_l^j, x_{j+1}) + \sum_{l=a}^{s_j}\sum_{j^*=1}^{s_j-j+1}\beta_{j^*}^{c_l^j} M_{X_{n+2}|X_{n+1}\in I_{i^*}^i}(c_l^j, x_{j+1}) \tag{5.18}$$

The first term on the right-hand side of Equation (5.18) is related to the conditional probability masses to the right of t_a^j, corresponding to all $c_l^j < t_a^j$. The second term in Equation (5.18) is related to the conditional probability masses corresponding to all $c_l^j \geq t_a^j$, and as $\sum_{j^*=1}^{s_j-j+1}\beta_{j^*}^{c_l^j} = 1$, this second term is equal to $\sum_{l=a}^{s_j} M_{X_{n+2}|X_{n+1}\in I*i_{i*}}(c_l^j, x_{j+1})$. So Equation (5.17) can be rewritten as

$$Q_{X_{n+2}|X_{n+1}\in I_{i^*}^i}(t_a^j, x_{j+1}) = \sum_{j^*=a}^{s_j}\beta_{j^*+1}^{j} M_{X_{n+2}|X_{n+1}\in I_{i^*}^i}(x_j, x_{j+1}) + \sum_{l=1}^{a-1}\sum_{j^*=a}^{s_j-l+1}\beta_{j^*}^{c_l^j} M_{X_{n+2}|X_{n+1}\in I_{i^*}^i}(c_l^j, x_{j+1}) + \sum_{l=a}^{s_j} M_{X_{n+2}|X_{n+1}\in I_{i^*}^i}(c_l^j, x_{j+1}) \tag{5.19}$$

If we want to find the values of $\underline{\beta}^j$'s and $\underline{\beta}^{c_l^j}$'s that minimize $Q_{X_{n+2}|X_{n+1}\in I_{i^*}^i}(t_a^j, x_{j+1})$, as given in Equation (5.19), then this can be achieved by assigning all conditional probability masses in (x_j, x_{j+1}) to the left of the t_a^j, that is

$$\sum_{j^*=a}^{s_j}\beta_{j^*+1}^{j} = 0\ , \qquad \sum_{j^*=0}^{a-1}\beta_{j^*+1}^{j} = 1$$

and

$$\sum_{j^*=1}^{a-1}\beta_{j^*}^{c_l^j} = 1\ , \qquad \sum_{j^*=a}^{s_j-l+1}\beta_{j^*}^{c_l^j} = 0$$

thus, the minimum value of $Q_{X_{n+2}|X_{n+1}\in I_{i^*}^i}(t_a^j, x_{j+1})$ in Equation (5.19) is

$$Q^{\min}_{X_{n+2}|X_{n+1}\in I_{i^*}^i}(t_a^j, x_{j+1}) = \sum_{l=a}^{s_j} M_{X_{n+2}|X_{n+1}\in I^i}(c_l^j, x_{j+1}) \tag{5.20}$$

If we want to find the values of $\underline{\beta}^j$'s and $\underline{\beta}^{c_l^j}$'s that maximise $Q_{X_{n+2}|X_{n+1}\in I_{i^*}^i}(t_a^j, x_{j+1})$, stated in Equation (5.19), then this can be achieved by assigning all conditional probability masses in the interval (x_j, x_{j+1}) to the right of t^j_a, that is

$$\sum_{j^*=a}^{s_j} \beta^j_{j^*+1} = 1 \ , \qquad \sum_{j^*=0}^{a-1} \beta^j_{j^*+1} = 0$$

and

$$\sum_{j^*=1}^{a-1} \beta^{c^j_l}_{j^*} = 0 \ , \qquad \sum_{j^*=a}^{s_j-l+1} \beta^{c^j_l}_{j^*} = 1$$

thus, the maximum value of $Q_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(t^j_a, x_{j+1})$ in Equation (5.19) is

$$\begin{aligned} Q^{\max}_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(t^j_a, x_{j+1}) &= M_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(x_j, x_{j+1}) + \sum_{l=1}^{a-1} M_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(c^j_l, x_{j+1}) \\ &\quad + \sum_{l=a}^{s_j} M_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(c^j_l, x_{j+1}) \\ &= M_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(x_j, x_{j+1}) + \sum_{l=1}^{s_i} M_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(c^j_l, x_{j+1}) \\ &= P_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(x_j, x_{j+1}) \end{aligned} \tag{5.21}$$

The probabilities, given by Equations (5.20) and (5.21), could be also denoted by $Q^{\min}_{X_{n+2}\in I^j_a|X_{n+1}\in I^i_{i^*}}$ and $Q^{\max}_{X_{n+2}\in I^j_a|X_{n+1}\in I^i_{i^*}}$, respectively, for convenience in order to be used later in Section 5.4.

Consequently, the NPI lower probability for the event $X_{n+2} > t$ given $X_{n+1} > t$, for $t \in [t^j_a, t^j_{a+1})$ with $j = 0, 1, \ldots, u+1$ and $a = 0, 1, \ldots, s_j$, is given by the following equation

$$\begin{aligned} \underline{P}(X_{n+2} > t|X_{n+1} > t) &= Q^{\min}_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(t^j_{a+1}, x_{j+1}) + \sum_{z=j+1}^{u+1} P_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(x_z, x_{z+1}) \\ &= \sum_{l=a+1}^{s_i} M_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(c^j_l, x_{j+1}) + \sum_{z=j+1}^{u+1} P_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(x_z, x_{z+1}) \end{aligned} \tag{5.22}$$

The corresponding NPI upper probability for the event $X_{n+2} > t$ given $X_{n+1} > t$, for $t \in [x_j, x_{j+1})$ with $j = 1, 2, \ldots, u+1$ and $a = 0, 1, \ldots, s_j$, is given by the following equation

$$\begin{aligned} \overline{P}(X_{n+2} > t|X_{n+1} > t) &= Q^{\max}_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(t^j_a, x_{j+1}) + \sum_{z=j+1}^{u+1} P_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(x_z, x_{z+1}) \\ &= \sum_{z=j}^{u+1} P_{X_{n+2}|X_{n+1}\in I^i_{i^*}}(x_z, x_{z+1}) \end{aligned} \tag{5.23}$$

Note that using the α approach through minimising and maximising Equation (5.5) is only considered to derive NPI lower and upper probabilities for the event $X_{n+1} > t$. And, using the β approach through minimising and maximising Equation (5.19) will only be considered to derive NPI conditional lower and upper probabilities for the event $X_{n+2} > t$ given $X_{n+1} > t$. However, both the α and β approaches through minimising and maximising Equations (5.5) and (5.19), respectively, must be used together in order to derive NPI lower and upper probabilities for the event $X_{n+1} > t$ and $X_{n+2} > t$, which will be presented in Section 5.4.

Next, the method presented in this section is illustrated with an example using the data set provided in Example 5.2.1.

Example 5.3.1 This example is provided to illustrate the rc-$A_{(n+1)}$ assumption for X_{n+2}, based on the rc-$A_{(n)}$ assumption for X_{n+1} [32], presented in Section 5.3. In particular, it shows how to derive the NPI conditional lower and upper probabilities for the event $X_{n+2} > t$ given $X_{n+1} > t$.

We consider the data set used in Example 5.2.1, for which we have $n = 4$ observations, including one right-censored observation, as shown in Figure 5.1. Let X_5 and X_6 denote the two future observations. In Example 5.2.1, the probability distribution for X_5 was partially specified by five M-function values associated with five intervals generated by the 4 observations, using Definition 5.2.1 (see Figure 5.2). Note that in Example 5.2.1, we applied rc-$A_{(4)}$ for X_5 with considering where the probability mass is for X_5, and based on this we apply rc-$A_{(5)}$ for X_6 in this example.

Given that X_5 falls in those five intervals created by the $n = 4$ data observations, i.e., $I^0 = (0, x_1), I^1_1 = (x_1, c^1_1), I^1_2 = (c^1_1, x_2), I^2 = (x_2, x_3)$ and $I^3 = (x_3, \infty)$, respectively, so there are five cases of which X_5 falls into these intervals. Then, we consider X_6 depending on X_5 being in a specific interval. This enables the probability distribution for X_6 to be partially specified by conditional M-function values assigned to six intervals formed by the 5 observations including X_5, using Definition 5.3.1 separately for each case.

As a result of applying Definition 5.3.1 with the assumption rc-$A_{(5)}$ given by Equations (5.14) and (5.15), these conditional M-function values for X_6 given $X_5 \in$

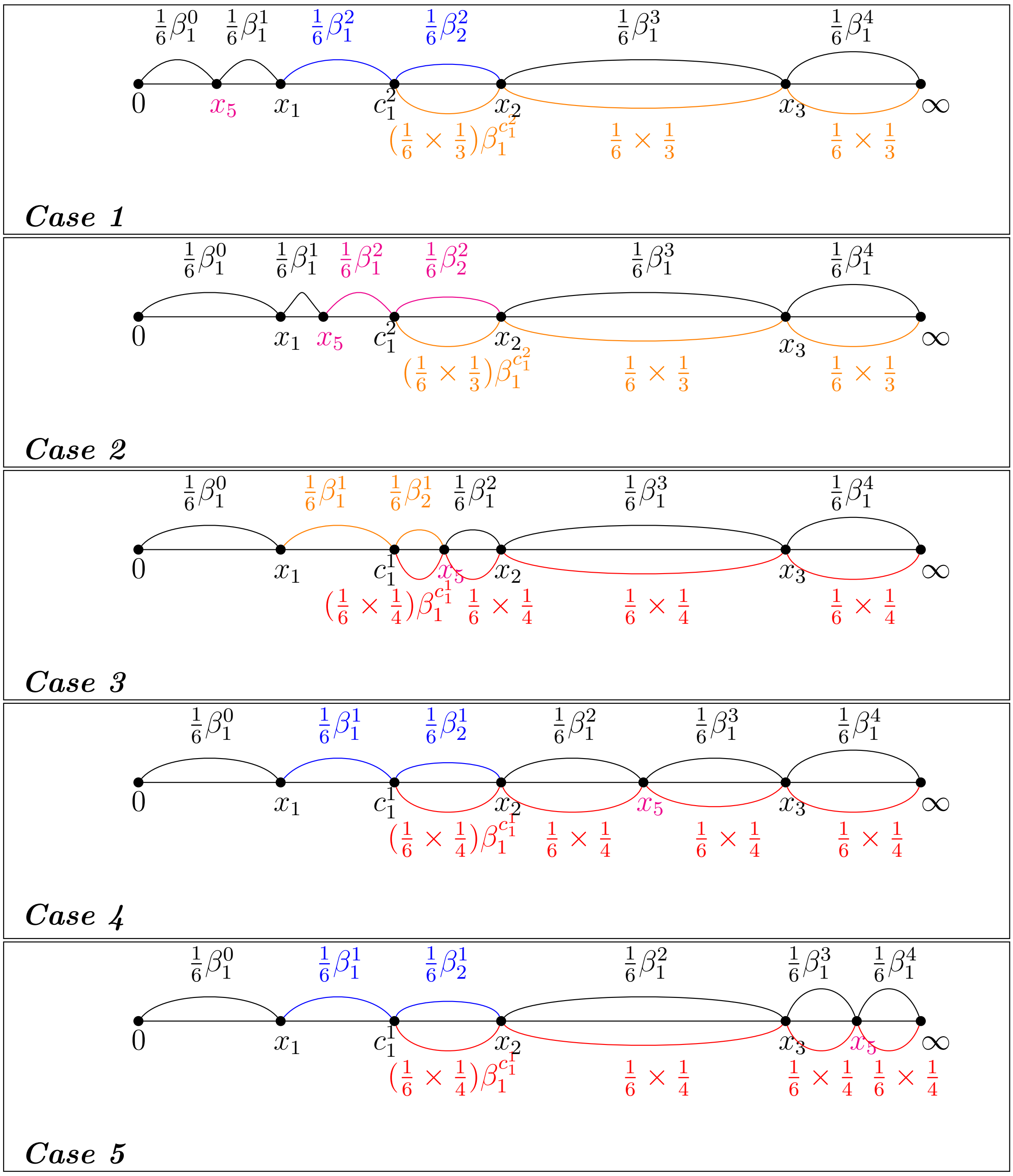

Figure 5.3: The conditional probabilities for $X_6|X_5$, Example 5.4.1

$\{I^0, I_1^1, I_2^1, I^2, I^3\}$, can be obtained as follows (see Figure 5.3).

Case 1: Given $X_5 \in I^0 = (0, x_1)$, the conditional M-function values for $X_6|X_5 \in I^0$, using Definition 5.3.1 with the assumption rc-$A_{(5)}$ given by Equations (5.14) and (5.15), are shown in the first box of Figure 5.3.

Since there is no censored observation in intervals $(0, x_5), (x_5, x_1)$, (x_2, x_3) and (x_3, ∞), respectively, the corresponding values β_1^0, β_1^1, β_1^3 and β_1^4 introduced to these intervals are equal to one, i.e. $\beta_1^0 = \beta_1^1 = \beta_1^3 = \beta_1^4 = 1$, as discussed in Section 5.3. For $c_1^2 \in (x_1, x_2)$ given $X_5 \in (0, x_1)$, the conditional M-function value $\frac{1}{6}$ that is assigned to interval (x_1, x_2) will be split up, based on Definition 5.3.1, and assigned to two sub-intervals with the M-function value $\frac{1}{6}\beta_1^2$ assigned to the sub-interval (x_1, c_1^2) and the M-function value $\frac{1}{6}\beta_2^2$ assigned to the sub-interval (c_1^2, x_2), where both β_1^2 and β_2^2 take values between 0 and 1, and $\beta_1^2 + \beta_2^2 = 1$. Also, based on Definition 5.3.1 and Equation (5.15), the M-function value $\frac{1}{18}\beta_1^{c_1^2}$ is assigned to the sub-interval (c_1^2, x_2), where $\beta_1^{c_1^2} = 1$. Thus, the conditional M-function values for the event $X_6|X_5 \in I^0$ are

$$M_{X_6|X_5 \in I^0}(0, x_5) = \frac{1}{6} \tag{5.24}$$

$$M_{X_6|X_5 \in I^0}(x_5, x_1) = \frac{1}{6} \tag{5.25}$$

$$M_{X_6|X_5 \in I^0}(x_1, c_1^2) = \frac{1}{6}\beta_1^2$$

$$M_{X_6|X_5 \in I^0}(c_1^2, x_2) = \frac{1}{6}\beta_2^2 + \frac{1}{18}\beta_1^{c_1^2}$$

$$M_{X_6|X_5 \in I^0}(x_2, x_3) = \frac{1}{6} + \frac{1}{18}$$

$$M_{X_6|X_5 \in I^0}(x_3, \infty) = \frac{1}{6} + \frac{1}{18}$$

where the total conditional probability mass for $X_6 \in (0, x_1)$ given $X_5 \in (0, x_1)$, given in Equations (5.24) and (5.25), is 1/6+1/6=2/6, see *Case 1* in the first box of Figure 5.3.

