Analysis of Recurrent Event Processes with Dynamic Models

by

© **Kunasekaran Nirmalkanna**

A thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Department of Mathematics and Statistics

Memorial University

May 2021

St. John's, Newfoundland and Labrador, Canada

Abstract

The analysis of past developments of processes through dynamic covariates is useful to understand the present and future of processes generating recurrent events. In this study, we consider two essential features of recurrent event processes through dynamic models. These features are related to monotonic trends and clustering of recurrent events, and frequently seen in medical studies. We discuss the estimation of these features through dynamic models for event counts. We also focus on the settings in which unexplained excess heterogeneity is present in the data. Furthermore, we show that the violation of the strong assumption of independent gap times may introduce substantial bias in the estimation of these features with models for event counts. To address these issues, we apply a copula-based estimation method for the gap time models. Our approach does not rely on the strong independent gap time assumption, and provides a valid estimation of model parameters. We illustrate the methods developed in this study with data on repeated asthma attacks in children. Finally, we propose some goodness-of-fit procedures as future research.

Lay summary

In many studies, subjects may experience an event of interest more than once over their observation periods. For example, individuals may experience repeated asthma attacks, older adults living in nursing homes, long-term care facilities may have repeat emergency department transfers or individuals may have more than one heart attack. Important characteristics of such recurrent events include the expected number of events over a time period and the rate of occurrence of events. These characteristics are defined as marginal characteristics in a sense that they ignore the effects of past event occurrences in a process. Therefore, models based on these characteristics are not adequate if there is an interest in understanding effects of the past of a process. In this study, we explore two frequently seen features of the recurrent event process. These features are related to trends and clustering of events over time. We consider modelling them under two approaches. In the first approach, we use the event intensity function to model event counts. In the second approach, our models based on the hazard functions for the times between subsequent event occurrences. Our study shows that these two approaches provide similar results only under specific settings, and we discuss the advantages and disadvantages of these approaches.

Acknowledgements

First, I would like to express my sincere gratitude to my advisor Dr. Candemir Cigsar for his guidance, supervision, patience, enthusiasm, motivation, positivity and immense knowledge during my Ph.D. program. He has provided constant support during my Ph.D. research and guided me throughout the preparation of this thesis. I am also very appreciative of the members of my Ph.D. supervisory committee's efforts and their feedback throughout my program's completion. Their valuable comments have helped to improve this thesis.

I wish to thank members of my thesis examining committee; Dr. Laurent Briollais, Dr. Zhaozhi Fan and Dr. J Concepcion Loredo-Osti for their insightful comments.

I would like to thank Dr. Yildiz Yilmaz, Dr. Alwell Julius Oyet and Dr. Asokan Mulayath Variyath, who have instructed me over the years and guided my education in Statistics. I sincerely acknowledge the financial support provided by the School of Graduate Studies and the Department of Mathematics & Statistics, Memorial University of Newfoundland in the forms of Graduate Fellowship, Teaching Assistants, and Graduate Assistants.

It is my great pleasure to thank my family and friends who encouraged and helped me during my Ph.D. program.

Statement of contribution

Dr. Candemir Cigsar proposed the research question that was investigated throughout this thesis. The overall study was jointly designed by Dr. Candemir Cigsar and Kunasekaran Nirmalkanna. The algorithms were implemented, the simulation studies were conducted and the manuscript was drafted by Kunasekaran Nirmalkanna. Dr. Candemir Cigsar supervised the study and contributed to the final manuscript.

Table of contents

List of tables

List of figures

Chapter 1

Introduction

This chapter consists of four sections. Section 1.1 includes a general introduction of the topics on which this thesis focuses. In Section 1.2, we provide a motivating example for the methods developed in this thesis. The data set considered in this section is analyzed in the remaining chapters of this thesis. We present a literature review in Section 1.3. In the last section, we outline the goals and scope of this thesis.

1.1 Introduction

Statistical methods and models for the analysis of complex event histories are needed in many fields of study including economics, insurance, public health, medicine, engineering, reliability and sociology (Cox and Lewis, 1966; Allison, 1984; Cook and Lawless, 2007; Aalen et al., 2008; Nelson, 2003). Throughout this thesis, we use the word "individual" as a generic term to define any object such as human, animal, electronic item and process unit, subject to experience certain events over its lifetime. The statistical analysis of event history data is usually based on multi state modeling (Cook and Lawless, 2018). Such an analysis allows us to understand and make inferences on the behaviour of processes, each staying and moving among a finite number of states over time or space. Survival analysis is a well-known example of event history analysis, where the processes or individuals are allowed to move from an alive state to a dead state (Aalen et al., 2008). Survival analysis contains statistical models and methods for analyzing data describing lifetimes, waiting times, or more generally times to the occurrence of a specified event, which can be death or failure of an individual in a study (Cox and Oakes, 1984; Lawless, 2003). These individuals could be humans or animals subjected to some treatment, as well as mechanical or electronic units. Such data, called survival data, can arise in various scientific fields including medicine, engineering, and demography.

Typically, the event of interest in an event history study is not experienced by a proportion of individuals in the sample during the follow-up period. The duration of event occurrence is said to be right-censored for those individuals, which leads to right-censored data. Such incomplete observations due to right-censoring must be addressed in the analysis of event history data. Also, there is usually an interest in the effects of explanatory variables on the timing of events or counts of events over a given time period. These variables may change their values throughout the follow-up of individuals. Such variables are called time-varying covariates. The usual statistical analysis of event history data are based on the assumption that the complete information on time-varying or time-fixed covariates are known. However, in applications, an event process can be influenced by its past developments. In literature, this phenomenon is referred to as the dynamic behaviour of an event process (see e.g. Aalen et al., 2008). As discussed throughout this thesis, it is a challenging but an important characteristic of modeling in the domain of the event history analysis. Another challenge with the analysis of event history data arises when individuals do not share the same baseline characteristics. In other words, they can be heterogeneous with respect to some characteristics. If there is excess heterogeneity present in the data, which cannot be explained with the available covariates, extended recurrent event models should be used to accommodate it. Such a situation makes inference in event history settings even more complicated when dynamic characteristics of processes are of interest because the dynamic characteristic may be confounded with the heterogeneity in some settings.

In this thesis, we focus on recurrent event processes, where a well defined event may occur more than once over a period of time. Event occurrence time can be considered as a time point, where the process moves from its current state to another state. Such time points are sometimes called time epochs (Feller, 1968). A recurrent process may depend on its history. Many models for recurrent events have been introduced and studied in the literature (e.g., Cox and Lewis, 1966; Cox and Isham, 1980; Daley and Vere-Jones, 2003, 2007). Cook and Lawless (2007) give a textbook length discussion of the statistical analysis of recurrent events. In some studies, there might be an interest only in the time to the first event occurrences in recurrent event processes. In this case, survival analysis techniques may be used to analyze time to first event occurrences. As discussed by Cox and Lewis (1966) and Cook and Lawless (2007), in many studies the purposes of analyzing recurrent event data include understanding and describing individual event processes, identifying and characterizing variations across a population of processes, comparing groups of processes, and estimating the effects of fixed and time-varying covariates and treatments on event occurrences.

Two fundamental features of recurrent event processes are time trends and clustering of events. There are various definitions of a time trend. For example, a time trend in a recurrent event process may be defined as a systematic change in the rate of event occurrences of a process over time. It can be either monotonic or non-monotonic. Monotonic trends are common in many applications and widely discussed in the literature. In recurrent event processes, trends may depend on calendar time and/or on the number of previous events. Modeling and the analysis of trends due to the number of previous events are crucial in many medical and reliability studies involving recurrent events. For example, inclusion of the number of previous events as a covariate in a process could be more informative than the calendar time (age) trend to predict an individual's future medical events, such as repeated hospitalizations, heart attacks and asthma attacks.

Clustering of events over time in a recurrent event process refers to a phenomenon when the number of event occurrences has sudden temporary hikes in short time intervals during the follow-up of process. In such settings, patterns of event clusters are observed for a given process. Carryover effects are defined as a particular type of event clustering feature, in which the occurrence of a condition on the process itself or an external condition causes a temporary increase in the number of events in a recurrent event process. The distinction between carryover effects due to external conditions and non-monotonic trends such as sinusoidal trends may become blurry in the sense that such trends may also cause clustering of events over time. However, patterns resulting from non-monotonic trends are usually more regular comparing with those resulting from external carryover effects.

In this thesis, we develop novel models and methods to simultaneously estimate carryover effects and monotonic trends in recurrent event processes. In the next section, we introduce data types and a data set used in this thesis to illustrate our models and methods.

1.2 Data Types and Motivating Examples

The data sets in recurrent event studies usually include times of occurrences of a well defined event or events in a continuum. The methods for the analysis of such data sets are based on either event counts or the elapsed times between successive events called gap times. Some studies include a single process or a relatively small number of processes, but processes included may generate a large number of events over their follow-up periods. The goal of such studies is usually to describe the patterns appearing in the data obtained from each process. The heterogeneity across processes is usually of little concern because each process generates data large enough to accommodate heterogeneity. Examples include data from stoppages in assembly lines, software fault detection and removal, and repeated incidences of injuries in manufacturing plants. On the other hand, in some studies, data are obtained from a large number of processes, each generating relatively small number of recurrent events. This type of processes commonly appears in medical or epidemiology studies, where a large number of individuals are usually included, and each may experience only a few clinical events repeatedly throughout their follow-up. The occurrence of asthma attacks in respiratory trials, epileptic seizures in neurology studies, and fractures in osteoporosis studies are notable examples of such cases. Throughout this thesis, we focus on studies in which a large number of individuals are included and each experiences a small number of events.

1.2.1 Simulated Data

For the sake of interpretation of the features of trends and clustering, we present a set of simulated data plotted in Figure 1.1 under different scenarios. The simulation procedure is explained in Section 2.5. The horizontal lines in the plots given in Figure 1.1 represent the individual processes and the cross symbols on the lines represent the event occurrences.

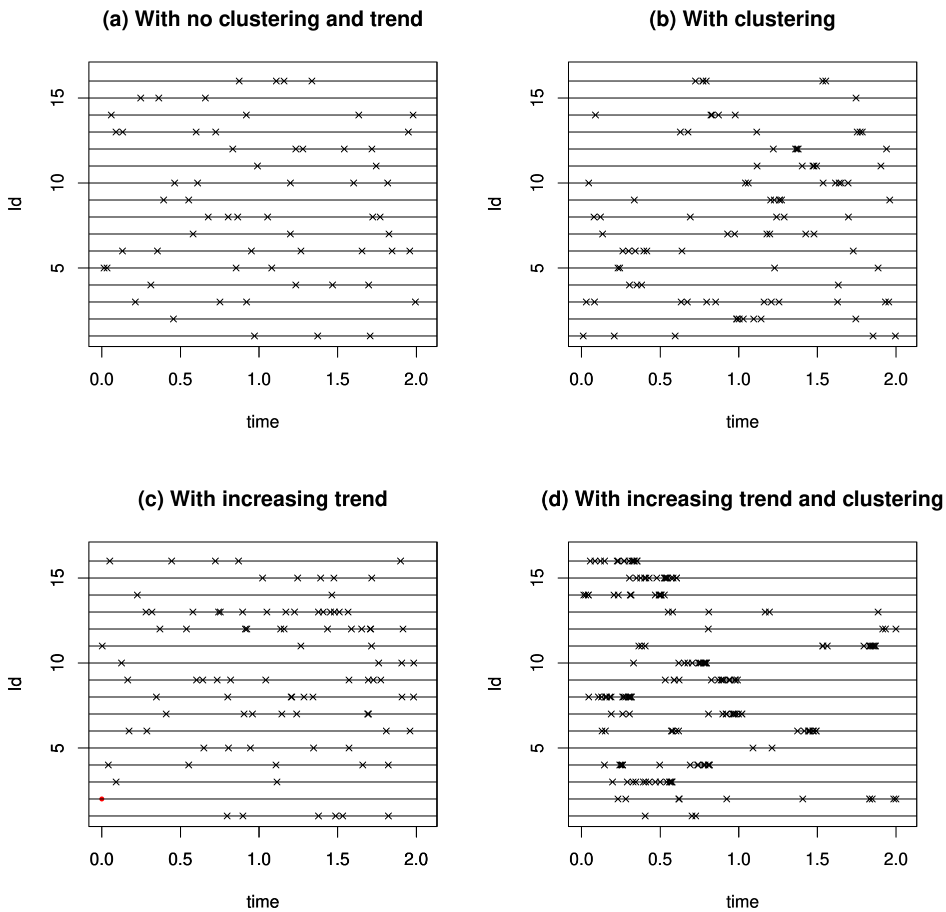

Figure 1.1: Dot plots of simulated data sets without any dynamic features (a), with clustering (b), with an increasing trend (c) and with both clustering and increasing trend (d).

The plot (a) in Figure 1.1 indicates that the repeated events do not show any clustering of events or any systematic pattern within the processes. Clustering of events can be observed in the plots (b) and (d) given in Figure 1.1 as many events occur soon after the previous one. Plots (c) and (d) in Figure 1.1 reveal that the gap times between successive events gradually decrease as time increases. A recurrent event model with a constant rate of events over time is suitable only for settings, in which the data do not indicates event clustering and trends. For other settings, more elaborate models are required.

We next introduce a data set from an asthma study in children. This data has been used in the thesis to illustrate the methods developed in the later chapters.

Table 1.1: Frequencies of observed and censored gap times in each control and treatment groups in asthma prevention trial data.

Gap-time		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Control	observed	119	82	62	45	37	28	22	18	11	11	8	7	5	3	2	1
	censored	0	37	19	17	8	8	6	4	6	0	3	1	2	1	0	0
Treatment	observed	113	63	38	27	20	17	15	12	9	8	3	2	2	1	1	1
	censored	0	50	25	8	6	2	2	3	2	1	4	1	0	1	0	0

1.2.2 Asthma Prevention Trial

Asthma is a chronic inflammatory disease of the airways. It causes consecutive episodes of wheezing, chest tightness or shortness of breath, commonly referred to as asthma attacks. The symptoms can be mild to severe and intermittent to chronic. Asthma is the most common chronic disease in children. Duchateau et al. (2003) present a dataset from an asthma prevention trial in infants. They considered a population with a high risk of asthma and focused on infants only. They sampled 232 infants with six months of age who had not yet experienced any asthma attacks. Each infant was allocated to either a placebo control group or an active drug treatment group and followed up for approximately 18 months. The treatment group included 113 infants, and the control group included 119 infants.

The primary purpose of the study was to assess the effect of a drug treatment on the occurrence of asthma attacks. Besides, the evolution of the asthma recurrent event rate over time, as well as how the presence of an event influences the event rate were also of interest. Since an asthma attack can be longer than one day and a patient is not considered at risk of having another asthma attack over that time period, the timescale of the study should be arranged accordingly.

In Table 1.1, the observed and censored number of gap times in each control and treatment group in asthma prevention trial data are given. There is one individual in the control group with 38 asthma attacks and one individual in the treatment group with 20 asthma attacks. For the illustrative purpose, we plotted subsets of asthma data for each groups in Figure 1.2.

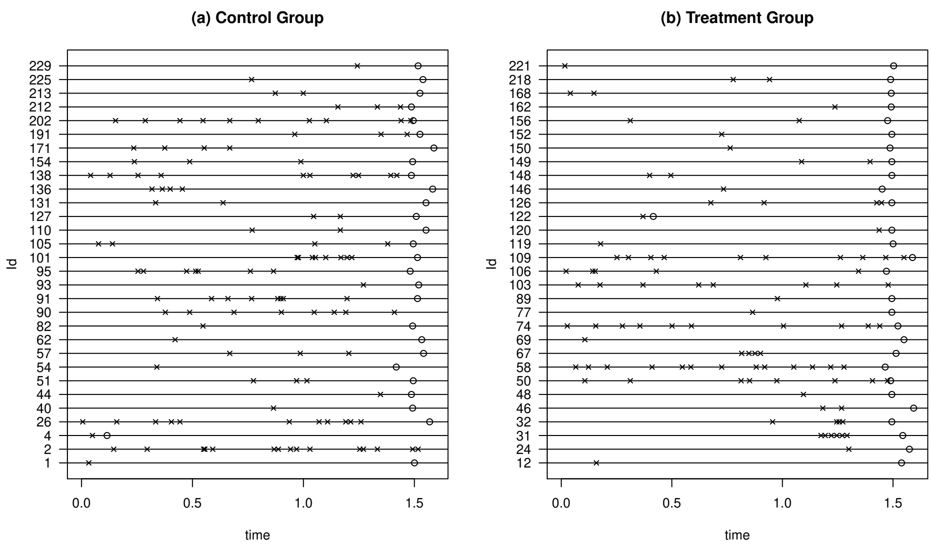

Figure 1.2: Dot plots of randomly selected subsets of asthma data for control group (a) and treatment group (b). The "×" symbol indicates the event time and the "o" symbol indicates the censoring time.

1.3 Literature Review

Statistical analysis of recurrent event data has started by analyzing data generated by single processes or populations. Early examples include occurrence of earthquakes, the emission of particles from a radioactive source, occurrence of accidents and cases of a disease in a human population. The event occurrences in these examples were considered as point processes. There are excellent books on this topic providing extensive probabilistic extensions and many examples of applications (see, for example, Cox and Lewis, 1966; Cox and Isham, 1980; Rigdon and Basu, 2000; Daley and Vere-Jones, 2003, 2007). Later, the modeling and analysis of recurrent events for various processes or systems have extensively developed (Andersen et al., 1993; Lawless, 1995; Cook and Lawless, 2007; Amorim and Cai, 2015). As discussed in Section 1.1, since unexplained heterogeneity across individual processes are common in studies, methods accommodating such heterogeneity have become important. Such developments were applied to medical data sets (Byar, 1980; Gail et al., 1980; Prentice et al., 1981), social science (Allison, 1984; Blossfeld and Rohwer, 2001), and product or equipment reliability (Nelson, 1988; Lawless et al., 1992).

Several regression models have been proposed to estimate the effects of covariates on the risk of recurrent events. The Cox proportional hazards model (Cox, 1972a) is the most popular approach for analyzing survival data, which estimates the effects of the factors that influence the rate of a particular event happening. It is a semiparametric multiplicative model, where a unit increment in a factor is multiplicative to the event rate. The Cox proportional hazards model has been extended in several ways to adapt the recurrent event settings (Andersen and Gill, 1982; Wei et al., 1989). The Cox-type proportional rate or intensity models that presume multiplicative covariate effects on the baseline rate or intensity function of recurrent event processes are also extensively studied in the literature (Prentice et al., 1981; Pepe and Cai, 1993; Lawless and Nadeau, 1995; Lin et al., 2000). Lin et al. (1998) studied an accelerated mean model, where covariates alter the timescale of the cumulative mean function. The accelerated rate or intensity model that formulates covariate effects to change the timescale directly on the baseline rate or intensity function was also studied by Chang and Wang (1999) and Ghosh (2004). Sun and Su (2008) proposed a general class of regression models for recurrent event processes that covers most of the above mentioned models.

Another way of analyzing recurrent events is through the gap times between successive events. Consider a recurrent event process where the process resets as in the initial stage of it immediately after the occurrence of each event and each gap time is identically distributed. Such processes are referred to as renewal processes. Methods for the analysis of renewal processes are mostly based on the methods of survival analysis, which were discussed broadly by Kalbfleisch and Prentice (2002) and Lawless (2003). Chang and Wang (1999) and Therneau and Grambsch (2000) implemented some models of gap time analyses using Cox models.

There has been a recent interest in the analysis of recurrent events through dynamic models (Peña and Hollander, 2004; Peña, 2006; Aalen et al., 2008; Cook and Lawless, 2013). Dynamic models typically include dynamic covariates, which are basically time-dependent internal covariates (Kalbfleisch and Prentice, 2002) such as the elapsed time since last event or the number of previous events in a process. Following the motivation of Aalen et al. (2008), we use the term *dynamic* to emphasize their role as explicitly picking up the past developments in processes.

The assumption of independent waiting (gap) times within individuals is a strong one, and should be carefully checked in all studies including recurrent events (Cook and Lawless, 2007). The models in Cigsar and Lawless (2012) and Cook and Lawless (2013) may result in biased estimates in some settings when the independent gap times assumption is violated. Cox and Lewis (1966) discussed methods for checking independence when no covariates are present. Another approach to deal with dependent gap times is to apply a random effects model (Cook and Lawless, 2007; Golzy and Carter, 2019). However, it does not give plausible results with dynamic covariates because the random effects may confound with some dynamic features in the recurrent event processes. We, therefore, apply a copula based estimation method first proposed by Lawless and Yilmaz (2011) which incorporates the dependency among gap times within individuals. For more than two gap times, more complicated copula parameterizations were introduced by Barthel et al. (2018). However, their modelling approach may result in a large number of nuisance parameters.

1.4 The Goal and Summary of the Thesis

In this thesis, we explore some features of recurrent event processes through dynamic models. Dynamic models are needed in understanding the recurrent event processes when the effect of past developments on the present or future evolution of a process is of interest. The inherent nature of dynamic covariates and their complex relations, dependent gap times, censoring and heterogeneity make modeling and inference challenging issues when dynamic covariates are present.

In this study, we discuss the assessment of two important features of recurrent event processes through dynamic intensity models. These features are monotonic trends and carryover effects. Trend is a frequently seen feature of recurrent event processes. There are various definitions of a trend. Broadly speaking, it can be defined as a systematic variation in the intensity function of a recurrent event process. In this thesis, we focus on monotonic trends due to the number of previous events in a process. Modelling such type of trends has mostly been overlooked in the literature. However, they may summarize valuable prognostic information about event processes, especially when limited external explanatory information about systematic variations in the intensity function is available. Carryover effects are defined as transient effects caused by previous event occurrences or some external conditions experienced by event processes. Their presence results in the clustering of events in time or sparsity of events.

In this study, our primary goal is to discuss simultaneous estimation of these features with dynamic recurrent event models. We, therefore, consider two modelling approaches. In the first approach, we carry out the maximum likelihood estimation method through a parametric class of multiplicative intensity models for event counts. In the second approach, we consider a gap time modelling approach under the maximum likelihood estimation method. In this approach, we address the serial dependency between gap times with copula models so that our approach allows us to obtain a more detailed understanding of the aforementioned dynamic features. We also hypothesize that these two approaches do not provide the same conclusions in most applied settings. Heterogeneity and correlation within gap times among individuals may affect the estimation of dynamic features in recurrent event data. Our novel approach can handle these issues simultaneously. We believe that outcome of this thesis will be instrumental in the analysis of the dynamic features of recurrent events so that a deep understanding of the event generating mechanisms about recurrent event processes can be obtained.

The remaining chapters of this thesis can be summarized as follows. We present a brief theoretical background in Chapter 2. In Chapter 3 and 4, we consider models for transient carryover effects and number of previous events, discuss the possible extensions of the model to deal with unexplained heterogeneity. To this end, we consider count-based models and gap times based models in Chapters 3 and 4, respectively. We discuss the issue and the solution when the effect of the number of previous events is included in the model and verify the asymptotic properties of the estimates of the model parameters. In Chapter 4, we discuss the issues with count-based models when the serial gap time dependency is present, and show why we propose a copula based models, and present simulation results on their properties, and examine the studies on asthma in infants. In many applications, the number of events is very small per individual, especially in medical data. Therefore, in Chapter 4, we first focus only on the first two gap times, and in later part of this chapter, we consider more than two gap times with more complicated dependent structures. Some goodness-of-fit procedures for checking the model adequacy for the models introduced in Chapter 4 are introduced in Section 5.2.4 as a future research. Chapter 5 contains concluding remarks and other future extensions.

Chapter 2

Theoretical Background

In this chapter, we introduce the notation frequently used in this thesis and provide a background on technical concepts essential in the remaining part of the thesis. In Section 2.1, we present the terminology and notation. Some basic statistical approaches to analyze the recurrent event data from a single or multiple identical processes are given in Section 2.2. The methods to deal with the non-identical processes are given in Section 2.3. We next introduce copula models to handle the dependency among random variables in Section 2.4. The simulation algorithms used to generate independent and dependent data are provided in Section 2.5.

2.1 Terminology and Notation

In this thesis, we are interested in a specific type of random phenomenon, which produces random occurrences of a well defined event as points over a time axis. In this context, a *point process* is defined as a collection of points randomly located on an underlying mathematical space such as the real line, the Cartesian plane, or on a more abstract space (Cox and Isham, 1980). We, therefore, adopt models and methods from the theory of point processes. Points represent times of events. Other than their locations, they are indistinguishable. In this section, we introduce a standard notation frequently used in the point processes framework (see e.g. Cox and Isham, 1980; Cook and Lawless, 2007).

Let $T_1, T_2, \ldots$, where $0 < T_1 < T_2 < \cdots$, represent times of occurrence of an event in a point process. The random variable T_k, $k = 1, 2, \ldots$, is called the k^{th} *event time*, and its observed value is denoted by t_k, $k = 1, 2, \ldots$. We next define the *waiting time* or *gap time* of a point process. Let the random variables W_1, W_2, $\ldots$ be the elapsed times between successive event occurrences in a point process with the convention that $W_1 = T_1$. The variable W_k, where $W_k = T_k - T_{k-1}$, $k = 1, 2, \ldots$, is then called the k^{th} waiting time or gap time and, by convention, $T_0 = 0$. The number of events occurring over the time period $(s, t]$, for $s < t$, is denoted by $N(s, t)$. When the starting point of observation is zero, i.e. $s = 0$, the number of events occurring over $(0, t]$ can be written as $N(t)$ instead of writing as $N(0, t)$. Let $\{N(t); t \geq 0\}$ denotes the corresponding *counting process* with the condition that $N(0) = 0$. An extensive discussion on the construction of a point process on the real line through the counting measure can be found in Daley and Vere-Jones (2003, Section 3.1). Throughout this thesis, we assume that the counting process $\{N(t); t \geq 0\}$ has the following property:

$$\Pr\left\{N(t) - N(t^-) > 1\right\} = 0, \tag{2.1}$$

where $N(t^-)$ is the number of events occurring over the time period $(0, t)$. That is, at any time point t, at most one event can occur. The property (2.1) is referred to as *orderliness* property of a counting process (Cox and Isham, 1980, Section 2.3).

The mean and rate functions are two important marginal characteristics of a counting process. The *mean function* of a counting process is defined as $\mu(s, t) = E\{N(s, t)\}$, where $0 \leq s < t$. That is, the mean function $\mu(s, t)$ gives the expected cumulative number of events occurring over the interval $(s, t]$ for any $s < t$. For convenience, we define the notation $\mu(t)$ as the expected cumulative number of events occurring over the interval $(0, t]$; that is, $\mu(t) = E\{N(t)\}$. Let $\delta N(t)$ denote the number of events in a short interval $[t, t+\delta t)$, where δt denotes a small time increment. In notation, $\delta N(t) = N((t + \delta t)^-) - N(t^-)$. The *rate function* of a counting process is then defined by

$$\rho(t) = \lim_{\delta t \to 0} \frac{\Pr\{\delta N(t) = 1\}}{\delta t}. \tag{2.2}$$

Specifically, for a sufficiently small δt, $\rho(t)\,\delta t \approx E\{\delta N(t)\}$. Since the mean function $\mu(t)$ denotes the expected cumulative number of events in $(0, t]$, we have the relation that $\mu(t) = E\{N(t)\} = \int_0^t \rho(s)ds$ and $\rho(t) = d\mu(t)/dt$, assuming the derivative exists.

Another crucial concept in modeling and analysis of recurrent event processes is the intensity function. The *intensity function* of a counting process $\{N(t), t > 0\}$ is mathematically defined by

$$\lambda[t|\mathcal{H}(t)] = \lim_{\delta t \to 0} \frac{\Pr\{\delta N(t) = 1|\mathcal{H}(t)\}}{\delta t}, \qquad t > 0, \tag{2.3}$$

where the *process history* $\mathcal{H}(t) = \{N(s); 0 \le s < t\}$ contains the information of past events in $[0, t)$. The intensity function of a counting process in continuous time completely specifies an event process (Cook and Lawless, 2007). From now on, we assume that all counting processes are in continuous time. We note that, in the limit as δt approaches zero, $\delta N(t)$ is a 0-1 valued binary random variable. Thus, $E\{\delta N(t)|\mathcal{H}(t)\}$ can be approximated by $\lambda[t|\mathcal{H}(t)]\delta t$ for a small δt.

For an orderly counting process $\{N(t); t \ge 0\}$, the probability that "n events occur in $[t, t+\delta t)$" can be written in terms of the intensity function defined in (2.3) as follows. Let $\{N(t), t > 0\}$ be a counting process and $\lambda[t|\mathcal{H}(t)]$ be its intensity function. Then, for $n = 0, 1, 2, \ldots,$

$$\Pr\{\delta N(t) = n|\mathcal{H}(t)\} = \begin{cases} 1 - \lambda[t|\mathcal{H}(t)]\delta t + o(\delta t), & if \;\; n = 0, \\ \lambda[t|\mathcal{H}(t)]\delta t + o(\delta t), & if \;\; n = 1, \\ o(\delta t), & otherwise, \end{cases} \tag{2.4}$$

where the notation o denotes the order of magnitude in a sense that, $g(t) = o(t)$ means that $g(t)/t \to 0$ as $t \to 0$. The likelihood function of the outcome "exactly n events occur at times $t_1 < t_2 < \cdots < t_n$ over the observation interval $[\tau_0, \tau]$", conditional on the history at time τ_0 can be derived by using the result (2.4), which gives

$$\left(\prod_{j=1}^{n} \lambda[t_j|\mathcal{H}(t_j)]\right) \exp\left\{-\int_{\tau_0}^{\tau} \lambda[u|\mathcal{H}(u)]du\right\}. \tag{2.5}$$

The derivation of the likelihood function (2.5) can be found in Andersen et al. (1993, Section II.7) and Cook and Lawless (2007, Section 2.1).

We can also show from the result (2.4) that the probability of outcome "no event occurs in $(s, t]$" conditional on the history $\mathcal{H}(s^+)$ is given by

$$\exp\left\{-\int_{s}^{t} \lambda[u|\mathcal{H}(u)]du\right\}, \tag{2.6}$$

where $\mathcal{H}(u) = \{\mathcal{H}(s^+), N(s, u) = 0\}$ and $\mathcal{H}(s^+)$ is the history of the process up to and including time s. It can be derived from (2.6) that the conditional probability of the j^{th} gap time is greater than w, given the value t_{j-1} and the history up to t_{j-1}, as

$$\Pr\left\{W_j > w | T_{j-1} = t_{j-1}, \mathcal{H}(t_{j-1})\right\} = \exp\left\{-\int_{t_{j-1}}^{t_{j-1}+w} \lambda[u|\mathcal{H}(u)]du\right\}. \tag{2.7}$$

We use the result (2.7) to generate realizations of a recurrent event process for a given intensity function as explained in Section 2.5.

Another important function is the *hazard function* of a gap time W, which is defined as

$$h(w) = \lim_{\delta w \to 0} \frac{\Pr\left\{w \leq W < w + \delta w | W \geq w\right\}}{\delta w}, \quad w \geq 0. \tag{2.8}$$

Let $F(w)$ be the cumulative distribution function (c.d.f), and $f(w) = dF(w)/dw$ be the probability density function (p.d.f) of the gap time W. Then, it can be shown that $h(w) = f(w)/S(w), w \geq 0$, where $S(w) = 1 - F(w)$ is the *survival function* of the gap time W.

An important relationship between the counting process $\{N(t); t \geq 0\}$ and the event time $T_i, i = 1, 2 \ldots,$ is that

$$N(t) \geq j \Leftrightarrow T_j \leq t, \qquad t > 0, j = 1, 2, \ldots. \tag{2.9}$$

Consequently, the probability of the number of events by time t is greater than or equal to j is the same as the probability of j^{th} event occurs before or at time t; that is,

$$\Pr\left\{N(t) \geq j\right\} = \Pr\left\{T_j \leq t\right\}, \qquad t > 0, j = 1, 2, \ldots. \tag{2.10}$$

2.2 Recurrent Event Processes

Poisson and renewal processes are considered as fundamental stochastic processes that provide models for recurrent events (Cook and Lawless, 2007). In this section, we introduce them and discuss some of their features. Event counts over specified time intervals are often used to analyze data from recurrent events. Poisson processes provide natural models for describing event counts in many settings. Models based on gap times provide another class of models for the analysis of recurrent events. These models are especially useful when the prediction of the next event time or the effects of interventions after the occurrence of an event are of interest in a study. Renewal processes and their ramifications provide mathematical models that are widely used to model gap times. Before introducing the Poisson and renewal processes and other related stochastic processes, we next discuss two important features of recurrent event processes.

2.2.1 Some Fundamental Features of Recurrent Event Processes

An important feature of recurrent event processes is the presence or absence of time trends. Providing a general definition of the trend is not an easy task in recurrent event analysis. This issue is discussed by Ascher and Feingold (1984, Section 9) in recurrent events settings and by White and Granger (2011) in time series analysis settings. Simple plots can be useful to reveal possible trends in data in some settings, but due to the elusive nature of a trend, testing statistical hypothesis of the presence or absence of trends, in general, is challenging.

Monotonic (increasing or decreasing) trends are common in applications. These trends are usually related to stochastic ageing (Lawless et al., 2012). There are also non-monotonic trends, such as seasonal trends. Trends may also occur due to factors related to the number of previous events $N(t^-)$ prior to t in a process (Kvist et al., 2008). Aalen et al. (2008) suggest the use of $N(t^-)$ as a covariate in the model to check the presence or absence of frailty. A significant $N(t^-)$ in the model may represent a monotonic trend in the waiting times between event occurrences. We use and discuss such definition in Chapter 3 and 4 in this thesis.

Event clustering is another common feature of recurrent event processes. In particular, events from a single recurrent event process may cluster together in time. Carryover effects arise in settings where the intensity of an event is temporarily increased or decreased following the occurrence of an external condition or an event (Lindsey, 2004; Whitaker et al., 2006; Cigsar and Lawless, 2012). The presence of carryover effects in processes may result in some forms of clustering of events. Both internal and external conditions may be associated with carryover effects. We include these conditions as covariates in the recurrent event models, and investigate their effects on the intensity functions. We focus on internal covariates, which is mathematically a more challenging situation. An application of carryover effects due to external covariates in vaccination studies is given by Whitaker et al. (2006). Simple plots can be used to perceive clustering and trends in studies, especially, when each process in the study provides a large number of events over its follow-up. However, when there is a large number of processes under observation, and the number of events per process is small, the detection of trends and carryover effects can be problematic. It should be noted that this situation is often the case in many medical and epidemiological studies, and was the motivation behind our study. In the thesis, we develop models to assess the presence or absence of time trends and carryover effects as internal factors to the processes under observation.

2.2.2 Poisson Processes

The Poisson processes can be characterized in various mathematically equivalent ways. Characterization and many properties of Poisson processes can be found in Daley and Vere-Jones (2003). Our goal in this section is to present some results that are used in the remaining parts of the thesis. Any counting process $\{N(t); t \geq 0\}$ possessing the following properties is a Poisson process with the rate function $\rho(t)$.

- *Property 1:* The number of events occurred at the initiation time is zero; that is, $N(0) = 0$.

- *Property 2:* For any time points a, b, c and d such that $0 \leq a < b \leq c < d$, the random variables $N(a, b)$ and $N(c, d)$ are independent.

- *Property 3:* For any time points s and t such that $0 \leq s < t$, the random variable $N(s, t)$ has a Poisson distribution with mean $\mu(s, t) = \mu(t) - \mu(s)$, where $\mu(t) = \int_0^t \rho(u)du$.

A stochastic process is called *Markov process* if, at any time $t > 0$, the conditional probability of an arbitrary future event given the entire past of the process equals the conditional probability of that future event given only the value of the process at time t. In other words, conditional on the present value of the process, the distribution of the increments of a Markov process does not depend on the past of the process. Such a process is said to have the *Markov property*. Poisson processes also have Markov property due to the independent increments in non overlapping intervals, which is the Property 2 given above.

The consequence of the third property is that, if $\{N(t); t \geq 0\}$ is a Poisson process with the corresponding mean function $\mu(t)$, then

$$\Pr\{N(s,t) = n\} = \frac{[\mu(s,t)]^n}{n!} \exp\{-\mu(s,t)\}, \quad n = 0, 1, 2, \ldots, \tag{2.11}$$

for any $0 \leq s < t$. A proof of the claim that the above properties characterize that the counting process $\{N(t); t \geq 0\}$ is a Poisson process with the rate function $\rho(t)$, $t \geq 0$, can be found in Cook and Lawless (2007, Section 2.2.1). For general Poisson processes, the marginal survival function of n^{th} event occurrence time is

$$\begin{aligned} \Pr(T_n > t) &= \Pr(N(t) < n), \\ &= \sum_{k=0}^{n-1} \frac{\mu(t)^k \exp\{-\mu(t)\}}{k!}, \\ &= \int_{\mu(t)}^{\infty} \frac{x^{n-1} \exp\{-x\}}{(n-1)!}\, dx, \qquad \text{for } n \geq 1. \end{aligned} \tag{2.12}$$

The last equality can be proven from the recurrence relationship of incomplete gamma function $\Gamma(n, x)$; that is,

$$\Gamma(n, x) = (n-1)\Gamma(n-1, x) + x^{n-1} \exp\{-x\}, \tag{2.13}$$

where $\Gamma(n, x) = \int_x^{\infty} t^{n-1} \exp\{-t\}\, dt$ is the incomplete gamma function (Abramowitz and Stegun, 1948, Section 6.5). For a counting process with an unbounded mean function $\mu(t)$, $t \geq 0$, the expected value of the n^{th} occurrence time T_n can be written as

$$E(T_n) = \int_0^{\infty} \mu^{-1}(x) \frac{x^{n-1} \exp\{-x\}}{(n-1)!} dx, \qquad \text{for } n \geq 1, \tag{2.14}$$

where $\mu^{-1}(t)$ is the inverse function of $\mu(t)$.

Suppose that $\{N(t); t \geq 0\}$ is a Poisson process with the mean function $\mu(t)$, $t \geq 0$. Let $t_1, t_2, \ldots, t_n$ be ordered arbitrary times, where $t_0 = 0 < t_1 - \delta_1 < t_1 < t_2 - \delta_2 < t_2 < \ldots < t_n - \delta_n < t_n$. Also, let $\Pr(t_1 - \delta_1 < T_1 \leq t_1, \ldots, t_n - \delta_n < T_n \leq t_n)$ and $\Pr\{N(t_0, t_1 - \delta_1) = 0, N(t_1 - \delta_1, t_1) = 1, \ldots, N(t_{n-1}, t_n - \delta_n) = 0, N(t_n - \delta_n, t_n) = 1\}$

are denoted by $\Pr(t_i - \delta_i < T_i \le t_i; i = 1, \ldots, n)$ and $\Pr\{N(t_{i-1}, t_i - \delta_i) = 0,$ $N(t_i - \delta_i, t_i) = 1; \quad i = 1, \ldots, n\}$, respectively. Then, the joint probability of the recurrent event times T_i observed within the interval $(t_i - \delta_i, t_i]$ for small values of the δ_i, for $i = 1, \ldots, n$, is

$$\begin{aligned} \Pr(t_i - \delta_i < T_i \le t_i; i = 1, \ldots, n) &= \Pr\{N(t_{i-1}, t_i - \delta_i) = 0, N(t_i - \delta_i, t_i) = 1; \\ &\qquad\qquad i = 1, \ldots, n\}, \\ &= \left\{\prod_{i=1}^{n} [\mu(t_i) - \mu(t_i - \delta_i)]\right\} \exp\{-\mu(t_n)\}, \end{aligned} \tag{2.15}$$

where the last equality follows from the properties of Poisson processes. Hence, the joint density function of the recurrent event times $T_1, T_2, \ldots, T_n$ of the Poisson process $\{N(t); t \ge 0\}$ with the rate function $\rho(t)$, $t \ge 0$, is

$$\lim_{\max(\delta_i) \to 0} \frac{\Pr(t_i - \delta_i < T_i \le t_i; i = 1, \ldots, n)}{\prod_{i=1}^{n} \delta_i} = \left\{\prod_{i=1}^{n} \rho(t_i)\right\} \exp\{-\mu(t_n)\}, \tag{2.16}$$

where $0 < t_1 < \cdots < t_n$ and the limit is taken as δ_i approaches zero for all $i = 1, \ldots, n$. Consequently, the joint density function of the gap times $W_1, W_2, \ldots, W_n$ is given by

$$f_{1:n}(w_1, \ldots, w_n) = \prod_{i=1}^{n} \rho w_{\cdot i} \exp\{-\mu w_{\cdot n}\}, \quad w_i > 0, i = 1, \ldots, n, \tag{2.17}$$

where $w_{\cdot i} = \sum_{j=1}^{i} w_j$ and $w_{\cdot n} = \sum_{j=1}^{n} w_j$. From (2.17) the conditional density function of W_{n+1} given the previous gap times is

$$f_{n+1|1:n}(w|w_1, \ldots, w_n) = \rho(t_n + w) \exp\{-[\mu(t_n + w) - \mu(t_n)]\}, \quad w > 0, \tag{2.18}$$

where $t_n = w_1 + \cdots + w_n$ and $n = 0, 1, 2, \ldots$. The conditional survival function of W_{n+1} given the previous gap times is

$$S_{n+1|1:n}(w|w_1, \ldots, w_n) = \exp\{\mu(t_n) - \mu(t_n + w)\}, \quad w > 0, \tag{2.19}$$

where $n = 1, 2, \ldots$. The marginal survival function of W_{n+1}; that is, $S_{W_{n+1}}(w) =$

$\Pr(W_{n+1} > w)$, is given by

$$\begin{aligned} S_{n+1}(w) &= \int_0^\infty \Pr(W_{n+1} > w | T_n = t) \Pr(N(t^-) = n-1) \rho(t) dt, \\ &= \int_0^\infty \frac{\mu(t)^{n-1} \exp\{-\mu(t+w)\}}{(n-1)!} \rho(t) dt, \quad w > 0, \end{aligned} \tag{2.20}$$

where $n = 0, 1, 2, \ldots$, and $\exp\{-\mu(w)\}$ for $n = 0$.

Poisson processes can also be characterized through their intensity functions. A counting process $\{N(t); t \geq 0\}$ is said to be a Poisson process if its intensity function is of the form

$$\lambda[t|\mathcal{H}(t)] = \rho(t), \qquad t \geq 0, \tag{2.21}$$

where $\rho(t)$ is a positive valued rate function on $[0, \infty)$ (Cook and Lawless, 2007, Section 2.2.1). The model in (2.21) implies that the Poisson process $\{N(t); t \geq 0\}$ has the Markov property because the event occurrence rate $\rho(t)$ does not depend on the history $\mathcal{H}(t)$. A Poisson process is called a *homogeneous Poisson process* when $\lambda[t|\mathcal{H}(t)]$ given in (2.22) is constant over time; otherwise, it is called a *non-homogeneous Poisson process*. It can be shown from the marginal survival function given in (2.20) that if a Poisson process is a homogeneous Poisson process with rate function α then the gap times have an exponential distribution with mean $1/\alpha$, $\alpha > 0$.

The Poisson process has some limitations to apply in some settings. In such cases, more general intensity-based models can be used. An important class of intensity models is of multiplicative form in which, for example, the intensity function (2.3) takes the form

$$\lambda[t|\mathcal{H}(t)] = \lambda_0(t) \exp[\boldsymbol{\psi}' \boldsymbol{W}^*(t)], \qquad t > 0, \tag{2.22}$$

where λ_0 is an age-specific baseline intensity function which can be specified parametrically or non-parametrically, $\boldsymbol{W}^*(t)$ is a $q \times 1$ vector of processes that is allowed to contain functions of the event history $\mathcal{H}(t)$ as well as external covariates, and $\boldsymbol{\psi}$ is a $q \times 1$ vector of parameters. Let $\mathcal{W}(t) = \{\boldsymbol{W}^*(s); 0 \leq s \leq t\}$ denote the history of covariates of interest to the time t. We use the notation $\boldsymbol{\mathcal{W}}(\infty)$ to denote the complete path information of $\mathcal{W}(t)$. For a given counting process $\{N(t); t \geq 0\}$, we let $\mathcal{H}(t) = \{N(s), \boldsymbol{W}^*(u); 0 \leq s < t, 0 \leq u \leq t\}$ and $\mathcal{H}(\infty) = \{N(s), \boldsymbol{W}^*(u); 0 \leq s < t, u \geq 0\}$. Now suppose that the intensity function of the counting process $\{N(t); t \geq 0\}$ is defined by $\lambda[t|\mathcal{H}(t)] = \lambda_0(t) g\left[\boldsymbol{W}^*(t); \boldsymbol{\psi}\right]$, $t > 0$, where g is a non-negative function. The inclusion of covariates requires $\lambda[t|\mathcal{H}(\infty)] = \lambda[t|\mathcal{H}(t)]$ for all $t > 0$. This condition merely states that the vector $\boldsymbol{W}^*(t)$ includes functions of events occurred prior to t and external covariates in the sense defined by Kalbfleisch and Prentice (2002, Section 6.3). More technical discussion on how to include covariate in recurrent event process are discussed by Cook and Lawless (2007, Section 2.2.2) and Aalen et al. (2008, Section 2.2.7). It should be noted that if there is one or more time-varying external covariates included in $\boldsymbol{W}^*(t)$, the process $\{N(t); t \geq 0\}$ is a non-homogeneous Poisson process. When internal covariates are also included in $\boldsymbol{W}^*(t)$, the corresponding process is then called a *modulated Poisson process* (Cox, 1972b). With the non-parametric specification of the baseline intensity function $\lambda_0(t)$, the model (2.22) is called a *semi-parametric* model.

2.2.3 Renewal Processes

Characterization of a renewal process is usually based on the gap times. A stochastic process with independent and identically distributed (i.i.d.) waiting times $W_1, W_2, \ldots$ is called a *renewal process*. The intensity function of a renewal process is given by

$$\lambda[t|\mathcal{H}(t)] = h(B(t)), \qquad t \geq 0, \tag{2.23}$$

where the hazard function h is defined in (2.8) and $B(t) = t - T_{N(t^-)}$ is called *the backward recurrence time*, which is the elapsed time since the last event time strictly before time t. Modeling and analysis of recurrent events can also be based on gap times in this setting. The hazard function of the k^{th} gap time is given by

$$h_k(w) = \frac{f_k(w)}{S_k(w)}, \qquad w > 0,\ k = 1, 2, \ldots, \tag{2.24}$$

where $S_k(w) = \Pr(W_k \geq w)$ and $f_k(w)$ are the survival function and p.d.f. for the k^{th} gap time W_k, respectively. The hazard function $h_k(w)$ can also be considered as the probability of the occurrence of the k^{th} event in an infinitesimally small time period $[t_{k-1} + w, t_{k-1} + w + dw)$, given that the individual is under observation until time $t_{k-1} + w,\ k = 1, 2, \ldots$, where dw denotes infinitesimal time increment.

In a reliability setting, the renewal processes are called *perfect repair* models in which after every repair, the system is stochastically considered as a new system. It should be noted that when the W_k are i.i.d. exponential random variables with mean $1/\alpha$, where $\alpha > 0$, the renewal process is equivalent to a homogeneous Poisson process with rate function α.

Fixed covariates or external time-varying covariates can also be included in a renewal process model. Let $\boldsymbol{x}$ and $\boldsymbol{x}(t)$ be $q \times 1$ vector of fixed and time varying covariates and $\boldsymbol{\psi}$ be the $q \times 1$ vector parameters for the corresponding vector of covariates. Survival regression models such as the proportional hazards model and accelerated failure time model are applicable for modeling the gap times in renewal processes when fixed covariates are present. The resulting conditional hazard functions of the k^{th} gap time, given the covariates $\boldsymbol{x}$, is of the form $h_k(w|\boldsymbol{x}) = h_{0k}(w)\exp(\boldsymbol{x'\psi})$ and $h_k(w|\boldsymbol{x}) = h_{0k}[w\exp(\boldsymbol{x'\psi})]\exp(\boldsymbol{x'\psi})$ for proportional hazards model and accelerated failure time model, respectively. When the covariates $\boldsymbol{x}$ contain both external covariates and functions of t or $\mathcal{H}(t)$, the models are called *modulated renewal processes* (Cox, 1972a; Cook and Lawless, 2007).

Since the density function of the k^{th} gap time W_k, denoted by $f_k(w|\boldsymbol{x})$, is the first partial derivative of $-S_k(w|\boldsymbol{x})$ with respective to w, $h_k(w|\boldsymbol{x})$ can be written as

$$h_k(w|\boldsymbol{x}) = \frac{-S_k'(w|\boldsymbol{x})}{S_k(w|\boldsymbol{x})} = -\frac{\partial}{\partial w}\log(S_k(w|\boldsymbol{x})), \quad k = 1, 2, \dots\ , \tag{2.25}$$

where $S_k'(w|\boldsymbol{x}) = \frac{\partial}{\partial w}S_k(w|\boldsymbol{x})$. By integrating both sides of (2.25), we can obtain the following relationship;

$$S_k(w|\boldsymbol{x}) = \exp(-\Lambda_k(w|\boldsymbol{x})), \quad k = 1, 2, \dots, \tag{2.26}$$

where $\Lambda_k(w|\boldsymbol{x}) = \int_0^w h_k(u|\boldsymbol{x})du$, which is called as the *cumulative hazard function* of the k^{th} gap time W_k. From (2.25) and (2.26), it is easy to see that the density function of the k^{th} gap time, conditional on the covariates $\boldsymbol{x}$, is given by

$$f_k(w|\boldsymbol{x}) = h_k(w|\boldsymbol{x})\exp(-\Lambda_k(w|\boldsymbol{x})), \quad k = 1, 2, \dots\ . \tag{2.27}$$

The result in (2.27) is useful to build a relation between the p.d.f. and the cumulative hazard function.

2.2.4 Other Stochastic Processes

Poisson and renewal processes establish the foundation of modeling and methods to analyze recurrent event data. Although they are very useful in many studies, they may not be adequate in some settings. In this section, we introduce some other modeling approaches for recurrent event data.

2.2.4.1 Delayed Renewal Process

An important extension of the renewal processes is the class of delayed renewal processes. The counting process $\{N(t); t \geq 0\}$ is called a *delayed renewal process* if the gap times $W_1, W_2, \ldots,$ are independent and the first gap time W_1 has a different distribution from the identically distributed remaining gap times $W_2, W_3, \ldots,$ (Feller, 1982; Ross, 1996). Delayed renewal process models are adequate to fit many medical data; for example, repeated heart attacks or repeated asthma attacks. In such data, defining the starting point of the recurrent event process could be complicated and delayed renewal processes provide flexibility in modeling such data. Also, in some studies, it is naturally inappropriate to assume that the first gap time between the starting point of the study and the initial event and the gap times between succeeding events are identically distributed (e.g. see Cigsar and Lawless, 2012). Some studies focus only on repeated events rather than on the whole process (e.g. see Gruneir et al., 2018). In such cases, the starting point of the process can be defined as the first event time and apply the renewal process model for the following subsequent gap times.

2.2.4.2 Self Exciting Process

Self-exciting process is first introduced by Hawkes (1971) and it is sometimes referred to as Hawkes Process. The complete intensity function of the self-exciting process takes the linear form

$$\lambda[t|\mathcal{H}(t)] = \alpha + \int_0^t \mathcal{K}(t-u)dN(u), \tag{2.28}$$

where $\alpha \geq 0$, $\mathcal{K}(u) \geq 0$ and $\int_0^\infty \mathcal{K}(u)du < 1$. A special case of particular interest is when

$$\mathcal{K}(u) = \gamma e^{-\beta u}, \tag{2.29}$$

where $\gamma \geq 0$ and $\beta \geq 0$. In this case, the intensity function makes a jump after every event time and decays exponentially in time until the occurrence of next event. This type of models are frequently used in financial analysis (Aït-Sahalia et al., 2015), earthquake studies (Ogata, 1988) and immigration-birth studies (Oakes, 1975). The model in (2.28) is in additive form because the effect of contributions of past events on the baseline rate function α is included in the model as an additive term. A Cox type multiplicative form of extension of (2.28) can also be used to model medical data (see e.g. Chen and Chen, 2014; Kim et al., 2019). The extension has the form

$$\lambda[t|\mathcal{H}(t)] = \lambda_0[t; \boldsymbol{\eta}] \exp\left[\boldsymbol{\psi}'\boldsymbol{x} + \int_0^t \mathcal{K}(t-u)dN(u)\right]. \tag{2.30}$$

With the specification (2.29), the semiparametric version of the model (2.30) assumes no self-exciting effects from previous events, and becomes the Andersen–Gill semiparametric model (Andersen and Gill, 1982) with time-independent covariates $\boldsymbol{x}$ when $\gamma = 0$ and $\beta \geq 0$. When $\gamma \neq 0$ and $\beta > 0$, the model (2.30) indicates that more recent events have stronger effects than more distant events, whereas $\gamma \neq 0$ and $\beta = 0$ result in again a regular Cox-type model with the total number of events up to time t, $N(t^-)$, as a time-varying covariate.

2.2.5 Likelihood Based Procedures

We now introduce the likelihood function for recurrent event processes. Suppose that m independent event processes are under observation and the i^{th} process, $i = 1, \ldots, m$, is observed over the fixed time interval $[\tau_{0i}, \tau_i]$, called the observation window. In the interval $[\tau_{0i}, \tau_i]$, the values of τ_{0i} and τ_i are the start and end-of-follow-up times of the i^{th} process, respectively.

Let $\{N_i(t); t \geq 0\}$ be a counting process with the intensity function $\lambda_i\left[t|\mathcal{H}_i(t)\right]$, $i = 1, \ldots, m$. The likelihood function of the outcome "$N_i(\tau_{0i}, \tau_i) = n_i$ events occurred at times $t_{i1} < \ldots < t_{in_i}$ over the interval $[\tau_{0i}, \tau_i]$, $i = 1, \ldots, m$," is

$$L = \prod_{i=1}^{m} L_i, \tag{2.31}$$

where

$$L_i = \left\{ \prod_{j=1}^{n_i} \lambda_i \left[t_{ij} | \mathcal{H}_i(t_{ij}) \right] \right\} \exp\left(- \int_{\tau_{0i}}^{\tau_i} \lambda_i \left[u | \mathcal{H}_i(u) \right] du \right). \tag{2.32}$$

The term in curly brackets in (2.32) corresponds to the likelihood of having n_i events at event times t_{ij}, $j = 1, \ldots, n_i$ for the i^{th} process. The second term, $\exp\left(- \int_{\tau_{0i}}^{\tau_i} \lambda_i \left[u | \mathcal{H}_i(u) \right] du \right)$, corresponds to the likelihood of having no events over (τ_{0i}, τ_i) for the i^{th} process, except in those t_{ij}, $j = 1, \ldots, n_i$, time points. An extensive explanation to the construction of (2.32) with product-integration and Taylor series expansion is given by Andersen et al. (1993, Section II.1).

