8-9-2014

Non- and Semi-parametric Bayesian Inference with Recurrent Events and Coherent Systems Data

A. K. M. Fazlur Rahman
University of South Carolina - Columbia

Non- and Semi-parametric Bayesian Inference with Recurrent Events and Coherent Systems Data

by

A. K. M. Fazlur Rahman

Bachelor of Science
Jahangirnagar University 1997

Master of Science
Jahangirnagar University 1999

Master of Arts
Ball State University 2009

Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy in

Statistics

College of Arts and Sciences

University of South Carolina

2014

Accepted by:

Edsel A. Peña, Major Professor

John M. Grego, Committee Member

Timothy E. Hanson, Committee Member

Lianming Wang, Committee Member

Alexander C. McLain, Committee Member

Lacy Ford, Vice Provost and Dean of Graduate Studies

Acknowledgments

I would like to express my deepest heartfelt gratitude to my advisor, Dr. Edsel Peña, for his patient guidance, inspiration, caring, and excellent research support during my dissertation research. Dr. Peña, you have set an example of excellence as a researcher, instructor, and mentor in the spirit of adventure.

I acknowledge support from National Science Foundation (NSF) Grant DMS 1106435, National Institutes of Health (NIH) R01 CA154731 and P20 RR17698.

I would like to thank all the committee members, Dr. John M. Grego, Dr. Timothy E. Hanson, Dr. Lianming Wang, and Dr. Alexander C. McLain, for their careful review of my dissertation, insightful comments, and inspiring discussions. Definitely I learned a lot from all of you.

I would like to thank Dr. James Lynch for his guidance, inspiration, and insightful comments on many potential research problems during one-to-one and research group meetings on Fridays.

I am grateful to fellow research group members Bereket Kindo, Lillian Wanda, and Piaomu Liu, for many discussions in weekly meetings on Fridays. I would like to thank the Department of Statistics, the College of Arts and Sciences, and the Graduate School at USC for their financial support to attend many statistical conferences and workshops which are very important for developing this dissertation. I want to express my gratitude to all the faculty and staff in the Department of Statistics for their continuous assistance and encouragement. I am thankful to have all of you in this endeavor.

I would like to express special thanks to my wife, Kazi Runa, my son, Farhan

Ishraque Rahman, and daughter, Farheen Shumaila Rahman for their endless love and support.

Abstract

This dissertation deals with non- and semi-parametric Bayesian inference of gap-time distribution with recurrent event data and simultaneous inference of component and system reliabilities of coherent systems data. Recurrent event data arise from a wide variety of studies/fields such as clinical trials, epidemiology, public health, biomedicine (e.g. repeated heart attack, repeated tumor occurrences of a cancer patient). In Chapter 2 we develop nonparametric Bayes and empirical Bayes estimators of the survivor function $\bar{F} = 1 - F$, of the gap-time distribution by assigning a Dirichlet process prior on F. We develop a closed form estimator of $\bar{F}$ as well as a procedure to sample from the posterior measure and thus construct point-wise credible intervals. Semiparametric Bayesian inference of the gap-time survivor function with the effect of covariates of a correlated recurrent event in the presence of censoring is considered in Chapter 3. A frailty model is considered to allow the association between inter-occurrence gap-times. We assign a gamma process prior on the baseline cumulative hazard function Λ_0 and parametric prior distributions on the finite dimensional parameters associated with covariates and frailties. We derive the conditional posterior distributions of the unknown parameters of interest from the joint posterior distribution and employ Gibbs sampling techniques to obtain samples from the joint posterior distribution. In Chapter 4 we focus on nonparametric Bayesian inference of reliability of coherent systems which are prevalent in many settings such as in mechanical, engineering, military, and financial systems. In our nonparametric Bayesian approach we assign independent partition-based Dirichlet (PBD) priors, on the component distribution functions. A simultaneous inference procedure of compo-

nent and system reliabilities is developed. Bayesian paradigm provides a more general estimator in the sense that we can recover corresponding nonparametric estimators as a limiting case of our developed estimators both in recurrent event and reliability settings.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Analysis of Recurrent Event Data

Determination of an unknown distribution is one of the most important problems in Statistics. In many studies we only observe one event for a subject and inference is based on a single event setup. For example, in medical research an HIV patient is observed from the beginning of the treatment to the occurrence of a particular condition or death and the event time is denoted by T. Various models and methods (parametric, non-/semi-parametric, and Bayesian) have been considered for the analysis of survival data (lifetime data) based on single event settings and their asymptotic properties are well established. However, in many situations an event (e.g. tumor occurrences, heart attack) could occur repeatedly for the same subject over the monitoring period. Such an event is called a recurrent event in the literature.

Recurrent event data are prevalent in a wide variety of studies/fields such as clinical trials, epidemiology, public health, biomedicine, psychology, reliability, and engineering. For instance, in clinical trials/medical research we could observe recurrent events when a patient diagnosed with cancer tends to relapse over time or when a patient is repeatedly admitted to a hospital. Examples of recurrent events in the reliability engineering settings, are repeated failures of a deployed (mechanical/electronic) system and the occurrences of cracks in concrete structures, among others. In recurrent event analysis, interest usually lies in modeling the dependence of the occurrence of recurrent events on the covariates, so that inference can be carried

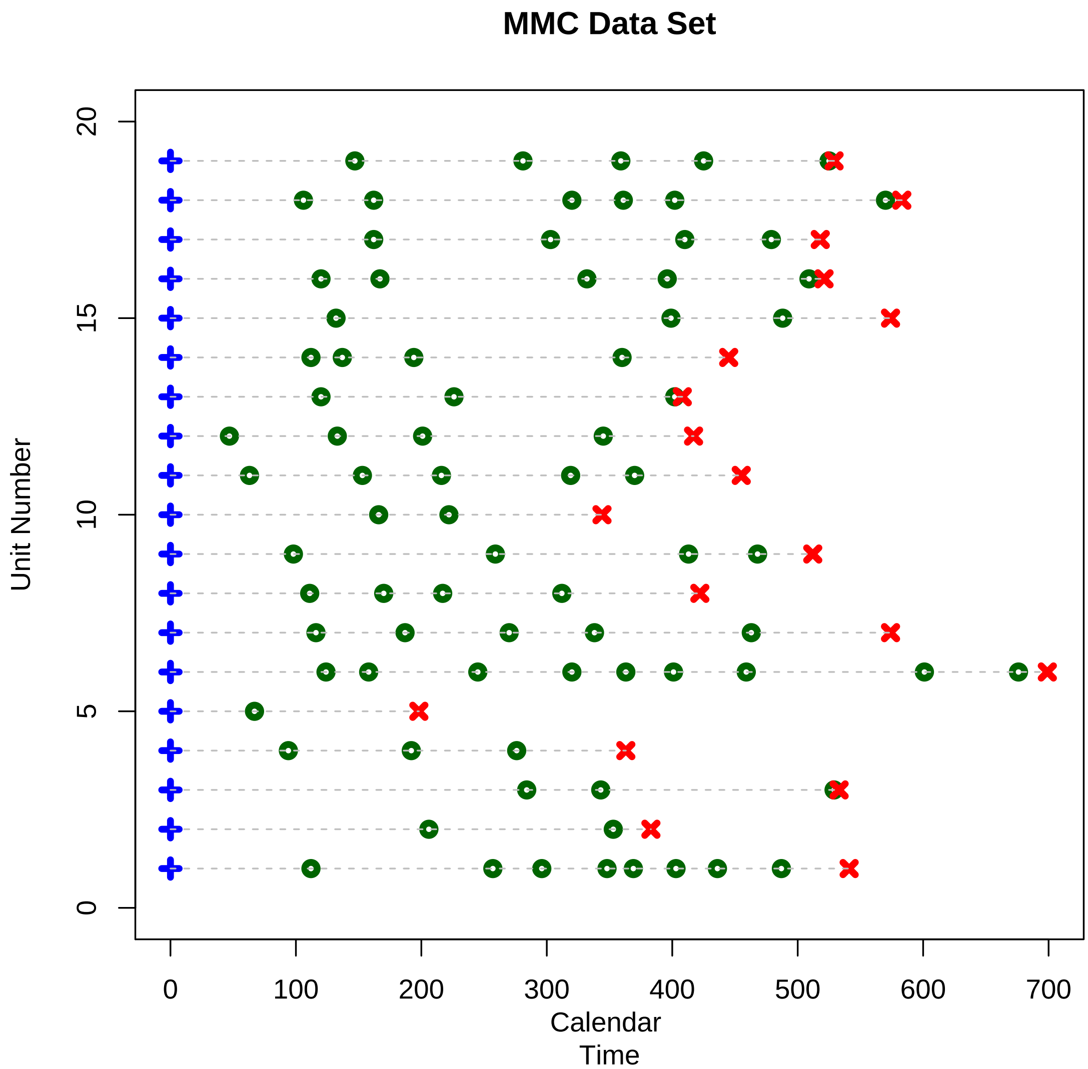

Figure 1.1: Pictorial representation of the MMC periods in minutes, (Censored: Red Crossed Sign)

out on the effects of covariates on the recurrent event process.

In the recurrent event data accrual scheme, a subject is monitored over a period $[0, \tau]$, where τ is some administrative time or study termination time. The monitoring period could be random, governed by an unknown distribution $G(t) = \mathscr{P}(\tau \leq t)$. Event-times are indexed by the calendar time s as well as the inter-event gap-times

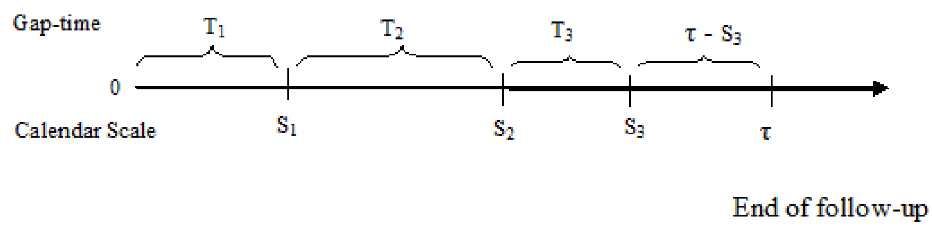

Figure 1.2: Pictorial representation of the occurrences of a recurrent event for a subject with $K = 3$ event occurrences.

Table 1.1: General form of data accrual in recurrent events

Unit	# Occurrences: K_i	Successive Gap-Times	Length of Study Period
1	K_1	$T_{11}, T_{12}, \ldots, T_{1K_1}$	τ_1
2	K_2	$T_{21}, T_{22}, \ldots, T_{2K_2}$	τ_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	K_n	$T_{n1}, T_{n2}, \ldots, T_{nK_n}$	τ_n

t. Denote the calendar times of event occurrences by

$$0 \equiv S_0 < S_1 < \ldots < S_K,$$

where S_k is the time of k-th event, and the inter-event gap-times by

$$T_1, T_2, \ldots, T_K \quad \text{with} \quad T_k = S_k - S_{k-1}, k = 1, 2, 3, \ldots.$$

The number of event occurrences is

$$K = \max\{k \in \{0, 1, 2, \ldots\} : \ S_k \leq \tau\}.$$

Note that the last event (event that traverses τ) is right censored by $\tau - S_K$, that is, $T_{K+1} \geq \tau - S_K$. Figure 1.1 is a pictorial representation of the occurrences of a recurrent event of 19 individuals from a study concerning the small bowel motility (see Husebye et al. (1990)), while Figure 1.2 is a pictorial representation of recurrent event occurrences for a subject with K=3. This recurrent event data, hereafter referred as the MMC data, will be used to illustrate the estimators developed in Chapter 2.

Table 1.2: Random Observables

Unit	Vector of Observables
1	$D_1 = (K_1, T_{11}, T_{12}, \ldots, T_{1K_1}, \tau_1 - S_{1K_1})$
2	$D_2 = (K_2, T_{21}, T_{22}, \ldots, T_{2K_2}, \tau_2 - S_{2K_2})$
$\vdots$	$\vdots$
n	$D_n = (K_n, T_{n1}, T_{n2}, \ldots, T_{nK_n}, \tau_n - S_{nK_n})$

A layout of data accrual is given in Table 1.1 when n units are in the study. The end of monitoring times $\tau_i, i = 1, 2, \ldots, n$ are IID from a common distribution function G. Random observables for the n units in the recurrent event setup is as given in Table 1.2

1.1.1 Models for Recurrent Event Analysis

There are several models and methods such as the complete intensity approach based counting process (Andersen and Gill (1982)), the marginal rate approach (Pepe and Cai (1993); Wei et al. (1989)), and the inter-event gap-times approach (e.g., Peña et al. (2001)) used for the analysis of recurrent event data. The main difference between various proposed methods is the function that is modeled or the parameter of interest (see Miloslavsky et al. (2004a)). Assume the covariate process $\mathbf{X}(s)$ for an arbitrary subject over the monitoring period $[0, \tau]$ is observable and denoted by $\{\mathbf{X}(s) = (X_1(s), X_2(s), \ldots, X_q(s)) : s \leq \tau\}$. To review some of the commonly used models in the literature, we assume $s \leq \tau$ and let

$$N(s) = \sum_{j=1}^{\infty} I(S_j \leq s) \text{ and } Y(s) = I(\tau \geq s),$$

which are also known as the counting process or the jump process and the "at-risk" process, respectively. The full information just before s is denoted by

$$\mathscr{H}_{s-} = \{\mathscr{N}(t), \mathscr{Y}(t), \mathscr{X}(t) : t < s\},$$

which is also known as the filtration or history of the stochastic process just before time s, where $\mathscr{N}(s) = \{N(t) : t \leq s\}$, $\mathscr{Y}(s) = \{Y(t) : t \leq s\}$, and $\mathscr{X}(s) = \{\mathbf{X}(t) : t \leq s\}$.

1.1.1.1 AG model (1982)

The intensity process of the counting process $N(t)$ is defined as

$$E[dN(t) \mid \mathscr{H}_{t-}] = Y(t)\lambda(t \mid \mathscr{H}_{t-}), \tag{1.1}$$

where $\lambda(t \mid \mathscr{H}_{t-})$ is the instantaneous probability of the process $N(t)$ jumping at time t conditional on $\mathscr{H}_{t-}$. The AG (Andersen and Gill (1982)) model is the most commonly used model for the intensity of a continuous counting process and is described in detail in Gill (1980), Andersen and Gill (1982), and Andersen et al. (1993) and is given by

$$\lambda(t \mid \mathscr{H}_{t-}) \equiv \lambda(t \mid X(t)) = \lambda_0(t) \exp(\beta^t X(t)), \tag{1.2}$$

where $\lambda_0(t) = \lim_{h \to 0} \frac{P(T < t+h | T \geq t)}{h}$ is the baseline hazard rate function of gap-time (baseline) distribution function F_0 at time t and β is a vector of regression coefficients. Note that from the formulation (1.1), it is clear that the intensity function depends on the past history of N. If the intensity process (e.g., 1.2) does not depend on N then the successive gap-times are assumed to be independent. The dependence between recurrent events can be addressed by incorporating frailty models in the full intensity process. In the frailty model one of the common assumptions is that given the unobserved frailty, successive gap-times are independent. Alternative models such as marginal models and proportional rate models have been developed by some authors to avoid the specification of dependence structure.

1.1.1.2 Marginal Models

Wei et al. (1989) (WLW) considered a marginal model in their analysis of the bladder cancer data. Their method models the marginal distribution of each failure time without considering the dependency structure among event times for a subject. In their approach each event or event type is modeled as a separate strata. The intensity function for the kth event is

$$\lambda_k(s \mid X(s)) = \lambda_{0k}(s) \exp[\beta^t X_k(s)], k = 1, 2, \ldots, K,$$

where $X_k(s)$ denotes the value of the covariate vector $X(s)$ at the kth event occurrence. Wei et al. (1989) provides event specific estimates as well as overall estimates. The overall estimate is the weighted average of the event-specific estimates such that the corresponding weighted average of the robust variance is the smallest possible. However, Cook and Lawless (1997) pointed out that the 'WLW' model is only valid when censoring is independent of event type or event occurrences. This approach is cumbersome when the number of recurrent event occurrences K is large. In the 'WLW' method subjects who are censored before $(k-1)$th event occurrence can be at risk for the kth event without having experienced $(k-1)$th event which seems unrealistic and difficult to interpret. Pepe and Cai (1993) included the term $N_{k-1}(t-) = 1$ in the filtration for the right-censored data to avoid the problem of being at risk of having kth event without having experienced $(k-1)$th event.

1.1.1.3 Proportional Rate Model

Lawless and Nadeau (1995a), Lawless et al. (2001), and Lin et al. (2000) are among others who considered modeling the rate of recurrent events by using a proportional rates model. The parameter of interest of their approach is the rate of $N(t)$ and is defined as

$$E[N(t) \mid \mathscr{H}^*_{t-}] = Y(t)m(t) \tag{1.3}$$

where $\mathscr{H}_{s-}^* = \{\mathscr{X}(t) : t \leq s\}$ and $m(t)$ is the rate of event occurrences in the process of $N(t)$. Lin et al. (2000) consider the proportional rate model given by

$$E[N(t) \mid \mathscr{H}_{t-}^*] = Y(t)m_0(t)\exp(\beta\gamma^*(t))$$

where m_0 is a nonnegative baseline rate function and $\gamma^*(t)$ is a known function of $\mathscr{H}_{t-}^*$.