Based on these conditional M-function values, we can derive the conditional probability for the event $X_6 \in (x_1, x_2)$ given $X_5 \in (0, x_1)$, by summing the probability masses assigned to the sub-intervals (x_1, c_1^2) and (c_1^2, x_2), so $P_{X_6|X_5 \in I^0}(x_1, x_2) = \frac{1}{6}\beta_1^2 + \frac{1}{6}\beta_2^2 + \frac{1}{18}\beta_1^{c_1^2} = \frac{1}{6}(\beta_1^2 + \beta_2^2) + \frac{1}{18}\beta_1^{c_1^2}$, and as discussed in Section 5.3, $\beta_1^2 + \beta_2^2 = 1$ and $\beta_1^{c_1^2} = 1$, so $P_{X_6|X_5 \in I^0}(x_1, x_2) = \frac{1}{6} + \frac{1}{18} = \frac{4}{18}$. Moreover, the conditional probabilities for X_6 to be in intervals (x_2, x_3) or (x_3, ∞), given $X_5 \in (0, x_1)$, are $P_{X_6|X_5 \in I^0}(x_2, x_3) = P_{X_6|X_5 \in I^0}(x_3, \infty) = \frac{1}{6} + \frac{1}{18} = \frac{4}{18}$.

Then from *Case 1* in which $X_5 \in (0, x_1)$, we now consider the event $X_6 > t$ given $X_5 > t$, where $t \in (0, x_1)$ (see the first box of Figure 5.3). By assigning all conditional probability masses that must be within (t, ∞), Equation (5.22) is used to determine the NPI lower conditional probability for the event $X_6 > t$ given $X_5 > t$, where $t \in (0, x_1)$. Thus

$$\begin{aligned}
\underline{P}(X_6 > t | X_5 > t) &= Q^{\min}_{X_6|X_5 \in I^0}(0, x_1) + \sum_{z=2}^{4} P_{X_6|X_5 \in I^0}(x_z, x_{z+1}) \\
&= M_{X_6|X_5 \in I^0}(x_5, x_1) + P_{X_6|X_5 \in I^0}(x_1, x_2) + P_{X_6|X_5 \in I^0}(x_2, x_3) \\
&+ P_{X_6|X_5 \in I^0}(x_3, \infty) \\
&= \frac{1}{6} + \frac{4}{18} + \frac{4}{18} + \frac{4}{18} = \frac{5}{6}
\end{aligned}$$

where the value of $Q^{\min}_{X_6|X_5 \in I^0}(0, x_1)$ is obtained by using Equation (5.20).

The NPI upper conditional probability for the event $X_6 > t$ given $X_5 > t$, where $t \in (0, x_1)$, is derived by assigning all conditional probability masses that could be within (t, ∞) using Equation (5.23). Thus

$$\begin{aligned}
\overline{P}(X_6 > t | X_5 > t) &= Q^{\max}_{X_6|X_5 \in I^0}(0, x_1) + \sum_{z=2}^{4} P_{X_6|X_5 \in I^0}(x_j, x_{j+1}) \\
&= P_{X_6|X_5 \in I^0}(0, x_1) + P_{X_6|X_5 \in I^0}(x_1, x_2) + P_{X_6|X_5 \in I^0}(x_2, x_3) \\
&+ P_{X_6|X_5 \in I^0}(x_3, \infty) \\
&= \frac{2}{6} + \frac{4}{18} + \frac{4}{18} + \frac{4}{18} = 1
\end{aligned}$$

where the value of $Q^{\max}_{X_6|X_5 \in I^0}(0, x_1)$ is obtained by using Equation (5.21).

Case 2: Given $X_5 \in I_1^1 = (x_1, c_1^1)$, the conditional M-function values for $X_6 | X_5 \in I_0^1$, using Definition 5.3.1 with the assumption rc-$A_{(5)}$ given by Equations (5.14) and (5.15), are shown in the second box of Figure 5.3.

Due to the fact that no censoring is involved in intervals $(0, x_1), (x_1, x_5)$, (x_2, x_3) and (x_3, ∞), respectively, the values β_1^0, β_1^1, β_1^3 and β_1^4 corresponding to these intervals are equal to 1, as stated in Section 5.3.

By using Equation (5.14), based on the assumption rc-$A_{(5)}$, the conditional M-function value for $X_6 \in (x_1, x_5) | X_5 \in I_1^1$ is $\frac{1}{6}$. For $c_1^2 \in (x_5, x_2)$ given $X_5 \in I_1^1$, the conditional M-function value $\frac{1}{6}$ that is assigned to interval (x_5, x_2) will be split up and assigned to two sub-intervals with the M-function value $\frac{1}{6}\beta_1^2$ assigned to the sub-interval (x_5, c_1^2) as well as the M-function value $\frac{1}{6}\beta_2^2$ assigned to the sub-interval (c_1^2, x_2), where both β_1^2 and β_2^2 take values between 0 and 1, and $\beta_1^2 + \beta_2^2 = 1$. Also, based on Definition 5.3.1 and Equation (5.15), the M-function value $\frac{1}{18}\beta_1^{c_1^2}$ is assigned to the sub-interval (c_1^2, x_2), where $\beta_1^{c_1^2} = 1$. Thus, the conditional M-function values for the event $X_6|X_5 \in I_1^1$ are

$$
\begin{aligned}
M_{X_6|X_5\in I_1^1}(0, x_1) &= \frac{1}{6} \\
M_{X_6|X_5\in I_1^1}(x_1, x_5) &= \frac{1}{6} \qquad (5.26)\\
M_{X_6|X_5\in I_1^1}(x_5, c_1^2) &= \frac{1}{6}\beta_1^2 \qquad (5.27)\\
M_{X_6|X_5\in I_1^1}(c_1^2, x_2) &= \frac{1}{6}\beta_2^2 + \frac{1}{18}\beta_1^{c_1^2} \\
M_{X_6|X_5\in I_1^1}(x_2, x_3) &= \frac{1}{6} + \frac{1}{18} \\
M_{X_6|X_5\in I_1^1}(x_3, \infty) &= \frac{1}{6} + \frac{1}{18}
\end{aligned}
$$

where the total conditional probability mass for $X_6 \in (x_1, c_1^1)$ given $X_5 \in (x_1, c_1^1)$, given in Equations (5.26) and (5.27), is 1/6 $(1+\beta_1^2)$, where $\beta_1^2 \in [0, 1]$, see *Case 2* in the second box of Figure 5.3.

From *Case 2*, where $X_5 \in I_1^1 = (x_1, c_1^1)$, we use Equation (5.22) to derive the NPI conditional lower probability for the event $X_6 > t$ given $X_5 > t$, where $t \in (x_1, c_1^1)$, as

$$
\begin{aligned}
\underline{P}(X_6 > t|X_5 > t) &= Q^{\min}_{X_6|X_5\in I_1^1}(x_1, x_2) + \sum_{z=3}^{4} P_{X_6|X_5\in I_1^1}(x_z, x_{z+1}) \\
&= M_{X_6|X_5\in I_1^1}(x_5, x_2) + P_{X_6|X_5\in I_1^1}(x_2, x_3) + P_{X_6|X_5\in I_1^1}(x_3, \infty) \\
&= \frac{4}{18} + \frac{4}{18} + \frac{4}{18} = \frac{2}{3}
\end{aligned}
$$

where the value of $Q^{\min}_{X_6|X_5\in I_1^1}(x_1, x_2)$ is obtained by using Equation (5.20), i.e., $Q^{\min}_{X_6|X_5\in I_1^1}(x_1, x_2) = \frac{1}{6}\beta_1^2 + \frac{1}{6}\beta_2^2 + \frac{1}{18}\beta_1^{c_1^2} = \frac{1}{6}(\beta_1^2 + \beta_2^2) + \frac{1}{18}\beta_1^{c_1^2}$. And for $\beta_1^2 + \beta_2^2 = 1$ and $\beta_1^{c_1^2} = 1$, $Q^{\min}_{X_6|X_5\in I_1^1}(x_1, x_2) = \frac{1}{6} + \frac{1}{18} = \frac{4}{18}$.

The NPI upper conditional probability for the event $X_6 > t$ given $X_5 > t$, where $t \in (x_1, c_1^1)$, is derived by using Equation (5.23) as follows.

$$
\begin{aligned}
\overline{P}(X_6 > t|X_5 > t) &= Q^{\max}_{X_6|X_5\in I_1^1}(x_1, x_2) + \sum_{z=3}^{4} P_{X_6|X_5\in I_1^1}(x_z, x_{z+1}) \\
&= P_{X_6|X_5\in I_1^1}(x_1, x_2) + P_{X_6|X_5\in I_1^1}(x_2, x_3) + P_{X_6|X_5\in I_1^1}(x_3, \infty) \\
&= \frac{7}{18} + \frac{4}{18} + \frac{4}{18} = \frac{5}{6}
\end{aligned}
$$

where the value of $Q^{\max}_{X_6|X_5\in I_1^1}(x_1, x_2)$ is obtained by using Equation (5.21), i.e., $Q^{\max}_{X_6|X_5\in I_1^1}(x_1, x_2) = P_{X_6|X_5\in I_1^1}(x_1, x_5) + P_{X_6|X_5\in I_1^1}(x_5, x_2) = \frac{1}{6} + \frac{4}{18} = \frac{7}{18}$.

Case 3: Given $X_5 \in I_2^1 = (c_1^1, x_2)$, the conditional M-function values for $X_6|X_5 \in I_2^1$, using Definition 5.3.1 with the assumption rc-$A_{(5)}$ given by Equations (5.14) and (5.15), are shown in the third box of Figure 5.3.

The β_1^0, β_1^2, β_1^3 and β_1^4 values corresponding to the intervals $(0, x_1), (x_5, x_2)$, (x_2, x_3) and (x_3, ∞), respectively, are equal to 1, since there no censoring is involved in these intervals. For $c_1^1 \in (x_1, x_5)$ given $X_5 \in I_2^1$, the conditional M-function value $\frac{1}{6}$ that is assigned to interval (x_1, x_5) will be split up and assigned to two sub-intervals with the M-function value $\frac{1}{6}\beta_1^1$ assigned to the sub-interval (x_1, c_1^1) as well as the M-function value $\frac{1}{6}\beta_2^1$ assigned to the sub-interval (c_1^1, x_5), where both β_1^1 and β_2^1 take values between 0 and 1, and $\beta_1^1 + \beta_2^1 = 1$. Also, based on Definition 5.3.1 and Equation (5.15), the M-function value $\frac{1}{24}\beta^{c_1^1}$ is assigned to the sub-interval (c_1^1, x_5), where $\beta^{c_1^1} = 1$. Thus, the conditional M-function values for the event $X_6|X_5 \in I_2^1$ are

$$
\begin{aligned}
M_{X_6|X_5\in I_2^1}(0, x_1) &= \frac{1}{6} \\
M_{X_6|X_5\in I_2^1}(x_1, c_1^1) &= \frac{1}{6}\beta_1^1 \\
M_{X_6|X_5\in I_2^1}(c_1^1, x_5) &= \frac{1}{6}\beta_2^1 + \frac{1}{24}\beta_1^{c_1^1} \qquad (5.28)\\
M_{X_6|X_5\in I_2^1}(x_5, x_2) &= \frac{1}{6} + \frac{1}{24} \qquad (5.29)\\
M_{X_6|X_5\in I_2^1}(x_2, x_3) &= \frac{1}{6} + \frac{1}{24} \\
M_{X_6|X_5\in I_2^1}(x_3, \infty) &= \frac{1}{6} + \frac{1}{24}
\end{aligned}
$$

where the total conditional probability mass for $X_6 \in (c_1^1, x_2)$ given $X_5 \in (c_1^1, x_2)$, given in Equations (5.28) and (5.29), is 1/6 (β_2^1 +1/4 $\beta_1^{c_1^1}$ + 5/4), where $\beta_2^1 \in [0, 1]$ and $\beta_1^{c_1^1} = 1$, see *Case 3* in the third box of Figure 5.3.

From *Case 3*, where $X_5 \in I_2^1 = (c_1^1, x_2)$, we use Equation (5.22) to derive the NPI lower conditional probability for the event $X_6 > t$ given $X_5 > t$, where $t \in (c_1^1, x_2)$, as

$$
\begin{aligned}
\underline{P}(X_6 > t|X_5 > t) &= Q^{\min}_{X_6|X_5\in I_2^1}(x_1, x_2) + \sum_{z=3}^{4} P_{X_6|X_5\in I_2^1}(x_z, x_{z+1}) \\
&= M_{X_6|X_5\in I_2^1}(x_5, x_2) + P_{X_6|X_5\in I_2^1}(x_2, x_3) + P_{X_6|X_5\in I_2^1}(x_3, \infty) \\
&= \frac{5}{24} + \frac{5}{24} + \frac{5}{24} = \frac{5}{8}
\end{aligned}
$$

where the value of $Q^{\min}_{X_6|X_5 \in I^1_2}(x_1, x_2)$ is obtained by using Equation (5.20), i.e., $Q^{\min}_{X_6|X_5 \in I^1_2}(x_1, x_2) = \frac{1}{6} + \frac{1}{24} = \frac{5}{24}$.

The NPI upper conditional probability for the event $X_6 > t$ given $X_5 > t$, where $t \in (c^1_1, x_2)$, is derived by using Equation (5.23) as follows.

$$\begin{aligned}\overline{P}(X_6 > t|X_5 > t) &= Q^{\max}_{X_6|X_5 \in I^1_2}(x_1, x_2) + \sum_{z=3}^{4} P_{X_6|X_5 \in I^1_2}(x_z, x_{z+1}) \\ &= P_{X_6|X_5 \in I^1_2}(x_1, x_2) + P_{X_6|X_5 \in I^1_2}(x_2, x_3) + P_{X_6|X_5 \in I^1_2}(x_3, \infty) \\ &= \frac{10}{24} + \frac{5}{24} + \frac{5}{24} = \frac{5}{6}\end{aligned}$$

where the value of $Q^{\max}_{X_6|X_5 \in I^1_2}(x_1, x_2)$ is obtained by using Equation (5.21), i.e., $Q^{\max}_{X_6|X_5 \in I^1_2}(x_1, x_2) = P_{X_6|X_5 \in I^1_2}(x_1, x_5) + P_{X_6|X_5 \in I^1_2}(x_5, x_2) = \frac{5}{24} + \frac{5}{24} = \frac{10}{24}$.

Case 4: Given $X_5 \in I^2 = (x_2, x_3)$, the conditional M-function values for $X_6|X_5 \in I^2$, using Definition 5.3.1 with the assumption rc-$A_{(5)}$ given by Equations (5.14) and (5.15), are shown in the fourth box of Figure 5.3.

The β^0_1, β^2_1, β^3_1 and β^4_1 values corresponding to the intervals $(0, x_1), (x_2, x_5)$, (x_5, x_3) and (x_3, ∞), respectively, are equal to 1, since there no censoring is involved in these intervals. For $c^1_1 \in (x_1, x_2)$ given $X_5 \in I^2$, the conditional M-function value $\frac{1}{6}$ that is assigned to interval (x_1, x_2) will be split up and assigned to two sub-intervals with the M-function value $\frac{1}{6}\beta^1_1$ assigned to the sub-interval (x_1, c^1_1) as well as the M-function value $\frac{1}{6}\beta^1_2$ assigned to the sub-interval (c^1_1, x_2), where both β^1_1 and β^1_2 take values between 0 and 1, and $\beta^1_1 + \beta^1_2 = 1$. Also, based on Definition 5.3.1 and Equation (5.15), the M-function value $\frac{1}{24}\beta^{c^1_1}$ is assigned to the sub-interval (c^1_1, x_2), where $\beta^{c^1_1} = 1$. Thus, the conditional M-function values for the event $X_6|X_5 \in I^2$ are

$$\begin{aligned}M_{X_6|X_5 \in I^2}(0, x_1) =& \frac{1}{6} \\ M_{X_6|X_5 \in I^2}(x_1, c^1_1) =& \frac{1}{6}\beta^1_1 \\ M_{X_6|X_5 \in I^2}(c^1_1, x_2) =& \frac{1}{6}\beta^1_2 + \frac{1}{24}\beta^{c^1_1}_1 \\ M_{X_6|X_5 \in I^2}(x_2, x_5) =& \frac{1}{6} + \frac{1}{24} \qquad (5.30) \\ M_{X_6 X_5 \in I^2}(x_5, x_3) =& \frac{1}{6} + \frac{1}{24} \qquad (5.31) \\ M_{X_6|X_5 \in I^2}(x_3, \infty) =& \frac{1}{6} + \frac{1}{24}\end{aligned}$$

where the total conditional probability mass for $X_6 \in (x_2, x_3)$ given $X_5 \in (x_2, x_3)$, given in Equations (5.30) and (5.31), is 5/24+5/24=10/24, see *Case 4* in the fourth box of Figure 5.3.