It should be noted that the likelihood function (2.31) is valid for observation schemes, in which follow-up of processes can start and end at various prespecified τ_{i0} and τ_i times, respectively. The requirement of continuously follow-up and prespecified τ_{i0} and τ_i values can be restrictive in many real-life applications. For example, a process may not be continuously at risk of experiencing an event. The likelihood function can be extended to deal with such situations as explained next. Let the random variable $Y_i(t), t > 0$, takes the value 1 when the counting process $\{N_i(t); t > 0\}$ is under observation and at risk of having an event at time t; otherwise, it takes the value 0. The random variable $Y_i(t)$ is called *at risk indicator function* of the i^{th} process. We then define a process $\{Y_i(t); t \geq 0\}$. Various observation schemes and scenarios can be adapted by the use of the process $\{Y_i(t); t \geq 0\}$. For example, when the counting process $\{N_i(t); t \geq 0\}$ is being continuously observed over the interval $[\tau_{0i}, \tau_i]$, where τ_0 and τ are prespecified quantities, and the process $\{N_i(t); t > 0\}$ is under risk of having an event all over the observation window $[\tau_{0i}, \tau_i]$, then $Y_i(t) = I(\tau_{0i} \leq t \leq \tau_i)$, which is deterministic. In this setting, however, the likelihood function in (2.31) is the same. In more general settings, the likelihood function is given by

$$L = \prod_{i=1}^{m} \left\{ \prod_{j=1}^{n_i} \lambda_i \left[t_{ij} | \mathcal{H}_i(t_{ij}) \right] \, \exp\left(- \int_{0}^{\infty} Y_i(u) \, \lambda_i \left[u | \mathcal{H}_i(u) \right] du \right) \right\}, \tag{2.33}$$

where we assume that the processes $\{N_i(t); t \geq 0\}$ and $\{Y_i(t); t \geq 0\}$ are independent for $i = 1, \ldots, m$. An excellent discussion of the generalization of the likelihood function to (2.33) can be found in (Cook and Lawless, 2007, Section 2.6).

External covariates in the model (2.22) can be time fixed or time-dependent. We consider covariates as processes. If observable external covariates $\boldsymbol{x}(t)$ are related to event occurrences, they can be incorporated in the model by extending the process history $\mathcal{H}(t)$ to include covariate information. This extension is referred to as innovation theorem, more detail about this theorem is available in Section 2.4.2 in Andersen et al. (1993) and Section 2.2.7 in Aalen et al. (2008). Following the notation in Cook and Lawless (2007), we let $\boldsymbol{x}^{(t)} = \{\boldsymbol{x}(s) : 0 \leq s \leq t\}$ denote the history of the external covariates over $[0, t]$, and $\boldsymbol{x}^{(\infty)}$ denote the complete covariate path. Unless stated otherwise, we assume that probabilities are conditional on the covariate path of the external covariates and that their values are included in the initial information $\mathcal{H}(0)$ for convenience so that $\lambda[t|\mathcal{H}(t)]$ depends only on $\boldsymbol{x}^{(t)}$, not other way around.

In renewal process setting, the likelihood contribution by the i^{th} individual in a renewal process is formulated as

$$L_i = \left\{ \prod_{j=1}^{n_i} f_j(w_{ij}|\boldsymbol{x}_i) \right\} \; S_{n_i+1}(w_{i,n_i+1}|\boldsymbol{x}_i), \tag{2.34}$$

by assuming that n_i renewals occurred and the ${n_i}^{\text{th}}$ gap time is right-censored. If the observation terminates after ${n_i}^{\text{th}}$ event, that is, if $w_{i,n_i+1} = 0$, the survival function of the w_{i,n_i+1} $S_{n_i+1}(w_{i,n_i+1}|\boldsymbol{x}_i)$ in (2.34) becomes one.

2.3 Heterogeneity in Recurrent Event Processes

Individuals in a study cohort are usually heterogeneous with respect to some characteristics. In such cases, even after conditioning on the available covariate information, some models may not be adequate for accommodating such unexplained heterogeneity. For example, if event counts are of interest, then a Poisson process may not be adequate to address the excess heterogeneity in the counts of events across individuals. This issue is especially common in medical studies when the subjects are human (Cook and Lawless, 2007).

There are two common ways to deal with unexplained heterogeneity in recurrent events settings. Following our notation in the previous section, a natural extension of the model (2.22) is given by

$$\lambda_i[t|\mathcal{H}(t)] = \lambda_{0i}(t;\boldsymbol{\eta}_i)\ \exp[\boldsymbol{\psi}'\boldsymbol{W}_{\boldsymbol{i}}^*(t)], \qquad i=1,\ldots,m,\ t>0, \tag{2.35}$$

where $\lambda_{0i}(t;\boldsymbol{\eta}_i)$ is a subject specific baseline intensity function indexed with the vector of parameters $\boldsymbol{\eta}_i$. For example, a constant intensity for the baseline intensity function can be specified for the i^{th} individual as $\lambda_{0i}(t;\boldsymbol{\eta}) = \nu_i$, where ν_i is a positive valued parameter to be estimated. This type of models is called a *fixed effects* model. Although fixed effects models can be useful in some applied settings, they may suffer from the nuisance parameter problem when the number of events per individual over a fixed follow-up period is small and the number of individuals m in the data is large. In such cases, the maximum likelihood estimators of model parameters may become inconsistent (Cigsar and Lawless, 2012).

Unexplained heterogeneity across individuals can be addressed with *random effects* models as well. In this case, given $\nu_1, \nu_2, \ldots, \nu_m$, the conditional intensity function of the i^{th} individual, $i=1,\ldots,m$, takes the form

$$\lambda_i[t|\mathcal{H}_i(t),\nu_i] = \nu_i\ \exp[\boldsymbol{\psi}'\boldsymbol{W}_{\boldsymbol{i}}^*(t)], \qquad t>0, \tag{2.36}$$

where $\nu_1, \nu_2, \ldots, \nu_m$, are positive-valued independent and identically distributed (i.i.d.) unobservable random variables. The ν_i are assumed to follow a distribution. Because of its mathematical tractability, the gamma distribution with mean 1 and variance $\phi > 0$ is a popular choice for the distribution of the ν_i. In this case, the probability density function (p.d.f.) of $\nu_i,\ i=1,\ldots,m$, is then given by

$$g(\nu;\phi) = \frac{\nu^{\phi^{-1}-1}\exp(-\nu/\phi)}{\phi^{\phi^{-1}}\Gamma(\phi^{-1})}, \qquad 0<\nu<\infty,\ \phi>0, \tag{2.37}$$

which is, in notation, $\nu_i \sim Gamma(1,\phi)$, for $i=1,\ldots,m$.

The fixed effects model (2.35) with the baseline intensity function ν_i and the random effects model (2.36) may look similar. However, an important difference here is that the fixed effects model (2.35) is a marginal model in the sense that it only conditions on the history $\mathcal{H}_i(t)$, but the random effects model (2.36) conditions on the history $\mathcal{H}_i(t)$ and the value of the unobservable random effects ν_i. Under the assumption that the ν_i are i.i.d. gamma random variables with mean 1 and variance ϕ, the resulting marginal intensity function is given by

$$\lambda_i[t \mid \mathcal{H}_i(t)] = \frac{(1+\phi N_i(t^-))}{\left[1+\phi \int\limits_0^\tau \exp[\boldsymbol{\psi}' \boldsymbol{W}_{\boldsymbol{i}}^*(u)]du\right]} \exp[\boldsymbol{\psi}' \boldsymbol{W}_{\boldsymbol{i}}^*(t)]. \tag{2.38}$$

The derivation of (2.38) is available in Appendix A.

Estimation in the random effects models is usually carried out after integrating out the random effects from the likelihood function. In this case, the resulting likelihood contribution by the i^{th} individual to the likelihood function $L = \prod_{i=1}^m L_i$ is

$$L_i = \int_0^\infty \left\{ \prod_{j=1}^{n_i} \lambda_i \left[t_{ij} | \mathcal{H}_i(t_{ij}), \nu_i\right] \right\} \exp\left(- \int\limits_0^\infty Y_i(u) \lambda_i \left[u | \mathcal{H}_i(u), \nu_i\right] du \right) g(\nu_i; \phi) d\nu_i. \tag{2.39}$$

In the renewal processes setting, conditional hazard functions of the k^{th} gap time W_{ik} for the i^{th} individual can be extended as

$$h_{ik}(w|\boldsymbol{x}_i, \nu_i) = \nu_i h_{0k}(w) \exp(\boldsymbol{x}_i' \boldsymbol{\psi}), \qquad k = 1, 2, \ldots; w > 0, \tag{2.40}$$

for proportional hazards model. The likelihood contribution by the i^{th} individual is then given by

$$L_i = \int_0^\infty \left\{ \prod_{j=1}^{n_i} f_j(w_{ij}|\boldsymbol{x}_i, \nu_i) \right\} S_{n_i+1}(w_{i,n_i+1}|\boldsymbol{x}_i, \nu_i)\; g(\nu_i; \phi) d\nu_i. \tag{2.41}$$

The model (2.41) is called the proportional hazards frailty model. More information on these model is given by Cook and Lawless (2007, Section 4.2.2).

2.4 Dependence Concepts

In recurrent event settings, models based on renewal processes assume that the gap times between event times are independent. This assumption is a strong one and rarely true in many applications unless a complete renewal is assumed after each event occurrence (Cook and Lawless, 2007, Section 4.1). Two common methods to deal with the dependency between gap times in recurrent event settings are based on the random effects models and copula models. In this section, we discuss the models based on copulas and related concepts. A more detailed discussion of the topics can be found in Nelsen (2006) and Joe (2014).

Let X_1 and X_2 be two random variables with cumulative distribution functions (c.d.f.) $F_1(x_1)$ and $F_2(x_2)$, respectively, and probability density functions (p.d.f.) $f_1(x_1)$ and $f_2(x_2)$, respectively. Let $F_{1,2}(x_1, x_2)$ be the joint c.d.f. of the random vector (X_1, X_2). Then, the joint distribution of (X_1, X_2) can be written as a product of their marginal distributions only when the random variables X_1 and X_2 are independent. That is, $F_{1,2}(x_1, x_2) = F_1(x_1)F_2(x_2)$ for any values of x_1 and x_2 within their domains. If X_1 and X_2 are not independent, we can replace one of the marginal distributions with the conditional distribution of the other random variable, given the other variable. In this case, we can write that $F_{1,2}(x_1, x_2) = F_1(x_1) \times F_{2|1}(x_2|X_1 = x_1)$ or $F_{1,2}(x_1, x_2) = F_{1|2}(x_1|X_2 = x_2) \times F_2(x_2)$, where $F_{2|1}(x_2|X_1 = x_1) = \int_{-\infty}^{x_2} f_{1,2}(x_1, u)/F_1(x_1)du$, $F_{1|2}(x_1|X_2 = x_2) = \int_{-\infty}^{x_1} f_{1,2}(u, x_2)/F_2(x_2)du$ and $f_{1,2}(x_1, x_2) = \frac{\partial^2}{\partial x_1 \partial x_2} F_{1,2}(x_1, x_2)$. This is one way of formulating the dependency among two variables X_1 and X_2. This procedure can be extended to more than two random variables case.

The *Pearson correlation coefficient*, denoted by ρ_P, is a simplified measurement of the linear dependence between two random variables coming from a bivariate normal distribution. For a pair of continuous random variables (X_1, X_2), the formula of the Pearson correlation coefficient is

$$\rho_P = \frac{E(X_1X_2) - E(X_1)E(X_2)}{\sigma_1\sigma_2}, \tag{2.42}$$

where $E(X_k) = \int x_k \, dF_k(x_k), k = 1, 2,$ $E(X_1X_2) = \int x_1x_2 \, dF_{1,2}(x_1, x_2)$, $\sigma_1 > 0$ and $\sigma_2 > 0$ are the standard deviations of the random variables X_1 and X_2, respectively. The range of possible values for ρ_P is $[-1, 1]$. It is well-known that the ρ_P is not a good measure of dependence for many bivariate distributions when the two variables are not linearly related (Joe, 1997, Section 2.1.9). Therefore, nonparametric measures based on concordance were developed. Two observations (x_{i1}, x_{i2}) and (x_{j1}, x_{j2}), $i \neq j$, are called concordant if $(x_{i1} - x_{j1})(x_{i2} - x_{j2}) > 0$ and discordant if $(x_{i1} - x_{j1})(x_{i2} - x_{j2}) < 0$. Kendall's tau and Spearman's rho, as defined next, are the most frequently used measures of association for two random variables based on concordance and discordance.

Kendall's tau is the probability of concordance minus the probability of discordance between two random vectors $(X_{i1} - X_{j1})$ and $(X_{i2} - X_{j2})$; that is

$$\begin{aligned} \tau_K &= \Pr((X_{i1} - X_{j1})(X_{i2} - X_{j2}) > 0) - \Pr((X_{i1} - X_{j1})(X_{i2} - X_{j2}) < 0), \\ &= 4 \int F_{1,2}(x_1, x_2) dF_{1,2}(x_1, x_2) - 1, \\ &= 4E[F_{1,2}(X_1, X_2)] - 1. \end{aligned} \tag{2.43}$$

The range of Kendall's tau τ_K is $[-1, 1]$. The result in the second line of (2.43) can be found in Joe (1997).

Spearman's rho is defined as the Pearson correlation coefficient of $F_1(X_1)$ and $F_2(X_2)$. It is proportional to the probability of concordance minus the probability of discordance for the two vectors (X_{i1}, X_{i2}) and (X_{j1}, X_{k2}), $i \neq j \neq k$, such that the joint distribution function of (X_{i1}, X_{i2}) is $F_{1,2}(x_1, x_2)$ and the joint distribution function of (X_{j1}, X_{k2}) is $F_1(x_1)F_2(x_2)$. Since $F_1(X_1)$ and $F_2(X_2)$ are standard uniform random variables with expectations and variances are 1/2 and 1/12, respectively, the Spearman's rho is given by

$$\begin{aligned} \rho_S &= 3[\Pr((X_{i1} - X_{j1})(X_{i2} - X_{k2}) > 0) - \Pr((X_{i1} - X_{j1})(X_{i2} - X_{k2}) < 0)], \\ &= 12 \int\int F_1(x_1)F_2(x_2) dF_{1,2}(x_1, x_2) - 3, \\ &= 12 \int\int F_{1,2}(x_1, x_2) dF_1(x_1) dF_2(x_2) - 3. \end{aligned} \tag{2.44}$$

The range of Spearman's rho ρ_S is $[-1, 1]$. These two measures are invariant with respect to strictly increasing transformations of the random variables X_1 and X_2 (Joe, 1997). The relationship between Kendall's tau and Spearman's rho measures of association and other details are given in Joe (1997) and Nelsen (2006).

Another important concept is tail dependence. *Tail dependence* measures the dependence between two continuous random variables X_1 and X_2 in the lower-quadrant and upper-quadrant tails of their distribution functions F_1 and F_2, respectively (Nelsen, 2006, Section 5.4). The lower-tail dependence and upper-tail dependence are defined as follows. Let x_{1p} and x_{2p} are the $100p^{\text{th}}$ percentile of the distributions of X_1 and X_2, respectively. That is, $\Pr(X_2 \leq x_{2p}) = p$, where $p \in [0, 1]$

and $\Pr(X_1 \leq x_{1p}) = p$. If it exists, the lower tail dependence parameter λ_L is then defined as the limit of the conditional probability that X_2 is less than or equal to x_{2p}, given that X_1 is less than or equal to x_{1p}, as p approaches to 0. That is,

$$\lim_{p \to 0^+} \Pr(X_2 \leq x_{2p} | X_1 \leq x_{1p}) = \lambda_L. \tag{2.45}$$

Similarly, the upper tail dependence parameter λ_U is the limit, if it exists, of the conditional probability that X_2 is greater than the x_{2p} given that X_1 is greater than the x_{1p} as p approaches to 1, i.e.

$$\lim_{p \to 1^-} \Pr(X_2 > x_{2p} | X_1 > x_{1p}) = \lambda_U, \tag{2.46}$$

where x_{1p} and x_{2p} are same as defined above.

2.4.1 Copula Models

Another way of formulating dependencies is through copula functions. The word "copula" derives from the Latin verb *copulare*, meaning "to join together." Sklar (1959, 1973) introduced the term "copula", and obtained several characterizations for k-dimensional copulas in the context of probabilistic measure spaces. Copulas are functions used to construct a joint distribution function by combining the marginal distributions. Theory and applications of copula functions are available in Joe (1997) and Nelsen (2006).

Let the notation $[0,1]^k$ denote the set of all k-dimensional vectors, in which each element is in $[0,1]$. A k-variate copula is a function $C(u_1, \ldots, u_k)$, where $(u_1, \ldots, u_k) \in [0,1]^k$, with the following properties.

1. The margins of C are uniform if all the other arguments are equal to 1. That is, for the i^{th} argument, we have

$$C(1, \ldots, 1, u_i, 1, \ldots, 1) = u_i, \qquad i = 1, 2, \ldots, k.$$

2. The k-variate copula C is zero if one of the arguments is zero. For example, if

the i^{th} argument is zero, we have

$$C(u_1, \ldots, u_{i-1}, 0, u_{i+1}, \ldots, u_k) = 0, \qquad i = 1, 2, \ldots, k.$$

3. The k-variate copula C is k-non-decreasing. That is, for any two k-dimensional vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ in $[0,1]^k$ that satisfy

$$V_C([\boldsymbol{a}, \boldsymbol{b}]) \geq 0, \qquad \boldsymbol{a} \leq \boldsymbol{b},$$

where $V_C([\boldsymbol{a}, \boldsymbol{b}]) = \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} C(\boldsymbol{u}) = \Delta_{a_k}^{b_k} \Delta_{a_{k-1}}^{b_{k-1}} \ldots \Delta_{a_2}^{b_2} \Delta_{a_1}^{b_1} C(\boldsymbol{u})$, $\boldsymbol{u} = (u_1, \ldots, u_k)'$ and $\Delta_{a_r}^{b_r} C(\boldsymbol{u}) = C(u_1, \ldots, u_{r-1}, b_r, u_{r+1}, \ldots, u_k) - C(u_1, \ldots, u_{r-1}, a_r, u_{r+1}, \ldots, u_k)$. Here we denote $\boldsymbol{a} \leq \boldsymbol{b}$ when $a_r \leq b_r$ for all $r, 1 \leq r \leq k$.

Let $W(u_1, \ldots, u_k) = \max\left\{1 - k + \sum_{i=1}^{k} u_i, 0\right\}$ and $M(u_1, \ldots, u_k) = \min\{u_1, \ldots, u_k\}$. The Fréchet-Hoeffding Theorem (see Nelsen, 2006, p. 47) states that for any Copula $C : [0,1]^k \to [0,1]$ and any $(u_1, \ldots, u_k) \in [0,1]^k$ the inequality $W(u_1, \ldots, u_k) \leq C(u_1, \ldots, u_k) \leq M(u_1, \ldots, u_k)$ holds. The function W and M are called lower and upper Fréchet-Hoeffding bounds, respectively.

The Sklar's Theorem (Sklar, 1959) states that, if the marginal distribution functions $F_i(x_i)$, $i = 1, \ldots, k$, are continuous, there exists a unique copula C such that

$$\Pr(X_1 \leq x_1, \ldots, X_k \leq x_k) = F(x_1, \ldots, x_k) = C(F_1(x_1), \ldots, F_k(x_k)). \tag{2.47}$$

Similarly, if the marginal survival functions $S_i(x_i)$, $i = 1, \ldots, k$, are continuous, there exists a unique copula $\breve{C}$ (Georges et al., 2001) such that,

$$\Pr(X_1 > x_1, \ldots, X_k > x_k) = S(x_1, \ldots, x_k) = \breve{C}(S_1(x_1), \ldots, S_k(x_k)). \tag{2.48}$$

When $k = 2$, the relationship between C and $\breve{C}$ can be shown as follows.

$$\begin{aligned}
C(F_1(x_1), F_2(x_2)) &= \Pr(X_1 \leq x_1, X_2 \leq x_2), \\
&= 1 - \Pr(X_1 > x_1) - \Pr(X_2 > x_2) + \Pr(X_1 > x_1, X_2 > x_2), \\
&= 1 - (1 - F_1(x_1)) - (1 - F_2(x_2)) + \breve{C}(S_1(x_1), S_2(x_2)), \\
&= F_1(x_1) + F_2(x_2) - 1 + \breve{C}((1 - F_1(x_1)), (1 - F_2(x_2))).
\end{aligned}$$

Copula techniques have some attractive properties. For example, the marginal distributions of the random variables $X_1, X_2, \ldots, X_k$ can be defined from different families. This property is especially useful in the gap time analysis of recurrent events when the gap times are assumed to have different characteristics. Another useful property is that the dependence structure can be investigated independently from the marginal distributions. Furthermore, copulas are invariant under strictly increasing transformations of the margins. That is, for example in the bivariate case, if $X_2 = G_1(X_1)$ and $Y_2 = G_2(Y_1)$, where G_1 and G_2 are strictly increasing functions, then (X_1, Y_1) and (X_2, Y_2) have the same copula. Since τ_K and ρ_S in (2.43) and (2.44), respectively, are invariant to strictly increasing transformations, they can be used as summary measures of dependence for bivariate copulas. Let C be a bivariate copula of the random variables of U_1 and U_2. The Kendall's tau can be expressed as (see Nelsen, 2006, p. 159)

$$\begin{aligned} \tau_K &= 4 \int \int C(u_1, u_2) dC(u_1, u_2) - 1, \\ &= 4E[C(U_1, U_2)] - 1, \end{aligned} \tag{2.49}$$

and the Spearman's rho becomes (see Nelsen, 2006, p. 167)

$$\rho_S = 12 \int \int C(u_1, u_2) du_1 du_2 - 3. \tag{2.50}$$

The tail dependencies can also be formalized as follows (Joe, 1997, Section 2.1). If a bivariate copula C of the random variables U_1 and U_2 is such that

$$\lim_{u \to 0^+} \Pr(U_2 \leq u | U_1 \leq u) = \lim_{u \to 0^+} \frac{C(u, u)}{u} = \lambda_L \tag{2.51}$$

exists, then C has lower tail dependence if $\lambda_L \in (0, 1]$ and no lower tail dependence if $\lambda_L = 0$. The upper tail dependency can be defined as follows. If

$$\lim_{u \to 1^-} \Pr(U_2 > u | U_1 > u) = \lim_{u \to 1^-} = \frac{1 - 2u + C(u, u)}{1 - u} = \lambda_U \tag{2.52}$$

exists, then the copula C has an upper tail dependence if $\lambda_U \in (0, 1]$ and no upper tail dependence if $\lambda_U = 0$.

A detailed list of one and two-parameter bivariate copula functions can be found in

Joe (1997, Section 5). In the remainder part of the this section, we introduce bivariate copula classes which are relevant to our study, and discuss some of their properties. These copula models are used in the following chapters.

2.4.1.1 Archimedean Copulas

A bivariate Archimedean copula of the random variables U_1 and U_2 (Genest and MacKay, 1986) can be written in the form

$$C(u_1, u_2) = \varphi^{[-1]}[\varphi(u_1) + \varphi(u_2)], \tag{2.53}$$

where φ, called a *generator function* of corresponding copula function, is a strictly decreasing convex function on $(0, 1]$ to $[0, \infty]$ satisfying $\varphi(1) = 0$. The pseudo-inverse function of φ is given by

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t), & 0 \leq t \leq \varphi(0), \\ 0, & \varphi(0) \leq t \leq \infty. \end{cases} \tag{2.54}$$

A bivariate Archimedean copula contains all information about the 2-dimensional dependence structure between random variables U_1 and U_2 through a univariate generator function φ. A lengthy discussion of Archimedean copulas are given by Joe (1997, Section 4.2) and Nelsen (2006, Section 4.3). Archimedean copula models can be naturally derived from bivariate frailty models since φ^{-1} is the Laplace transform of the underlying frailty distribution (Oakes, 1989; Joe, 1997).

2.4.1.2 One-Parameter Copula Models

We next introduce some frequently used one-parameter Archimedean families. The Clayton family (Clayton, 1978) has the form

$$C_\phi(u_1, u_2) = \left(u_1^{-\phi} + u_2^{-\phi} - 1\right)^{-1/\phi}, \quad \phi \in [-1, \infty)\backslash\{0\}. \tag{2.55}$$

Its generator function is

$$\varphi_\phi(t) = t^{-\phi} - 1, \qquad 0 < t \leq 1. \tag{2.56}$$

Since $\varphi_\phi^{-1}(v) = (v+1)^{-1/\phi}$ is the Laplace transform of a gamma distribution with the shape parameter $1/\phi$ and the scale parameter 1, the bivariate gamma shared frailty model leads to the Clayton survivor copula model (Goethals et al., 2008). It should be noted that, with the copula parameter ϕ in (2.55), the Kendall's tau coefficient is given by

$$\tau_K(\phi) = \frac{\phi}{\phi+2}, \qquad \phi \in [-1,\infty)\backslash\{0\}. \tag{2.57}$$

The random variables U_1 and U_2 are positively associated when $\phi > 0$ and the dependence between them increases as the value of the parameter ϕ increases. The independent copula is obtained when the parameter ϕ approaches zero and the Fréchet-Hoeffding upper bound is obtained as the parameter ϕ approaches infinity. The lower tail dependence parameter of the Clayton copula is $\lambda_L = 2^{-1/\phi}$, whereas there is no upper tail dependence.

The Gumbel-Hougaard family (Gumbel, 1960) is another family of copula models, which is of the form

$$C_\theta(u_1,u_2) = \exp\left[-\left\{(-\log u_1)^\theta + (-\log u_2)^\theta\right\}^{1/\theta}\right], \quad \theta \geq 1. \tag{2.58}$$

Its generator function is

$$\varphi_\theta(t) = (-\log t)^\theta, \qquad 0 < t \leq 1. \tag{2.59}$$

Since $\varphi_\theta^{-1}(v) = \exp\left(-v^{1/\theta}\right)$ is the Laplace transform of a positive stable distribution, the positive stable shared frailty model leads to the Gumbel-Hougaard survivor copula model (Duchateau and Janssen, 2007, Section 4.4). It can be shown that, in this case, the Kendall's tau coefficient is

$$\tau_K(\theta) = \frac{\theta-1}{\theta}, \qquad \theta \geq 1. \tag{2.60}$$

The Kendall's tau coefficient τ_θ indicates an increasing dependence between two random variables as the value of the parameter θ increases. The independent copula is obtained as θ approaches 1 and the Fréchet-Hoeffding upper bound is obtained as $\theta \to \infty$. The upper tail dependence parameter is $\lambda_U = 2 - 2^{1/\theta}$, and there is no lower tail dependence.

The next bivariate copula family that we introduce is the Frank family (Frank, 1979), which has the form

$$C_\nu(u_1,u_2) = -\frac{1}{\nu}\log\left[1+\frac{\left(e^{-\nu u_1}-1\right)\left(e^{-\nu u_2}-1\right)}{\left(e^{-\nu}-1\right)}\right], \quad \nu \in \mathbb{R}\backslash 0, \tag{2.61}$$

with the generator function

$$\varphi_\nu(t) = \log\left[\frac{e^{-\nu}-1}{e^{-\nu t}-1}\right], \qquad 0 < t \leq 1, \tag{2.62}$$

and the Kendall's tau coefficient

$$\tau_K(\nu) = 1 + 4\frac{D_1(\nu)-1}{\nu}, \qquad \nu \in \mathbb{R}\backslash 0, \tag{2.63}$$

where D_1 is the first Debye function; that is, $D_1(\nu) = \int_0^\nu \frac{t}{\nu(e^t-1)}dt$ (Abramowitz and Stegun, 1948, Section 27.1) and $\mathbb{R}$ is denoted as the set of real numbers. The random variables U_1 and U_2 are positively associated when $\nu > 0$ and negatively associated when $\nu < 0$. The independent copula is obtained as $\nu \to 0$. The Fréchet-Hoeffding upper and lower bounds are obtained as $\nu \to \infty$ and as $\nu \to -\infty$, respectively.

It should be noted that Clayton copula has the potential to capture the lower tail dependency and Gumbel-Hougaard copulas can capture the upper tail dependency, whereas Frank copulas do not classify either tail behavior (Embrechts et al., 2003).

2.4.1.3 Two-Parameter Copula Models

Two or more-parameter copula families provide greater flexibility for fitting data since they can capture more than one type of dependence. When such a family includes some of the well-known one-parameter copula families such as Clayton, Frank and Gumbel-Hougaard, testing of those models can be easily performed, for example, by applying the model expansion technique (Lawless and Yilmaz, 2011). We provide an example below.

A bivariate two-parameter family of the form of an Archimedean copula is

$$C_{\phi,\theta}(u_1,u_2) = \left\{\left[\left(u_1^{-\phi}-1\right)^\theta + \left(u_2^{-\phi}-1\right)^\theta\right]^{1/\theta}+1\right\}^{-1/\phi}, \phi > 0 \text{ and } \theta \geq 1. \tag{2.64}$$

This family includes the Clayton and Gumbel-Hougaard families as special cases. In particular, it reduces to the Clayton family when $\theta = 1$, and becomes the Gumbel-Hougaard family as $\phi \to 0$. Its generator function is

$$\psi_{\phi,\theta}(t) = \left(t^{-\phi} - 1\right)^{\theta}, \qquad 0 < t \leq 1, \tag{2.65}$$

and the Kendall's tau is

$$\tau_K(\phi, \theta) = 1 - \frac{2}{\theta(\phi + 2)}, \qquad \phi > 0 \text{ and } \theta \geq 1. \tag{2.66}$$

The dependence increases as the parameters θ and/or ϕ increase. The independent copula $u_1 u_2$ is obtained as $\phi \to 0$ and $\theta \to 1$ and the Fréchet upper bound is obtained as $\phi \to \infty$ or $\theta \to \infty$.

Scatter plots of different type of copula data are useful to understand more about the requirement of different types of copulas. Figure 2.1 presents the scatter plots of various copula dependent bivariate random variables. Here, we considered standard exponential distribution for the marginals. In Figure 2.1, each pair of plots are based on the same simulated data. The difference is the right side plots are plotted against two random variables, say W_1 versus W_2, whereas the left side plots are plotted against their corresponding cumulative probabilities, respectively, say $U_1 = F_1(W_1)$ versus $U_2 = F_2(W_2)$.

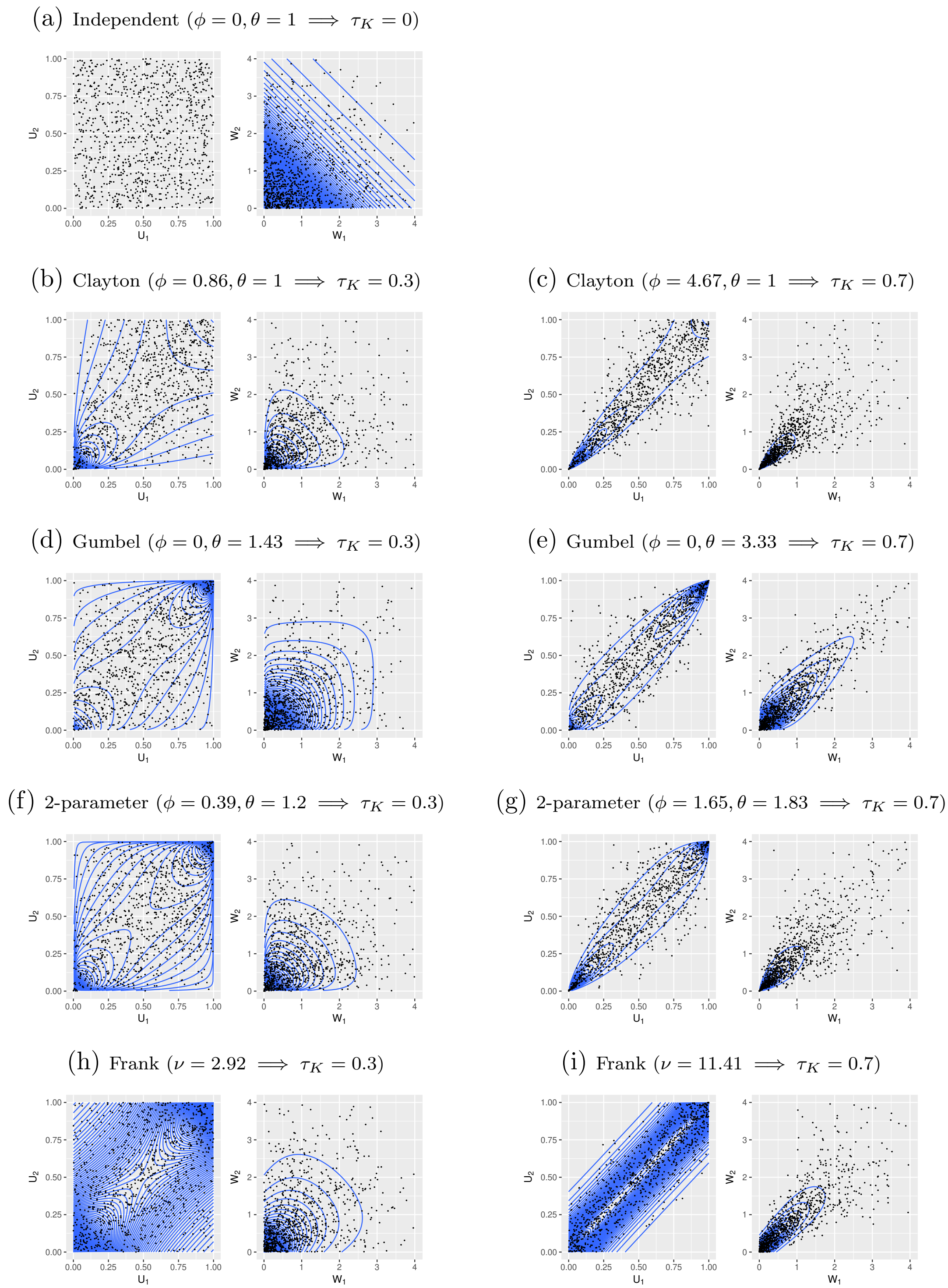

Figure 2.1: Scatter plots of simulated copula dependent (n=1,000) uniform random variables (U_1, U_2) and their corresponding standard exponential quantiles (W_1, W_2) with the contour plots of underlying theoretical densities for weak ($\tau_K = 0.3$) and strong ($\tau_K = 0.7$) dependence.

Notably, the points in the scatter plot of cumulative probabilities of observed bivariate data under proper marginal assumption should be evenly distributed within unit square when those two random variables are independent as seen in Figure 2.1a. In many statistical approaches, such a property is essential because they assume the independence. However, in many real data application, those approaches may lead to wrong conclusions, in particular, in the marginal estimates due to unavoidable dependency. Copula-based models play an important role in modeling when dependency involved. From the plots, we can see that when the positive dependency increases (i.e., when $\tau_K > 0$ increases) the points in the scatter plots of cumulative probabilities move near to line of equality.

2.4.1.4 Gaussian Copula and t Copula

Let Φ denote the distribution function of a random variable from the univariate standard normal distribution. Then, the bivariate Gaussian copula with single parameter ρ is given by

$$C(u_1, u_2; \rho) = \Phi_\rho\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2)\right), \tag{2.67}$$

where Φ_ρ is a bivariate normal distribution function with the 2×1 zero mean vector and a 2×2 correlation matrix, in which the diagonal elements are equal to 1 and off-diagonal elements are equal to ρ.

Similarly, let $t_{\nu,\rho}$ be the cumulative distribution function of the bivariate t-distribution with the 2×1 zero mean vector and a 2×2 correlation matrix with 1 as diagonal elements and ρ as off-diagonal elements, and the parameter ν denotes the degrees of freedom. Further, let t_ν denote the cumulative distribution function of the univariate t-distribution with degrees of freedom ν. The bivariate Student or t-copula with parameters ν and ρ is given by

$$C(u_1, u_2; \nu, \rho) = t_{\nu,\rho}\left(t_\nu^{-1}(u_1), t_\nu^{-1}(u_2)\right). \tag{2.68}$$

All the copulas as mentioned above so far can be extended to incorporate more than two variables.

2.4.2 Pair-Copula Constructions

The joint and marginal conditional density functions of bivariate random variables (W_1, W_2) can be written in terms of copula densities as

$$\begin{aligned} f_{12}(w_1, w_2) &= \frac{\partial^2 F_{12}(w_1, w_2)}{\partial w_1 \partial w_2} = c_{12}(F_1(w_1), F_2(w_2)) \times f_1(w_1) \times f_2(w_2), \\ f_{2|1}(w_2|w_1) &= \frac{f_{12}(w_1, w_2)}{f_1(w_1)} = c_{12}(F_1(w_1), F_2(w_2)) \times f_2(w_2), \end{aligned} \tag{2.69}$$

where

$$c_{12}(F_1(w_1), F_2(w_2)) = \frac{\partial^2 C(F_1(w_1), F_2(w_2))}{\partial F_1(w_1) \partial F_2(w_2)}, \tag{2.70}$$

called *pair-copula density function*, and $f_i(w_i)$ is the marginal density function of $w_i, i = 1, 2$. Consider a three-variate joint density function $f(w_1, w_2, w_3)$. Using the relationship between density and copula functions, it can be decomposed as follows

$$\begin{aligned} f(w_1, w_2, w_3) &= f_{3|12}(w_3|w_1, w_2) \times f_{2|1}(w_2|w_1) \times f_1(w_1), \\ &= c_{13|2}(F_{1|2}(w_1|w_2), F_{3|2}(w_3|w_2)) \;\; c_{23}(F_2(w_2), F_3(w_3)) \;\; f_3(w_3) \\ &\quad \times c_{12}(F_1(w_1), F_2(w_2)) \;\; f_2(w_2) \\ &\quad \times f_1(w_1), \end{aligned} \tag{2.71}$$

where the function $c_{13|2}$ is conditional on w_2, and the arguments of $c_{13|2}$ also conditional on w_2, which is called conditioning variable. When the number of variable increases, the number of conditioning variables also increases. Thus, it may result in a large number of parameters to be estimated. It can be assumed (Hobæk Haff et al., 2010) that all the pair-copulas depend on the conditioning variables only through the two conditional distribution functions which being their arguments, and not directly. That is, for example, a simplified version of the pair copula construction for a three-variate joint density function can be written by

$$\begin{aligned} f(w_1, w_2, w_3) &= c_{13|2}(F_{1|2}(w_1|w_2), F_{3|2}(w_3|w_2)) \\ &\quad \times c_{12}(F_1(w_1), F_2(w_2)) \;\; c_{23}(F_2(w_2), F_3(w_3)) \\ &\quad \times f_1(w_1) \;\; f_2(w_2) \;\; f_3(w_3). \end{aligned} \tag{2.72}$$

Any k-variate joint density function can be written in terms of marginal and pair-copula density functions. The decomposition given in (2.72) is not unique for $f(w_1, w_2, w_3)$. There are two more decompositions that can be made. For high-dimensional distributions, the number of possible pair-copula constructions increases drastically. For example, there are 240 different constructions for a five-dimensional density (Aas et al., 2009). Bedford and Cooke (2001) introduced a graphical method to organize such possible decompositions and denoted as the *regular vine (R-vine)*. To form an R-vine, a sequence of trees $\mathcal{V} = (T_1, \ldots, T_{k-1})$ has to fulfill the following conditions.

1. T_1 is a tree with nodes $N_1 = 1, \ldots, k$ and edges E_1.
2. For $i \geq 2$, T_i is a tree with nodes $N_i = E_{i-1}$ and edges E_i.
3. If two nodes in T_{i+1} are joint by an edge, the corresponding edges in T_i must share a common node.

The condition is sometimes referred to as the *proximity condition*. There are two important classes of vines; the *canonical vine (C-vine)* and the *D-vine*. Figure 2.2 and Figure 2.3 show k-dimensional C-vine and D-vine pair copula decompositions, respectively. Both consist of $k-1$ trees T_j, $j = 1, \ldots, k-1$, where T_j has $k+1-j$ nodes and $k-j$ edges. Every edge corresponds to a pair-copula density. The whole decomposition is defined by the $k(k-1)/2$ edges and the marginal densities of each variable.

The k-dimensional density corresponding to a C-vine is given by

$$f(w_1, \ldots, w_k) = \prod_{l=1}^{k} f_l(w_l) \prod_{j=1}^{k-1} \prod_{i=1}^{k-j} c_{j,j+i|1:j-1} \left[F(w_j|\boldsymbol{w}_{1:j-1}), F(w_{j+i}|\boldsymbol{w}_{1:j-1}) \right], \quad (2.73)$$

where

$$F(w_{j+i}|\boldsymbol{w}_{1:j-1}) = \frac{\partial C_{j+i,j-1|1:j-2} \left[F(w_{j+i}|\boldsymbol{w}_{1:j-2}), F(w_{j-1}|\boldsymbol{w}_{1:j-2}) \right]}{\partial F(w_{j-1}|\boldsymbol{w}_{1:j-2})}, \quad i = 0, \ldots, k-j, \quad (2.74)$$

and a D-vine is given by

$$f(w_1, \ldots, w_k) = \prod_{l=1}^{k} f_l(w_l) \prod_{j=1}^{k-1} \prod_{i=1}^{k-j} c_{i,(i+j)|i+1:i+j-1} \left[F(w_i|\boldsymbol{w}_{i+1:i+j-1}), F(w_{i+j}|\boldsymbol{w}_{i+1:i+j-1}) \right]. \quad (2.75)$$

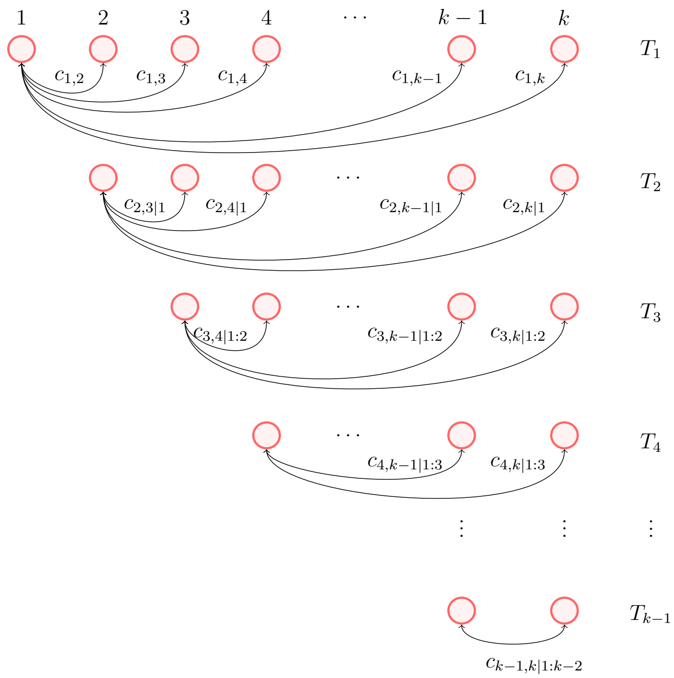

Figure 2.2: A C-vine with k variables, $k-1$ trees and $k(k-1)/2$ edges. Each edge corresponds to a pair-copula density.

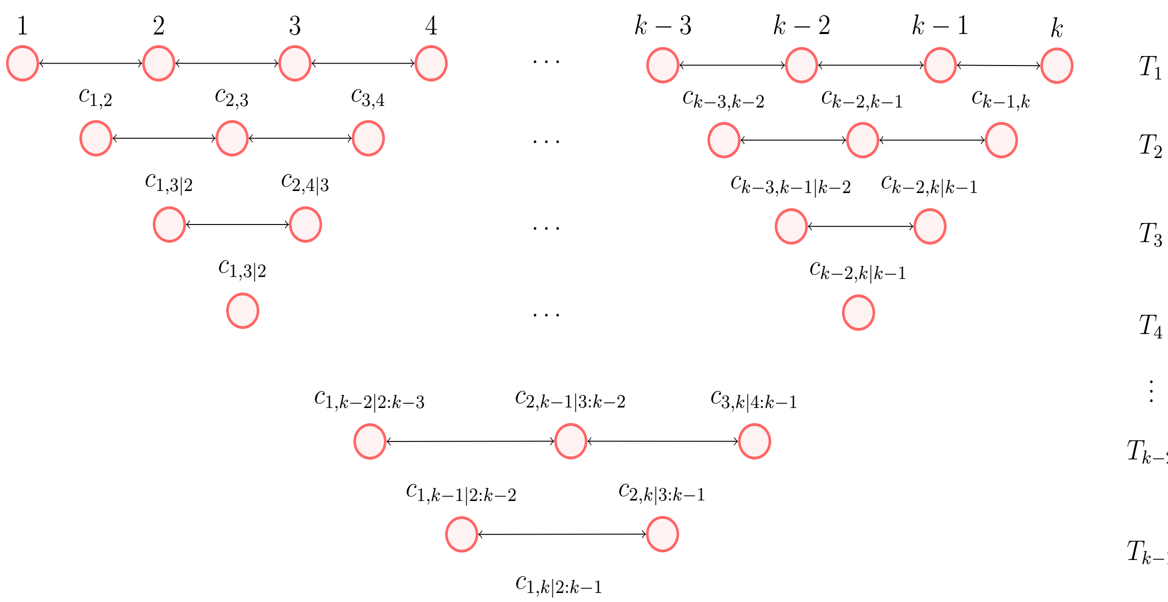

Figure 2.3: A D-vine with k variables, $k-1$ trees and $k(k-1)/2$ edges. Each edge corresponds to a pair-copula density.

where

$$F(w_i|\boldsymbol{w}_{i+1:i+j-1}) = \frac{\partial C_{i,i+j|i+1:i+j-1}\left[F(w_i|\boldsymbol{w}_{i+1:i+j-1}), F(w_{i+j}|\boldsymbol{w}_{i+1:i+j-1})\right]}{\partial F(w_{i+j}|\boldsymbol{w}_{i+1:i+j-1})}, \quad (2.76)$$

and

$$F(w_{i+j}|\boldsymbol{w}_{i+1:i+j-1}) = \frac{\partial C_{i,i+j|i+1:i+j-1}\left[F(w_i|\boldsymbol{w}_{i+1:i+j-1}), F(w_{i+j}|\boldsymbol{w}_{i+1:i+j-1})\right]}{\partial F(w_i|\boldsymbol{w}_{i+1:i+j-1})}. \quad (2.77)$$

Note that here we denote the vector $(w_j, \ldots, w_k)$ by $\boldsymbol{w}_{j:k}$ where $j < k$. A general formula for R-vine structure of $f(w_1, \ldots, w_k)$ is available in Dißmann et al. (2013). The arguments in the pair copulas are conditional on intermediate variables for D-vine structure, whereas in C-vine they are conditional on non-intermediate variables. In recurrent event context, those pair copulas are conditional on the previous gap times when we consider C-vine structure. It is more intuitive to pick the C-vine over the D-vine if our focus is on an event history analysis. Because the C-vine copula parameters in tree $T_i, i = 2, \ldots, k-1$ give the strength of dependence between W_i and $W_j, j = i+1, \ldots, k$ given $W_1, \ldots, W_{i-1}$.

2.5 Simulation Procedures

In this section, we introduce how to simulate realizations of a recurrent event process with a given intensity function. Our goal with the simulated data is to check whether the parameters are correctly estimated when the correct model is fitted and to study the behavior of the estimates of the parameters when the model is misspecified. In a recurrent event process, the gap times can be either independent or dependent. We therefore, present event generating algorithms under independence and dependence cases separately. In real data, it is critical to identify the dependency structure of the gap times. It can be either the current gap time depending on all the previous gap time or depending only on the previous gap time. This thesis mainly focuses on the latter one, and the corresponding data generation algorithm is presented in Section 2.5.2.

2.5.1 Simulation of a Serially Independent Event Process

Let $\{N(t); t \geq 0\}$ be a counting process with an associated intensity function $\lambda[t|\mathcal{H}(t)]$. If we let

$$E_j = \int_{t_{j-1}}^{t_{j-1}+W_j} \lambda[t|\mathcal{H}(t)]dt, \qquad j = 1, 2, \ldots, \tag{2.78}$$

where the W_j are the gap times generated by the process $\{N(t); t \geq 0\}$, then, given t_{j-1} and $H(t_{j-1})$, each random variable E_j has an exponential distribution with mean 1 (Cook and Lawless, 2007). This result follows from the fact that $\Pr\{W_j > t_{j-1} + w \; |T_{j-1} = t_{j-1}, H(t_{j-1})\} = \exp\left\{-\int_{t_{j-1}}^{t_{j-1}+w} \lambda[t|\mathcal{H}(t)]dt\right\}, j = 1, 2, \ldots,$ and so $U_j = \exp(-E_j)$ has a standard uniform distribution. The algorithm to generate event times of a recurrent event process for a given intensity function in this thesis is given as follows:

1. Set $j = 1$ and $t_0 = 0$.
2. Generate U_j from a standard uniform distribution.
3. Use the transformation $E_j = -\log(U_j)$.
4. Calculate the j^{th} event time T_j by solving $E_j = \int_{t_{j-1}}^{T_j} \lambda[t|\mathcal{H}(t)]dt$ for T_j.
5. If $T_j \leq \tau$, let $t_j = T_j$ and advance j by 1. Then, return to the second step. Otherwise, stop the loop. If $j = 1$, no events occurred over $[0, \tau]$. If $j > 1$, the recurrent event times observed over $[0, \tau]$ are given by $t_1, \ldots, t_n$, where $n = j-1$.

By repeating the above algorithm m times, we can generate recurrent event times from m identical processes with the intensity function $\lambda[t|\mathcal{H}(t)]$. The algorithm can be extended to generate nonidentical processes with a given intensity function $\lambda_i[t|\mathcal{H}_i(t), \nu_i]$, $i = 1, \ldots, m$. Note that this model is a random effects model, which is discussed in Section 2.3. The data generation algorithm becomes:

1. Set $j = 1$ and $t_{i0} = 0$.
2. Generate the values of the random effect ν_i from a given distribution.
3. Generate U_{ij} from a standard uniform distribution.
4. Use the transformation $E_{ij} = -\log(U_{ij})$.

5. Calculate the j^{th} event time T_{ij} for the i^{th} individual by solving $E_{ij} = \int_{t_{i(j-1)}}^{T_{ij}} \lambda_i[t|\mathcal{H}_i(t), \nu_i]dt$ for T_{ij}.

6. If $T_{ij} \leq \tau$, let $t_{ij} = T_{ij}$ and advance j by 1. Then, return to the third step. Otherwise, stop the loop. If $j = 1$, no events occurred over $[0, \tau]$. If $j > 1$, the recurrent event times observed over $[0, \tau]$ are given by $t_{i1}, \ldots, t_{in_i}$, where $n_i = j - 1$.

2.5.2 Simulation of a Serially Dependent Event Process

This subsection introduces algorithms to generate recurrent event times, where the subsequent gap times are serially dependent. Let $W_j, j = 1, 2, \ldots$ be j^{th} gap time of a recurrent event process where $W_j = T_j - T_{j-1}$ with $T_0 = 0$. For each gap time, let $h_j(w) = \lim_{s \to 0} \Pr(W_j < w + s | W_j \geq w)/s$, $w > 0$, be the hazard function of W_j, $j = 1, 2, \ldots$. The data generation algorithm is as follows.

1. Set $j = 1$ and $t_0 = 0$.

2. Calculate the j^{th} event time T_j where $T_j = t_{j-1} + W_j$.

   (a) When $j = 1$

      i. Generate U_1 from a standard uniform distribution.

      ii. Calculate the W_1 such that $W_1 = F_1^{-1}(U_1)$ where

$$F_1(w_1) = 1 - \exp\left\{ - \int_0^{w_1} h_1(s)\mathrm{d}s \right\}.$$

   (b) When $j \geq 2$

   Note: For a given copula function $C(u_{j-1}, u_j) = C(U_{j-1} \leq u_{j-1}, U_j \leq u_j)$, the conditional distribution of U_j given $U_{j-1} = u_{j-1}$: $c_{u_{j-1}}(u_j) = C(U_j \leq u_j | U_{j-1} = u_{j-1}) = \dfrac{\partial C(u_{j-1}, u_j)}{\partial u_{j-1}}$ is a standard uniform distribution.

      i. Generate U^* from a standard uniform distribution.

      ii. Calculate the U_j such that $U_j = c_{u_{j-1}}^{-1}(U^*)$

      iii. Calculate the W_j such that $W_j = F_j^{-1}(U_j)$ where

$$F_j(w_j) = 1 - \exp\left\{ - \int_0^{w_j} h_j(s)\mathrm{d}s \right\}.$$

3. If $T_j \leq \tau$, let $t_j = T_j$ and increase j by 1. Then, return to the second step. Otherwise, stop the loop. If $j = 1$, no events occurred over $[0, \tau]$. If $j > 1$, the recurrent event times observed over $[0, \tau]$ are given by $t_1, \ldots, t_n$, where $n = j - 1$.

To generate nonidentical processes with individual specific hazard functions $h_j(w|\nu_i)$, $j = 1, 2, \ldots$, $i = 1, \ldots, m$, we modify the above algorithm as below:

1. Set $j = 1$ and $t_{i0} = 0$.
2. Generate ν_i from a given distribution.
3. Calculate the j^{th} event time for the i^{th} individual T_{ij} where $T_{ij} = t_{i(j-1)} + W_{ij}$.
   (a) When $j = 1$
      i. Generate U_{i1} from a standard uniform distribution.
      ii. Calculate the W_{i1} such that $W_{i1} = F_1^{-1}(U_{i1})$ where
$$F_1(w_{i1}) = 1 - \exp\left\{-\int_0^{w_{i1}} h_1(s|\nu_i)\mathrm{d}s\right\}.$$
   (b) When $j \geq 2$
      i. Generate U^* from a standard uniform distribution.
      ii. Calculate the U_{ij} such that $U_{ij} = c^{-1}_{u_{i(j-1)}}(U^*)$
      iii. Calculate the W_{ij} such that $W_j = F_{ij}^{-1}(U_{ij})$ where
$$F_j(w_{ij}) = 1 - \exp\left\{-\int_0^{w_{ij}} h_{ij}(s|\nu_i)\mathrm{d}s\right\}.$$
4. If $T_{ij} \leq \tau$, let $t_{ij} = T_{ij}$ and advance j by 1. Then, return to the third step. Otherwise, stop the loop. If $j = 1$, no events occurred over $[0, \tau]$. If $j > 1$, the recurrent event times observed over $[0, \tau]$ are given by $t_{i1}, \ldots, t_{in_i}$, where $n_i = j - 1$.