1.1.1.4 A General Class of Models

Peña and Hollander (2004) proposed a general class of models for recurrent event analysis with the intensity function given by

$$\lambda(s \mid W = w, X(s) = x(s)) = w\lambda_0[\mathscr{E}(s)]\rho[N_i^\dagger(s-);\alpha]\psi[\beta^t x(s)].$$

This general class of models simultaneously incorporates the effect of covariates through the link function ψ, association among the event occurrences for a subject through the unobserved frailty W, the effect of accumulated event occurrences through $\rho(\cdot;\alpha)$, and the effect of intervention after event occurrences through the effective age processes $\mathscr{E}(\cdot)$. They considered semiparametric inference for a general class of models for recurrent event data based on gap-time formulation. In the estimation procedure Peña and Hollander (2004) and Peña et al. (2007) used modern tools such as counting process and martingale theory to develop the estimators with their asymptotic properties. In a simple setting (when gap-times are IID) with $\psi[\beta^t X] = \exp[\beta^t X]$ the intensity function associated with kth gap-time distribution is given by

$$\lambda_k(s \mid X) = \lambda_0(s - S_{k-1})\exp[\beta^t X].$$

In this case the at risk process is $Y(s) = I\{S_{k-1} < s \leq S_k\}$.

Some of the above models have been compared using real and simulated data, yielding different results as illustrated by Therneau and Hamilton (1997) and Therneau and Grambsch (2000), among others. The framework of the AG model is more widely used for multiple event and recurrent event analysis due to its efficiency as Therneau and Grambsch (2000) concluded. However, as Kelly and Lim (2000) pointed out, it is unclear which model is suitable for the analysis of recurrent events but they prefer the framework of the AG model over the marginal model.

1.1.2 Bayesian Inference with Recurrent Event Data

The classical form of making inference is to assume that F belongs to a parametric family characterized by a finite number of parameters and then estimate the parameters using the likelihood approach. However, the resulting inference will be erroneous if the assumed parametric form is misspecified. In contrast, the Bayesian paradigm provides more general estimators in that nonparametric Bayes estimators often are in the form of a linear combination of prior means and the classical nonparametric estimators. Consequently, we can obtain the corresponding nonparametric estimators as a limiting case of Bayes estimators. In addition, Bayes estimators are robust with respect to misspecification of prior measures or distributions. Susarla and Van Ryzin (1976) studied the nonparametric Bayesian estimation of the survival function in single event settings by assigning a Dirichlet process prior on F. One can view Susarla and Van Ryzin (1976)'s estimator as a Bayesian counterpart of the Kaplan and Meier (1958) estimator of $\bar{F}$. Kalbfleisch (1978) considered nonparametric Bayesian analysis of single event survival time data by assigning a gamma processes prior on the baseline cumulative hazard function $\Lambda_0(t) = \int_0^t \lambda_0(u)du$.

Nonparametric Bayesian inference with recurrent event data has not yet been completely developed. Sinha (1993) considered semiparametric Bayesian inference of multiple events time data, while Ouyang et al. (2013) considered semiparametric Bayesian inference with recurrent event data. The main interest of their analysis is to assess the effects of treatments on the event time. However, in this dissertation we develop nonparametric and semiparametric Bayesian inference of the inter-event gap-time distribution with recurrent event data.

First, we consider nonparametric Bayes (NPB) and empirical Bayes (NPEB) estimation of the inter-event gap-time survivor function. NPB and NPEB estimators are developed under the assumption that the gap-times are IID from some common distribution F and gap-times and monitoring times are mutually independent. In our nonparametric Bayesian approach, we assume that F has a Dirichlet process prior with parameter α, a non-null finite measure; see Ferguson (1973) and Sethuraman (1994) for the early development of the Dirichlet process prior. The resulting posterior first moment of $\bar{F}$ is our nonparametric Bayes estimator of $\bar{F}$ under the integrated squared-error loss function

$$L(\hat{F}, F) = \int [\hat{F}(t) - F(t)]^2 dt, \tag{1.4}$$

where $\hat{F}(t)$ is an estimator of $F(t)$.

Assigning a Dirichlet process prior on F also requires us to specify a parameter α, which is a measure such that $\alpha(-\infty, t]/\alpha(\Re)$ is the prior mean of $F(t)$. However, instead of specifying α based on subjective beliefs, we may estimate it using the observed data. The resulting Bayes estimator of $\bar{F}$ is referred to as an empirical Bayes estimator (cf.,Robbins (1956)). We also developed an empirical Bayes estimator of the gap-time distribution function in recurrent event settings. Details of the nonparametric Bayes and empirical Bayes inference procedure of the gap-time distribution are presented in Chapter 2.

In biomedical applications IID assumptions are somewhat restrictive as the gap-times could be correlated. Moreover, in many biomedical/epidemiological applications or clinical trials the primary interest is whether the treatment is effective in reducing event occurrences. Thus, there is a need to develop methodology for the correlated gap-times to assess the effect of the treatment on the gap-times.

Secondly, we consider a semiparametric Bayesian inference of correlated gap-times with recurrent event data. To model the correlation between gap-times for a subject we consider a frailty model. In our approach we assume an unobserved frailty random variable $W \mid \nu \sim Ga(\nu, \nu)$ with unit mean and variance $1/\nu$. It is assumed that gap-times and monitoring times are mutually independent. Given the unobserved frailty variable $W = w$, we assume gap-times $T_j, j = 1, 2, \ldots,$ are IID from a distribution

$$\bar{F}(t \mid W = w) = \bar{F}_0(t)^w = \exp(-w\Lambda_0(t)),$$

where $\bar{F}_0(t)$ and $\Lambda_0(t)$ are the baseline survival and cumulative hazard function, respectively. Denote by X the observable covariate of interest. We consider the intensity function defined by

$$\lambda(t \mid W = w, \mathbf{X} = \mathbf{x}) = \lambda_0(t) w \exp(\beta^T \mathbf{x}).$$

The parameters of interest are Λ_0, β, and ν. We model Λ_0 non-parametrically, in particular, assign it a gamma process prior and assume parametric priors for β and ν.

To recall the definition of the gamma process prior let $V(t)$ be a stochastic process with $V(\infty) < \infty$ and $0 \equiv t_0 < t_1 < \ldots < t_M < t_{M+1} = \infty$ are partition points of $[0, \infty)$. Assume that the increments $\{V(t_1), V(t_2) - V(t_1), \ldots, V(t_{M+1}) - V(t_M)\}$ are independent and have independent gamma distributions with shape parameters $\{\zeta(t_1), \zeta(t_2) - \zeta(t_1), \ldots, \zeta(t_{M+1}) - \zeta(t_M)\}$, where ζ is a nondecreasing function with $\zeta(\infty) < \infty$. Then $V(t)$ is called a gamma process.

Following Kalbfleisch (1978) and Sinha (1993) we assign a gamma process prior on the baseline cumulative hazard function $\Lambda_0(t)$. Assume the prior distributions of β and ν are a multivariate normal distribution and a gamma distribution with known hyper-parameters, respectively. Conditional posterior distributions of Λ_0, β, and ν are derived to facilitate the Gibbs sampling procedure. Credible intervals of the parameters of interest are obtained from the posterior samples. From our Bayes estimator of the baseline cumulative hazard function we can recover a Breslow-Aalen type estimator of the baseline cumulative hazard function as a limiting case. We illustrate our methodology by analyzing two real data sets. Chapter 3 contains the details of the semiparametric Bayesian inference procedure with recurrent event data.

1.2 Reliability of a Coherent System

It is of interest to assess the risk of failure and the reliability of systems in many settings, for instance, in mechanical, engineering, military, and financial systems. The statistical analysis of the reliability of technical systems emerged just after World War I when it was used to compare the operational safety of one-, two-, and four-engine airplanes. The theoretical basis of the statistical methods for analyzing the quality of industrial components and systems was laid down by Walter Shewhart, Harold F. Dodge, Harry G. Romig, and Walter Deming at the beginning of the 1930's. However, such methods were not brought into use to any great extent until the beginning of World War II. Prior to the advent of the use of statistical reliability and quality control methods, it was the case that even systems with a large number of high quality components often did not achieve their desired reliability (cf., Rausand and Høyland (2004)). After World War II more sophisticated products/systems were produced with an increasing number of parts/components, for example, televisions, computers, intercontinental ballistic missiles, spacecraft, communication satellites, passenger and military aircraft. Consequently, there is a critical need to be able to assess the risk and reliability of complicated systems. For more details about the history and development of reliability theory, see Knight (1991) and Villemeur (1992).

Precise and reliable knowledge of the performance of deployed systems enables an informed assessment of risk and failure of the system that could potentially save lives, enhance wealth, and prevent destruction. It is therefore imperative to have probabilistic and statistical inferential methods to assess the risk and reliability of systems.

Assume n identical systems each with K independent components are in the study.

The structure function of a reliability system is defined by $\phi : \{0,1\}^K \to \{0,1\}$ such that $\phi(\mathbf{x})$ indicates whether the system is in a functioning state ($\phi(\mathbf{x}) = 1$) or is in a failed state ($\phi(\mathbf{x}) = 0$). Denote the component lifetime survivor functions by $\bar{F}_j(t) = Pr\{T_j > t\}, j = 1, 2, \ldots, K$. If the component lifetimes are independent, then the system survivor function could be expressed in terms of the system's reliability function via

$$\bar{F}_\phi(t) = h_\phi(\bar{F}_1(t), \bar{F}_2(t), \ldots, \bar{F}_K(t)). \tag{1.5}$$

More details of the structure function of a coherent system and the reliability function is given in Chapter 4.

In the nonparametric Bayesian framework we develop an estimation procedure by assigning a partition-based Dirichlet process prior (Sethuraman and Hollander (2009)) on F when system lifetime data are available. Similarly in the autopsy model when component lifetime data are available we estimate the distribution for each component by assigning an independent PBD prior. Therefore, simultaneous estimation of component reliabilities and system reliability is performed by plugging in the estimates of component reliabilities in (1.5). We evaluate the performance of the developed nonparametric Bayes estimators through simulation studies in terms of biases and RMSE's functions and compare with the corresponding nonparametric Doss et al. (1989) estimators. The Doss et al. (1989) estimator of the system reliability function is a limiting case of our proposed estimator.

We now outline the contents of this dissertation. In Chapter 2 we develop a methodology for nonparametric Bayes estimation of the gap-time distribution with recurrent event data. Chapter 2 is a pre-print version of the article published by the Journal of Nonparametric Statistics in 2014, available at http://www.tandfonline.com /doi/full/10.1080/10485252.2014.906744. In Chapter 3 we develop a semiparametric Bayesian inference of gap-time distribution with recurrent event data. In Chapter 4 we develop a nonparametric Bayesian inference of reliability of coherent systems.

Technical results of Chapter 2 and Chapter 4 are gathered in Chapter 2 and Chapter 4 appendix, respectively. Chapter 2 appendix also contains the copyright permission to reprint the Chapter 2.

Chapter 2

Nonparametric Bayes Estimation of Gap-Time Distribution with Recurrent Event Data[1]

Abstract

Nonparametric Bayes estimation of the gap-time survivor function governing the time to occurrence of a recurrent event in the presence of censoring is considered. In our Bayesian approach, the gap-time distribution, denoted by F, has a Dirichlet process prior with parameter α. We derive nonparametric Bayes (NPB) and empirical Bayes (NPEB) estimators of the survivor function $\bar{F} = 1 - F$ and construct point-wise credible intervals. The resulting Bayes estimator of $\bar{F}$ extends that based on single-event right-censored data, and the PL-type estimator is a limiting case of this Bayes estimator. Through simulation studies, we demonstrate that the PL-type estimator has smaller biases but higher root-mean-squared errors (RMSE's) than those of the NPB and the NPEB estimators. Even in the case of a mis-specified prior measure parameter α, the NPB and the NPEB estimators have smaller RMSE's than the PL-type estimator, indicating robustness of the NPB and NPEB estimators. In addition, the NPB and NPEB estimators are smoother (in some sense) than the PL-type estimator.

[1] A.K.M. F. Rahman, J. Lynch and E. A. Peña. 2014. *Journal of Nonparametric Statistics*. http://www.tandfonline.com/doi/full/10.1080/10485252.2014.906744.
Reprinted here with permission of publisher.

2.1 Introduction

One of the most important problems in Statistics is the determination of an unknown distribution, equivalently an unknown probability measure, governing a probability space. In its most classical form, there is a random sample $T_1, T_2, \ldots, T_n$ which are independent and identically distributed (IID) from a distribution F which is only known to belong to the space of distribution functions $\mathfrak{F}$. The nonparametric estimator of F is the empirical distribution function (EDF) defined by

$$\widehat{F}_n(t) = \frac{1}{n}\sum_{i=1}^{n} I\{T_i \le t\}, t \in \Re,$$

where $I\{\cdot\}$ is the indicator function. From the works of Kolmogorov (1933), Smirnov (1948), Glivenko (1933), Cantelli (1933), and Donsker (1952), the consistency of $\widehat{F}_n$ to F and the weak convergence of the process $W_n(t) = \{\sqrt{n}[\widehat{F}_n(t) - F(t)] : t \in \Re\}$ to a zero-mean Gaussian process are well-established.

Starting with the seminal work of Kaplan and Meier (1958), the estimation of F was also undertaken when the sample data are right-censored. Thus, instead of observing $T_1, T_2, \ldots, T_n$, one is only able to observe $(Z_1, \delta_1), (Z_2, \delta_2), \ldots, (Z_n, \delta_n)$, where $\delta_i \in \{0, 1\}$ with the interpretation that $\delta_i = 1$ means that $T_i = Z_i$, while $\delta_i = 0$ means that $T_i > Z_i$. In this setting, the nonparametric estimator of F is the so-called product-limit estimator (PLE), also called the Kaplan-Meier estimator (KME), given by

$$\widehat{F}_n(t) = 1 - \prod_{v \le t}\left[1 - \frac{\Delta N(v)}{Y(v)}\right], t \in \Re,$$

where $\prod$ means product-integral, and where the processes $N = \{N(t) : t \in \Re\}$ and $Y = \{Y(t) : t \in \Re\}$ are defined via

$$N(t) = \sum_{i=1}^{n} I\{Z_i \le t; \delta_i = 1\} \quad \text{and} \quad Y(t) = \sum_{i=1}^{n} I\{Z_i \ge t\}.$$

From the works of Kaplan and Meier (1958), Efron (1967), Breslow and Crowley (1974), and Gill (1980), finite-sample and asymptotic properties of this PLE are well-established, in particular, its consistency and the weak convergence to a Gaussian process of the process W_n.

Ferguson (1973) introduced the Dirichlet process probability measure on the space of distribution functions. This Dirichlet process was then used as a prior measure over the space $\mathfrak{F}$, and a nonparametric Bayes estimator of F was developed in the complete data setting where $T_1, T_2, \ldots, T_n$ are observed. This Bayes estimator, which is obtained under an integrated squared-error loss function, is a convex combination of the empirical distribution function and the prior estimate of F under the Dirichlet process prior. With right-censored data, Susarla and Van Ryzin (1976) obtained the nonparametric Bayes estimator of F, also under integrated squared-error loss, when a Dirichlet process prior is assigned on F.

Recurrent event data sets arise from a wide variety of studies/fields such as clinical trials, epidemiology, public health, biomedicine, psychology, reliability and engineering. Examples are repeated tumor occurrences in cancer patients (Byar (1980)), successive seizures in epileptic patients (Albert (1991)), recurrent small bowel motility in gastroenterology study (Aalen and Husebye (1991)), and repeated warranty claims or failures in manufactured equipments (Kalbfleisch et al. (1991)). The development of statistical models and methods for analyzing recurrent event data is of crucial importance. For instance, Proschan (1963), Gill (1980), Vardi (1982), Sellke (1988), Aalen and Husebye (1991), McClean and Devine (1995), Soon and Woodroofe (1996), Wang and Chang (1999), Ghosh and Lin (2000), Peña et al. (2001), Cook and Lawless (2002), Nelson (2003), Lindqvist (2006), Peña et al. (2007), Stocker and Peña (2007), Adekpedjou et al. (2010), and Gjessing et al. (2010), have dealt with inferential problems with recurrent event data. However, nonparametric Bayesian inference with recurrent event data has not yet been completely developed. Susarla and Van Ryzin (1976) studied nonparametric Bayesian estimation of the survival function given a right-censored data in single event settings using a Dirichlet process prior. Sethuraman and Hollander (2009) considered nonparametric Bayes estimation for repair models using a partitioned-based prior. In addition, Robbins (1956), first formally introduced the notion of an empirical Bayes approach. Following Robbins (1956), Korwar and Hollander (1976) and Susarla and Van Ryzin (1978) dealt with a nonparametric setup of empirical Bayes estimation of survival function for single event complete and right-censored observations, respectively. Significant works on a parametric empirical Bayes approach appeared in a series of paper by Efron and Morris (1972, 1973, 1976).