From *Case 4*, where $X_5 \in I^2 = (x_2, x_3)$, we use Equation (5.22) to derive the NPI lower conditional probability for the event $X_6 > t$ given $X_5 > t$, where $t \in (x_2, x_3)$, as

$$\begin{aligned}\underline{P}(X_6 > t|X_5 > t) &= Q^{\min}_{X_6|X_5\in I^2}(x_2, x_3) + \sum_{z=4}^{4} P_{X_6|X_5\in I^2}(x_z, x_{z+1}) \\ &= M_{X_6|X_5\in I^2}(x_5, x_3) + P_{X_6|X_5\in I^2}(x_3, \infty) \\ &= \frac{5}{24} + \frac{5}{24} = \frac{5}{12}\end{aligned}$$

where the value of $Q^{\min}_{X_6|X_5\in I^2}(x_2, x_3)$ is obtained by using Equation (5.20), i.e., $Q^{\min}_{X_6|X_5\in I^2}(x_2, x_3) = \frac{1}{6} + \frac{1}{24} = \frac{5}{24}$.

The NPI upper conditional probability for the event $X_6 > t$ given $X_5 > t$, where $t \in (x_2, x_3)$, is derived by using Equation (5.23), as follows.

$$\begin{aligned}\overline{P}(X_6 > t|X_5 > t) &= Q^{\max}_{X_6|X_5\in I^2}(x_2, x_3) + \sum_{z=4}^{4} P_{X_6|X_5\in I^2}(x_z, x_{z+1}) \\ &= P_{X_6|X_5\in I^2}(x_2, x_5) + P_{X_6|X_5\in I^2}(x_5, x_3) + P_{X_6|X_5\in I^2}(x_3, \infty) \\ &= \frac{5}{24} + \frac{5}{24} + \frac{5}{24} = \frac{5}{8}\end{aligned}$$

where the value of $Q^{\max}_{X_6|X_5\in I^2}(x_2, x_3)$ is obtained by using Equation (5.21), i.e., $Q^{\max}_{X_6|X_5\in I^2}(x_2, x_3) = P_{X_6|X_5\in I^2}(x_2, x_5) + P_{X_6|X_5\in I^2}(x_5, x_3) = \frac{5}{24} + \frac{5}{24} = \frac{10}{24}$.

Case 5: Given $X_5 \in I^3 = (x_3, \infty)$, the conditional M-function values for $X_6|X_5 \in I^3$, using Definition 5.3.1 with the assumption rc-$A_{(4+1)}$ given by Equations (5.14) and (5.15), are shown in the fifth box of Figure 5.3.

The β^0_1, β^2_1, β^3_1 and β^4_1 values corresponding to the intervals $(0, x_1), (x_2, x_3)$, (x_3, x_5) and (x_5, ∞), respectively, are equal to 1, since there no censoring is involved in these intervals. For $c^1_1 \in (x_1, x_2)$ given $X_5 \in I^3$, the conditional M-function value $\frac{1}{6}$ that is assigned to interval (x_1, x_2) will be split up and assigned to two sub-intervals with the M-function value $\frac{1}{6}\beta^1_1$ assigned to the sub-interval (x_1, c^1_1) as well as the M-function value $\frac{1}{6}\beta^1_2$ assigned to the sub-interval (c^1_1, x_2), where both β^1_1 and β^1_2 take values between 0 and 1, and $\beta^1_1 + \beta^1_2 = 1$. Also, based on Definition 5.3.1 and Equation (5.15), the M-function value for X_6, $\frac{1}{24}\beta^{c^1_1}$, is assigned to the

$t \in (.)$	$\underline{P}(X_6 > t \mid X_5 > t)$	$\overline{P}(X_6 > t \mid X_5 > t)$
$(0, x_1)$	$\frac{5}{6}$	1
(x_1, c_1^1)	$\frac{2}{3}$	$\frac{5}{6}$
(c_1^1, x_2)	$\frac{5}{8}$	$\frac{5}{6}$
(x_2, x_3)	$\frac{5}{12}$	$\frac{5}{8}$
(x_3, ∞)	$\frac{5}{24}$	$\frac{5}{12}$

Table 5.2: Lower and upper conditional probabilities for the event $(X_6 > t \mid X_5 > t)$.

sub-interval (c_1^1, x_2), where $\beta^{c_1^1} = 1$. Thus, the conditional M-function values for the event $X_6 \mid X_5 \in I^3$ are

$$\begin{aligned}
M_{X_6|X_5 \in I^3}(0, x_1) =& \frac{1}{6} \\
M_{X_6|X_5 \in I^3}(x_1, c_1^1) =& \frac{1}{6}\beta_1^1 \\
M_{X_6|X_5 \in I^3}(c_1^1, x_2) =& \frac{1}{6}\beta_2^1 + \frac{1}{24}\beta_1^{c_1^1} \\
M_{X_6|X_5 \in I^3}(x_2, x_3) =& \frac{1}{6} + \frac{1}{24} \\
M_{X_6|X_5 \in I^3}(x_3, x_5) =& \frac{1}{6} + \frac{1}{24} \qquad (5.32) \\
M_{X_6|X_5 \in I^3}(x_5, \infty) =& \frac{1}{6} + \frac{1}{24} \qquad (5.33)
\end{aligned}$$

where the total conditional probability mass for $X_6 \in (x_3, \infty)$ given $X_5 \in (x_3, \infty)$, given in Equations (5.32) and (5.33), is 5/24+5/24=10/24, see *Case 5* in the fifth box of Figure 5.3.

From *Case 5*, where $X_5 \in I^3 = (x_3, \infty)$, we use Equation (5.22) to derive the NPI lower conditional probability for the event $X_6 > t$ given $X_5 > t$, where $t \in (x_3, \infty)$, as

$$\underline{P}(X_6 > t \mid X_5 > t) = Q^{\min}_{X_6|X_5 \in I^3}(x_3, \infty) = M_{X_6|X_5 \in I^3}(x_5, \infty) = \frac{5}{24}$$

where the value of $Q^{\min}_{X_6|X_5 \in I^2}(x_2, x_3)$ is obtained by using Equation (5.20).

The NPI upper conditional probability for the event $X_6 > t$ given $X_5 > t$, where

$t \in (x_3, \infty)$, is derived by using Equation (5.23), as follows.

$$\overline{P}(X_6 > t | X_5 > t) = Q^{\max}_{X_6|X_5 \in I^3}(x_3, \infty) = P_{X_6|X_5 \in I^3}(x_3, x_5) + P_{X_6|X_5 \in I^3}(x_5, \infty)$$
$$= \frac{5}{24} + \frac{5}{24} = \frac{5}{12}$$

where the value of $Q^{\max}_{X_6|X_5 \in I^3}(x_2, x_3)$ is obtained by using Equation (5.21). Therefore, the NPI lower and upper conditional probabilities for the event that $X_6 > t$ given $X_5 > t$, are given in Table 5.2.

The values of the NPI lower and upper probabilities at observations are easily derived from Table 5.2, using the fact that the lower probability is continuous from the left at all observations, given by Equation (5.22), and the upper probability is continuous from the right at event times, given by Equation (5.23). An effect of conditioning on the second future observation X_5 to be in the final interval (x_3, ∞), the NPI lower probability for $X_6 \in (x_3, \infty)$ is positive which is given by the M-function value $\frac{5}{24}$ that assigned to the sub-interval (x_5, ∞).

In the next section, we present NPI lower and upper probabilities for the event $X_{n+1} > t$ and $X_{n+2} > t$, based on the results presented in Sections 5.2 and 5.3.

5.4 Lower and upper probabilities for $X_{n+1} > t$, $X_{n+2} > t$

This section derives the NPI lower and upper probabilities for the event that both future observations X_{n+1} and X_{n+2} are greater than time $t > 0$ [49]. The notation used in this section follow those introduced in Sections 5.2 and 5.3. Let $I^i_{i^*} = (t^i_{i^*}, t^i_{i^*+1})$ be an interval created by the n data observations, $i = 0, 1, 2, \ldots, u$ and $i^* = 1, 2, \ldots, s_i$, that is we have $n+1$ intervals created by the data, and let $I^i = (x_i, x_{i+1})$ be the ith interval created by two consecutive failures and $I^i_a = (t^i_a, x_{i+1})$, $i = 0, 1, \ldots, u$ and $a = 0, 1, \ldots, s_i$. Furthermore, let $M_{X_{n+1} \in I^j_{j^*}}$ be the M-function values for X_{n+1}, based on the assumption rc-$A_{(n)}$ [32], as defined in Definition 5.2.1, where $j = 0, 1, \ldots, u$ and $j^* = 1, 2, \ldots, s_j$. Let $P_{X_{n+1} \in I^j}$ be the probabilities for X_{n+1} to belong to the intervals $I^j = (x_j, x_{j+1})$ as given by Equation (5.2). Let

$M_{X_{n+2} \in I^k_{k^*} | X_{n+1} \in I^j_{j^*}}$ be the conditional M-function values for $X_{n+2} \in I^k = (x_k, x_{k+1})$, $k = 0, 1, \ldots, u$, $k^* = 1, 2, \ldots, s_k$, based on the assumption rc-$A_{(n+1)}$, as defined in Definition 5.3.1. Let $P_{X_{n+2} \in I^k | X_{n+1} \in I^j}$ be the conditional probabilities for the event $\{X_{n+2} \in I^k | X_{n+1} \in I^j\}$, as given by Equation (5.16).

Using the results in Sections 5.2 and 5.3, the derivation of the NPI lower and upper probabilities for the joint event $X_{n+1} > t$ and $X_{n+2} > t$, for all $t > 0$, is presented as follows.

First, we derive the NPI upper probability for the event that $X_{n+1} > t$ and $X_{n+2} > t$, for $t \in [x_i, x_{i+1})$, $i = 0, 1, \ldots, u$.

Theorem 5.4.1 The NPI upper probability is derived by summing all probability masses that can be to the right of t. This means all M-function values assigned to intervals $I^k_{k^*}, I^j_{j^*} \in \{I^i_a, \ldots, I^i_{s_i}, I^{i+1}_0, \ldots, I^{i+1}_{s_{i+1}}, \ldots, I^u_{s_u}\}$ will lead to the following NPI upper probability

$$\overline{P}(X_{n+1} > t, X_{n+2} > t) = \sum_{j=i}^{u} \sum_{k=i}^{u} P_{X_{n+2} \in I^k | X_{n+1} \in I^j} P_{X_{n+1} \in I^j} \tag{5.34}$$

Proof: There are four terms of summations that, when added together, lead to derive the NPI upper probability for the event $X_{n+1} > t$ and $X_{n+2} > t$. We refer to these terms as J_1, J_2, J_3, and J_4, stated in Equations (5.35), (5.36), (5.37), and (5.38), respectively, which are illustrated in detail below.

First, we sum over $I^k_{k^*} \in \{I^{i+1}_0, \ldots, I^{i+1}_{s_{i+1}}, \ldots, I^u_{s_u}\}$ and $I^j_{j^*} \in \{I^{i+1}_0, \ldots, I^{i+1}_{s_{i+1}}, \ldots, I^u_{s_u}\}$, which is equivalent to summing over the intervals I^k and I^j for $k, j \in \{i+1, \ldots, u\}$. This will lead to constant probabilities using Equations (5.2) and (5.16), respectively, so these probabilities are not functions of the α's or β's, so no optimisation is required here. We can write these summations terms as

$$\begin{aligned}
J_1 &= \sum_{j=i+1}^{u} \sum_{k=i+1}^{u} \sum_{j^*=0}^{s_j} \sum_{k^*=0}^{s_k} M_{X_{n+2} \in I^k_{k^*} | X_{n+1} \in I^j_{j^*}} M_{X_{n+1} \in I^j_{j^*}} \\
&= \sum_{j=i+1}^{u} \sum_{k=i+1}^{u} \left[\sum_{j^*=0}^{s_j} M_{X_{n+1} \in I^j_{j^*}} \right] \left[\sum_{k^*=0}^{s_k} M_{X_{n+2} \in I^k_{k^*} | X_{n+1} \in I^j_{j^*}} \right] \\
&= \sum_{j=i+1}^{u} \sum_{k=i+1}^{u} P_{X_{n+2} \in I^k | X_{n+1} \in I^j} P_{X_{n+1} \in I^j}
\end{aligned} \tag{5.35}$$

where summing all M-function values for $X_{n+1} \in I^j_{j^*}$ as well as summing up all conditional M-function values for $X_{n+2} \in I^k_{k^*} | X_{n+1} \in I^j_{j^*}$, in the second equality, lead to the probabilities for the event $X_{n+1} \in I^j$, as well as to the conditional probability masses for the event $X_{n+1} \in I^j | X_{n+1} \in I^j$, for $j = i+1, \ldots, u$ and $k = i+1, \ldots, u$, as in the third equality. Thus, we have advanced from the second equality to the third equality by using Equations (5.2) and (5.16), respectively.

Secondly, we sum over $I^k_{k^*} \in \{I^i_a, \ldots, I^i_{s_i}\}$ and $I^j_{j^*} \in \{I^{i+1}_0, \ldots, I^{i+1}_{s_{i+1}}, \ldots, I^u_{s_u}\}$, which will lead to a function of the β's only, so we need to maximise this function. This leads to

$$\begin{aligned}
J_2 &= \sum_{j=i+1}^{u} \sum_{j^*=0}^{s_j} \sum_{k^*=a}^{s_i} M_{X_{n+2} \in I^k_{k^*} | X_{n+1} \in I^j_{j^*}} M_{X_{n+1} \in I^j_{j^*}} \\
&= \sum_{j=i+1}^{u} \sum_{j^*=0}^{s_j} M_{X_{n+1} \in I^j_{j^*}} \sum_{k^*=a}^{s_i} M_{X_{n+2} \in I^k_{k^*} | X_{n+1} \in I^j_{j^*}} \\
&= \sum_{j=i+1}^{u} Q^{\max}_{X_{n+2} \in I^i_a | X_{n+1} \in I^j} P_{X_{n+1} \in I^j} \\
&= \sum_{j=i+1}^{u} P_{X_{n+2} \in I^i | X_{n+1} \in I^j} P_{X_{n+1} \in I^j}
\end{aligned} \tag{5.36}$$

where in the third equality, the function $Q^{\max}_{X_{n+2} \in I^i_a | X_{n+1} \in I^j_{j^*}}$ is considered to maximise the conditional probability mass for $X_{n+2} \in I^i_a = (t^i_a, x_{i+1})$ given $X_{n+1} \in I^j_{j^*}$, where $j = i+1, \ldots, u$, by assigning all conditional M-function values within the interval $I^i = (x_i, x_{i+1})$ to the right of t^i_a. This leads to the conditional probability mass for the event $X_{n+2} \in I^i$ given $X_{n+1} \in I^j$, where $I^i = (x_i, x_{i+1})$ and $j = i+1, \ldots, u$. Then, we are able to move from the third equality to the fourth equality via the product of the conditional probability $P_{X_{n+2} \in I^i | X_{n+1} \in I^j}$ and the probability mass for the event that $X_{n+1} \in I^j$, where $I^i = (x_i, x_{i+1})$, $I^j = (x_j, x_{j+1})$ and $j = i+1, \ldots, u$.