Chapter 3

Analysis of Recurrent Events Using Dynamic Models for Event Counts

Past event occurrences during the evolution of a recurrent process may alter the probability of experiencing new events. In this chapter, we introduce a modulated Poisson process model for event counts. This model can incorporate two important features of recurrent event processes, event clustering and trend, as dynamic covariates. We discuss large sample properties of maximum likelihood estimators of model parameters, as well as some issues related to unexplained heterogeneity and other difficulties in the estimation of these features.

The remaining part of this chapter is organized as follows. The model formulation and issues regarding the existing approaches are discussed in Section 3.1 for identical processes. We consider the non-identical processes case in Section 3.2. Section 3.3 includes a summary of set of simulation studies under various settings to understand the issues related to models for event counts. In the last section of this chapter, we analyze a real data set to illustrate the methods introduced in this chapter.

3.1 A Dynamic Model for Carryover Effects and Number of Previous Events

As discussed in the first chapter, carryover effects and the number of previous events in a recurrent event process may cause clustering of events over time and some sort of monotonic trend in the event intensity function, respectively. A flexible model describing these effects as dynamic covariates along with other time-varying covariates can be defined through a modulated Poisson process as explained below.

Suppose that $\{N(t); t \geq 0\}$ is a a counting process with intensity function of the multiplicative form defined by $\lambda[t|\mathcal{H}(t)] = \lambda_0(t) g\left[\boldsymbol{W}^*(t); \boldsymbol{\psi}\right]$, where g is a non-negative function, $\boldsymbol{W}^*(t)$ is a vector of covariates and $\boldsymbol{\psi}$ is a vector of parameters. In this study, we let $g(t) = \exp(t)$, $t > 0$. The vector of covariates $\boldsymbol{W}^*(t)$ includes functions of the history $\mathcal{H}(t)$. It may also include external covariates. Let $\boldsymbol{W}^*(t) = (\boldsymbol{H}'(t), \boldsymbol{Z}'(t), \boldsymbol{x}'(t))'$, where the vector $\boldsymbol{H}(t)$ includes nondecreasing functions of the number of previous events $N(t^-)$, the vector $\boldsymbol{Z}(t)$ includes 0-1 valued terms that temporarily take value of 1 after event occurrences, and $\boldsymbol{x}(t)$ is the vector of external covariates. Note that components of the vectors $\boldsymbol{H}(t)$ and $\boldsymbol{Z}(t)$ are functions of the history $\mathcal{H}(t)$, including information on number of previous events and time since last event in the process $\{N(t); t \geq 0\}$, respectively. Therefore, as discussed in Section 2.2.2, the vector $\boldsymbol{W}^*(t)$ can be included in a recurrent event model and the likelihood based inference given in Chapter 2 can be applied to the vector of parameters $\boldsymbol{\psi}$.

In the remainder of this chapter, we specify $\boldsymbol{W}^*(t) = (N(t^-), Z(t), \boldsymbol{x}'(t))'$, where $Z(t) = I[N(t^-) > 0]\, I[B(t) \leq \Delta]$, $B(t) = t - T_{N(t^-)}$ is the backward recurrence time, I is a typical 0-1 valued indicator function and Δ is a positive valued prespecified quantity, called the *risk window*. With this specification of $\boldsymbol{W}^*(t)$, trends due to number of previous events and carryover effects can be investigated through the intensity function

$$\lambda[t|\mathcal{H}(t)] = \lambda_0(t) \exp\left[\gamma\, N(t^-) + \beta\, Z(t) + \boldsymbol{\xi}'\, \boldsymbol{x}(t)\right], \qquad t > 0, \tag{3.1}$$

where γ, β and $\boldsymbol{\xi}$ are any real valued regression parameters. The baseline intensity function $\lambda_0(t)$ in the model (3.1) can be parametrically or non-parametrically specified. The model (3.1) is simple enough to apply the likelihood based inference methods but quite useful to investigate these dynamic features. For example, Cigsar and Lawless (2012) used the model (3.1) without the trend component $N(t^-)$ to develop partial score tests for the presence of carryover effects in recurrent event processes.

The choice of a value for the risk window Δ is an important issue with the model (3.1). This is usually based on the background information, in particular when there are not too many events experienced per individual during their follow-ups. As an alternative, the quantity Δ can be treated as an unknown parameter to be estimated. However, as discussed by Cigsar and Lawless (2012), the profile likelihood for Δ is flat with respect to mild changes in the values of Δ. As a result, estimability may become an issue in some settings. Furthermore, Cigsar and Lawless (2012) showed through a simulation study that mild misspecification of the value of Δ does not impact the inference on the model parameters. In essence, the value of Δ should be small comparing with the average time between events in a process. It should be noted that the choice of Δ is also an issue with the carryover effects due to external factors (Farrington and Whitaker, 2006). Either internal or external carryover effects are of interest, a sensitivity analysis can be conducted if uncertainty related to the choice of a value for Δ is present (Xu et al., 2011; Cigsar and Lawless, 2012). We would like to note that, in Section 5.2.4, we discuss a semiparametric method, which may give some insight on the value of Δ by using the data.

As discussed later in this chapter, unexplained heterogeneity and trend due to number of previous events may confound with carryover effects, and make it difficult to investigate carryover effects in some cases. To discuss these issues and introduce our methodology, we parametrically specify the baseline intensity function $\lambda_0(t)$ in (3.1) with an unknown constant α, where $\alpha > 0$. Figure 3.1 illustrates the shapes of three different versions of the log transformed intensity function (3.1) fitted by using a simulated data set as t increases.

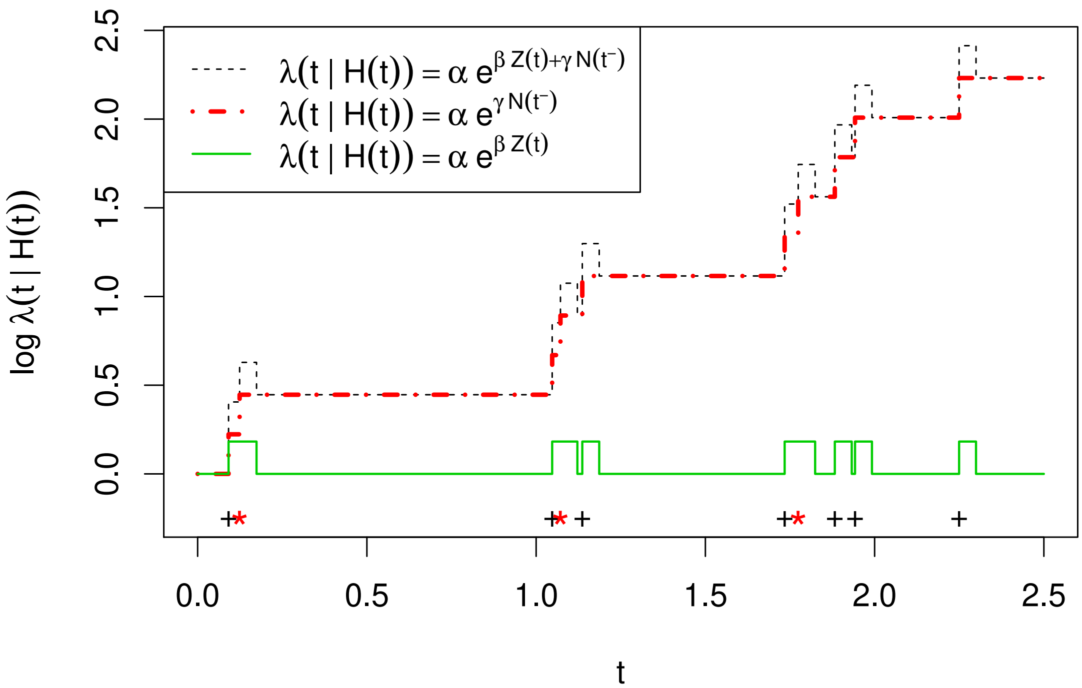

Figure 3.1: Realizations of three different dynamic intensity functions for an event process over the interval $[0, 2.5]$. The '+' signs in the plot indicate the events and the '$*$' signs are for the events which occur within the Δ time period after an event.

Suppose that there are m independent processes under observation. Let $\boldsymbol{\theta} = (\alpha, \gamma, \beta, \boldsymbol{\xi}')'$, where $\boldsymbol{\xi}$ is a $(q-3) \times 1$ vector of parameters. The log-likelihood function $\ell(\boldsymbol{\theta}) = \log L(\boldsymbol{\theta})$ with the model (3.1), where $\lambda_0(t) = \alpha$ for $t > 0$, is given by

$$\begin{aligned} \ell(\boldsymbol{\theta}) = & \sum_{i=1}^{m} \sum_{j=1}^{n_i} \log(\alpha) + \gamma \sum_{i=1}^{m} \sum_{j=1}^{n_i} N_i(t_{ij}^-) + \beta \sum_{i=1}^{m} \sum_{j=1}^{n_i} Z_i(t_{ij}) \\ & + \boldsymbol{\xi}' \sum_{i=1}^{m} \sum_{j=1}^{n_i} \boldsymbol{x}_i(t_{ij}) - \sum_{i=1}^{m} \int_0^\infty Y_i(u) \, \alpha \, \mathrm{e}^{\left[\gamma N_i(u^-) + \beta Z_i(u) + \boldsymbol{\xi}_i' \boldsymbol{x}(u)\right]} du, \end{aligned} \tag{3.2}$$

where $Y_i(t)$ is the at risk indicator at time t for the i^{th} process. Let $(\partial/\partial\boldsymbol{\theta})\ell(\boldsymbol{\theta})$ be the gradient of $\ell(\boldsymbol{\theta})$. Then $\boldsymbol{U}(\boldsymbol{\theta}) = [(\partial/\partial\boldsymbol{\theta})\ell(\boldsymbol{\theta})]$ is a $q \times 1$ vector of score functions with components

$$
\begin{aligned}
U_\alpha(\boldsymbol{\theta}) &= \frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha} = \sum_{i=1}^{m}\sum_{j=1}^{n_i} \frac{1}{\alpha} - \sum_{i=1}^{m} \int_0^\infty Y_i(t) \frac{1}{\alpha} \lambda_i \left[t|\mathcal{H}_i(t)\right] \mathrm{d}t, \\
U_\gamma(\boldsymbol{\theta}) &= \frac{\partial \ell(\boldsymbol{\theta})}{\partial \gamma} = \sum_{i=1}^{m}\sum_{j=1}^{n_i} N_i(t_{ij}^-) - \sum_{i=1}^{m} \int_0^\infty Y_i(t) N_i(t^-) \lambda_i \left[t|\mathcal{H}_i(t)\right] \mathrm{d}t, \\
U_\beta(\boldsymbol{\theta}) &= \frac{\partial \ell(\boldsymbol{\theta})}{\partial \beta} = \sum_{i=1}^{m}\sum_{j=1}^{n_i} Z_i(t_{ij}) - \sum_{i=1}^{m} \int_0^\infty Y_i(t) Z_i(t) \lambda_i \left[t|\mathcal{H}_i(t)\right] \mathrm{d}t, \\
\boldsymbol{U}_{\boldsymbol{\xi}}(\boldsymbol{\theta}) &= \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\xi}} = \sum_{i=1}^{m}\sum_{j=1}^{n_i} \boldsymbol{x}_i(t_{ij}) - \sum_{i=1}^{m} \int_0^\infty Y_i(t) \boldsymbol{x}_i(t) \lambda_i \left[t|\mathcal{H}_i(t)\right] \mathrm{d}t.
\end{aligned}
\tag{3.3}
$$

The maximum likelihood estimates $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ can be obtained by solving the system of score equations $\boldsymbol{U}(\boldsymbol{\theta}) = \mathbf{0}$, where $\mathbf{0}$ is a $q \times 1$ vector of zeros. It should be noted that the existence of $\hat{\boldsymbol{\theta}}$ as a unique maximizer of $\ell(\boldsymbol{\theta})$ is discussed by Cook and Lawless (2007, pp. 201–202). The condition required for it is that $\sum_{i=1}^{m} Y_i(t) \boldsymbol{W}_i^{**}(t) \boldsymbol{W}_i^{**\prime}(t)$, where $\boldsymbol{W}_i^{**}(t) = (1, N_i(t^-), Z_i(t), \boldsymbol{x}_i'(t))'$, should be positive definite for at least one value of $t > 0$.

We let $\boldsymbol{I}(\boldsymbol{\theta}) = [-(\partial^2/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}')\ell(\boldsymbol{\theta})]$ denote the $q \times q$ information matrix and $\boldsymbol{\mathcal{I}}(\boldsymbol{\theta}) = E(\boldsymbol{I}(\boldsymbol{\theta})) = E\left[-(\partial/\partial\boldsymbol{\theta}')\boldsymbol{U}(\boldsymbol{\theta})\right] = E\left[\boldsymbol{U}(\boldsymbol{\theta})\boldsymbol{U}'(\boldsymbol{\theta})\right]$ denote the $q \times q$ Fisher or expected information matrix, which is the covariance matrix of the score vector $\boldsymbol{U}(\boldsymbol{\theta})$. Model based variance estimates of the maximum likelihood estimators are available from $\boldsymbol{I}^{-1}(\boldsymbol{\theta})$ or $\boldsymbol{\mathcal{I}}^{-1}(\boldsymbol{\theta})$ by replacing $\boldsymbol{\theta}$ by $\hat{\boldsymbol{\theta}}$. As discussed in Section 3.1.1, under mild regularity conditions, $\hat{\boldsymbol{\theta}}$ is a consistent estimator of $\boldsymbol{\theta}$ and the asymptotic distribution of the $q \times 1$ random vector $\sqrt{m}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$, as $m \to \infty$, is a q-dimensional multivariate normal distribution with $q \times 1$ vector of zeros and $q \times q$ covariance matrix $\boldsymbol{\mathcal{I}}^{-1}(\boldsymbol{\theta}_0)$, where $\boldsymbol{\theta_0}$ is the $q \times 1$ vector of true values of parameters; that is, in notation, $\sqrt{m}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \sim N_q(\mathbf{0}, \boldsymbol{\mathcal{I}}^{-1}(\boldsymbol{\theta}_0))$ as $m \to \infty$, under mild regularity conditions.

An alternative to the model (3.1) which simultaneously includes trends due to number of previous events and carryover effects is

$$
\lambda[t|\mathcal{H}(t)] = \lambda_0(t) + \gamma\, N(t^-) + \beta\, Z(t) + \boldsymbol{\xi}'\, \boldsymbol{x}(t), \qquad t > 0. \tag{3.4}
$$

The model in (3.4) is of additive form. Since the intensity function $\lambda[t|\mathcal{H}(t)]$ is a positive quantity, γ and β in (3.4) should also be restricted to positive values. Because of that reason, it assumes only a positive monotonic trend due to the number of previous events and event clustering in the data, whereas the model (3.1) can capture both positive ($\gamma > 0$) and negative ($\gamma < 0$) trends, as well as event clustering ($\beta > 0$) and event sparsity ($\beta < 0$) in the data. The model (3.4) or its ramifications can, however, be useful in some applications (Section 8.3.2 in Aalen et al., 2008; Simpson, 2013).

A rather technical issue related to the model (3.1) is that it may be *dishonest* because of the dynamic covariate $N(t^-)$. A definition of dishonest processes can be found in Cox and Miller (1965, p. 163). Dishonest models may subject to explosion in a finite time interval. This issue has been discussed by Aalen et al. (2008, Section 8.6.3) in the context of dynamic models. We conducted a small simulation study to illustrate this issue with the model (3.1). The details of the simulation study can be found in Appendix B. The results of this study shows that the possibility of explosion is increasing as the values of β, Δ and/or τ increase only when the $\gamma > 0$. A necessary and sufficient condition for models to be nonexplosive is called *the Feller condition*, which can be applied in our case as follows. Let the intensity function only depend on $N(t^-)$; that is, $\lambda_i[t|\mathcal{H}_i(t)] = g(N_i(t^-))$ for a nonnegative function g and $t > 0$. Then, the process is nonexplosive on finite intervals if and only if $\sum_{k=1}^{\infty} g(k)^{-1}$ diverges (Feller, 1968, p. 453). For an exponential form $\lambda_i[t|\mathcal{H}_i(t)] = \exp(\gamma N_i(t^-))$, it is clear that $\sum_{k=1}^{\infty} \exp(-\gamma k)$ converges for all positive values of γ, which implies that there is a positive probability for an explosion on finite time intervals. If γ takes negative values, then models with dynamic covariate $N(t^-)$ are honest. In this study, we do not impose any restriction on the values that γ can take. However, we can achieve an honest process for all values of γ by making some adjustment in the function g. We use the trimmed version of $N(t^-)$ in this study to overcome the issue of dishonest behaviour of the event generating process. That is, we choose $g(N(t^-)) = \exp(\gamma N^*(t^-))$, where

$$N^*(t^-) = N(t^-)\, I[N(t^-) \leq c] + c\, I[N(t^-) > c], \qquad t > 0, \tag{3.5}$$

and the cutoff point c is a prespecified positive integer. It is clear that $\sum_{k=1}^{\infty} \exp\{-\gamma\, [k\, I(k \leq c) + c\, I(k > c)]\}$ diverges for any positive finite integer c. The value of cutoff point c may depend on the context of the study. We would like to note that the issue of explosion is not too restrictive for the settings considered in this thesis.

This is because, the observed number of events per individual in the applications considered in our study is small as in many medical and epidemiological studies. Nonetheless, we replace $N(t^-)$ with $N^*(t^-)$ in model (3.1), and take $c = 20$ just to guarantee non-explosion. It should be noted that Cook and Lawless (2013) encounter a similar problem, and they choose $c = 20$ to analyze trends due to $N(t^-)$ in a medical data set. There are other choices for the function $g(N(t^-))$ such as $\exp(\gamma N(t^-)/t)$ or $\exp(\gamma \log N(t^-))$ (Aalen et al., 2008, Section 8.6.3). However, the practical interpretation of their corresponding parameter γ is not as straightforward as that of γ in $\exp(\gamma N(t^-))$ to understand the trend due to number of previous events in recurrent event processes.

3.1.1 Large Sample Properties

The large sample properties of the maximum likelihood estimators in the context of counting processes require five regularity conditions (Conditions A–E), which are available in Andersen et al. (1993, pp. 420–421). For the completeness of the discussion, we give these conditions in Appendix C. We would like to note that the following discussion requires basic knowledge from the theory of counting process martingales and stochastic integration. The required mathematical tools and concepts can be found in Andersen et al. (1993, Chapter 2).

Conditions A and E are regularity conditions concerning the continuity, boundedness and convergence of log-likelihood derivatives, similar to those Cramér type regularity conditions found in the classical case (see Cox and Hinkley, 1974, p. 281). Condition A states that the log-likelihood function is three times differentiable with respect to model parameters by interchanging the order of integration and differentiation. This condition is required for the Taylor series expansion of the score vector $\boldsymbol{U}(\boldsymbol{\theta})$ for any $\boldsymbol{\theta}$ in the parameter space around the q-dimensional vector of true values of parameter $\boldsymbol{\theta_0}$. Condition B ensures that the variances (predictable variation processes) of score functions

$$a_m^{-2} \sum_{i=1}^{m} \int_0^\infty \left\{ \frac{\partial}{\partial \theta_k} \lambda_i \left[u | \mathcal{H}_i(u); \boldsymbol{\theta_0} \right] \right\} \left\{ \frac{\partial}{\partial \theta_l} \lambda_i \left[u | \mathcal{H}_i(u); \boldsymbol{\theta_0} \right] \right\} \times Y_i(u) \lambda_i \left[u | \mathcal{H}_i(u); \boldsymbol{\theta_0} \right] du, \tag{3.6}$$

converge, in probability to deterministic functions $\sigma_{kl}(\boldsymbol{\theta_0})$, $(k, l = 1, 2, \ldots, q)$, for some sequence $(a_m)_{m=1}^{\infty}$ of positive constants increasing to infinity as $m \to \infty$. Typically, the normalizing constants a_m^2 can be considered as the sample size m.

Condition C is required to show that the jumps of martingales or stochastic integrals with respect to these martingales approach zero as the normalizing constants a_m approaches infinity. That is, for all $\epsilon > 0$ and $k = 1, \ldots, q$, the sum

$$a_m^{-2} \sum_{i=1}^{m} \int_0^{\infty} \left\{ \frac{\partial}{\partial \theta_k} \lambda_i \left[u | \mathcal{H}_i(u); \boldsymbol{\theta_0} \right] \right\}^2 I \left(\left| \frac{\partial}{\partial \theta_k} \lambda_i \left[u | \mathcal{H}_i(u); \boldsymbol{\theta_0} \right] \right| > a_m \epsilon \right) \times Y_i(u) \lambda_i \left[u | \mathcal{H}_i(u); \boldsymbol{\theta_0} \right] du, \tag{3.7}$$

should converge in probability to 0 as $m \to \infty$. Condition D is that the matrix constructed with the elements $\sigma_{kl}(\boldsymbol{\theta_0})$ defined in Condition B should be positive definite. Condition E is regarding the boundedness of the third derivative of the log-likelihood function. This condition needs to be satisfied in order to show that the remainder term in a Taylor expansion of log-likelihood function is negligible.

It should be noted that Conditions A to E should be verified separately for every model considered in order to obtain large sample properties of the maximum likelihood estimators. We now verify them for a specific case of the model discussed in the previous section. Since our primary goal is to investigate the parameters related to internal covariates, we ignore external covariates for brevity. However, results can be extended to that case as well. The intensity function of the counting process $\{N_i^*(t); t \geq 0\}$, $i = 1, \ldots, m$, as defined in (3.5), with the parameter vector $\boldsymbol{\theta} = (\alpha, \beta, \gamma)'$ can be written as

$$\lambda_i \left[t | \mathcal{H}_i(t); \boldsymbol{\theta} \right] = \alpha \exp \left\{ \beta Z_i(t) + \gamma N_i^*(t^-) \right\}, \qquad t \geq 0, \qquad i = 1, \ldots, m, \tag{3.8}$$

where $\alpha > 0$, $\beta \in \mathbb{R}$ and $\gamma \in \mathbb{R}$ are the parameters to be estimated. Let $\bar{N}_i^*(t) = \int_0^t Y_i(u) dN_i^*(u)$, where $Y_i(u)$ is the at-risk indicator. The corresponding intensity function of the observed counting process $\left\{ \bar{N}_i^*(t); t \geq 0 \right\}$, $i = 1, \ldots, m$, is then

$$\bar{\lambda}_i \left[t | \bar{\mathcal{H}}_i(t); \boldsymbol{\theta} \right] = Y_i(t) \alpha \exp \left\{ \beta Z_i(t) + \gamma N_i^*(t^-) \right\}, \qquad t \geq 0, \qquad i = 1, \ldots, m, \tag{3.9}$$

where $\bar{\mathcal{H}}_i(t) = \left\{ \bar{N}_i(s), Y_i(s); 0 \leq s < t \right\}$. We assume that the predictable at-risk process $\{Y_i(t); t \geq 0\}$ and $\{N_i^*(t); t \geq 0\}$ are conditionally independent given the history. The likelihood obtained from m independent individuals is

$$L(\alpha,\beta,\gamma)=\prod_{i=1}^{m}\left[\prod_{j=1}^{n_i}\alpha \mathrm{e}^{\left\{\beta Z_i(t_{ij})+\gamma N_i^*(t_{ij}^-)\right\}}\right]\mathrm{e}^{-\int_0^\tau Y_i(u)\alpha_0 \mathrm{e}^{\left\{\beta Z_i(u)+\gamma N_i^*(u^-)\right\}}du}, \tag{3.10}$$

where τ is the maximum of follow-up times of all the individuals; i.e. $\tau = \max\{\tau_1,\ldots,\tau_m\}$. The likelihood function (3.10) is constructed using the observed data $\{(t_{i1},\ldots,t_{in_i};\tau_i); i=1,\ldots,m\}$ on the joint multivariate process $\left\{\bar{N}_i^*(t), Y_i(t); i=1,\ldots,m,\ t\geq 0\right\}$ as discussed in Section 2.1.

The first three elements U_α, U_γ and U_β in the score vector defined in (3.3) are differentiable with respect to α, γ and β, respectively. Therefore, Condition A is satisfied for the model (3.8). In order to show that Condition B is fulfilled for the model (3.8), we need to show that, as $m\to\infty$,

$$\frac{1}{m}\sum_{i=1}^{m}\int_0^\tau Y_i(t)\frac{1}{\alpha_0}\ \exp\left\{\beta_0 Z_i(t)+\gamma_0 N_i^*(t^-)\right\}dt \xrightarrow{p} \sigma_{\alpha\alpha}, \tag{3.11}$$

$$\frac{1}{m}\sum_{i=1}^{m}\int_0^\tau Y_i(t)\ Z_i(t)^2\ \alpha_0\ \exp\left\{\beta_0 Z_i(t)+\gamma_0 N_i^*(t^-)\right\}dt \xrightarrow{p} \sigma_{\beta\beta}, \tag{3.12}$$

$$\frac{1}{m}\sum_{i=1}^{m}\int_0^\tau Y_i(t)\ N_i^*(t^-)^2\ \alpha_0\ \exp\left\{\beta_0 Z_i(t)+\gamma_0 N_i^*(t^-)\right\}dt \xrightarrow{p} \sigma_{\gamma\gamma}, \tag{3.13}$$

$$\frac{1}{m}\sum_{i=1}^{m}\int_0^\tau Y_i(t)\ Z_i(t)\ \exp\left\{\beta_0 Z_i(t)+\gamma_0 N_i^*(t^-)\right\}dt \xrightarrow{p} \sigma_{\alpha\beta}, \tag{3.14}$$

$$\frac{1}{m}\sum_{i=1}^{m}\int_0^\tau Y_i(t)\ N_i^*(t^-)\ \exp\left\{\beta_0 Z_i(t)+\gamma_0 N_i^*(t^-)\right\}dt \xrightarrow{p} \sigma_{\alpha\gamma}, \tag{3.15}$$

$$\frac{1}{m}\sum_{i=1}^{m}\int_0^\tau Y_i(t)\ Z_i(t)\ N_i^*(t^-)\ \alpha_0\ \exp\left\{\beta_0 Z_i(t)+\gamma_0 N_i^*(t^-)\right\}dt \xrightarrow{p} \sigma_{\beta\gamma}, \tag{3.16}$$

where the notation $\xrightarrow{p}$ represents convergence in probability. The integrals on the left hand sides of all the convergence conditions given from (3.11) to (3.16) above have integrands in the form of step functions. Therefore, they can be rewritten as sums of rectangles. For example, when τ is sufficiently large for the underlying counting process $\{N_i^*(t); t\geq 0\}$, $i=1,\ldots,m$, the integral $\int_0^\tau \exp\{\beta_0 Z_i(t)+\gamma_0 N_i^*(t^-)\}dt$ is almost surely equivalent to

$$W_{i1}+\sum_{k=2}^{N_i(\tau)+1}\mathrm{e}^{((k\wedge c)-1)\gamma_0}\left\{W_{ik}+\left(\mathrm{e}^{\beta_0}-1\right)(W_{ik}\wedge\Delta)\right\}, \tag{3.17}$$

where $a\wedge b=\min\{a,b\}$ and c is defined in (3.5). Similarly, for the observable process $\left\{\left(\bar{N}_i(t),Y_i(t)\right);i=1,\ldots,m;t\geq 0\right\}$, the integral $\int_0^\tau Y_i(t)\exp\{\beta_0 Z_i(t)+\gamma_0 N_i^*(t^-)\}dt$ can be expressed as

$$\widetilde{W}_{i1}+\sum_{k=2}^{\bar{N}_i(\tau)+1}\mathrm{e}^{((k\wedge c)-1)\gamma_0}\left\{\widetilde{W}_{ik}+\left(\mathrm{e}^{\beta_0}-1\right)(\widetilde{W}_{ik}\wedge\Delta)\right\}, \tag{3.18}$$

almost surely, where $\widetilde{W}_{ik}=\int_{t_{i(k-1)}}^{t_{i(k-1)}+W_{ik}}Y_i(t)dt$. Under the conditionally independent assumption, following the weak law of large numbers, we can easily show that, as $m\to\infty$, the random variable (3.18) converges in probability to

$$E(\widetilde{W}_1)+\sum_{k=2}^{E(\bar{N}(\tau))+1}\mathrm{e}^{((k\wedge c)-1)\gamma_0}\left\{E(\widetilde{W}_k)+\left(\mathrm{e}^{\beta_0}-1\right)E(\widetilde{W}_k\wedge\Delta)\right\}, \tag{3.19}$$

where $E(\widetilde{W}_k)=\int_{t_{(k-1)}}^{t_{(k-1)}+E(W_k)}E(Y(t))dt$, $E(\widetilde{W}_k\wedge\Delta)=\int_{t_{(k-1)}}^{t_{(k-1)}+E(W_k\wedge\Delta)}E(Y(t))dt$ and

$$\begin{aligned}E(W_k)&=\frac{\mathrm{e}^{\Delta\alpha_0\mathrm{e}^{\{(k-1)\gamma_0+\beta_0 I(k>1)\}}}-1+\mathrm{e}^{\beta_0 I(k>1)}}{\alpha_0\ \exp\left[(k-1)\gamma_0+\beta_0 I(k>1)+\Delta\alpha_0\mathrm{e}^{\{(k-1)\gamma_0+\beta_0 I(k>1)\}}\right]},\\E(W_k\wedge\Delta)&=\frac{\mathrm{e}^{\Delta\alpha_0\mathrm{e}^{\{(k-1)\gamma_0+\beta_0 I(k>1)\}}}-1+\Delta-\alpha_0\Delta\mathrm{e}^{\{(k-1)\gamma_0+\beta_0 I(k>1)\}}}{\alpha_0\ \exp\left[(k-1)\gamma_0+\beta_0 I(k>1)+\Delta\alpha_0\mathrm{e}^{\{(k-1)\gamma_0+\beta_0 I(k>1)\}}\right]}.\end{aligned} \tag{3.20}$$

Therefore, we obtain that the convergence result stated in (3.11) is equivalent to

$$\sigma_{\alpha\alpha}=\frac{E(\widetilde{W}_1)}{\alpha_0}+\frac{1}{\alpha_0}\sum_{k=2}^{E(\bar{N}(\tau))+1}\left\{E(\widetilde{W}_k)+\left(\mathrm{e}^{\beta_0}-1\right)E(\widetilde{W}_k\wedge\Delta)\right\}\mathrm{e}^{((k\wedge c)-1)\gamma_0}. \tag{3.21}$$

Similarly, we can show that

$$\sigma_{\gamma\gamma} = \alpha_0 E(\widetilde{W}_1) + \sum_{k=2}^{E(\bar{N}(\tau))+1} \alpha_0((k \wedge c) - 1)^2 \left\{E(\widetilde{W}_k) + \left(e^{\beta_0} - 1\right) E(\widetilde{W}_k \wedge \Delta)\right\} e^{((k\wedge c)-1)\gamma_0}, \tag{3.22}$$

$$\sigma_{\alpha\gamma} = E(\widetilde{W}_1) + \sum_{k=2}^{E(\bar{N}(\tau))+1} ((k \wedge c) - 1) \left\{E(\widetilde{W}_k) + \left(e^{\beta_0} - 1\right) E(\widetilde{W}_k \wedge \Delta)\right\} e^{((k\wedge c)-1)\gamma_0}, \tag{3.23}$$

$$\sigma_{\beta\beta} = \alpha_0 e^{\beta_0} \sum_{k=2}^{E(\bar{N}(\tau))+1} E(\widetilde{W}_k \wedge \Delta) e^{((k\wedge c)-1)\gamma_0}, \tag{3.24}$$

$$\sigma_{\alpha\beta} = e^{\beta_0} \sum_{k=2}^{E(\bar{N}(\tau))+1} E(\widetilde{W}_k \wedge \Delta) e^{((k\wedge c)-1)\gamma_0}, \tag{3.25}$$

and

$$\sigma_{\beta\gamma} = \alpha_0 e^{\beta_0} \sum_{k=2}^{E(\bar{N}(\tau))+1} ((k \wedge c) - 1) E(\widetilde{W}_k \wedge \Delta) e^{((k\wedge c)-1)\gamma_0}. \tag{3.26}$$

Combining the results in (3.11) – (3.16), the convergence can be rewritten as

$$\frac{1}{m} \sum_{i=1}^{m} \int_0^\tau Y_i(t)\ \boldsymbol{A}_i(t) \alpha_0\ \exp\left\{\beta_0 Z_i(t) + \gamma_0 N_i^*(t^-)\right\} dt \xrightarrow{p} \boldsymbol{\Sigma}, \tag{3.27}$$

where $\boldsymbol{A}_i(t) = \boldsymbol{b}_i(t)\boldsymbol{b}_i'(t)$, $\boldsymbol{b}_i(t) = (1/\alpha_0, Z_i(t), N_i^*(t^-))'$ and

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{\alpha\alpha} & \sigma_{\alpha\beta} & \sigma_{\alpha\gamma} \\ \sigma_{\alpha\beta} & \sigma_{\beta\beta} & \sigma_{\beta\gamma} \\ \sigma_{\alpha\gamma} & \sigma_{\beta\gamma} & \sigma_{\gamma\gamma} \end{bmatrix}. \tag{3.28}$$

For any 3-dimensional vector $\boldsymbol{a}$ such that $\boldsymbol{a}$ contains all non-zero values, the quadratic form, $\boldsymbol{a}'\boldsymbol{A}_i(t)\boldsymbol{a} > 0$. Therefore, the matrix $\boldsymbol{\Sigma}$ is a positive definite matrix, which satisfies Condition D.

For Condition C, it is obvious that $I(m^{-1/2} > \epsilon)$ converges to 0 as $m \to \infty$ for any constant $\epsilon > 0$. Thus, we can show that

$$\frac{1}{m}\int_0^\tau \sum_{i=1}^m \frac{1}{\alpha_0^2} I\left(\left|\frac{1}{\sqrt{m}}\frac{1}{\alpha_0}\right| > \epsilon\right) Y_i(s)\alpha_0 \exp\left\{\beta_0 Z_i(s) + \gamma_0 N_i^*(s^-)\right\} ds \xrightarrow{p} 0, \tag{3.29}$$

$$\frac{1}{m}\int_0^\tau \sum_{i=1}^m Z_i(s)^2 I\left(\left|\frac{1}{\sqrt{m}} Z_h(s)\right| > \epsilon\right) Y_i(s)\alpha_0 \exp\left\{\beta_0 Z_i(s) + \gamma_0 N_i^*(s^-)\right\} ds \xrightarrow{p} 0, \tag{3.30}$$

$$\frac{1}{m}\int_0^\tau \sum_{i=1}^m N_i^*(s^-)^2 I\left(\left|\frac{1}{\sqrt{m}} N_i^*(s^-)\right| > \epsilon\right) Y_i(s)\alpha_0 \exp\left\{\beta_0 Z_i(s) + \gamma_0 N_i^*(s^-)\right\} ds \xrightarrow{p} 0, \tag{3.31}$$

as $m \to \infty$. The convergence results in (3.29) – (3.31) prove that the Condition C is satisfied for the model (3.8).

To show that Condition E is satisfied, we need to introduce the following weak assumption. For any fixed $\tau > 0$,

$$\frac{1}{m}\sum_{i=1}^m \int_0^t Y_i(u)\, du \xrightarrow{p} r(t), \qquad 0 \le t \le \tau, \tag{3.32}$$

as $m \to \infty$, where $r(t)$ is a positive constant for any $t \in [0, \tau]$. Now suppose that $M_1 \le \alpha \le M_2$ for some $M_2 > M_1 > 0$, and $\beta \le M_3$ and $\gamma \le M_4$ for some $M_3, M_4 \in \mathbb{R}$. The required supremum norms of $\frac{\partial^3}{\partial\theta_j\partial\theta_l\partial\theta_r}\lambda_i\left[t|\mathcal{H}_i(t);\boldsymbol{\theta}\right]$ and $\frac{\partial^3}{\partial\theta_j\partial\theta_l\partial\theta_r}\log\lambda_i\left[t|\mathcal{H}_i(t);\boldsymbol{\theta}\right]$ are bounded by $2/M_1^3$ and $M_2c^2\exp(M_3 + cM_4)$ respectively for any $\theta_j, \theta_l, \theta_r \in \{\alpha, \beta, \gamma\}$, which do not depend on any parameter in $\{\alpha, \beta, \gamma\}$ for any m. Consequently, using (3.11) and (3.32), we have

$$\frac{1}{m}\sum_{i=1}^m \int_0^\tau M_2c^2\exp(M_3 + cM_4)Y_i(u)du \xrightarrow{p} M_2c^2\exp(M_3 + cM_4)r(\tau), \tag{3.33}$$

$$\frac{1}{m}\sum_{i=1}^m \int_0^\tau 1/\alpha_0^4 Y_i(u)\alpha_0 \exp\left\{\beta Z_i(u) + N_i^*(u^-)\right\} du \xrightarrow{p} \alpha_0^{-2}\sigma_{\alpha\alpha}, \tag{3.34}$$

$$\frac{1}{m}\sum_{i=1}^{m}\int_0^{\tau} 2/M_1^3 Y_i(u)\alpha_0 \exp\left\{\beta_0 Z_i(u)+\gamma_0 N_i^*(u^-)\right\} du \xrightarrow{p} (2\alpha_0^2/M_1^3)\sigma_{\alpha\alpha}, \tag{3.35}$$

as $m \to \infty$. For any $\epsilon > 0$,

$$\frac{1}{m}\sum_{i=1}^{m}\int_0^{\tau} 2/M_1^3 I\left\{\frac{\sqrt{2/M_1^3}}{\sqrt{m}} > \epsilon\right\} Y_i(u)\alpha_0 \exp\left\{\beta_0 Z_i(u)+\gamma_0 N_i^*(u^-)\right\} du \xrightarrow{p} 0 \tag{3.36}$$

as $m \to \infty$. These convergence results conclude that Condition E is satisfied for the model in (3.8). Hence, the model in (3.8) satisfies all the conditions (A–E) specified in Andersen et al. (1993, pp. 420–421) to obtain the asymptotic properties of the maximum likelihood estimator $\hat{\boldsymbol{\theta}} = (\hat{\alpha}, \hat{\beta}, \hat{\gamma})$. That is, the vector of maximum likelihood estimators $\hat{\boldsymbol{\theta}}$ is a consistent estimator of $\boldsymbol{\theta_0}$, and $\sqrt{m}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$ convergences in distribution to a 3-dimensional multivariate Normal distribution with the 3×1 mean zero vector and 3×3 covariance matrix $\boldsymbol{\Sigma}$, where $\boldsymbol{\Sigma}$ is defined in (3.28) as $m \to \infty$. The covariance matrix $\boldsymbol{\Sigma}$ can be consistently estimated by plugging in the maximum likelihood estimators in $\boldsymbol{\Sigma}$; that is, $\boldsymbol{\Sigma}(\hat{\alpha}, \hat{\beta}, \hat{\gamma})$ in notation. The proofs of these assertions can be found in Andersen et al. (1993, Section 6.1.2). The likelihood ratio and score statistics enjoy the usual large sample properties. For example, as $m \to \infty$, the score vector under $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ satisfies

$$\frac{1}{\sqrt{m}}\boldsymbol{U}(\boldsymbol{\theta}_0) \xrightarrow{\mathcal{D}} N_3(\mathbf{0}, \boldsymbol{\Sigma}(\boldsymbol{\theta}_0)), \tag{3.37}$$

where the notation $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution. A score test can be then developed from (3.37) for testing $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$. Similarly, partial score tests can also be developed for $H_0 : \beta = \beta_0;\ \alpha \in \mathbb{R}^+, \gamma \in \mathbb{R}$ and $H_0 : \gamma = \gamma_0;\ \alpha \in \mathbb{R}^+, \gamma \in \mathbb{R}$.

3.2 Extensions of the Model to Deal with Unexplained Heterogeneity

In recurrent event studies, including multiple events, there is often an excess heterogeneity in the event counts across individuals. Environmental factors and differences in the individual characteristics usually result in such a heterogeneity. If these variables are not included as a covariate in a model, there can be excess variation in the intensity functions across processes. As a result, assumptions of a common baseline intensity function for all processes become inadequate. We now consider an extension of the model (3.1) to deal with unexplained heterogeneity in a cohort. To this end, we introduce a random effects model with the intensity function

$$\lambda_i[t|\mathcal{H}_i(t),\nu_i] = \alpha\nu_i \exp\left[\gamma\, N_i(t^-) + \beta\, Z_i(t) + \boldsymbol{\xi}'\,\boldsymbol{x}(t)\right], \qquad i=1,\ldots,m,\ t>0, \tag{3.38}$$

where $\nu_1, \nu_2, \ldots, \nu_m$ are positive-valued i.i.d. unobservable random variables. In this study, we assume that the ν_i have a gamma distribution with mean 1 and variance ϕ. Let G and g denote the c.d.f and p.d.f. of the ν_i. As discussed in Section 2.3, the likelihood function for m independent individuals is then given by $L(\alpha,\beta,\gamma,\boldsymbol{\xi},\phi) = \prod_{i=1}^m L_i(\alpha,\beta,\gamma,\boldsymbol{\xi},\phi)$, where

$$L_i(\alpha,\beta,\gamma,\boldsymbol{\xi},\phi) = \int_0^\infty \left\{\prod_{j=1}^{n_i} \lambda_i\left[t_{ij}|\mathcal{H}_i(t_{ij}),\nu_i\right]\ \mathrm{e}^{-\int\limits_0^\infty Y_i(u)\lambda_i[u|\mathcal{H}_i(u),\nu_i]du}\right\} g(\nu_i;\phi)d\nu_i. \tag{3.39}$$

Note that, since the ν_i are unobservable, the likelihood contribution $L_i(\alpha,\beta,\gamma,\boldsymbol{\xi},\phi)$ in (3.39) is obtained by integrating out the random effect ν_i. The resulting log-likelihood function is given by

$$\begin{aligned}\ell(\alpha,\beta,\gamma,\boldsymbol{\xi},\phi) =& \sum_{i=1}^m \left[\log\{\Gamma(n_i+\phi^{-1})\} - \log\{\Gamma(\phi^{-1})\} + n_i\log(\phi)\right]\\ &+ n_{..}\log\alpha + \beta\sum_{i=1}^m\sum_{j=1}^{n_i} Z_i(t_{ij}) + \gamma\sum_{i=1}^m\sum_{j=1}^{n_i} N_i(t_{ij}^-) + \boldsymbol{\xi}'\sum_{i=1}^m\sum_{j=1}^{n_i}\boldsymbol{x}_i(t_{ij})\\ &- \sum_{i=1}^m (n_i+\phi^{-1})\log\left(1+\phi\int\limits_0^\tau \alpha\mathrm{e}^{\left\{\beta Z_i(u)+\gamma N_i(u^-)+\boldsymbol{\xi}'\boldsymbol{x}_i(u)\right\}}du\right),\end{aligned} \tag{3.40}$$

where $n_{..} = \sum_{i=1}^m n_i$. We maximize the log-likelihood function (3.40) with the `nlminb` function in `optimx` `R` package to obtain the estimates of the parameters α, β, γ, ϕ and $\boldsymbol{\xi}$, and their standard errors.

It should be noted that random effects models to address the unexplained heterogeneity may cause unnatural clustering of events over time. This issue is especially important when the goal of a study is to estimate the effects of dynamic covariates because random effects models may lead to wrong conclusions in such situations. To discuss this issue, we now consider the random effects model (3.38) without carryover effects; that is,

$$\lambda_i[t|\mathcal{H}_i(t), \nu_i] = \nu_i \, \exp\left[\gamma \, N_i(t^-)\right], \qquad i = 1, \ldots, m, \; t > 0. \tag{3.41}$$

After a little algebra, it can be shown that the marginal intensity function is of the form

$$\lambda_i[t|\mathcal{H}_i(t)] = \frac{[1 + \phi N_i(t^-)] \exp\left[\gamma \, N_i(t^-)\right]}{1 + \phi \int\limits_0^t \exp\left[\gamma \, N_i(u^-)\right] du}, \qquad t > 0. \tag{3.42}$$

The derivation of such a marginal intensity function in (3.42) for the general case is available in Appendix A. In (3.42), for $\gamma > 0$, the numerator increases at each event occurrence because of the term $N_i(t^-)$. In other words, the intensity makes a jump up at each event occurrence time because of the terms in the numerator. After each jump, the intensity function smoothly decreases until the occurrence of the next event as t increases because of the integral term in the denominator (3.42). Carryover effects also cause similar type of behavior on the intensity function. Therefore, carryover effects and unexplained heterogeneity may confound in some settings. Cigsar and Lawless (2012) conducted a simulation study to investigate the effect of confounding of heterogeneity with carryover effects. Results of their study show that the carryover effect is over estimated when the heterogeneity is ignored on a model while it is present in the data. We also discuss the confounding issue under the model (3.42) through a simulation study in the next section. Our conclusion is similar to that of Cigsar and Lawless (2012). A lengthy discussion on the effects of unexplained heterogeneity in the context of survival analysis can be found in Aalen et al. (2008, Chapter 6).

3.3 Simulation Studies

Our goal in this section is to present the results of two simulation studies conducted to investigate confounding issues related to dynamic covariates. Since our main focus in this thesis is not on the dynamic models for event counts, we do not study the finite sample properties of the estimators of the model parameters, such as when the asymptotic normality satisfactorily holds. However, we provide some insight on this issue in the next chapter, where we introduce methods based on gap times.

We conducted our first simulation study to explore the issue of confounding between carryover effects and unexplained heterogeneity as briefly discussed in the previous section. In simulations, we assumed that everyone is observed over $[0, \tau]$, and we took three values of τ ($\tau = 1, 2$ and 5). For each τ, we generated 2,000 realizations of m ($m = 100$ and 200) processes with the intensity function (3.38), where we took $\beta = 0$ and $\boldsymbol{\xi} = \mathbf{0}$. That is, we generated realizations of the counting process $\{N_i(t); t \geq 0\}$ with the intensity function

$$\lambda_i[t|\mathcal{H}_i(t)] = \nu_i \exp\left[\gamma\, N_i(t^-)\right], \quad i = 1, \ldots, m,\ t > 0, \tag{3.43}$$

where we considered three values for γ ($\gamma = 0, 0.05$ and 0.095), and two levels of heterogeneity across m processes; a low level ($\phi = 0.1$) and a high level ($\phi = 2$). For each generated realization of m processes under the scenarios with combinations of (τ, m, γ, ϕ), we used the following model to fit the data.

$$\lambda_i[t|\mathcal{H}_i(t)] = \alpha\ \exp\left[\gamma\, N_i(t^-) + \beta I(N_i(t^-) > 0) I(B_i(t) \leq \Delta)\right], \quad i = 1, \ldots, m,\ t > 0, \tag{3.44}$$

where $B_i(t)$ is the backward recurrence time for the i^{th} process at time t and $\Delta = 0.05$. For each realization, we obtained estimates of the parameters in the model (3.44) by maximizing the corresponding log-likelihood function with the `optimx` package in R. The average of the estimates over 2,000 simulation runs and their corresponding empirical standard deviations (within parenthesis) are presented in Table 3.1.

The results in Table 3.1 shows that, as the heterogeneity parameter ϕ increases, the bias in the estimate of the carryover effect β increases. As a result, the fitted modulated Poisson process model (3.44) gives a wrong conclusion that there is a significant carryover effect present in the data. For example, when the model (3.44) fitted to the data generated with $\phi = 2$, $\beta = 0$, $\gamma = 0.095$ and $\tau = 5$, the empirical means of estimates of β (= 0.290 and 0.301 when $m = 100$ and 200, respectively) do not include zero within its two standard deviations interval. For the corresponding scenarios when $\phi = 0.1$, the empirical means of estimates of β (= 0.028 and 0.038 when $m = 100$ and 200, respectively) include zero within its 2 standard deviations interval. It is also noted from the results in Table 3.1 that the bias in the estimates of β is getting larger as the sample size m increases.

Table 3.1: Empirical means and standard deviations (within parenthesis) of parameter estimates in the model (3.44) based on 2,000 simulation runs; data generated from the model (3.43) over $[0, \tau]$.

Estimates	$\tau = 1$		$\tau = 2$		$\tau = 5$	
	m=100	m=200	m=100	m=200	m=100	m=200
Data generated with $\phi = 0.1$, $\beta = 0$ and $\gamma = 0$						
$\bar{\hat{\alpha}}$	0.972 (0.120)	0.966 (0.080)	0.944 (0.090)	0.938 (0.062)	0.897 (0.065)	0.894 (0.046)
$\bar{\hat{\gamma}}$	0.050 (0.129)	0.062 (0.092)	0.054 (0.058)	0.058 (0.041)	0.041 (0.019)	0.042 (0.013)
$\bar{\hat{\beta}}_{\Delta=0.05}$	-0.177 (1.491)	-0.013 (0.356)	-0.032 (0.354)	0.006 (0.234)	0.013 (0.211)	0.035 (0.140)
Data generated with $\phi = 0.1$, $\beta = 0$ and $\gamma = 0.05$						
$\bar{\hat{\alpha}}$	0.972 (0.115)	0.966 (0.081)	0.945 (0.086)	0.939 (0.063)	0.903 (0.060)	0.899 (0.041)
$\bar{\hat{\gamma}}$	0.100 (0.116)	0.111 (0.083)	0.100 (0.051)	0.106 (0.035)	0.084 (0.013)	0.085 (0.009)
$\bar{\hat{\beta}}_{\Delta=0.05}$	-0.161 (1.349)	-0.011 (0.340)	-0.010 (0.322)	0.006 (0.215)	0.028 (0.167)	0.034 (0.115)
Data generated with $\phi = 0.1$, $\beta = 0$ and $\gamma = 0.095$						
$\bar{\hat{\alpha}}$	0.973 (0.116)	0.966 (0.080)	0.947 (0.085)	0.943 (0.061)	0.912 (0.055)	0.910 (0.038)
$\bar{\hat{\gamma}}$	0.142 (0.113)	0.154 (0.075)	0.142 (0.041)	0.146 (0.028)	0.122 (0.008)	0.122 (0.006)
$\bar{\hat{\beta}}_{\Delta=0.05}$	-0.108 (1.098)	-0.005 (0.314)	-0.004 (0.288)	0.011 (0.198)	0.028 (0.120)	0.038 (0.085)
Data generated with $\phi = 2$, $\beta = 0$ and $\gamma = 0$						
$\bar{\hat{\alpha}}$	0.732 (0.117)	0.742 (0.087)	0.688 (0.102)	0.694 (0.070)	0.571 (0.066)	0.572 (0.046)
$\bar{\hat{\gamma}}$	0.337 (0.070)	0.321 (0.055)	0.208 (0.039)	0.198 (0.027)	0.130 (0.010)	0.129 (0.007)
$\bar{\hat{\beta}}_{\Delta=0.05}$	0.246 (0.414)	0.310 (0.282)	0.296 (0.279)	0.331 (0.188)	0.281 (0.158)	0.299 (0.110)
Data generated with $\phi = 2$, $\beta = 0$ and $\gamma = 0.05$						
$\bar{\hat{\alpha}}$	0.780 (0.135)	0.793 (0.097)	0.767 (0.108)	0.773 (0.075)	0.604 (0.067)	0.603 (0.048)
$\bar{\hat{\gamma}}$	0.316 (0.079)	0.290 (0.063)	0.191 (0.029)	0.184 (0.017)	0.153 (0.009)	0.153 (0.006)
$\bar{\hat{\beta}}_{\Delta=0.05}$	0.269 (0.606)	0.358 (0.281)	0.266 (0.245)	0.285 (0.168)	0.216 (0.123)	0.233 (0.092)
Data generated with $\phi = 2$, $\beta = 0$ and $\gamma = 0.095$						
$\bar{\hat{\alpha}}$	0.838 (0.149)	0.860 (0.109)	0.825 (0.114)	0.826 (0.074)	0.620 (0.068)	0.616 (0.046)
$\bar{\hat{\gamma}}$	0.284 (0.085)	0.251 (0.058)	0.197 (0.019)	0.194 (0.010)	0.180 (0.008)	0.180 (0.005)
$\bar{\hat{\beta}}_{\Delta=0.05}$	0.380 (0.422)	0.465 (0.283)	0.256 (0.217)	0.278 (0.148)	0.290 (0.128)	0.301 (0.092)

Another issue is that the parameter γ representing the trend due to the number of previous events is significantly overestimated when τ and/or ϕ increases. For example, when the model (3.44) was fitted to the data generated with $m = 200$, $\beta = 0$, $\gamma = 0.095$ and $\tau = 5$, the empirical means of estimates of γ (= 0.122 and 0.180 when $\phi = 0.1$ and 2, respectively) increases as ϕ increases. The results given in Table 3.1, therefore, indicate that when the unexplained heterogeneity in dynamic models for event counts is ignored, significant bias in the estimates of dynamic covariates may occur in some settings. As a consequence, we recommend conducting a detailed analysis for the presence of excess heterogeneity if there is an interest in modelling dynamic covariates.

We next present the design and results of our second simulation study conducted to check whether carryover effects are confounded with trends due to $N(t^{-})$. If the term $N(t^{-})$ is of interest along with carryover effects, then carryover effects may confound with an increasing trend due to $N(t^{-})$ as well. This issue can be explained as follows. As the observation period increases, models with an increasing trend due to $N(t^{-})$ eventually start generating events in which the expectation of gap times will be shorter than a given carryover effect period. In such cases, models based on modulated Poisson processes may spuriously reveal the presence of carryover effects. This is also an issue for the gap time models discussed in Chapter 4 because carryover effects may become unidentifiable in such cases. Our simulation study presented here also shows some insight when such a case could happen. We explore this phenomenon through simulation studies in two settings; (i) identical processes, and (ii) nonidentical processes. To prevent explosion in the number of events, instead of using $N_i(t^{-})$, we used the term $N_i^*(t^{-})$ as defined in (3.5), where we fixed the value of c at 20 throughout this section. We considered three models to generate the data. These models define the scenarios explained below.

In the first scenario, we simulated 2,000 runs of the model

$$\lambda_i[t|\mathcal{H}_i(t)] = \alpha, \qquad t > 0, \tag{3.45}$$

for each combination of m $(= 10, 50, 100$ and $250)$ and τ $(= 1, 2$ and $5)$. We took $\alpha = 2$ in the model (3.45). Thus, the data generated did not include carryover effects or trend. As discussed in Section 3.2, we used `optimx` package in `R` to fit the following four models with the same generated data for each of 2,000 simulation runs.