In this paper, we consider nonparametric Bayes estimation of the inter-event gap-time survivor function when recurrent event data is available. For our purpose we assume that the successive inter-event times (gap-times) for the ith unit, denoted by $\{T_{ij}, j = 1, 2, \ldots\}$, are independent and identically distributed (IID) nonnegative random variables with a common distribution function F. However, the ith unit will only be observed over $[0, \tau_i]$ where $\tau_1, \tau_2, \ldots, \tau_n$ are IID with a common distribution function G. It is assumed that τ_i and $\{T_{ij}, j = 1, 2, \ldots\}$ are mutually independent. For the ith unit the number of observed event occurrences is

$$K_i = \max\{k \in \{0, 1, \ldots\} : S_{ik} \leq \tau_i\},$$

where $S_{i0} = 0$ and $S_{ik} = \sum_{j=1}^{k} T_{ij}, k = 1, 2, \ldots,$ and the observable random vector is

$$\mathbf{D}_i = (\tau_i, K_i, T_{i1}, T_{i2}, \ldots, T_{iK_i}, \tau_i - S_{iK_i}).$$

It should be emphasized that this observable random vector has more complicated distributional characteristics owing to the sum-quota accrual constraint which states that for each $i = 1, 2, \ldots, n$,

$$S_{iK_i} \leq \tau_i < S_{iK_i} + T_{iK_i+1}.$$

In particular, given $K_i \geq 1$, T_{i1} does not any more have an $\bar{F}$ survival function. Aside from this, $(T_{i1}, T_{i2}, \ldots, T_{iK_i})$, given $K_i = k_i$, are not any more independent of each other since $\sum_{j=1}^{K_i} T_{ij} \leq \tau_i$. This distributional complexity distinguishes recurrent-event problems from the conventional single event settings. Note that $\tau_i - S_{iK_i}$ is the right-censoring variable for T_{iK_i+1}. This is the same recurrent event model considered in Peña et al. (2001) where a nonparametric PL-type estimator of F was obtained. Their estimator extended the single event Kaplan and Meier (1958) product-limit estimator to the recurrent event setting.

In our nonparametric Bayesian approach, we assume that F has a Dirichlet process prior with parameter α, a non-null finite measure; see Ferguson (1973) and Sethuraman (1994) for the early developments of the Dirichlet process prior. The idea implemented in deriving our nonparametric Bayes estimator of $\bar{F}$ mimics that of Susarla and Van Ryzin (1976), where the first step is to update the Dirichlet process given all the complete observations. In the second step we then compute the posterior moments of $\bar{F}$ when the right-censored observations are also given. The resulting posterior first moment of $\bar{F}$ is our nonparametric Bayes estimator of $\bar{F}$ under an integrated squared-error loss function. In addition, instead of specifying the prior parameter α, we also employ empirical Bayes ideas where we utilize the data to estimate α. The estimator resulting from this approach is referred to as an empirical Bayes estimator; see Robbins (1956) and Casella (1985). It will be seen that both the nonparametric Bayes and empirical Bayes estimators have smaller root-mean-squared errors (RMSE's) than the PL-type estimator. Results of simulation studies will further indicate that the nonparametric Bayes and empirical Bayes estimators are robust in the sense that they do not suffer severely from a mis-specified prior measure parameter α. Moreover, the RMSE's of the empirical Bayes estimator is smaller than that of the PL-type estimator even in this mis-specified case.

We outline the contents of this chapter. In Section 2.2, we recall some results about Dirichlet processes and then derive the nonparametric Bayes estimator of $\bar{F} = 1 - F$ and construct point-wise credible intervals of $\bar{F}$. Section 2.3 establishes that the

PL-type estimator is a limiting case of the nonparametric Bayes estimator. It is also established that the Bayes estimator is a linear combination of the prior mean distribution $\bar{\alpha} = \alpha/\alpha(\Re_+)$ and the PL-type estimator. In Section 2.4, results of our simulation studies are presented. In this section the biases and root-mean-squared errors (RMSE's) of the Bayes and the empirical Bayes estimators are compared with those of the PL-type estimator. Biases and RMSE's of Bayes and empirical Bayes estimators under a mis-specified prior are also examined, which will demonstrate the robustness of the Bayesian estimator. Section 2.5 provides an illustration using the gastroenterology data in Aalen and Husebye (1991). Section 2.6 provides some concluding thoughts. An appendix gathers the technical proofs.

2.2 Nonparametric Bayes Estimation

2.2.1 Mathematical Setup

In the sequel, all random entities will be defined on a basic probability space given by $(\Omega, \mathscr{A},)$, where $\mathscr{A}$ is a σ-field of subsets of the space Ω and is a probability measure on the measurable space $(\Omega, \mathscr{A})$. Generic elements of Ω will be denoted by ω's. $\mathfrak{F}$ will denote the collection of all distribution functions on $\Re$ whose supports are subsets of $[0, \infty)$, and this is endowed with a σ-field of subsets, $\mathscr{F}$, generated by the finite-dimensional cylinder sets. $\mathscr{D}(\alpha)$ will denote a Dirichlet process on $(\mathfrak{F}, \mathscr{F})$ with parameter α, a non-null finite measure on $(\Re_+, \mathscr{B}_+)$, as introduced in Ferguson (1973). The number of units or subjects in the study will be denoted by n. We now formally describe the Bayesian statistical model of interest.

Let $F : (\Omega, \mathscr{A}) \to (\mathfrak{F}, \mathscr{F})$ and $G : (\Omega, \mathscr{A}) \to (\mathfrak{F}, \mathscr{F})$ be $\mathfrak{F}$-valued stochastic processes which are independent and with $F \sim \mathscr{D}(\alpha)$ and $G \sim \mathscr{D}(\alpha')$, where α and α' are non-null finite measures on $(\Re_+, \mathscr{B}_+)$. For $j = 1, 2, \ldots; i = 1, 2, \ldots, n$, let $T_{ij} : (\Omega, \mathscr{A}) \to (\Re_+, \mathscr{B}_+)$ and $\tau_i : (\Omega, \mathscr{A}) \to (\Re_+, \mathscr{B}_+)$, such that, given (F, G),

$\{T_{ij}\}$s are independent and identically distributed (IID) with distribution F, $\{\tau_i\}$s are IID with distribution G, and $\{T_{ij}\}$s and $\{\tau_i\}$s are independent. We do not observe completely the $\{T_{ij}\}$s, rather we observe, for each $i = 1, 2, \ldots, n$,

$$\mathbf{D}_i = (\tau_i, K_i, T_{i1}, T_{i2}, \ldots, T_{iK_i}, \tau_i - S_{iK_i}), \tag{2.1}$$

where $S_{i0} \equiv 0$, $S_{ik} = \sum_{j=1}^{k} T_{ij}, k = 1, 2, \ldots;$ and $K_i = \max\{k \in \{0, 1, 2, \ldots\} : S_{ik} \leq \tau_i\}$.

The observable $\mathbf{D}_i$ is the recurrent event data as considered in Peña et al. (2001). The independent Dirichlet processes $\mathscr{D}(\alpha)$ and $\mathscr{D}(\alpha')$, which are probability measures governing F and G, serve as the nonparametric priors on the underlying distributions F and G in the context of the Bayesian setting considered in this paper. Our goal is to obtain a nonparametric Bayes estimator of F, given $\mathbf{D} = (\mathbf{D}_1, \mathbf{D}_2, \ldots, \mathbf{D}_n)$, under the integrated squared-error loss function

$$L(\widehat{F}, F) = \int [\widehat{F}(t) - F(t)]^2 dw(t), \tag{2.2}$$

where $w : [0, \infty) \to \Re_+$ is a pre-specified weight function and $\widehat{F}(t)$ is an estimator of $F(t)$. In our approach we will consider $w(t) = t$, so that under the loss function (2.2), the posterior mean function is our Bayes estimator of $F(\cdot)$, the estimator that yields the minimum Bayes risk.

We recall the definition of a Dirichlet process and some of its basic properties as obtained in Ferguson (1973).

Definition 1: Let α be a non-null finite measure on $(\Re_+, \mathscr{B}_+)$, where $\mathscr{B}_+$ is the Borel σ-field on $[0, \infty)$. We say that the random probability measure P on $(\Re_+, \mathscr{B}_+)$ is a Dirichlet process with parameter α if for every $k = 1, 2, \ldots$ and measurable partition $(B_1, B_2, \ldots, B_k)$ of $\Re_+$, the joint distribution of the random vector $(P(B_1), \ldots, P(B_k))$ is a Dirichlet distribution denoted by $\mathfrak{D}(\alpha(B_1), \ldots, \alpha(B_k)) \triangleq \mathfrak{D}(\alpha)$. The associated random distribution function $F(t) \doteq P((-\infty, t]), t \in \Re_+$, is then also said to be a Dirichlet process.

Definition 2: Let P be a random probability measure on $(\Re_+, \mathscr{B}_+)$. Then $(T_1, T_2, \ldots, T_n)$ is said to be a random sample of size n from P if, for any $m = 1, 2, \ldots,$ and measurable sets $A_1, A_2, \ldots, A_m, C_1, C_2, \ldots, C_n$, we have

$$\mathscr{P}\{T_1 \in C_1, \ldots, T_n \in C_n \mid P(A_1), \ldots, P(A_m), P(C_1), \ldots, P(C_n)\} = \prod_{j=1}^{n} P(C_j), \text{ a.s..}$$

Result 1: Let P be a Dirichlet process on $(\Re_+, \mathscr{B}_+)$ with parameter α. If T is a sample of size one from P, then, for any $A \in \mathscr{B}_+$,

$$\mathscr{P}(T \in A) = \frac{\alpha(A)}{\alpha(\Re_+)}.$$

Result 2: Let P be a Dirichlet process on $(\Re_+, \mathscr{B}_+)$ with parameter α and $T_1, T_2, \ldots, T_n$ be a sample of size n from P. Then, the conditional probability measure of P, given $T_1, T_2, \ldots, T_n$, is a Dirichlet process with parameter $\alpha + \sum_{i=1}^{n} \delta_{T_i}$. That is, if $P \sim \mathscr{D}(\alpha)$, then

$$P \mid (T_1, T_2, \ldots, T_n) \sim \mathscr{D}\left(\alpha + \sum_{i=1}^{n} \delta_{T_i}\right), \text{ where } \delta_T(A) = I\{T \in A\}.$$

2.2.2 Nonparametric Bayes Estimator of $\bar{F}$

Recall that random observables for the n subjects are $\mathbf{D} = (\mathbf{D}_1, \mathbf{D}_2, \ldots, \mathbf{D}_n)$, where $\mathbf{D}_i$ is defined in (2.1). Observe that $\tau_i - S_{iK_i}$ is the right-censoring random variable for T_{iK_i+1}, so that the last observation for each unit is always right-censored. Clearly, $T_{iK_i+1} \in [T_i^*, \infty)$, where $T_i^* = \tau_i - S_{iK_i}, i = 1, 2, \ldots, n$. We use $\{t_i^*\}$'s as the realized values of $\{T_i^*\}$'s. For $i = 1, 2, \ldots, n$, we define

$$\delta_{ij} = 1, j = 1, 2, \ldots, K_i;\ \delta_{iK_i+1} = 0.$$

One may represent the observable random variables for all the n units according to

$$\{(\delta_{ij}, T_{ij}), \quad i = 1, 2, \ldots, n; \quad j = 1, 2, \ldots, K_i, K_i + 1\}.$$

That is, if δ_{ij} =1, then T_{ij}'s are complete (uncensored) observations, while if δ_{ij}=0, then T_{ij} is a right-censored observation.

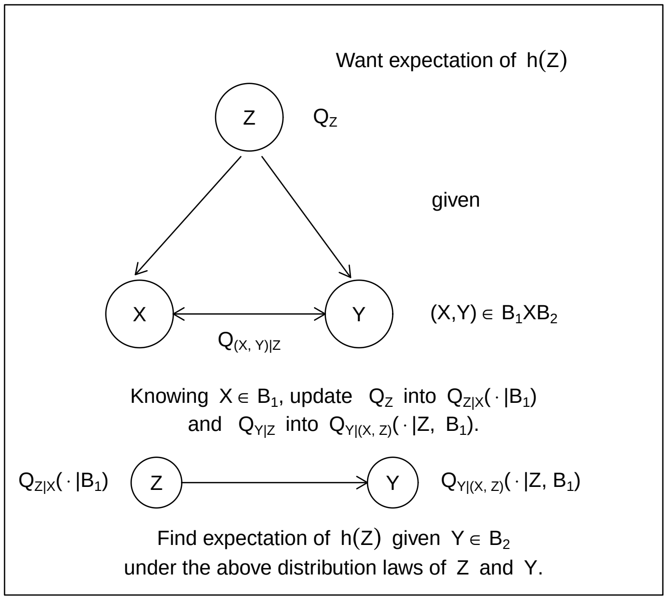

Figure 2.1: Graphical representation of Lemma 1.

We update the prior measure by deriving the distribution of P, given $(\tau_i, K_i = k_i, T_{i1} = t_{i1}, T_{i2} = t_{i2}, \ldots, T_{iK_i} = t_{ik_i}, T_{ik_i+1} > \tau_i - S_{ik_i}; i = 1, 2, \ldots, n)$, or, equivalently, $P \mid (T_{i1} = t_{i1}, T_{i2} = t_{i2}, \ldots, T_{iK_i} = t_{ik_i}, S_{ik_i} \leq \tau_i, T_{ik_i+1} \in [t_i^*, \infty); i = 1, 2, \ldots, n)$. This will be done in two steps. Step 1 uses the crucial Proposition 1, which demonstrates that the posterior measure is still a Dirichlet process under the renewal process setting given all the uncensored (complete) observations. In order to prove Proposition 1, we first prove a more general lemma on conditional expectation. A special case of the Lemma will then be applied to establish Proposition 1. Using the results in Proposition 1, Proposition 2, Proposition 3, and Proposition 4 below, Theorem 1 will establish the nonparametric Bayes estimator (2.5). The proofs of these results are gathered in the Chapter 2 appendix.

Lemma 1: Let $(\Omega, \mathscr{A}, P)$ be a probability space and let $(Z, X, Y) : (\Omega, \mathscr{A}) \to (\mathscr{Z} \times \mathscr{X} \times \mathscr{Y}, \sigma(\mathscr{Z} \times \mathscr{X} \times \mathscr{Y}))$. Let Q_Z be the probability measure induced by Z, and let $Q_{(X,Y)|Z}$ be a version of the conditional joint probability measure of (X, Y) given Z. Let $h : (\mathscr{Z}, \sigma(\mathscr{Z})) \to (\Re_1, \mathscr{B}_1)$ with $E[|h(Z)|] < \infty$. Then, for $B_1 \times B_2 \in \sigma(\mathscr{X} \times \mathscr{Y})$,

$$\begin{aligned} &\mathbf{E}_{\substack{Z\sim Q_Z(\cdot) \\ (X,Y)|Z\sim Q_{(X,Y)|Z}(\cdot,\cdot|z)}} [h(Z) \mid (X, Y) \in (B_1 \times B_2)] \\ &= \mathbf{E}_{\substack{Z\sim Q_{Z|X}(\cdot|B_1) \\ Y|(Z,X)\sim Q_{Y|(Z,X)}(\cdot|z,B_1)}} [h(Z) \mid Y \in B_2] \,, \end{aligned}$$

where $Q_{Z|X}(\cdot \mid B_1)$ is a conditional probability measure of Z given $X \in B_1$, and $Q_{Y|(Z,X)}(\cdot \mid z, B_1)$ is a conditional probability measure of Y given $Z = z$ and $X \in B_1$. Corollary 1 follows immediately from Lemma 1 with $X \perp Y \mid Z$.

Corollary 1: Under the condition of Lemma 1, if (X, Y) are independent conditional on Z, then

$$\begin{aligned} &\mathbf{E}_{\substack{Z\sim Q_Z(\cdot) \\ (X,Y)|Z\sim Q_{(X,Y)|Z}(\cdot,\cdot|z)}} [h(Z) \mid (X, Y) \in (B_1 \times B_2)] \\ &= \mathbf{E}_{\substack{Z\sim Q_{Z|X}(\cdot|B_1) \\ Y|Z\sim Q_{Y|Z}(\cdot|Z)}} [h(Z) \mid Y \in B_2] \,. \end{aligned}$$

Corollary 2 is a specific version of the Corollary 1, which will be applied to prove the Proposition 1.

Corollary 2: Let P be a random probability measure on $(\Re_+, \mathscr{B}_+)$ which is a Dirichlet process with parameter α, denoted by $\mathscr{D}(\alpha) \equiv \mathscr{D}_\alpha$. Let T_1, T_2 be random variables such that, given P, T_1 and T_2 are independent with $(T_1 \mid P) \sim P$ and $(T_2 \mid P) \sim P$. Let h be a measurable function of P with $\int |h(P)| \mathscr{D}_\alpha(dP) < \infty$. Then,

$$\begin{aligned} &\mathbf{E}_{P\sim\mathscr{D}_\alpha;(T_1,T_2)|P \overset{iid}{\sim} P} [h(P) \mid (T_1 = t_1, T_2 \in B_2)] \\ &= \mathbf{E}_{P\sim\mathscr{D}_{\alpha+\delta_{t_1}};\ T_2|P\sim P} [h(P) \mid T_2 \in B_2] \,. \end{aligned}$$

In the sequel, we use an abbreviated notation where $\mathbf{E}_{\substack{P\sim\mathscr{D}_\alpha \\ T_1,T_2,\dots\ |P \overset{iid}{\sim} P}} [\dots] \equiv \mathbf{E}_{P\sim\mathscr{D}_\alpha} [\dots]$. So in particular, $\mathbf{E}_{\substack{P\sim\mathscr{D}_{\alpha^*} \\ T_{K+1}|P\sim P}} [\dots] \equiv \mathbf{E}_{P\sim\mathscr{D}_{\alpha^*}} [\dots]$.