The function $Q^{\max}_{X_{n+2}\in I^i_a|X_{n+1}\in I^j_{j^*}}$, which is a function of the β's only, is maximised by using Equation (5.21).

Thirdly, we sum over $I^k_{k^*} \in \{I^{i+1}_0, \ldots, I^{i+1}_{s_{i+1}}, \ldots, I^u_{s_u}\}$ and $I^j_{j^*} \in \{I^i_a, \ldots, I^i_{s_i}\}$, which will lead to a function of the α's only, so we need to maximise this function. This leads to

$$\begin{aligned}
J_3 &= \sum_{k=i+1}^{u} \sum_{j^*=a}^{s_i} \sum_{k^*=0}^{s_i} M_{X_{n+2}\in I^k_{k^*}|X_{n+1}\in I^i_{j^*}} M_{X_{n+1}\in I^i_{j^*}} \\
&= \sum_{k=i+1}^{u} \sum_{j^*=a}^{s_i} P_{X_{n+2}\in I^k|X_{n+1}\in I^i_{j^*}} M_{X_{n+1}\in I^i_{j^*}} \\
&= \sum_{k=i+1}^{u} P_{X_{n+2}\in I^k|X_{n+1}\in I^i_{j^*}} Q^{\max}_{X_{n+1}\in I^i_a} \\
&= \sum_{k=i+1}^{u} P_{X_{n+2}\in I^k|X_{n+1}\in I^i} P_{X_{n+1}\in I^i} \qquad (5.37)
\end{aligned}$$

where in the third equality, the function $Q^{\max}_{X_{n+1}\in I^i_a}$ is considered to maximise the probability mass for $X_{n+1} \in I^j_a = (t^j_a, x_{j+1})$, by assigning all M-function values within the interval $I^j = (x_j, x_{j+1})$ to the right of t^i_a, using Equation (5.7). This leads to the probability mass for the event $X_{n+1} \in I^i = (x_i, x_{i+1})$. This has advanced from the third equality to the fourth equality through the product of the conditional probabilities for the event that $X_{n+2} \in I^k|X_{n+1} \in I^i$, and the probability mass for the event that $X_{n+1} \in I^i$, where $k = i+1, \ldots, u$.

Finally, we sum over $I^k_{k^*} \in \{I^i_a, \ldots, I^i_{s_i}\}$ and $I^j_{j^*} \in \{I^i_a, \ldots, I^i_{s_i}\}$, which will lead to functions of the α's and β's, so we need to maximise both functions. This leads to

$$\begin{aligned}
J_4 &= \sum_{j^*=a}^{s_i} \sum_{k^*=a}^{s_i} M_{X_{n+2}\in I^i_{k^*}|X_{n+1}\in I^i_{j^*}} M_(X_{n+1} \in I^i_{j^*}) \\
&= Q^{\max}_{X_{n+2}\in I^i_a|X_{n+1}\in I^i_{j^*}} Q^{\max}_{X_{n+1}\in I^i_a} \\
&= P_{X_{n+2}\in I^i|X_{n+1}\in I^i} P_{X_{n+1}\in I^i} \qquad (5.38)
\end{aligned}$$

where in the second equality, the function $Q^{\max}_{X_{n+1}\in I^i_a}$ is considered to maximise the probability mass for $X_{n+1} \in I^i_a = (t^i_a, x_{i+1})$, by assigning all M-function values within the interval $I^j = (x_j, x_{j+1})$ to the right of t^i_a, using Equation (5.7), which leads to the probability mass $P_{X_{n+1}\in I^i}$. Also, the function $Q^{\max}_{X_{n+2}\in I^i_a|X_{n+1}\in I^i_{j^*}}$ is considered to maximise the conditional probability mass for $X_{n+2} \in I^i_a = (t^i_a, x_{i+1})$ given $X_{n+1} \in I^i_{j^*}$, where $j^* = 1, \ldots, s_j$, by assigning all conditional M-function values within the interval $I^i = (x_i, x_{i+1})$ to the right of t^i_a, using Equation (5.21), which leads to the conditional probability mass $P_{X_{n+2} \in I^i | X_{n+1} \in I^i}$.

As a result, the NPI upper probability for the event that $X_{n+1} > t$ and $X_{n+2} > t$, for $t \in [x_i, x_{i+1})$, $i = 0, 1, \ldots, u$, and for all $t > 0$, is obtained by summing the values of $J_1, \ldots, J_4$, i. e. $\overline{P}(X_{n+1} > t, X_{n+2} > t) = J_1 + J_2 + J_3 + J_4$. Thus, the proof is complete.

□

Next, we derive the NPI lower probability for the event that $X_{n+1} > t$ and $X_{n+2} > t$, for $t \in [t^i_a, t^i_{a+1})$, $i = 0, 1, \ldots, u$ and $a = 0, 1, \ldots, s_i$.

Theorem 5.4.2 This NPI lower probability is derived by summing all probability masses that must be assigned to the right of t^i_{a+1}. This means all M-function values assigned to intervals $I^k_{k^*}, I^j_{j^*} \in \{I^i_{a+1}, \ldots, I^i_{s_i}, I^{i+1}_0, \ldots, I^{i+1}_{s_{i+1}}, \ldots, I^u_{s_u}\}$. This leads to

$$\underline{P}(X_{n+1} > t, X_{n+2} > t) = \sum_{j=i}^{u} \sum_{j^*=a+1}^{s_j} \sum_{k=i}^{u} \sum_{k^*=a+1}^{s_k} M_{X_{n+2} \in I^k_{k^*} | X_{n+1} \in I^j_{j^*}} M_{X_{n+1} \in I^j_{j^*}} \quad (5.39)$$

where we must start from $a + 1$; that is, we start from the first right-censored observation up to s_i within the interval I^i.

Proof: There are four terms of summations that, when added together, lead to derive the NPI lower probability for the event $X_{n+1} > t$ and $X_{n+2} > t$. We refer to these terms as K_1, K_2, K_3, and K_4, stated in Equations (5.40), (5.41), (5.42), and (5.43), respectively, which are illustrated in detail below.

First, similar to the summations in the derivation of the NPI upper probability for this event, given in Equation (5.35), we sum over $I^k_{k^*} \in \{I^{i+1}_0, \ldots, I^{i+1}_{s_{i+1}}, \ldots, I^u_{s_u}\}$ and $I^j_{j^*} \in \{I^{i+1}_0, \ldots, I^{i+1}_{s_{i+1}}, \ldots, I^u_{s_u}\}$, which will lead to constant probabilities using Equations (5.2) and (5.16), respectively, so these probabilities are not functions of the α's or β's, so no optimisation is required here. We can write these summations terms as

$$K_1 = J_1 = \sum_{j=i+1}^{u} \sum_{k=i+1}^{u} P_{X_{n+2} \in I^k | X_{n+1} \in I^j} P_{X_{n+1} \in I^j} \qquad (5.40)$$

Secondly, we sum over $I^k_{k^*} \in \{I^i_{a+1}, \ldots, I^i_{s_i}\}$ and $I^j_{j^*} \in \{I^{i+1}_0, \ldots, I^{i+1}_{s_{i+1}}, \ldots, I^u_{s_u}\}$, which will lead to a function of the β's only, so we need to minimise this function. This leads to

$$\begin{aligned} K_2 &= \sum_{j=i+1}^{u} \sum_{j^*=0}^{s_j} \sum_{k^*=a+1}^{s_i} M_{X_{n+2} \in I^i_{k^*} | X_{n+1} \in I^j_{j^*}} M_{X_{n+1} \in I^j_{j^*}} \\ &= \sum_{j=i+1}^{u} \sum_{j^*=0}^{s_j} M_{X_{n+1} \in I^j_{j^*}} \sum_{k^*=a+1}^{s_i} M_{X_{n+2} \in I^i_{k^*} | X_{n+1} \in I^j_{j^*}} \\ &= \sum_{j=i+1}^{u} \sum_{j^*=0}^{s_j} M_{X_{n+1} \in I^j_{j^*}} Q^{\min}_{X_{n+2} \in I^i_a | X_{n+1} \in I^j_{j^*}} \\ &= \sum_{j=i+1}^{u} Q^{\min}_{X_{n+2} \in I^i_a | X_{n+1} \in I^j} P_{X_{n+1} \in I^j} \end{aligned} \qquad (5.41)$$

where in the third equality, the function $Q^{\min}_{X_{n+2} \in I^i_a | X_{n+1} \in I^j_{j^*}}$ is considered to minimise the conditional probability mass for $X_{n+2} \in I^i_a = (t^i_a, x_{i+1})$ given $X_{n+1} \in I^j_{j^*}$, where $j = i+1, \ldots, u$, by assigning all conditional M-function values within the interval $I^i = (x_i, x_{i+1})$ to the left of t^i_a. This leads to the conditional probability mass $Q^{\min}_{X_{n+2} \in I^i_a | X_{n+1} \in I^j}$; that is we sum all the conditional probability mass for the event $X_{n+2} \in (c^i_{i^*}, x_{i+1})$ given $X_{n+1} \in I^j$, where $i = 0, 1, 2, \ldots, u$, $i^* = 1, 2, \ldots, s_i$ and $j = i+1, \ldots, u$. Then, we are able to move from the third equality to the fourth equality via the product of the conditional probability $Q^{\min}_{X_{n+2} \in I^i_a | X_{n+1} \in I^j}$, and the probability masses for the event that $X_{n+1} \in I^j$. The function $Q^{\min}_{X_{n+2} \in I^i_a | X_{n+1} \in I^j_{j^*}}$, which is a function of the β's only, is minimised by using Equation (5.20).

Thirdly, we sum over $I^k_{k^*} \in \{I^{i+1}_0, \ldots, I^{i+1}_{s_{i+1}}, \ldots, I^u_{s_u}\}$ and $I^j_{j^*} \in \{I^i_{a+1}, \ldots, I^i_{s_i}\}$, which will lead to a function of the α's only, so we need to minimise this function.

This leads to

$$K_3 = \sum_{k=i+1}^{u} \sum_{j^*=a+1}^{s_i} \sum_{k^*=0}^{s_k} M_{X_{n+2}\in I^k_{k^*}|X_{n+1}\in I^i_{j^*}} M_{X_{n+1}\in I^i_{j^*}}$$
$$= \sum_{k=i+1}^{u} \sum_{j^*=a+1}^{s_i} P_{X_{n+2}\in I^k|X_{n+1}\in I^i_{j^*}} M_{X_{n+1}\in I^i_{j^*}}$$
$$= \sum_{k=i+1}^{u} P_{X_{n+2}\in I^k|X_{n+1}\in I^i_a} Q^{\min}_{X_{n+1}\in I^i_a} \quad (5.42)$$

where in the third equality, the function $Q^{\min}_{X_{n+1}\in I^i_a}$ is considered to minimise the probability mass for $X_{n+1} \in I^i_a = (t^i_a, x_{i+1})$, by assigning all M-function values within the interval $I^j = (x_j, x_{j+1})$ to the left of t^i_a. This leads to the probability mass for the event $X_{n+1} \in (c^i_{i^*}, x_{i+1})$, using Equation (5.6). This enables us to obtain the product of the conditional probabilities for the event that $X_{n+2} \in I^k|X_{n+1} \in I^i_a$, and the probability mass $Q^{\min}_{X_{n+1}\in I^i_a}$, where $k = i+1, \ldots, u$ and $I^i_a = (t^i_a, x_{i+1})$.

Finally, we sum over $I^k_{k^*} \in \{I^i_{a+1}, \ldots, I^i_{s_i}\}$ and $I^j_{j^*} \in \{I^i_{a+1}, \ldots, I^i_{s_i}\}$, which will lead to functions of the α's and β's, so we need to minimise both functions. This leads to

$$K_4 = \sum_{j^*=a+1}^{s_i} \sum_{k^*=a+1}^{s_i} M_{X_{n+2}\in I^i_{k^*}|X_{n+1}\in I^i_{j^*}} M_{X_{n+1}\in I^i_{j^*}}$$
$$= Q^{\min}_{X_{n+2}\in I^i_a|X_{n+1}\in I^i_a} Q^{\min}_{X_{n+1}\in I^i_a} \quad (5.43)$$

where in the second equality, the function $Q^{\min}_{X_{n+1}\in I^i_a}$ is considered to minimise the probability mass for $X_{n+1} \in I^i_a = (t^i_a, x_{i+1})$, by assigning all M-function values within the interval $I^i = (x_i, x_{i+1})$ to the left of t^i_a, using Equation (5.6). Also, the function $Q^{\min}_{X_{n+2}\in I^i_a|X_{n+1}\in I^i_a}$ is considered to minimise the conditional probability mass for $X_{n+2} \in I^i_a = (t^i_a, x_{i+1})$ given $X_{n+1} \in I^i_a$, by assigning all conditional M-function values within the interval $I^i = (x_i, x_{i+1})$ to the left of t^i_a, using Equation (5.20).

As a result, the NPI lower probability for the event that $X_{n+1} > t$ and $X_{n+2} > t$, for $t \in [t^i_a, t^i_{a+1})$, $i = 0, 1, \ldots, u$ and $a = 0, 1, \ldots, s_i$, and for all $t > 0$, is obtained by summing the values of $K_1, \ldots, K_4$, i. e. $\underline{P}(X_{n+1} > t, X_{n+2} > t) = K_1 + K_2 + K_3 + K_4$. Thus, the proof is complete.

$\square$

	$X_6 \in I^0 = (0, x_1)$	$X_6 \in I_1^1 = (x_1, c_1^1)$	$X_6 \in I_2^1 = (c_1^1, x_2)$	$X_6 \in I^2 = (x_2, x_3)$	$X_6 \in I^3 = (x_3, \infty)$	Total
$X_5 \in I^0 = (0, x_1)$	$2 \cdot \frac{1}{30}$	$\frac{1}{30}\beta_1^2$	$\frac{1}{30}(\beta_2^2 + \frac{1}{3})$	$\frac{1}{30} \cdot \frac{4}{3}$	$\frac{1}{30} \cdot \frac{4}{3}$	$\frac{1}{5}$
$X_5 \in I_1^1 = (x_1, c_1^1)$	$\frac{1}{30}\alpha_1^1$	$\frac{1}{30}\alpha_1^1(1 + \beta_1^2)$	$\frac{1}{30}\alpha_1^1(\beta_2^2 + \frac{1}{3})$	$\frac{1}{30} \cdot \frac{4}{3}\alpha_1^1$	$\frac{1}{30} \cdot \frac{4}{3}\alpha_1^1$	$\frac{1}{5}\alpha_1^1$
$X_5 \in I_2^1 = (c_1^1, x_2)$	$\frac{1}{30}(\alpha_2^1 + \frac{1}{3})$	$\frac{1}{30}\beta_1^1(\alpha_2^1 + \frac{1}{3})$	$\frac{1}{30}(\alpha_2^1 + \frac{1}{3})(\beta_2^1 + \frac{3}{2})$	$\frac{1}{30}(\alpha_2^1 + \frac{1}{3})(\frac{5}{4})$	$\frac{1}{30}(\alpha_2^1 + \frac{1}{3})(\frac{5}{4})$	$\frac{1}{5}\alpha_2^1 + \frac{1}{15}$
$X_5 \in I^2 = (x_2, x_3)$	$\frac{1}{30} \cdot \frac{4}{3}$	$\frac{1}{30} \cdot \frac{4}{3}\beta_1^1$	$\frac{1}{30} \cdot \frac{4}{3}(\beta_2^1 + \frac{1}{4})$	$\frac{1}{30} \cdot \frac{10}{3}$	$\frac{1}{30} \cdot \frac{5}{3}$	$\frac{4}{15}$
$X_5 \in I^3 = (x_3, \infty)$	$\frac{1}{30} \cdot \frac{4}{3}$	$\frac{1}{30} \cdot \frac{4}{3}\beta_1^1$	$\frac{1}{30} \cdot \frac{4}{3}(\beta_2^1 + \frac{1}{4})$	$\frac{1}{30} \cdot \frac{5}{3}$	$\frac{1}{30} \cdot \frac{10}{3}$	$\frac{4}{15}$

Table 5.3: Joint probability of X_5 and X_6, according to Example 5.4.1

The following example illustrates the NPI lower and upper probabilities for the events $X_5 > t$ and $X_6 > t$, in particular it shows the steps leading to these lower and upper probabilities in Theorems 5.4.1 and 5.4.2.