M1: Model 1 is a homogeneous Poisson model with the intensity function $\lambda_i[t|\mathcal{H}_i(t)] = \alpha$, $i = 1, \ldots, m$,

M2: Model 2 is a carryover effects model with the intensity function $\lambda_i[t|\mathcal{H}_i(t)] = \alpha \exp[\beta\, Z_i(t)]$, $i = 1, \ldots, m$,

M3: Model 3 is a trend model with the intensity function $\lambda_i[t|\mathcal{H}_i(t)] = \alpha \exp[\gamma\, N_i^*(t^{-})]$, $i = 1, \ldots, m$, and

M4: Model 4 is a hybrid model with the intensity function $\lambda_i[t|\mathcal{H}_i(t)] = \alpha \exp[\gamma\, N_i^*(t^{-}) + \beta\, Z_i(t)]$, $i = 1, \ldots, m$.

We fixed Δ at 0.02 in the models M2 and M4. Note that, in this scenario, the first model (M1) is the correctly specified model and other models (M2, M3 and M4) are misspecified. The empirical means of the parameter estimates based on 2,000 simulation runs and their empirical standard errors are reported for the four fitted models.

Table 3.2: Empirical means and standard deviations (within parenthesis) of parameter estimates of 4 different models when the data generated from a model with $\alpha = 2,\ \beta = 0$ and $\gamma = 0$ in (3.44) with 2,000 simulations.

Fitted		Simulation											
		$\tau = 1$				$\tau = 2$				$\tau = 5$			
		m=10	m=50	m=100	m=250	m=10	m=50	m=100	m=250	m=10	m=50	m=100	m=250
M1	$\bar{\hat{\alpha}}$	2.001	2.002	1.998	1.999	1.997	1.999	1.999	1.998	1.998	2.000	1.998	2.000
		(0.444)	(0.199)	(0.144)	(0.089)	(0.309)	(0.141)	(0.098)	(0.063)	(0.200)	(0.092)	(0.064)	(0.040)
M2	$\bar{\hat{\alpha}}$	2.003	2.003	1.999	1.999	1.998	2.000	2.000	1.998	1.999	2.000	1.999	2.000
		(0.454)	(0.204)	(0.148)	(0.090)	(0.314)	(0.143)	(0.100)	(0.065)	(0.204)	(0.094)	(0.065)	(0.041)
	$\bar{\hat{\beta}}$	-9.665	-0.531	-0.079	-0.024	-4.517	-0.078	-0.041	-0.016	-0.594	-0.028	-0.013	-0.005
		(10.795)	(2.990)	(0.414)	(0.239)	(8.732)	(0.413)	(0.276)	(0.167)	(3.196)	(0.235)	(0.169)	(0.106)
M3	$\bar{\hat{\alpha}}$	2.208	2.045	2.020	2.006	2.156	2.037	2.018	2.004	2.091	2.015	2.006	2.003
		(0.680)	(0.279)	(0.191)	(0.117)	(0.520)	(0.221)	(0.151)	(0.094)	(0.359)	(0.158)	(0.109)	(0.070)
	$\bar{\hat{\gamma}}$	-0.101	-0.019	-0.010	-0.003	-0.034	-0.008	-0.004	-0.001	-0.007	-0.001	-0.001	0.000
		(0.250)	(0.092)	(0.062)	(0.038)	(0.102)	(0.040)	(0.028)	(0.017)	(0.028)	(0.013)	(0.009)	(0.006)
M4	$\bar{\hat{\alpha}}$	2.205	2.044	2.020	2.006	2.154	2.037	2.018	2.004	2.091	2.015	2.007	2.003
		(0.680)	(0.280)	(0.192)	(0.117)	(0.518)	(0.221)	(0.151)	(0.094)	(0.361)	(0.159)	(0.109)	(0.070)
	$\bar{\hat{\gamma}}$	-0.101	-0.019	-0.010	-0.003	-0.034	-0.008	-0.004	-0.001	-0.007	-0.001	-0.001	0.000
		(0.255)	(0.093)	(0.063)	(0.039)	(0.103)	(0.041)	(0.028)	(0.017)	(0.028)	(0.013)	(0.009)	(0.006)
	$\bar{\hat{\beta}}$	-9.653	-0.521	-0.074	-0.023	-4.491	-0.075	-0.040	-0.016	-0.551	-0.029	-0.013	-0.005
		(10.834)	(2.988)	(0.422)	(0.242)	(8.713)	(0.415)	(0.278)	(0.168)	(2.949)	(0.235)	(0.169)	(0.106)

Table 3.2 contains the results of the simulation study from the first scenario. When the sample size and/or the follow-up period increase, the bias in the empirical mean is decreasing in any fitted model. For example, when the model M4 fitted to the data generated in first scenario with $\alpha = 2$ and $\tau = 1$, the empirical means of estimates of α (= 2.205, 2.044, 2.020 and 2.006 when $m = 10$, 50, 100 and 250, respectively) gets closer to the true value of α as m increases. Also, when the model M4 fitted to the data generated in first scenario with $\alpha = 2$ and $m = 10$, the empirical means of estimates of α (= 2.205, 2.154 and 2.091 when $\tau = 1$, 2 and 5, respectively) gets closer to the true value of α as τ increases. The empirical standard deviations of parameter estimates are decreasing as sample size m and/or the follow-up period τ increase in any fitted model. For example, when the model M4 fitted to the data generated in first scenario with $\alpha = 2$ and $\tau = 1$, the empirical standard deviations of estimates of α (= 0.680, 0.280, 0.192 and 0.117 when $m = 10$, 50, 100 and 250, respectively)

decreases α as m increases. Also when the model M4 fitted to the data generated in first scenario with $\alpha = 2$ and $m = 10$, the empirical means of estimates of α (= 0.680, 0.518 and 0.361 when $\tau = 1$, 2 and 5, respectively) decreases as τ increases. These results indicate that the parameter estimates are consistent as m increases and also τ increases. In general, parameter estimates and their standard errors of α in all models are within two standard errors of the correct values. All models considered in Table 3.2 revealed the correct conclusion that there is no carryover effect or trend in the data.

In the second scenario, we generated data from the model

$$\lambda_i[t|\mathcal{H}_i(t)] = \alpha \exp\left[\beta\, Z_i(t)\right], \qquad t > 0, \tag{3.46}$$

where $\alpha = 1$ and $\beta = 0.693$. We fixed the duration of the carryover effect (i.e. Δ) at 0.02. Note that in this scenario, the generated data included carryover effect but there was no trend due to the number of previous events. Other than the data generation process, the simulation design of the second scenario was the same with that of the first scenario. We present the results of simulations for the second scenario in Table 3.3. Note that in this case, the fitted model M2 is the correctly specified model whereas other fitted models (M1, M3 and M4) are misspecified.

Table 3.3: Empirical means and standard deviations (within parenthesis) of parameter estimates of 4 different models when the data generated from a model with $\alpha = 1, \beta = 0.693$, $\gamma = 0$ and $\Delta = 0.02$ in (3.44) with 2,000 simulations.

Fitted		Simulation											
		$\tau = 1$				$\tau = 2$				$\tau = 5$			
		m=10	m=50	m=100	m=250	m=10	m=50	m=100	m=250	m=10	m=50	m=100	m=250
M1	$\bar{\hat{\alpha}}$	1.018	1.022	1.019	1.018	1.020	1.020	1.019	1.019	1.018	1.020	1.019	1.019
		(0.331)	(0.144)	(0.102)	(0.064)	(0.230)	(0.099)	(0.073)	(0.046)	(0.145)	(0.064)	(0.046)	(0.030)
M2	$\bar{\hat{\alpha}}$	0.999	1.003	1.000	0.998	1.000	1.001	1.000	0.999	0.998	1.001	0.999	0.999
		(0.325)	(0.141)	(0.100)	(0.063)	(0.225)	(0.097)	(0.072)	(0.045)	(0.142)	(0.063)	(0.045)	(0.029)
	$\bar{\hat{\beta}}$	-12.360	-2.198	0.138	0.640	-8.170	0.234	0.617	0.670	-2.186	0.632	0.666	0.684
		(9.664)	(6.990)	(2.949)	(0.355)	(10.121)	(2.658)	(0.402)	(0.233)	(7.017)	(0.344)	(0.233)	(0.147)
M3	$\bar{\hat{\alpha}}$	1.116	1.034	1.015	1.005	1.116	1.027	1.016	1.008	1.077	1.023	1.014	1.012
		(0.433)	(0.172)	(0.122)	(0.076)	(0.350)	(0.135)	(0.094)	(0.059)	(0.236)	(0.101)	(0.069)	(0.044)
	$\bar{\hat{\gamma}}$	-1.137	-0.029	0.005	0.023	-0.089	-0.005	0.004	0.011	-0.018	0.000	0.002	0.003
		(4.411)	(0.187)	(0.130)	(0.078)	(0.245)	(0.087)	(0.058)	(0.037)	(0.072)	(0.030)	(0.021)	(0.013)
M4	$\bar{\hat{\alpha}}$	1.110	1.030	1.012	1.002	1.109	1.022	1.010	1.003	1.066	1.013	1.005	1.002
		(0.431)	(0.171)	(0.122)	(0.076)	(0.347)	(0.135)	(0.093)	(0.059)	(0.232)	(0.099)	(0.068)	(0.044)
	$\bar{\hat{\gamma}}$	-1.328	-0.060	-0.027	-0.008	-0.105	-0.020	-0.010	-0.003	-0.022	-0.004	-0.002	-0.001
		(4.760)	(0.193)	(0.135)	(0.081)	(0.249)	(0.088)	(0.059)	(0.038)	(0.072)	(0.029)	(0.020)	(0.013)
	$\bar{\hat{\beta}}$	-11.286	-2.159	0.154	0.644	-8.075	0.241	0.623	0.672	-2.191	0.632	0.666	0.684
		(10.347)	(6.996)	(2.959)	(0.366)	(10.077)	(2.679)	(0.408)	(0.235)	(7.194)	(0.345)	(0.234)	(0.148)

In this scenario, our conclusions are similar to those given in the previous scenario for all the models and cases considered. The simulation results for the models M3 and M4 show that trend due to number of previous events is not significant in these models. Therefore, we conclude that, when the data generated from a model with carryover effects but without the dynamic trend covariate $N^*(t^-)$, carryover effects and effects of $N^*(t^-)$ do not confound.

In the last scenario, we used the model

$$\lambda_i[t|\mathcal{H}_i(t)] = \alpha \exp\left[\gamma\, N_i^*(t^-)\right], \qquad t > 0, \tag{3.47}$$

where $\alpha = 1$ and $\gamma = 0.223$, to generate data. Other than the data generation process we used the same simulation design given in the first scenario. Thus, the correct model in this scenario is the model M3, which does not include carryover effects.

Table 3.4: Empirical means and standard deviations (within parenthesis) of parameter estimates of 4 different models when the data generated from a model with $\alpha = 1, \beta = 0$ and $\gamma = 0.223$ in (3.44) with 2,000 simulations.

Fitted		Simulation											
		$\tau = 1$				$\tau = 2$				$\tau = 5$			
		m=10	m=50	m=100	m=250	m=10	m=50	m=100	m=250	m=10	m=50	m=100	m=250
M1	$\bar{\hat{\alpha}}$	1.153	1.157	1.155	1.156	1.521	1.525	1.526	1.530	14.730	14.726	14.665	14.642
		(0.393)	(0.183)	(0.128)	(0.081)	(0.664)	(0.298)	(0.211)	(0.131)	(4.918)	(2.158)	(1.521)	(0.965)
M2	$\bar{\hat{\alpha}}$	1.146	1.149	1.147	1.147	1.427	1.430	1.431	1.431	5.278	5.193	5.160	5.155
		(0.392)	(0.180)	(0.128)	(0.080)	(0.438)	(0.196)	(0.139)	(0.087)	(1.561)	(0.672)	(0.472)	(0.300)
	$\bar{\hat{\beta}}$	-13.479	-3.762	-0.593	0.211	-6.226	0.628	0.925	1.104	2.506	2.519	2.524	2.522
		(9.277)	(8.012)	(3.851)	(0.604)	(9.766)	(1.696)	(0.658)	(0.450)	(0.228)	(0.104)	(0.074)	(0.047)
M3	$\bar{\hat{\alpha}}$	1.115	1.030	1.013	1.007	1.094	1.028	1.014	1.004	1.008	1.003	1.001	1.000
		(0.405)	(0.164)	(0.116)	(0.073)	(0.291)	(0.110)	(0.074)	(0.045)	(0.130)	(0.058)	(0.041)	(0.026)
	$\bar{\hat{\gamma}}$	-0.491	0.171	0.200	0.211	0.156	0.206	0.215	0.220	0.223	0.223	0.223	0.223
		(3.268)	(0.138)	(0.090)	(0.055)	(0.146)	(0.036)	(0.021)	(0.011)	(0.007)	(0.003)	(0.002)	(0.001)
M4	$\bar{\hat{\alpha}}$	1.112	1.029	1.012	1.007	1.091	1.026	1.013	1.004	1.007	1.003	1.001	1.000
		(0.405)	(0.164)	(0.116)	(0.074)	(0.292)	(0.111)	(0.075)	(0.045)	(0.131)	(0.058)	(0.041)	(0.026)
	$\bar{\hat{\gamma}}$	-0.520	0.172	0.200	0.212	0.158	0.208	0.216	0.221	0.223	0.223	0.223	0.223
		(3.365)	(0.142)	(0.092)	(0.056)	(0.148)	(0.038)	(0.023)	(0.012)	(0.009)	(0.004)	(0.003)	(0.002)
	$\bar{\hat{\beta}}$	-13.198	-4.021	-0.869	-0.076	-6.725	-0.206	-0.056	-0.019	-0.008	0.000	0.000	0.000
		(9.598)	(8.017)	(3.841)	(0.601)	(9.596)	(1.529)	(0.294)	(0.166)	(0.106)	(0.045)	(0.032)	(0.020)

Table 3.4 contains the results of the simulation studies in the third scenario. For models M1, M3 and M4, our conclusions from Table 3.4 are similar to those presented in Table 3.2. However, the results from the misspecified models (M1 and M2) indicate significant carryover effects when τ increases. For example, when the model M2 fitted to the data generated in third scenario with $\alpha = 1$, $\gamma = 0.223$ and $m = 250$, the

empirical means of estimates of β (= 1.104 and 2.522 when $\tau = 2$ and 5, respectively) are significantly biased and the bias increases as τ increases. These results show that, when the data include trend but not carryover effects and if the trend is ignored in the model, there might be wrong conclusions about the presence of carryover effects. The expanded model M4 includes both trend due to $N(t^-)$ and carryover effects. In this model, carryover effects are not significant in the cases considered in the simulation study. Therefore, it indicates that the model M4 does not suffer from confounding issues between carryover effects and trends due to the previous number of events.

In the second setting of this simulation study, we investigated the issue of confounding between carryover effects and trend due to number of previous events in heterogeneous processes. We considered three models to generate the data. These models define the scenarios explained below. In the first scenario, we simulated 1,000 runs of the model

$$\lambda_i[t|\mathcal{H}_i(t), \nu_i] = \alpha\nu_i, \qquad t > 0, \tag{3.48}$$

where the ν_i are i.i.d. gamma random variables with mean 1 and variance ϕ, for each combination of m (= 50, 100 and 250), τ (= 1 and 2) and ϕ (= 0.1 and 2). We took $\alpha = 2$ in the model (3.48). Thus, the data generated did not include carryover effects or trend. We then fitted four models in each scenario;

M1*: Model 1 is a heterogeneous null model with the intensity function $\lambda_i[t|\mathcal{H}_i(t), \nu_i] = \alpha\nu_i$, $i = 1, \ldots, m$,

M2*: Model 2 is a heterogeneous carryover effects model with the intensity function $\lambda_i[t|\mathcal{H}_i(t), \nu_i] = \alpha\nu_i \exp[\beta\, Z_i(t)]$, $i = 1, \ldots, m$,

M3*: Model 3 is a heterogeneous trend model with the intensity function $\lambda_i[t|\mathcal{H}_i(t), \nu_i] = \alpha\nu_i \exp[\gamma\, N_i(t^-)]$, $i = 1, \ldots, m$, and

M4*: Model 4 is a heterogeneous hybrid model with the intensity function $\lambda_i[t|\mathcal{H}_i(t), \nu_i] = \alpha\nu_i \exp[\gamma\, N_i(t^-) + \beta\, Z_i(t)]$, $i = 1, \ldots, m$.

Table 3.5 contains the results of the simulation study from the first scenario for nonidentical processes. In general, empirical estimates and their standard errors of α in all models are within two standard errors of the correct values. All models considered in Table 3.5 reveal the correct conclusion that there is no carryover effect or trend in the data.

Table 3.5: Empirical means and standard deviations (within parenthesis) of parameter estimates of 4 different models when the data generated from a model with $\alpha = 2$ and $\phi = (0.1, 2)$ in (3.48) with 1,000 simulations.

Fitted		Simulation											
		$\phi = 0.1$						$\phi = 2$					
		$\tau = 1$			$\tau = 2$			$\tau = 1$			$\tau = 2$		
		m=50	m=100	m=250	m=50	m=100	m=250	m=50	m=100	m=250	m=50	m=100	m=250
M1*	$\bar{\hat{\alpha}}$	1.995	2.000	1.997	1.998	2.003	1.995	1.987	2.007	1.998	2.004	2.008	1.997
		(0.218)	(0.156)	(0.096)	(0.166)	(0.123)	(0.075)	(0.443)	(0.331)	(0.205)	(0.421)	(0.307)	(0.189)
	$\bar{\hat{\phi}}$	0.102	0.101	0.098	0.096	0.097	0.099	2.008	1.978	2.005	1.989	2.028	2.007
		(0.109)	(0.080)	(0.055)	(0.070)	(0.050)	(0.031)	(0.691)	(0.430)	(0.275)	(0.528)	(0.370)	(0.227)
M2*	$\bar{\hat{\alpha}}$	1.997	2.002	1.997	1.997	2.004	1.995	1.991	2.013	1.998	2.008	2.011	1.997
		(0.221)	(0.160)	(0.098)	(0.168)	(0.125)	(0.076)	(0.451)	(0.335)	(0.208)	(0.424)	(0.310)	(0.190)
	$\bar{\hat{\phi}}$	0.104	0.103	0.098	0.096	0.097	0.099	2.012	1.985	2.006	1.992	2.031	2.008
		(0.112)	(0.082)	(0.056)	(0.070)	(0.050)	(0.031)	(0.704)	(0.434)	(0.277)	(0.530)	(0.373)	(0.227)
	$\bar{\hat{\beta}}$	-0.497	-0.127	-0.018	-0.049	-0.041	-0.003	-0.145	-0.050	-0.009	-0.046	-0.024	-0.007
		(2.746)	(1.015)	(0.234)	(0.390)	(0.264)	(0.157)	(1.408)	(0.249)	(0.155)	(0.271)	(0.170)	(0.108)
M3*	$\bar{\hat{\alpha}}$	2.100	2.062	2.022	2.034	2.017	2.007	2.096	2.044	2.017	2.045	2.031	2.002
		(0.377)	(0.267)	(0.166)	(0.277)	(0.214)	(0.128)	(0.542)	(0.376)	(0.236)	(0.476)	(0.346)	(0.213)
	$\bar{\hat{\phi}}$	0.149	0.133	0.113	0.106	0.101	0.103	2.114	2.013	2.025	2.014	2.043	2.010
		(0.195)	(0.145)	(0.092)	(0.095)	(0.071)	(0.047)	(0.795)	(0.489)	(0.311)	(0.564)	(0.390)	(0.241)
	$\bar{\hat{\gamma}}$	-0.032	-0.019	-0.007	-0.004	0.000	-0.002	-0.016	-0.005	-0.003	-0.003	-0.002	0.000
		(0.118)	(0.092)	(0.059)	(0.048)	(0.037)	(0.024)	(0.063)	(0.034)	(0.021)	(0.022)	(0.014)	(0.009)
M4*	$\bar{\hat{\alpha}}$	2.101	2.064	2.022	2.034	2.018	2.007	2.102	2.048	2.018	2.050	2.035	2.003
		(0.380)	(0.267)	(0.167)	(0.278)	(0.216)	(0.129)	(0.552)	(0.379)	(0.238)	(0.480)	(0.350)	(0.213)
	$\bar{\hat{\phi}}$	0.150	0.134	0.113	0.106	0.101	0.102	2.118	2.018	2.025	2.019	2.046	2.010
		(0.196)	(0.146)	(0.093)	(0.095)	(0.071)	(0.048)	(0.805)	(0.492)	(0.312)	(0.568)	(0.393)	(0.240)
	$\bar{\hat{\gamma}}$	-0.032	-0.019	-0.007	-0.004	0.000	-0.002	-0.016	-0.005	-0.003	-0.003	-0.002	0.000
		(0.119)	(0.093)	(0.059)	(0.048)	(0.037)	(0.024)	(0.063)	(0.035)	(0.021)	(0.022)	(0.014)	(0.009)
	$\bar{\hat{\beta}}$	-0.447	-0.129	-0.019	-0.054	-0.044	-0.003	-0.157	-0.056	-0.012	-0.053	-0.027	-0.008
		(2.449)	(1.014)	(0.234)	(0.391)	(0.264)	(0.157)	(1.406)	(0.250)	(0.155)	(0.271)	(0.170)	(0.108)

In the second scenario, we generated data from the model

$$\lambda_i[t|\mathcal{H}_i(t), \nu_i] = \alpha\nu_i \exp\left[\beta Z_i(t)\right], \qquad t > 0, \tag{3.49}$$

where $\alpha = 1$ and $\beta = 0.693$. We fixed the duration of the carryover effect (i.e. Δ) at 0.02. Note that in this scenario, the generated data included carryover effect but there was no trend due to the number of previous events. Other than the data generation process, the simulation design of the second scenario was the same with that of the first scenario. We present the results of simulations for the second scenario in Table 3.6. Note that in this case, the fitted model M2* is the correctly specified model whereas other fitted models (M1*, M3* and M4*) are misspecified.

Table 3.6: Empirical means and standard deviations (within parenthesis) of parameter estimates of 4 different models when the data generated from a model with $\alpha = 1, \beta = 0.693$, $\phi = (0.1, 2)$ and $\Delta = 0.02$ in (3.49) with 1,000 simulations.

Fitted		Simulation											
		$\phi = 0.1$						$\phi = 2$					
		$\tau = 1$			$\tau = 2$			$\tau = 1$			$\tau = 2$		
		m=50	m=100	m=250	m=50	m=100	m=250	m=50	m=100	m=250	m=50	m=100	m=250
M1*	$\bar{\hat{\alpha}}$	1.019	1.020	1.023	1.021	1.021	1.021	1.055	1.051	1.060	1.053	1.064	1.057
		(0.153)	(0.109)	(0.068)	(0.112)	(0.080)	(0.050)	(0.271)	(0.200)	(0.122)	(0.248)	(0.180)	(0.113)
	$\bar{\hat{\phi}}$	0.164	0.155	0.148	0.124	0.124	0.122	2.175	2.188	2.207	2.138	2.141	2.142
		(0.194)	(0.148)	(0.101)	(0.117)	(0.088)	(0.059)	(0.968)	(0.625)	(0.379)	(0.705)	(0.478)	(0.287)
M2*	$\bar{\hat{\alpha}}$	1.000	0.999	1.000	1.000	1.000	0.999	0.999	0.995	1.002	0.996	1.006	0.998
		(0.150)	(0.106)	(0.067)	(0.110)	(0.079)	(0.048)	(0.248)	(0.182)	(0.111)	(0.225)	(0.164)	(0.102)
	$\bar{\hat{\phi}}$	0.136	0.120	0.106	0.104	0.102	0.098	1.987	1.996	2.006	2.007	2.006	2.000
		(0.179)	(0.137)	(0.091)	(0.107)	(0.082)	(0.055)	(0.913)	(0.583)	(0.358)	(0.670)	(0.452)	(0.271)
	$\bar{\hat{\beta}}$	-1.015	0.368	0.668	0.398	0.590	0.677	0.076	0.609	0.679	0.614	0.653	0.690
		(5.178)	(1.956)	(0.339)	(1.767)	(0.747)	(0.223)	(3.282)	(0.721)	(0.208)	(0.723)	(0.242)	(0.145)
M3*	$\bar{\hat{\alpha}}$	1.136	1.069	1.043	1.067	1.048	1.029	1.123	1.087	1.065	1.096	1.075	1.054
		(0.277)	(0.175)	(0.112)	(0.193)	(0.132)	(0.085)	(0.352)	(0.238)	(0.141)	(0.290)	(0.206)	(0.127)
	$\bar{\hat{\phi}}$	0.344	0.238	0.187	0.163	0.155	0.131	2.340	2.283	2.212	2.219	2.157	2.128
		(0.452)	(0.294)	(0.196)	(0.195)	(0.151)	(0.095)	(1.218)	(0.790)	(0.472)	(0.807)	(0.553)	(0.319)
	$\bar{\hat{\gamma}}$	-0.133	-0.054	-0.019	-0.025	-0.015	-0.003	-0.025	-0.016	-0.001	-0.012	-0.002	0.001
		(0.263)	(0.195)	(0.131)	(0.117)	(0.089)	(0.057)	(0.154)	(0.094)	(0.046)	(0.056)	(0.033)	(0.019)
M4*	$\bar{\hat{\alpha}}$	1.124	1.056	1.029	1.051	1.031	1.010	1.077	1.044	1.022	1.047	1.028	1.007
		(0.272)	(0.170)	(0.108)	(0.188)	(0.128)	(0.082)	(0.320)	(0.219)	(0.129)	(0.263)	(0.187)	(0.115)
	$\bar{\hat{\phi}}$	0.324	0.216	0.160	0.148	0.137	0.111	2.208	2.146	2.071	2.114	2.052	2.018
		(0.444)	(0.284)	(0.187)	(0.187)	(0.146)	(0.090)	(1.167)	(0.747)	(0.447)	(0.768)	(0.523)	(0.303)
	$\bar{\hat{\gamma}}$	-0.155	-0.073	-0.036	-0.031	-0.020	-0.006	-0.036	-0.026	-0.011	-0.016	-0.006	-0.002
		(0.265)	(0.195)	(0.130)	(0.116)	(0.088)	(0.056)	(0.151)	(0.092)	(0.045)	(0.054)	(0.031)	(0.018)
	$\bar{\hat{\beta}}$	-1.080	0.432	0.669	0.393	0.584	0.676	0.055	0.602	0.676	0.604	0.647	0.688
		(5.286)	(1.520)	(0.341)	(1.774)	(0.771)	(0.223)	(3.347)	(0.725)	(0.210)	(0.724)	(0.243)	(0.145)

In this simulation scenario, we observed similar outcomes to those given in the previous scenario for all the models and cases. The simulation results for the models M3* and M4* show that trend due to number of previous events is not significant in these models model. This result suggests that when the data generated from a model with carryover effects but without the dynamic covariate $N^*(t^-)$, carryover effects and effects of $N^*(t^-)$ do not confound.

In the last scenario, we generated data from the model

$$\lambda_i[t|\mathcal{H}_i(t), \nu_i] = \alpha \nu_i \exp\left[\gamma \, N_i^*(t^-)\right], \qquad t > 0, \tag{3.50}$$

where $\alpha = 1$ and $\gamma = 0.095$. Other than the data generation process we used the same simulation design given in the first scenario for nonidentical processes. Thus, the correct model in this scenario is the model M3*, which does not include carryover effects.

Table 3.7: Empirical means and standard deviations (within parenthesis) of parameter estimates of 4 different models when the data generated from a model with $\alpha = 1, \gamma = 0.095$ and $\phi = (0.1, 2)$ in (3.50) with 1,000 simulations.

Fitted		Simulation											
		$\phi = 0.1$						$\phi = 2$					
		$\tau = 1$			$\tau = 2$			$\tau = 1$			$\tau = 2$		
		m=50	m=100	m=250	m=50	m=100	m=250	m=50	m=100	m=250	m=50	m=100	m=250
M1*	$\bar{\hat{\alpha}}$	1.058	1.058	1.060	1.131	1.131	1.133	1.254	1.254	1.265	1.701	1.703	1.710
		(0.164)	(0.116)	(0.074)	(0.134)	(0.098)	(0.060)	(0.415)	(0.302)	(0.202)	(0.623)	(0.427)	(0.266)
	$\bar{\hat{\phi}}$	0.231	0.231	0.228	0.244	0.246	0.257	2.759	2.818	2.865	3.185	3.292	3.330
		(0.232)	(0.178)	(0.115)	(0.148)	(0.110)	(0.070)	(1.286)	(0.902)	(0.561)	(1.025)	(0.714)	(0.458)
M2*	$\bar{\hat{\alpha}}$	1.059	1.058	1.060	1.130	1.131	1.133	1.225	1.222	1.230	1.590	1.586	1.593
		(0.167)	(0.116)	(0.075)	(0.134)	(0.099)	(0.060)	(0.381)	(0.269)	(0.179)	(0.538)	(0.368)	(0.227)
	$\bar{\hat{\phi}}$	0.233	0.232	0.227	0.243	0.247	0.256	2.697	2.741	2.773	3.045	3.128	3.159
		(0.235)	(0.182)	(0.117)	(0.149)	(0.111)	(0.071)	(1.224)	(0.835)	(0.507)	(0.953)	(0.662)	(0.423)
	$\bar{\hat{\beta}}$	-4.388	-1.291	-0.070	-0.843	-0.129	-0.017	-0.899	0.009	0.172	0.168	0.340	0.357
		(8.000)	(4.615)	(0.743)	(3.814)	(0.975)	(0.272)	(4.315)	(1.330)	(0.255)	(1.444)	(0.198)	(0.113)
M3*	$\bar{\hat{\alpha}}$	1.115	1.054	1.027	1.042	1.023	1.007	1.068	1.026	1.015	1.014	0.992	0.992
		(0.257)	(0.164)	(0.104)	(0.166)	(0.113)	(0.075)	(0.285)	(0.192)	(0.120)	(0.233)	(0.161)	(0.101)
	$\bar{\hat{\phi}}$	0.316	0.220	0.159	0.144	0.124	0.109	2.222	2.109	2.066	1.992	1.973	1.971
		(0.427)	(0.280)	(0.182)	(0.168)	(0.123)	(0.083)	(1.076)	(0.685)	(0.397)	(0.675)	(0.452)	(0.289)
	$\bar{\hat{\gamma}}$	-0.044	0.024	0.063	0.070	0.080	0.092	0.061	0.079	0.087	0.088	0.093	0.094
		(0.246)	(0.175)	(0.118)	(0.090)	(0.066)	(0.044)	(0.120)	(0.061)	(0.030)	(0.028)	(0.015)	(0.008)
M4*	$\bar{\hat{\alpha}}$	1.116	1.055	1.027	1.042	1.024	1.007	1.068	1.026	1.015	1.013	0.991	0.992
		(0.258)	(0.165)	(0.104)	(0.167)	(0.114)	(0.075)	(0.287)	(0.192)	(0.120)	(0.233)	(0.161)	(0.101)
	$\bar{\hat{\phi}}$	0.316	0.220	0.159	0.144	0.125	0.109	2.221	2.109	2.065	1.991	1.972	1.971
		(0.428)	(0.281)	(0.181)	(0.169)	(0.124)	(0.083)	(1.085)	(0.689)	(0.398)	(0.676)	(0.452)	(0.290)
	$\bar{\hat{\gamma}}$	-0.044	0.025	0.062	0.070	0.081	0.092	0.063	0.080	0.087	0.088	0.094	0.094
		(0.247)	(0.176)	(0.119)	(0.090)	(0.066)	(0.044)	(0.121)	(0.063)	(0.031)	(0.029)	(0.016)	(0.009)
	$\bar{\hat{\beta}}$	-3.934	-1.017	-0.068	-0.662	-0.149	-0.033	-1.005	-0.167	-0.029	-0.152	-0.017	-0.012
		(7.317)	(3.818)	(0.477)	(3.120)	(0.964)	(0.272)	(4.185)	(1.306)	(0.226)	(1.413)	(0.181)	(0.101)

Table 3.7 contains the results of the simulation studies for the third scenario for nonidentical processes. For models M1*, M3* and M4*, our conclusions from Table 3.7 are similar to those presented in Table 3.5. However, the results from the misspecified models (M1*, M2* and M4*) indicate significant carryover effects when τ increases. For example, when the model M2* fitted to the data generated in third scenario with $\alpha = 1$, $\gamma = 0.095$ and $m = 250$, the empirical means of estimates of β (= 0.172 and 0.357 when $\tau = 1$ and 2, respectively) are biased and the bias increases as τ increases. Note that in this specific example, the empirical means of estimates of ϕ (= 2.773 and 3.159 when $\tau = 1$ and 2, respectively) are biased and the bias increases as τ increases. These results show that when the data includes trend but not carryover effects and if the trend is ignored in the model, there might be wrong conclusions about the presence of carryover effects. The expanded model M4* includes both trend due to number of previous events and carryover effects. In this model, carryover effects are not significant. Therefore, the simulation results indicate that the model M4* does

not suffer from confounding issues between carryover effects and trends due to the previous number of events.

Briefly, simulations for the nonidentical processes revealed that if the population is heterogeneous, similar to the conclusion of the first setting, model misspecification may result in wrong conclusions about carryover effects when the trend is present. The correctly specified model does not suffer from such an issue.

3.4 Application: Recurrent Asthma Attacks in Children

We analyze the data from a prevention trial in infants with a high risk of asthma (Duchateau et al., 2003). A brief introduction of this data is available in Section 1.2.2. We examine the data with some simple graphical illustrations to understand the asthma data before we fit the models which are developed in this chapter. For this purpose, we choose the plots of the *cumulative mean function* and the *cumulative variance function* (Cook and Lawless, 2007, p. 2). These plots require all the processes to be observed in a fixed time interval. But in asthma data, the individuals have different lengths of observation periods. Therefore, we consider the individuals with at least 420 days observation period. Consequently, we select only 101 individuals from 113 individuals in the treatment group and 108 individuals from 119 individuals in the control group. We plot cumulative mean functions and the cumulative variance functions for each group separately and also for the full data using those chosen individuals for the period of 420 days (1.286 years).

When the data does not have excess heterogeneity, the cumulative mean and variance functions are close together. The plots in Figure 3.2 show that the cumulative variance functions increase more exponentially than the cumulative mean functions in each group, as well as in the full data. This result indicates that there is excess heterogeneity among the individuals in the asthma data.

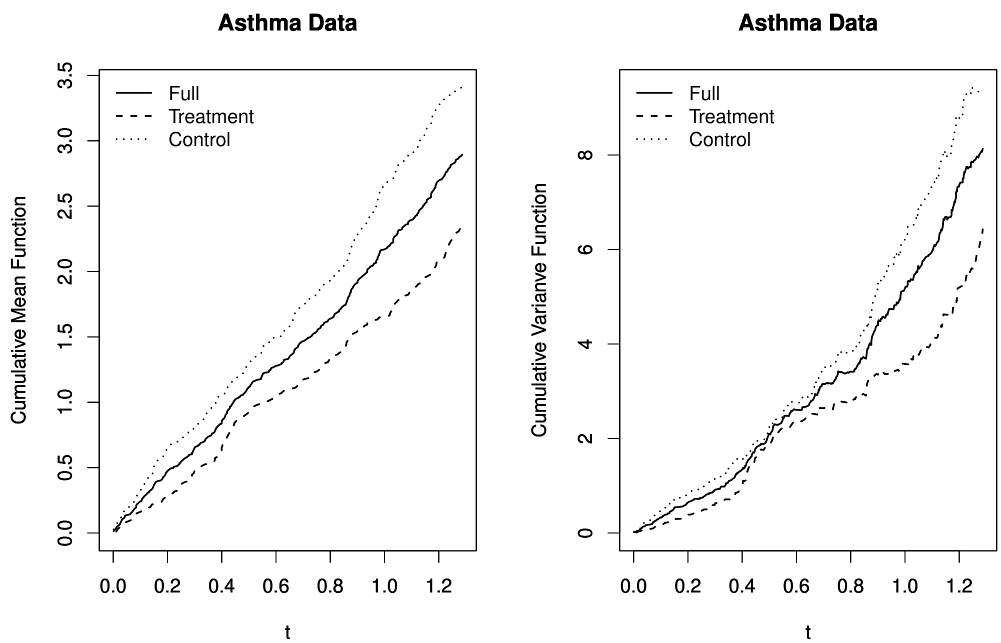

Figure 3.2: Cumulative mean and variance functions of the asthma data.

We consider two different modeling procedures based on the modulated Poisson process model (3.1) (MPP) and the random effects model (3.38) (REM). In each procedure we fit four different models; (1) the null model, which has neither trend nor carryover effects, (2) the carryover effects model, which does not consider the trend effect, (3) the trend model, which does not consider the carryover effects effect, and (4) the hybrid model, which consider both carryover effects and trend due to the number of previous events. Parameter estimates of the models for the group–wise data are presented in Table 3.8 with their corresponding Akaike information criterion (AIC) values (Akaike, 1974). As suggested in Cook and Lawless (2013), we pick carryover effect period as two months, i.e., $\Delta = 56$ days $= 0.1533$ year.

Table 3.8: Parameter estimates and standard errors (within parenthesis) from count-based models for the asthma data.

Models		Treatment group						Control group					
		$\hat{\alpha}$	$\hat{\phi}$	$\hat{\gamma}$	$\hat{\beta}$	$\ell(\hat{\boldsymbol{\theta}})$	AIC	$\hat{\alpha}$	$\hat{\phi}$	$\hat{\gamma}$	$\hat{\beta}$	$\ell(\hat{\boldsymbol{\theta}})$	AIC
Null	MPP	2.1219 (0.1158)				-83.2183	170.4366	2.8832 (0.1312)				28.4411	-52.8821
	REM	2.3061 (0.2089)	0.5428 (0.1080)			-35.8561	77.7122	3.0966 (0.2656)	0.5881 (0.1021)			116.2783	-226.5565
Carryover	MPP	1.3732 (0.1041)			1.3162 (0.1092)	-16.2204	38.4408	1.8265 (0.1223)			1.1421 (0.0913)	104.8940	-203.7880
	REM	1.5188 (0.1416)	0.2515 (0.0775)		1.0450 (0.1258)	-3.1817	14.3635	2.1528 (0.1950)	0.3420 (0.0796)		0.7544 (0.1070)	141.0022	-274.0044
Trend	MPP	1.4682 (0.1010)		0.2012 (0.0158)		-27.6926	61.3851	2.1822 (0.1146)		0.1044 (0.0065)		98.1608	-190.3217
	REM	1.6628 (0.1570)	0.2069 (0.0931)	0.1357 (0.0274)		-23.5912	55.1825	2.6568 (0.2410)	0.4170 (0.1002)	0.0380 (0.0128)		120.6328	-233.2655
Hybrid	MPP	1.2124 (0.0958)		0.1276 (0.0196)	0.9820 (0.1252)	**1.4679**	**5.0641**	1.6409 (0.1116)		0.0774 (0.0078)	0.8837 (0.0983)	**137.4363**	**-266.8726**
	REM	1.2822 (0.1230)	0.0751 (0.0708)	0.0997 (0.0300)	0.9569 (0.1287)	**2.2849**	**5.4301**	1.9361 (0.1824)	0.2180 (0.0826)	0.0329 (0.0140)	0.7439 (0.1087)	**143.7813**	**-277.5626**

Note that the hybrid versions of MPP and REM models give the largest maximum log-likelihood values with smallest AIC values for both groups. This result indicates that the model with both the carryover effects and trends, due to the number of previous events, adequately fits both groups in the asthma data. For the hybrid model, we fixed the carryover effects period as $\Delta = 56$ days $= 0.1533$ year and it needs to be justified. For this purpose, we choose a wide range of plausible values for Δ and fit the hybrid MPP and hybrid REM models. The summaries of the fitted models are given in Table 3.9. We compare the maximum values of the log-likelihood for each model and see that the value of Δ (among those shown) best supported by the data is $\Delta = 56$ days in both the treatment and control groups. A similar conclusion was also given by Cigsar and Lawless (2012).

Table 3.9: Parameter estimates and standard errors (within parenthesis) from count-based hybrid models for the asthma data with different values of carryover effects period.

Δ (in days)	Model	Treatment group					Control group				
		$\hat{\alpha}$	$\hat{\phi}$	$\hat{\gamma}$	$\hat{\beta}$	$\ell(\hat{\boldsymbol{\theta}})$	$\hat{\alpha}$	$\hat{\phi}$	$\hat{\gamma}$	$\hat{\beta}$	$\ell(\hat{\boldsymbol{\theta}})$
7	MPP	1.4604 (0.1001)		0.1799 (0.0171)	0.6590 (0.1846)	-22.0818	2.1467 (0.1137)		0.0908 (0.0074)	0.6202 (0.1490)	105.8516
	REM	1.6251 (0.1511)	0.1751 (0.0917)	0.1248 (0.0283)	0.5681 (0.1874)	-19.4512	2.5689 (0.2279)	0.3809 (0.0963)	0.0332 (0.0128)	0.4466 (0.1458)	124.9639
14	MPP	1.4117 (0.0985)		0.1573 (0.0178)	0.9091 (0.1467)	-10.8721	2.0363 (0.1123)		0.0819 (0.0074)	0.8093 (0.1179)	118.6873
	REM	1.5214 (0.1398)	0.1139 (0.0873)	0.1190 (0.0300)	0.8403 (0.1525)	-9.8184	2.4122 (0.2106)	0.3284 (0.0916)	0.0308 (0.0128)	0.6157 (0.1201)	132.7901
28	MPP	1.3177 (0.0966)		0.1385 (0.0187)	1.0025 (0.1298)	-0.6171	1.8670 (0.1114)		0.0784 (0.0076)	0.8500 (0.1028)	129.1304
	REM	1.3791 (0.1273)	0.0617 (0.0761)	0.1156 (0.0310)	0.9682 (0.1360)	-0.1940	2.2062 (0.1955)	0.2700 (0.0875)	0.0316 (0.0133)	0.6695 (0.1103)	138.3192
42	MPP	1.2612 (0.0961)		0.1328 (0.0192)	0.9807 (0.1261)	0.4972	1.7468 (0.1115)		0.0781 (0.0077)	0.8558 (0.0989)	133.3978
	REM	1.3266 (0.1243)	0.0677 (0.0721)	0.1076 (0.0302)	0.9518 (0.1305)	1.1102	2.0655 (0.1886)	0.2401 (0.0851)	0.0327 (0.0137)	0.6953 (0.1085)	140.7412
56	MPP	1.2124 (0.0958)		0.1276 (0.0196)	0.9820 (0.1252)	**1.4679**	1.6409 (0.1116)		0.0774 (0.0078)	0.8837 (0.0983)	**137.4363**
	REM	1.2822 (0.1230)	0.0751 (0.0708)	0.0997 (0.0300)	0.9569 (0.1287)	**2.2849**	1.9361 (0.1824)	0.2180 (0.0826)	0.0329 (0.0140)	0.7439 (0.1087)	**143.7813**
70	MPP	1.2063 (0.0970)		0.1320 (0.0197)	0.8853 (0.1259)	-3.9601	1.6542 (0.1153)		0.0813 (0.0077)	0.7689 (0.0987)	128.1531
	REM	1.2932 (0.1274)	0.0947 (0.0743)	0.0978 (0.0300)	0.8584 (0.1293)	-2.7309	2.0020 (0.1959)	0.2500 (0.0864)	0.0328 (0.0137)	0.6150 (0.1095)	136.2797
84	MPP	1.2207 (0.0991)		0.1405 (0.0197)	0.7578 (0.1268)	-10.4106	1.5809 (0.1160)		0.0810 (0.0078)	0.7936 (0.0998)	130.0665
	REM	1.3278 (0.1337)	0.1160 (0.0784)	0.0996 (0.0300)	0.7289 (0.1306)	-8.7034	1.9144 (0.1927)	0.2427 (0.0850)	0.0321 (0.0139)	0.6551 (0.1106)	138.0973

In the next chapter, we introduce some gap times based dynamic models for the recurrent event data. We fit those models for the asthma data and review how the count-based and gap times based models are useful in grabbing information in the data.

Chapter 4

Analysis of Recurrent Events Using Copula Models with Dynamic Covariates

There is often an interest in making inferences on some features of recurrent event processes over a specific inter-arrival time of a recurrent event process. Models and methods based on event counts may become inadequate and lead to wrong conclusions in such cases, in particular when there is a significant correlation between inter-arrival times in a process. In this chapter, we consider the estimation of carryover effects and trends due to the number of previous events in recurrent event processes through gap time models. For this purpose, we use copula models to determine the association between gap times. To lessen the complexity of modelling caused by the dependent gap times, we start our discussion with the first two gap times in this chapter, and we consider possible extensions for subsequent gap times later in Section 4.5.

We introduce copula based approach in Section 4.1, and formulate that approach for the first two gap times in Section 4.2. We then discuss the drawbacks in the count-based models when the consecutive gap times are dependent, through simulation study in Section 4.3. The Section 4.4 consists of real data analysis with the model introduced in Section 4.2. Next, we introduce copula models for more than two gap times in Section 4.5. We extended the models in Section 4.6 to deal with non-identical processes. Section 4.7 includes a summary of a simulation study to emphasize that our approach performs well under various settings. In the last section of this chapter, we analyze a real data set with models introduced in Section 4.5.

4.1 A Copula Based Analysis of the Dynamic Models for Recurrent Events

Inference for carryover effects and trends due to number of previous events in a counting process can be based on models for event counts as described in the previous chapter. Even though such models can be very useful in some situations, they have important limitations. First of all, event count models are not effective in the analysis of dynamic features of recurrent event processes when there are small number of events experienced by the most of the individuals in a study. This issue has been discussed by Cigsar and Lawless (2012) in the context of hypothesis testing. Another limitation of the count models is that the estimates of the dynamic features can be used for specific gap times only under special circumstances. In particular, these circumstances require the gap times of a process to be independent. This is an important drawback because the independence assumption rarely holds in many applied settings. In such settings, models based on event counts may result in a substantial bias in the estimation of dynamic covariates, and lead to misleading conclusions. We also discussed issues related to heterogeneity of dynamic covariates in the previous chapter. To address these limitations of the models based on event counts, we focus on gap time models in this chapter. The use of gap time models to model carryover effects has been briefly mentioned by Cigsar and Lawless (2012), but it has not investigated. There is a vast literature on the use of gap time models for the analysis of recurrent events. Most of the previous research has concentrated on renewal processes or some ramification of them. In this study, we introduce a copula based method to model the association between inter-arrival times within a recurrent event process. An extension of this approach can address the issues related to the dependency among gap times within individuals, as well as handle the unexplained heterogeneity across individuals, and does not suffer from the issues related to confounding as discussed in the previous chapter.

Carryover effects and trends related to $N(t^-)$ can also be modeled through mixture type of distributions of the gap times $W_k,\ k = 2, 3, \ldots$, in which a substantial mass is given over the risk window Δ after each event occurrence. Let $\{N(t); t \geq 0\}$ be a counting process of a recurrent event process and define the corresponding intensity function as

$$\lambda[t|\mathcal{H}(t)] = \alpha \exp\left[\gamma\, N(t^-) + \beta\, Z(t)\right], \qquad t > 0, \tag{4.1}$$

where α is the positive constant baseline intensity, γ is the trend effect due to the number of previous events, β is the effect of carryover effects, and $N(t^-)$ and $Z(t)$ are the number of previous events at time t and the backward recurrent time at time t, respectively. In this case, the model (4.1) can be equivalently represented by the hazard functions of the gap times W_k, $k = 1, 2, \ldots$, as follows. Let $h_k(w) = \lim_{s\to 0} \Pr(W_k < w + s|W_k \geq w)/s$, $w > 0$, be the hazard function of the k^{th} gap time W_k, $k = 1, 2, \ldots$. Then, for $w > 0$,

$$h_k(w) = \begin{cases} \alpha, & k = 1, \\ \alpha \mathrm{e}^{[(k-1)\gamma+\beta]}\, I(w \leq \Delta) + \alpha \mathrm{e}^{[(k-1)\gamma]}\, I(w > \Delta), & k = 2, 3, \ldots, \end{cases} \tag{4.2}$$

can be considered as the marginal hazard function of the k^{th} gap time W_k. The hazard function (4.2) constitutes neither a renewal process nor a delayed renewal process. For $k = 2, 3, \ldots$, the discrete mixture form of the hazard function (4.2) imposes a mass of magnitude $\alpha \exp[(k-1)\gamma + \beta]$ for the hazard of a new event over the carryover effect period Δ after each event occurrence. Note that the mass after the carryover effect period is $\alpha \exp[(k-1)\gamma]$. For positive values of β, this specification of the hazard function results in serial clustering of events due to carryover effects. Whereas for negative values of β, the hazard function results in an intermittent sparsity of events within a process. The model (4.2) can also incorporate monotonic trends related to the number of previous events. Note that, depending on the value of γ, the hazard function (4.2) monotonically increases or decreases as k increases, which results in a monotonic trend due to $N(t^-)$ in gap times. The effects of these dynamic covariates can be investigated through the model (4.2). For example, an estimation procedure or a test, by considering the hypothesis $H_0 : \beta = 0$, for carryover effects can be developed when there is a need for dependence on the number of previous events in the model.

An important issue with the development of inference procedures based on gap time models is the induced dependent censoring, which is often overlooked in the analysis of gap times in recurrent event studies. This issue arises when the gap times are not independent. For example, suppose that in a study including m independent individuals, the i^{th} individual is under observation over the interval

$[0, \tau_i]$, $i = 1, \ldots, m$. The potential censoring time for this individual is then $\tau_i^* = \tau_i - \min(W_{i1}, \tau_i)$, which is induced for the second gap time W_{i2}. Now note that if W_{i1} and W_{i2} are not independent, τ_i^* and W_{i2} are not independent too, even if τ_i is a fixed value. In this case, ignoring the dependence between W_{i1} and W_{i2} may lead to wrong conclusions in the marginal analysis of the second gap time W_{i2}. Induced dependent censoring can similarly arise for the subsequent gap times W_{ij}, $j = 3, 4, \ldots$, as well. Another important challenge is the non-identifiability of the marginal distributions of the second or subsequent gap times. This issue arises because the second or subsequent gap times are observable only if the preceding gap times of an individual are not censored. We use copula models in this study to address these issues, which cannot be readily handled by the models for event counts.

Copula functions are useful to model the joint distribution of dependent gap times. An alternative method to deal with the dependent gap times is to use individual-specific random effects to address associations among gap times (Duchateau et al., 2003). However, this method is not useful in our study because, random effects may confound with carryover effects, especially when the observation periods of individuals are not long enough. Also, random effects models do not always provide marginal distributions for gap times in simple forms. A major advantage of using copula based approach is that specific types of distributions can be used to model the marginal distributions of gap times, which can be specified according to modeling needs.

4.2 Bivariate Copula Models for the First and Second Gap Times

In this section, for simplicity we first introduce our copula based model and methodology for the first two gap times of recurrent event processes. We consider the third and subsequent gap times situation in Section 4.5. It should be noted that even though it is easier to deal with the first two gap times, there are many applications in which the second gap time is of direct interest.

Suppose that there are m independent individuals in a study. For $i = 1, \ldots, m$ and $j = 1, 2, \ldots$, let the positive valued random variable W_{ij} be the j^{th} gap time from the

i^{th} individual and let another positive valued random variable C_i be the independent right censoring time for the i^{th} individual. We focus on a sequence of first and second gap times $(W_{i1}, W_{i2}),\ i = 1, \ldots, m$. Let $(w_{i1}, w_{i2}) = (\min(W_{i1}, C_i), \min(W_{i2}, C_i - w_{i1}))$ and $(\delta_{i1}, \delta_{i2}) = (I[W_{i1} \leq C_i], I[W_{i1} + W_{i2} \leq C_i])$, respectively, be the observed gap times and their event indicators for the i^{th} individual, $i = 1, \ldots, m$, where I is an indicator function. The bivariate distribution function of W_1 and W_2 is denoted by $F(w_1, w_2)$, where $w_1 > 0$ and $w_2 > 0$. The marginal distribution functions of W_1 and W_2 are given by $F_1(w_1) = F(w_1, \infty)$ and $F_2(w_2) = F(\infty, w_2)$, respectively. With this setup, the likelihood function of the observed data $\{(w_{i1}, w_{i2}, \delta_{i1}, \delta_{i2}) : i = 1, \ldots, m\}$ can be written as

$$L = \prod_{i=1}^{m} \left[\frac{\partial^2 F(w_{i1}, w_{i2})}{\partial w_{i1} \partial w_{i2}} \right]^{\delta_{i1}\delta_{i2}} \left[\frac{\partial F_1(w_{i1})}{\partial w_{i1}} - \frac{\partial F(w_{i1}, w_{i2})}{\partial w_{i1}} \right]^{\delta_{i1}(1-\delta_{i2})} [1 - F_1(w_{i1})]^{(1-\delta_{i1})}. \tag{4.3}$$

As discussed in Section 2.4.1, a copula model for the random pair (W_1, W_2) is defined as the joint distribution function $P(W_1 \leq w_1, W_2 \leq w_2) = C(F_1(w_1), F_2(w_2))$, where $C(u_1, u_2)$ is a bivariate copula function specifying a bivariate distribution function on the unit square having uniform marginal distributions. The likelihood function (4.3) can be rewritten in terms of $C(F_1(w_1), F_2(w_2))$ by replacing $F(w_{i1}, w_{i2})$ with $C(F_1(w_1), F_2(w_2))$ in the likelihood function (4.3).