Proposition 1: Let P be a Dirichlet process on $(\Re_+, \mathscr{B}_+)$ with parameter α. Given

P, let $T_{ij}, j = 1, 2, \ldots;\ i = 1, 2, \ldots, n$, be IID from P. Let $\tau_i, i = 1, 2, \ldots, n$, be IID from G. For $i = 1, 2, \ldots, n$, let $K_i = \max\{k \in \{0, 1, 2, \ldots\} :\ S_{ik} = \sum_{j=1}^{k} T_{ij} \le \tau_i\}$. For a measurable function $h : (\mathfrak{F}, \mathscr{F}) \to (\Re, \mathscr{B})$ with $\mathbf{E}_{P \sim \mathscr{D}(\alpha)} \mid h(P) \mid < \infty$, we have

$$\mathbf{E}_{P \sim \mathscr{D}(\alpha)}[h(P) \mid (\tau_i, K_i = k_i, T_{i1} = t_{i1}, T_{i2} = t_{i2}, \ldots, T_{ik_i} = t_{ik_i}, i = 1, 2, \ldots, n)]$$
$$= \mathbf{E}_{P \sim \mathscr{D}(\alpha^*)}[h(P) \mid (T_{ik_i+1} \in [t_i^*, \infty), i = 1, 2, \ldots, n)] I(S_{ik_i} \le \tau_i),$$

where $\alpha^* = \alpha + \sum_{i=1}^{n} \sum_{j=1}^{k_i} \delta_{T_{ij}}$.

In step 2, instead of deriving the posterior distribution and then obtaining the moments of $h(P)$, we directly derive the moments of $h(P)$, given all the right-censored observations, i.e. the moments of $h(P) \mid (T_{ik_i+1} \in [t_i^*, \infty); i = 1, 2, \ldots, n)$, where P is a Dirichlet process with parameter $\alpha^* = \alpha + \sum_{i=1}^{n} \sum_{j=1}^{k_i} \delta_{t_{ij}}$. The conditional distribution is invariant under any permutation of the censored observations. More specifically, we obtain the conditional moments of $h(\bar{F}(u)) \mid (T_{ik_i+1} \in [t_i^*, \infty); i = 1, 2, \ldots, n)$, where $\bar{F}(u) = P([u, \infty))$. We define the processes $Y^+ = \{Y^+(t) : t \in \Re\}$ and $Y = \{Y(t) : t \in \Re\}$ via

$$Y^+(t) = \sum_{i=1}^{n} \sum_{j=1}^{\infty} I(T_{ij} > t, \tau_i - S_{ij-1} > t) = \sum_{i=1}^{n} \left[\sum_{j=1}^{K_i} I(T_{ij} > t) + I(\tau_i - S_{iK_i} > t) \right] \tag{2.3}$$

and

$$Y(t) = \sum_{i=1}^{n} \sum_{j=1}^{\infty} I(T_{ij} \ge t, \tau_i - S_{ij-1} \ge t) = \sum_{i=1}^{n} \left[\sum_{j=1}^{K_i} I(T_{ij} \ge t) + I(\tau_i - S_{iK_i} \ge t) \right] \tag{2.4}$$

where $Y^+(t)$ and $Y(t)$, respectively, denote the number of events (censored and uncensored) such that event times are strictly greater than $t \in \Re$ and greater than or equal to $t \in \Re$. In the statements below, $\mathbf{E} \equiv \mathbf{E}_{P \sim \mathscr{D}(\alpha^*)}$ and $T^*_{(j)}$'s are the ordered censored times.

Proposition 2: Let $0 = T^*_{(0)} < T^*_{(1)} < \ldots < T^*_{(m)} < T^*_{(m+1)} = \infty$ be the partition points on $\Re_+ = [0, \infty)$ and $\lambda_1, \ldots, \lambda_m$ be nonnegative integers. Then

$$c\, \mathbf{E}\left[\prod_{j=1}^{m} (P[T_j^*, \infty))^{\lambda_j} \right] = \prod_{j=1}^{m} B\left(\beta_j, \sum_{r=j+1}^{m+1} (\beta_r + \lambda_{r-1}) \right)$$

where $\beta_j = \alpha^*[T^*_{(j-1)}, T^*_{(j)})$ for $j = 1, 2, \ldots, m{+}1$, $c = \frac{\prod_{j=1}^{m+1}\Gamma(\beta_j)}{\Gamma(\alpha^*(\Re_+))}$, and $B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.

Proposition 3: Let $F = 1 - \bar{F}$ be a Dirichlet process with parameter $\alpha^* = \alpha + \sum_{i=1}^{n}\sum_{j=1}^{k_i}\delta_{t_{ij}}$. Then, for $u \in \Re_+$,

$$\mathscr{P}\{\bar{F}(u) \geq w | T_{ik_i+1} \in [T^*_i, \infty), i = 1, 2, \ldots, n\} = \frac{\mathbf{E}[I_{[F(u)\leq 1-w]}\prod_{i=1}^{n} P[T^*_i, \infty)]}{\mathbf{E}[\prod_{i=1}^{n} P[T^*_i, \infty)]}.$$

Proposition 4: If $T^*_{(l-1)} \leq u < T^*_{(l)}$ for $l = 1, 2, \ldots, m+1$, and $T^*_{(0)} = 0$ and $T^*_{(m+1)} = \infty$, then, for $\nu = 1, 2, \ldots$

$\mathbf{E}\left[(\bar{F}(u))^{\nu} | T_{ik_i+1} \in [T^*_i, \infty), i = 1, 2, \ldots, n\right]$

$$= \prod_{s=0}^{\nu-1}\left\{\left[\frac{\alpha(u,\infty) + s + Y^+(u)}{\alpha(\Re_+) + s + N}\right]\left[\prod_{j=1}^{l}\left\{\frac{\alpha[T^*_{(j)},\infty) + s + Y(T^*_{(j)})}{\alpha[T^*_{(j)},\infty) + s + Y(T^*_{(j)}) - \lambda_j}\right\}\right]\right\}$$

where $N = n + \sum_{i=1}^{n} k_i$ and λ_j is the number of events censored at time $T^*_{(j)}$.

Theorem 1, which provides the nonparametric Bayes estimator, follows immediately from Proposition 4 by letting $\nu = 1$.

Theorem 1: Let $F = 1 - \bar{F}$ have a Dirichlet process prior on $(\Re_+, \mathscr{B}_+)$ with parameter α. Then, under the loss function (4.8) with $w(t) = t$, the nonparametric Bayes (NPB) estimator of the survival function $\bar{F}(u)$ is given by

$\hat{\bar{F}}_{NPB}(u) = \mathbf{E}[\bar{F}(u) \mid T_{ij} = t_{ij}, T_{iK_i+1} \in [T^*_i, \infty), i = 1, \ldots, n; j = 1, \ldots, K_i]$

$$= \left\{\left[\frac{\alpha(u,\infty) + Y^+(u)}{\alpha(\Re_+) + N}\right]\prod_{j=1}^{l}\left\{\frac{\alpha[T^*_{(j)},\infty) + Y(T^*_{(j)})}{\alpha[T^*_{(j)},\infty) + Y(T^*_{(j)}) - \lambda_j}\right\}\right\} \tag{2.5}$$

for $u \in [T^*_{(l-1)}, T^*_{(l)})$, $l = 1, 2, \ldots, m+1$.

Though the proposed estimator (2.5) is notationally similar in form to the Susarla and Van Ryzin (1976) estimator in the single event setting, it does have intrinsic differences since in the recurrent event setting the data structure is different than for the single event setting. For instance, the risk process $Y^+(t)$ and $Y(t)$ as defined in (2.3) and (2.4), respectively are different than those of the Susarla and Van Ryzin

(1976). Moreover, frequentist properties of the estimator $\hat{\bar{F}}_{NPB}$ in (2.5), such as biases, variances, and RMSE's, are quite different, owing to the sum-quota constraint. For instance, this constraint induces dependencies on the observed gap-times for each unit and renders the K_i's to become informative about F. See, for example, the impact of these distributional properties for the PL-type estimator in Peña et al. (2001).

Theorem 2 develops the posterior variance of $\bar{F}(u)$ which is also useful for constructing credible intervals for $\bar{F}(u)$.

Theorem 2: Let $F = 1 - \bar{F}$ have a Dirichlet process prior on $(\Re_+, \mathscr{B}_+)$ with parameter α. Then, the posterior variance of the survival function $\bar{F}(u)$ is given by

$$\sigma^2_{NPB}(u) = \hat{\bar{F}}_{NPB}(u) \times$$
$$\left[\left\{\frac{\alpha(u,\infty) + Y^+(u) + 1}{\alpha(\Re_+) + N + 1}\right\} \prod_{j=1}^{l} \left\{\frac{\alpha[T^*_{(j)},\infty) + Y(T^*_{(j)}) + 1}{\alpha[T^*_{(j)},\infty) + Y(T^*_{(j)}) + 1 - \lambda_j}\right\} - \hat{\bar{F}}_{NPB}(u)\right] \quad (2.6)$$

for $u \in [T^*_{(l-1)}, T^*_{(l)})$, $l = 1, 2, \ldots, m+1$.

2.2.3 Posterior Measure

In the derivation of the Bayes estimator in (2.5), we avoided obtaining directly the posterior measure of $\bar{F}$, given the recurrent event data. Instead, we obtained the estimator by directly computing the posterior expected value of $\bar{F}$, and the resulting estimator is in closed-form and is exact. There is a way, however, to obtain an approximation of the Bayes estimate via sampling from the posterior measure. The posterior measure of $\bar{F}$ turns out to be a mixture of Dirichlet measures, though the mixing coefficients are somewhat cumbersome to obtain. In Theorem 3 below we present a representation of this posterior measure which will enable us to readily take samples from this posterior measure. We could then obtain the mean, standard deviations, and percentiles of these posterior samples, and via this approach we are able to construct approximate credible intervals for $\bar{F}$.

Theorem 3: Let $P \sim \mathscr{D}(\alpha)$ on $(\Re_+, \mathscr{B}_+)$ and $\mathscr{B}^* = (B_1, B_2, \ldots, B_m, B_{m+1})$ be a measurable partition of $\Re_+$. Then the posterior measure of P, given $\{T_{ij} = t_{ij}, T_{iK_i+1} \in [t_i^*, \infty), i = 1, 2, \ldots, n; j = 1, 2, \ldots, K_i\}$, satisfies

$$\mathscr{P}\left\{P(\mathscr{B}^*) \in \mathbf{B} \mid T_{ij} = t_{ij}, T_{ik_i+1} \in [t_i^*, \infty), i = 1, 2, \ldots, n; j = 1, 2, \ldots, k_i\right\}$$
$$\propto \int_{\mathbf{B}} y_{m+1}^{\alpha_{m+1}^*-1} \prod_{l=1}^{m} \left[y_l^{\alpha_l^*-1} \left(1 - \sum_{j=1}^{l} y_j\right)\right] d\mathbf{y},$$

where $P(\mathscr{B}^*) \equiv (P(B_1), P(B_2), \ldots, P(B_{m+1}))$ and $P(B_l) = y_l, \alpha_l^* = \alpha^*(B_l), B_l = [t_{(l-1)}^*, t_{(l)}^*), l = 1, 2, \ldots, m+1$ with $t_{(0)}^* = 0$ and $t_{(m+1)}^* = \infty$.

Note that if λ_l is the number of observations censored at $t_{(l)}^*$, then

$$\mathscr{P}\left\{P(\mathscr{B}^*) \in \mathbf{B} \mid T_{ij} = t_{ij}, T_{ik_i+1} \in [t_i^*, \infty), i = 1, 2, \ldots, n\right\}$$
$$\propto \int_{\mathbf{B}} y_{m+1}^{\alpha_{m+1}^*-1} \prod_{l=1}^{m} \left[y_l^{\alpha_l^*-1} \left(1 - \sum_{j=1}^{l} y_j\right)^{\lambda_l}\right] d\mathbf{y}. \quad (2.7)$$

Therefore, the density function associated with posterior measure (2.7) is

$$h^*(\mathbf{y}) \propto y_{m+1}^{\alpha_{m+1}^*-1} \prod_{l=1}^{m} \left[y_l^{\alpha_l^*-1} \left(1 - \sum_{j=1}^{l} y_j\right)^{\lambda_l}\right], \quad (2.8)$$

which is also proportional to the so called generalized Dirichlet distribution (see Connor and Mosimann (1969)). We may now take samples from $h^*(\mathbf{y})$ given in (2.8) to construct point-wise credible intervals for $\bar{F}(u)$. To sample from the posterior measure, following Grego et al. (2013), we consider the transformations

$$Z_l = Y_l + Y_{l+1} + \ldots + Y_{m+1}, l = 1, 2, \ldots, m+1.$$

Define,

$$W_l = \frac{Z_{l+1}}{Z_l}, l = 1, 2, \ldots, m.$$

Simplification yields that

$$Y_1 = 1 - W_1, Y_2 = W_1(1 - W_2), \ldots, Y_m = (1 - W_m) \prod_{j=1}^{m-1} W_j,\ Y_{m+1} = \prod_{j=1}^{m} W_j. \quad (2.9)$$

Straight-forward derivations show that $W_1, W_2, \ldots, W_m$ have independent Beta distributions with

$$W_1 \sim Beta(A_1, \alpha_1^*), W_2 \sim Beta(A_2, \alpha_2^*), \ldots, W_m \sim Beta(A_m, \alpha_m^*), \tag{2.10}$$

where $A_j = \alpha^*[t^*_{(j)}, \infty) + \sum_{j=1}^m \lambda_j, j = 1, 2, \ldots, m$. One may now take samples of $W_1, W_2, \ldots, W_m$, and then obtain $Y_1, Y_2, \ldots, Y_{m+1}$ using (4.14). The approximate posterior mean and point-wise credible intervals of $\bar{F}(u)$ could then be obtained and constructed, respectively.

Another analytical representation of the nonparametric Bayes estimator of F is given below. For any $j \in \{1, 2, \ldots, m, m+1\}$, let $u \in B_j$ and we want to estimate $F(u) \equiv P((-\infty, u])$. Note that for any $A \in \mathscr{B}_+, P(A) = \sum_{k=1}^{m+1} [P(B_k)P(A \mid B_k)]$. Since $P \sim \mathscr{D}(\alpha)$, then

$$P(\mathscr{B}^*) \text{ and } P_{B_1}, P_{B_2}, \ldots, P_{B_m}, P_{B_{m+1}} \text{ are independent and } P_{B_k} \sim \mathscr{D}(\alpha_{B_k}),$$

with $\alpha_{B_k}(A) = \alpha(A \cap B_k)$. Therefore, the nonparametric Bayes estimator of $F(u)$ is

$$\widehat{F}(u) = \sum_{l=1}^{j-1} \mathbf{E}[Y_l] + \mathbf{E}[Y_j] \left\{ \frac{\alpha^*(B_j \cap (0, u])}{\alpha^*(B_j)} \right\}, \tag{2.11}$$

where the expectation is with respect to the posterior measure given in Theorem 3,

$$\mathbf{E}[Y_j] = \mathbf{E}\left[(1 - W_j) \prod_{l=1}^{j-1} W_l\right] = \left[\frac{\alpha_j^*}{A_j + \alpha_j^*}\right] \prod_{l=1}^{j-1} \left[\frac{A_l}{A_l + \alpha_l^*}\right], \quad j = 1, 2, \ldots, m,$$

and

$$\mathbf{E}[Y_{m+1}] = \prod_{l=1}^{m} \left[\frac{A_l}{A_l + \alpha_l^*}\right].$$

2.2.4 Empirical Bayes Estimator of the Survival Function $\bar{F}$

Consider $(P_{ij}, T_{ij}),\ j = 1, 2, \ldots; i = 1, \ldots, n$, be a sequence of pairs of independent random elements. The random probability measures $\{P_{ij}\}$ have common prior probability measure P given by Dirichlet process $\mathscr{D}(\alpha)$. Given P_{ij}, T_{ij} has probability measure P_{ij}; see Robbins (1956), Korwar and Hollander (1976), and Susarla and Van Ryzin (1978) for the early developments and nonparametric setup of empirical Bayes approach. For $j = 1, 2, \ldots; i = 1, \ldots, n$, the common marginal survival function of the $\{T_{ij}\}$ is $\mathscr{P}(T_{ij} > t) = \mathbf{E}[\mathscr{P}(T_{ij} > t) \mid P_{ij}] = \mathbf{E}[P_{ij}(t, \infty)]$ $= \alpha(t, \infty)/\alpha(\Re_+) = \bar{\alpha}(t, \infty)$. Our goal here is to estimate $\bar{\alpha}$ empirically, where $\alpha = \alpha(\Re_+)\bar{\alpha}$. In life testing/survival analysis problems one of the common choices of $\bar{\alpha}$ is the Weibull or exponential survivor function; see Susarla and Van Ryzin (1976), Sethuraman and Hollander (2009). For example, let $\alpha(t, \infty) = \beta \exp(-t\theta)$. Then the marginal survivor function of T_{11}, $\mathscr{P}(T_{11} > t) = \alpha(t, \infty)/\alpha(\Re_+) = \bar{\alpha}(t, \infty) = \exp(-t\theta)$. For the recurrent event data, the analogue of the empirical estimator of survivor function is the Peña et al. (2001) PL-type estimator. We estimate θ by equating some percentile of the PL-type estimator survival curve to the $\bar{\alpha}$, since the PL-type estimator is estimating the common marginal survival function of $\{T_{ij}\}$. For instance, let M^* be the median estimate from the PL-type estimator survival curve. Then solving $\exp(-M^*\theta) = .5$, yields an estimate of θ, denoted by $\hat{\theta}$. One can have a maximum likelihood (ML) estimate of θ from the likelihood associated with the prior mean function. However, we observed that with an empirical estimate of θ, the resulting empirical Bayes estimator is robust, whereas with a ML estimate of θ, the resulting empirical Bayes estimator is not robust with respect to a mis-specified prior. This will be seen in the simulation studies section.