Example 5.4.1 Consider again the data set used in Examples 5.2.1 and 5.3.1, for which we have $n = 4$ observations, including one right-censored observation. Based on the probability masses for X_5, presented in Figure 5.2, and the conditional probability masses for $X_6|X_5$, in Figure 5.3, the joint probability masses for X_5 and X_6 are given in Table 5.3. Note that from Examples 5.2.1 and 5.3.1, $\alpha_1^{c_1^1} = 1$, $\beta_1^{c_1^2} = 1$, and $\beta_1^{c_1^1} = 1$.

From Table 5.3, the upper probability for the event that $X_5 > t$ and $X_6 > t$ when $t \in (c_1^1, x_2)$ can be calculated by summing all probabilities represented by blue cells as well as maximising all probability masses represented by green, purple, and red cells. We refer to these summations terms as J_1, J_2, J_3, and J_4, as given by Equations (5.35), (5.36), (5.37) and (5.38), respectively. These summations are illustrated in detail below.

Considering $t \in (c_1^1, x_2)$, we first sum over I^2 and I^3, respectively, where X_6 is in intervals I^2, I^3 given X_5 is in these intervals I^2, I^3, respectively. This will lead to constant probabilities which are represented by the blue cells in Table 5.3, which is not a function of the α's or β's, so no optimisation is required here. These summations are derived by using Equation (5.35), as

$$\begin{aligned} J_1 &= P_{X_6\in I^2|X_5\in I^2}P_{X_5\in I^2} + P_{X_6\in I^2|X_5\in I^3}P_{X_5\in I^3} + P_{X_6\in I^3|X_5\in I^2}P_{X_5\in I^2} \\ &+ P_{X_6\in I^3|X_5\in I^3}P_{X_5\in I^3} \\ &= \frac{1}{9} + \frac{1}{18} + \frac{1}{18} + \frac{1}{9} = \frac{1}{3}. \end{aligned}$$

By using Equation (5.36), we sum over the case where X_6 is in interval $I_2^1 = (c_1^1, x_2)$, given X_5 is in intervals I^2 and I^3, respectively, represented by the red cells in Table 5.3. This will lead to a function of the β's only, so we need to maximise this function. This leads to

$$\begin{aligned} J_2 &= Q^{\max}_{X_6\in I_2^1|X_5\in I^2}P_{X_5\in I^2} + Q^{\max}_{X_6\in I_2^1|X_5\in I^3}P_{X_5\in I^3} \\ &= \frac{2}{45}(\beta_2^1 + \frac{1}{4}) + \frac{2}{45}(\beta_2^1 + \frac{1}{4}) = \frac{2}{45}(2\beta_2^1 + \frac{1}{2}) \end{aligned}$$

Here, by using Equation (5.21), the function $Q_{X_6\in I_2^1|X_5\in I^i} = 2\beta_2^1 + \frac{1}{2}$, for $i = 2, 3$, which is a function of the β's only, is maximised by assigning all conditional M-function values within the interval $I^1 = (x_1, x_2)$ to the right of c_1^1. This is achieved when $\beta_2^1 = 1$, so $\beta_1^1 = 0$ and $Q^{\max}_{X_6\in I_2^1|X_5\in I^i} = 2+\frac{1}{2} = \frac{5}{2}$. Consequently, $J_2 = \frac{5}{2}\times\frac{2}{45} = \frac{1}{9}$.

Using Equation (5.37), we sum over the case where X_6 is in intervals I^2 and I^3, given X_5 is in interval $I_2^1 = (c_1^1, x_2)$, respectively, represented by the purple cells in Table 5.3. This will lead to a function of the α's only, so we need to maximise this function. This leads to

$$\begin{aligned} J_3 &= P_{X_6\in I^2|X_5\in I_2^1}Q^{\max}_{X_5\in I_2^1} + P_{X_6\in I^3|X_5\in I_2^1}Q^{\max}_{X_5\in I_2^1} \\ &= \frac{1}{24}(\alpha_2^1 + \frac{1}{3}) + \frac{1}{24}(\alpha_2^1 + \frac{1}{3}) = \frac{1}{24}(2\alpha_2^1 + \frac{2}{3}) \end{aligned}$$

Here, the function $Q_{X_5\in I_2^1} = 2\alpha_2^1+\frac{2}{3}$, which is a function of the α's only, is maximised, using Equation (5.7), when $\alpha_2^1 = 1$, so $\alpha_2^1 = 0$ and $Q^{\max}_{X_5\in I_2^1} = 2+\frac{2}{3} = \frac{8}{3}$. Consequently, $J_3 = \frac{1}{24}\times\frac{8}{3} = \frac{1}{9}$.

Finally, by using Equation (5.38), we sum over the case where X_6 given X_5 are both in the interval $I_2^1 = (c_1^1, x_2)$, represented by the green cells in Table 5.3. This will lead to functions of the α's and β's, so we need to maximise both functions.

This leads to

$$
\begin{aligned}
J_4 &= Q^{\max}_{X_6 \in I^1_2 | X_5 \in I^1_2} Q^{\max}_{X_5 \in I^1_2} \\
&= \frac{1}{30}(\beta^1_2 + \frac{3}{2})(\alpha^1_2 + \frac{1}{3})
\end{aligned}
$$

By using Equation (5.7), the function $Q_{X_5 \in I^1_2} = \alpha^1_2 + \frac{1}{3}$ is maximised when $\alpha^1_2 = 1$, so $\alpha^1_1 = 0$ and $Q^{max}_{X_5 \in I^1_2} = 1 + \frac{1}{3} = \frac{4}{3}$. And the function $Q_{X_6 \in I^1_2 | X_5 \in I^1_2} = \beta^1_2 + \frac{3}{2}$ is maximised, using Equation (5.21), by assigning all conditional M-function values within the interval $I^1 = (x_1, x_2)$ to the right of c^1_1, i.e., when $\beta^1_2 = 1$, so $\beta^1_1 = 0$ and $Q^{max}_{X_6 \in I^1_2 | X_5 \in I^1_2} = 1 + \frac{3}{2} = \frac{5}{2}$. Consequently, $J_4 = \frac{1}{30} \times \frac{5}{2} \times \frac{4}{3} = \frac{1}{9}$.

As a result, the NPI upper probability for the events $X_5 > t$ and $X_6 > t$, for $t \in (c^1_1, x_2)$, is obtained by summing $J_1 + J_2 + J_3 + J_4$, that is $1/3 + 1/9 + 1/9 + 1/9 = 2/3$. Thus, the NPI upper probability for the events $X_5 > t$ and $X_6 > t$, where $t \in (c^1_1, x_2)$, is $\overline{P}(X_5 > t, X_6 > t) = \frac{2}{3}$ (see Table 5.4). The NPI upper probabilities for the events $X_5 > t$ and $X_6 > t$, for t in other intervals are given in Table 5.4, these have all been derived similarly using corresponding values of $J_1, \ldots, J_4$.

The NPI lower probability for the event that $X_5 > t$ and $X_6 > t$ when $t \in (x_1, c^1_1)$ can be calculated by summing all probabilities represented by blue cells as well as minimising all probability masses represented by green, purple, and red cells of the Table 5.3. We refer to these summations terms as K_1, K_2, K_3 and K_4, as given by Equations (5.40), (5.41), (5.42) and (5.43), respectively. These summations are illustrated in detail below.

First, and similar to the summation of the upper case which was represented by Equation (5.35), we sum over I^2 and I^3, respectively, where X_6 is in intervals I^2, I^3 given X_5 is in these intervals I^2, I^3, respectively. This will lead to constant probabilities which are represented by the blue cells in Table 5.3, which is not a function of the α's or β's, so no optimisation is required here. These summations are derived by using Equation (5.40), that is $K_1 = J_1 = 1/3$.

By using Equation (5.41), we sum over the case where X_6 is in interval I^1_2, given X_5 is in intervals I^2 and I^3, respectively, represented by the red cells in Table 5.3. This will lead to a function of the β's only, so we need to minimise this function.

$t \in (.)$	$\underline{P}(X_5 > t, X_6 > t)$	$\overline{P}(X_5 > t, X_6 > t)$
$(0, x_1)$	$\frac{2}{3}$	1
(x_1, c_1^1)	$\frac{2}{5}$	$\frac{2}{3}$
(c_1^1, x_2)	$\frac{1}{3}$	$\frac{2}{3}$
(x_2, x_3)	$\frac{1}{9}$	$\frac{1}{3}$
(x_3, ∞)	0	$\frac{1}{9}$

Table 5.4: NPI lower and upper probabilities for the event $(X_5 > t, X_6 > t)$, Example 5.4.1.

This leads to

$$\begin{aligned} K_2 &= Q^{min}_{X_6 \in I_2^1 | X_5 \in I^2} P_{X_5 \in I^2} + Q^{min}_{X_6 \in I_2^1 | X_5 \in I^3} P_{X_5 \in I^3} \\ &= \frac{2}{45}(2\beta_2^1 + \frac{1}{2}) \end{aligned}$$

Here, by using Equation (5.20), the function $Q_{X_6 \in I_2^1 | X_5 \in I^2} = 2\beta_2^1 + \frac{1}{2}$, for $i = 2, 3$, which is a function of the β's only, is minimised by assigning all conditional M-function values within the interval $I^1 = (x_1, x_2)$ to the left of c_1^1. This is achieved when $\beta_2^1 = 0$, so $\beta_1^1 = 1$ and $Q^{min}_{X_6 \in I_2^1 | X_5 \in I^2} = \frac{1}{2}$. Consequently, $K_2 = \frac{2}{45} \times \frac{1}{2} = \frac{1}{45}$.

Using Equation (5.42), we sum over the case where X_6 is in intervals I^2 and I^2, given X_5 is in interval I_2^1, respectively, represented by the purple cells in Table 5.3. This will lead to a function of the α's only, so we need to minimise this function. This leads to

$$\begin{aligned} K_3 &= P_{X_6 \in I^2 | X_5 \in I_2^1} Q^{min}_{X_5 \in I_2^1} + P_{X_6 \in I^3 | X_5 \in I_2^1} Q^{min}_{X_5 \in I_2^1} \\ &= \frac{1}{24}(2\alpha_2^1 + \frac{2}{3}) \end{aligned}$$

Here, the function $Q_{X_5 \in I_2^1} = 2\alpha_2^1 + \frac{2}{3}$, which is a function of the α's only, is minimised, using Equation (5.6), when $\alpha_2^1 = 0$, so $\alpha_1^1 = 1$ and $Q^{min}_{X_5 \in I_2^1} = \frac{2}{3}$. Consequently, $K_3 = \frac{1}{24} \times \frac{2}{3} = \frac{1}{36}$.

Finaly, by using Equation (5.43), we sum over the case where X_6 given X_5 are both in interval $I_2^1 = (c_1^1, x_2)$, represented by the green cells in Table 5.3. This will

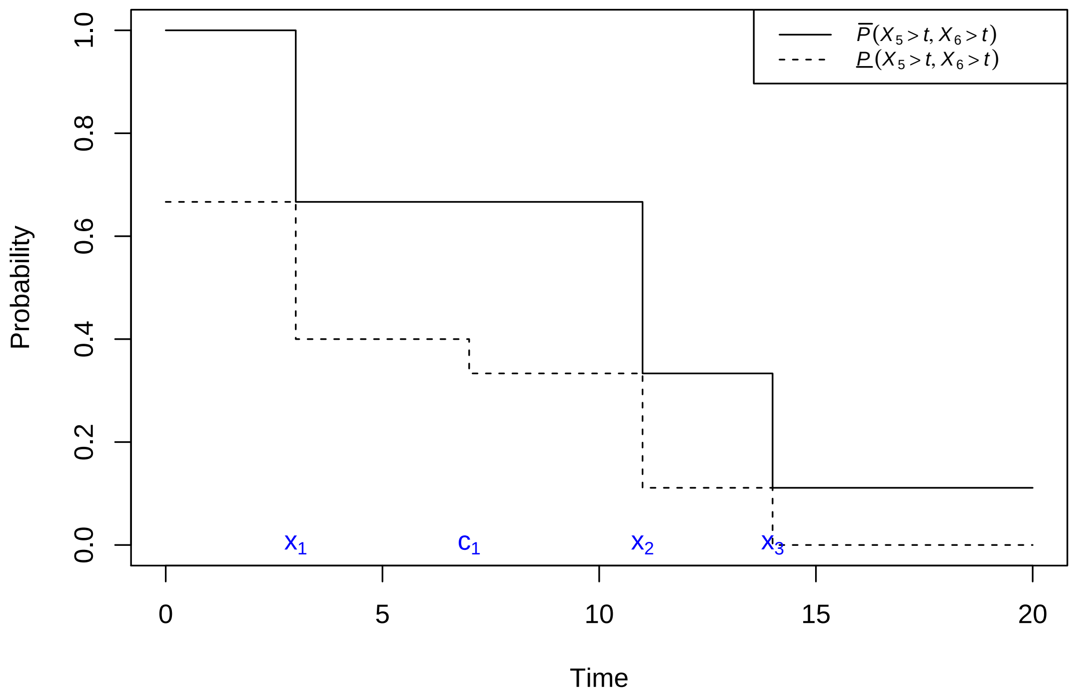

Figure 5.4: NPI lower and upper probabilities for event $X_5 > t$ and $X_6 > t$, Example 5.4.1.

lead to functions of the α's and β's, so we need to minimise both functions. This leads to

$$\begin{aligned} K_4 &= Q^{min}_{X_6 \in I^1_2 | X_5 \in I^1_2} Q^{min}_{X_5 \in I^1_2} \\ &= \frac{1}{30}(\beta^1_2 + \frac{3}{2})(\alpha^1_2 + \frac{1}{3}) \end{aligned}$$

By using Equation (5.6), the function $Q_{X_5 \in I^1_2} = \alpha^1_2 + \frac{1}{3}$ is minimised when $\alpha^1_2 = 0$, so $\alpha^1_1 = 1$ and $Q^{min}_{X_5 \in I^1_2} = \frac{1}{3}$. And the function $Q_{X_6 \in I^1_2 | X_5 \in I^1_2} = \beta^1_2 + \frac{3}{2}$ is minimised, using Equation (5.20), by assigning all conditional M-function values within the interval $I^1 = (x_1, x_2)$ to the left of c^1_1, i.e., when $\beta^1_2 = 0$, so $\beta^1_1 = 1$ and $Q^{min}_{X_6 \in I^1_2 | X_5 \in I^1_2} = \frac{3}{2}$. Consequently, $K_4 = \frac{1}{30} \times \frac{3}{2} \times \frac{1}{3} = \frac{1}{60}$.