Our primary goal in this chapter is to develop models for some dynamic behaviours of recurrent event processes through gap times, in which we allow dependency. For this purpose, we use the model (4.2) for the first two gap times and combine the likelihood function (4.3) to relax the independent assumption between the first two gap times. From the hazard function (4.2), we can show that the corresponding cumulative distribution function of the k^{th} gap time W_k is of the form

$$F_k(w_k) = 1 - \exp\left[-\alpha \mathrm{e}^{(k-1)\gamma} \left\{ w_k + I(k > 1) \min(w_k, \Delta) \left(\mathrm{e}^{\beta} - 1\right) \right\} \right], \qquad w_k > 0. \tag{4.4}$$

The likelihood function (4.3) is then given by

$$L = \prod_{i=1}^{m} \left[\frac{\partial^2 C_{\alpha^*}(F_1(w_{i1}), F_2(w_{i2}))}{\partial w_{i1} \partial w_{i2}} \right]^{\delta_{i1}\delta_{i2}} \left[\frac{\partial F_1(w_{i1})}{\partial w_{i1}} - \frac{\partial C_{\alpha^*}(F_1(w_{i1}), F_2(w_{i2}))}{\partial w_{i1}} \right]^{\delta_{i1}(1-\delta_{i2})} [1 - F_1(w_{i1})]^{(1-\delta_{i1})}, \tag{4.5}$$

where α^* is the copula parameter, $F_1(w_{i1}) = 1 - \exp[-\alpha w_{i1}]$, $w_{i1} > 0$ and $F_2(w_{i2}) = 1 - \exp\left[-\alpha e^{\gamma}\left\{w_{i2} + \min(w_{i2}, \Delta)\left(e^{\beta} - 1\right)\right\}\right]$, $w_{i2} > 0$. For an Archimedean copula, defined in Section 2.4.1, the likelihood function (4.5) can be rewritten in terms of its generator function as

$$L = \prod_{i=1}^{m} \left[-\frac{\ddot{\varphi}_{\alpha^*}\left(\varphi_{\alpha^*}^{-1}[\varphi_{\alpha^*}(F_{i1}) + \varphi_{\alpha^*}(F_{i2})]\right)}{\dot{\varphi}_{\alpha^*}\left(\varphi_{\alpha^*}^{-1}[\varphi_{\alpha^*}(F_{i1}) + \varphi_{\alpha^*}(F_{i2})]\right)^3} \dot{\varphi}_{\alpha^*}(F_{i1}) \dot{\varphi}_{\alpha^*}(F_{i2}) f_{i1} f_{i2} \right]^{\delta_{i1}\delta_{i2}} \left[f_{i1} - \frac{\dot{\varphi}_{\alpha^*}(F_{i1}) f_{i1}}{\dot{\varphi}_{\alpha^*}\left(\varphi_{\alpha^*}^{-1}[\varphi_{\alpha^*}(F_{i1}) + \varphi_{\alpha^*}(F_{i2})]\right)} \right]^{\delta_{i1}(1-\delta_{i2})} [1 - F_{i1}]^{1-\delta_{i1}}, \tag{4.6}$$

where $\dot{\varphi}_{\alpha^*} = \frac{\partial \varphi_{\alpha^*}(t)}{\partial t}$, $\ddot{\varphi}_{\alpha^*} = \frac{\partial^2 \varphi_{\alpha^*}(t)}{\partial t^2}$, f_{ij} and F_{ij} are $f_j(w_{ij})$ and $F_j(w_{ij})$, respectively. Then, the log-likelihood function $\ell = \log L$ is given by

$$\begin{aligned}
&\sum_{i=1}^{m} \delta_{i1}\delta_{i2} \log\left[\ddot{\varphi}_{\alpha^*}\left(\varphi_{\alpha^*}^{-1}[\varphi_{\alpha^*}(F_{i1}) + \varphi_{\alpha^*}(F_{i2})]\right)\right] \\
&\quad - \sum_{i=1}^{m} 3\delta_{i1}\delta_{i2} \log\left[-\dot{\varphi}_{\alpha^*}\left(\varphi_{\alpha^*}^{-1}[\varphi_{\alpha^*}(F_{i1}) + \varphi_{\alpha^*}(F_{i2})]\right)\right] \\
&\quad + \sum_{i=1}^{m} \delta_{i1}\delta_{i2} \log\left[-\dot{\varphi}_{\alpha^*}(F_{i1})\right] + \sum_{i=1}^{m} \delta_{i1}\delta_{i2} \log\left[-\dot{\varphi}_{\alpha^*}(F_{i2})\right] \\
&\quad + \sum_{i=1}^{m} \delta_{i1} \log[f_{i1}] + \sum_{i=1}^{m} \delta_{i1}\delta_{i2} \log[f_{i2}] \\
&\quad + \sum_{i=1}^{m} \delta_{i1}(1-\delta_{i2}) \log\left[1 - \frac{\dot{\varphi}_{\alpha^*}(F_{i1})}{\dot{\varphi}_{\alpha^*}\left(\varphi_{\alpha^*}^{-1}[\varphi_{\alpha^*}(F_{i1}) + \varphi_{\alpha^*}(F_{i2})]\right)}\right] \\
&\quad + \sum_{i=1}^{m} (1-\delta_{i1}) \log[1 - F_{i1}].
\end{aligned} \tag{4.7}$$

Let $\boldsymbol{\theta}$ be the vector of parameters such that $\boldsymbol{\theta} = (\alpha, \beta, \gamma, \alpha^*)'$ and let $U_\alpha(\boldsymbol{\theta})$, $U_\beta(\boldsymbol{\theta})$, $U_\gamma(\boldsymbol{\theta})$ and $U_{\alpha^*}(\boldsymbol{\theta})$ be the first derivatives of the log-likelihood function $\ell(\boldsymbol{\theta})$ with respect to α, β, γ and α^*, respectively. We can obtain the maximum likelihood estimates $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ by solving the system of score equations $\boldsymbol{U}(\hat{\boldsymbol{\theta}}) = \mathbf{0}$, where $\boldsymbol{U}(\boldsymbol{\theta}) = (U_\alpha(\boldsymbol{\theta}), U_\beta(\boldsymbol{\theta}), U_\gamma(\boldsymbol{\theta}), U_{\alpha^*}(\boldsymbol{\theta}))'$ and $\mathbf{0}$ is a 4×1 vector of zeros. Let $\boldsymbol{\theta_0} = (\alpha_0, \beta_0, \gamma_0, \alpha_0^*)'$ denote the vector of the true values of the parameters in $\boldsymbol{\theta}$. Under the regularity conditions (see Cox and Hinkley, 1974, p. 281), we can show that, as $m \to \infty$, $m^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta_0})$ is asymptotically $N_4(\mathbf{0}, \boldsymbol{\mathcal{I}}(\boldsymbol{\theta_0}))$, where the $(l, r)^{th}$ entry of the 4×4 matrix $\boldsymbol{\mathcal{I}}(\boldsymbol{\theta_0})$ is given by $m^{-1} E\left[-\partial^2 \ell(\boldsymbol{\theta_0})/\partial\theta_l \partial\theta_r\right]$, $l, r = 1, 2, 3, 4$, and $(\theta_1, \theta_2, \theta_3, \theta_4)' = (\alpha, \beta, \gamma, \alpha^*)'$. This result follows from the standard large sample maximum likelihood theory, and discussed by Lawless and Yilmaz (2011) for sequentially observed data. Under mild regularity conditions, since the vector $\hat{\theta} = (\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\alpha}^*)'$ includes consistent estimators of the parameters in $\boldsymbol{\theta}$, the matrix $\boldsymbol{\mathcal{I}}(\boldsymbol{\theta_0})$ can be consistently estimated by $\boldsymbol{I}(\hat{\boldsymbol{\theta}})$, where the $(l, r)^{th}$ entry of the 4×4 matrix $\boldsymbol{I}(\hat{\boldsymbol{\theta}})$ is given by $m^{-1}[-\partial^2 \ell(\hat{\boldsymbol{\theta}})/\partial\theta_l \partial\theta_r]$, $l, r = 1, 2, 3, 4$. We maximize the log-likelihood function (4.7) with the `nlminb` function in `optimx` `R` package to obtain the estimates of the parameters α, β, γ and α^*, and their standard errors. In the next section, we present the results of a simulation study which provides some insights on the assessment of these asymptotic results for finite sample sizes.

4.3 Simulation Studies

In this section, we present the results of two simulation studies conducted to investigate the effect of dependence, an issue related to the model developed in Section 3.1, and to evaluate the finite-sample performance of the model developed in Section 4.2. The count-based model, which is developed in Section 3.1 has an important restriction; that is, it assumes that the gap times between recurrent events are independent of each other. But in real data, this independent assumption between gap times within a process may be invalid. In such a case, the performance of the count-based model developed in Section 3.1 is questionable. To investigate this issue, we conducted a simulation study and the results are provided below. In the later part of this section, we provide a simulation study which evaluates the finite-sample performance of the copula model with dynamic covariates for the first two gap times which is developed in Section 4.2.

In the first simulation study, we generate recurrent event processes, where the consecutive gap times are dependent within processes. For the data generation process, we use the algorithm given in Section 2.5.2. The Clayton copula is chosen in this algorithm to generate serially dependent data. In simulations, we assumed that everyone is observed over $[0, \tau]$, and we took two values of τ ($\tau = 2$ and 5). For each τ, we generated 1,000 realizations of m ($m = 50$ and 100) processes with the hazard function (4.2), where we considered two values of β ($\beta = 0.693$ and 1.609), two values of γ ($\gamma = 0.095$ and 0.223) and a single value for α ($\alpha = 1$). We fixed the duration of the carryover effect (i.e. Δ) at 0.0513. We specified two values for Clayton copula parameter ϕ_c ($\phi_c = 1.3333$ and 4.6667) so that the Kendall's tau τ_K was equal to 0.4 and 0.6, respectively. We then fit the count-based modulated Poison process model (3.44) to the generated serially dependent data. The intensity function of this model is

$$\lambda_i[t|\mathcal{H}_i(t)] = \alpha \, \exp\left[\gamma \, N_i(t^-) + \beta I(N_i(t^-) > 0) I(B_i(t) \leq \Delta)\right], \quad i = 1, \ldots, m, \; t > 0.$$

For each realization, we obtained estimates of the parameters in the above model by maximizing the corresponding log-likelihood function with the `optimx` package in R. The averages of the estimates over 1,000 simulation runs and their corresponding empirical standard deviations (within parenthesis) and averages of standard errors (within square brackets) are presented in Table 4.1. We also report the corresponding empirical means and standard deviations of estimated cumulative probabilities of second gap time for 20^{th}, 40^{th}, 60^{th} and 80^{th} percentile points in Table 4.2.

Table 4.1: Empirical means (and standard deviations) [and means of standard errors] of parameter estimates when the modulated Poisson process model fitted to the dependent gap times data with B=1,000 number of simulation.

Data Generated	Fitted	Simulation							
		$\tau_K = 0.4$				$\tau_K = 0.7$			
		$\tau = 2$		$\tau = 5$		$\tau = 2$		$\tau = 5$	
		m=50	m=100	m=50	m=100	m=50	m=100	m=50	m=100
$\alpha = 1$	$\bar{\hat{\alpha}}$	1.2020	1.2023	1.0720	1.0687	1.5231	1.5063	1.2046	1.1922
		(0.1535)	(0.1058)	(0.1064)	(0.0744)	(0.2287)	(0.1604)	(0.1627)	(0.1136)
		[0.1155]	[0.0806]	[0.0715]	[0.0503]	[0.1250]	[0.0876]	[0.0764]	[0.0536]
$\gamma = 0.095$	$\bar{\hat{\gamma}}$	0.0991	0.0972	0.0938	0.0936	0.1065	0.1064	0.0969	0.0970
		(0.0238)	(0.0144)	(0.0097)	(0.0066)	(0.0162)	(0.0116)	(0.0103)	(0.0073)
		[0.0157]	[0.0097]	[0.0062]	[0.0043]	[0.0060]	[0.0042]	[0.0042]	[0.0030]
$\beta = 0.693$	$\bar{\hat{\beta}}$	1.0423	1.0926	0.8265	0.8451	1.2948	1.3398	1.1296	1.1500
		(0.3364)	(0.2242)	(0.1742)	(0.1156)	(0.3262)	(0.2119)	(0.2082)	(0.1502)
		[0.1844]	[0.1249]	[0.1002]	[0.0697]	[0.1058]	[0.0728]	[0.0646]	[0.0453]
$\alpha = 1$	$\bar{\hat{\alpha}}$	1.1793	1.1725	1.0586	1.0566	1.4554	1.4408	1.2062	1.2008
		(0.1380)	(0.0981)	(0.0944)	(0.0660)	(0.1958)	(0.1410)	(0.1482)	(0.1044)
		[0.1097]	[0.0769]	[0.0672]	[0.0474]	[0.1225]	[0.0860]	[0.0755]	[0.0532]
$\gamma = 0.095$	$\bar{\hat{\gamma}}$	0.0817	0.0823	0.0918	0.0915	0.0817	0.0821	0.0872	0.0869
		(0.0145)	(0.0098)	(0.0071)	(0.0050)	(0.0125)	(0.0094)	(0.0087)	(0.0062)
		[0.0095]	[0.0062]	[0.0045]	[0.0032]	[0.0050]	[0.0035]	[0.0038]	[0.0027]
$\beta = 1.609$	$\bar{\hat{\beta}}$	1.8766	1.8912	1.6697	1.6749	2.1043	2.1171	1.8258	1.8353
		(0.1990)	(0.1372)	(0.0964)	(0.0694)	(0.1943)	(0.1376)	(0.1182)	(0.0870)
		[0.1385]	[0.0965]	[0.0710]	[0.0500]	[0.0975]	[0.0684]	[0.0574]	[0.0403]
$\alpha = 1$	$\bar{\hat{\alpha}}$	1.2034	1.2028	0.7029	0.6955	1.5809	1.5757	0.5106	0.5125
		(0.1321)	(0.0922)	(0.0633)	(0.0462)	(0.2194)	(0.1558)	(0.0569)	(0.0407)
		[0.1029]	[0.0727]	[0.0467]	[0.0328]	[0.1331]	[0.0938]	[0.0377]	[0.0267]
$\gamma = 0.223$	$\bar{\hat{\gamma}}$	0.2005	0.2001	0.1673	0.1661	0.1720	0.1711	0.1437	0.1438
		(0.0147)	(0.0103)	(0.0099)	(0.0069)	(0.0174)	(0.0116)	(0.0086)	(0.0061)
		[0.0079]	[0.0055]	[0.0035]	[0.0024]	[0.0059]	[0.0041]	[0.0033]	[0.0023]
$\beta = 0.693$	$\bar{\hat{\beta}}$	0.9731	0.9781	2.1385	2.1695	1.3448	1.3622	2.9417	2.9306
		(0.2653)	(0.1841)	(0.2638)	(0.1872)	(0.3702)	(0.2529)	(0.2569)	(0.1818)
		[0.1573]	[0.1103]	[0.0630]	[0.0443]	[0.1336]	[0.0932]	[0.0618]	[0.0434]
$\alpha = 1$	$\bar{\hat{\alpha}}$	0.6542	0.6448	0.1103	0.1099	0.4580	0.4512	0.0883	0.0875
		(0.1108)	(0.0720)	(0.0141)	(0.0100)	(0.0725)	(0.0513)	(0.0112)	(0.0075)
		[0.0664]	[0.0463]	[0.0095]	[0.0067]	[0.0504]	[0.0352]	[0.0080]	[0.0056]
$\gamma = 0.223$	$\bar{\hat{\gamma}}$	0.1843	0.1833	0.1739	0.1738	0.1586	0.1581	0.1669	0.1669
		(0.0107)	(0.0069)	(0.0042)	(0.0029)	(0.0099)	(0.0069)	(0.0068)	(0.0046)
		[0.0043]	[0.0030]	[0.0030]	[0.0021]	[0.0037]	[0.0026]	[0.0031]	[0.0022]
$\beta = 1.609$	$\bar{\hat{\beta}}$	2.7852	2.8111	4.7474	4.7498	3.6660	3.6862	5.1163	5.1207
		(0.3432)	(0.2246)	(0.1733)	(0.1216)	(0.3055)	(0.2182)	(0.2062)	(0.1386)
		[0.1040]	[0.0723]	[0.0671]	[0.0474]	[0.1013]	[0.0714]	[0.0714]	[0.0504]

The estimates of modulated Poisson process model are completely biased when the gap times between recurrent events are dependent even for the small value of Kendall's $\tau_K(= 0.4)$. For example, when the modulated Poisson process model (3.44) fitted to the data generated with $m = 100$, $\tau_K = 0.4$, $\alpha = 1$, $\beta = 1.609$, $\gamma = 0.223$ and $\tau = 2$, the empirical means of estimates of α, β and γ are $0.6448, 2.8111$ and 0.1833, respectively.

Table 4.2: Empirical means (and standard deviations) of estimated cumulative probabilities for the quantiles of W_2 when the modulated Poisson process model fitted to the dependent gap times data with B=1,000 number of simulation.

Data Generated	True Value	Simulation							
		$\tau_K = 0.4$				$\tau_K = 0.7$			
		$\tau = 2$		$\tau = 5$		$\tau = 2$		$\tau = 5$	
		m=50	m=100	m=50	m=100	m=50	m=100	m=50	m=100
$\alpha = 1$	$F_2(0.15) = 0.2$	0.2834	0.2874	0.2271	0.2278	0.3892	0.3925	0.2920	0.2921
		(0.0506)	(0.0349)	(0.0264)	(0.0183)	(0.0606)	(0.0413)	(0.0379)	(0.0277)
$\gamma = 0.095$	$F_2(0.41) = 0.4$	0.4922	0.4955	0.4312	0.4316	0.6061	0.6072	0.4983	0.4973
		(0.0490)	(0.0335)	(0.0329)	(0.0229)	(0.0526)	(0.0366)	(0.0437)	(0.0315)
$\beta = 0.693$	$F_2(0.78) = 0.6$	0.6867	0.6896	0.6304	0.6307	0.7864	0.7868	0.6903	0.6892
		(0.0451)	(0.0309)	(0.0352)	(0.0246)	(0.0435)	(0.0307)	(0.0441)	(0.0316)
	$F_2(1.41) = 0.8$	0.8618	0.8642	0.8224	0.8229	0.9237	0.9244	0.8631	0.8628
		(0.0323)	(0.0221)	(0.0285)	(0.0200)	(0.0259)	(0.0183)	(0.0329)	(0.0236)
$\alpha = 1$	$F_2(0.04) = 0.2$	0.2894	0.2909	0.2218	0.2222	0.4092	0.4100	0.2826	0.2836
		(0.0485)	(0.0350)	(0.0251)	(0.0177)	(0.0618)	(0.0438)	(0.0381)	(0.0277)
$\gamma = 0.095$	$F_2(0.26) = 0.4$	0.5013	0.5025	0.4273	0.4276	0.6279	0.6282	0.4989	0.4998
		(0.0497)	(0.0362)	(0.0322)	(0.0225)	(0.0563)	(0.0402)	(0.0455)	(0.0328)
$\beta = 1.609$	$F_2(0.63) = 0.6$	0.6874	0.6881	0.6257	0.6261	0.7901	0.7901	0.6896	0.6904
		(0.0425)	(0.0310)	(0.0330)	(0.0230)	(0.0422)	(0.0304)	(0.0432)	(0.0310)
	$F_2(1.26) = 0.8$	0.8585	0.8592	0.8186	0.8191	0.9202	0.9205	0.8622	0.8632
		(0.0300)	(0.0217)	(0.0265)	(0.0184)	(0.0244)	(0.0177)	(0.0312)	(0.0224)
$\alpha = 1$	$F_2(0.13) = 0.2$	0.2706	0.2696	0.3479	0.3511	0.4060	0.4057	0.4609	0.4587
		(0.0421)	(0.0299)	(0.0402)	(0.0287)	(0.0706)	(0.0484)	(0.0486)	(0.0346)
$\gamma = 0.223$	$F_2(0.36) = 0.4$	0.4792	0.4787	0.4618	0.4631	0.6136	0.6131	0.5297	0.5278
		(0.0413)	(0.0295)	(0.0277)	(0.0197)	(0.0526)	(0.0361)	(0.0383)	(0.0273)
$\beta = 0.693$	$F_2(0.68) = 0.6$	0.6754	0.6756	0.5891	0.5887	0.7880	0.7881	0.6118	0.6105
		(0.0393)	(0.0280)	(0.0192)	(0.0137)	(0.0403)	(0.0282)	(0.0285)	(0.0203)
	$F_2(1.24) = 0.8$	0.8545	0.8554	0.7406	0.7391	0.9228	0.9237	0.7201	0.7194
		(0.0293)	(0.0207)	(0.0174)	(0.0127)	(0.0241)	(0.0170)	(0.0209)	(0.0149)
$\alpha = 1$	$F_2(0.04) = 0.2$	0.3648	0.3683	0.4155	0.4161	0.5247	0.5276	0.4612	0.4603
		(0.0547)	(0.0365)	(0.0246)	(0.0173)	(0.0648)	(0.0459)	(0.0422)	(0.0287)
$\gamma = 0.223$	$F_2(0.20) = 0.4$	0.5375	0.5404	0.5464	0.5472	0.6819	0.6850	0.5943	0.5937
		(0.0503)	(0.0339)	(0.0269)	(0.0190)	(0.0592)	(0.0422)	(0.0450)	(0.0307)
$\beta = 1.609$	$F_2(0.53) = 0.6$	0.6425	0.6429	0.5654	0.5660	0.7334	0.7350	0.6079	0.6071
		(0.0316)	(0.0217)	(0.0251)	(0.0178)	(0.0464)	(0.0330)	(0.0430)	(0.0294)
	$F_2(1.08) = 0.8$	0.7687	0.7675	0.5959	0.5964	0.8025	0.8025	0.6300	0.6290
		(0.0237)	(0.0158)	(0.0226)	(0.0162)	(0.0323)	(0.0229)	(0.0400)	(0.0274)

The estimates of cumulative probabilities for the quantiles from modulated Poisson process models are also biased when the the gap times between recurrent events are dependent. We can see that the true values of the cumulative probabilities are not within two standard deviations interval from the empirical mean of the estimates of cumulative probabilities. For example, when the modulated Poisson process model (3.44) fitted to the data generated with $m = 100$, $\tau_K = 0.4$, $\alpha = 1$, $\beta = 1.609$, $\gamma = 0.223$ and $\tau = 5$, the empirical means of estimates of cumulative probabilities of second gap time for 20^{th}, 40^{th}, 60^{th} and 80^{th} percentile points are 0.4161, 0.5472, 0.5660 and 0.5964, respectively, and the empirical standard deviations of estimates are 0.0173, 0.0190, 0.0178 and 0.0162, respectively. These results reveal that the dependency among the gap times is a serious problem for the estimation of dynamic covariates effects through models for event counts. In particular, the modulated Poisson process model discussed in Chapter 3 is not adequate to fit the data when there is a dependency between gap times. Therefore, we next considered the copula based approach to handle this issue.

Similarly, in the second simulation study, we generated recurrent event processes where the consecutive gap times are dependent within processes. We used the same values as in the first simulation study for the parameters to generate 1,000 realizations for each combination of parameters. We then fit the gap time-based model developed in Section 4.2 only to the first two gap times. For each realization, we obtained estimates of the parameters in the model by maximizing the corresponding log-likelihood function (4.7) with the `optimx` package in R. The averages of the estimates over 1,000 simulation runs and their corresponding empirical standard deviations (within parenthesis) and averages of standard errors (within square brackets) are presented in Table 4.3. We also report the corresponding empirical means and standard deviations of the estimated cumulative probabilities of second gap time for 20^{th}, 40^{th}, 60^{th} and 80^{th} percentile points in Table 4.4.

The estimated bias in the estimates of copula based models are not significant when the gap times between recurrent events are dependent. We observed that the true values of the parameters are within two standard deviations interval from the empirical mean of the parameter estimates. For example, when the copula model fitted to the data generated with $m = 100$, $\tau_K = 0.4$, $\alpha = 1$, $\beta = 1.609$, $\gamma = 0.223$ and $\tau = 2$, the empirical means of estimates of α, β and γ are 1.0068, 1.6104 and 0.2175, respectively, and the empirical standard deviations of the estimates are 0.1082, 0.2311 and 0.1765, respectively. Through this simulation study, we can conclude that the asymptotic properties are still valid for finite samples.

Table 4.3: Empirical means (and standard deviations) [and means of standard errors] of parameter estimates when the Copula based model fitted to the dependent gap times data with B=1,000 number of simulation and the model uses only first two gap times.

Data Generated	Fitted	Simulation							
		$\tau_K = 0.4 \Rightarrow \phi_C = 1.3333$				$\tau_K = 0.7 \Rightarrow \phi_C = 4.6667$			
		$\tau = 2$		$\tau = 5$		$\tau = 2$		$\tau = 5$	
		m=50	m=100	m=50	m=100	m=50	m=100	m=50	m=100
$\alpha = 1$	$\bar{\hat{\alpha}}$	1.0204	1.0076	1.0244	1.0120	1.0192	1.0076	1.0235	1.0143
		(0.1571)	(0.1101)	(0.1472)	(0.0970)	(0.1564)	(0.1063)	(0.1427)	(0.0986)
		[0.1532]	[0.1072]	[0.1417]	[0.0990]	[0.1532]	[0.1073]	[0.1371]	[0.0960]
$\gamma = 0.095$	$\bar{\hat{\gamma}}$	0.0818	0.0908	0.0958	0.0934	0.0899	0.0920	0.0963	0.0948
		(0.2196)	(0.1558)	(0.1731)	(0.1190)	(0.1197)	(0.0845)	(0.0994)	(0.0728)
		[0.2157]	[0.1526]	[0.1701]	[0.1204]	[0.1203]	[0.0852]	[0.1004]	[0.0712]
$\beta = 0.693$	$\bar{\hat{\beta}}$	0.6195	0.6753	0.6364	0.6682	0.6730	0.6932	0.6824	0.6899
		(0.8185)	(0.2948)	(0.5617)	(0.2625)	(0.5305)	(0.1425)	(0.1939)	(0.1323)
		[0.8921]	[0.2845]	[0.5457]	[0.2579]	[0.2048]	[0.1430]	[0.1848]	[0.1288]
	$\bar{\hat{\phi}}_C$	1.4293	1.3853	1.4078	1.3803	4.9338	4.8117	4.9069	4.8078
		(0.4138)	(0.2759)	(0.3653)	(0.2463)	(0.8720)	(0.5912)	(0.8168)	(0.5412)
		[0.3838]	[0.2631]	[0.3478]	[0.2405]	[0.8406]	[0.5785]	[0.7791]	[0.5381]
$\alpha = 1$	$\bar{\hat{\alpha}}$	1.0184	1.0063	1.0254	1.0135	1.0183	1.0076	1.0220	1.0130
		(0.1521)	(0.1071)	(0.1457)	(0.1001)	(0.1553)	(0.1082)	(0.1424)	(0.0989)
		[0.1530]	[0.1071]	[0.1426]	[0.0997]	[0.1528]	[0.1071]	[0.1382]	[0.0968]
$\gamma = 0.095$	$\bar{\hat{\gamma}}$	0.0846	0.0898	0.0970	0.0906	0.0824	0.0927	0.0987	0.0955
		(0.2575)	(0.1784)	(0.1902)	(0.1349)	(0.1832)	(0.1250)	(0.1396)	(0.0947)
		[0.2515]	[0.1778]	[0.1925]	[0.1363]	[0.1762]	[0.1248]	[0.1342]	[0.0952]
$\beta = 1.609$	$\bar{\hat{\beta}}$	1.5926	1.6102	1.5828	1.6057	1.6144	1.6086	1.5990	1.6077
		(0.3444)	(0.2351)	(0.2834)	(0.2046)	(0.2278)	(0.1531)	(0.1885)	(0.1256)
		[0.3367]	[0.2357]	[0.2877]	[0.2012]	[0.2191]	[0.1551]	[0.1787]	[0.1259]
	$\bar{\hat{\phi}}_C$	1.4057	1.3820	1.3976	1.3812	4.9511	4.7978	4.8804	4.7959
		(0.3802)	(0.2660)	(0.3576)	(0.2467)	(0.9011)	(0.5997)	(0.7961)	(0.5353)
		[0.3695]	[0.2561]	[0.3415]	[0.2375]	[0.8492]	[0.5822]	[0.7822]	[0.5421]
$\alpha = 1$	$\bar{\hat{\alpha}}$	1.0186	1.0076	1.0251	1.0151	1.0182	1.0080	1.0264	1.0141
		(0.1535)	(0.1079)	(0.1415)	(0.0992)	(0.1557)	(0.1087)	(0.1408)	(0.0997)
		[0.1529]	[0.1071]	[0.1418]	[0.0993]	[0.1530]	[0.1073]	[0.1372]	[0.0958]
$\gamma = 0.223$	$\bar{\hat{\gamma}}$	0.2117	0.2199	0.2212	0.2160	0.2176	0.2195	0.2206	0.2217
		(0.2202)	(0.1522)	(0.1722)	(0.1222)	(0.1255)	(0.0903)	(0.1051)	(0.0744)
		[0.2156]	[0.1525]	[0.1719]	[0.1219]	[0.1236]	[0.0877]	[0.1034]	[0.0733]
$\beta = 0.693$	$\bar{\hat{\beta}}$	0.6205	0.6767	0.6521	0.6835	0.6843	0.6944	0.6854	0.6925
		(0.9045)	(0.2794)	(0.3764)	(0.2561)	(0.2104)	(0.1430)	(0.1927)	(0.1272)
		[0.6362]	[0.2722]	[0.3577]	[0.2465]	[0.2015]	[0.1410]	[0.1808]	[0.1265]
	$\bar{\hat{\phi}}_C$	1.4219	1.3859	1.4083	1.3794	4.9457	4.8081	4.9095	4.8154
		(0.4025)	(0.2716)	(0.3606)	(0.2402)	(0.8697)	(0.5880)	(0.8071)	(0.5572)
		[0.3812]	[0.2622]	[0.3478]	[0.2403]	[0.8389]	[0.5759]	[0.7793]	[0.5397]
$\alpha = 1$	$\bar{\hat{\alpha}}$	1.0199	1.0068	1.0241	1.0145	1.0185	1.0088	1.0244	1.0119
		(0.1540)	(0.1082)	(0.1426)	(0.1011)	(0.1529)	(0.1092)	(0.1417)	(0.0958)
		[0.1530]	[0.1070]	[0.1423]	[0.0997]	[0.1526]	[0.1070]	[0.1384]	[0.0966]
$\gamma = 0.223$	$\bar{\hat{\gamma}}$	0.2114	0.2175	0.2244	0.2208	0.2097	0.2244	0.2237	0.2210
		(0.2517)	(0.1765)	(0.1921)	(0.1404)	(0.1919)	(0.1309)	(0.1438)	(0.0996)
		[0.2529]	[0.1787]	[0.1961]	[0.1388]	[0.1854]	[0.1315]	[0.1408]	[0.0997]
$\beta = 1.609$	$\bar{\hat{\beta}}$	1.5926	1.6104	1.5849	1.6020	1.6145	1.6054	1.6017	1.6096
		(0.3305)	(0.2311)	(0.2788)	(0.1988)	(0.2320)	(0.1579)	(0.1858)	(0.1276)
		[0.3284]	[0.2299]	[0.2805]	[0.1966]	[0.2239]	[0.1588]	[0.1809]	[0.1274]
	$\bar{\hat{\phi}}_C$	1.3978	1.3770	1.3957	1.3700	4.9357	4.7808	4.8737	4.7919
		(0.3784)	(0.2583)	(0.3487)	(0.2342)	(0.8758)	(0.5936)	(0.8170)	(0.5371)
		[0.3660]	[0.2538]	[0.3405]	[0.2353]	[0.8436]	[0.5789]	[0.7815]	[0.5426]

The estimated bias in the estimates of cumulative probabilities for the quantiles from copula based models is also not significant in the cases considered in Table 4.4.

For example, when the copula model fitted to the data generated with $m = 100$, $\tau_K = 0.4$, $\alpha = 1$, $\beta = 1.609$, $\gamma = 0.223$ and $\tau = 5$, the empirical means of estimates of cumulative probabilities of second gap time for 20^{th}, 40^{th}, 60^{th} and 80^{th} percentile points are 0.2018, 0.4032, 0.6035 and 0.8019, respectively, with empirical standard deviations 0.0284, 0.0336, 0.0352 and 0.0322, respectively, while true values of those are 0.2, 0.4, 0.6 and 0.8, respectively. These results reveal that the problem due to dependency among the gap times is correctly handled by the copula based models.

Table 4.4: Empirical means (and standard deviations) of estimated cumulative probabilities for the quantiles of W_2 when the Copula based model fitted to the dependent gap times data with B=1,000 number of simulation and the model uses only first two gap times.

Data Generated	True Value	Simulation							
		$\tau_K = 0.4$				$\tau_K = 0.7$			
		$\tau = 2$		$\tau = 5$		$\tau = 2$		$\tau = 5$	
		m=50	m=100	m=50	m=100	m=50	m=100	m=50	m=100
$\alpha = 1$	$F_2(0.15) = 0.2$	0.2022	0.2014	0.2041	0.2015	0.2023	0.2012	0.2040	0.2025
		(0.0366)	(0.0265)	(0.0341)	(0.0236)	(0.0291)	(0.0204)	(0.0265)	(0.0181)
$\gamma = 0.095$	$F_2(0.41) = 0.4$	0.4028	0.4016	0.4065	0.4026	0.4032	0.4012	0.4062	0.4038
		(0.0560)	(0.0405)	(0.0446)	(0.0307)	(0.0510)	(0.0357)	(0.0425)	(0.0292)
$\beta = 0.693$	$F_2(0.78) = 0.6$	0.6004	0.6002	0.6062	0.6025	0.6017	0.6002	0.6059	0.6037
		(0.0696)	(0.0502)	(0.0523)	(0.0363)	(0.0627)	(0.0439)	(0.0508)	(0.0351)
	$F_2(1.41) = 0.8$	0.7956	0.7977	0.8029	0.8010	0.7979	0.7984	0.8029	0.8021
		(0.0643)	(0.0460)	(0.0463)	(0.0323)	(0.0565)	(0.0395)	(0.0447)	(0.0311)
$\alpha = 1$	$F_2(0.04) = 0.2$	0.2007	0.2015	0.2023	0.2021	0.2020	0.2009	0.2030	0.2023
		(0.0429)	(0.0312)	(0.0421)	(0.0302)	(0.0311)	(0.0224)	(0.0302)	(0.0211)
$\gamma = 0.095$	$F_2(0.26) = 0.4$	0.4031	0.4021	0.4057	0.4032	0.4030	0.4017	0.4059	0.4037
		(0.0550)	(0.0393)	(0.0475)	(0.0331)	(0.0500)	(0.0363)	(0.0419)	(0.0293)
$\beta = 1.609$	$F_2(0.63) = 0.6$	0.6020	0.6009	0.6067	0.6029	0.6010	0.6010	0.6063	0.6038
		(0.0697)	(0.0490)	(0.0514)	(0.0356)	(0.0659)	(0.0475)	(0.0504)	(0.0350)
	$F_2(1.26) = 0.8$	0.7964	0.7978	0.8037	0.8011	0.7959	0.7985	0.8032	0.8021
		(0.0672)	(0.0473)	(0.0460)	(0.0324)	(0.0632)	(0.0450)	(0.0459)	(0.0318)
$\alpha = 1$	$F_2(0.13) = 0.2$	0.2020	0.2015	0.2040	0.2023	0.2021	0.2013	0.2041	0.2025
		(0.0374)	(0.0276)	(0.0347)	(0.0245)	(0.0292)	(0.0211)	(0.0266)	(0.0186)
$\gamma = 0.223$	$F_2(0.36) = 0.4$	0.4029	0.4019	0.4064	0.4028	0.4030	0.4014	0.4063	0.4035
		(0.0561)	(0.0404)	(0.0433)	(0.0301)	(0.0509)	(0.0370)	(0.0418)	(0.0296)
$\beta = 0.693$	$F_2(0.68) = 0.6$	0.6005	0.6007	0.6061	0.6023	0.6013	0.6003	0.6059	0.6034
		(0.0697)	(0.0499)	(0.0509)	(0.0354)	(0.0631)	(0.0460)	(0.0501)	(0.0357)
	$F_2(1.24) = 0.8$	0.7957	0.7981	0.8028	0.8006	0.7975	0.7982	0.8029	0.8017
		(0.0641)	(0.0456)	(0.0453)	(0.0318)	(0.0573)	(0.0416)	(0.0442)	(0.0316)
$\alpha = 1$	$F_2(0.04) = 0.2$	0.2004	0.2015	0.2020	0.2018	0.2019	0.2012	0.2033	0.2019
		(0.0415)	(0.0308)	(0.0401)	(0.0284)	(0.0307)	(0.0223)	(0.0292)	(0.0203)
$\gamma = 0.223$	$F_2(0.20) = 0.4$	0.4022	0.4021	0.4048	0.4032	0.4030	0.4025	0.4059	0.4030
		(0.0550)	(0.0401)	(0.0483)	(0.0336)	(0.0486)	(0.0360)	(0.0415)	(0.0289)
$\beta = 1.609$	$F_2(0.53) = 0.6$	0.6017	0.6010	0.6061	0.6035	0.6011	0.6024	0.6064	0.6028
		(0.0673)	(0.0477)	(0.0510)	(0.0352)	(0.0644)	(0.0474)	(0.0494)	(0.0344)
	$F_2(1.08) = 0.8$	0.7966	0.7980	0.8032	0.8019	0.7961	0.7998	0.8032	0.8012
		(0.0652)	(0.0458)	(0.0460)	(0.0322)	(0.0623)	(0.0452)	(0.0452)	(0.0317)

4.4 Application: Recurrent Asthma Attacks in Children

We introduced the recurrent asthma attacks data set in Section 1.2.2 and analyzed it in Section 3.4 to provide estimates of carryover effects and trend due to the number of previous events. In the previous analysis, we utilize the dynamic models for event counts. In this section, we discuss the presence of carryover effects in the marginal distribution of the second gap time. The main purpose of this section is to illustrate the method introduced in Section 4.2, as well as compare our results with those found in Section 3.4. We would like to note that even though we use the asthma data set to illustrate the methods, there are recurrent event studies in which the main goal is to dynamically model the second gap time. For example, Gruneir et al. (2018) analyzed on observational administrative data set to identify the risk factors for repeat emergency department visits from nursing homes in Ontario. In such studies, there is a special interest in the analysis of the marginal distribution of second gap times. Unfortunately, we could not use this data set here because of privacy issues, but our methods can be applied to analyze such data sets.

It is worth underlying that our analysis in Section 3.4 was based on the models for event counts. With that approach, we detected a significant unexplained heterogeneity in the number of events across subjects in both control and treatment groups. To deal with such an unexplained heterogeneity, we apply a random effects model by assigning a subject specific random effect to every individual process under observation. We discuss issues with modeling dynamic covariates arising from this approach in Section 3.2. The analysis in this section focuses on the second gap time. Every subject has either one or two events. Therefore, the approach in this section does not suffer from the aforementioned type of unexplained heterogeneity regarding the observed number of event counts per process.

The recurrent asthma data set, given by Duchateau et al. (2003), is obtained from a prevention trial in infants with a high risk of having asthma attacks. More information and some descriptive statistics about the data can be found in Section 3.4. Subjects are randomly assigned either to an active drug group (treatment group) or to a placebo group (control group). Here we present our results separately for each group. An asthma attack experienced by subjects is defined as an event. The occurrence time of events were recorded in days after the start of follow-up, which was the day zero. As discussed in Section 3.4, the resolution of an asthma attack may last longer than a day. During their asthma attacks, subjects are considered as "risk-free". We therefore discard the risk-free periods from the data so that second and subsequent gap times start after the resolution of asthma attacks. We do not assess the effects of risk-free periods on the distribution of the second gap time here, but this can be done by including the duration of risk-free periods as a covariate in the model and then a test for its significance can be applied.

There are $m = 232$ subjects included in the data set. Among them, $m_1 = 119$ are in the control group and $m_2 = 113$ are in the treatment group. The data set used to fit models includes the values of $\{(w_{i1}, w_{i2}, \delta_i, x_i, \tau_i); i = 1, \ldots, m\}$, where w_{i1} is the first gap time, w_{i2} is either the second gap time or censored gap time; that is, $\tau_i - w_{i1}$, δ_i takes the value of 1 if w_{i2} is a complete gap time; otherwise, it is equal to 0, x_i is a treatment indicator ($x_i = 1$, if the i^{th} subject is in the treatment group; otherwise, $x_i = 0$), and τ_i is the end of follow-up time. For this specific data set, all of the w_{i1} are complete observations. There are 37 and 50 incomplete w_{i2} values in the control and treatment groups, respectively. We model the dependency between the first two gap times with the Clayton copula (CLY), the Gumbel-Hougaard copula (G_H) and the 2-parameter copula (2PR) models. These models are explained in Section 2.4.1. We also fit a model in which the first two gap times are assumed independent (IND).

Following the discussion in Section 4.1, the marginal hazard function of the first gap time W_1 is $h_1(w) = \alpha$, $w > 0$. The marginal hazard function of the second gap time W_2 is given by

$$h_2(w) = \alpha e^{\gamma+\beta} I(w \leq \Delta) + \alpha e^{\gamma} I(w > \Delta), \qquad w > 0. \tag{4.8}$$

We consider four cases of the hazard function (4.8) to fit the data. The first one is the "Null" model, in which we take $\gamma = 0$ and $\beta = 0$ in the model (4.8). The second model is a "Carryover" effects model. In this model, the value of the parameter γ in (4.8) is zero. The third model is the "Trend" model, which is the model (4.8) with $\beta = 0$. The last model is the "Hybrid" model, which includes both carryover effects and trend with the hazard function (4.8). We obtain the estimates of the parameters and their standard errors by maximizing the log-likelihood function (4.7) with the `nlminb` function in `optimx` `R` package. As suggested in our analysis presented in Section 3.4, we fixed the value of Δ in the model (4.8) at 56 days (≈ 0.1533 years).

We respectively present the results for the control and treatment groups in Tables 4.5 and 4.6, which show the maximum likelihood estimates of the model parameters $\boldsymbol{\theta} = (\alpha, \gamma, \beta_{\Delta=0.1533}, \phi_{C_{1,2}}, \theta_{G_{1,2}})'$, and their standard errors in parenthesis, estimated Kendall's tau $\hat{\tau}_{K_{1,2}}$ and the value of the maximum log-likelihood calculated at $\hat{\boldsymbol{\theta}} = (\hat{\alpha}, \hat{\gamma}, \hat{\beta}_{\Delta=0.1533}, \hat{\phi}_{C_{1,2}}, \hat{\theta}_{G_{1,2}})'$. Note that the Clayton and Gumbel-Hougaard copulas require the estimation of a single copula parameter ($\phi_{C_{1,2}}$ for CLY and $\theta_{G_{1,2}}$ for G_H), which represents the dependence level between the first two gap times, whereas the 2-parameter copula requires the estimation of both parameters $\phi_{C_{1,2}}$ and $\theta_{G_{1,2}}$ simultaneously.

In the control group 53 asthma attacks (out of 82) occurred within 56 days after the resolution of the first asthma attacks. Under the null model with no dependence (Null-IND), the expected number of asthma attack over that time period is 25.97 ($= 2.0656/365.25 \times 56 \times 82$). Under the standard normal approximations, all estimates of carryover effects presented in Table 4.5 are significant. For example, in the hybrid model with Gumbel-Hougaard (Hybrid-G_H) copula, $\hat{\beta}_{\Delta=0.1533} = 1.5093$, which gives 95% confidence interval (c.i.) with limits 1.0477 and 1.9709, and a Wald type test of $H_0 : \beta = 0$ gives a $p-$value close to 0 (< 0.005).

We next consider the presence of dependency between the first two gap times in the control group. To do this, we use the hybrid model with the Gumbel-Hougaard (Hybrid-G_H) copula because it has the smallest Akaike information criterion (AIC) value. In the hybrid model, independence can be tested by comparing the estimated log-likelihood values of Hybrid-G_H and Hybrid-IND copula models. A likelihood ratio statistic $\Lambda(\theta_{G_{1,2}} = 1) = 2(\ell_{\text{G\_H}} - \ell_{\text{IND}}) = 2(-31.0958 + 35.9767) = 9.7618$ can be used for testing $H_0 : \theta_{G_{1,2}} = 1$. It should be noted that the value of the parameter $\theta_{G_{1,2}}$ under the null hypothesis is a boundary point, and standard asymptotic theory cannot be applied (Self and Liang, 1987). However, a $p-$value can be calculated by using the limiting distribution $\Pr(\Lambda(\theta_{G_{1,2}} = 1) \leq q) = 0.5 + 0.5\Pr(\chi^2_{(1)} \leq q)$, which gives a $p-$value as $0.5\Pr(\chi^2_{(1)} \geq 9.7618) = 0.0009$ so that we reject $H_0 : \theta_{G_{1,2}} = 1$ in favor of $H_1 : \theta_{G_{1,2}} > 1$. Therefore, we conclude that there is a small level of statistically significant dependence between the first two gap times.

Table 4.5: Parameter estimates and standard errors (within parenthesis) from gap times models which consider only first two gap times for the control group in asthma data.

Models		Estimates						$\ell(\hat{\boldsymbol{\theta}})$
		$\hat{\alpha}$	$\hat{\gamma}$	$\hat{\beta}_{\Delta=0.1533}$	$\hat{\phi}_{C_{1,2}}$	$\hat{\theta}_{G_{1,2}}$	$\hat{\tau}_{K_{1,2}}$	
Null	IND	2.0656 (0.1457)						-55.1941
	CLY	1.9818 (0.1504)			0.2676 (0.0906)		0.1180 (0.0353)	-49.2286
	G_H	1.9984 (0.1568)				1.1373 (0.0740)	0.1208 (0.0572)	-53.0727
	2PR	1.9818 (0.1504)			0.2676 (0.0906)	≈ 1 (≈ 0)	0.1180 (0.0353)	-49.2286
Carryover	IND	1.7561 (0.1444)		0.8398 (0.1601)				-43.3101
	CLY	1.7545 (0.1452)		0.7459 (0.1577)	0.2561 (0.1028)		0.1135 (0.0404)	-39.2825
	G_H	1.7366 (0.1485)		0.8029 (0.1536)		1.1416 (0.0805)	0.1241 (0.0618)	-41.4453
	2PR	1.7545 (0.1452)		0.7459 (0.1577)	0.2561 (0.1028)	≈ 1 (≈ 0)	0.1135 (0.0404)	-39.2825
Trend	IND	2.1333 (0.1956)	-0.0773 (0.1435)					-55.0484
	CLY	2.0224 (0.1917)	-0.0486 (0.1377)		0.2644 (0.0905)		0.1168 (0.0353)	-49.1662
	G_H	2.0844 (0.1965)	-0.1052 (0.1339)			1.1426 (0.0742)	0.1248 (0.0568)	-52.7629
	2PR	2.0224 (0.1917)	-0.0486 (0.1377)		0.2644 (0.0905)	≈ 1 (≈ 0)	0.1168 (0.0353)	-49.1662
Hybrid	IND	2.1333 (0.1956)	-0.7401 (0.2071)	1.3853 (0.2310)				-35.9767
	CLY	2.1169 (0.1939)	-0.7382 (0.2063)	1.3306 (0.2331)	0.2724 (0.1084)		0.1199 (0.0420)	-31.9031
	G_H	2.1098 (0.1938)	-0.8495 (0.2003)	1.5093 (0.2355)		1.2489 (0.0942)	0.1993 (0.0604)	-31.0958
	2PR	2.1139 (0.1944)	-0.8332 (0.2092)	1.4787 (0.2574)	0.0521 (0.1795)	1.2060 (0.1690)	0.1919 (0.0659)	-31.0517

In the Hybrid-G_H model, the carryover effect is significant ($\hat{\beta}_{\Delta=0.1533} = 1.5093$, 95% c.i. (1.0477,1.9709), $p-$value ≈ 0), and it gives a relative risk (RR) of 4.52. The trend component in the Hybrid-G_H model is also significant ($\hat{\gamma} = -0.8495$, 95% c.i. (-1.2421,-0.4569), $p-$value ≈ 0). Note that the term $N(t^{-})$ in model (4.8) results in a change in the baseline hazard α to αe^{γ} for the second gap time. A comparison between the Hybrid-G_H model and the Carryover-G_H model also reveals that γ is significant (a likelihood ratio test with $p-$value ≈ 0). Furthermore, the estimate of the carryover effect β found under the event counts model (Hybrid-REM model) in Section 3.4 was 0.7439, which is quite smaller than the one found under the Hybrid-G_H model. Also, the parameter γ was not significant in the Hybrid-REM model of Section 3.4 ($p-$value $= 0.0184$ with a LR test). We therefore conclude that the analysis based on the gap time model gives different results from the analysis based on the event counts model for the control group.

Table 4.6: Parameter estimates and standard errors (within parenthesis) from gap times models which consider only first two gap times for the treatment group in asthma data.

Models		Estimates $\hat{\alpha}$	$\hat{\gamma}$	$\hat{\beta}_{\Delta=0.1533}$	$\hat{\phi}_{C_{1,2}}$	$\hat{\theta}_{G_{1,2}}$	$\hat{\tau}_{K_{1,2}}$	$\ell(\hat{\boldsymbol{\theta}})$
Null	IND	1.5084 (0.1137)						-103.6580
	CLY	1.5047 (0.1152)			0.0172 (0.0800)		0.0085 (0.0393)	-103.6342
	G_H	1.5084 (0.1137)				≈ 1 (≈ 0)	≈ 0 (≈ 0)	-103.6580
	2PR	1.5047 (0.1152)			0.0172 (0.0800)	≈ 1 (≈ 0)	0.0085 (0.0393)	-103.6342
Carryover	IND	1.3095 (0.1123)		0.8677 (0.1799)				-93.8326
	CLY	1.3095 (0.1123)		0.8672 (0.1808)	0.0026 (0.1003)		0.0013 (0.0500)	-93.8323
	G_H	1.3095 (0.1123)		0.8677 (0.1799)		≈ 1 (≈ 0)	≈ 0 (≈ 0)	-93.8326
	2PR	1.3095 (0.1123)		0.8676 (0.1797)	0.0004 (≈ 0)	≈ 1 (0.0004)	0.0002 (≈ 0)	-93.8326
Trend	IND	1.6988 (0.1598)	-0.3020 (0.1572)					-101.7659
	CLY	1.6982 (0.1630)	-0.3017 (0.1583)		0.0013 (0.0749)		0.0007 (0.0374)	-101.7657
	G_H	1.6988 (0.1598)	-0.3020 (0.1572)			≈ 1 (≈ 0)	≈ 0 (≈ 0)	-101.7659
	2PR	1.6988 (0.1597)	-0.3020 (0.1572)		≈ 0 (≈ 0)	≈ 1 (0.0003)	≈ 0 (≈ 0)	-101.7659
Hybrid	IND	1.6988 (0.1598)	-1.0144 (0.2288)	1.6218 (0.2617)				-81.7701
	CLY	1.6980 (0.1598)	-1.0150 (0.2288)	1.6235 (0.2621)	0.0177 (0.1148)		0.0088 (0.0564)	-81.7579
	G_H	1.6987 (0.1599)	-1.0430 (0.2320)	1.6499 (0.2669)		1.0502 (0.0913)	0.0478 (0.0828)	-81.6101
	2PR	1.6986 (0.1599)	-1.0430 (0.2319)	1.6499 (0.2669)	≈ 0 (0.0001)	1.0502 (0.0913)	0.0478 (0.0828)	-81.6101

We conducted a similar analysis for the treatment group. The results are presented in Table 4.6. In this case, the Hybrid-IND model gives the smallest AIC value (169.5402). Also, none of the copula parameters are significant in the models considered in Table 4.6. We therefore use the Hybrid-IND model to make inference on the covariates. Note that this model assumes independence between the first two gap times, and uses only the first two gap times to fit the data. In the Hybrid-MPP model used in Section 3.4, the results are based on all gap times. The parameter β in the Hybrid-IND model is significant ($\hat{\beta}_{\Delta=0.1533} = 1.6218$, 95% c.i. (1.1089,2.1347), $p-$value ≈ 0, and RR=5.06). The estimate of β in the Hybrid-MPP was 0.9820 (RR=2.67). We also observe in the Hybrid-IND model in Table 4.6 that the parameter γ is significant ($\hat{\gamma} = -1.0144$, 95% c.i. (-1.4628,-0.5659), $p-$value ≈ 0). The estimate of γ in the Hybrid-MPP model was 0.1276, which is in the opposite direction of the estimate found in the Hybrid-IND model. Our analysis here and Section 3.4 show that the results should be carefully interpreted if the gap times are of interest, and we recommend to use a specific gap time model in such applications.

In the final part of our analysis, we discuss the adequacy of the models considered in Table 4.5 and 4.6 to fit the data. For this purpose, we compare the estimates of the marginal survival functions of the second gap times with their corresponding step function estimates. To do this, we use a two-stage semi-parametric method, proposed by Lawless and Yilmaz (2011), for estimation of the marginal survival function of the second gap time. In this estimation method the marginal distributions are estimated with step functions, and combined with a parametric copula model so that a step function estimate, similar to the Kaplan-Meier estimate, of the marginal survival function of the second gap time (i.e. $S_2(t) = 1 - F_2(t), t > 0$) can be obtained. We present the results in Figure 4.1 as estimated survival functions $\hat{S}_2(t)$. For the comparison, we use all the models considered in Table 4.5 and 4.6 as well as in Table 3.8. For these models, we estimate $S_2(t)$ by plugging-in the corresponding estimates of the parameters.

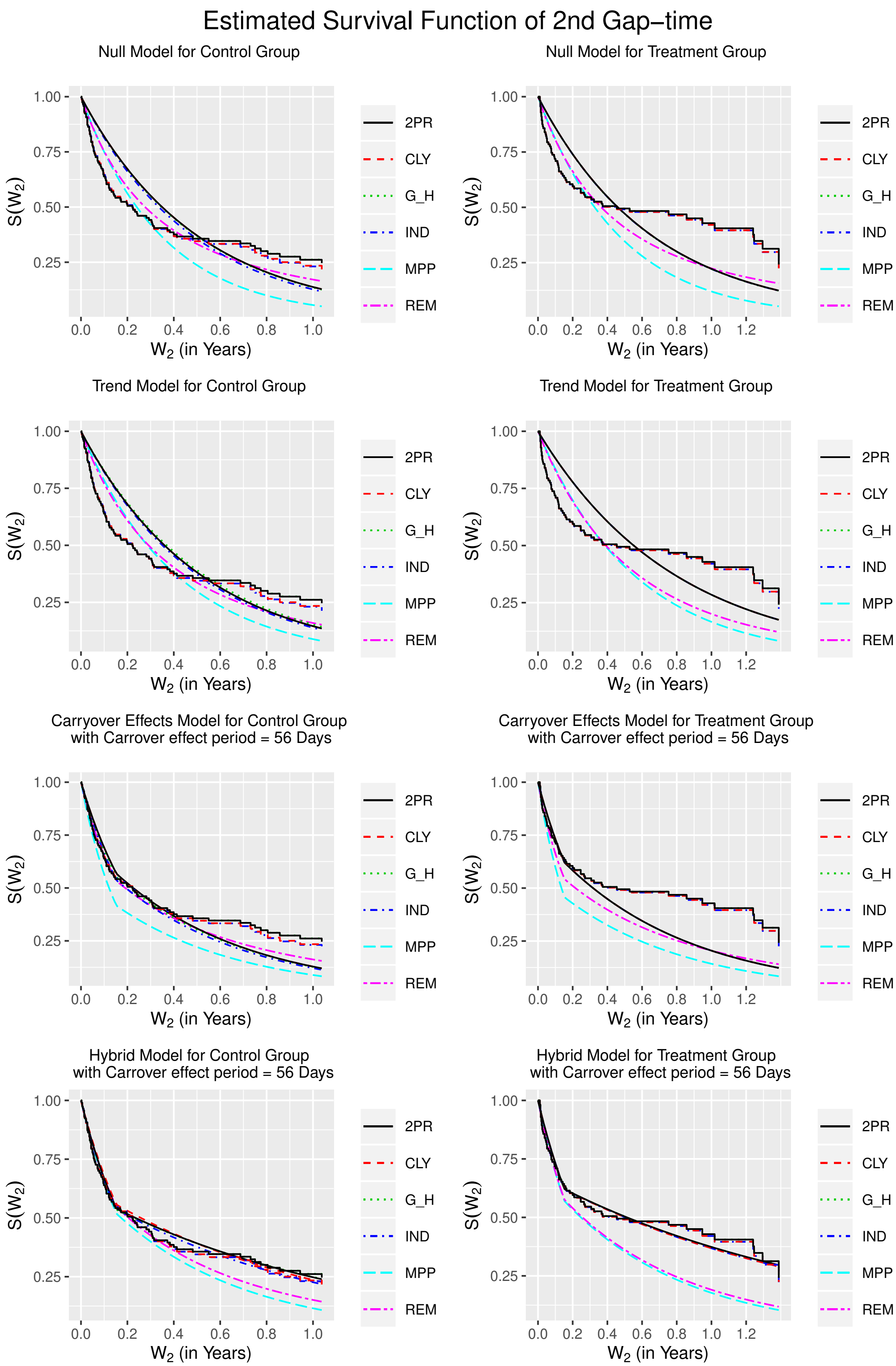

Figure 4.1: Modulated Poisson process (MPP) model based maximum likelihood, random effects model (REM) based maximum likelihood, copula-based maximum likelihood (2PR: 2 parameter copula, CLY: Clayton copula, G_H: Gumbel-Hougaard copula, IND: independent copula) and the corresponding two-stage semiparametric estimates of $S_2(t) = 1 - F_2(t)$ for the treatment and the control groups. The curves are from maximum likelihood estimates and the step functions are from two-stage semiparametric estimates.