We also estimate $\beta = \alpha(\Re_+)$, the precision of the prior belief. To this end, let $T_1, T_2, \ldots, T_M$ denote the observed uncensored events of $\{T_{ij}\}$ and assume that α is non-atomic. Now, $P \sim \mathscr{D}(\alpha)$ implies that $P \mid (T_1, \ldots, T_{M-1}) \sim \mathscr{D}(\alpha + \sum_{i=1}^{M-1} \delta_{T_i})$. Thus

$$\mathscr{P}(T_M \in \{T_1, \ldots, T_{M-1}\} \mid T_1, \ldots, T_{M-1}) = \frac{\alpha(T_1, \ldots, T_{M-1}) + M - 1}{\alpha(\Re_+) + M - 1}$$
$$= \frac{M-1}{\alpha(\Re_+) + M - 1}.$$

Therefore, $\mathscr{P}(T_M \notin \{T_1, \ldots, T_{M-1}\} \mid T_1, \ldots, T_{M-1}) = \frac{\alpha(\Re_+)}{\alpha(\Re_+)+M-1}$. We define Bernoulli random variables $D_1 = 1$ and for $M = 2, 3, \ldots,$

$$D_M = \begin{cases} 0 & \text{if} \quad T_M \in \{T_1, \ldots, T_{M-1}\} \\ 1 & \text{if} \quad T_M \notin \{T_1, \ldots, T_{M-1}\} \end{cases}.$$

Using Theorem 25.3 of Sethuraman (2008), we know that $D_1, \ldots, D_M$ are independent, and $\mathscr{P}(D_M = 1 \mid T_1, \ldots, T_{M-1}) = \frac{\alpha(\Re_+)}{\alpha(\Re_+)+M-1}$. From Theorem 27.1 of Sethuraman (2008), it follows that

$$\frac{\sum_{i=1}^{M} D_i}{log(M)} \to \alpha(\Re_+) \quad \text{with probability 1.} \tag{2.12}$$

As a result, an estimate of β is given by $\hat{\beta} = \sum_{i=1}^{M} D_i / log(M)$, where $\sum_{i=1}^{M} D_i$ is the number of distinct uncensored observations. Thus, the resulting α with the $\hat{\theta}$ and $\hat{\beta}$ is an empirical estimate of α, e.g., $\hat{\alpha}(t, \infty) = \hat{\beta} \exp(-t\hat{\theta}))$ and the corresponding empirical Bayes estimator of $\bar{F}$ is given by

$$\hat{\bar{F}}_{NPEB}(u) = \left\{ \frac{\widehat{\alpha}(u, \infty) + Y^+(u)}{\widehat{\alpha}(\Re_+) + N} \prod_{j=1}^{l} \left\{ \frac{\widehat{\alpha}[T^*_{(j)}, \infty) + Y(T^*_{(j)})}{\widehat{\alpha}[T^*_{(j)}, \infty) + Y(T^*_{(j)}) - \lambda_j} \right\} \right\}.$$

Empirical Bayes estimation offers some safeguards against the possible misspecification of the prior measure. Instead of specifying the parameter(s) of the prior measure, we empirically estimate them by utilizing observed data. It will be seen in the simulation studies that even in the case of a mis-specified prior measure, the empirical Bayes estimator of $\bar{F}$ demonstrates smaller root-mean-squared error than the Bayes and PL-type estimators.

2.3 PL-Type Estimator is Limit of Bayes Estimator

We can recover the PL-type estimator in the recurrent event setting from our nonparametric Bayes (NPB) estimator (2.5). To this end, let $T'_i, i = 1, 2, \ldots, N$, denote the ordered (increasing magnitude) observed values of $T_{ij}, j = 1, 2, \ldots, K_i$, and $T_{iK_i+1} = \tau_i - S_{iK_i}$, $i = 1, 2, \ldots, n$, so that $0 \le T'_1 \le T'_2 \le, \ldots, \le T'_N$. Let

$N^\dagger(w) = \sum_{r=1}^{N} I(T_r' \leq w, \delta_r = 1)$, where $\delta_r = 1$ if T_r' is an uncensored (complete) observation, and 0 otherwise. Then the PL-type estimator in Peña et al. (2001) is

$$\hat{\bar{F}}_{PLE}(u) = \prod_{w \leq u} \left\{ 1 - \frac{\Delta N^\dagger(w)}{Y(w)} \right\} = \prod_{w \leq u} \left\{ \frac{Y^+(w)}{Y(w)} \right\},$$

where $Y^+(w)$ and $Y(w)$ are defined in (2.3) and (2.4), respectively.

Theorem 4: If $\alpha(\Re_+) \to 0$, then $\hat{\bar{F}}_{NPB}(u) \to \hat{\bar{F}}_{PLE}(u)$.

Moreover, we can also express $\hat{\bar{F}}_{NPB}(u)$ as a linear combination of the $\hat{\bar{F}}_{PLE}(u)$ and the prior mean function, $\bar{\alpha}$, as follows:

$$\begin{aligned}
\hat{\bar{F}}_{NPB}(u) = & \left\{ \frac{\alpha(u,\infty) + Y^+(u)}{\alpha(\Re_+) + N} \prod_{j=1}^{l} \left\{ \frac{\alpha(T^*_{(j)},\infty) + Y(T^*_{(j)})}{\alpha(T^*_{(j)},\infty) + Y(T^*_{(j)}) - \lambda_j} \right\} \right\} \\
= & \, \bar{\alpha}(u,\infty) \left[\left\{ \frac{\alpha(\Re_+)}{\alpha(\Re_+) + N} \right\} \prod_{j=1}^{l} \left\{ \frac{\alpha(T^*_{(j)},\infty) + Y(T^*_{(j)})}{\alpha(T^*_{(j)},\infty) + Y(T^*_{(j)}) - \lambda_j} \right\} \right] \\
& + \hat{\bar{F}}_{PLE}(u) \left[\left\{ \frac{N}{\alpha(\Re_+) + N} \right\} \prod_{j=1}^{l} \left\{ \frac{Y(T^*_{(j)}) - \lambda_j}{Y(T^*_{(j)})} \frac{\alpha(T^*_{(j)},\infty) + Y(T^*_{(j)})}{\alpha(T^*_{(j)},\infty) + Y(T^*_{(j)}) - \lambda_j} \right\} \right],
\end{aligned}$$

where $\bar{\alpha}(u,\infty)$ and $\hat{\bar{F}}_{PLE}(u)$, denote the prior mean function and PL-type estimator, respectively.

From the above representation it is clear that if $\alpha(\Re_+)$ is small relative to N, little weight is given to the prior estimate of $\bar{F}$, hence $\hat{\bar{F}}_{NPB}(u)$ and $\hat{\bar{F}}_{NPEB}(u)$ will be dominated by the estimate based on the data ($\hat{\bar{F}}_{PLE}(u)$) rather than $\bar{\alpha}(u,\infty)$. Note also that with this representation, we are immediately able to get the result that

$$\hat{\bar{F}}_{PLE}(u) = \lim_{\alpha(\Re_+) \to 0} \hat{\bar{F}}_{NPB}(u).$$

2.4 Simulation Studies

Simulation studies were carried out to examine the biases and root-mean-squared errors (RMSE's) of the nonparametric Bayes estimator (labeled NPBayes) $\hat{\bar{F}}_{NPB}(u)$ and empirical Bayes estimator (labeled EmpBayes) $\hat{\bar{F}}_{NPEB}(u)$, as well as the Product-Limit estimator (labeled PLE) $\hat{\bar{F}}_{PLE}(u)$ under the IID inter-event time model. Simulated biases and RMSE's were obtained for equally spaced values of $u_i, i = 1, 2, \ldots, 20$, of duration time u over the interval [0,1] based on 1000 replications for n=20, where the inter-event times were generated from an exponential distribution with parameter $\theta = 6$. To compute $\hat{\bar{\bar{F}}}_{NPB}(u)$ we use the prior measure $\alpha(u, \infty) = \beta \exp(-(\theta u)^\gamma)$ with $\theta = 6, \gamma = 1$, and $\beta = 20$, [that is, β times an exponential survivor function with parameter θ] and where β may be viewed as the precision of the prior measure. Note that in this case $\bar{\alpha}$ will coincide with the true $\bar{F}$. Empirical estimates, denoted by $\hat{\theta}$ and $\hat{\gamma}$, of the parameters θ and γ of the prior measure α are obtained by equating the marginal survival function of T_1, $\mathscr{P}(T_1 > u) = \exp(-(\theta u)^\gamma)$ to the approximately 50^{th} and 25^{th} percentiles of the $\hat{\bar{\bar{F}}}_{PLE}(u)$ survival curve. One can choose different percentiles to equate with the marginal distribution since the resulting empirical estimates of $\bar{F}$ are not sensitive to the different choices of percentiles. It is also easy to have ML estimates of θ and γ, but the corresponding empirical Bayes estimator is not robust in the case of a mis-specified prior. More specifically the RMSE's of the resulting empirical Bayes estimator with the ML estimate(s) are higher than the empirical Bayes estimator with $\hat{\theta}$ and $\hat{\gamma}$ in the case of a mis-specified prior (graph is not included here). We kept a record of the empirical estimates of θ and γ in the simulation studies which demonstrates that $\hat{\theta} \dot{\sim} N(6.02, .34)$ and $\hat{\gamma} \dot{\sim} N(1.01, .15)$ for 1000 replicates where the true parameters are $\theta = 6$ and $\gamma = 1$. An estimate of β is obtained by $\hat{\beta} = \sum_{i=1}^{N} D_i / log(N)$ using (2.12). With the resulting estimate of θ, γ, and β, the empirical Bayes estimate, $\hat{\bar{\bar{F}}}_{NPEB}(u)$, is obtained by replacing $\alpha(u; \theta, \gamma)$ by $\widehat{\alpha}(u; \theta, \gamma) = \hat{\beta} \exp(-(u\hat{\theta})^{\hat{\gamma}})$ in (2.5). By examining Figure 2.2, it is evident that both the $\hat{\bar{\bar{F}}}_{NPB}(u)$ and $\hat{\bar{\bar{F}}}_{NPEB}(u)$ possess smaller biases and smaller RMSE's than the $\hat{\bar{\bar{F}}}_{PLE}(u)$ when the mean of the prior measure $\bar{\alpha}(u, \infty)$ coincides with the true $\bar{F}(u)$ or the mean of the prior measure does not differ significantly from the true distribution function.

We also investigated the biases and RMSE's in the case of a mis-specified prior

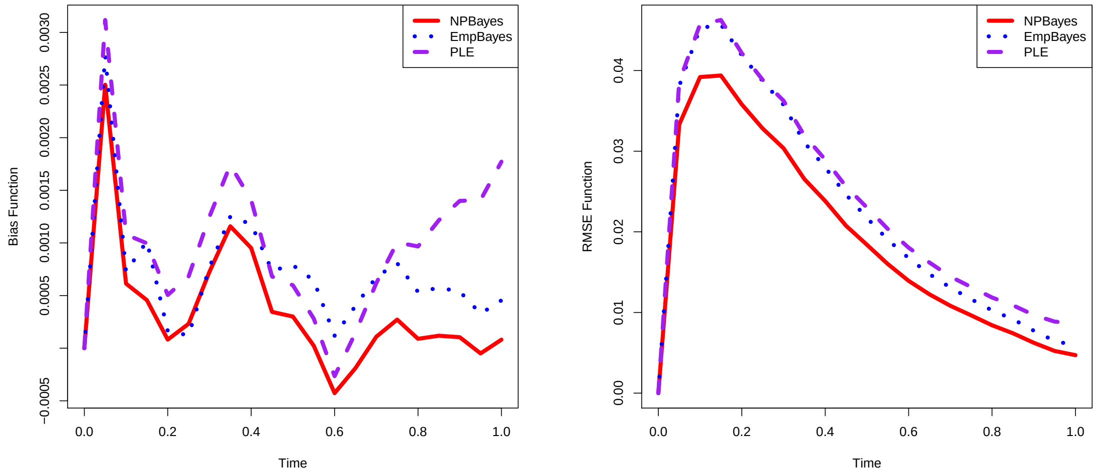

Figure 2.2: Simulated biases and RMSE's of $\hat{\bar{\bar{F}}}_{NPB}(u)$, $\hat{\bar{\bar{F}}}_{NPEB}(u)$, and $\hat{\bar{\bar{F}}}_{PLE}(u)$. Simulation parameters were $n = 20$ and $\theta = 6$ with 1,000 replications.

parameter, that is, when $\bar{\alpha}$ differs from the true data generating distribution $\bar{F}$. In particular, IID gap-times (inter-event times) were generated from the Weibull model with scale parameter $\theta = 6$ and shape parameter $\gamma = 2$. However, we assume a prior measure given by $\alpha(u, \infty) = \beta \exp(-\theta u)$ (β times an exponential survivor function with parameter θ) instead of $\alpha(u, \infty) = \beta \exp(-(\theta u)^{\gamma})$. Figure 2.3 compares the three estimators by demonstrating the effect of the precision parameter β on biases and RMSE's when the prior measure is mis-specified. By examining Figure 2.3, it is obvious that $\hat{\bar{\bar{F}}}_{PLE}$ exhibits negligible biases which are smaller in magnitude than the biases of $\hat{\bar{\bar{F}}}_{NPB}$ and approximately equal to the biases of $\hat{\bar{\bar{F}}}_{NPEB}$. However, with $\beta = 1$, the RMSE's of $\hat{\bar{\bar{F}}}_{NPB}(u)$ are smaller than $\hat{\bar{\bar{F}}}_{PLE}(u)$ for larger values u and almost identical for smaller values of u. In addition, the RMSE's of the empirical Bayes estimator $\hat{\bar{\bar{F}}}_{NPEB}(u)$ (Figure 2.3) are smaller than the $\hat{\bar{\bar{F}}}_{PLE}(u)$ for a mis-specified prior.

As suggested by one of the reviewers and also recommended by the associate editor, to investigate further the robustness of our proposed estimator we generate

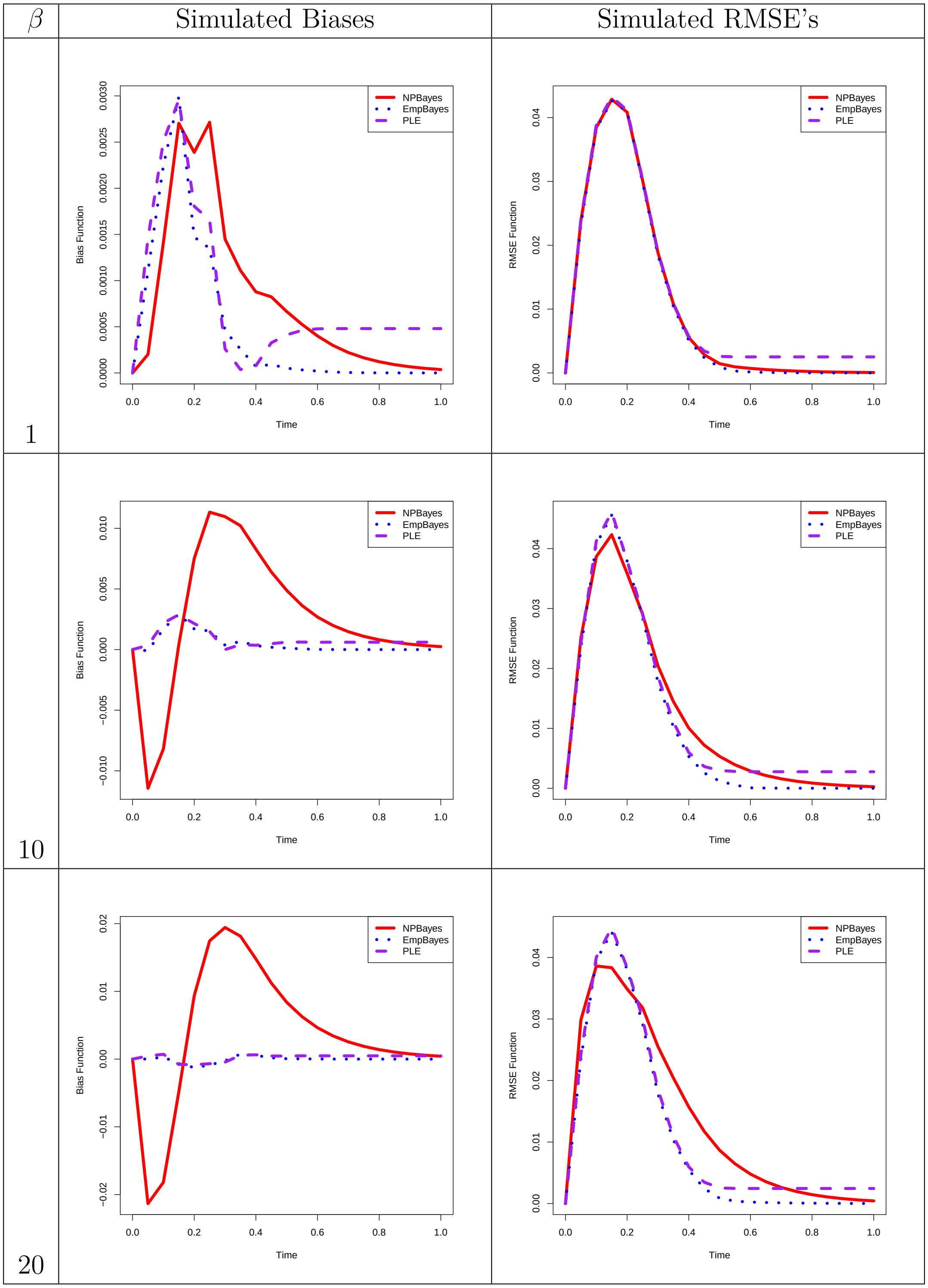

Figure 2.3: Simulated biases and RMSE's of $\widehat{\widehat{F}}_{NPB}(u)$, $\widehat{\widehat{F}}_{NPEB}(u)$, and $\widehat{\widehat{F}}_{PLE}(u)$. Simulation parameters were $n = 20$, $\theta = 6$ and $\gamma = 2$ (Weibull (6,2)) with 1,000 replications. Mis-specified prior parameters $\alpha(u, \infty) = \beta \exp(-\theta u)$, with $\theta = 6$ and $\beta = 1$, 10, 20 respectively.