As a result, the NPI lower probability for the events $X_5 > t$ and $X_6 > t$, for $t \in (x_1, c^1_1)$, is obtained by summing $K_1 + K_2 + K_3 + K_4$, that is $1/3 + /45 + /36 + 1/60 = 2/5$. Thus, the NPI lower probability for the events $X_5 > t$ and $X_6 > t$, where $t \in (x_1, c^1_1)$, is $\underline{P}(X_5 > t, X_6 > t) = \frac{2}{5}$ (see Table 5.4). The NPI lower probabilities for the events $X_5 > t$ and $X_6 > t$, for t in other intervals are given in Table 5.4,

$t \in (.)$	$\underline{P}(X_5 > t)$	$\overline{P}(X_5 > t)$	$\underline{P}(X_6 > t \| X_5 > t)$	$\overline{P}(X_6 > t \| X_5 > t)$	$\underline{P}(X_5 > t, X_6 > t)$	$\overline{P}(X_5 > t, X_6 > t)$
$(0, x_1)$	$\frac{4}{5}$	1	$\frac{5}{6}$	1	$\frac{2}{3}$	1
(x_1, c_1^1)	$\frac{3}{5}$	$\frac{4}{5}$	$\frac{2}{3}$	$\frac{5}{6}$	$\frac{2}{5}$	$\frac{2}{3}$
(c_1^1, x_2)	$\frac{8}{15}$	$\frac{4}{5}$	$\frac{5}{8}$	$\frac{5}{6}$	$\frac{1}{3}$	$\frac{2}{3}$
(x_2, x_3)	$\frac{4}{15}$	$\frac{8}{15}$	$\frac{5}{12}$	$\frac{5}{8}$	$\frac{1}{9}$	$\frac{1}{3}$
(x_3, ∞)	0	$\frac{4}{15}$	$\frac{5}{24}$	$\frac{5}{12}$	0	$\frac{1}{9}$

Table 5.5: NPI lower and upper probabilities for the events $(X_5 > t)$, $(X_6 > t|X_5 > t)$ and $(X_5 > t, X_6 > t)$, Example 5.4.1.

and shown in Figure 5.4, these have all been derived similarly using corresponding values of $K_1, \ldots, K_4$.

Based on using the results in Sections 5.2 and 5.3, we also get the same results of the derivation of the NPI lower and upper probabilities for the joint event $X_{n+1} > t$ and $X_{n+2} > t$, presented above, if we straightforwardly multiply the NPI lower and upper probabilities for the event $X_{n+1} > t$, given by Equations (5.8) and (5.9), respectively, in Section 5.2, with the NPI lower and upper conditional probabilities for the event that $X_{n+2} > t$ given $X_{n+1} > t$, given by Equations (5.22) and (5.23), respectively, in Section 5.3. So, for $t \in [t_a^i, t_{a+1}^i)$ with $i = 0, 1, \ldots, u$ and $a = 0, 1, \ldots, s_i$, the NPI lower probability for the joint event $X_{n+1} > t$ and $X_{n+2} > t$, is

$$\underline{P}(X_{n+1} > t, X_{n+2} > t) = \underline{P}(X_{n+2} > t|X_{n+1} > t)\underline{P}(X_{n+1} > t) \tag{5.44}$$

and for $t \in [x_i, x_{i+1})$ with $i = 0, 1, \ldots, u$, the corresponding NPI upper probability for the joint event $X_{n+1} > t$ and $X_{n+2} > t$, is

$$\overline{P}(X_{n+1} > t, X_{n+2} > t) = \overline{P}(X_{n+2} > t|X_{n+1} > t)\overline{P}(X_{n+1} > t) \tag{5.45}$$

Again, we consider the data set used in Examples 5.2.1 and 5.3.1, for which we have $n = 4$ observations, including one right-censored observation. If we multiply the results of the NPI lower and upper probabilities for the event $X_5 > t$, presented in Table 5.1, with the corresponding results of the NPI lower and upper conditional probabilities for the event $X_6 > t|X_5 > t$, presented in Table 5.2, then we get the same results of the NPI lower and upper probabilities for the joint event that $X_5 > t$ and $X_6 > t$, shown in Table 5.4 and Figure 5.4. The results presented in Table 5.5 illustrate this point clearly.

The following section illustrates how the proposed method, presented in this chapter, can be applied to the reliability of a series system.

5.5 Reliability of a series system

This section introduces an application for the proposed method in this chapter. It considers a system of three pairs of parallel components arranged in series, where the components in each parallel pair are both of types A, B, or C, as presented in Figure 5.5. For each type, 20 components were tested, leading to the observed failure times and right-censoring times presented in Table 5.6. The failure times of components of different types are assumed to be independent, while failure times of components of the same type are assumed to be exchangeable. Right-censoring is assumed to be non-informative with regard to the component's remaining time to failure.

Consider the following notation. Let m_A, m_B and m_C represent the number of components of Type A, B and C, respectively, so $m_A = m_B = m_C = 2$, with 20 observations for each type, so $n_A = n_B = n_C = 20$. In addition, let $X^A_{i,1}$ and $X^A_{i,2}$, for $i = 1, 2, \ldots, n_A$, let $X^B_{i,1}$ and $X^B_{i,2}$, for $i = 1, 2, \ldots, n_B$, and let $X^C_{i,1}$ and $X^C_{i,2}$, for $i = 1, 2, \ldots, n_C$, represent the two components of types A, B, C, respectively. Let $T^A_{n_A}$, $T^B_{n_B}$ and $T^C_{n_C}$ are denote the minimum of the two components in Types A, B and C, respectively, (e.g. $T^A_{n_A} = \min(X^A_{i,1}, X^A_{i,2})$, etc). Let $\underline{P}_{T^A_{n_A+1}, T^A_{n_A+2}}(t)$ and $\overline{P}_{T^A_{n_A+1}, T^A_{n_A+2}}(t)$ denote the NPI lower and upper probabilities for the event that the two future failure times of components of Type A are both greater than t, with similar notation for Types B and C.

The data for the components failure and right-censoring times, presented in Table 5.6, are derived via simulation. For each component of Type A, 20 failure times are simulated from the Weibull distribution with shape parameter 1.5 and scale parameter 1. Next, the minimum of these two components is obtained, that is

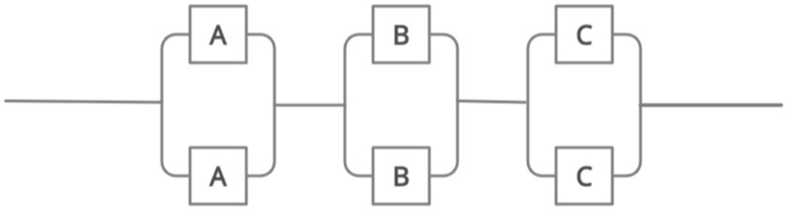

Figure 5.5: A series system with three types of components A, B and C, with two components of each type.

$T^A_{20} = \min(X^A_{i,1}, X^A_{i,2})$, for $i = 1, 2, \ldots, 20$. For each component of Type B, $X^B_{i,1}$ and $X^B_{i,2}$, 17 failure times, and three right-censoring times are simulated from the Weibull distribution with shape parameter 2 and scale parameter 1 and exponential distributions with a rate of 0.27, respectively. Now, the minimum of these two components is obtained, that is $T^B_{20} = \min(X^B_{i,1}, X^B_{i,2})$, for $i = 1, 2, \ldots, 20$. Also, for each component of Type C, that are $X^C_{i,1}$ and $X^C_{i,2}$, 13 failure times and seven right-censoring times are simulated from the Weibull distribution with shape parameter 3 and scale parameter 1 and exponential distribution with rate of 0.35, respectively. Next, the minimum of these two components is obtained, that is $T^C_{20} = \min(X^C_{i,1}, X^C_{i,2})$, for $i = 1, 2, \ldots, 20$. It should be noted that the presence of censoring events in each Type is simulated using the statistical software R.

In order to compute the reliability of the system for the data set shown in Table 5.6, the results presented in Section 5.4 will be first applied separately for each type of component T^A_{20}, T^B_{20} and T^C_{20}. This leads to the results of the NPI lower and upper probabilities for the event that the two future failure times of components of each type, separately, are both greater than t, as given in Table 5.7 and shown in Figure 5.6. For Type A, we derive the NPI lower and upper probabilities, that are $[\underline{P}, \overline{P}](T^A_{21} > t_A, T^A_{22} > t_A)$, for $t_A \in (0, \mathit{data(A)}, \infty)$ and for Type B, we derive the NPI lower and upper probabilities, that are $[\underline{P}, \overline{P}](T^B_{21} > t_B, T^B_{22} > t_B)$, for $t_B \in (0, \mathit{data(B)}, \infty)$, and finally for Type C, we derive the NPI lower and upper probabilities, that are $[\underline{P}, \overline{P}](T^C_{21} > t_C, T^C_{22} > t_C)$, for $t_C \in (0, \mathit{data(C)}, \infty)$.

Second, the reliability function of the whole system are derived by multiplying the corresponding intersection NPI lower and upper probabilities for each type presented

T_{20}^A		T_{20}^B		T_{20}^C	
0.090	0.461	0.115	0.490	> 0.050	0.593
0.147	0.464	> 0.150	0.496	> 0.161	0.602
0.216	0.472	0.185	0.533	> 0.172	0.604
0.224	0.536	> 0.262	> 0.630	0.257	0.607
0.332	0.552	0.343	0.640	> 0.349	0.693
0.342	0.786	0.401	0.647	0.377	0.728
0.356	0.903	0.421	0.654	> 0.421	0.750
0.377	0.937	0.437	0.729	0.522	0.957
0.388	1.036	0.442	0.852	> 0.539	0.966
0.431	1.400	0.450	1.282	0.563	> 0.976

Table 5.6: Simulated data with the three types of components A, B and C (> indicates a right-censored observation).

in Table 5.7, with the emphasis that the exact values of the t's in this table differ for the different systems. The NPI lower and upper probabilities for the whole reliability system at time t, denoted as $\underline{P}_{T_{21}^S}(t)$ and $\overline{P}_{T_{21}^S}(t)$, respectively, are shown in Figure 5.7. So the reliability function of the whole system is calculated as $[\underline{P},\overline{P}](T_{21}^S > t, T_{22}^C > t) = [\underline{P},\overline{P}](T_{21}^A > t_A, T_{22}^A > t_A) \times [\underline{P},\overline{P}](T_{21}^B > t_B, T_{22}^B > t_B) \times [\underline{P},\overline{P}](T_{21}^C > t_C, T_{22}^C > t_C)$, for $t \in (0, data, \infty)$.

It is worth mentioning that the NPI for the joint event $X_{n+1} > t$ and $X_{n+2} > t$, presented in this chapter, takes into account the dependence between these two variables when there is limited information in the form of n observations in the data. It is of interest to see the effect of taking this dependence carefully into account. For this reason, we will compare the results followed the proposed method with those resulting from ignoring, mistakenly, the dependency between these two future observations, i.e., one would use the squared NPI lower and upper probabilities for the event $X_{n+1} > t$. Next, the results of the proposed method are compared with those if the dependence between the two future observations would be ignored. And

$t \in$	$\underline{P}_{T_{21}^A,T_{22}^A}(t)$	$\overline{P}_{T_{21}^A,T_{22}^A}(t)$	$\underline{P}_{T_{21}^B,T_{22}^B}(t)$	$\overline{P}_{T_{21}^B,T_{22}^B}(t)$	$\underline{P}_{T_{21}^C,T_{22}^C}(t)$	$\overline{P}_{T_{21}^C,T_{22}^C}(t)$
$(0,t_1)$	0.909	1	0.909	1	0.909	1
(t_1,t_2)	0.823	0.909	0.823	0.909	0.905	1
(t_2,t_3)	0.740	0.823	0.818	0.909	0.900	1
(t_3,t_4)	0.662	0.740	0.732	0.818	0.895	1
(t_4,t_5)	0.589	0.662	0.727	0.818	0.795	0.895
(t_5,t_6)	0.520	0.589	0.642	0.727	0.789	0.895
(t_6,t_7)	0.455	0.520	0.561	0.642	0.691	0.789
(t_7,t_8)	0.394	0.455	0.487	0.561	0.684	0.789
(t_8,t_9)	0.338	0.394	0.417	0.487	0.586	0.684
(t_9,t_{10})	0.286	0.338	0.353	0.417	0.579	0.684
(t_{10},t_{11})	0.238	0.286	0.294	0.353	0.482	0.579
(t_{11},t_{12})	0.195	0.238	0.241	0.294	0.395	0.482
(t_{12},t_{13})	0.156	0.195	0.193	0.241	0.316	0.395
(t_{13},t_{14})	0.121	0.156	0.150	0.193	0.246	0.316
(t_{14},t_{15})	0.091	0.121	0.144	0.193	0.184	0.246
(t_{15},t_{16})	0.065	0.091	0.103	0.144	0.132	0.184
(t_{16},t_{17})	0.043	0.065	0.069	0.103	0.088	0.132
(t_{17},t_{18})	0.026	0.043	0.041	0.069	0.053	0.088
(t_{18},t_{19})	0.013	0.026	0.021	0.041	0.026	0.053
(t_{19},t_{20})	0.004	0.013	0.007	0.021	0.009	0.026
(t_{20},∞)	0	0.004	0	0.007	0	0.026

Table 5.7: NPI lower and upper probabilities of Type A, Type B and Type C for the data in Table 5.6.

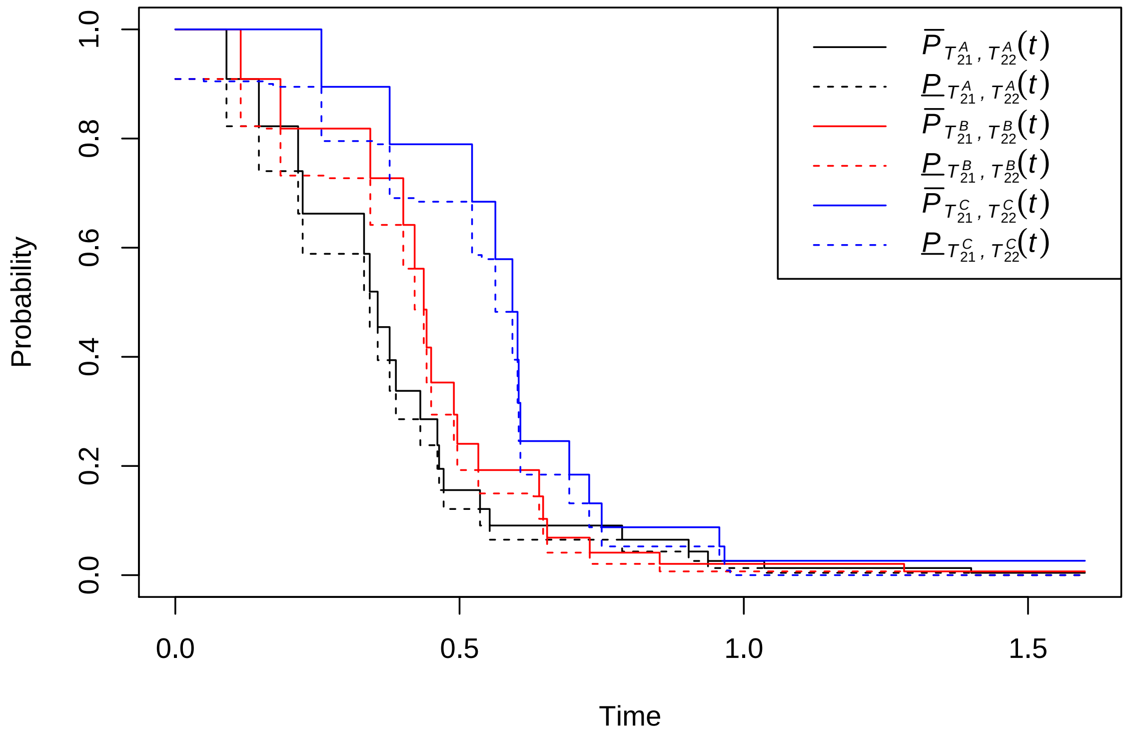

Figure 5.6: NPI lower and upper probabilities for Types A, B and C, respectively, of the series system in Table 5.6.