It is clear by comparing the step function estimates of $S_2(t)$ with the parametric estimates of $S_2(t)$ that the hybrid models fit the data the best, suggesting that the two existing dynamic features in the data are correctly captured by the models introduced in this section. Among the models under the group of hybrid models, the copula based models (CLY, G_H, 2PR) seem better fit than the models for event counts (MPP, REM) for both the control and treatment groups.

A final note here is that the two-stage semiparametric estimation of $S_2(t)$ can be used to get a rough idea about the value for the duration of the carryover effect. For example, in the plots given in Figure 4.1, a choice of $\Delta = 0.1533$ years (=56 days) seems appropriate for the asthma data in both control and treatment groups.

4.5 Copula Models for Series of Gap Times

Our goal in this section is to extend the discussion of Section 4.2 to the third and subsequent gap times. We will follow a similar setup given in Section 4.2.

Suppose that there are m independent individuals in a study. Let W_{ij} be the j^{th} gap time from the i^{th} individual and let C_i be the independent right censoring time for the i^{th} individual, where $i = 1, \ldots, m$ and $j = 1, 2, \ldots$. We focus on the whole sequence of the gap times $(W_{i1}, W_{i2}, \ldots, W_{in_i})$, $i = 1, \ldots, m$ with the joint distribution function

$$\Pr(W_{i1} \le w_1, W_{i2} \le w_2, \ldots, W_{in_i} \le w_{n_i}), \tag{4.9}$$

defined for every set of positive real numbers $w_1, w_2, \ldots, w_{n_i}$. To deal with the censoring, we let $w_{ij} = \min(W_{ij},\ C_i - T_{i,j-1})$ and $\delta_{ij} = I[T_{ij} \le C_i]$ respectively be the observed gap time and its event indicator obtained from the i^{th} individual, where $T_{ij} = \sum_{l=1}^{j} W_{il}$ is the j^{th} event time of the i^{th} individual, and I is a $0-1$ valued indicator function.

To model serial dependency, we need to assume a structure for the dependency between the gap times within a subject. We first start with a dependent structure, which is useful in many applications and mathematically tractable. In this section, we therefore consider the case in which only the two consecutive gap times are dependent. That is, the k^{th} gap time depends only on the $(k-1)^{\text{st}}$ gap time and does not depend on any other previous ($(k-2)^{\text{nd}}$,...,1^{st}) gap times. We refer to this dependence structure as the Markov dependence structure, a term sometimes used in time series analysis to describe similar type of dependency (Joe, 1997, Chapter 8). Note that the Markov dependence structure is a special case of the D-Vine copula. A brief introduction of the D-Vine copula is available in Section 2.4.2.

We denote the bivariate distribution function of random variables $W_{ij} \geq 0$ and $W_{i,j+1} \geq 0$ as $F_{j,j+1}(w_{ij}, w_{i,j+1})$, the conditional distribution function of $W_{i,j+1} \geq 0$ given $W_{ij} = w_{ij}$ as $F_{j+1|j}(w_{i,j+1}|w_{ij})$ and the marginal distribution function of W_{ij} as $F_j(w_{ij}) = F_{j-1,j}(\infty, w_2)$. The corresponding bivariate joint density, conditional density and the marginal density functions are denoted by $f_{j,j+1}(w_{ij}, w_{i,j+1})$, $f_{j+1|j}(w_{i,j+1}|w_{ij})$ and $f_j(w_{ij})$, respectively. Then, the likelihood of the observed data $\{(w_{i1}, \ldots, w_{i,n_i+1}, \delta_{i1}, \ldots, \delta_{i,n_i+1}) : i = 1, \ldots, m\}$ for m independent processes is

$$L = \prod_{i=1}^{m} L_i, \tag{4.10}$$

where n_i is the number of events observed during the observation period for the i^{th} individual and

$$\begin{aligned} L_i = {} & f_1(w_{i1})^{\delta_{i1}} \left[1 - F_1(w_{i1})\right]^{1-\delta_{i1}} \\ & \times \left[\left\{ \prod_{j=2}^{n_i} f_{j|j-1}(w_{ij}|w_{i,j-1}) \right\}^{\delta_{i2}} \left[1 - F_{n_i+1|n_i}(w_{i,n_i+1}|w_{in_i})\right]^{1-\delta_{i,n_i+1}} \right]^{\delta_{i1}} . \end{aligned} \tag{4.11}$$

Following the discussion in Section 4.1, we focus on the discrete mixture model (4.2), where the marginal hazard function of the k^{th} gap time is given by

$$h_k(w) = \begin{cases} \alpha, & k = 1, \\ \alpha \mathrm{e}^{[(k-1)\gamma+\beta]}\, I(w \leq \Delta) + \alpha \mathrm{e}^{[(k-1)\gamma]}\, I(w > \Delta), & k = 2, 3, \ldots, \end{cases} \tag{4.12}$$

for any $w > 0$. For $k = 1, 2, \ldots$, the corresponding marginal cumulative distribution function and density function of k^{th} gap time are given by

$$F_k(w) = 1 - \exp\left[-\alpha \mathrm{e}^{(k-1)\gamma} \left\{w + I(k > 1) \min(w, \Delta) \left(\mathrm{e}^{\beta} - 1\right)\right\}\right], \quad w > 0, \tag{4.13}$$

and

$$f_k(w) = \begin{cases} \alpha \ \exp\left[(k-1)\gamma + \beta I(k>1) - \alpha e^{(k-1)\gamma}\left\{w e^{\beta I(k>1)}\right\}\right], & w \le \Delta, \\ \alpha \ \exp\left[(k-1)\gamma - \alpha e^{(k-1)\gamma}\left\{w + \Delta\left(e^{\beta I(k>1)} - 1\right)\right\}\right], & w > \Delta, \end{cases} \tag{4.14}$$

respectively. According to Sklar's Theorem (Nelsen, 2006, p. 21), for any $w_1, w_2 > 0$, there exists a unique copula function such that

$$F(w_1, w_2) = C_{\alpha^*}(F_1(w_1), F_2(w_2)), \tag{4.15}$$

where the subscript α^* represents the copula parameter to be estimated. As discussed in Section 2.4.1, there are many parametric copula functions available to use in (4.15). We choose the Clayton copula

$$C_{\alpha^*}(u_1, u_2) = \left(u_1^{-\alpha^*} + u_2^{-\alpha^*} - 1\right)^{-1/\alpha^*}, \qquad u_1, u_2 \in [0,1] \tag{4.16}$$

where $\alpha^* \in [-1, \infty)\backslash\{0\}$. This is because it is an excellent candidate to investigate lower tail dependency (Embrechts et al., 2003), and a popular choice in survival studies (Lawless and Yilmaz, 2011). Clayton copula can handle both positive and negative associations without any rotation, whereas most of the other copulas have to be rotated to capture the negative association. We would like to note that copula assumption is testable. Other choices include Gumbel-Hougaard copula, Frank copula and Joe copula.

Since we pick the copula function to model the joint distribution, the conditional density and distribution functions in (4.10) can be replaced by their corresponding copula versions. The full likelihood function which contains all the available gap times is then

$$\begin{aligned} L = \prod_{i=1}^{m} & f_1(w_{i1})^{\delta_{i1}} \left[1 - F_1(w_{i1})\right]^{1-\delta_{i1}} \\ & \times \left[\left\{ \prod_{j=2}^{n_i} \frac{\partial^2 C_{\alpha^*}(F_j(w_{ij}), F_{j-1}(w_{i,j-1}))}{\partial F_j(w_{ij}) \partial F_{j-1}(w_{i,j-1})} f_j(w_{ij}) \right\}^{\delta_{i2}} \right. \\ & \left. \times \left[1 - \frac{\partial C_{\alpha^*}(F_{n_i+1}(w_{i,n_i+1}), F_{n_i}(w_{in_i}))}{\partial F_{n_i}(w_{in_i})} \right]^{1-\delta_{i,n_i+1}} \right]^{\delta_{i1}}. \end{aligned} \tag{4.17}$$

The Archimedean copulas including the Clayton copula (4.16) can be too restrictive when $k \geq 3$ because it assigns a single parameter α^* to explain all the consecutive dependency between the gap times. An alternative is to assign separate parameters α_j^*, $j = 1, \ldots, n^*$, where $n^* = \max(n_1, \ldots, n_m)$, for each consecutive dependencies. In this case, the parameter α_j^* corresponds to the copula parameter which captures the dependency between the j^{th} and $(j-1)^{\text{st}}$ gap times. An important remark is, however, that an estimability issue may arise. For example, if there exists only one individual with the highest number of events, then the copula parameter $\alpha_{n^*}^*$ corresponding to the last two gap times may arise estimable issue due to lack of information. Therefore, we modify the likelihood function (4.17) to include only the first fixed number of consecutive gap times in a way that the parameters in (4.17) can be estimated from the data. To do this, we need to specify a value for k in the hazard function (4.12). Suppose that k is fixed at such a value. Then, the likelihood function for first $k(> 2)$ gap times for m independent individuals can be written as

$$\begin{aligned} L = &\prod_{i=1}^{m} f_1(w_{i1})^{\delta_{i1}} \left[1 - F_1(w_{i1})\right]^{1-\delta_{i1}} \\ &\times \left[\prod_{j=2}^{\min(n_i+1,k)} \left\{ \frac{\partial^2 C_{\alpha_{j-1}^*}(F_j(w_{ij}), F_{j-1}(w_{i,j-1}))}{\partial F_j(w_{ij}) \partial F_{j-1}(w_{i,j-1})} f_j(w_{ij}) \right\}^{\delta_{ij}} \right. \\ &\left. \times \left\{ 1 - \frac{\partial C_{\alpha_{j-1}^*}(F_j(w_{ij}), F_{j-1}(w_{i,j-1}))}{\partial F_{j-1}(w_{i,j-1})} \right\}^{1-\delta_{ij}} \right]^{\delta_{i1}} . \end{aligned} \tag{4.18}$$

For an Archimedean copula, the likelihood function (4.18) can be rewritten in terms of its generator function as,

$$\begin{aligned} L = &\prod_{i=1}^{m} f_{i1}^{\delta_{i1}} \left[1 - F_{i1}\right]^{1-\delta_{i1}} \\ &\times \left[\prod_{j=2}^{\min(n_i+1,k)} \left\{ - \frac{\ddot{\varphi}_{\alpha_{j-1}^*}\left(\varphi_{\alpha_{j-1}^*}^{-1}[\varphi_{\alpha_{j-1}^*}(F_{i,j-1}) + \varphi_{\alpha_{j-1}^*}(F_{ij})]\right)}{\dot{\varphi}_{\alpha_{j-1}^*}\left(\varphi_{\alpha_{j-1}^*}^{-1}[\varphi_{\alpha_{j-1}^*}(F_{i,j-1}) + \varphi_{\alpha_{j-1}^*}(F_{ij})]\right)^3} \dot{\varphi}_{\alpha_{j-1}^*}(F_{i,j-1}) \dot{\varphi}_{\alpha_{j-1}^*}(F_{ij}) f_{ij} \right\}^{\delta_{ij}} \right. \\ &\left. \times \left\{ 1 - \frac{\dot{\varphi}_{\alpha_{j-1}^*}(F_{i,j-1})}{\dot{\varphi}_{\alpha_{j-1}^*}\left(\varphi_{\alpha_{j-1}^*}^{-1}[\varphi_{\alpha_{j-1}^*}(F_{i,j-1}) + \varphi_{\alpha_{j-1}^*}(F_{ij})]\right)} \right\}^{1-\delta_{ij}} \right]^{\delta_{i1}} , \end{aligned} \tag{4.19}$$

where $\dot{\varphi}_{\alpha^*} = \frac{\partial \varphi_{\alpha^*}(t)}{\partial t}$, $\ddot{\varphi}_{\alpha^*} = \frac{\partial^2 \varphi_{\alpha^*}(t)}{\partial t^2}$, f_{ij} and F_{ij} are $f_j(w_{ij})$ and $F_j(w_{ij})$, respectively.

In case there are external covariates of interest in a study, the probabilistic characteristics of gap times including appropriate functions f_{ij} and F_{ij} are defined conditionally on the value of covariates. As discussed in Section 2.2.3, the model (4.12) can be extended to incorporate external covariates. Assuming there is enough information about covariates available in data for each gap time, the likelihood function (4.19) can be used to make inference about covariates.

4.6 Extensions of the Model for the Heterogeneous Data

The issue of excess heterogeneity in the rate functions of event count models has been discussed in Section 3.2. In this section, we investigate the impacts of such a scenario on the estimation of dynamic covariates with gap time models.

As discussed in Section 3.2, a useful random effects model is given with the following intensity function

$$\lambda_i[t|\mathcal{H}_i(t), \nu_i] = \alpha\nu_i \exp\left[\gamma\, N_i(t^-) + \beta\, Z_i(t)\right], \qquad i = 1, \ldots, m,\ t > 0, \tag{4.20}$$

where $\nu_1, \nu_2, \ldots, \nu_m$ are positive-valued i.i.d. random effects from a gamma distribution with mean 1 and variance ϕ, where $\phi > 0$. External covariates can be included in (4.20).

The model (4.20) can be expressed with the conditional hazard function of W_{ik}, given the value of ν_i as follows. For $i = 1, \ldots, m$ and $w > 0$,

$$h_k(w|\nu_i) = \begin{cases} \alpha\nu_i, & k = 1, \\ \alpha\nu_i \mathrm{e}^{[(k-1)\gamma+\beta]}\, I(w \le \Delta) + \alpha\nu_i \mathrm{e}^{[(k-1)\gamma]}\, I(w > \Delta), & k = 2, 3, \ldots, \end{cases} \tag{4.21}$$

where $\nu_1, \nu_2, \ldots, \nu_m$ are independent and identically distributed unobservable random variables. Then, the conditional likelihood contribution by the i^{th} individual is

$$
\begin{aligned}
L_i(\boldsymbol{\theta}|\nu_i) =& f_1(w_{i1}|\nu_i)^{\delta_{i1}} \left[1 - F_1(w_{i1}|\nu_i)\right]^{1-\delta_{i1}} \\
& \left[\prod_{j=2}^{\min(n_i+1,k)} \left\{ \frac{\partial^2 C_{\alpha^*_{j-1}}(F_j(w_{ij}|\nu_i), F_{j-1}(w_{i,j-1}|\nu_i))}{\partial F_j(w_{ij}|\nu_i)\partial F_{j-1}(w_{i,j-1}|\nu_i)} f_j(w_{ij}|\nu_i) \right\}^{\delta_{ij}} \right. \\
& \left. \times \left\{ 1 - \frac{\partial C_{\alpha^*_{j-1}}(F_j(w_{ij}|\nu_i), F_{j-1}(w_{i,j-1}|\nu_i))}{\partial F_{j-1}(w_{i,j-1}|\nu_i)} \right\}^{1-\delta_{ij}} \right]^{\delta_{i1}} .
\end{aligned}
\tag{4.22}
$$

where $\boldsymbol{\theta} = (\alpha, \beta, \gamma, \phi, \boldsymbol{\alpha^{*\prime}})'$.

From the conditional hazard function (4.21), we can respectively obtain the conditional marginal c.d.f. and p.d.f. of W_{ik} given ν_i as follows. For $i = 1, \ldots, m$, $k = 2, 3, \ldots$ and $w > 0$,

$$
F_k(w|\nu_i) = 1 - \exp\left[-\alpha\nu_i \mathrm{e}^{(k-1)\gamma}\left\{w + I(k > 1)\min(w, \Delta)\left(\mathrm{e}^{\beta} - 1\right)\right\}\right], \quad w > 0,
\tag{4.23}
$$

and

$$
f_k(w|\nu_i) = \begin{cases} \alpha\nu_i \ \exp\left[(k-1)\gamma + \beta I(k>1) - \alpha\nu_i \mathrm{e}^{(k-1)\gamma}\left\{w\mathrm{e}^{\beta I(k>1)}\right\}\right], & w \le \Delta, \\ \alpha\nu_i \ \exp\left[(k-1)\gamma - \alpha\nu_i \mathrm{e}^{(k-1)\gamma}\left\{w + \Delta\left(\mathrm{e}^{\beta I(k>1)} - 1\right)\right\}\right], & w > \Delta. \end{cases}
\tag{4.24}
$$

In model (4.21), we assume that ν_i are i.i.d. gamma random variables with mean 1 and variance ϕ. In this case the p.d.f. of ν_i, $i = 1, \ldots, m$, is given by

$$
g(\nu_i; \phi) = \frac{\nu_i^{\phi^{-1}-1} \exp(-\nu_i/\phi)}{\phi^{\phi^{-1}}\Gamma(\phi^{-1})}, \qquad 0 < \nu_i < \infty.
\tag{4.25}
$$

Estimation in the random effects models can be carried out after integrating out random effects from the conditional likelihood function given the value of random effects. In this case, the marginal likelihood function is

$$
L(\boldsymbol{\theta}) = \prod_{i=1}^{m} L_i(\boldsymbol{\theta}) = \prod_{i=1}^{m} \int_0^\infty L_i(\boldsymbol{\theta}|\nu_i) g(\nu_i; \phi) d\nu_i.
\tag{4.26}
$$

However, the integration in the marginal likelihood function (4.26) is quite messy. Therefore, it requires a software package to calculate numerical values.

We encountered issues with the R software during the optimization and calculation of the maximum likelihood estimates. The integration function returned with non-numerical values. To deal with this issue, we used the Gaussian quadrature method to obtain an approximation of the likelihood function (4.26). A similar application can be seen in Selvaratnam et al. (2017). In our case, Gauss–Laguerre quadrature approximation is appropriate. This procedure is given by the approximation

$$\int_0^\infty f(\nu^*)\exp(-\nu^*)d\nu^* \approx \sum_{l=1}^{r} \omega_l f(\nu_l^*), \tag{4.27}$$

where the nodes ν_l^* are the solutions of the r^{th} order Laguerre polynomial and the ω_l are suitably corresponding weights. The r^{th} order Laguerre polynomial $L_r(x)$ is $\frac{e^x}{r!}\frac{d^r}{dx^r}(e^x x^r)$ and the ω_l is $\frac{\nu_l^*}{(r+1)^2\left[L_{r+1}(\nu_l^*)\right]^2}$. By Theorem 5.1.9 of Brass and Petras (2011), the Gauss–Laguerre approximation converges to the exact integral as $r \to \infty$. In our case, for the i^{th} individual, $f(\nu_i^*)$ in left hand side of (4.27) can be expressed as

$$f(\nu_i^*;\boldsymbol{\theta},\phi) = L_i(\boldsymbol{\theta}|\phi\nu_i^*)\frac{(\nu_i^*)^{\phi^{-1}-1}}{\Gamma(\phi^{-1})}, \tag{4.28}$$

following the substitution of $\phi\nu_i^*$ in place of ν_i in (4.26). Then the resulting approximate likelihood function can be written as

$$L(\boldsymbol{\theta},\phi) \approx \prod_{i=1}^{m}\sum_{l=1}^{r} \omega_l f(\nu_{il}^*;\boldsymbol{\theta},\phi). \tag{4.29}$$

We use `nlminb` method in `optimx` package in R to obtain the estimates of the parameters and their standard errors by maximizing the log of the $L(\boldsymbol{\theta},\phi)$ in (4.29). As discussed in the next section through simulations, this method works fine in our case.

4.7 Simulation Results

In this section, we present the results of simulation studies conducted to investigate the bias and precision in the estimation of parameters in the dynamic gap times models discussed in Section 4.5 and 4.6. We first consider the identical processes case and then the nonidentical case. In each case, to generate data, we assume that the successive gap times have a Markov type of dependency; that is, the k^{th} gap time only depends on the immediately previous $(k-1)^{\text{th}}$ gap time for $k = 2, 3, \ldots$. We formulate such dependency in two ways. First, we assign a common copula parameter for the successive dependency. Second, we assign separate copula parameters for the successive dependency. Since our primary interest is in the investigation of two dynamic features of recurrent event processes, it is important to check which dependency formulation gives better results in terms of estimation of them. While doing this, we also fit the best model with less number of parameters for the dependency to facilitate the estimation procedure.

4.7.1 Identical Processes

In the identical processes case, we generated 2,000 realizations of serially dependent recurrent event processes using the algorithm given in Section 2.5.2 with the model

$$h_k(w) = \begin{cases} \alpha, & k = 1, \\ \alpha \mathrm{e}^{[(k-1)\gamma+\beta]}\, I(w \le \Delta) + \alpha \mathrm{e}^{[(k-1)\gamma]}\, I(w > \Delta), & k = 2, 3, \ldots, \end{cases} \tag{4.30}$$

where $h_k(w)$, $w > 0$, is the hazard function of the k^{th} gap time W_k. We generated the data from the model (4.30) with α, γ ($\alpha = 1$, $\gamma = 0.223$ and $\beta = 1.609$) so that both trend and carryover effects are involved in the data generation process. We fixed Δ at 0.05 and used (m, τ) combinations, where $m = 50, 100, 250, 500$ and $\tau = 1$ and 5. We fitted models for only up to the first five gap times. Since we fixed the follow-up period τ, a process is type 2 censored when the fifth event occurred before τ. Otherwise the process is censored at time τ. As discussed in Section 2.2.5, the likelihood function (4.18) is still valid in such cases. Also, see Cook and Lawless (2007, Section 2.6).

We generated the gap times so that they were serially dependent. The series of dependence parameters is denoted by $\boldsymbol{\phi_c} = (\phi_{c1}, \phi_{c2}, \phi_{c3}, \phi_{c4})$, and the corresponding vector of Kendall's tau values are denoted by $\boldsymbol{\tau_{\phi_c}} = (\tau_{\phi_{c1}}, \tau_{\phi_{c2}}, \tau_{\phi_{c3}}, \tau_{\phi_{c4}})$. The first value $\tau_{\phi_{c1}}$ in $\boldsymbol{\tau_{\phi_c}}$ corresponds to the dependency between the first gap time W_1 and the second gap time W_2, the second value $\tau_{\phi_{c2}}$ corresponds to the dependency between the second gap time W_2 and the third gap time W_3 and so on. We picked three sets of values for $\boldsymbol{\tau_{\phi_c}}$ so that the dependency was increasing, decreasing and constant. In this way, we checked whether the trend effect and dependency parameters confounded each other. For the increasing serial dependence parameters, we assigned $\boldsymbol{\phi_c} = (1.33, 2, 3, 4.67)$, and the corresponding Kendall's tau values were $\boldsymbol{\tau_{\phi_c}} = (0.4, 0.5, 0.6, 0.7)$. For the decreasing and constant serial dependence parameters, respectively, we assigned $\boldsymbol{\phi_c} = (4.67, 3, 2, 1.33)$ and $\boldsymbol{\phi_c} = (2, 2, 2, 2)$.

Based on the model (4.30), we fitted the following hybrid gap time model with Clayton copula dependency. The hybrid model:

$$h_k(w_i) = \begin{cases} \alpha, & k = 1, \\ \alpha \mathrm{e}^{[(k-1)\gamma+\beta]}\, I(w_i \leq \Delta) + \alpha \mathrm{e}^{[(k-1)\gamma]}\, I(w_i > \Delta), & k = 2, 3, \ldots, \end{cases} \tag{4.31}$$

for $i = 1, \ldots, m$ and $w_i > 0$.

We used the following abbreviations for two different parameterizations of dependence structure. The acronym DDP stands for difference dependence parameters, which defines the Markov type dependent structured model with different parameters for the dependence series, and SDP stands for single dependence parameter, which defines the same dependent structured model with only a single parameter for the dependence series. To compare DDP and SDP models with the classical approach, we also fitted the model with independent gap times assumption and denoted by IND. The estimates of the parameters in the models SDP and DDP estimates were obtained by maximizing the log of likelihood functions in (4.17) and (4.18), respectively. Estimates in the IND model were obtained by maximizing the log of the likelihood function in (4.17), where the copula function C_{α^*} is replaced with the independent copula function C_I. The independent copula has the form

$$C_I(u_1, u_2) = u_1 u_2, \;\; 0 \leq u_1, u_2 \leq 1, \tag{4.32}$$

where the subscript I denotes the independent. We used the `optimx` package in `R` to obtain the estimates of the parameters by maximizing the log-likelihood function.

Table 4.7: Empirical means (standard deviations) [means of standard errors] of parameter estimates from the hybrid model (4.31) under three different structures (DDP, SDP and IND). The data were generated from the model (4.30), where $\alpha = 1$, $\gamma = 0.223$, $\beta = 1.609$, $\Delta = 0.05$ and serial dependence copula parameters were $\boldsymbol{\phi_c} = (1.33, 2, 3, 4.67)$ with 2,000 simulations.

τ	Parameter	DDP				SDP				IND			
		m = 50	**m = 100**	**m = 250**	**m = 500**	**m = 50**	**m = 100**	**m = 250**	**m = 500**	**m = 50**	**m = 100**	**m = 250**	**m = 500**
5	$\bar{\hat{\alpha}}$	1.016398 (0.118120) [0.114362]	1.008957 (0.079490) [0.080214]	1.004273 (0.050192) [0.050496]	1.001837 (0.035139) [0.035620]	1.051388 (0.127742) [0.127373]	1.042954 (0.086036) [0.089407]	1.037665 (0.054443) [0.056300]	1.035646 (0.037859) [0.039740]	1.001024 (0.127374) [0.112408]	0.993567 (0.085629) [0.078820]	0.989303 (0.054127) [0.049625]	0.987159 (0.038137) [0.035015]
	$\bar{\hat{\gamma}}$	0.220168 (0.044311) [0.042722]	0.220924 (0.030855) [0.030161]	0.222019 (0.019559) [0.019073]	0.222417 (0.013861) [0.013480]	0.208616 (0.047315) [0.047725]	0.209088 (0.033295) [0.033693]	0.210137 (0.021007) [0.021311]	0.210196 (0.014718) [0.015066]	0.247133 (0.056182) [0.050596]	0.247175 (0.039443) [0.035687]	0.247853 (0.024831) [0.022548]	0.247750 (0.017159) [0.015944]
	$\bar{\hat{\beta}}$	1.597803 (0.158339) [0.154357]	1.600798 (0.108950) [0.108480]	1.598042 (0.068291) [0.068405]	1.600188 (0.047806) [0.048286]	1.553954 (0.172596) [0.142762]	1.557927 (0.119695) [0.100478]	1.556069 (0.076567) [0.063467]	1.558709 (0.053288) [0.044812]	1.598985 (0.219142) [0.155089]	1.608732 (0.155818) [0.109148]	1.612159 (0.100749) [0.068861]	1.617607 (0.070604) [0.048626]
	$\bar{\hat{\phi}}_{c1}$	1.384899 (0.341383) [0.332050]	1.360880 (0.233049) [0.230464]	1.346778 (0.146115) [0.144137]	1.340187 (0.100029) [0.101478]								
	$\bar{\hat{\phi}}_{c2}$	2.057964 (0.444063) [0.430645]	2.036171 (0.306353) [0.300794]	2.014858 (0.193676) [0.188203]	2.007926 (0.134532) [0.132662]								
	$\bar{\hat{\phi}}_{c3}$	3.090353 (0.586488) [0.573210]	3.040231 (0.402132) [0.398695]	3.012476 (0.248510) [0.249726]	3.006001 (0.176787) [0.176084]								
	$\bar{\hat{\phi}}_{c4}$	4.793780 (0.832738) [0.820700]	4.727589 (0.561755) [0.570011]	4.687248 (0.352407) [0.356476]	4.675602 (0.242824) [0.251102]								
	$\bar{\hat{\phi}}_{c}$					2.331883 (0.288948) [0.255345]	2.318773 (0.198903) [0.178882]	2.307897 (0.125509) [0.112343]	2.304723 (0.084936) [0.079224]				
1	$\bar{\hat{\alpha}}$	1.009573 (0.175283) [0.173159]	1.003833 (0.121710) [0.122218]	1.000659 (0.076973) [0.077197]	1.001855 (0.054138) [0.054641]	1.017354 (0.178636) [0.176725]	1.010616 (0.124461) [0.124530]	1.006742 (0.078921) [0.078616]	1.007758 (0.055057) [0.055637]	1.042573 (0.172448) [0.159731]	1.039442 (0.119892) [0.112658]	1.036838 (0.075226) [0.071102]	1.038456 (0.053251) [0.050316]
	$\bar{\hat{\gamma}}$	0.214991 (0.085182) [0.075572]	0.217162 (0.055444) [0.053432]	0.220955 (0.033951) [0.033744]	0.222402 (0.024128) [0.023837]	0.232543 (0.091928) [0.100685]	0.234741 (0.062437) [0.070278]	0.238355 (0.038937) [0.044220]	0.240144 (0.027504) [0.031204]	0.473643 (0.092449) [0.093307]	0.466745 (0.063506) [0.064971]	0.464624 (0.039793) [0.040751]	0.464186 (0.028110) [0.028704]
	$\bar{\hat{\beta}}$	1.597507 (0.225180) [0.214434]	1.600452 (0.156243) [0.149863]	1.600043 (0.095840) [0.094324]	1.598884 (0.066891) [0.066589]	1.546709 (0.224176) [0.196542]	1.552694 (0.156449) [0.137361]	1.551837 (0.097076) [0.086501]	1.550830 (0.067765) [0.061069]	1.780111 (0.329083) [0.271329]	1.802601 (0.231594) [0.188810]	1.812257 (0.142700) [0.118352]	1.811690 (0.098756) [0.083385]
	$\bar{\hat{\phi}}_{c1}$	1.407304 (0.408090) [0.387812]	1.373672 (0.275842) [0.267126]	1.345869 (0.168504) [0.165813]	1.334240 (0.118980) [0.116405]								
	$\bar{\hat{\phi}}_{c2}$	2.092604 (0.574602) [0.551762]	2.048931 (0.378923) [0.378281]	2.017373 (0.230980) [0.234796]	2.010217 (0.164282) [0.165165]								
	$\bar{\hat{\phi}}_{c3}$	3.218671 (0.913893) [0.817548]	3.089031 (0.572300) [0.547861]	3.037123 (0.338770) [0.338509]	3.024201 (0.240197) [0.237694]								
	$\bar{\hat{\phi}}_{c4}$	5.085861 (1.472392) [1.276935]	4.860941 (0.881007) [0.846859]	4.755834 (0.540230) [0.518953]	4.718583 (0.368912) [0.362481]								
	$\bar{\hat{\phi}}_{c}$					2.177989 (0.340550) [0.303648]	2.161402 (0.233430) [0.210722]	2.145689 (0.144598) [0.131663]	2.139835 (0.102050) [0.092735]				

Table 4.7 contains the results of the simulation study when the hybrid models without external covariates fitted to the data generated from the model (4.30). When the follow-up time τ is 5, the estimates in IND and SDP are slightly biased, but within 95% c.i. based on standard normal approximation. The estimate of the copula parameter in SDP is close to the value of the Clayton copula parameter corresponding to the average of the sequence of Kendall's tau values. The average of the sequence of Kendall's tau values is 0.55 and the corresponding Clayton copula parameter is

2.444. Even though the sequence of gap times has strong dependency within each other, the model with independent assumption (IND) gives slightly biased estimates for the parameters when the follow-up period τ is 5. We also notice that the variance estimates are larger than the corresponding DDP and SDP models. However, when the follow-up time reduces from 5 to 1, the estimates from the IND model are drastically biased. In particular, the model overestimates the effects of the internal covariates. It should be noted that, when the follow-up time is 5, on the average the observed percentage of the fifth gap times is 84.1%. When the follow-up period reduces to 1, the observed percentages of the first, second and up to fifth gap times are 63.2%, 42.5%, 33.8%, 29.6% and 27.5%, respectively. Follow-up time reduction does not affect much on the estimates of the parameters in the SDP model. Overall, the results in Table 4.7 indicate censoring rate affects the bias in the estimates of the parameters when the fitted model assumes the gap times are independent. In both follow-up time cases, DDP performs better and is not affected too much by the increased censoring.

We next present the results of a simulation study conducted to investigate the effects of time fixed external covariates on the estimates of the dynamic covariates. For this purpose, we include two time fixed covariates $\boldsymbol{x} = (x_1, x_2)'$ into the model (4.30) and define the hazard function of the k^{th} gap time W_k as

$$h_k(w) = \begin{cases} \alpha e^{\boldsymbol{\xi}'\boldsymbol{x}}, & k = 1, \\ \alpha e^{[(k-1)\gamma+\beta+\boldsymbol{\xi}'\boldsymbol{x}]}\, I(w \leq \Delta) + \alpha e^{[(k-1)\gamma+\boldsymbol{\xi}'\boldsymbol{x}]}\, I(w > \Delta), & k = 2, 3, \ldots, \end{cases} \tag{4.33}$$

where $\boldsymbol{\xi}' = (\xi_1, \xi_2)$ is a vector of parameters. Following the steps of the previous four simulation scenarios explained above, we included two binary external covariates with 1 and 0 as possible values, and their corresponding effect values are $\boldsymbol{\xi} = (0.5, 1)$.

Table 4.8 includes the simulation results when the hybrid model (4.33) fitted to the data generated with external covariate effects. The IND model underestimates the effects of external covariates, while it overestimates the effects of internal covariates when the follow-up time is small. For example, when the hybrid IND model fitted to the data generated with $m = 500$, $\tau_K = (0.4, 0.5, 0.6, 0.7)$, $\alpha = 1$, $\beta = 1.609$, $\gamma = 0.223$, $\xi_1 = 0.5$, $\xi_2 = 1$ and $\tau = 1$, the empirical means of estimates of ξ_1 and ξ_2 are 0.3535 and 0.6523, respectively, and the empirical standard deviations of the estimates are 0.0721 and 0.0751, respectively. However, when the follow-up time is large, the estimates in the IND model are not significantly different from the actual values. For example, when the hybrid IND model fitted to the data generated with $m = 500$, $\tau_K = (0.4, 0.5, 0.6, 0.7)$, $\alpha = 1$, $\beta = 1.609$, $\gamma = 0.223$, $\xi_1 = 0.5$, $\xi_2 = 1$ and $\tau = 5$, the empirical means of estimates of ξ_1 and ξ_2 are 0.4836 and 0.9809, respectively, and the empirical standard deviations of the estimates are 0.0625 and 0.0657, respectively. It should be noted that, when $\tau = 5$, on the average 95.7% of the fifth gap times were uncensored in this setting. When $\tau = 1$, this average was 84.1% for the first gap time, 71.7% for the second gap time, 64.9% for the third gap time, 60.9% for the fourth gap time and 58.6% for the fifth gap time. We also note that the estimates of the DDP model given in Table 4.7 and Table 4.8 are consistent. In other words, when the sample size increases the estimates converges to the correct parameter values, as well as the simulation standard deviations and the simulation means of standard errors decrease.

Table 4.8: Empirical means (standard deviations) [means of standard errors] of parameter estimates from the hybrid model (4.33) under three different structures (DDP, SDP and IND). The data were generated from the model (4.33), where $\alpha = 1$, $\gamma = 0.223$, $\beta = 1.609$, $\Delta = 0.05$, $\boldsymbol{\xi} = (0.5, 1)$ and serial dependence copula parameters were $\boldsymbol{\phi_c} = (1.33, 2, 3, 4.67)$ with 2,000 simulations.

τ	Parameter	DDP				SDP				IND			
		m = 50	m = 100	m = 250	m = 500	m = 50	m = 100	m = 250	m = 500	m = 50	m = 100	m = 250	m = 500
5	$\bar{\hat{\alpha}}$	1.033482 (0.184445) [0.175350]	1.012987 (0.124584) [0.120168]	1.004552 (0.076203) [0.075030]	1.002741 (0.053531) [0.052797]	1.079978 (0.194081) [0.190689]	1.057678 (0.131575) [0.130620]	1.048814 (0.080980) [0.081589]	1.047102 (0.057293) [0.057445]	1.043512 (0.210484) [0.151397]	1.024211 (0.139445) [0.104087]	1.019325 (0.086436) [0.065258]	1.018036 (0.061191) [0.045987]
	$\bar{\hat{\gamma}}$	0.223138 (0.045928) [0.041276]	0.222892 (0.029436) [0.029150]	0.223091 (0.018347) [0.018395]	0.222917 (0.012931) [0.012991]	0.202316 (0.044899) [0.046783]	0.202316 (0.030998) [0.032958]	0.202191 (0.019577) [0.020784]	0.201775 (0.013834) [0.014679]	0.235635 (0.053902) [0.051847]	0.232724 (0.038581) [0.036503]	0.230915 (0.024591) [0.023017]	0.229812 (0.017304) [0.016259]
	$\bar{\hat{\beta}}$	1.588864 (0.138350) [0.136579]	1.594845 (0.096715) [0.096288]	1.596294 (0.060591) [0.060687]	1.598787 (0.042878) [0.042884]	1.581618 (0.145550) [0.124917]	1.587449 (0.102396) [0.087988]	1.589511 (0.064444) [0.055506]	1.592302 (0.045555) [0.039214]	1.560753 (0.189632) [0.153178]	1.581001 (0.131957) [0.107712]	1.591471 (0.085594) [0.067856]	1.598136 (0.060202) [0.047925]
	$\bar{\hat{\xi}}_1$	0.508108 (0.178280) [0.175740]	0.505828 (0.121675) [0.121792]	0.502859 (0.075277) [0.076099]	0.501269 (0.052027) [0.053597]	0.505637 (0.180884) [0.177635]	0.503809 (0.123326) [0.123088]	0.500764 (0.076172) [0.076903]	0.499119 (0.053128) [0.054173]	0.489835 (0.210572) [0.133003]	0.489067 (0.142232) [0.092638]	0.485242 (0.088279) [0.058061]	0.483639 (0.062511) [0.040934]
	$\bar{\hat{\xi}}_2$	1.007976 (0.177128) [0.177556]	1.005128 (0.125343) [0.122961]	1.002442 (0.075739) [0.076819]	1.000950 (0.054524) [0.054104]	1.004700 (0.180139) [0.179476]	1.002307 (0.127605) [0.124287]	1.000248 (0.077432) [0.077660]	0.998918 (0.055598) [0.054705]	0.992851 (0.216038) [0.137477]	0.988669 (0.148006) [0.095763]	0.982456 (0.092519) [0.060038]	0.980946 (0.065691) [0.042338]
	$\bar{\hat{\phi}}_{c1}$	1.384459 (0.341769) [0.329760]	1.357495 (0.230727) [0.228720]	1.346019 (0.142516) [0.143046]	1.339904 (0.099597) [0.100681]								
	$\bar{\hat{\phi}}_{c2}$	2.043043 (0.427591) [0.414308]	2.029948 (0.295502) [0.289463]	2.009272 (0.184440) [0.181093]	2.004073 (0.130121) [0.127580]								
	$\bar{\hat{\phi}}_{c3}$	3.046319 (0.545054) [0.540406]	3.015407 (0.377134) [0.376408]	2.998815 (0.236244) [0.236160]	2.997469 (0.169317) [0.166677]								
	$\bar{\hat{\phi}}_{c4}$	4.818300 (0.811274) [0.773522]	4.722108 (0.541259) [0.533473]	4.680673 (0.333593) [0.333424]	4.669938 (0.233084) [0.234919]								
	$\bar{\hat{\phi}}_{c}$					2.372334 (0.265585) [0.240862]	2.361016 (0.178726) [0.168155]	2.353287 (0.113521) [0.105521]	2.351222 (0.079381) [0.074402]				
1	$\bar{\hat{\alpha}}$	1.013568 (0.274901) [0.260087]	1.008940 (0.185530) [0.182205]	1.005213 (0.114802) [0.114602]	1.003941 (0.079122) [0.080837]	1.038670 (0.289720) [0.271175]	1.032371 (0.194752) [0.189827]	1.027728 (0.120334) [0.119282]	1.026008 (0.082989) [0.084116]	1.324479 (0.318916) [0.236473]	1.324291 (0.211936) [0.165763]	1.322669 (0.133504) [0.104217]	1.321626 (0.091849) [0.073501]
	$\bar{\hat{\gamma}}$	0.220117 (0.069951) [0.052578]	0.220183 (0.040212) [0.037045]	0.221553 (0.023758) [0.023400]	0.222117 (0.016458) [0.016538]	0.206816 (0.062193) [0.062693]	0.208968 (0.041735) [0.043996]	0.210797 (0.025459) [0.027716]	0.211331 (0.017880) [0.019576]	0.302136 (0.070545) [0.069914]	0.299822 (0.047918) [0.048874]	0.298572 (0.029992) [0.030742]	0.298087 (0.021146) [0.021696]
	$\bar{\hat{\beta}}$	1.595266 (0.167498) [0.161545]	1.598562 (0.116975) [0.114055]	1.600247 (0.072101) [0.071968]	1.600537 (0.049175) [0.050877]	1.592532 (0.168032) [0.145880]	1.593768 (0.119828) [0.102782]	1.594751 (0.074174) [0.064825]	1.595090 (0.051334) [0.045826]	1.668462 (0.225918) [0.205402]	1.685422 (0.161489) [0.143653]	1.698019 (0.098970) [0.090408]	1.700927 (0.069781) [0.063816]
	$\bar{\hat{\xi}}_1$	0.526527 (0.255733) [0.239452]	0.515386 (0.169251) [0.166013]	0.505815 (0.103943) [0.103844]	0.501451 (0.073314) [0.073141]	0.523452 (0.259299) [0.243947]	0.512168 (0.171202) [0.169113]	0.501914 (0.105600) [0.105730]	0.498185 (0.074734) [0.074487]	0.376593 (0.247087) [0.161897]	0.365035 (0.163806) [0.112280]	0.356000 (0.101894) [0.070261]	0.353522 (0.072062) [0.049499]
	$\bar{\hat{\xi}}_2$	1.031310 (0.272013) [0.263312]	1.009756 (0.184692) [0.183114]	1.003025 (0.113979) [0.114815]	1.000636 (0.080743) [0.081008]	1.031810 (0.280124) [0.269874]	1.008522 (0.190155) [0.188029]	1.001799 (0.117847) [0.117953]	0.998969 (0.083945) [0.083240]	0.687758 (0.259544) [0.172347]	0.662842 (0.172502) [0.119189]	0.654994 (0.108612) [0.074508]	0.652252 (0.075090) [0.052503]
	$\bar{\hat{\phi}}_{c1}$	1.389403 (0.363924) [0.351716]	1.352328 (0.242188) [0.242258]	1.343838 (0.150466) [0.151469]	1.337987 (0.106302) [0.106621]								
	$\bar{\hat{\phi}}_{c2}$	2.056228 (0.472847) [0.459045]	2.030771 (0.323770) [0.318714]	2.012084 (0.198097) [0.199129]	2.005054 (0.140487) [0.140285]								
	$\bar{\hat{\phi}}_{c3}$	3.088357 (0.651132) [0.625616]	3.047526 (0.428378) [0.432833]	3.022438 (0.269883) [0.270399]	3.015654 (0.192775) [0.190565]								
	$\bar{\hat{\phi}}_{c4}$	4.890080 (1.003051) [0.925029]	4.774926 (0.643403) [0.632410]	4.704582 (0.385173) [0.392329]	4.681975 (0.272675) [0.275761]								
	$\bar{\hat{\phi}}_{c}$					2.298837 (0.277854) [0.256835]	2.282561 (0.194048) [0.178901]	2.276733 (0.120077) [0.112166]	2.273001 (0.085586) [0.079076]				

Overall, the simulation results show that, when the data includes trend but not carryover effects and if the trend component is ignored, there might be wrong conclusions about the presence of carryover effects as a result of the fact that the effects of trend and carryover effects may confound in such cases. The same phenomenon is observed in count-based models discussed in Chapter 3. Consequently, the hybrid model is recommended to overcome this issue. Based on the results given in Tables 4.7 and 4.8, we do not recommend to fit a model based on independent gap times when the assumption of independence of the serial gap times is in question. We also do not recommend the use of independent gap time models when the data has too many censored gap times. Hybrid DDP model is the ideal approach to make valid inferences about dynamic features because it can handle the aforementioned issues related to confounding and censoring. In the next subsection, we investigate the similar scenarios with nonidentical processes.

We would like to note that we also used three more versions of the model (4.30) to generate data. These models were the null model with only α, the carryover effects model with α and γ. We then fitted each generated data, including the data generated from the model (4.30) with α, β and γ, with four models; which include a null model with the hazard function $h_k(w_i) = \alpha, k = 1, 2, \ldots,$, the trend only model with the hazard function $h_k(w_i) = \alpha e^{[(k-1)\gamma]}, k = 1, 2, \ldots$, the carryover effects model $h_k(w_i) = I(k = 1)\alpha + I(k > 1)\left[\alpha e^{\beta} I(w_i \leq \Delta) + \alpha I(w_i > \Delta)\right]$, $k = 1, 2, \ldots$, as well as the hybrid model (4.31). Since the hybrid model performed better in all scenarios, and presenting the results require extensive number of tables, we only presented the results here when the hybrid model was used to fit the data generated from model (4.30). For other cases, we obtained similar results to those obtained with models for event counts discussed in Chapter 3. That is, the parameter estimates in null, carryover effects and trend models were biased when the model was misspecified. Also, we noticed that the trend effect and the dependency parameters do not confound each other. That is, serially increasing or decreasing the dependency among gap times within a process does not cause any problem in estimating the trend effect. Therefore, we present only the results where the hybrid model fitted to the data generated with increasing dependency for brevity.

4.7.2 Nonidentical Processes

For the nonidentical processes case, we generated 2,000 realizations of serially dependent recurrent event processes. We used the extended version of the algorithm presented in Section 2.5.2 to generate heterogeneity across individual processes. In the data generation process, we used the model

$$h_k(w|\nu_i) = \begin{cases} \alpha\nu_i, & k = 1, \\ \alpha\nu_i e^{[(k-1)\gamma+\beta]}\, I(w \leq \Delta) + \alpha\nu_i e^{[(k-1)\gamma]}\, I(w > \Delta), & k = 2, 3, \ldots, \end{cases} \tag{4.34}$$

where $h_k(w|\nu_i)$ is the conditional hazard function for the k^{th} gap time W_k given the value of the random effect ν_i. We generated the ν_i from $Gamma(1, \phi = 0.3), i = 1, \ldots, m$, distribution. For the data generation process, we considered the same scenarios given in Section 4.7.1. For brevity, we present the results in this section only for the model (4.34), where $\alpha = 1$, $\gamma = 0.223$, $\beta = 1.609$, and $\Delta = 0.05$. The value of $Var(\nu_i) = \phi$ fixed at 0.3, which generates a moderate level of heterogeneity in the data. Such amount of heterogeneity is frequently seen in epidemiology studies. We denote DDP_RE, SDP_RE and IND_RE as extended versions of DDP, SDP and IND with random effects, respectively. DDP_RE, SDP_RE and IND_RE models include heterogeneity parameter ϕ in addition to the parameters in DDP, SDP and IND models, respectively. Parameter estimates for DDP_RE and SDP_RE models are obtained by maximizing the log of the likelihood function in (4.29).

Table 4.9 contains the results of our simulation study when the hybrid model (4.34) fitted under three different dependence structures to the same data generated by using the model (4.34). As explained in the previous section, we used the vector of serial dependent copula parameters $\boldsymbol{\phi_c} = (1.33, 2, 3, 4.67)$ during data generation process. The results from DDP, SDP and IND models, which ignore the heterogeneity, are provided in the Appendix E (Table E.1). Due to heavy censoring, we fixed the value of τ at 2 instead of 1 for the scenarios with short follow-up time. On the average, the percentage of the uncensored fifth gap times is 73.2% when the follow-up period τ was 5, while the averages of that from the first to the fifth gap times were 79.2%, 61.9%, 52.2%, 46.7% and 43.5%, respectively, when $\tau = 2$.

Table 4.9: Empirical means (standard deviations) [means of standard errors] of parameter estimates from the hybrid model (4.34) under six different structures (DDP_RE, SDP_RE, IND_RE, DDP, SDP and IND). The data were generated from the model (4.34), where $\alpha = 1$, $\phi = 0.3$, $\gamma = 0.223$, $\beta = 1.609$, $\Delta = 0.05$ and serial dependence copula parameters were $\boldsymbol{\phi_c} = (1.33, 2, 3, 4.67)$ with 2,000 simulations.

τ	**Parameter**	**DDP_RE**			**SDP_RE**			**IND_RE**		
		m = 50	**m = 100**	**m = 250**	**m = 50**	**m = 100**	**m = 250**	**m = 50**	**m = 100**	**m = 250**
5	$\bar{\hat{\alpha}}$	1.016196 (0.211288) [0.197529]	1.006865 (0.151369) [0.134752]	1.003352 (0.101765) [0.076967]	1.035117 (0.233214) [0.210522]	1.023924 (0.164702) [0.144745]	1.021792 (0.118068) [0.081501]	2.147346 (0.796276) [0.415447]	2.109496 (0.615468) [0.298755]	2.041331 (0.453428) [0.187050]
	$\bar{\hat{\phi}}$	0.289380 (0.183554) [0.175930]	0.288900 (0.131766) [0.126017]	0.297002 (0.089583) [0.078908]	0.286234 (0.195182) [0.176158]	0.286143 (0.140300) [0.129328]	0.296764 (0.100351) [0.079584]	1.123655 (0.296526) [0.198531]	1.133023 (0.255090) [0.146108]	1.118198 (0.208501) [0.094734]
	$\bar{\hat{\gamma}}$	0.226499 (0.060008) [0.048729]	0.227365 (0.048622) [0.034254]	0.223719 (0.032579) [0.021390]	0.223902 (0.067345) [0.056884]	0.225145 (0.054487) [0.040033]	0.221237 (0.043659) [0.024966]	0.338703 (0.068122) [0.060847]	0.335241 (0.052672) [0.043182]	0.332029 (0.036354) [0.027329]
	$\bar{\hat{\beta}}$	1.596184 (0.180504) [0.166973]	1.597120 (0.131367) [0.116683]	1.603849 (0.086322) [0.073272]	1.548465 (0.196571) [0.154615]	1.552174 (0.141042) [0.108729]	1.563223 (0.098939) [0.068620]	1.004439 (0.214203) [0.186392]	1.015720 (0.161153) [0.131447]	1.035354 (0.110708) [0.082817]
	$\bar{\hat{\phi}}_{c1}$	1.396087 (0.364335) [0.346372]	1.379472 (0.260167) [0.240288]	1.366681 (0.165274) [0.150460]						
	$\bar{\hat{\phi}}_{c2}$	2.077308 (0.476862) [0.455414]	2.047919 (0.334494) [0.315989]	2.027988 (0.218193) [0.198118]						
	$\bar{\hat{\phi}}_{c3}$	3.116177 (0.664491) [0.616010]	3.063067 (0.464517) [0.426120]	3.044875 (0.298078) [0.266773]						
	$\bar{\hat{\phi}}_{c4}$	4.837381 (0.940617) [0.893513]	4.764811 (0.661821) [0.616494]	4.739905 (0.451328) [0.386306]						
	$\bar{\hat{\phi}}_{c}$				2.275622 (0.348199) [0.266008]	2.244922 (0.299015) [0.185634]	2.190441 (0.303837) [0.116167]			
2	$\bar{\hat{\alpha}}$	1.045828 (0.256094) [0.232232]	1.025911 (0.179759) [0.161871]	1.010925 (0.115526) [0.103093]	1.045722 (0.257162) [0.235410]	1.024650 (0.187091) [0.169190]	1.006087 (0.122215) [0.107637]	1.995725 (0.706824) [0.459655]	1.944900 (0.433764) [0.319486]	1.931458 (0.251850) [0.200326]
	$\bar{\hat{\phi}}$	0.326706 (0.291358) [0.261840]	0.313417 (0.237289) [0.208278]	0.305562 (0.158283) [0.144476]	0.310966 (0.296248) [0.251916]	0.293223 (0.243066) [0.206518]	0.284726 (0.168500) [0.147766]	1.185120 (0.235305) [0.256397]	1.201628 (0.154450) [0.180535]	1.218666 (0.085444) [0.113067]
	$\bar{\hat{\gamma}}$	0.230768 (0.092195) [0.064383]	0.229122 (0.065949) [0.046122]	0.227789 (0.047071) [0.029412]	0.243426 (0.105087) [0.081534]	0.241703 (0.078690) [0.059073]	0.241822 (0.056463) [0.037928]	0.351379 (0.083533) [0.082368]	0.348667 (0.058405) [0.057793]	0.347796 (0.036929) [0.036460]
	$\bar{\hat{\beta}}$	1.589455 (0.211752) [0.187930]	1.588256 (0.150681) [0.132669]	1.592205 (0.096632) [0.083742]	1.553707 (0.227059) [0.173776]	1.558300 (0.158980) [0.123071]	1.560068 (0.106723) [0.077190]	1.155654 (0.251471) [0.233072]	1.155504 (0.172864) [0.163092]	1.159704 (0.106817) [0.102614]
	$\bar{\hat{\phi}}_{c1}$	1.386633 (0.397376) [0.376233]	1.347352 (0.271715) [0.255931]	1.327292 (0.179115) [0.159375]						
	$\bar{\hat{\phi}}_{c2}$	2.090678 (0.540584) [0.511321]	2.035831 (0.365441) [0.351799]	2.013329 (0.226207) [0.219529]						
	$\bar{\hat{\phi}}_{c3}$	3.146723 (0.799983) [0.714914]	3.055447 (0.515133) [0.488998]	3.025653 (0.310251) [0.304903]						
	$\bar{\hat{\phi}}_{c4}$	4.990405 (1.244310) [1.089966]	4.815447 (0.827772) [0.735825]	4.742978 (0.519966) [0.456914]						
	$\bar{\hat{\phi}}_{c}$				2.206292 (0.403840) [0.293439]	2.177059 (0.335884) [0.205985]	2.162604 (0.287609) [0.127940]			

The results in Table 4.9 show that the empirical estimates of β from SDP_RE models are slightly biased, but the corresponding 95% normal approximation based confidence intervals cover the true value of the parameter. For example, when the hybrid SDP_RE model fitted to the data generated with $m = 250$, $\alpha = 1$, $\phi = 0.3$, $\beta = 1.609$, $\Delta = 0.05$, $\gamma = 0.223$ and $\tau_K = (0.4, 0.5, 0.6, 0.7)$, the empirical means of estimates of β (= 1.5632 and 1.5601 when $\tau = 5$ and 2) are slightly biased. However, assuming the normality properties hold for the estimates, the corresponding 95% empirical confidence intervals ($(1.369, 1.757)$ and $(1.351, 1.769)$ when $\tau = 5$ and 2) cover the true value of the parameter β. We also notice that the empirical means of the trend effect estimates from SDP_RE models are slightly biased when the follow-up period is short; however, the corresponding 95% confidence intervals cover the true value of the parameter. For example, when the hybrid SDP_RE model fitted to the data generated with $m = 250$, $\alpha = 1$, $\phi = 0.3$, $\beta = 1.609$, $\Delta = 0.05$, $\gamma = 0.223$ and $\tau_K = (0.4, 0.5, 0.6, 0.7)$, the empirical means of estimates of γ (= 0.221 and 0.242 when $\tau = 5$ and 2) are slightly biased when the follow-up period is short. However, using the normality assumption, the constructed 95% empirical confidence intervals ($(0.136, 0.307)$ and $(0.131, 0.352)$ when $\tau = 5$ and 2) cover the true value of the parameter γ.