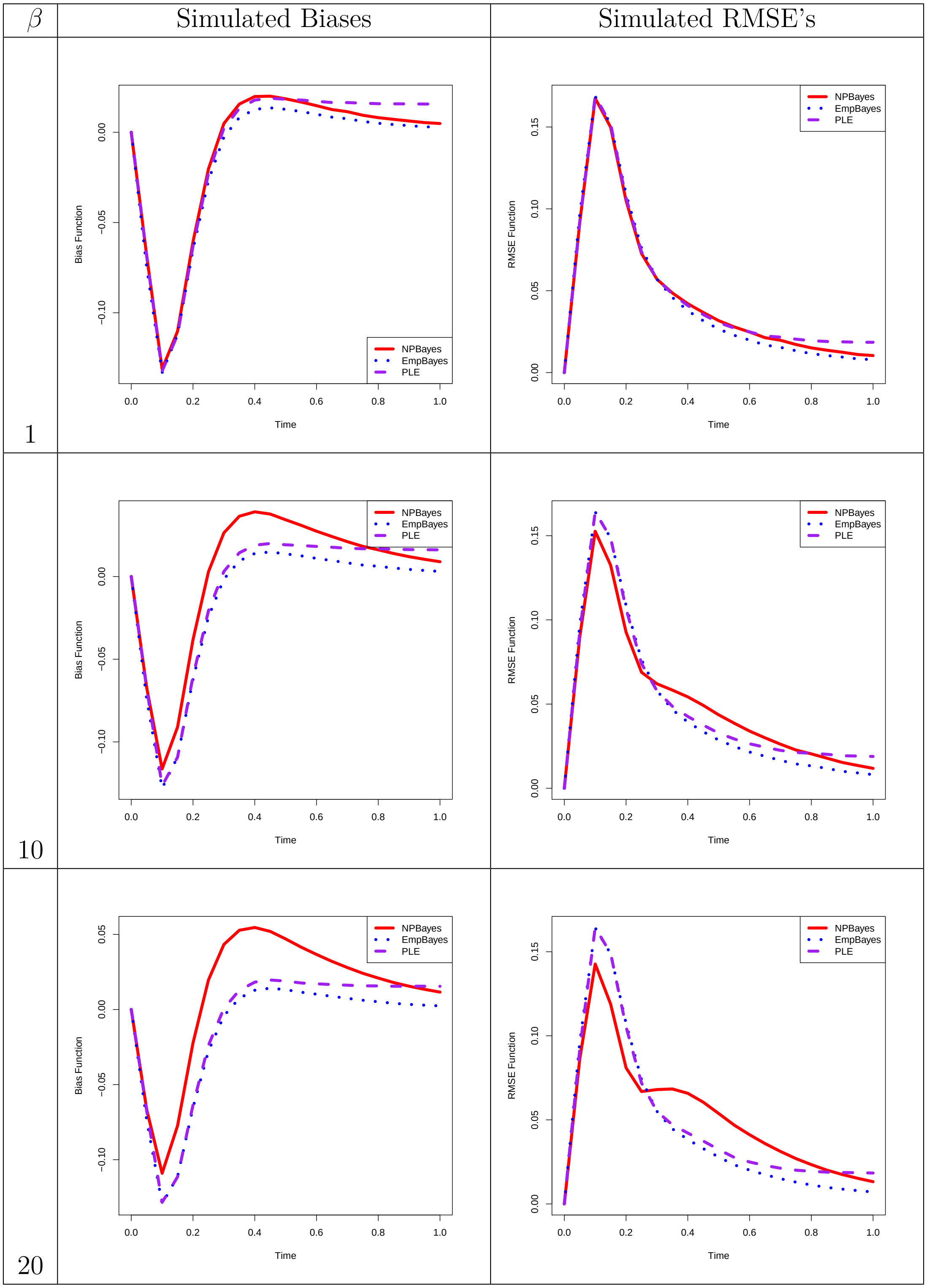

Figure 2.4: Data: Correlated gap-times with sample size $n = 20$ and 1,000 replications. Simulated biases and RMSE's of the estimators $\widehat{\widehat{F}}_{NPB}(u)$, $\widehat{\widehat{F}}_{NPEB}(u)$ and $\widehat{\widehat{F}}_{PLE}(u)$. Mis-specified prior parameters $\alpha(u,\infty) = \beta\exp(-\theta u)$, with $\theta = 3$ and $\beta =$ 1, 10, 20 respectively.

data such that for each unit gap-times are correlated by using a frailty model. More precisely, let $W_1, W_2, \ldots, W_n$ be a vector of IID positive-valued random variables from $Gamma(\nu, \nu)$ with unit mean and variance $1/\nu$. Given the unobserved frailty variable W_i, we assume that the successive inter-event times (gap-times) for the ith unit, denoted by $\{T_{ij}, j = 1, 2, \ldots\}$, are IID nonnegative random variables with a common distribution function $F(\cdot \mid W_i)$ with $\bar{F}(u \mid W) = [\bar{F}_0(u)]^W$, where $\bar{F}_0(u)$ is the baseline survivor function. Note that unconditionally, the gap-times are dependent with the marginal survivor function of T_{ij} being

$$\bar{F}(u) = E[\bar{F}_0(u)^W] = \left[\frac{\nu}{\nu + \Lambda_0(u)}\right]^{\nu},$$

where $\Lambda_0(u)$ is the baseline cumulative hazard function of $\bar{F}_0$. For the ith unit, given W_i, gap-times are generated from a Weibull distribution with shape parameter 2 and scale parameter $1/6W_i$. For Biases and RMSE's functions we use our proposed estimator based on the assumptions that gap-times $\{T_{ij}\}$ are IID from F and hence the model is mis-specified as the generated gap-times are correlated. In addition, we assume a prior measure given by $\alpha(u, \infty) = \beta \exp(-\theta u)$ with $\theta = 3$ (β times an exponential survivor function with parameter θ) instead of $\alpha(u, \infty) = \beta \exp(-(\theta u)^{\gamma})$ with $\theta = 6$ and $\gamma = 2$. Thus, we have model mis-specification as well as mean function of the prior measure mis-specification. Even with these two types of mis-specifications, Figure 2.4 demonstrates that $\hat{\bar{F}}_{NPB}(u)$ and $\hat{\bar{F}}_{NPEB}(u)$ have equal or smaller RMSE's than $\hat{\bar{F}}_{PLE}(u)$ for smaller values of the precision parameter β. Moreover, $\hat{\bar{F}}_{NPEB}(u)$ demonstrates smaller RMSE's than $\hat{\bar{F}}_{PLE}(u)$ even for larger values of precision parameter. This simulation study further demonstrates that the proposed estimators, in particular, $\hat{\bar{F}}_{NPEB}(u)$, are robust estimators.

When the prior measure α is mis-specified, a larger magnitude of the precision parameter β produces larger biases and RMSE's for the $\hat{\bar{F}}_{NPB}(u)$. Therefore, if one is not confident about his/her prior measure belief, a smaller value of β can offset the possible effect of a mis-specification. Simulation results presented in Figure 2.3 and

Figure 2.4 demonstrate that the nonparametric Bayes and empirical Bayes are robust estimators in the sense that they do not suffer significantly due to a mis-specification of the prior measure.

2.5 Application to Small Bowel Motility Data

We implemented the three survivor function estimators $\hat{\bar{\bar{F}}}_{NPB}(u)$, $\hat{\bar{\bar{F}}}_{NPEB}(u)$, and $\hat{\bar{\bar{F}}}_{PLE}(u)$ for the gastroenterology data from a study concerning the small bowel motility (see Husebye et al. (1990)). The data and its description are available in Aalen and Husebye (1991), where they estimated the mean length of the migratory motor complex (MMC) period. Peña et al. (2001) also applied their PL-type estimator, $\hat{\bar{\bar{F}}}_{PLE}(u)$, to the same data. In order to apply the estimators to the gastroenterology data, it is necessary to check the IID assumption. However, the renewal assumption for each subject in the above study is reasonable as established by Aalen and Husebye (1991, p. 1229) where they stated that the "consecutive MMC periods for each individual appear (to be) approximate renewal process."

We computed the survivor function estimates $\hat{\bar{\bar{F}}}_{NPB}(u)$, $\hat{\bar{\bar{F}}}_{NPEB}(u)$, and $\hat{\bar{\bar{F}}}_{PLE}(u)$ of the inter-event time distribution for the gastroenterology data (MMC data). The resulting estimates are presented in Figure 2.5. Though the graphical representation of all the three estimates of the survivor function looks very similar, a magnified view (graph is not given here) demonstrates that the estimates of $\hat{\bar{\bar{F}}}_{NPB}(u)$ and $\hat{\bar{\bar{F}}}_{NPEB}(u)$ are smoother than the estimate of $\hat{\bar{\bar{F}}}_{PLE}(u)$ in the sense that the former two are non-step functions with smaller jump sizes than the latter. To obtain the nonparametric Bayes survivor function estimate we assumed a Dirichlet process with prior measure $\alpha(u, \infty) = \beta \exp\{-(u/\theta)^{\gamma}\}$ with $\beta = 20$, $\theta = 120$, and $\gamma = 2$ (Figure 2.5, left graph). One can choose any other values of the parameters θ and γ with the associated value of the precision parameter β. For empirical estimates, equating the marginal survivor function of $T, \mathscr{P}(T > t) = \exp\{-(u/\theta)^{\gamma}\}$ with the 50^{th} and 25^{th}

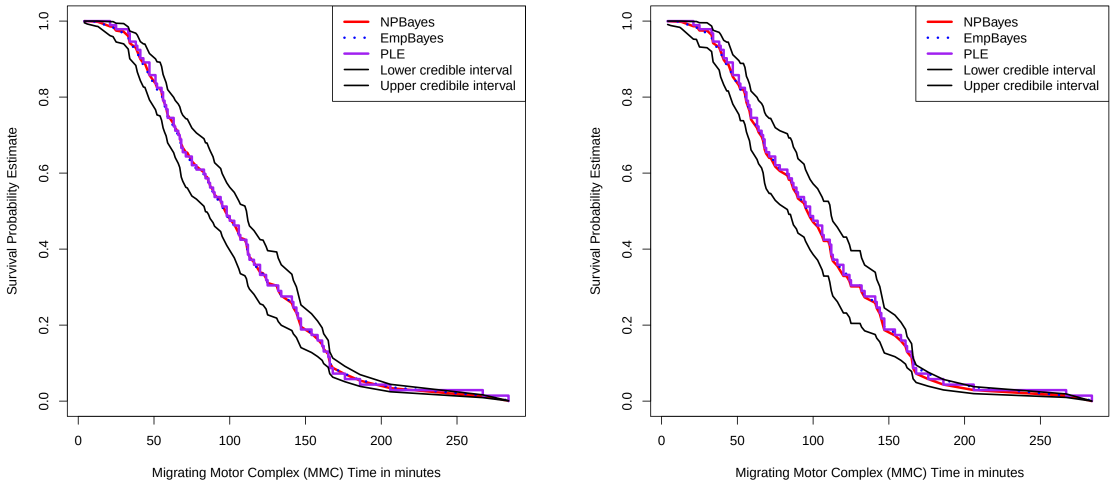

Figure 2.5: Survival function of NPBayes, Empirical Bayes and PLE and 95% pointwise credible intervals for the *MMC data set* with prior $\alpha(u, \infty) = \beta \exp\{-(u/\theta)^{\gamma}\}$ with $(\beta, \theta, \gamma) = (20, 120, 2)$ (left side graph), $(\beta, \theta, \gamma) = (1, 60, 1)$ (right side graph), and $\hat{\alpha}(u, \infty) = \hat{\beta} \exp\{-(u/\hat{\theta})^{\hat{\gamma}}\}$ with $\hat{\beta} = 15$, $\hat{\theta} = 119$, and $\hat{\gamma} = 1.76$.

percentiles of the $\hat{\bar{\bar{F}}}_{PLE}(u)$ survival curve we obtain $\hat{\theta} \approx 119$ and $\hat{\gamma} \approx 1.76$. Again using (2.12) we estimate β by $\hat{\beta} = \sum_{i=1}^{N} D_i / log(N) \approx 15$, where $\sum_{i=1}^{N} D_i$ denotes the number of distinct failure times. The resulting estimate $\hat{\bar{\bar{F}}}_{NPB}(u)$ with the estimated parameters, $\hat{\beta}, \hat{\theta}$, and $\hat{\gamma}$ is the empirical Bayes estimate $\hat{\bar{\bar{F}}}_{NPEB}(u)$.

Using the estimates of the survival curve of the inter-event time distribution, we obtained the corresponding mean MMC period length (in minutes) to be $\mu(\hat{\bar{\bar{F}}}_{PLE})$ $= 104.12$, $\mu(\hat{\bar{\bar{F}}}_{NPB}) = 102.83$, and $\mu(\hat{\bar{\bar{F}}}_{NPEB}) = 104.76$. For this data set, the three methods of analysis yielded almost the same estimates for the mean MMC period length. We also used another prior parameter $\alpha(u, \infty) = \beta \exp\{-(u/\theta)^{\gamma}\}$ with $\beta = 1$, $\theta = 60$, and $\gamma = 1$, which is significantly different from the previous prior measure. However, the resulting nonparametric Bayes estimate, empirical Bayes estimate, and PL-type estimate look almost similar as shown in Figure 2.5 (right graph). In this choice of parameters we assumed that we are not confident enough about our prior measure and hence we set the precision of the prior parameter to be small, namely

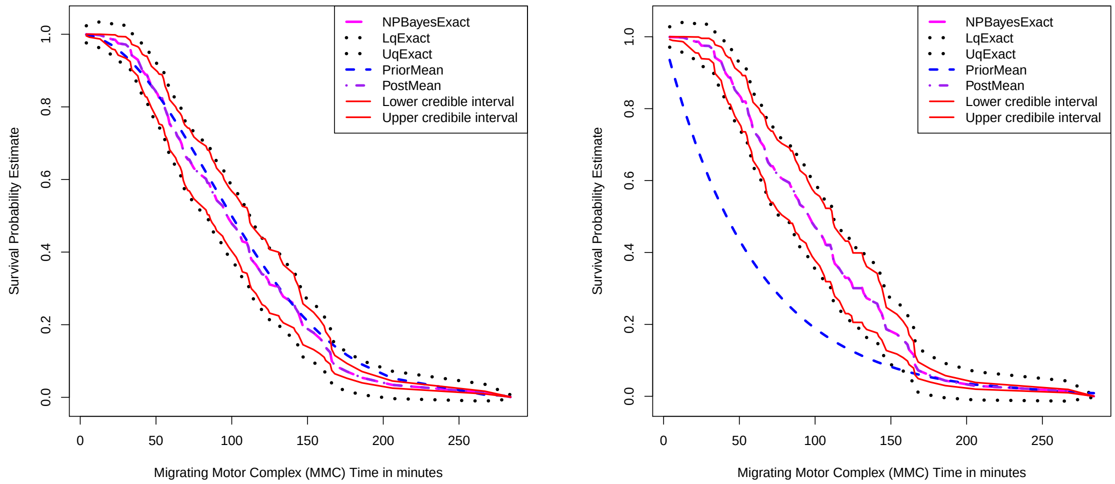

Figure 2.6: Survival function of NPBayes (labeled NPBayesExact), prior mean function (labeled PriorMean), posterior mean function (labeled PostMean, based on sample), and 95% pointwise credible intervals (dotted line: using exact variance, solid line: using sample from posterior) for the *MMC data set* with prior $\alpha(u, \infty) = \beta \exp\{-(u/\theta)^\gamma\}$ with $(\beta, \theta, \gamma) = (20, 120, 2)$ (left side graph), $(\beta, \theta, \gamma) = (1, 60, 1)$ (right side graph).

$\beta = 1$. The mean MMC period length of $\hat{\bar{F}}_{NPB}$ for this choice of prior is 103.60, which is close to the other estimates.

In Figure 2.5, 95% point-wise credible intervals are also plotted which are based on 1000 samples of W's and thus of Y's from the posterior measure given in Theorem 3. In Figure 2.6, we observe that the posterior mean of $\bar{F}$ based on the samples from the posterior measure utilizing (4.14), (2.10), and (2.11) and from the closed-form expression of the nonparametric Bayes estimator (2.5) of $\bar{F}$ coincides with each other as expected. We also compute point-wise credible interval (dotted lines) using $\hat{\bar{F}}(t) \pm 1.96\hat{\sigma}$, where $\hat{\sigma}$ as given in (2.6). It turns out that the width of the point-wise credible interval is narrower when we take samples from the posterior measure rather than using the exact variance formula (2.6). Figure 2.6 demonstrates that credible intervals based on samples from the posterior measure is more precise and accurate.