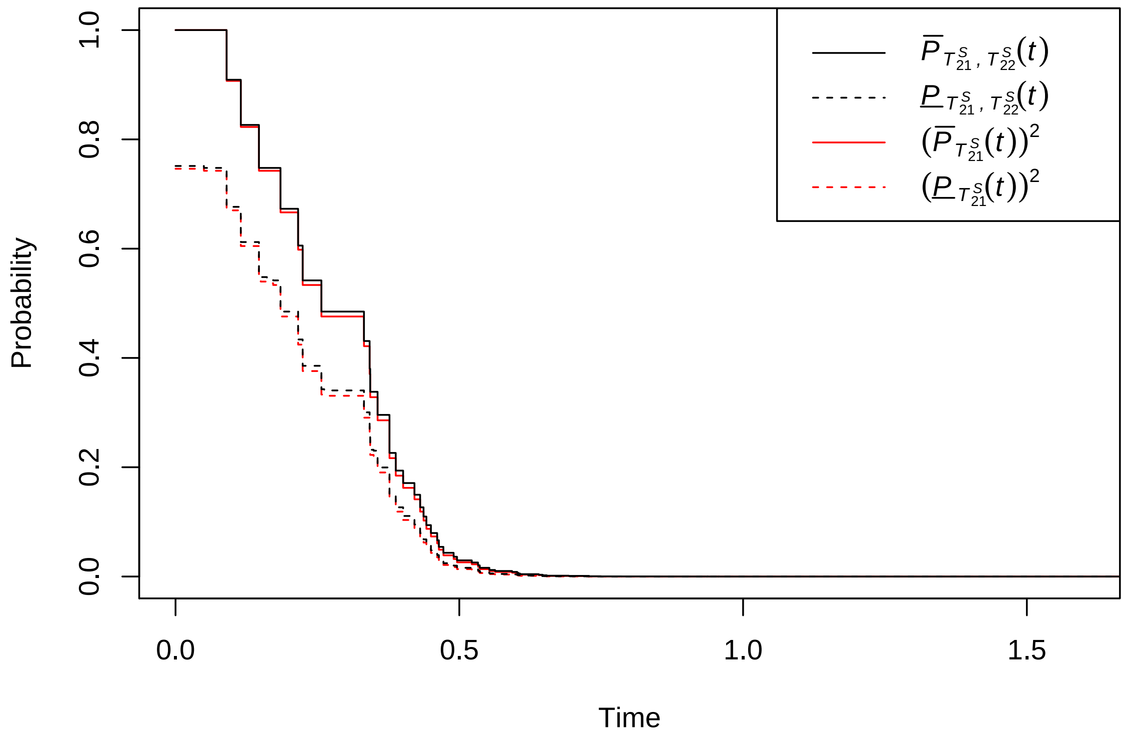

Figure 5.7: NPI lower and upper probabilities for the whole series system.

$(\underline{P}_{T^S_{21}}(t))^2$ and $(\overline{P}_{T^S_{21}}(t))^2$ represent the NPI lower and upper probabilities based on the wrong assumption of independence of the two future observations per type of component, as shown in Figure 5.7.

Figure 5.7 shows that the proposed method in this chapter provides lower and upper probabilities $\underline{P}_{T^S_{21}}(t)$ and $\overline{P}_{T^S_{21}}(t)$ of the system failure time, that are never smaller than the incorrect ones via the squared lower and upper probabilities $(\underline{P}_{T^S_{21}}(t))^2$ and $(\overline{P}_{T^S_{21}}(t))^2$. And they only are equal at the start $(\overline{P}_{T^S_{21}}(t) = (\overline{P}_{T^S_{21}}(t))^2 = 1)$ or end $(\underline{P}_{T^S_{21}}(t) = (\underline{P}_{T^S_{21}}(t))^2 = 0$). While the differences between the lower and upper probabilities may only be small, it should be remarked that for more than two future observations, the differences will be larger. Detailed investigation is left as a topic for future research, as it requires the development of the NPI approach for more than two future observations in case of right-censored data.

5.6 Concluding remarks

This chapter has developed NPI for two future observations for data including right-censored observations [49], in particular considering the event that these two future observations are greater than time t. The rc-$A_{(n)}$ assumption [32], without any other assumptions, is used for the first future observation. Then, the rc-$A_{(n+1)}$ assumption provides a partially specified predictive probability distribution for the second future observation conditioned on the first future observation.

The analytical approach with the α's and β's, presented in chapter, enabled us to find the NPI lower and upper probabilities for any event involving the next two future observations, by minimising and maximising over the α's and β's. The NPI lower and upper probabilities were derived explicitly for the event that the two future observations are both greater than time t but that the method can be used for general events.

By extending the NPI to two future observations with right-censored data, we have taken the dependency between these two variables into account when there is a limited amount of information in the form of n observations in the data. We have compared the results of the proposed method with those obtained when ignoring the dependence between these two future observations. Our results have been applied to system reliability of a small series system containing three types of components, each of which contains multiple components of the same type, in order to demonstrate the practical benefit of this work.

This work has shown that the analytical approach will be very complicated for more than two future observations. One way forward can be linked to actually sample the first future observation, using the M-function values, given by Equations (2.31) and (2.32), for X_{n+1} and an assumption for the distribution of these probabilities within the intervals. Then this samples future observation can be added to the data set and the next one can similarly be sampled. Resulting inferences will of course depend on the assumed distribution per interval, but computationally it will be straightforward. This approach would be related to NPI Bootstrap [17, 28] and smoothed bootstrap for right-censored data [3].

Chapter 6

Conclusions

In this thesis, we have introduced three contributions, which are described in Chapters 3, 4, and 5 with respect to the NPI with right-censored data.

In Chapter 3, we introduced a new approach on which we used the largest observed value within a data set, including right-censored observations, as the end point of support. The new approach allowed us to derive the probability for the event of interest that the actual lifetime corresponding to a right-censored observation would exceed the largest observed value. We extended the new approach to derive the probability for the event that the actual lifetime corresponding to a right-censored observation would exceed the j_{th} largest observed value, as long as it past the largest censored observation in the data set. We applied the proposed methods to the full Supercentenarian data set, but separately for the women and the men. [8]. Our investigation of the Supercentenarian data showed that since the probabilities that somebody would survive the largest observed age were quite high, we do not think that it is appropriate for analysis of extreme values to assume that the largest value is the end-point of support. NPI cannot be used for predicting observations beyond the largest observation due to its weak assumptions, this would require additional distributional assumptions.

In Chapter 4, we assume that time is discrete and consists of the number of events and the number right-censorings that take place at discrete time points. Then, we introduced the NPI as an alternative to the actuarial estimator using NPI for Bernoulli data [20], as discussed in Section 2.3 of Chapter 2. The NPI alternatives to the actuarial estimator is used to derive the NPI lower and upper probabilities for the event that all future observations are greater than a particular discrete time. This development is compared to the theory of NPI for grouped data, established by Coolen and Yan [65]. NPI for Bernoulli data [20] was also used to obtain the NPI lower and upper probabilities for such events of interest such as for the event that at least a future observation out of multiple future observations will survive for all t_j. Then, the results are applied to discrete time system reliability using survival signatures combined with NPI for Bernoulli data [25].

In Chapter 5, we developed NPI for two future observations for right-censored data [49], based on the rc-$A_{(n)}$ assumption [32], without any other additional assumptions. We have proposed an analytical approach to partially specify the probability distribution for the first future observation by the M-function values as well as for the second future observation conditioned on where it can be the first future observation, by the conditional M-function values, respectively. This approach with the α's and β's enabled us to find the NPI lower and upper probabilities for any event of interest involving the next two future observations, by minimizing and maximising over the these α's and $beta$'s. Therefore, the NPI lower and upper probabilities are derived for the event that the two future observations are greater than time t. Next, in this chapter, we were interested in investigating the effect of neglecting the dependence between the future observations, so we compared the results based on the proposed method with those resulting from ignoring the dependency between these two future observations. To conclude this chapter, we implemented our proposed method using a reliability system.

A further topic of interest for further study is related to our method presented in Chapter 5, such as extending the NPI for multiple future observations for right-censored data for the event that all these multiple future observations are greater than time t, and apply the results for some reliability systems. Also, NPI for two future observations for right-censored data [49], presented in Chapter 5, will be considered for at least one more event of interest.

Appendix A

R codes

A.1 On exceedance of the largest observed value

```
X.c <- function(X) {
ifelse(length(X[X[, 2] == 0, ]) > 2, x1 <- X[X[, 2] ==0, ][, 1],
x1 <- X[X[, 2] == 0, ][1])
return(x1)
}
##(Run) the following code to show all lifetime observations:
Xt0 <- function(X) {
Y <- c(0, X[, 1])
return(Y)
}

#reading the old women data set:
data<-read.csv(file.choose(),header=T)

##(Run) the following code to read the data
#which including right-censord obs.:

data=SupercentWomen_data1
data
n=nrow(data)
m=3                 #(m >= 1) future observations
cens=X.c(data)      #the censored observations
tot=Xt0(data)       #all observations
```

```
###################################
#First largest observation

#compute (Ncr -1/(Ncr)) term
#(Ncr is the # of individuals in the risk set
#just prior to the lifetime c(r))
cens[[1]][1]
tot[[2]][1]
product=NULL
for(i in 1:nrow(cens)){
product[i]=(sum(cens[[1]][i]<=tot[[2]])-1)/(sum(cens[[1]][i]
<=tot[[2]]))
}
product
###################################
#Second largest observation

product10nd=NULL
for(i in 1:nrow(cens)){
product10nd[i]=(sum(tot>cens[i,1])-1)/((sum(tot>cens[i,1]))+1)
}
product10nd
###################################
#Third largest observation

product10rd=NULL
for(i in 1:nrow(cens)){
product10rd[i]=(sum(tot>cens[i,1])-2)/((sum(tot>cens[i,1]))+1)
}
product10rd
###################################
#compute the product of the previous products
#for the actual data
prod(product)
#compute the prob that at least one of
#the censored will live longer than 122:
```

```
prob1=1-(prod(product))
prob1
#compute the probability that at least one of
#either the previous censored obs.
#or the multiple future obs. will live longer than 122 yrs old:
n=1580
m=2
prob=1-(prod(product)*(n/(n+m)))      #(general form)
prob

m=seq(0,20000,by=1)
#pry=1-(10/(10+m))
n=nrow(data)
plot(m,prob,type="l",ylim=c(0,1))
abline(v=13009, h=0.9288780)
```

A.2 NPI alternative to the actuarial estimator

```
#Create the Data set :
time=c(0,1,2,3,4,5,6,7,8,9,10,11)
n_tj=c(374,284,208,157,120,95,79,66,62,54,47,0)
c_tj=c(0,0,0,0,12,5,9,9,3,5,5,0)
r=c(374,374,284,208,145,115,86,70,63,57,49,47)
data.frame(time,n_tj,c_tj)
n=374
# To compute Upper NPI alternative to the actuarial estimator:
p_n <- function(n_tj, r, m) {
v1 <- n_tj + 1:m  # - 1
v2 <- r + 1:m
v <- v1 / v2
upr <- prod(v)
return(upr)
}
p_n_vec <- function(n_tj_vec, r_vec, m) {
n_tj_len <- length(n_tj_vec)
```

```
p_n_index <- function(i) {
n_tj <- n_tj_vec[i]
r <- r_vec[i]
upr <- p_n(n_tj = n_tj, r = r, m = m)
return(upr)
}
upr_vec <- sapply(1:n_tj_len, FUN = p_n_index,
simplify = TRUE, USE.NAMES = TRUE)
return(upr_vec)
}
p_n(n_tj = n_tj[1], r = r[1], m = 5)
#p_n_vec(n_tj_vec = n_tj, r_vec = r, m = 1)
upr_m <- p_n_vec(n_tj_vec = n_tj, r_vec = r, m = 5)
upr_m

# For Upper survival based on upper NPI alternative:
U_prob_m=NULL
for(i in 1:(length(time))){
U_prob_m[i]=prod(upr_m[1:i])
}
U_prob_m

#  To compute Lower NPI alternative to the actuarial estimator:
p_m <- function(n_tj, r, m) {
v1 <- n_tj + 1:m - 1
v2 <- r + 1:m
v <- v1 / v2
lwr <- prod(v)
return(lwr)
}
p_m_vec <- function(n_tj_vec, r_vec, m) {
n_tj_len <- length(n_tj_vec)
p_m_index <- function(i) {
n_tj <- n_tj_vec[i]
r <- r_vec[i]
lwr <- p_m(n_tj = n_tj, r = r, m = m)
return(lwr)
```

```
}
lwr_vec <- sapply(1:n_tj_len, FUN = p_m_index,
simplify = TRUE, USE.NAMES = TRUE)
return(lwr_vec)
}
p_m(n_tj = n_tj[1], r = r[1], m = 5)
#p_m_vec(n_tj_vec = n_tj, r_vec = r, m = 2)
lwr_m <- p_m_vec(n_tj_vec = n_tj, r_vec = r, m = 5)
lwr_m

# # For lower survival based on lower NPI alternative:
l_prob_m=NULL
for(i in 2:(length(time))){
l_prob_m[i]=prod(lwr_m[2:i])
}
l_prob_m
```

```
# To compute $x-out-of-y$:
#upper
m=3
n=c(9,7,4,1)
s=c(8,5,2,0)
y=2
#y=c(0,1,2,3)
upper=NULL
for(i in 1:length(n)){
upper[i]=(1/choose(n[i]+m,n[i]))*
((choose(s[i]+y,s[i])*choose(n[i]-s[i]+m-y,n[i]-s[i])))}
upper

#lower
x=2
n=7
s=5
y=2
comp=rep(NA,x)
d1=(choose(n+y,n))^(-1)
```

```
for (i in 1:x) {
comp[i]=choose((s+(i-1)-1),s-1)*choose((n-s+y-(i-1)),n-s)
}
1-d1*sum(comp)
# Function
fun1= function(n,s,y,x) {
comp=rep(NA,x)
d1=(choose(n+y,n))^(-1)
for (i in 1:x) {
comp[i]=choose((s+(i-1)-1),s-1)*choose((n-s+y-(i-1)),n-s)
}
return(1-d1*sum(comp))
}
fun1(n=9,s=8,y=1,x=1)
# full Column
for (i in 1:3) {
print(fun1(n=9,s=8,y=i,x=1) )
}
# full lower matrix with fixed n and s
y=c(1:10)
PL=matrix(NA,nrow=length(y),ncol=length(y))
for (i in y) {
for (j in y) {
PL[i,j]=fun1(n=7,s=5,y=i,x=j)
}
}
PL
```