Overall, the model (4.34) with independent gap time assumption overestimates the baseline constant intensity function α, heterogeneity parameter and trend effect due to the number of previous events, as well as underestimates the carryover effect. For example, when the hybrid IND_RE model fitted to the data generated with $m = 250$, $\alpha = 1$, $\phi = 0.3$, $\beta = 1.609$, $\Delta = 0.05$, $\gamma = 0.223$, $\tau_K = (0.4, 0.5, 0.6, 0.7)$ and $\tau = 5$, the empirical means of estimates of α, ϕ, β and γ are $2.041, 1.008, 0.332$ and 1.035, respectively, and the empirical standard deviations of the estimates are $0.453, 0.209, 0.036$ and 0.111, respectively. The corresponding empirical biases of estimates of α, ϕ, β and γ are $1.041, 0.708, 0.109$ and 0.565, respectively, and the corresponding 95% normal approximation based confidence intervals do not include the true values of the parameters.

The results presented in Table E.1 in Appendix E reveal that the models (DDP, SDP and IND), which do not consider the heterogeneity among individuals, underestimate the baseline constant intensity function α, and overestimate the trend effect γ. For example, when the hybrid DDP model fitted to the data generated with $m = 250$, $\alpha = 1$, $\phi = 0.3$, $\beta = 1.609$, $\Delta = 0.05$, $\gamma = 0.223$, $\tau_K = (0.4, 0.5, 0.6, 0.7)$ and

$\tau = 5$, the empirical mean of the estimate of α is 0.7566 with the empirical standard deviation 0.0477 and the empirical mean of the estimates of γ is 0.2527 with the empirical standard deviation 0.0221. The models DDP and IND slightly overestimate the carryover effect, while SDP slightly underestimates it. For example, when the hybrid DDP, SDP and IND models fitted to the data generated with $m = 250$, $\alpha = 1$, $\phi = 0.3$, $\beta = 1.609$, $\Delta = 0.05$, $\gamma = 0.223$, $\tau_K = (0.4, 0.5, 0.6, 0.7)$ and $\tau = 5$, the carryover effect β is slightly overestimated ($= 1.6362$ and 1.7597) for DDP and IND models, but it is slightly underestimated ($= 1.5631$) under the SDP model. To sum up, the results obtained by fitting the model IND show that the assumption of independent gap times in the same recurrent event process or ignoring the heterogeneity across the processes may lead to the wrong conclusions about the dynamic features of recurrent event processes in the scenarios considered in Table E.1.

Our next simulation study was conducted to investigate the bias and precision of the estimates of dynamic covariates when models include external covariates. For this purpose, we define the conditional hazard function of the k^{th} gap time W_k given the random effect ν_i as

$$h_k(w|\nu_i) = \begin{cases} \alpha\nu_i \mathrm{e}^{\boldsymbol{\xi}'\boldsymbol{x_i}}, & k = 1, \\ \alpha\nu_i \mathrm{e}^{[(k-1)\gamma+\beta+\boldsymbol{\xi}'\boldsymbol{x_i}]}\, I(w \leq \Delta) + \alpha\nu_i \mathrm{e}^{[(k-1)\gamma+\boldsymbol{\xi}'\boldsymbol{x_i}]}\, I(w > \Delta), & k = 2, 3, \ldots, \end{cases} \tag{4.35}$$

for $i = 1, \ldots, m$ and $w > 0$. Following the settings in the identical process case, we include two binary external covariates with 0 and 1 as possible values. The values of the covariates $\boldsymbol{x_i} = (x_{i1}, x_{i2})'$ are generated from a Bernoulli distribution with a success probability 0.5. Their corresponding effect values are $\boldsymbol{\xi} = (0.5, 1)$. Table 4.10 includes the simulation results when the hybrid models with heterogeneity parameter and external covariates fitted to the data generated with model (4.35) and dependence parameter values $\boldsymbol{\phi_c} = (1.33, 2, 3, 4.67)$. The results from DDP, SDP and IND models are provided in the Appendix E (Table E.2). On the average, the percentage of the observed fifth gap times was 89.1% under the model (4.35) with dependence parameter values $\boldsymbol{\phi_c} = (1.33, 2, 3, 4.67)$ when $\tau = 5$. When $\tau = 1$,the percentages of the observed gap times from the first one to the fifth one were 78.1%, 65.5%, 59.1%, 55.5% and 53.4%, respectively.

Table 4.10: Empirical means (standard deviations) [means of standard errors] of parameter estimates from the hybrid model (4.35) under six different structures (DDP_RE, SDP_RE, IND_RE, DDP, SDP and IND). The data were generated from the model (4.35), where $\alpha = 1$, $\phi = 0.3$, $\gamma = 0.223$, $\beta = 1.609$, $\Delta = 0.05$, $\boldsymbol{\xi} = (0.5, 1)$ and serial dependence copula parameters were $\boldsymbol{\phi_c} = (1.33, 2, 3, 4.67)$ with 2,000 simulations.

τ	**Parameter**	**DDP_RE**			**SDP_RE**			**IND_RE**		
		m = 50	**m = 100**	**m = 250**	**m = 50**	**m = 100**	**m = 250**	**m = 50**	**m = 100**	**m = 250**
	$\bar{\hat{\alpha}}$	1.030876	1.014960	1.002261	1.063573	1.042943	1.028418	2.076028	1.971891	1.892719
5		(0.302089)	(0.205850)	(0.132553)	(0.319591)	(0.215387)	(0.136999)	(1.175115)	(0.803525)	(0.511012)
		[0.271428]	[0.184191]	[0.103649]	[0.288791]	[0.195097]	[0.109629]	[0.570495]	[0.402130]	[0.255093]
	$\bar{\hat{\phi}}$	0.256684	0.278750	0.286444	0.256273	0.279680	0.288064	0.935005	0.960168	0.969267
		(0.153813)	(0.107808)	(0.067071)	(0.155924)	(0.109623)	(0.068486)	(0.275099)	(0.225405)	(0.168302)
		[0.123645]	[0.095277]	[0.060720]	[0.126194]	[0.096307]	[0.061684]	[0.150753]	[0.114396]	[0.079929]
	$\bar{\hat{\gamma}}$	0.222137	0.223382	0.223032	0.210161	0.211401	0.211750	0.340436	0.340311	0.339039
		(0.044886)	(0.032026)	(0.020900)	(0.047853)	(0.034207)	(0.022742)	(0.061387)	(0.045597)	(0.029179)
		[0.043980]	[0.030948]	[0.019585]	[0.050783]	[0.035720]	[0.022539]	[0.058348]	[0.041509]	[0.026500]
	$\bar{\hat{\beta}}$	1.605100	1.605225	1.613074	1.588268	1.588365	1.595506	1.102886	1.106954	1.112526
		(0.144105)	(0.102876)	(0.070502)	(0.152883)	(0.108372)	(0.072938)	(0.191500)	(0.141959)	(0.094983)
		[0.144207]	[0.101809]	[0.064353]	[0.131736]	[0.092856]	[0.059299]	[0.175262]	[0.124616]	[0.080067]
	$\bar{\hat{\xi}}_1$	0.500889	0.499067	0.498197	0.501521	0.500874	0.500106	0.537038	0.531303	0.529876
		(0.300572)	(0.208926)	(0.134411)	(0.307383)	(0.212112)	(0.135095)	(0.494181)	(0.372798)	(0.243647)
		[0.275446]	[0.194775]	[0.120046]	[0.280368]	[0.198159]	[0.122628]	[0.281643]	[0.210984]	[0.142991]
	$\bar{\hat{\xi}}_2$	0.988740	0.997231	0.998506	0.984789	0.996856	0.996111	1.047171	1.055822	1.062450
		(0.293991)	(0.207802)	(0.137210)	(0.299154)	(0.211384)	(0.137392)	(0.485764)	(0.372945)	(0.254795)
		[0.277271]	[0.196827]	[0.121634]	[0.282859]	[0.199901]	[0.124025]	[0.284951]	[0.213765]	[0.143979]
	$\bar{\hat{\phi}}_{c1}$	1.400922	1.366854	1.356257						
		(0.349069)	(0.238258)	(0.147617)						
		[0.337075]	[0.233487]	[0.144739]						
	$\bar{\hat{\phi}}_{c2}$	2.081750	2.039112	2.015195						
		(0.439551)	(0.311581)	(0.186065)						
		[0.429945]	[0.298958]	[0.186568]						
	$\bar{\hat{\phi}}_{c3}$	3.090370	3.045809	3.004955						
		(0.567789)	(0.390532)	(0.246617)						
		[0.566137]	[0.394728]	[0.247088]						
	$\bar{\hat{\phi}}_{c4}$	4.789192	4.721390	4.654737						
		(0.809323)	(0.558339)	(0.358821)						
		[0.807625]	[0.561893]	[0.351561]						
	$\bar{\hat{\phi}}_{c}$				2.378212	2.355100	2.339323			
					(0.270528)	(0.190153)	(0.120610)			
					[0.249747]	[0.172512]	[0.109727]			
	$\bar{\hat{\alpha}}$	1.033036	1.017533	1.004788	1.046996	1.033676	1.015623	2.059858	1.934480	1.874837
1		(0.388520)	(0.251082)	(0.154592)	(0.398750)	(0.259576)	(0.160152)	(1.300778)	(0.713480)	(0.409195)
		[0.333230]	[0.227737]	[0.135403]	[0.345406]	[0.235021]	[0.141161]	[0.719821]	[0.484065]	[0.294979]
	$\bar{\hat{\phi}}$	0.286509	0.297062	0.296726	0.279964	0.291264	0.292402	1.081102	1.113471	1.122785
		(0.262158)	(0.202543)	(0.132384)	(0.265816)	(0.206197)	(0.136172)	(0.257841)	(0.174573)	(0.103976)
		[0.217140]	[0.179715]	[0.120268]	[0.215734]	[0.179374]	[0.121961]	[0.225477]	[0.164224]	[0.106108]
	$\bar{\hat{\gamma}}$	0.224563	0.223548	0.223426	0.220668	0.220468	0.221461	0.321815	0.319482	0.317119
		(0.065070)	(0.042966)	(0.029382)	(0.071324)	(0.046123)	(0.031640)	(0.075043)	(0.051706)	(0.032392)
		[0.057622]	[0.040939]	[0.025816]	[0.071100]	[0.050429]	[0.031854]	[0.080890]	[0.057003]	[0.035988]
	$\bar{\hat{\beta}}$	1.603634	1.604690	1.603859	1.591152	1.591652	1.591040	1.323123	1.319620	1.322257
		(0.177843)	(0.123274)	(0.076939)	(0.180033)	(0.123984)	(0.077454)	(0.242470)	(0.169152)	(0.104335)
		[0.172768]	[0.121926]	[0.076901]	[0.153827]	[0.108316]	[0.068300]	[0.240621]	[0.169042]	[0.106582]
	$\bar{\hat{\xi}}_1$	0.521898	0.510730	0.507021	0.516564	0.506087	0.503463	0.523978	0.513414	0.507529
		(0.363108)	(0.243812)	(0.152308)	(0.368579)	(0.248473)	(0.155935)	(0.559334)	(0.389862)	(0.235281)
		[0.330631]	[0.230077]	[0.141435]	[0.333280]	[0.231890]	[0.143845]	[0.350883]	[0.256396]	[0.164353]
	$\bar{\hat{\xi}}_2$	1.023761	1.008724	1.004916	1.022385	1.000012	1.001846	1.026333	1.026240	1.021231
		(0.370772)	(0.249418)	(0.158340)	(0.374276)	(0.250097)	(0.162293)	(0.552395)	(0.384383)	(0.238230)
		[0.341567]	[0.238999]	[0.148310]	[0.346411]	[0.242283]	[0.151380]	[0.357476]	[0.258303]	[0.165245]
	$\bar{\hat{\phi}}_{c1}$	1.403351	1.372622	1.353575						
		(0.377582)	(0.258441)	(0.154038)						
		[0.368126]	[0.254351]	[0.158468]						
	$\bar{\hat{\phi}}_{c2}$	2.072585	2.022296	2.000367						
		(0.518116)	(0.333927)	(0.208708)						
		[0.489770]	[0.334789]	[0.208758]						
	$\bar{\hat{\phi}}_{c3}$	3.126600	3.056077	3.007607						
		(0.709540)	(0.479093)	(0.289544)						
		[0.671157]	[0.461621]	[0.286216]						
	$\bar{\hat{\phi}}_{c4}$	4.882913	4.760599	4.687949						
		(1.065824)	(0.711608)	(0.430816)						
		[0.994533]	[0.681860]	[0.422507]						
	$\bar{\hat{\phi}}_{c}$				2.297185	2.274592	2.261772			
					(0.304402)	(0.206649)	(0.131334)			
					[0.272013]	[0.188598]	[0.117909]			

The hybrid model under DDP_RE and SDP_RE dependence structures consistently estimate both external and internal covariate effects. The estimated bias of the external covariate effects $\boldsymbol{\xi}$ from IND_RE model is not significant in each case, but internal covariate effects γ and β are significantly biased. For example, when the hybrid IND_RE model fitted to the data generated with $m = 250$, $\alpha = 1$, $\phi = 0.3$, $\beta = 1.609$, $\Delta = 0.05$, $\gamma = 0.223$, $\boldsymbol{\xi} = (0.5, 1)$, $\tau_K = (0.4, 0.5, 0.6, 0.7)$ and $\tau = 5$, the empirical means of estimates of α, ϕ, β, γ, ξ_1 and ξ_2 are $1.893, 0.969, 0.339, 1.113, 0.530$ and 1.062, respectively, and the empirical standard deviations of the estimates are $0.511, 0.168, 0.029, 0.095, 0.244$ and 0.255, respectively.

Similar to the previous simulation results presented in Table E.2, the estimated bias in the parameter estimates from the models (DDP, SDP and IND) ignoring heterogeneity across individuals is significant in each case. For example, when the hybrid DDP model fitted to the data generated with $m = 250$, $\alpha = 1$, $\phi = 0.3$, $\beta = 1.609$, $\Delta = 0.05$, $\gamma = 0.223$, $\boldsymbol{\xi} = (0.5, 1)$, $\tau_K = (0.4, 0.5, 0.6, 0.7)$ and $\tau = 5$, the empirical means of the estimates of α, β, γ, ξ_1 and ξ_2 are $0.768, 0.231, 1.677, 0.416$ and 0.819 with the empirical standard deviations $0.086, 0.030, 0.067, 0.132$ and 0.131.

Similar to what we observe for the IND model in Table E.2, the results from the model IND in Table E.2 also show that the assumption of independent gap times coming from the same individual or ignoring the heterogeneity across individuals may lead to the wrong conclusions about the dynamic features and external effects in recurrent event processes.

4.8 Application: Recurrent Asthma Attacks in Children

Following the analysis presented in Section 4.4, we use the data from a prevention trial in infants with a high risk of asthma to illustrate the methods considered in Sections 4.5 and 4.6. The data set is discussed in Section 1.2.2. The study started after random allocation of six months old subjects to a placebo control group or an active drug treatment group. There were 483 asthma attacks among 119 children in the control group and 336 asthma attacks among 113 children in the treatment group, during 18 months of follow-up. Our main aim here is to check whether the occurrence of an asthma attack triggers subsequent asthma attacks, and to investigate whether there is a persistent effects of the number of previous events on the future event occurrences. We, therefore, consider the hybrid gap time models (4.12) and (4.20) under DDP_RE, SDP_RE, IND_RE, DDP, SDP, and IND dependence models for each group separately. We would like to remind that RE means, that the model includes random effects to deal with unexplained heterogeneity across individuals. The acronym DDP stands for difference dependence parameters, which defines the Markov type dependent structured model with different parameters for the dependence series, SDP stands for single dependence parameter, which defines the same dependent structured model with only a single parameter for the dependence series and IND stands for a model with independent gap times. We use a Clayton copula to model the dependency between consecutive gap times of asthma attacks. Our analysis includes only the first five gap times of each individual. The percentages of complete first to fifth gap times are respectively, 100%, 68.9%, 52.1%, 37.8%, 31.1% in the treatment group and 100%, 55.7%, 33.6%, 23.9%, 17.7% in the control group. The model parameter estimates and standard errors are provided in Table 4.11. To obtain the maximum likelihood estimates of the parameters, we minimize the corresponding negative log of likelihood functions with respect to the parameters with the `nlminb` function in `optimx` `R` package. The `R` package returns the Hessian matrix at the minimum point. We inverted that matrix to obtain the corresponding standard errors for the parameter estimates.

Table 4.11, shows that the heterogeneity parameter ϕ is not significant in control and treatment groups. Therefore, we conclude that both groups share the same characteristics within the groups. This result suggests that the marginal hazard model (4.12) given in Section 4.5 is more adequate to model the gap times than the hazard function with random effects model introduced in Section 4.6. Especially, in the treatment group, the maximum likelihood estimates of the parameters in DDP_RE, SDP_RE and IND_RE are respectively very similar to those obtained from the DDP, SDP, and IND models. Overall, the maximum likelihood estimates of the heterogeneity parameter ϕ in the control group models are higher than those obtained from the treatment group models, but they are not significant. We therefore, consider the DDP, SDP, and IND models in each group. Based on the AIC, we base our analysis on the DDP model in the treatment group and on SDP model in the control group.

Table 4.11: Parameter estimates (and standard errors) of the models which are fitted to the asthma data.

Parameter	Treatment Group						Control Group					
	DDP_RE	SDP_RE	IND_RE	DDP	SDP	IND	DDP_RE	SDP_RE	IND_RE	DDP	SDP	IND
$\hat{\alpha}$	1.3485	1.3055	1.3521	1.3374	1.2944	1.2362	1.9749	1.9637	2.1310	1.7622	1.7626	1.7426
	(0.1286)	(0.1263)	(0.1540)	(0.1264)	(0.1240)	(0.1138)	(0.2483)	(0.2430)	(0.2531)	(0.1528)	(0.1525)	(0.1461)
$\hat{\phi}$	0.0156	0.0158	0.1236				0.1233	0.1188	0.2174			
	(0.0079)	(0.0110)	(0.0911)				(0.1003)	(0.0990)	(0.0956)			
$\hat{\gamma}$	-0.1192	-0.0184	-0.0110	-0.1173	-0.0155	0.0410	-0.0538	-0.0474	-0.0673	-0.0059	-0.0041	0.0156
	(0.0909)	(0.0772)	(0.0736)	(0.0913)	(0.0776)	(0.0650)	(0.0655)	(0.0620)	(0.0570)	(0.0542)	(0.0524)	(0.0479)
$\hat{\beta}$	1.1375	1.1263	1.1508	1.1372	1.1261	1.1776	0.7692	0.7687	0.7879	0.7943	0.7957	0.8727
	(0.1581)	(0.1501)	(0.1495)	(0.1580)	(0.1500)	(0.1457)	(0.1297)	(0.1288)	(0.1277)	(0.1274)	(0.1266)	(0.1218)
$\hat{\phi}_{c1}$	-0.0187			-0.0084			0.2072			0.2513		
	(0.1062)			(0.1059)			(0.1083)			(0.1028)		
$\hat{\phi}_{c2}$	0.2790			0.2879			0.1866			0.2121		
	(0.1798)			(0.1786)			(0.1703)			(0.1664)		
$\hat{\phi}_{c3}$	0.7470			0.7553			0.3764			0.4060		
	(0.2371)			(0.2357)			(0.2179)			(0.2135)		
$\hat{\phi}_{c4}$	0.7325			0.7426			0.3030			0.3044		
	(0.2998)			(0.2971)			(0.2034)			(0.1983)		
$\hat{\phi}_{c}$		0.1990			0.2104			0.2386			0.2708	
		(0.0988)			(0.0976)			(0.0819)			(0.0771)	
$\ell(\hat{\boldsymbol{\theta}})$	-61.2769	-67.0163	-68.5027	-61.2065	-66.9872	-69.6988	-13.7098	-14.0632	-18.9350	-14.6078	-14.9126	-22.7457
AIC	138.5538	144.0327	145.0055	**136.4129**	141.9743	145.3976	43.4197	38.1264	45.8700	43.2157	**37.8252**	51.4914

In the treatment group with the DDP model, the trend effect γ is not significant ($\hat{\gamma} = -0.1173$, 95% c.i. (-0.2962, 0.0616), $p-$value $= 0.1989$). The carryover effect β is significant ($\hat{\beta} = 1.1372$, 95% c.i. (0.8275, 1.4469), $p-$value ≈ 0). Comparing these results, with the results given in Table 4.6, we observe that the trend is not significant in the model but it is significant in Table 4.6, and the carryover effect is still significant but it is higher in Table 4.6 when we consider only first two gap times.

In the control group, we consider the results under the SDP model presented in Table 4.11. The trend effect is not significant in this model ($\hat{\gamma} = -0.0041$, 95% c.i. (-0.1068, 0.0986), $p-$value $= 0.9376$). The carryover effect β is significant ($\hat{\beta} =$ 0.7943, 95% c.i. (0.5462, 1.0424), $p-$value ≈ 0). Comparing with the results given in Table 4.5, the estimate of the trend is not significant in the SDP model but it is significant in Table 4.5, and the estimate of the carryover effects is smaller in the SDP model.

Comparison of the results between the treatment and control groups in Table 4.11 reveals that the baseline intensity of the treatment group is smaller than that of the control group. That is, the overall rate of the asthma attack is smaller in the treatment group. In both groups, the trend due to previous number of events is not significant. However, the carryover effect in the treatment group is larger than that in the control group. After every asthma attack, the intensity is increased multiplicatively by 3.119 and 2.216 for 56 days in treatment and control groups, respectively. Since in both groups the trend effects due to the number of previous events are insignificant, after 56 days from each attack the intensities reset to the corresponding baseline intensity values.

A final note in this section is that the estimates of the carryover effect β in Table 4.11 reflect on average effect obtained by using the first five gap times. The model can be easily extend to obtain estimates of such effects in each gap times by considering a separate carryover effects parameter $\beta_k, k = 2, \ldots, 5$, so that, if there is an interest, more detailed analysis of the carryover effects can be developed.

Chapter 5

Summary and Future Work

This chapter includes a summary of the thesis and possible extensions of the research carried out in the previous chapters. The future work section consists of a lengthy discussion of some prominent goodness-of-fit tests for the adequacy of the models discussed in Chapter 4.

5.1 Summary and Conclusions

Dynamic models are instrumental to make inferences on features of recurrent event processes when dependence on the past in a process is not ignorable. Comparing with models ignoring the past developments, dynamic models provide deeper insight into underlying event-generating processes. Dynamic models include dynamic covariates, which are time varying functions of the history of a process. The inherent nature of such covariates creates complex inter-relationships. Furthermore, dependent gap times, censoring, and heterogeneity make modeling and inference challenging issues when dynamic covariates are of interest. Thus, elaborate intensity modeling is required for making valid inferences.

In this study, we explored two crucial features of the recurrent event processes through dynamic models. These features are serial clustering of events over time and monotonic trends. Carryover effects define a situation, where the occurrence of an event causes a temporary increase, or in some cases decrease, in event intensity functions of recurrent event processes. Therefore, presence of carryover effects may result in event clustering or, in some cases, sparsity of events. Trends may depend on calendar time or the number of previous events in recurrent event processes. Former type of trends has been much discussed in the literature. In this thesis, we focused on monotonic trends due to the number of previous events, which is a dynamic covariate. Recurrent event data can be analyzed under two broad classes of modeling approaches. The first approach is through count-based models, in which the event counts of recurrent processes are modeled through rate or intensity functions. In the second approach, gap times between sequential event occurrence times are modeled by using hazard functions specified for gap times. An important outcome of our study is that these two modeling approaches do not necessarily produce same results. In particular, if the interest is to make inference on gap times, the inference based on count models may result in wrong conclusions, especially when there is significant dependency between gap times. As detailed below, we first considered the count-based modeling approach to make inferences on the dynamic covariates in Chapter 3. We then examined the gap time modeling approach, which allows us to estimate the effects of those dynamic covariates for specific gap times. Since the independent gap times assumption is not realistic in most practical situations, we modeled the joint distribution of gap times with copulas in Chapter 4. We analyzed a data set from an asthma prevention trial in young children to illustrate the methods developed in this thesis.

Our focus in Chapter 3 was on the simultaneous modeling of carryover effects and number of previous events as dynamic covariates through count-based models. To this end, we introduced a parametric multiplicative intensity model. We first discussed the technical issue of dishonest processes, which may arise in some settings because of the trend component in our model. We presented the results of a simulation study showing that our model may suffer from this issue. We, therefore, replaced the trend component with its trimmed version, which resolves the dishonest process issue and does not affect the estimation of model parameters in most applied settings. The estimation of the model parameters was then carried out by using the maximum likelihood method. We analytically investigated the large sample properties of the maximum likelihood estimators by validating the regularity conditions stated by Andersen et al. (1993). As showed in Section 3.1.1, the maximum likelihood estimators are consistent, and have the usual asymptotic normal distribution under the regularity conditions stated in Appendix C. We next introduced a random effects model to deal with unexplained excess heterogeneity. With the count-based modeling approach, the random effect models create an issue of confounding between carryover effects and heterogeneity in some settings. The results of a simulation study presented in Section 3.3 revealed that, if there is excess heterogeneity across the baseline intensity functions of individuals in a study, it is important to address the heterogeneity in the dynamic count-based models. This issue is more pronounced in the early stages of follow-up of individuals when there are not too many events experienced by individuals. It should be noted that, in most of the epidemiology studies, individuals experience a small number of events during their follow-ups. Section 3.3 also includes the results of simulation studies conducted to investigate the issue of confounding of the carryover effects and trends due to previous number of events in homogeneous and heterogeneous processes. Simulations for both identical and nonidentical processes revealed that some model misspecifications might result in wrong conclusions about carryover effects when the trend is present. The hybrid versions of the models do not suffer from such an issue. Finally, in the last section of this chapter, we analyzed the asthma data with dynamic count-based models.

Because of the above-mentioned shortcomings of the dynamic models based on count data, we considered models for gap times of recurrent event processes in Chapter 4. In Section 4.1, we introduced the hazard functions of gap times to capture the carryover effects and monotonic trends due to previous number of events. These marginal hazard models create a discrete mixture distribution for the second and subsequent gap times. They are equivalent to the distribution of gap times obtained from the dynamic model (3.1) introduced in Section 3.1 only when the baseline intensity function is constant and gap times are independent. To extend our approach, we used copulas to model the dependence between successive gap times. Copulas have the important advantage of modeling dependence over other methods that marginal distributions of gap times and dependence structures can be separately specified. This allows us to estimate the effects of dynamic covariates over a specific gap time, which cannot be obtained by the models based on event counts unless the gap times are assumed to be independent. In Section 4.2, we introduced the maximum likelihood estimation method based on the first two-gap times. In this section our goal is to estimate the carryover effects in the marginal distribution of the second gap time. In this specific case, the trend component was considered as an adjustment for the baseline intensity function after the first event time. Since the method is based on the maximization of the likelihood function, the standard large sample properties follow under the Cramér-type regularity conditions (see Cox and Hinkley, 1974, p. 281). In Section 4.3, we presented the results of a simulation study conducted to investigate the aforementioned issues arising from event count models. We observed that, when the gap times within processes are dependent to each other, the count-based approach provides significantly biased results for the estimation of the model parameters, while the gap time based approach provides valid inference results. Therefore, we recommend using the copula-based method introduced in Section 4.2 to assess the effects of the dynamic features in the marginal distribution of the second gap times. An illustrative example was given in Section 4.4.

In Section 4.5, we extended the model introduced in Section 4.2 to incorporate third and subsequent gap times. We assumed a Markov type dependence structure among gap times within a recurrent event process for the models with more than three gap times. In Section 4.6, we further extended the model in Section 4.5, and defined as random effects model to deal with heterogeneity among the individual processes. We encountered some computational issues while maximizing the log likelihood function for the random effects models. Therefore, we used the Gauss–Laguerre quadrature approximation method to obtain an approximation of the corresponding likelihood function. We noticed that this approximation provides valid estimations for the parameters through many simulation studies. Based on the simulation results given in Section 4.7, we observed that the hybrid models are the ideal approach to make valid conclusions about dynamic features because it can handle the issues related to confounding and censoring. In Section 4.8, we illustrated the methods discussed in Sections 4.5 and 4.6 by analyzing the asthma data.

5.2 Future Work

In this section, we introduce some research topics to be investigated in the future.

5.2.1 Separate Carryover Effects for the Gap Times

In this thesis, we assumed that carryover effects are the same for the serial gap times in a recurrent event process. In other words, the estimate of carryover effects parameter

β represented an overall average of carryover effects among the gap times within a process. Thus, the estimate of the effect was the same for each gap time. The reason why we assumed this parameterization was to facilitate the comparison of the results obtained from count-based models in Chapter 3 with those obtained from gap time models in Chapter 4. The models for the gap times can be easily extended to include carryover effects separately for each gap time. To do this, we can assign different carryover effects parameters β_k for each gap time W_k, $k = 2, 3, \ldots$. Then, the hazard function of the k^{th} gap time can be defined as

$$h_k(w) = \begin{cases} \alpha, & k = 1, \\ \alpha e^{[(k-1)\gamma + \beta_k]} \, I(w \leq \Delta) + \alpha e^{[(k-1)\gamma]} \, I(w > \Delta), & k = 2, 3, \ldots. \end{cases} \tag{5.1}$$

A similar estimation method as explained in Chapter 4 can be then applied.

5.2.2 Choice of the Risk Window

Either for external or internal carryover effects, methods proposed in the literature, including ours, need the specification of a value for the length of the risk window Δ. An important issue, which may lead to biased estimates of β in models (3.1) and (3.38), is related to the misspecification of Δ. For a given data set, a sensitivity analysis can be conducted to identify an appropriate Δ value. However, this approach requires model fitting for various values of Δ. A scan test for identifying optimal risk windows was proposed by Xu et al. (2013) for external carryover effects. This test depends on the Poisson model assumption, and departures from this assumption may lead to wrong conclusions. The copula method discussed in our study allows us to obtain a nonparametric estimator of the distributions of the marginal gap times. Therefore, a formal procedure for identification of optimal risk windows can be developed. We will investigate this issue as a future work.

5.2.3 Terminating Events

In some studies, the observation of a process is terminated by the occurrence of an event, called the terminating event. A typical example is the death of an individual during the follow-up for recurrent events. This situation should not be confused with the censoring concept discussed in Chapter 2. If terminating events are not addressed correctly, the methods may lead to wrong conclusions about dynamic features of recurrent event processes. Terminating events are usually modeled as an absolute state in stochastic processes, and models are extended accordingly (Cook and Lawless, 1997; Andersen et al., 2019). As a future work, we will extend our study by developing the proposed models to incorporate terminating events.

5.2.4 Goodness-of-fit Procedures

In Chapter 4, we introduced gap time based models for capturing dynamic features in the recurrent event processes. Here, we discuss some possible methods to check model adequacy for such gap time models. We first discuss adequacy for the models using the first two gap times. We next extend our discussion to deal with more than two gap times. The models in Chapter 4 are defined by two sets of parameters. One set, called marginal parameters, defines the marginal characteristics of gap times. The other set, called copula parameters, specifies the dependency between gap times. Marginal parameters are the main objective of the entire thesis. Therefore, they are the parameters of interest in this chapter. The parameters for dependency specification are considered as nuisance parameters.

Let's consider the copula model for the first two gap times in a recurrent event process. Following the likelihood function (4.5), the likelihood function for the data $\{(w_{i1}, w_{i2}, \delta_{i1}, \delta_{i2}) : 1 = 1, \ldots, m\}$ is given by

$$L(\boldsymbol{\beta}^*, \boldsymbol{\alpha}^*) = \prod_{i=1}^{m} \left[\frac{\partial^2 C(F_1(w_{i1}; \boldsymbol{\beta_1^*}), F_2(w_{i2}; \boldsymbol{\beta_2^*}); \boldsymbol{\alpha}^*)}{\partial w_{i1} \partial w_{i2}} \right]^{\delta_{i1}\delta_{i2}} \left[\frac{\partial F_1(w_{i1}; \boldsymbol{\beta_1^*})}{\partial w_{i1}} - \frac{\partial C(F_1(w_{i1}; \boldsymbol{\beta_1^*}), F_2(w_{i2}; \boldsymbol{\beta_2^*}); \boldsymbol{\alpha}^*)}{\partial w_{i1}} \right]^{\delta_{i1}(1-\delta_{i2})} [1 - F_1(w_{i1}; \boldsymbol{\beta_1^*})]^{(1-\delta_{i1})}, \tag{5.2}$$

where $\boldsymbol{\beta_1^*}$ and $\boldsymbol{\beta_2^*}$ are the vectors of parameters of marginal distributions of gap times W_1 and W_2, respectively, and $\boldsymbol{\alpha}^*$ is the vector of copula parameters. For the first and second gap times, $\boldsymbol{\beta_1^*}$ and $\boldsymbol{\beta_2^*}$ can be specified as α and $(\alpha, \beta, \gamma)'$, respectively.

According to the choice of copula, the vector $\boldsymbol{\alpha}^*$ includes the parameters from copula families, which are introduced in Section 2.4. In Sections 5.2.4.1 and 5.2.4.2, we provide two procedures for checking adequacy of the models with first two gap times. For models with more than two gap times, a testing procedure is provided in Section 5.2.4.3 to check whether the Markov type dependency structure in the serial gap times is adequate or not. We want to note that we give some introductions to the model adequacy checks and leave the study of large sample properties of those tests as a future work. As suggested by Stute et al. (1993), we use bootstrap procedures for the tests presented in Sections 5.2.4.1 and 5.2.4.2. In particular, we apply the bootstrap procedure given in Appendix D to calculate the corresponding $p-$values of the tests.

5.2.4.1 Two-Stage Test Procedure

In this section, we introduce a two-stage goodness-of-fit procedure. In the first stage, the adequacy of copula choice is examined. In the second stage, the adequacy of the fitted marginal distribution of second gap time W_2 is examined. Lawless and Yilmaz (2011) have introduced semiparametric estimation of copula models for the first two gap times. They model the joint distribution of the successive gap times by using copula functions, and provide two semi-parametric estimation procedures, in which copula parameters are estimated without parametric assumptions on the marginal distributions. Their marginal distributions are specified by using discrete hazard parametrization method, which provides robust estimates of the marginal distributions. Lawless and Yilmaz (2011) define discrete hazard parametrization $\boldsymbol{\lambda_k^*} = (\lambda_{1k}^*, \ldots, \lambda_{rk}^*)$ for the k^{th} gap time W_k as $\lambda_{1k}^* = F_k(w_{(l)k}^*)$ and $\lambda_{lk}^* = (F_k(w_{(l)k}^*) - F_k(w_{(l-1)k}^*))/(1 - F_k(w_{(l-1)k}^*))$ for $l = 2, \ldots, r$, where $w_{(1)k}^* < w_{(2)k}^* < \cdots < w_{(r)k}^*$ are the distinct observed w_{ik}'s with $\delta_{ik} = 1$, for $i = 1, \ldots, m$, where $r \leq m$. Consequently, the marginal distribution function of the k^{th} gap time W_k can be defined as a function of λ_{lk}^*'s where $F_k(w_{ik}) = 1 - \prod_{l: w_{(l)k}^* \leq w_{ik}} (1 - \lambda_{lk}^*)$ and $F_k(w_{ik}^-) = 1 - \prod_{l: w_{(l)k}^* < w_{ik}} (1 - \lambda_{lk}^*)$. The non-parametric estimates of the marginal distributions of the gap times can be obtained by plugging in the discretized versions of the c.d.f.'s into the likelihood function (5.2) and maximizing the log likelihood function (5.2) with respect to $\boldsymbol{\lambda_1^*}$, $\boldsymbol{\lambda_2^*}$ and $\boldsymbol{\alpha}^*$.

The above-mentioned nonparametric estimates of marginal distributions can be used to check the performance of our models. Lawless and Yilmaz (2011) consider two estimation procedures. In the first procedure, the Kaplan-Meier method is used to estimate the distribution of the first gap time, whereas in the second procedure, a discrete hazard parameterization is used. Due to computational efficiency, we pick the former procedure, which is a two-stage estimation procedure, to develop model checks in this section.

Stage 1: Testing for the Copula Choice

We implement a two-stage procedure to develop goodness-of-fit tests. In the first stage, we check the adequacy of the specification of the copula model by applying a model expansion technique (Lawless, 2003, Section 10.2.2). The alternative hypothesis includes an expanded family of the copula models, which contains the copula model in the null hypothesis as a special case. For example, suppose a copula model has the vector of parameters $\boldsymbol{\alpha}^* = (\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*)$ and a model of special interest is obtained when $\boldsymbol{\alpha}_1^* = \boldsymbol{\alpha}_{10}^*$. A test for the composite null hypothesis $H_0 : \boldsymbol{\alpha_1^*} = \boldsymbol{\alpha_{10}^*}$ against the composite alternative hypothesis $H_1 : \boldsymbol{\alpha_1^*} \neq \boldsymbol{\alpha_{10}^*}$ can be then developed for the adequacy of the copula model based on the value of $\boldsymbol{\alpha_1^*}$ under the null hypothesis. We propose to test such hypotheses with the likelihood ratio statistic

$$\Lambda_{stage1}(\boldsymbol{\alpha}_{10}^*) = -2\log\frac{L(\hat{\boldsymbol{\lambda}}_{\mathbf{2}}^*(\boldsymbol{\alpha}_{10}^*), \boldsymbol{\alpha}_{10}^*, \hat{\boldsymbol{\alpha}}_2^*(\boldsymbol{\alpha}_{10}^*))}{L(\hat{\boldsymbol{\lambda}}_{\mathbf{2}}^*, \hat{\boldsymbol{\alpha}}^*)}, \tag{5.3}$$

where $\hat{\boldsymbol{\lambda}}_{\mathbf{2}}^*(\boldsymbol{\alpha}_{10}^*)$ and $\hat{\boldsymbol{\alpha}}_2^*(\boldsymbol{\alpha}_{10}^*)$ are obtained by maximizing $L(\boldsymbol{\lambda}_{\mathbf{2}}^*, \boldsymbol{\alpha}^*)$ in (5.2) with $\boldsymbol{\alpha}_1^* = \boldsymbol{\alpha}_{10}^*$. As suggested by Lawless and Yilmaz (2011), we use the bootstrap method provided in Appendix D to calculate the $p-$values for the test statistics.

Since the Clayton (2.55) and the Gumbel-Hougaard (2.58) copula functions are special cases of two parameter copula function (2.64), we consider the copula based likelihood function (5.2) with (2.64) to develop test for the choice of copula. That is, the alternative hypothesis includes the two parameter copula model

$$C_{\alpha_2^*,\alpha_1^*}(u_1,u_2) = \left\{\left[\left(u_1^{-\alpha_2^*}-1\right)^{\alpha_1^*} + \left(u_2^{-\alpha_2^*}-1\right)^{\alpha_1^*}\right]^{1/\alpha_1^*} + 1\right\}^{-1/\alpha_2^*}, \alpha_1^* \geq 1 \text{ and } \alpha_2^* > 0, \tag{5.4}$$

where α_1^* and α_2^* are copula parameters. To test whether the dependency among first two gap times is Clayton, the null and alternative hypotheses can be specified as $H_0 : \alpha_1^* = 1$ & $\alpha_2^* \in (0, \infty)$ and $H_1 : \alpha_1^* \neq 1$ & $\alpha_2^* \in (0, \infty)$, respectively. Based on (5.3), we can use the likelihood test statistic

$$\Lambda^{(1)}_{stage1}(\alpha_{10}^* = 1) = -2 \log \frac{L(\hat{\boldsymbol{\lambda}}_{\mathbf{2}}^{*}(\alpha_{10}^*), \alpha_{10}^*, \hat{\alpha_2^*}(\alpha_{10}^*))}{L(\hat{\boldsymbol{\lambda}}_{\mathbf{2}}^{*}, \hat{\alpha_1^*}, \hat{\alpha_2^*})}, \tag{5.5}$$

to test the null hypothesis. Similarly, for the Gumbel-Hougaard dependency, the null and alternative hypotheses can be specified as $H_0 : \alpha_2^* = 0$ & $\alpha_1^* \in (1, \infty)$ and $H_1 : \alpha_2^* \neq 0$ & $\alpha_1^* \in (1, \infty)$, respectively. For testing this hypothesis, we propose the likelihood ratio test statistic

$$\Lambda^{(2)}_{stage1}(\alpha_{20}^* = 0) = -2 \log \frac{L(\hat{\boldsymbol{\lambda}}_{\mathbf{2}}^{*}(\alpha_{20}^*), \alpha_{20}^*, \hat{\alpha_1^*}(\alpha_{20}^*))}{L(\hat{\boldsymbol{\lambda}}_{\mathbf{2}}^{*}, \hat{\alpha_1^*}, \hat{\alpha_2^*})}. \tag{5.6}$$

To test whether the first two gap times are independent, the null and alternative hypotheses can be specified as $H_0 : \boldsymbol{\alpha}^* = (\alpha_1^*, \alpha_2^*) = (1, 0)$ and $H_1 : \boldsymbol{\alpha}^* = (\alpha_1^*, \alpha_2^*) \neq (1, 0)$, respectively. In this case, we use the following likelihood ratio statistic for testing the null hypothesis.

$$\Lambda^{(3)}_{stage1}(\alpha_{10}^* = 1, \alpha_{20}^* = 0) = -2 \log \frac{L(\hat{\boldsymbol{\lambda}}_{\mathbf{2}}^{*}(\alpha_{10}^*, \alpha_{20}^*), \alpha_{10}^*, \alpha_{20}^*)}{L(\hat{\boldsymbol{\lambda}}_{\mathbf{2}}^{*}, \hat{\alpha_1^*}, \hat{\alpha_2^*})}. \tag{5.7}$$

It should be noted that, in all three tests considered above, the parameters in $\boldsymbol{\lambda}_{\mathbf{2}}^{*}$ are nuisance. The consistency of the estimation method has been discussed with simulations by Lawless and Yilmaz (2011). In Step 3 of the bootstrap method given in Appendix D, the estimates of marginal distribution functions for the first and second gap times are the Kaplan-Meier estimate and $F_2(w_{i2}; \hat{\boldsymbol{\lambda}}_{\mathbf{2}}^{*}) = 1 - \prod_{l: w^*_{(l)2} \leq w_{i2}} (1 - \hat{\lambda}^*_{l2})$, respectively.

For the illustrative purpose, we apply this procedure for the asthma data. The estimated $p-$values for the test statistic $\Lambda^{(i)}_{stage1}$, $i = 1, 2$ and 3, under three null hypotheses are presented in Table 5.1.

Table 5.1: The estimated $p-$values in Stage 1 for the asthma data.

H_0	Test Statistics	Treatment Group	Control Group
Clayton	$\Lambda^{(1)}_{stage1}$	0.3795	0.2425
Gumbel-Hougaard	$\Lambda^{(2)}_{stage1}$	0.1981	0.1423
Independent	$\Lambda^{(3)}_{stage1}$	0.2405	0.0220

Each $p-$value in Table 5.1 is obtained by generating 2,000 bootstrap samples. There is no enough evidence against the independent assumption for the treatment group, with an estimated $p-$value 0.2405 for the likelihood ratio statistic $\Lambda^{(3)}_{stage1}$ given in (5.7). There is strong evidence against the independent assumption for the control group, with estimated $p-$values 0.0220 for the likelihood ratio statistic (5.7). There is no evidence against both Clayton and Gumbel-Hougaard models, with estimated $p-$values 0.2425 and 0.1423 for the test statistics (5.5) and (5.6), respectively.

Stage 2: Testing for the Specification of the Marginal Model for Gap Times

In the second stage, we consider the adequacy of the marginal model specification for the gap times. Since the model is based on the first two gap times, our focus here is on the adequacy of the model for the second gap time. We use several test statistics to measure the discrepancy between the fitted marginal c.d.f. of the second gap time and its corresponding nonparametric counterpart obtained from a procedure developed in Lawless and Yilmaz (2011). To construct a test in the second stage, we define the null and alternative hypotheses as

$$H_0 : F_2 \in \{F_2(\cdot;\boldsymbol{\beta}^*);\boldsymbol{\beta}^* \in \Theta_0\} \qquad \text{vs.} \qquad H_1 : F_2 \notin \{F_2(\cdot;\boldsymbol{\beta}^*);\boldsymbol{\beta}^* \in \Theta_0\}.$$

By restricting the parameter space Θ_0 in the null hypothesis, we can test whether the estimated c.d.f. of the second gap time has a particular form or not.

For the asthma data, we assume exponential distribution with a constant rate α for the first gap time W_1. By recalling our model in (4.12), we can derive the desired c.d.f. of the second gap time W_2 as

$$F_2(w;\boldsymbol{\beta}^*) = 1 - \exp\left[-\alpha e^{\gamma}\left\{w + \min(w,\Delta)\left(e^{\beta}-1\right)\right\}\right], \quad w > 0, \tag{5.8}$$

where $\boldsymbol{\beta^*} = (\alpha, \beta, \gamma)'$. We assume four different marginal models for the second gap time W_2; (i) a null model with c.d.f.

$$F_2(w; \boldsymbol{\beta^*}) = 1 - \exp\left[-\alpha w\right], \quad w > 0, \tag{5.9}$$

(ii) a trend model with

$$F_2(w; \boldsymbol{\beta^*}) = 1 - \exp\left[-\alpha \mathrm{e}^{\gamma} w\right], \quad w > 0, \tag{5.10}$$

(iii) a carryover effects model with

$$F_2(w; \boldsymbol{\beta^*}) = 1 - \exp\left[-\alpha\left\{w + \min(w, \Delta)\left(\mathrm{e}^{\beta} - 1\right)\right\}\right], \quad w > 0, \tag{5.11}$$

and (iv) a hybrid model with

$$F_2(w; \boldsymbol{\beta^*}) = 1 - \exp\left[-\alpha \mathrm{e}^{\gamma}\left\{w + \min(w, \Delta)\left(\mathrm{e}^{\beta} - 1\right)\right\}\right], \quad w > 0. \tag{5.12}$$

Let $F_2(w; \hat{\boldsymbol{\beta}}^*)$ be the estimated c.d.f. of the second gap time from our model and $\hat{F}_{m2}(w)$ be the corresponding nonparametric counterpart obtained from a procedure developed by Lawless and Yilmaz (2011). We propose to use the following test statistics to measure the discrepancy between the fitted and nonparametric estimates of the distribution functions in the second stage for the marginal model adequacy check.

1. The Kolmogorov-Smirnov (KS) test statistic

$$D_m^* = \sqrt{m} \sup_w \left| \hat{F}_{m2}(w) - F_2(w; \hat{\boldsymbol{\beta}}^*) \right|. \tag{5.13}$$

2. The Cramér-Von Mises (CM) test statistic

$$C_m^* = m \int_0^\infty [\hat{F}_{m2}(w) - F_2(w; \hat{\boldsymbol{\beta}}^*)]^2 dF_2(w; \hat{\boldsymbol{\beta}}^*). \tag{5.14}$$

3. The Anderson-Darling (AD) test statistic

$$A_m^* = m \int_0^\infty \frac{[\hat{F}_{m2}(w) - F_2(w; \hat{\boldsymbol{\beta}}^*)]^2}{F_2(w; \hat{\boldsymbol{\beta}}^*)(1 - F_2(w; \hat{\boldsymbol{\beta}}^*))} dF_2(w; \hat{\boldsymbol{\beta}}^*). \tag{5.15}$$

4. The Pearson's chi-squared (CS) test statistic

$$\chi^{2*}_{m} = \sum_{l=1}^{L} \frac{\left[O(w_{(l-1)2} - w_{(l)2}) - E(w_{(l-1)2} - w_{(l)2})\right]^2}{E(w_{(l-1)2} - w_{(l)2})}, \tag{5.16}$$

where

$$O(w_{(l-1)2} - w_{(l)2}) = m\left[\hat{F}_{m2}(w_{(l)2}) - \hat{F}_{m2}(w_{(l-1)2})\right], \tag{5.17}$$

$$E(w_{(l-1)2} - w_{(l)2}) = m\left[F_2(w_{(l)2}; \hat{\boldsymbol{\beta}}^{*}) - F_2(w_{(l-1)2}; \hat{\boldsymbol{\beta}}^{*})\right] \tag{5.18}$$

and $0 = w_{(0)2} < w_{(1)2} < \cdots < w_{(L)2} = \max(C - W_1)$ be the boundaries of L number of partitions of the second gap time W_2. For the illustrative purpose, we choose eight partitions.

To calculate the $p-$values, we can use the bootstrap method provided in Appendix D. In Step 2 of the bootstrap procedure, we need to specify $\widetilde{\boldsymbol{\alpha}^{*}}$ according to the conclusion obtained from the first stage.

5.2.4.2 Model Adequacy Tests Based on Conditional Marginal Distribution of Gap Times

Lin et al. (1999) introduced a consistent nonparametric estimator for the joint distribution function of times between successive events when the follow-up time is subject to right censoring. Let's denote a bivariate semi-survival function of W_1 and W_2 by $H(w_1, w_2) = \Pr(W_1 \leq w_1, W_2 > w_2)$ for $w_1, w_2 > 0$. A simple nonparametric estimator (Lin et al., 1999) for $H(w_1, w_2)$ can be written as

$$\tilde{H}(w_1, w_2) = \frac{1}{m}\sum_{i=1}^{m} \frac{I[W_{i1} \leq w_1, W_{i2} > w_2, C_i > W_{i1} + w_2]}{G(W_{i1} + w_2)}, \qquad \begin{array}{l} 0 < w_1 \leq C_{\max}, \\ 0 < w_2 \leq C_{\max} - w_1, \end{array} \tag{5.19}$$

where $G(c)$ is the survival function of the censoring time C and $C_{\max}$ is the maximum follow-up time. The survival function G in the denominator of (5.19) can be estimated by the Kaplan-Meier method based on the data $\{(w_{i1}, 1 - \delta_{i1}), i = 1, \ldots, m\}$ or $\{(\min(w_{i1} + w_{i2}, c_i), 1 - \delta_{i2}), i = 1, \ldots, m\}$. The corresponding estimator for the joint distribution function of W_1 and W_2 is then

$$\tilde{F}(w_1, w_2) = \tilde{H}(w_1, 0) - \tilde{H}(w_1, w_2), \tag{5.20}$$

where $\tilde{H}(w_1, 0)$ is an estimator of $F_1(w_1)$. Now, let's denote the conditional distribution of the second gap time W_2 given that the first gap time W_1 is less than a fixed point w_1 by

$$F_{2|1}(w_2|w_1) = \Pr(W_2 \leq w_2|W_1 \leq w_1) = F_{12}(w_1, w_2)/F_1(w_1), \qquad w_1 > 0, w_2 > 0, \tag{5.21}$$

where $F_{12}(w_1, w_2)$ is the joint c.d.f. of W_1 and W_2, and $F_1(w_1)$ is the c.d.f. of W_1 and we assume $F_1(w_1) > 0$. From (5.20), a nonparametric estimator of $F_{2|1}(w_2|w_1)$ can be written as $1 - \left[\tilde{H}(w_1, w_2)/\tilde{H}(w_1, 0)\right]$, for all $\tilde{H}(w_1, 0) > 0$. We construct tests by comparing the estimated model $F_{2|1}(w_2|w_1)$ with its nonparametric counterpart.

A goodness-of-fit test can be developed for the following null and alternative hypotheses.

$$H_0 : F_{2|1} \in \left\{F_{2|1}\left(\cdot|w_1; \boldsymbol{\alpha}^*, \boldsymbol{\beta}^*\right); \boldsymbol{\alpha}^* \in \Theta_{01}, \boldsymbol{\beta}^* \in \Theta_{02}\right\}$$
$$\text{vs.} \qquad H_1 : F_{2|1} \notin \left\{F_{2|1}\left(\cdot|w_1; \boldsymbol{\alpha}^*, \boldsymbol{\beta}^*\right); \boldsymbol{\alpha}^* \in \Theta_{01}, \boldsymbol{\beta}^* \in \Theta_{02}\right\}. \tag{5.22}$$

The nonparametric counterpart of conditional distribution of the second gap time in our model can be obtained with the method given by Lin et al. (1999). We, therefore, consider the tests based on the following modified test statistics, which measure the discrepancy between estimated conditional distribution $F_{2|1}(w|w_1; \hat{\boldsymbol{\beta}}^*)$ and its nonparametric counterpart $\hat{F}_{m2|1}(w|w_1)$. Note that the nonparametric estimator $F_{2|1}(w_2|w_1)$ is valid only for $w_1 + w_2 \leq C$, where C is the censoring time (Lin et al., 1999). Therefore, the results of the tests are based only on the conditional distribution of W_2 within the interval $(0, C_{\max} - w_1)$. We can use the following modified test statistics for the goodness-of-fit testing procedures.