2.6 Concluding Remarks

The nonparametric Bayes estimator $\hat{\bar{F}}_{NPB}(u)$ developed in this paper extended the single-event Susarla and Van Ryzin (1976) nonparametric Bayes estimator and serves as a nonparametric Bayesian counterpart of the PLE in Peña et al. (2001) under the recurrent event settings. Both the nonparametric Bayes and empirical Bayes estimators are smoother than the PL-type estimator in that the jumps at the time of events (uncensored time), are smaller for the nonparametric Bayes and empirical Bayes than the PL-type estimator. Also, the PL-type estimator is a step function, while the nonparametric Bayes and empirical Bayes estimators are non-step (piece-wise non-constant) function with jump discontinuities at the time of events. In addition, RMSE's are smaller for Bayes and empirical Bayes estimators than the PL-type estimator when the mean of the prior is approximately close to the true distribution function, or, for other choice of prior with smaller value of precision parameter β (e.g. $\beta = 1$).

To compute the PL-type estimator $\hat{\bar{F}}_{PLE}(u)$, one need not know the exact times of censoring (in the ordered data set), but rather one only needs the number of observations censored in between two uncensored observations. One of the consequences of this is that, given the PL-type estimator, it is not possible to recover the exact censoring times. In contrast, to compute $\hat{\bar{F}}_{NPB}(u)$ and $\hat{\bar{F}}_{NPEB}(u)$, one needs the number observations censored in between uncensored observations as well as the *exact* times of censoring. Thus, the Bayes and empirical Bayes estimators accommodate the *exact* times of censoring. Note that given the nonparametric Bayes estimator, it is possible to recover the distinct censoring times.

Except in the simulation studies, we did not consider here the models for correlated inter-event gap-times since the manuscript is already long. In a forthcoming manuscript we will be considering nonparametric Bayesian inference with correlated gap-times and in the presence of covariates.

Chapter 3

Semiparametric Bayes Inference of Gap-Time Distribution with Recurrent Event Data [1]

Abstract

Recurrent event data arise from a wide variety of studies/fields such as clinical trials, epidemiology, public health, biomedicine (e.g. repeated heart attack, repeated tumor occurrences of a cancer patient). Semiparametric Bayes inference of the gap-time survivor function with the effect of covariates of a correlated recurrent event in the presence of censoring is considered. A frailty model is considered to allow the association between inter-occurrence gap-times. We assume that for a subject or unit given the unobserved frailty variable $W = w$, the inter-occurrence gap-times $\{T_j,\ j \geq 1\}$ are IID with some distribution function $F(\cdot \mid W = w)$. In our procedure, we assign a Gamma process prior on the baseline cumulative hazard function Λ_0 and parametric prior distributions on the finite dimensional parameters associated with covariates and frailty. We derive the conditional posterior distributions from the joint posterior distribution of the unknown parameters of interest and employ Gibbs sampler techniques to obtain samples from the joint posterior distribution. Simulation studies demonstrate the effectiveness of the developed method. The Peña et al.'s (2001) estimator of $\bar{F}$ for correlated recurrent event data without covariates is a special case of our developed estimator with the precision parameter of the gamma process prior tending to zero.

[1] A.K.M. Fazlur Rahman and Edsel A. Peña. To be submitted to *Biometrics.*

3.1 Introduction

Recurrent events frequently arise in biomedical studies involving subjects with some treatable diseases (e.g. asthma, leukemia, tumors) with or without immediate risk of death. Examples of recurrent events are re-occurrence of a tumor in bladder cancer patients (Byar (1980)), repeated mammary tumor occurrences in carcinogenesis studies (Gail et al. (1980)), successive seizures in epileptic patients (Albert (1991)), and repeated hospitalizations of a patient with cardiovascular disease. The primary research interest in recurrent event analysis is to investigate whether the treatment is effective in reducing the hazard rate of further re-occurrence of an event.

Several models and methods have been considered for the analysis of recurrent event data. These models include complete intensity approach (e.g., Andersen et al. (1993); Prentice et al. (1981)), the marginal rate approach (Pepe and Cai (1993); Lawless and Nadeau (1995b); Lin et al. (2000)), and the inter-event gap-time approach (e.g. Peña et al. (2001); Peña et al. (2007)). The main differences between the various methods proposed is the function that is modeled or the parameter of interest. Methods based on the gap-time formulation are intuitively more appealing because they address questions, such as the distribution of time to next event occurrence for a subject who has already experienced some events.

The analysis of recurrent event data based on gap-times yields estimates of regression coefficients as well as an estimate of the gap-time distribution function. In many biomedical/epidemiological applications we may only be interested in the effects of covariates on the event occurrences. However, in reliability settings one may also be interested in the baseline survivor or baseline cumulative hazard function as well as the effect of covariates on the hazard rate of event occurrences. The Bayesian paradigm provides more general estimators in the sense that nonparametric Bayes estimators often are in the form of a linear combination of the prior measures and

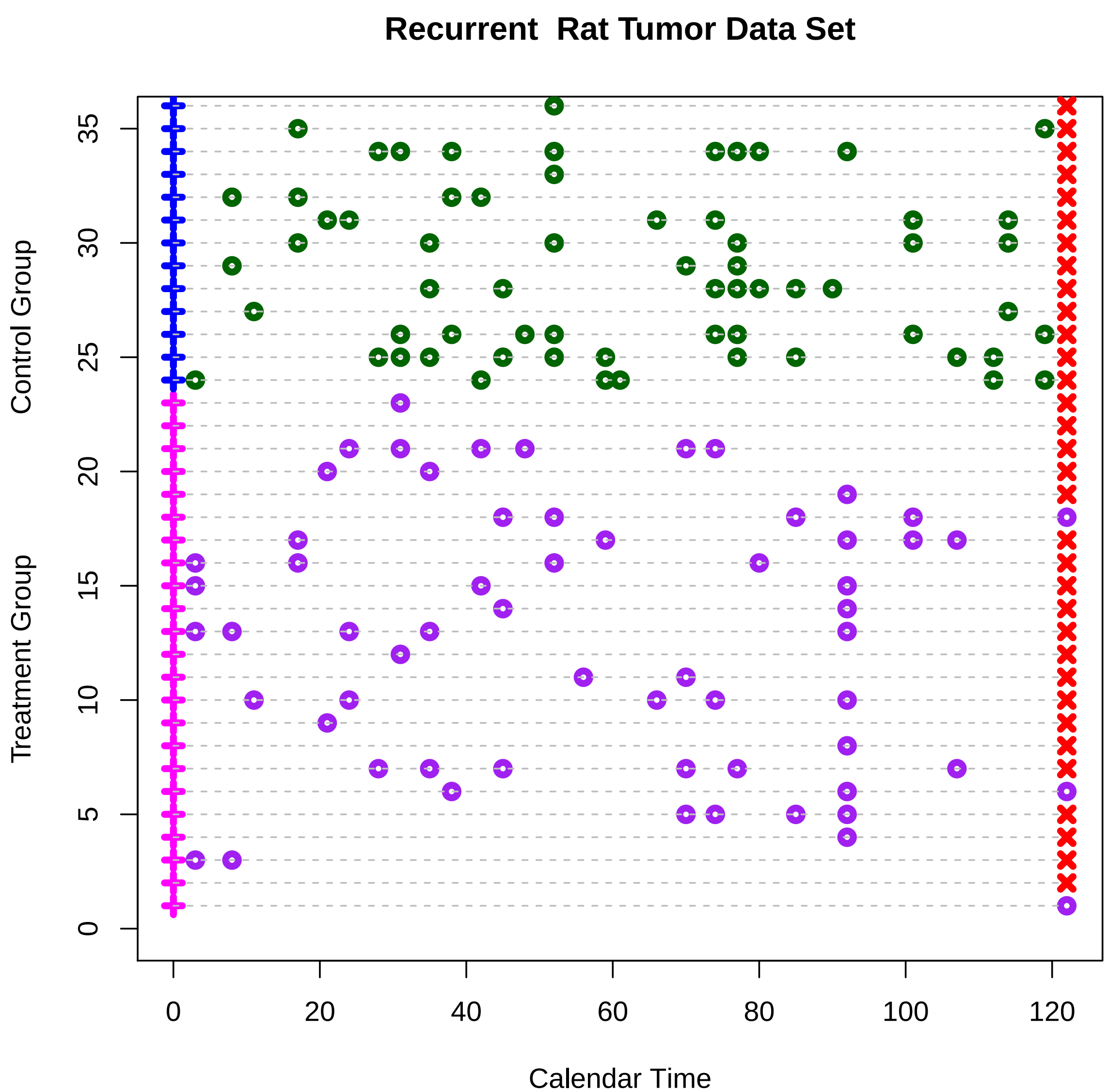

Figure 3.1: Graphical representation of a recurrent event mammary tumors dataset (Gail et al. (1980)). A sample of the mammary tumors data where first 23 units are in the treatment group and the remaining are in the control group.

the corresponding nonparametric estimators. Rahman et al. (2014) considered nonparametric Bayes estimation of a gap-time distribution with recurrent event data assuming the gap-times are independent and identically distributed (IID) from some distribution function F. In biomedical applications IID assumptions are somewhat restrictive as the gap-times could be correlated.

In this chapter, we consider semiparametric Bayesian inference of correlated gap-times with recurrent event data. Consider an arbitrary subject (subscript omitted) in the study for the occurrence of a recurrent event over the monitoring period $[0, \tau]$, where τ could be a pre-specified study termination time or some other random termination time. Denote by $S_1, S_2, S_3, \ldots,$ the successive calendar times of event occurrences and $T_1, T_2, T_3, \ldots,$ the successive inter-event gap-times of the event occurrences, so that

$$T_j = S_j - S_{j-1}, j = 1, 2, 3, \ldots, \text{ with } S_0 \equiv 0.$$

Denote by

$$K = \max\{k \in \{0, 1, 2, \ldots, \} : S_k \leq \tau\}$$

the number of event occurrences for a subject within the monitoring period. It is assumed that gap-times and monitoring times are mutually independent. Let W be an unobserved random variable from the Gamma distribution with shape and rate parameter equal to ν. A frailty model is considered to accommodate the correlation between gap-times of the recurrent event for a subject. That is, given $W = w$, we assume that gap-times $T_j, j = 1, 2, 3, \ldots,$ are IID from $F(\cdot \mid W = w)$. Denote the q-dimensional observable covariate vector by $\mathbf{X} = (X_1, X_2, \ldots, X_q)$.

We consider the intensity function defined by

$$\lambda(t \mid W = w, \mathbf{X} = \mathbf{x}) = \lambda_0(t) w \exp(\beta^T \mathbf{x}),$$

where $\lambda_0(t)$ is the baseline hazard function of the distribution F_0 and β is a vector of regression coefficients. The parameters of interest are the baseline cumulative hazard function $\Lambda_0(t) = \int_0^t \lambda_0(u)du$, regression parameter associated with covariates β, and the frailty parameter ν. The Baseline hazard function $\lambda_0(\cdot)$ is specified non-parametrically while β and ν are finite dimensional parameters. In our Bayesian approach, following Kalbfleisch (1978) and Sinha (1993), we assign a gamma process prior on $\Lambda_0(\cdot)$ and parametric prior distributions such as the multivariate normal and Gamma on β and ν, respectively. The conditional posterior distributions of $\Lambda_0(\cdot), \beta$, and ν are derived from the joint posterior distribution. Though W is unobservable, we need to replace it by an estimator because the conditional distribution of $\Lambda_0(\cdot)$, and thus the closed form posterior mean of $\Lambda_0(\cdot)$, involve W. We employ Gibbs sampling techniques to sample from the joint posterior distribution. Under an integrated squared error loss function posterior means are our Bayes estimators. Credible intervals of the parameters of interest are readily available from the posterior samples.

The Breslow-Aalen type estimators of the baseline cumulative hazard function can be recovered as a limiting case of our proposed estimator. The proposed estimators are robust in the sense that the parameter estimates is not sensitive to choice of the prior distribution. Simulation studies demonstrate the effectiveness of the developed procedure. Biases and root mean-squared errors of the baseline survivor functions are also examined for different combinations of parameters and sample sizes.

The rest of this chapter is organized as follows. In Section 2, we describe the frailty model for the gap-time distribution. In Section 3, we define notation, derive likelihood and conditional posterior distribution of the parameters of interest with MCMC tools. Section 4 includes simulation studies where we assess performance of the proposed estimators in terms of their bias, standard deviation and MCMC convergence. In Section 5, we use bladder cancer data analyzed by Wei et al. (1989) and Gail et al. (1980)'s mammary tumors data to illustrate our methodology. Section 6 provides some concluding thoughts.

3.2 Frailty Model

Consider n units are in the study and let $W_1, W_2, \ldots, W_n$ be independent and identically distributed (IID) positive-valued random variables known as frailties from a parametric distribution $H(w;\nu)$. More precisely, the W_i's are assumed to be an IID gamma variables with unit mean and variance $1/\nu$; that is $W_i \sim Ga(\nu,\nu), i = 1,2,\ldots,n$. Note that the mean of the W_i is taken to be 1 to make ν identifiable. Given the unobserved frailty variable W_i, we assume that the successive inter-event times (gap-times) for the ith unit, denoted by $\{T_{ij}, j = 1,2,\ldots\}$, are IID nonnegative random variables with a common distribution function

$$\bar{F}(t \mid W = w) = \left[\bar{F}_0(t)\right]^w = \exp(-w\Lambda_0(t)),$$

where $\bar{F}_0(t)$ and $\Lambda_0(t) = \int_0^t \lambda_0(u)du$ with $\lambda_0(t) = \frac{dF_0(t)}{\bar{F}_0(t)}$ are the baseline survivor function and cumulative hazard function, respectively. Note that conditionally, gap-times $\{ T_{ij} \}$ are independent, but unconditionally they are dependent. The ith unit will be observed over $[0,\tau_i]$ where $\tau_1, \tau_2, \ldots, \tau_n$ are IID with a common distribution function G. It is assumed that τ_i and $\{T_{ij}, j = 1,2,\ldots\}$ are mutually independent. The marginal survivor function of $\{T_{ij} \}$ is therefore given by

$$\bar{F}(t) = E[\bar{F}_0(t)^W] = \left[\frac{\nu}{\nu + \Lambda_0(t)}\right]^\nu .$$

A smaller value of ν is an indication of stronger correlation between gap-times whereas a larger value of ν suggests weaker correlation between gap-times. For instance, when gap-times are governed by the exponential distribution with $\Lambda_0(t) = \eta t$, then,

$$Corr(T_{11}, T_{12}) = 1/\nu, \nu > 2.$$

Similarly, in the presence of observable covariates $\mathbf{X} = \mathbf{x}$, given $W = w$, we assume $\{T_{ij}, j = 1,2,\ldots,\}$ are IID with common distribution

$$\begin{aligned}\bar{F}(t \mid W = w, \mathbf{X} = \mathbf{x}) &= \left[\bar{F}_0(t)\right]^{w\exp(\beta^T\mathbf{x})} \\ &= \exp\left[-w\exp(\beta^T\mathbf{x})\Lambda_0(t)\right].\end{aligned}$$

Therefore we consider the intensity function defined by

$$\lambda_i(t \mid W_i, X_i) = W_i \lambda_0(t) \exp(\beta^T X_i), \tag{3.1}$$

where β is a vector of parameters associated with the covariates. For the ith unit the number of observed event occurrences is

$$K_i = \max\{k \in \{0, 1, \ldots\} : S_{ik} \leq \tau_i\},$$

where $S_{i0} = 0$ and $S_{ik} = \sum_{j=1}^{k} T_{ij}, k = 1, 2, \ldots,$ and the observable random vector is

$$D_i^* = (\tau_i, \mathbf{X}_i, K_i, T_{i1}, T_{i2}, \ldots, T_{iK_i}, \tau_i - S_{iK_i}).$$

3.3 Bayesian Inference

3.3.1 Likelihood

In general we consider the case $s < \infty$ and let $B_i(v) = v - S_{iN_i^\dagger(v-)}$ be the backward recurrence time. Define the counting process and "at risk" processes via

$$N_i^\dagger(s) = \sum_{j=1}^{\infty} I\{S_{ij} \leq s, S_{ij} \leq \tau_i\} \text{ and } Y_i^\dagger(s) = I\{\tau_i \geq s\},$$

respectively. Let $\mathbf{D}^* = \{D_i^*\}_{i=1}^n$. Then, following Jacod (1975) or section II.7 of Andersen et al. (1993) define the conditional likelihood function over $[0, s]$ by

$$L(s \mid \mathbf{D}^*, W_1, W_2, \ldots, W_n)$$

$$= \prod_{i=1}^{n} \left[\left\{ \prod_{v \in [0,s]} \left[W_i \exp(\beta^T X_i) \lambda_0(B_i(v)) \right]^{N_i^\dagger(\Delta v)} \right\} \exp\left\{ -W_i \exp(\beta^T X_i) \int_0^s Y_i^\dagger(v) \lambda_0(B_i(v)) dv \right\} \right]. \tag{3.2}$$

Since in our Bayesian procedure we will assign a Gamma process prior on $\Lambda_0(\cdot)$ we need to further simplify our likelihood function for posterior calculations.