A.3 NPI for two future observations

The R codes for calculating M-function values and NPI survival for the first future observation were developed by Maturi [51]. Here, we extend it for 2 future observations.

```
rm(list=ls())
# calculate the lower and upper survival functions
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 <- Inf, Y <- c(X.u(X),
Inf))
return(Y)
}
Xt0 <- function(X) {
Y <- c(0, X[, 1])
return(Y)
}

# calculate the product terms to use later for Mfun and prob
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)
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)
}
Prob <- function(X) {
Y <- Xu1(X)
ny <- length(Y)
P4 <- NULL
for (i in 1:ny) {
P4[i] <- cond(X, Y[i])
}
return(P4)
}

#Mfun(Xdata11)
# Towards 2 future observations: Examples in Chapter 5
LUsur<-function(X){
LS<-NULL; US<-NULL
t<-Xt0(X); t0<-c(Xt0(X),Inf)
ts<-c(0,X.u(X))
MM<-cbind(t,Mfun(X))
PP<-cbind(ts,Prob(X))
# Lower
for (j in 1:length(t0)){
if(length(MM[MM[,1]>=t0[j],])==0) LS[j]<-0
```

```
ifelse( length(MM[MM[,1]>=t0[j],])==2,
LS[j]<-MM[MM[,1]>=t0[j],][2], LS[j]<- sum(MM[MM[,1]>=t0[j],][,2]))}
# Upper
for (j in 1:length(t0)){
tmax<-NULL
tmax<-max(ts[ts<=t0[j]])
ifelse(length(PP[ts>=tmax,])==2,US[j]<-PP[ts>=tmax,][[2]],US[j]
<-apply(PP[ts>=tmax,],2,sum)[[2]])}
return(list(LS,US))}

#create the data:
d<-c(1,0,1,1)
x<-1:4
Xdata<-cbind(x,d)
LUsur(Xdata)

x0<-1:5
d1<-c(1,1,0,1,1)
d2<-c(1,1,0,1,1)
d3<-c(1,0,1,1,1)
d4<-c(1,0,1,1,1)
d5<-c(1,0,1,1,1)

#lower
LUsur(cbind(x0,d1))[[1]][2]* LUsur(Xdata)[[1]][2]
LUsur(cbind(x0,d2))[[1]][3]* LUsur(Xdata)[[1]][3]
LUsur(cbind(x0,d3))[[1]][4]* LUsur(Xdata)[[1]][4]
LUsur(cbind(x0,d4))[[1]][5]* LUsur(Xdata)[[1]][5]
LUsur(cbind(x0,d5))[[1]][6]* LUsur(Xdata)[[1]][6]
#upper
LUsur(cbind(x0,d1))[[2]][1]* LUsur(Xdata)[[2]][1]
LUsur(cbind(x0,d2))[[2]][2]* LUsur(Xdata)[[2]][2]
LUsur(cbind(x0,d3))[[2]][3]* LUsur(Xdata)[[2]][3]
LUsur(cbind(x0,d4))[[2]][4]* LUsur(Xdata)[[2]][4]
LUsur(cbind(x0,d5))[[2]][5]* LUsur(Xdata)[[2]][5]
# ex2
d<-c(1,0,0,0,1)
```

```
x<-1:5
Xdata<-cbind(x,d)
LUsur(Xdata)

x0<-1:6
d1<-c(1,1,0,0,0,1)
d2<-c(1,1,0,0,0,1)
d3<-c(1,0,1,0,0,1)
d4<-c(1,0,0,1,0,1)
d5<-c(1,0,0,0,1,1)
d6<-c(1,0,0,0,1,1)

#lower
LUsur(cbind(x0,d1))[[1]][2]* LUsur(Xdata)[[1]][2]
LUsur(cbind(x0,d2))[[1]][3]* LUsur(Xdata)[[1]][3]
LUsur(cbind(x0,d3))[[1]][4]* LUsur(Xdata)[[1]][4]
LUsur(cbind(x0,d4))[[1]][5]* LUsur(Xdata)[[1]][5]
LUsur(cbind(x0,d5))[[1]][6]* LUsur(Xdata)[[1]][6]
LUsur(cbind(x0,d6))[[1]][7]* LUsur(Xdata)[[1]][7]
#upper
LUsur(cbind(x0,d1))[[2]][1]* LUsur(Xdata)[[2]][1]
LUsur(cbind(x0,d2))[[2]][2]* LUsur(Xdata)[[2]][2]
LUsur(cbind(x0,d3))[[2]][3]* LUsur(Xdata)[[2]][3]
LUsur(cbind(x0,d4))[[2]][4]* LUsur(Xdata)[[2]][4]
LUsur(cbind(x0,d5))[[2]][5]* LUsur(Xdata)[[2]][5]
LUsur(cbind(x0,d6))[[2]][6]* LUsur(Xdata)[[2]][6]

## THE END ##
```

Bibliography

[1] Aboalkhair A.M. (2012). Nonparametric predictive inference for system reliability. *PhD-thesis, University of Durham. URL https://maths.durham.ac.uk/stats/people/fc/thesis-AA.pdf S.*

[2] Aboalkhair A.M., Coolen F.P.A. and MacPhee I.M. (2014). Nonparametric predictive inference for reliability of a k-out-of-m: G system with multiple component types. *Reliability Engineering & System Safety*, 131, 298–304.

[3] Al-Luhayb A. (2021). Smoothed bootstrap methods for right-censored data and bivariate data. *PhD-thesis, University of Durham. URL http://etheses.dur.ac.uk/14096/1/Al-Luhayb000724649.pdf.*

[4] Al-nefaiee A.H. (2014). Nonparametric predictive inference for system failure time. *PhD-thesis, University of Durham. URL https://maths.durham.ac.uk/stats/people/fc/thesis-AAN.pdf.*

[5] Al-Nefaiee A.H. and Coolen F.P.A. (2013). Nonparametric predictive inference for system failure time based on bounds for the signature. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 227, 513–522.

[6] Allison P.D. (1982). Discrete-time methods for the analysis of event histories. *Sociological Methodology*, 13, 61–98.

[7] Alves I.F. and Neves C. (2014). Estimation of the finite right endpoint in the Gumbel domain. *Statistica Sinica*, pp. 1811–1835.

[8] Alves I.F., Neves C. and Rosário P. (2017). A general estimator for the right endpoint with an application to supercentenarian women's records. *Extremes*, 20, 199–237.

[9] Arts G.R.J. and Coolen F.P.A. (2008). Two nonparametric predictive control charts. *Journal of Statistical Theory and Practice*, 2, 499–512.

[10] Arts G.R.J., Coolen F.P.A. and van der Laan P. (2004). Nonparametric predictive inference in statistical process control. *Quality Technology & Quantitative Management*, 1, 201–216.

[11] Aslett L.J.M. (2012). ReliabilityTheory: Tools for structural reliability analysis. *Elsevier. R package, URL http://www.louisaslett.com/.*

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

[13] Augustin T., Coolen F.P.A., De Cooman G. and Troffaes M.C. (2014). *Introduction to Imprecise Probabilities.* Wiley, Chichester.

[14] Baker R. (2010). Multinomial nonparametric predictive inference: selection, classification and subcategory data. *PhD-thesis, University of Durham. URL https://maths.durham.ac.uk/stats/people/fc/thesis-RB.pdf.*

[15] Berkson J. and Gage R.P. (1950). Calculation of survival rates for cancer. In *Proceedings of the staff meetings. Mayo Clinic*, volume 25, pp. 270–286.

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

[17] BinHimd S. (2014). Nonparametric predictive methods for bootstrap and test reproducibility. *PhD-thesis, University of Durham. URL https://maths.durham.ac.uk/stats/people/fc/thesis-SBH.pdf.*

[18] Boole G. (1854). *An Investigation of the Laws of Thought: On which are Founded the Mathematical Theories of Logic and Probabilities*, volume 2. Walton and Maberly, Dover.

[19] Coolen F.P.A. (1994). On Bernoulli experiments with imprecise prior probabilities. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 43, 155–167.

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

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

[22] Coolen F.P.A. and Augustin T. (2005). Learning from multinomial data: a nonparametric predictive alternative to the Imprecise Dirichlet Model. In *ISIPTA*, pp. 125–134. Proceedings of the Fourth International Symposium on Imprecise Probabilities and Their Applications, F.G. Cozman, R. Nau and T. Seidenfeld (Eds).

[23] Coolen F.P.A. and Coolen-Maturi T. (2013). Generalizing the signature to systems with multiple types of components. In *Complex systems and dependability*, pp. 115–130. Springer, Berlin.

[24] Coolen F.P.A. and Coolen-Maturi T. (2015). Predictive inference for system reliability after common-cause component failures. *Reliability Engineering & System Safety*, 135, 27–33.

[25] Coolen F.P.A., Coolen-Maturi T. and Al-Nefaiee A.H. (2014). Nonparametric predictive inference for system reliability using the survival signature. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk & Reliability*, 228, 437–448.

[26] Coolen F.P.A. and Coolen-Schrijner P. (2003). A nonparametric predictive method for queues. *European Journal of Operational Research*, 145, 425–442.

[27] Coolen F.P.A. and Coolen-Schrijner P. (2006). Nonparametric predictive subset selection for proportions. *Statistics & Probability Letters*, 76, 1675–1684.

[28] Coolen F.P.A. and Himd S. (2020). Nonparametric predictive inference bootstrap with application to reproducibility of the two-sample Kolmogorov–Smirnov test. *Journal of Statistical Theory and Practice*, 14, 1–13.

[29] Coolen F.P.A. and Maturi T.A. (2010). Nonparametric predictive inference for order statistics of future observations. In *Combining Soft Computing and Statistical Methods in Data Analysis*, pp. 97–104. Proceedings 5th International Conference on Soft Methods in Probability and Statistics, Oviedo (Spain), Sept 2010. Springer, Berlin.

[30] Coolen F.P.A. and Van der Laan P. (2001). Imprecise predictive selection based on low structure assumptions. *Journal of Statistical Planning and Inference*, 98, 259–277.

[31] Coolen F.P.A. and Yan K.J. (2003). Nonparametric Predictive Comparison of Two Groups of Lifetime Data. In *Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications*, pp. 148–161. (eds J.M. Bernard, T. Seidenfeld and M. Zaffalon) Carlton Scientific.

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

[33] Coolen-Maturi T. and Coolen F.P. (2014). Nonparametric predictive inference for combined competing risks data. *Reliability Engineering & System Safety*, 126, 87–97.

[34] Coolen-Maturi T.A., 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, 515–538.

[35] De Finetti B. (1974). *Theory of Probability: A Critical Introductory Treatment*. Wiley, London.

[36] Desu M.M. and Raghavarao D. (2003). *Nonparametric Statistical Methods for Complete and Censored Data.* Chapman and Hall/CRC Press.

[37] Fisher R.A. and Tippett L.H.C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 24, pp. 180–190. Cambridge University Press.

[38] Gnedenko B. (1943). Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of Mathematics*, 4, 423–453.

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

[40] Hill B.M. (1988). De Finetti's Theorem, Induction, and Bayesian nonparametric predictive inference (with discussion). In *Bayesian Statistics*, pp. 211–241. Oxford University Press.

[41] Hill B.M. (1992). Bayesian nonparametric prediction and statistical inference. In *Bayesian Analysis in Statistics and Econometrics*, pp. 43–94. Springer, New York, NY.

[42] Hill B.M. (1993). Parametric models for A(n): splitting processes and mixtures. *Journal of the Royal Statistical Society. Series B*, pp. 423–433.

[43] Janurová K. and Briš R. (2014). A nonparametric approach to medical survival data: Uncertainty in the context of risk in mortality analysis. *Reliability Engineering & System Safety*, 125, 145–152.

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

[45] Kyburg H.E. (1998). Interval-valued probabilities. *Imprecise Probabilities Project, Citeseer.*

[46] Lagakos S.W. (1979). General right censoring and its impact on the analysis of survival data. *Biometrics*, pp. 139–156.

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

[48] Machin D., Cheung Y.B. and Parmar M. (2006). *Survival Analysis: A Practical Approach*. Wiley, New York.

[49] Mahnashi A.M., Coolen-Maturi T. and Coolen F.P.A. (2019). Generalizing nonparametric predictive inference for right-censored data to two future observations. *Proceedings of the 29th European Safety and Reliability Conference (ESREL)*, Hannover (Germany). pp. 2671–2678.

[50] Masyn K.E. (2003). Discrete-time survival mixture analysis for single and recurrent events using latent variables. *PhD-thesis, University of California, Los Angeles. URL https://www.statmodel.com/download/masyndissertation.pdf.*

[51] Maturi T.A. (2010). Nonparametric Predictive Inference for Multiple Comparisons. *PhD-thesis, University of Durham. URL https://maths.durham.ac.uk/stats/people/fc/thesis-AAN.pdf.*

[52] Maturi T.A., Coolen-Schrijner P. and Coolen F.P.A. (2009). Nonparametric predictive pairwise comparison for real-valued data with terminated tails. *International Journal of Approximate Reasoning*, 51, 141–150.

[53] Meeker W.Q. and Escobar L.A. (1998). *Statistical Methods for Reliability Data*. Wiley, New York.

[54] Navarro J., Samaniego F.J., Balakrishnan N. and Bhattacharya D. (2008). On the application and extension of system signatures in engineering reliability. *Naval Research Logistics*, 55, 313–327.

[55] Peressini A.F. (2016). Imprecise probability and chance. *Erkenntnis*, 81, 561–586.

[56] R Core Team (2013). R: A Language and Environment for Statistical Computing. *Vienna, Austria. URL http://www.R-project.org/.*

[57] Samaniego F.J. (1985). On closure of the IFR class under formation of coherent systems. *IEEE Transactions on Reliability*, 34, 69–72.

[58] Samaniego F.J. (2007). *System Signatures and Their Applications in Engineering Reliability*. Springer, Berlin, Germany.

[59] Shafer G. (1976). *A Mathematical Theory of Evidence*. Princeton University Press, New Jersey.

[60] Singer J.D. and Willett J.B. (1993). It's about time: Using discrete-time survival analysis to study duration and the timing of events. *Journal of Educational Statistics*, 18, 155–195.

[61] Walley P. (1991). *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London.

[62] Weichselberger K. (1995). Axiomatic foundations of the theory of interval-probability. In: V. Mammitzsch and H. Schneeweiss (eds.). *Proceedings of the Second Gauβ Symposion, Section B*, pp. 47–64.

[63] Weichselberger K. (2000). The theory of interval-probability as a unifying concept for uncertainty. *International Journal of Approximate Reasoning*, 24, 149–170.

[64] Weichselberger K. (2001). *Elementare Grundbegriffe einer Allgemeineren Wahrscheinlichkeitsrechnung I. Intervall-wahrscheinlichkeit als Umfassendes Konzept*. Physika, Heidelberg.

[65] Yan K.J. (2002). Nonparametric predictive inference with right-censored data. *PhD-Thesis, Durham University. URL https://maths.durham.ac.uk/stats/people/fc/thesis-KJY.pdf.*