1. The modified Kolmogorov-Smirnov (KS*) test statistic

$$D_m^*(w_1) = \frac{\sqrt{m}}{\hat{F}_{m2|1}(C_{\max} - w_1|w_1)} \sup_w \left|\hat{F}_{m2|1}(w|w_1) - F_{2|1}(w|w_1; \hat{\boldsymbol{\beta}}^*)\right|. \tag{5.23}$$

2. The modified Cramér-Von Mises (CM*) test statistic

$$C_m^*(w_1) = \frac{m}{\hat{F}_{m2|1}(C_{\max} - w_1|w_1)} \int_0^{C_{\max}-w_1} [\hat{F}_{m2|1}(w|w_1) - F_{2|1}(w|w_1; \hat{\boldsymbol{\beta}}^*)]^2 dF_{2|1}(w|w_1; \hat{\boldsymbol{\beta}}^*). \tag{5.24}$$

3. The modified Anderson-Darling (AD*) test statistic

$$A_m^*(w_1) = \frac{m}{\hat{F}_{m2|1}(C_{\max} - w_1|w_1)} \int_0^{C_{\max}-w_1} \frac{[\hat{F}_{m2|1}(w|w_1) - F_{2|1}(w|w_1; \hat{\boldsymbol{\beta}}^*)]^2}{F_{2|1}(w|w_1; \hat{\boldsymbol{\beta}}^*)[1 - F_{2|1}(w|w_1; \hat{\boldsymbol{\beta}}^*)]} dF_{2|1}(w|w_1; \hat{\boldsymbol{\beta}}^*). \tag{5.25}$$

4. The modified Pearson's chi-squared (CS*) test statistic

$$\chi^{2*}_{m}(w_1) = \frac{1}{\hat{F}_{m2|1}(C_{\max} - w_1|w_1)} \sum_{l=1}^{L} \frac{\left[O(w_{(l-1)2} - w_{(l)2}) - E(w_{(l-1)2} - w_{(l)2})\right]^2}{E(w_{(l-1)2} - w_{(l)2})}, \tag{5.26}$$

where

$$O(w_{(l-1)2} - w_{(l)2}) = m\left[\hat{F}_{m2|1}(w_{(l)2}|w_1) - \hat{F}_{m2|1}(w_{(l-1)2}|w_1)\right], \tag{5.27}$$

$$E(w_{(l-1)2} - w_{(l)2}) = m\left[F_{2|1}(w_{(l)2}|w_1; \hat{\boldsymbol{\beta}}^*) - F_{2|1}(w_{(l-1)2}|w_1; \hat{\boldsymbol{\beta}}^*)\right] \tag{5.28}$$

and $0 = w_{(0)2} < w_{(1)2} < \cdots < w_{(L)2} = C_{\max} - w_1$ be the boundaries of L number of partitions of the second gap time W_2.

By restricting the parameter space Θ_{01} and Θ_{02} in the null hypothesis given in (5.22), similar to the test developed in Section 5.2.4.1, we can test the adequacy of the assumed specification of the conditional c.d.f. of the second gap time by comparing the estimated conditional c.d.f. of the second gap time with its nonparametric counterpart. An advantage of this procedure over the procedure developed in Section 5.2.4.1 is that this procedure simultaneously check the adequacy of the copula specification and the marginal model specification. On the other hand, this test requires the specification of a value, say w_1, for the first gap time W_1. It may give different conclusions for different choices of w_1.

5.2.4.3 Testing for Markov Dependence

In Chapter 4, we assumed Markov type dependence structure to deal with situations, in which more than two gap times are serially observed. This assumption requires a justification. We noticed that the Markov dependence structure is a special case of the D-vine dependence structure. D-vine copulas are discussed in Section 2.4.2. The assumption of Markov dependency can be tested by using a model expansion technique as follows. The likelihood of the observed data $\{(w_{i1}, \ldots, w_{i,n_i+1}, \delta_{i1}, \ldots, \delta_{i,n_i+1}) : i = 1, \ldots, m\}$ under a D-vine dependence structure is given by

$$L_D = \prod_{i=1}^{m} f(\boldsymbol{w}_{i,1:n_i}) F(w_{i,n_i+1}|\boldsymbol{w}_{i,1:n_i})^{(1-\delta_{i,n_i+1})}, \tag{5.29}$$

where w_{ij} is the j^{th} gap time of the i^{th} process, δ_{ij} is the censoring indicator for w_{ij} and for $k = 2, \ldots, n_i$,

$$f(\boldsymbol{w}_{1:k}) = \prod_{l=1}^{k} f_l(w_l) \prod_{q=1}^{k-1} \prod_{p=1}^{k-q} c_{p,(p+q)|p+1:p+q-1} \left[F(w_p|\boldsymbol{w}_{p+1:p+q-1}), F(w_{p+q}|\boldsymbol{w}_{p+1:p+q-1}) \right]. \tag{5.30}$$

In the joint p.d.f. (5.30), the function $c_{p,(p+q)|p+1:p+q-1}$ is defined in Section 2.4.2 and

$$F(w_p|\boldsymbol{w}_{p+1:p+q-1}) = \frac{\partial C_{p,p+q|p+1:p+q-1} \left[F(w_p|\boldsymbol{w}_{p+1:p+q-1}), F(w_{p+q}|\boldsymbol{w}_{p+1:p+q-1}) \right]}{\partial F(w_{p+q}|\boldsymbol{w}_{p+1:p+q-1})}, \tag{5.31}$$

and

$$F(w_{p+q}|\boldsymbol{w}_{p+1:p+q-1}) = \frac{\partial C_{p,p+q|p+1:p+q-1} \left[F(w_p|\boldsymbol{w}_{p+1:p+q-1}), F(w_{p+q}|\boldsymbol{w}_{p+1:p+q-1}) \right]}{\partial F(w_p|\boldsymbol{w}_{p+1:p+q-1})}. \tag{5.32}$$

To illustrate this procedure, let's consider the first three gap times. By assuming the D-vine dependence structure among serially observed gap times, the likelihood contribution from the first three gap times from m independent individuals can be written as

$$L_{D3} = \prod_{i=1}^{m} f(\boldsymbol{w}_{i,1:\min(n_i,3)}) F(w_{i,\min(n_i+1,3)}|\boldsymbol{w}_{i,1:\min(n_i,2)})^{(1-\delta_{i,\min(n_i+1,3)})}. \tag{5.33}$$

With the Markov dependence assumption, using the observed data $\{(w_{i1}, \ldots, w_{\min(n_i,3)}, \delta_{i1}, \ldots, \delta_{i,\min(n_i+1,3)}) : i = 1, \ldots, m\}$, the likelihood function from the first three gap times from m independent individuals is given by

$$
\begin{aligned}
L_{M3} = \prod_{i=1}^{m} f_1(w_{i1})^{\delta_{i1}} \left[1 - F_1(w_{i1})\right]^{1-\delta_{i1}} \\
\times \left[\prod_{j=2}^{\min(n_i+1,3)} \left\{ \frac{\partial^2 C_{\alpha^*_{j-1}}(F_j(w_{ij}), F_{j-1}(w_{i,j-1}))}{\partial F_j(w_{ij}) \partial F_{j-1}(w_{i,j-1})} f_j(w_{ij}) \right\}^{\delta_{ij}} \right. \\
\left. \times \left\{ 1 - \frac{\partial C_{\alpha^*_{j-1}}(F_j(w_{ij}), F_{j-1}(w_{i,j-1}))}{\partial F_{j-1}(w_{i,j-1})} \right\}^{1-\delta_{ij}} \right]^{\delta_{i1}} .
\end{aligned}
\tag{5.34}
$$

M3 in (5.34) stands for the model with only the first three gap times, which assumes the Markov dependency among the gap times, whereas D3 in (5.33) represents the same model under the D-vine dependency. In its simple form, we then consider testing H_0 : Markov dependency against H_1 : D-vine dependency. To do this, we propose a likelihood ratio test statistic, which is given by

$$
\Lambda_{MD} = -2 \log \left(\frac{L_{M3}}{L_{D3}} \right). \tag{5.35}
$$

When the interest is on the first three gap times, the null and alternative hypotheses can be defined as $H_0 : \phi_{31} = 0$ vs. $H_1 : \phi_{31} > 0$. Note that in this case, D3 includes ϕ_{31} and all the parameters included in M3. We fitted D3 and M3 for the first three gap times obtained in the asthma study. The maximum likelihood estimates of model parameters are presented in Table 5.2.

For the treatment group, the likelihood ratio statistic is $\Lambda_{MD}(\phi_{31} = 0) = 2(\ell_{\text{D3}} - \ell_{\text{M3}}) = 2(-88.5560 + 88.8107) = 0.5094$. A $p-$value can be calculated by using the limiting distribution $\Pr(\Lambda_{MD}(\phi_{31} = 0) \le q) = 0.5 + 0.5 \Pr(\chi^2_{(1)} \le q)$ (Self and Liang, 1987), which gives a $p-$value as $0.5 \Pr(\chi^2_{(1)} \ge 0.5094) = 0.2377$, so that we do not reject $H_0 : \phi_{31} = 0$ in favor of $H_1 : \phi_{31} > 0$. Therefore, we conclude that M3 is adequate for the treatment group in asthma data. Similarly for the control group, the likelihood ratio statistic is $\Lambda_{MD}(\phi_{31} = 0) = 2(\ell_{\text{D3}} - \ell_{\text{M3}}) = 2(-42.4963 + 42.5704) = 0.1482$, which gives a $p-$value as $0.5 \Pr(\chi^2_{(1)} \ge 0.1482) = 0.3501$. We, therefore, do not reject $H_0 : \phi_{31} = 0$ in favor of $H_1 : \phi_{31} > 0$. Therefore, we conclude that M3 is also adequate for the control group in asthma data.

Table 5.2: Parameter estimates (and standard errors) of the models for the first three gap times from the asthma data.

Parameter	Treatment Group		Control Group	
	D3	M3	D3	M3
$\hat{\alpha}$	1.5187	1.5268	1.9970	1.9995
	(0.1423)	(0.1434)	(0.1794)	(0.1790)
$\hat{\gamma}$	-0.4829	-0.4760	-0.2841	-0.2890
	(0.1227)	(0.1233)	(0.1013)	(0.1004)
$\hat{\beta}$	1.3329	1.3276	0.9139	0.9134
	(0.1854)	(0.1848)	(0.1591)	(0.1594)
$\hat{\phi}_{21}$	0.0082		0.2617	
	(0.1135)		(0.1056)	
$\hat{\phi}_{31}$	0.0802		-0.0284	
	(0.1231)		(0.0700)	
$\hat{\phi}_{32}$	0.4070		0.3814	
	(0.1760)		(0.1819)	
$\hat{\phi}_{c1}$		0.0074		0.2653
		(0.1125)		(0.1053)
$\hat{\phi}_{c2}$		0.3963		0.3864
		(0.1770)		(0.1820)
$\ell(\hat{\boldsymbol{\theta}})$	-88.5560	-88.8107	-42.4963	-42.5704

This test can be extended to check the adequacy of the Markov dependency among serial gap times for more than three gap times. In that case, under the alternative hypothesis, the model has more than one excess parameters than that of under the null hypothesis. Consequently, the asymptotic distribution of the test statistic (5.35) may be complicated. We will investigate such tests in the future.

Bibliography

[1] Aalen, O. O., Borgan, O., and Gjessing, H. K. a. (2008). *Survival and Event History Analysis: A Process Point of View*. Springer, New York.

[2] Aas, K., Czado, C., Frigessi, A., and Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, 44(2):182 – 198.

[3] Abramowitz, M. and Stegun, I. A. (1948). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55. US Government printing office.

[4] Aït-Sahalia, Y., Cacho-Diaz, J., and Laeven, R. J. (2015). Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics*, 117(3):585–606.

[5] Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723.

[6] Allison, P. D. (1984). *Event history analysis: Regression for longitudinal event data*. Number 46. Sage.

[7] Amorim, L. D. and Cai, J. (2015). Modelling recurrent events: a tutorial for analysis in epidemiology. *International journal of epidemiology*, 44(1):324–333.

[8] Andersen, P. K., Angst, J., and Ravn, H. (2019). Modeling marginal features in studies of recurrent events in the presence of a terminal event. *Lifetime data analysis*, pages 1–15.

[9] Andersen, P. K., Borgan, O., Gill, R. D., and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer-Verlag, New York.

[10] Andersen, P. K. and Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. *The annals of statistics*, pages 1100–1120.

[11] Ascher, H. and Feingold, H. (1984). *Repairable systems reliability: modeling, inference, misconceptions and their causes*. M. Dekker New York.

[12] Barthel, N., Geerdens, C., Czado, C., and Janssen, P. (2018). Dependence modeling for recurrent event times subject to right-censoring with d-vine copulas. *Biometrics*.

[13] Bedford, T. and Cooke, R. M. (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence*, 32(1):245–268.

[14] Blossfeld, H.-P. and Rohwer, G. (2001). *Techniques of event history modeling: New approaches to casual analysis*. Psychology Press, Taylor & Francis.

[15] Brass, H. and Petras, K. (2011). *Quadrature theory: the theory of numerical integration on a compact interval*. Number 178. American Mathematical Society.

[16] Byar, D. P. (1980). The veterans administration study of chemoprophylaxis for recurrent stage I bladder tumours: comparisons of placebo, pyridoxine and topical thiotepa. In *Bladder tumors and other topics in urological oncology*, pages 363–370. Springer.

[17] Chang, S.-H. and Wang, M.-C. (1999). Conditional regression analysis for recurrence time data. *Journal of the American Statistical Association*, 94(448):1221–1230.

[18] Chen, F. and Chen, K. (2014). Modeling event clustering using the m-memory cox-type self-exciting intensity model. *International Journal of Statistics and Probability*, 3(3):126.

[19] Cigsar, C. and Lawless, J. F. (2012). Assessing transient carryover effects in recurrent event processes, with application to chronic health conditions. *The Annals of Applied Statistics*, 6(4):1641–1663.

[20] Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, 65(1):141–151.

[21] Cook, R. J. and Lawless, J. F. (1997). Marginal analysis of recurrent events and a terminating event. *Statistics in medicine*, 16(8):911–924.

[22] Cook, R. J. and Lawless, J. F. (2007). *The Statistical Analysis of Recurrent Events*. Springer, New York.

[23] Cook, R. J. and Lawless, J. F. (2013). Concepts and tests for trend in recurrent event processes. *Journal of The Iranian Statistical Society*, 12(1):35–69.

[24] Cook, R. J. and Lawless, J. F. (2018). *Multistate models for the analysis of life history data*. Chapman and Hall/CRC.

[25] Cox, D. R. (1972a). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2):187–202.

[26] Cox, D. R. (1972b). The statistical analysis of dependencies in point processes. *Stochastic Point Processes. Wiley: New York*, pages 55–66.

[27] Cox, D. R. and Hinkley, D. V. (1974). *Theoretical statistics*. Chapman and Hall/CRC, New York.

[28] Cox, D. R. and Isham, V. (1980). *Point Processes*. Chapman and Hall, London.

[29] Cox, D. R. and Lewis, P. A. W. (1966). *The Statistical Analysis of Series of Events*. Chapman and Hall, London.

[30] Cox, D. R. and Miller, H. D. (1965). *The theory of stochastic processes*. CRC press.

[31] Cox, D. R. and Oakes, D. (1984). *Analysis of Survival Data*, volume 362 of *Monographs on Statistics and Applied Probability*. Chapman & Hall/CRC.

[32] Daley, D. J. and Vere-Jones, D. (2003). *An introduction to the theory of point processes, volume 1: Elementary theory and methods*.

[33] Daley, D. J. and Vere-Jones, D. (2007). *An introduction to the theory of point processes: volume II: general theory and structure*. Springer Science & Business Media.

[34] Dißmann, J., Brechmann, E., Czado, C., and Kurowicka, D. (2013). Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics & Data Analysis*, 59:52 – 69.

[35] Duchateau, L. and Janssen, P. (2007). *The frailty model*. Springer Science & Business Media.

[36] Duchateau, L., Janssen, P., Kezic, I., and Fortpied, C. (2003). Evolution of recurrent asthma event rate over time in frailty models. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 52(3):355–363.

[37] Embrechts, P., Lindskog, F., and Mcneil, A. (2003). Chapter 8 - Modelling Dependence with Copulas and Applications to Risk Management. In Rachev, S. T., editor, *Handbook of Heavy Tailed Distributions in Finance*, volume 1 of *Handbooks in Finance*, pages 329 – 384. North-Holland, Amsterdam.

[38] Farrington, C. P. and Whitaker, H. J. (2006). Semiparametric analysis of case series data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 55(5):553–594.

[39] Feller, W. (1968). *An introduction to probability theory and its applications*, volume 1. John Wiley & Sons, New York, 3 edition.

[40] Feller, W. (1982). *An introduction to probability theory and its applications*, volume 2. John Wiley & Sons, New York, 2 edition.

[41] Frank, M. J. (1979). On the simultaneous associativity of $F(x, y)$ and $x + y - F(x, y)$. *Aequationes mathematicae*, 19:194–226.

[42] Gail, M. H., Santner, T. J., and Brown, C. C. (1980). An analysis of comparative carcinogenesis experiments based on multiple times to tumor. *Biometrics*, 36(2):255–266.

[43] Genest, C. and MacKay, J. (1986). The joy of copulas: Bivariate distributions with uniform marginals. *The American Statistician*, 40(4):280–283.

[44] Georges, P., Lamy, A., Nicolas, E., Quibel, G., and Roncalli, T. (2001). Multivariate survival modelling: A unified approach with copulas. *Working paper, Crédit Lyonnais*.

[45] Ghosh, D. (2004). Accelerated rates regression models for recurrent failure time data. *Lifetime Data Analysis*, 10(3):247–261.

[46] Goethals, K., Janssen, P., and Duchateau, L. (2008). Frailty models and copulas: similarities and differences. *Journal of Applied Statistics*, 35(9):1071–1079.

[47] Golzy, M. and Carter, R. L. (2019). Generalized frailty models for analysis of recurrent events. *Journal of Statistical Planning and Inference*, 200:213–222.

[48] Gruneir, A., Cigsar, C., Wang, X., Newman, A., Bronskill, S. E., Anderson, G. M., and Rochon, P. A. (2018). Repeat emergency department visits by nursing home residents: a cohort study using health administrative data. *BMC geriatrics*, 18(1):157.

[49] Gumbel, E. J. (1960). Distributions de valeurs extrêmes en plusieurs dimensions. *Publications de lInstitut de Statistique de lUniversité de Paris*, 9:171–173.

[50] Hawkes, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90.

[51] Hobæk Haff, I., Aas, K., and Frigessi, A. (2010). On the simplified pair-copula construction - simply useful or too simplistic? *Journal of Multivariate Analysis*, 101(5):1296 – 1310.

[52] Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman & Hall/CRC, London.

[53] Joe, H. (2014). *Dependence Modeling with Copulas*. Chapman & Hall/CRC, London.

[54] Kalbfleisch, J. D. and Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data*. John Wiley and Sons, New York, 2 edition.

[55] Kim, J. I., Lin, F.-C., and Fine, J. P. (2019). On the self-triggering cox model for recurrent event data. *Statistics in Medicine*, 38(22):4240–4252.

[56] Kvist, K., Andersen, P. K., Angst, J., and Kessing, L. V. (2008). Repeated events and total time on test. *Statistics in Medicine*, 27(19):3817–3832.

[57] Lawless, J. (1995). The analysis of recurrent events for multiple subjects. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 44(4):487–498.

[58] Lawless, J. and Nadeau, J. (1995). Nonparametric estimation of cumulative mean functions for recurrent events. *Technometrics*, 37:158–168.

[59] Lawless, J. F. (2003). *Statistical models and methods for lifetime data*. John Wiley & Sons, 2nd edition.

[60] Lawless, J. F., Cigsar, C., and Cook, R. J. (2012). Testing for monotone trend in recurrent event processes. *Technometrics*, 54(2):147–158.

[61] Lawless, J. F., Kalbfleisch, J. D., and Blumenthal, S. (1992). *Some Issues in the Collection and Analysis of Field Reliability Data*, pages 141–152. Springer Netherlands, Dordrecht.

[62] Lawless, J. F. and Yilmaz, Y. E. (2011). Semiparametric estimation in copula models for bivariate sequential survival times. *Biometrical Journal*, 53(5):779–796.

[63] Lin, D., Wei, L., and Ying, Z. (1998). Accelerated failure time models for counting processes. *Biometrika*, 85(3):605–618.

[64] Lin, D. Y., Sun, W., and Ying, Z. (1999). Nonparametric estimation of the gap time distributions for serial events with censored data. *Biometrika*, 86(1):59–70.

[65] Lin, D. Y., Wei, L.-J., Yang, I., and Ying, Z. (2000). Semiparametric regression for the mean and rate functions of recurrent events. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62(4):711–730.

[66] Lindsey, J. K. (2004). *Statistical analysis of stochastic processes in time*, volume 14. Cambridge University Press.

[67] Nelsen, R. B. (2006). *An Introduction to Copulas*. Springer, New York, 2 edition.

[68] Nelson, W. (1988). Graphical analysis of system repair data. *Journal of Quality Technology*, 20(1):24–35.

[69] Nelson, W. B. (2003). *Recurrent events data analysis for product repairs, disease recurrences, and other applications*, volume 10. SIAM.

[70] Oakes, D. (1975). The markovian self-exciting process. *Journal of Applied Probability*, 12(1):69–77.

[71] Oakes, D. (1989). Bivariate survival models induced by frailties. *Journal of the American Statistical Association*, 84(406):487–493.

[72] Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical association*, 83(401):9–27.

[73] Peña, E. A. (2006). Dynamic modeling and statistical analysis of event times. *Statistical Science*, 21(4):487–500.

[74] Peña, E. A. and Hollander, M. (2004). *Models for Recurrent Events in Reliability and Survival Analysis*, pages 105–123. Springer US, Boston, MA.

[75] Pepe, M. S. and Cai, J. (1993). Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates. *Journal of the American statistical Association*, 88(423):811–820.

[76] Prentice, R. L., Williams, B. J., and Peterson, A. V. (1981). On the regression analysis of multivariate failure time data. *Biometrika*, 68(2):373–379.

[77] Rigdon, S. E. and Basu, A. P. (2000). *Statistical methods for the reliability of repairable systems*. Wiley New York.

[78] Ross, S. M. (1996). *Stochastic processes*. John Wiley & Sons, 2 edition.

[79] Self, S. G. and Liang, K.-Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*, 82(398):605–610.

[80] Selvaratnam, S., Oyet, A. J., Yi, Y., and Gadag, V. (2017). Estimation of a generalized linear mixed model for response-adaptive designs in multi-centre clinical trials. *Canadian Journal of Statistics*, 45(3):310–325.

[81] Simpson, S. E. (2013). A positive event dependence model for self-controlled case series with applications in postmarketing surveillance. *Biometrics*, 69(1):128–136.

[82] Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8:229–231.

[83] Sklar, A. (1973). Random variables, joint distribution functions, and copulas. *Kybernetika*, 9:449–460.

[84] Stute, W., Manteiga, W. G., and Quindimil, M. P. (1993). Bootstrap based goodness-of-fit-tests. *Metrika*, 40(1):243–256.

[85] Sun, L. and Su, B. (2008). A class of accelerated means regression models for recurrent event data. *Lifetime data analysis*, 14(3):357–375.

[86] Therneau, T. M. and Grambsch, P. M. (2000). *Modeling Survival Data: Extending the Cox Model*. Springer Science & Business Media.

[87] Wei, L.-J., Lin, D. Y., and Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of the American statistical association*, 84(408):1065–1073.

[88] Whitaker, H. J., Paddy, F. C., Spiessens, B., and Musonda, P. (2006). Tutorial in biostatistics: the selfcontrolled case series method. *Statistics in Medicine*, 25(10):1768–1797.

[89] White, H. and Granger, C. W. J. (2011). Consideration of trends in time series. *Journal of Time Series Econometrics*, 3(1).

[90] Xu, S., Hambidge, S. J., McClure, D. L., Daley, M. F., and Glanz, J. M. (2013). A scan statistic for identifying optimal risk windows in vaccine safety studies using selfcontrolled case series design. *Statistics in Medicine*, 32(19):3290–3299.

[91] Xu, S., Zhang, L., Nelson, J. C., Zeng, C., Mullooly, J., McClure, D., and Glanz, J. (2011). Identifying optimal risk windows for selfcontrolled case series studies of vaccine safety. *Statistics in Medicine*, 30(7):742–752.

Appendix A

Derivation of Marginal Intensity Function in Random Effects Models

In this appendix, we derive the marginal density function $\lambda_i\left[t \mid \mathcal{H}_i(t)\right]$ when it is assumed that the i.i.d. random effects ν_i in the conditional intensity function $\lambda_i\left[t \mid \mathcal{H}_i(t), \nu_i\right]$ follow a gamma distribution with mean 1 and variance ϕ.

Suppose that $\{N_i(t), t \geq 0\}$ is a continuous time counting process with the intensity function $\lambda_i\left[t \mid \mathcal{H}_i(t)\right]$, where $\mathcal{H}_i(t) = \{N_i(s), 0 \leq s < t\}$, $i = 1, \ldots, m$, is the history of the counting process. Let $\Delta N_i(t)$ denote the number of events in $[t, t + \Delta t)$ for the counting process $\{N_i(t), t \geq 0\}$. Then, as $\Delta t \to 0$,

$$\begin{aligned}
\lambda_i\left[t \mid \mathcal{H}_i(t)\right]\ \Delta t = \Pr\left\{\Delta N_i(t) = 1 \mid \mathcal{H}_i(t)\right\} &= \frac{\Pr\left\{\Delta N_i(t) = 1, \mathcal{H}_i(t)\right\}}{\Pr\left\{\mathcal{H}_i(t)\right\}}, \\
&= \frac{\int_0^\infty \Pr\left\{\Delta N_i(t) = 1 \mid \mathcal{H}_i(t), \nu_i\right\} \Pr\left\{\mathcal{H}_i(t)|\nu_i\right\} g(\nu_i)\, d\nu_i}{\int_0^\infty \Pr\left\{\mathcal{H}_i(t)|\nu_i\right\} g(\nu_i)\, d\nu_i},
\end{aligned} \tag{A.1}$$

where $g(\nu_i)$ is the gamma density function of the random effect ν_i. Note that any probability notation with $\mathcal{H}_i(t)$ denotes the probability of observing a given realization of $\{N_i(t), t \geq 0\}$ over $[0, t)$. Thus, if we let $N_i(t^-) = n_i$, we observe n_i events at times $t_{i1} < \ldots < t_{in_i}$ over $[0, t)$.

Let $\lambda_i\left[t \mid \mathcal{H}_i(t), \nu_i\right] = \nu_i \eta_i(t)$, where $\eta_i(t) = \exp[\boldsymbol{\psi}' \boldsymbol{W}_i^*(t)]$, $\boldsymbol{W}^*(t)$ is a $q \times 1$ vector of processes that is allowed to contain functions of the event history $\mathcal{H}_i(t)$ as well as external covariates, and $\boldsymbol{\psi}$ is a $q \times 1$ vector of parameters. The integrand in the denominator of (A.1) can be written as

$$\Pr\{\mathcal{H}_i(t)|\nu_i\}\, g(\nu_i) = \left\{\prod_{j=1}^{N_i(t^-)} \nu_i \eta_i(t_{ij}) \;\exp\left(-\nu_i \int_0^t \eta_i(u)du\right)\right\} \frac{\nu_i^{\phi^{-1}-1}\exp(-\nu_i/\phi)}{\phi^{\phi^{-1}}\Gamma(\phi^{-1})},$$

$$= \left\{\nu_i^{[N_i(t^-)+\phi^{-1}-1]} \;\exp\left(-\nu_i\left[\int_0^t \eta_i(u)du + \frac{1}{\phi}\right]\right)\right\} \frac{\prod_{j=1}^{N_i(t^-)} \eta_i(t_{ij})}{\phi^{\phi^{-1}}\Gamma(\phi^{-1})}. \tag{A.2}$$

By taking the integral of (A.2) with respect to ν_i, we find that the denominator of (A.1) can be written as

$$\int_0^\infty \Pr\{\mathcal{H}_i(t)|\nu_i\}\, g(\nu_i)d\nu_i = \frac{\prod_{j=1}^{N_i(t^-)} \eta_i(t_{ij})}{\phi^{\phi^{-1}}\Gamma(\phi^{-1})} \times \frac{\Gamma(N_i(t^-)+\phi^{-1})}{\left[\int_0^t \eta_i(u)du + \frac{1}{\phi}\right]^{(N_i(t^-)+\phi^{-1})}}. \tag{A.3}$$

Similarly, from (A.2), the numerator of (A.1) is equal to

$$\int_0^\infty \nu_i\eta_i(t)\Delta t \Pr\{\mathcal{H}_i(t)|\nu_i\}\, g(\nu_i)d\nu_i = \frac{\eta_i(t)\Delta t \prod_{j=1}^{N_i(t^-)} \eta_i(t_{ij})}{\phi^{\phi^{-1}}\Gamma(\phi^{-1})} \times \frac{\Gamma(N_i(t^-)+\phi^{-1}+1)}{\left[\int_0^t \eta_i(u)du + \frac{1}{\phi}\right]^{(N_i(t^-)+\phi^{-1}+1)}}. \tag{A.4}$$

From the results given in (A.3) and (A.4) the marginal intensity function for the process $\{N_i(t), t \geq 0\}$, $i = 1, \ldots, m$, can be written as

$$\lambda_i\left[t \mid \mathcal{H}_i(t)\right] = \frac{(1+\phi N_i(t^-))}{\left[1+\phi \int_0^t \eta_i(u)du\right]}\eta_i(t), \qquad t \geq 0. \tag{A.5}$$

Appendix B

Explosion of Dynamic Recurrent Event Processes

In this appendix, we present details and results of a simulation study conducted to investigate the issue of dishonest process $\{N(t); t \geq 0\}$ with the intensity function (3.1). For the data generation, we considered the intensity function (3.1) with $\lambda_0(t)$ fixed at a constant value α ($\alpha = 1$) and without external covariates $\boldsymbol{x}(t)$. We generated data under 36 scenarios with various values of γ, β, Δ and τ as denoted in Table B.1. With every combination of $(\gamma, \beta, \Delta, \tau)$, we simulated 10,000 realizations under four different models. In the first model, we used the aforementioned dishonest process with the intensity function $\lambda[t|\mathcal{H}(t)] = \exp[\gamma\, N(t^-) + \beta\, Z(t)]$. We replaced the dynamic covariate $N(t^-)$ with its trimmed version $N^*(t^-)$ given in (3.5) with cutoff points $c = 100, 50$ and 20 in the second, third and fourth models, respectively, so that we were able to compare the results in an empirical setting. Note that when $c = \infty$ in the model (3.5), $N^*(t^-) = N(t^-)$. In this empirical setting, once the process reached 1,000 events it was considered as an event explosion, and simulation algorithm stopped generating further events. Number of explosions and the maximum number of events in 10,000 realizations for each scenario are reported in Table B.1.

Table B.1: Number of exploded processes and the maximum number of events in 10,000 realization.

γ	β	Δ	τ	Model 1; $c = \infty$ Exploded	Model 1; $c = \infty$ Maximum	Model 2; $c = 100$ Exploded	Model 2; $c = 100$ Maximum	Model 3; $c = 50$ Exploded	Model 3; $c = 50$ Maximum	Model 4; $c = 20$ Exploded	Model 4; $c = 20$ Maximum
0.000	0.693	0.05	1	0	8	0	7	0	7	0	7
			2	0	10	0	9	0	10	0	11
			5	0	16	0	19	0	16	0	18
		0.10	1	0	9	0	8	0	8	0	8
			2	0	11	0	11	0	12	0	11
			5	0	16	0	19	0	20	0	20
	1.609	0.05	1	0	11	0	9	0	9	0	10
			2	0	14	0	14	0	14	0	13
			5	0	24	0	21	0	20	0	19
		0.10	1	0	14	0	12	0	12	0	12
			2	0	19	0	22	0	18	0	17
			5	0	27	0	26	0	29	0	27
0.095	0.693	0.05	1	0	9	0	9	0	8	0	11
			2	0	15	0	16	0	16	0	14
			5	95	1,000	91	1,000	0	395	0	46
		0.10	1	0	10	0	11	0	11	0	10
			2	0	29	0	25	0	17	0	21
			5	279	1,000	265	1,000	0	564	0	62
	1.609	0.05	1	0	19	0	22	0	20	0	25
			2	16	1,000	14	1,000	0	354	0	41
			5	2,063	1,000	2,028	1,000	368	1,000	0	121
		0.10	1	3	1,000	3	1,000	0	153	0	35
			2	203	1,000	218	1,000	0	684	0	66
			5	5,118	1,000	5,126	1,000	1,906	1,000	0	183
0.223	0.693	0.05	1	8	1,000	6	1,000	3	1,000	0	73
			2	407	1,000	400	1,000	393	1,000	0	248
			5	6,970	1,000	6,938	1,000	6,982	1,000	0	799
		0.10	1	31	1,000	34	1,000	16	1,000	0	73
			2	730	1,000	722	1,000	725	1,000	0	315
			5	7,636	1,000	7,724	1,000	7,669	1,000	0	773
	1.609	0.05	1	344	1,000	362	1,000	342	1,000	0	320
			2	2,334	1,000	2,424	1,000	2,433	1,000	0	806
			5	8,724	1,000	8,724	1,000	8,770	1,000	4,501	1,000
		0.10	1	1,344	1,000	1,369	1,000	1,408	1,000	0	366
			2	4,655	1,000	4,632	1,000	4,538	1,000	0	790
			5	9,403	1,000	9,425	1,000	9,379	1,000	6,612	1,000

It is clear that when $\gamma = 0$, processes in the first 12 scenarios in Table B.1 do not explode even when the follow-up time τ increases. The processes with $\gamma > 0$ start to explode when the follow-up time τ increases. It is also noted that while $\gamma > 0$, an increment in carryover effects β and/or carryover effects period Δ also results in an increase in the number of explosions. When $\gamma > 0$, the number of explosions reduces while the cutoff value c decreases for a given $\gamma, \beta, \Delta, \tau$ combination. Some other suggestions, by modifying $N(t^-)$ with time t, to handle such explosions due to $N(t^-)$ are available in Aalen et al., (2008, Section 8.6.3). The coefficients of those modified $N(t^-)$ may not be easily interpretable. Therefore, for our purpose, $N^*(t^-)$ given in (3.5) is the best option to handle dishonest processes.

Appendix C

Regularity Conditions

In this appendix, we state the regularity conditions discussed in Section 3.1.1. This appendix is excerpt directly taken from Andersen et al. (1993, pp. 420–421). We only present it here for the completeness of the thesis.

Consider a counting process $\{N(t); t \geq 0\}$ with intensity process $\lambda^*[t|\mathcal{H}(t);\boldsymbol{\theta}]$ specified by a q-dimensional parameter $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_q) \in \boldsymbol{\Theta}$, an open subset of the q-dimensional Euclidean space. The true value of the vector of parameter $\boldsymbol{\theta}$ is denoted by $\boldsymbol{\theta}_0 = (\theta_{10}, \ldots, \theta_{q0})$. Let the sequence of counting processes $\boldsymbol{N^{(m)}} = (N_1, \ldots, N_m)$, have the associated intensity processes $\boldsymbol{\lambda^{(m)}} = (\lambda_1, \ldots, \lambda_m)$ of the parametric form $\lambda_h(t;\boldsymbol{\theta}_0) = Y_h(t)\lambda_h^*[t|\mathcal{H}(t);\boldsymbol{\theta}_0]$, $h = 1, 2, \ldots, m$, where $Y_h(t)$ is the at-risk indicator function of the h^{th} process and $\boldsymbol{\theta}_0$ is assumed to be the same for all m.

(A) There exists a neighborhood $\boldsymbol{\Theta}_0$ of $\boldsymbol{\theta}_0$ such that for all m, h and $\boldsymbol{\theta}_0 \in \boldsymbol{\Theta}_0$, and almost all $t \in \mathcal{T}$, the partial derivatives of $\lambda_h(t;\boldsymbol{\theta})$ and $\log \lambda_h(t;\boldsymbol{\theta})$ of the first, second, and third order with respect to $\boldsymbol{\theta}$ exist and are continuous in $\boldsymbol{\theta}$ for $\boldsymbol{\theta} \in \boldsymbol{\Theta}_0$. Moreover, the log-likelihood function

$$\ell(\boldsymbol{\theta}) = \int_0^\infty \sum_{h=1}^m \log \lambda_h(t;\boldsymbol{\theta}) dN_h(t) - \int_0^\infty \sum_{h=1}^m \lambda_h(t;\boldsymbol{\theta}) dt$$

may be differentiated three times with respect to $\boldsymbol{\theta} \in \boldsymbol{\Theta}_0$ by interchanging the order of integration and differentiation.

(B) There exist a sequence $\{a_m\}$ of non-negative constants increasing to infinity as

$m \to \infty$ and finite functions $\sigma_{jl}(\boldsymbol{\theta})$ defined on $\boldsymbol{\Theta}_0$ such that for all j, l

$$a_m^{-2} \int_0^\infty \sum_{h=1}^m \left\{ \frac{\partial}{\partial \theta_j} \log \lambda_h(t; \boldsymbol{\theta}_0) \right\} \left\{ \frac{\partial}{\partial \theta_l} \log \lambda_h(t; \boldsymbol{\theta}_0) \right\} \lambda_h(t; \boldsymbol{\theta}_0) dt \xrightarrow{\boldsymbol{P}} \sigma_{jl}(\boldsymbol{\theta})$$

as $m \to \infty$.

(C) For all h, j and all $\epsilon > 0$, we have that

$$a_m^{-2} \int_0^\infty \sum_{h=1}^m \left\{ \frac{\partial}{\partial \theta_j} \log \lambda_h(t; \boldsymbol{\theta}_0) \right\}^2 I \left(\left| a_m^{-1} \frac{\partial}{\partial \theta_j} \log \lambda_h(t; \boldsymbol{\theta}_0) \right| > \epsilon \right) \lambda_h(t; \boldsymbol{\theta}_0) dt \xrightarrow{\boldsymbol{P}} 0$$

as $n \to \infty$.

(D) The matrix $\boldsymbol{\Sigma} = \{\sigma_{jl}(\boldsymbol{\theta}_0)\}$ with $\sigma_{jl}(\boldsymbol{\theta}_0)$ defined in Condition B is positive definite.

(E) For any m and each $h = 1, 2, ..., k$ there exist predictable processes G_{hm} and H_{hm} not depending on $\boldsymbol{\theta}$ such that for all $t \in \mathcal{T}$

$$\sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_0} \left| \frac{\partial^3}{\partial \theta_j \partial \theta_l \partial \theta_r} \lambda_h(t, \boldsymbol{\theta}) \right| \le G_{hm}(t),$$

and

$$\sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_0} \left| \frac{\partial^3}{\partial \theta_j \partial \theta_l \partial \theta_r} \log \lambda_h(t, \boldsymbol{\theta}) \right| \le H_{hm}(t),$$

for all j, l, r. Moreover

$$a_m^{-2} \int_0^\infty \sum_{h=1}^m G_{hm}(t) dt, \qquad a_m^{-2} \int_0^\infty \sum_{h=1}^m H_{hm}(t) \lambda_h(t; \boldsymbol{\theta}_0) dt$$

as well as (for all j, l)

$$a_m^{-2} \int_0^\infty \sum_{h=1}^m \left\{ \frac{\partial^2}{\partial \theta_j \partial \theta_l} \log \lambda_h(t; \boldsymbol{\theta}_0) \right\}^2 \lambda_h(t; \boldsymbol{\theta}_0) dt$$

all converge in probability to finite quantities as $m \to \infty$, and, for all $\epsilon > 0$,

$$a_m^{-2} \int_0^\infty \sum_{h=1}^m H_{hm}(t) I \left(a_m^{-1} H_{hm}(t)^{1/2} > \epsilon \right) \lambda_h(t; \boldsymbol{\theta}_0) dt \xrightarrow{\boldsymbol{P}} 0.$$

Appendix D

Bootstrap Algorithm to Calculate $p-$value of Goodness-of-fit Test

We use following steps to calculate the bootstrap $p-$values for goodness-of-fit tests, which are introduced in Section 5.2.4.

Step 1 : Generate $\{U_{i1}; i = 1, \ldots, m\}$ from standard uniform distribution.

Step 2 : Generate $\{U_{i2}^*; i = 1, \ldots, m\}$ from standard uniform distribution and calculate $\{U_{i2}; i = 1, \ldots, m\}$ by plugging in the value U_{i2}^* into the function $c_{i1}^{-1}(\cdot)$, where $c_{i1}(U_{i2}) = \frac{\partial C(U_1, U_{i2}; \widetilde{\boldsymbol{\alpha}^*})}{\partial U_1}\Big|_{U_1 = U_{i1}}$, for $i = 1, \ldots, m$, and $C(\cdot, \cdot; \widetilde{\boldsymbol{\alpha}^*})$ is the assumed copula model under null hypothesis and $\widetilde{\boldsymbol{\alpha}^*} = (\boldsymbol{\alpha}_{10}^*, \hat{\boldsymbol{\alpha}_2^*}(\boldsymbol{\alpha}_{10}^*))$. We denote $\boldsymbol{\alpha}_{10}^*$ as the vector of copula parameters under null hypothesis and $\hat{\boldsymbol{\alpha}_2^*}$ as the vector of second set of copula parameters which is estimated from data.

Step 3 : Obtain $W_{i1}^* = \hat{F}_1^{-1}(U_{i1})$ and $W_{i2}^* = \hat{F}_2^{-1}(U_{i2})$.

Step 4 : From the Kaplan-Meier estimate based on the data $\{(\min(W_{i1} + W_{i2}, C_i), 1 - \delta_{i2}), i = 1, \ldots, m\}$, generate censoring times $\{C_i^*; i = 1, \ldots, m\}$.

Step 5 : Compute $w_{i1}^* = \min(W_{i1}^*, C_i^*)$, $w_{i2}^* = \min(W_{i2}^*, C_i^*)$, $\delta_{i1}^* = I(W_{i1}^* = w_{i1}^*)$ and $\delta_{i2}^* = I(W_{i2}^* = w_{i2}^*)$.

Step 6 : Obtain the test statistic Λ^* for the bootstrap sample $\{(w_{i1}^*, w_{i2}^*, \delta_{i1}^*, \delta_{i2}^*), i = 1, \ldots, m\}$ according to the given estimation procedure.

Step 7 : Steps 1 to 6 are repeated B times and the $p-$value is estimated as the proportion of times that $\Lambda^* \geq \Lambda^{obs}$, where Λ^{obs} is the observed value of the test statistics with the original sample.

Appendix E

Simulation Results

Table E.1: Table 4.9 continued.

τ	Parameter	DDP			SDP			IND		
		m = 50	**m = 100**	**m = 250**	**m = 50**	**m = 100**	**m = 250**	**m = 50**	**m = 100**	**m = 250**
5	$\bar{\hat{\alpha}}$	0.766524 (0.108115) [0.092860]	0.761530 (0.076525) [0.065319]	0.756559 (0.047761) [0.041046]	0.772814 (0.114899) [0.099818]	0.767190 (0.081221) [0.070207]	0.761467 (0.050934) [0.044126]	0.773402 (0.112847) [0.089090]	0.769724 (0.079807) [0.062678]	0.764681 (0.049301) [0.039388]
	$\bar{\hat{\phi}}$									
	$\bar{\hat{\gamma}}$	0.253810 (0.052973) [0.047723]	0.252671 (0.034988) [0.033692]	0.252706 (0.022098) [0.021298]	0.261812 (0.056482) [0.053499]	0.260370 (0.038847) [0.037698]	0.260573 (0.024560) [0.023836]	0.319079 (0.062898) [0.053779]	0.316764 (0.044073) [0.037915]	0.316576 (0.028240) [0.023959]
	$\bar{\hat{\beta}}$	1.628480 (0.172938) [0.163924]	1.632454 (0.119920) [0.114939]	1.636216 (0.073887) [0.072437]	1.552996 (0.185656) [0.151522]	1.558489 (0.129294) [0.106347]	1.563107 (0.079898) [0.067083]	1.732822 (0.246961) [0.165165]	1.748336 (0.171412) [0.116047]	1.759707 (0.107711) [0.073153]
	$\bar{\hat{\phi}}_{c1}$	1.433288 (0.335946) [0.322937]	1.411751 (0.226792) [0.224157]	1.393854 (0.137519) [0.140050]						
	$\bar{\hat{\phi}}_{c2}$	2.064704 (0.447509) [0.428529]	2.038810 (0.306447) [0.298536]	2.014926 (0.187461) [0.186567]						
	$\bar{\hat{\phi}}_{c3}$	3.024794 (0.609829) [0.572276]	2.972593 (0.410777) [0.396560]	2.951333 (0.249327) [0.248478]						
	$\bar{\hat{\phi}}_{c4}$	4.611030 (0.858628) [0.814574]	4.529512 (0.573166) [0.562437]	4.492233 (0.356109) [0.351851]						
	$\bar{\hat{\phi}}_{c}$				2.263331 (0.283761) [0.246669]	2.252328 (0.194502) [0.172377]	2.241646 (0.117205) [0.108211]			
2	$\hat{\alpha}$	0.865173 (0.142063) [0.126720]	0.861891 (0.098776) [0.089487]	0.857989 (0.061187) [0.056467]	0.865732 (0.145190) [0.130759]	0.861253 (0.100805) [0.092173]	0.856557 (0.062661) [0.058095]	0.878929 (0.136679) [0.114858]	0.876945 (0.093838) [0.080990]	0.873888 (0.058548) [0.051078]
	$\bar{\hat{\phi}}$									
	$\bar{\hat{\gamma}}$	0.251832 (0.067807) [0.062953]	0.252861 (0.047439) [0.044585]	0.252970 (0.028524) [0.028176]	0.275641 (0.071737) [0.076002]	0.276789 (0.049489) [0.053447]	0.277271 (0.031033) [0.033659]	0.424235 (0.074516) [0.070263]	0.420823 (0.051606) [0.049212]	0.418549 (0.033056) [0.030994]
	$\bar{\hat{\beta}}$	1.620165 (0.199771) [0.187760]	1.619229 (0.135614) [0.131739]	1.621338 (0.083999) [0.082911]	1.555182 (0.208895) [0.174677]	1.556224 (0.142781) [0.122666]	1.558706 (0.089681) [0.077176]	1.729927 (0.282201) [0.207799]	1.736729 (0.195083) [0.145324]	1.750290 (0.120312) [0.091394]
	$\bar{\hat{\phi}}_{c1}$	1.416096 (0.368686) [0.351680]	1.381588 (0.246260) [0.243178]	1.364130 (0.150476) [0.151804]						
	$\bar{\hat{\phi}}_{c2}$	2.072396 (0.506729) [0.484846]	2.031061 (0.337134) [0.335087]	2.009863 (0.210580) [0.209170]						
	$\bar{\hat{\phi}}_{c3}$	3.050009 (0.718185) [0.667385]	2.975801 (0.475090) [0.458018]	2.947966 (0.282158) [0.285609]						
	$\bar{\hat{\phi}}_{c4}$	4.742540 (1.075324) [0.991397]	4.595521 (0.715245) [0.673768]	4.527523 (0.439167) [0.417564]						
	$\bar{\hat{\phi}}_{c}$				2.191437 (0.310299) [0.273024]	2.171123 (0.212406) [0.190448]	2.163040 (0.130763) [0.119350]			

Table E.2: Table 4.10 continued.

τ	Parameter	DDP			SDP			IND		
		m = 50	**m = 100**	**m = 250**	**m = 50**	**m = 100**	**m = 250**	**m = 50**	**m = 100**	**m = 250**
5	$\hat{\alpha}$	0.797294	0.777530	0.768584	0.823132	0.801033	0.792169	0.866317	0.848265	0.840443
		(0.202746)	(0.135535)	(0.085828)	(0.216007)	(0.143864)	(0.091044)	(0.196919)	(0.130878)	(0.083795)
		[0.151455]	[0.103670]	[0.064430]	[0.160817]	[0.109901]	[0.068371]	[0.127718]	[0.087607]	[0.054571]
	$\bar{\hat{\phi}}$									
	$\bar{\hat{\gamma}}$	0.228785	0.230406	0.231220	0.220116	0.222052	0.222529	0.237667	0.234281	0.231902
		(0.048155)	(0.035812)	(0.030022)	(0.050880)	(0.037016)	(0.023234)	(0.062206)	(0.044570)	(0.027808)
		[0.043228]	[0.030471]	[0.019237]	[0.049127]	[0.034615]	[0.021840]	[0.053198]	[0.037410]	[0.023574]
	$\bar{\hat{\beta}}$	1.661556	1.670755	1.676655	1.628881	1.635319	1.641358	1.841433	1.875344	1.893743
		(0.145988)	(0.100738)	(0.066746)	(0.149401)	(0.105840)	(0.068406)	(0.216196)	(0.154294)	(0.096742)
		[0.139511]	[0.098381]	[0.062049]	[0.126692]	[0.089351]	[0.056389]	[0.153589]	[0.107708]	[0.067747]
	$\bar{\hat{\xi}}_1$	0.439306	0.421204	0.415707	0.437014	0.418556	0.412342	0.405559	0.392309	0.389156
		(0.311991)	(0.213357)	(0.131608)	(0.317065)	(0.217393)	(0.133993)	(0.256200)	(0.175617)	(0.109307)
		[0.202203]	[0.139778]	[0.087233]	[0.205133]	[0.141660]	[0.088411]	[0.136720]	[0.094854]	[0.059348]
	$\bar{\hat{\xi}}_2$	0.850691	0.835170	0.818941	0.842124	0.827080	0.808947	0.792789	0.784531	0.773147
		(0.301957)	(0.210512)	(0.130891)	(0.307271)	(0.214208)	(0.133045)	(0.252935)	(0.178368)	(0.111166)
		[0.203551]	[0.140606]	[0.087812]	[0.206437]	[0.142420]	[0.088922]	[0.138902]	[0.096411]	[0.060315]
	$\bar{\hat{\phi}}_{c1}$	1.456949	1.426108	1.410890						
		(0.326006)	(0.221832)	(0.138180)						
		[0.313435]	[0.216471]	[0.135083]						
	$\bar{\hat{\phi}}_{c2}$	2.096348	2.065368	2.042895						
		(0.412597)	(0.291688)	(0.179977)						
		[0.400031]	[0.277631]	[0.173419]						
	$\bar{\hat{\phi}}_{c3}$	3.026396	2.986202	2.952597						
		(0.528662)	(0.360264)	(0.236917)						
		[0.519827]	[0.361047]	[0.225377]						
	$\bar{\hat{\phi}}_{c4}$	4.573162	4.508235	4.459524						
		(0.755398)	(0.519093)	(0.341714)						
		[0.727055]	[0.503293]	[0.313820]						
	$\bar{\hat{\phi}}_c$				2.361555	2.340167	2.327183			
					(0.246748)	(0.170715)	(0.107077)			
					[0.225219]	[0.156306]	[0.097800]			
1	$\hat{\alpha}$	0.937958	0.930842	0.926231	0.952784	0.943226	0.937400	1.207596	1.205703	1.201127
		(0.297249)	(0.196154)	(0.122391)	(0.305443)	(0.202684)	(0.126055)	(0.315396)	(0.210328)	(0.128674)
		[0.257181]	[0.180519]	[0.113297]	[0.265791]	[0.185653]	[0.116348]	[0.225145]	[0.157386]	[0.098608]
	$\bar{\hat{\phi}}$									
	$\bar{\hat{\gamma}}$	0.245864	0.244454	0.245041	0.247558	0.249923	0.250689	0.341514	0.336377	0.333246
		(0.074904)	(0.041232)	(0.027339)	(0.064087)	(0.044470)	(0.028898)	(0.075344)	(0.050472)	(0.031259)
		[0.056518]	[0.039894]	[0.025197]	[0.068077]	[0.047678]	[0.030037]	[0.074038]	[0.051580]	[0.032388]
	$\bar{\hat{\beta}}$	1.636317	1.639947	1.640113	1.599821	1.600288	1.601079	1.819124	1.834027	1.842796
		(0.179541)	(0.122295)	(0.076629)	(0.177576)	(0.122997)	(0.077079)	(0.251761)	(0.174703)	(0.107617)
		[0.170354]	[0.120132]	[0.075884]	[0.153392]	[0.107884]	[0.068124]	[0.219331]	[0.152897]	[0.095992]
	$\bar{\hat{\xi}}_1$	0.447882	0.433607	0.427004	0.444877	0.429600	0.422870	0.311890	0.298199	0.294031
		(0.316825)	(0.212596)	(0.130643)	(0.321575)	(0.215390)	(0.132681)	(0.272196)	(0.180656)	(0.110035)
		[0.270557]	[0.187730]	[0.117197]	[0.276462]	[0.191380]	[0.119424]	[0.169911]	[0.117283]	[0.073272]
	$\bar{\hat{\xi}}_2$	0.900230	0.883786	0.876240	0.896114	0.878384	0.871224	0.587467	0.572492	0.566612
		(0.324548)	(0.216547)	(0.136885)	(0.329307)	(0.220550)	(0.140723)	(0.271569)	(0.180986)	(0.112552)
		[0.286269]	[0.199963]	[0.125346]	[0.293232]	[0.204664]	[0.128404]	[0.177308]	[0.122580]	[0.076566]
	$\bar{\hat{\phi}}_{c1}$	1.421767	1.391892	1.372936						
		(0.357832)	(0.246312)	(0.147385)						
		[0.350409]	[0.242011]	[0.150982]						
	$\bar{\hat{\phi}}_{c2}$	2.059850	2.014979	1.995279						
		(0.496211)	(0.317018)	(0.199251)						
		[0.461451]	[0.317650]	[0.198335]						
	$\bar{\hat{\phi}}_{c3}$	3.040611	2.987921	2.943371						
		(0.667062)	(0.449053)	(0.273538)						
		[0.627335]	[0.431591]	[0.268247]						
	$\bar{\hat{\phi}}_{c4}$	4.692080	4.578988	4.515403						
		(1.009182)	(0.657609)	(0.397317)						
		[0.912366]	[0.623064]	[0.386734]						
	$\bar{\hat{\phi}}_c$				2.260792	2.241922	2.229871			
					(0.282315)	(0.195006)	(0.121473)			
					[0.254695]	[0.176763]	[0.110707]			