Following Peña et al. (2000), we can write

$$
\begin{aligned}
&\int_0^s Y_i^\dagger(v)\lambda_0(B_i(v))dv \\
&= \int_0^s I(\tau_i \geq v)\lambda_0(B_i(v))dv \\
&= \sum_{j=1}^{N_i^\dagger((s\wedge\tau_i)-)} \int_{S_{ij-1}}^{S_{ij}} I(\tau_i \geq v)\lambda_0(v)dv + \int_{S_{iN_i^\dagger((s\wedge\tau_i)-)}}^{s\wedge\tau_i} I(\tau_i \geq v)\lambda_0(v)dv \\
&= \sum_{j=1}^{N_i^\dagger((s\wedge\tau_i)-)} \int_0^{T_{ij}} I(T_{ij} \geq v)\lambda_0(v)dv \\
&\quad + \int_0^{(s\wedge\tau_i)-S_{iN_i^\dagger}((s\wedge\tau_i)-)} I((s\wedge\tau_i) - S_{iN_i^\dagger}((s\wedge\tau_i)-) \geq v)\lambda_0(v)dv \\
&= \int_0^s \left[\sum_{j=1}^{N_i^\dagger((s\wedge\tau_i)-)} I\{T_{ij} \geq v\} + I\{(s\wedge\tau_i) - S_{iN_i^\dagger}((s\wedge\tau_i)-) \geq v\}\right]\lambda_0(v)dv \\
&= \int_0^s Y_i(s,v)\lambda_0(v)dv,
\end{aligned}
$$

where

$$Y_i(s,v) = \sum_{j=1}^{N_i^\dagger((s\wedge\tau_i)-)} I\{T_{ij} \geq v\} + I\{(s\wedge\tau_i) - S_{iN_i^\dagger((s\wedge\tau_i)-)} \geq v\}.$$

When $s \to \infty$,

$$Y_i(s,v) \to \sum_{j=1}^{K_i} I\{T_{ij} \geq v\} + I\{\tau_i - S_{iK_i} \geq v\} \equiv Y_i(v).$$

In fact the above relation holds true when $s \geq \tau_i$. Let $t_{(1)}, t_{(2)}, \ldots, t_{(M)}$ be the M partition points on $\Re_+ = (0, \infty)$ with $t_{(0)} \equiv 0$ and $t_{(M+1)} \equiv \infty$ such that $Y_i(s,v)$ is constant within each subinterval $(t_{(j-1)}, t_{(j)}]$. Define

$$\Lambda_0(\Delta t_{(j)}) = \Lambda_0(t_{(j)}) - \Lambda_0(t_{(j-1)}), j = 1, 2, \ldots, M, M+1.$$

Then,

$$\int_0^s Y_i(s,v)\lambda_0(v)dv = \sum_{j=1}^{M+1} Y_i(s, t_{(j)})\Lambda_0(\Delta t_{(j)})$$

Similarly, we can write the product integral as a finite product as follows.

$$\prod_{v\in[0,s]} \left[W_i \exp(\beta^T X_i)\lambda_0(B_i(v))\right]^{N_i^\dagger(\Delta v)} = \prod_{j=1}^{M+1} \left[W_i \exp(\beta^T X_i)\Lambda_0(\Delta t_{(j)})\right]^{N_i^\dagger(\Delta t_{(j)})},$$

where

$$N_i^\dagger(\Delta t_{(j)}) = N_i^\dagger(t_{(j)}) - N_i^\dagger(t_{(j-1)})$$

Let $s \geq \max_{1\leq i\leq n}\{\tau_i\}$. Then, using the above notation we can rewrite the likelihood (3.2) function as follows

$$L(\Lambda_0(\cdot), \beta, \nu \mid \mathbf{D}^*, W_1, W_2, \ldots, W_n) = \prod_{i=1}^{n} L_i, \tag{3.3}$$

where

$$\begin{aligned} L_i \equiv L_i(\Lambda_0(\cdot), \beta, \nu \mid D_i^*, W_i) \\ = \left[\left\{\prod_{j=1}^{M+1}\left[W_i \exp(\beta^T X_i)\Lambda_0(\Delta t_{(j)})\right]^{N_i^\dagger(\Delta t_{(j)})}\right\}\right. \\ \left.\exp\left\{-W_i \exp(\beta^T X_i)\sum_{j=1}^{M+1} Y_i(t_{(j)})\Lambda_0(\Delta t_{(j)})\right\}\right]. \end{aligned} \tag{3.4}$$

3.3.2 Prior Specifications and Conditional Posteriors

The unknown parameters of interest are $(\Lambda_0(\cdot), \nu, \beta)$, however, we also derive the conditional distribution of the unobservable frailty vector $W_1, W_2, \ldots, W_n$ as those will be involved in the Bayes estimator of $\Lambda_0(\cdot)$. Following Kalbfleisch (1978), we assume $\Lambda_0(\cdot)$ has a gamma process prior

$$\Lambda_0(\cdot) \sim \mathscr{G}_{c,\Lambda_0^*(\cdot)},$$

where $\Lambda_0^*(\cdot)$ is a completely known mean intensity function and c is a precision of the prior measure. Then,

$$\Lambda_0(\Delta t_{(j)}) \sim Ga(c\Lambda_0^*(\Delta t_{(j)}), c).$$

We use the notation $\pi(\Lambda_0(\Delta t_{(j)}))$ to indicate the prior distribution of $\Lambda_0(\Delta t_{(j)})$ and so on. That is

$$\pi(\Lambda_0(\Delta t_{(j)})) \equiv Ga(c\Lambda_0^*(\Delta t_{(j)}), c), j = 1, 2, \ldots, M+1. \tag{3.5}$$

We consider the prior of ν as a Gamma distribution with a known shape γ and scale η and the prior of β as the multivariate normal distribution with a known mean vector μ_β and a variance-covariance matrix Σ_β denoted, respectively, by

$$\pi(\nu) \equiv Ga(\gamma, \eta) \tag{3.6}$$

and

$$\pi(\beta) \equiv N_q(\mu_\beta, \Sigma_\beta). \tag{3.7}$$

Using (3.3) and priors (3.5), (3.6), and (3.7) we define the joint posterior distribution of $\{\Lambda_0(\cdot), \nu, \beta\}$ via,

$$p(\Lambda_0(\cdot), \mathbf{W}, \nu, \beta \mid D^*) \equiv \prod_{i=1}^{n} \left[L_i(\Lambda_0(\cdot), \nu, \beta \mid W_i)\right] \pi(\Lambda_0(\cdot))\pi(\nu)\pi(\beta). \tag{3.8}$$

Let

$$r_i = \exp(\beta^T \mathbf{X}_i) \int_0^\infty Y_i(v)\Lambda_0(dv) \equiv \exp(\beta^T \mathbf{X}_i) \sum_{j=1}^{M+1} \left[Y_i(t_{(j)})\Lambda_0(\Delta t_{(j)})\right]$$

and

$$r_i^*(t_{(j)}) = \exp(\beta^T \mathbf{X}_i) Y_i(t_{(j)}).$$

Define $N(\Delta t_{(j)}) = \sum_{i=1}^n \left[N_i^\dagger(t_{(j)}) - N_i^\dagger(t_{(j-1)})\right]$. Clearly $N_i^\dagger(s) = K_i$ for $s \geq \tau_i$. Then, the conditional posterior distributions follow:

$$p(\Lambda_0(\Delta t_{(j)}) \mid \mathbf{W}, \nu, \beta) \\ \propto Ga(N(\Delta t_{(j)}) + c\Lambda_0^*(\Delta t_{(j)}), c + \sum_{i=1}^{n} W_i r_i^*(t_{(j)})), j = 1, 2, \ldots, M+1. \tag{3.9}$$

Under an integrated squared-error loss function we obtain

$$\widetilde{\Lambda}_0(t \mid \mathbf{W}, \beta, \nu) = \sum_{j=1}^{M+1} \left[\frac{N(\Delta t_{(j)}) + c\Lambda_0^*(\Delta t_{(j)})}{\sum_{i=1}^n W_i r_i^*(t_{(j)}) + c}\right] I(t_{(j)} \leq t) \tag{3.10}$$

This is not yet an estimator of $\Lambda_0(t)$ since W_i's are unknown. However, we can obtain $\widehat{W}_i$ to replace W_i using the following conditional posterior distribution.

$$p(W_i \mid \Lambda_0(\cdot), \nu, \beta) \\ \propto W_i^{K_i} \exp(-W_i r_i) g_{W_i}(w_i \mid \nu) \propto Ga(\nu + K_i, \nu + r_i), i = 1, 2, \ldots, n, \tag{3.11}$$

where $g_{W_i}(w_i \mid \nu) \propto W_i^{\nu-1}\exp\{-W_i\nu\}$. Clearly W_i can be updated via

$$\widehat{W}_i = \frac{\nu + K_i}{\nu + r_i}, i = 1, 2, \ldots, n, \text{ the posterior mean of } W_i. \tag{3.12}$$

Thus, we have a closed form estimator of $\Lambda_0(t \mid \cdot)$ given by

$$\widehat{\Lambda}_0(t \mid \beta, \nu) = \sum_{j=1}^{M+1} \left[\frac{N(\Delta t_{(j)}) + c\Lambda_0^*(\Delta t_{(j)})}{\sum_{i=1}^{n} \widehat{W}_i r_i^*(t_{(j)}) + c} \right] I(t_{(j)} \leq t) \tag{3.13}$$

We can recover the Breslow-Aalen type estimator of the baseline cumulative hazard function from our estimator by letting the precision parameter $c \to 0$. The conditional posterior distribution of the frailty parameter ν is given by

$$p(\nu \mid \Lambda_0(\cdot), \mathbf{W}, \beta) \propto L_m(\nu, \beta, \Lambda_0)\pi(\nu), \tag{3.14}$$

where the marginal likelihood L_m is defined by

$$L_m \equiv \prod_{i=1}^{n} \left[\int_0^\infty L_i(\lambda_0(\cdot), \beta, \nu \mid D_i, W_i = w_i) g_W(w) dw \right] \propto \prod_{i=1}^{n} \left[\frac{\Gamma(K_i + \nu)}{(r_i + \nu)^{K_i + \nu}} \frac{\nu^\nu}{\Gamma(\nu)} \right].$$

The conditional posterior distribution of β is given by

$$p(\beta \mid W, \Lambda_0(\cdot), \nu) \propto \exp\left[\sum_{i=1}^{n} K_i(\beta^T X_i) - \sum_{i=1}^{n} W_i r_i \right] \pi(\beta), \tag{3.15}$$

where $\pi(\beta)$ is as defined in (3.7).

3.3.3 MCMC Sampling

The conditional posterior distribution of Λ_0 and $W_i, i = 1, 2, \ldots, n$ as given in (3.9) and (3.11), respectively, are in closed form and thus easy to draw samples or one can update the estimators (3.10) and (3.12) in the Gibbs sampling algorithm. For ν we employ Metropolis-Hastings (MH) algorithm to sample from the conditional posterior (3.14). We sample β from the conditional posterior distribution (3.15) employing adaptive rejection sampling (ARS) (Gilks and Wild (1992)) since (3.15) is a log-concave function. The MCMC algorithm follows:

- **Step 0**: Start with initial

$$\Lambda_0^{(0)}(\cdot), W^{(0)}, \nu^{(0)}, \text{and } \beta^{(0)}$$

- **Step 1**: Update

$$\widehat{\Lambda}_0(\cdot \mid W^{(0)}, \beta^{(0)}).$$

Alternatively we could take samples from (3.9).

- **Step 2**: Using MH algorithm we sample ν from

$$p(\nu \mid \Lambda_0(\cdot), W, \beta) \propto L_m(\nu, \beta, \Lambda_0(\cdot))\pi(\nu).$$

More precisely, we propose

$$\nu^{cand} \equiv \nu' = exp(log(\nu^{curr}) + \epsilon), \quad \text{where } \epsilon \sim N(0, \sigma_\nu^2),$$

that is, the proposal density of ν is a log-Normal density,

$$q(\nu' \mid \nu, \sigma_\nu^2) \equiv lnN(log(\nu^{curr}), \sigma_\nu^2),$$

and accept it with probability

$$\min\left\{1, \frac{\prod_{i=1}^n \left[\frac{\Gamma(K_i+\nu')}{(r_i+\nu')^{K_i+\nu'}} \frac{(\nu')^{\nu'}}{\Gamma(\nu')}\right] \pi(\nu') q(\nu \mid \nu', \sigma_\nu^2)}{\prod_{i=1}^n \left[\frac{\Gamma(K_i+\nu)}{(r_i+\nu)^{K_i+\nu}} \frac{(\nu)^{\nu}}{\Gamma(\nu)}\right] \pi(\nu) q(\nu' \mid \nu, \sigma_\nu^2)}\right\}$$

where $\nu^{curr} \equiv \nu$. Usually σ_ν^2 is chosen small for example $\sigma_\nu^2 = 0.5^2$.

- **Step 3**: Update $W_i, i = 1, 2, \ldots, n$ using

$$\widehat{W}_i = \frac{\nu + k_i}{\nu + r_i}.$$

Alternatively we could take samples from (3.11).

- **Step 4**: Sample β from

$$p(\beta \mid W, \Lambda_0(\cdot), \nu) \propto \exp\left[\sum_{i=1}^n K_i \exp(\beta^T X_i) - \sum_{i=1}^n W_i r_i\right] \pi(\beta)$$

employing adaptive rejection sampling (ARS) since the conditional posterior density is a log-concave function. Repeat steps 1-4 until convergence.

Bias, variance, and credible intervals of the parameter of interest are immediately available from the posterior samples.

3.4 Simulation Studies

3.4.1 Simulation Design

Simulation studies are performed to evaluate our proposed method and properties of the parameter estimators numerically. More precisely, the goals of these studies are: (i) to examine the effects of the sample size (n) on the distributional properties of the estimators; (ii) examine bias, variance, and root mean-squared errors (RMSEs) of the baseline survivor function (pointwise); and (iii) examine the convergence of the MCMC algorithm. We describe the settings for different simulation parameters with their priors.

Baseline cumulative hazard function with priors: We generate the baseline cumulative hazard function $\Lambda_0(\cdot)$ associated with the gap-time from the Weibull distribution with unit scale and shape parameter, $\gamma \in \{1.1, 0.9\}$. We specify $\Lambda_0(\cdot)$ nonparametrically and consider $\pi(\Lambda_0(\Delta t)) \equiv Ga(c\Lambda_0^*(\Delta t), c)$, where $\Lambda_0^*(t) = t$ (cumulative hazard function from the unit exponential distribution) and $c = 0.1$. In this case the prior is misspecified. We also assign another prior which is a naive estimator of $\Lambda_0(t)$ given by

$$\Lambda_0^*(t) = \sum_{j=1}^{M+1} \left[\frac{N(\Delta t_{(j)})}{\sum_{i=1}^{n} Y_i(t_{(j)})} \right] I(t_{(j)} \leq t). \tag{3.16}$$

Covariates with priors for regression coefficients: We generate two covariates X_1, and X_2, where X_1 is a binary variate taking 0 or 1 with probability 0.5 and $X_2 \sim N(0, 0.5^2)$ with X_1 and X_2 independent. The regression parameters are set to be $(\beta_1, \beta_2) = (-1, 1)$. Priors of β_1 and β_2 are denoted by $\pi(\beta_1) \equiv N(-.5, 1)$ and $\pi(\beta_2) \equiv N(.5, 1)$, respectively. Note that

$$p(\beta_1 \mid W, \Lambda_0, \beta_2, \nu) \propto \exp\left[\beta_1 K_{trt} - \exp(\beta_1) r_{trt}\right] \pi(\beta) = \theta^{K_{trt}} \exp[-\theta\ r_{trt}] \pi(\beta), \tag{3.17}$$

where $\theta \equiv \exp(\beta_1)$, K_{trt} = total number of events occurrences for the treatment ($X_1 = 1$) group and $r_{trt} = \sum_{i\in\{trt\}} \left[W_i \exp(\beta_2 X_{2i}) \int_0^\infty Y_i(v)\Lambda_0(dv)\right]$. With a non-informative flat prior for β_1, the above conditional distribution (3.17) is proportional to the Gamma distribution with shape K_{trt} and rate r_{trt}. Similarly,

$$p(\beta_2 \mid W, \Lambda_0, \beta_1, \nu) \propto \exp\left[\beta_2 \sum_{i=1}^{n} K_i X_{2i} - \sum_{i=1}^{n} [\exp(\beta_2 X_{2i}) r_{1i}]\right] \pi(\beta), \tag{3.18}$$

where $r_{1i} = W_i \exp(\beta_1 X_{1i}) \int_0^\infty Y_i(v)\Lambda_0(dv)$. Clearly, both (3.17) and (3.18) are log-concave function. Thus, in the MCMC sampling we employ adaptive rejection sampling (Gilks and Wild (1992)) algorithm for the log-concave function to sample from (3.17) and (3.18).

Frailty component with a prior: The parameter ν of the gamma distribution generating the frailty variable W is set to be $\{2, 5\}$. Prior for ν is defined by $\pi(\nu) \equiv Ga(2, .5)$.

Censoring variate: We consider censoring variable $\tau_i, i = 1, 2, \ldots, n$ are IID from a continuous uniform distribution $U(1, 5)$.

Sample size: We choose $n \in \{100, 200, 300\}$ to examine the impact of different sample sizes on bias, variance, and convergence of the parameter estimators.

For each combination of these simulation parameters, we compute the mean and standard deviation of the estimates of the parameters and averaged pointwise biases and RMSEs for the baseline survivor function based on 100 data sets. We set the end of monitoring time to $\tau = 5$ and thus we only need to compute the cumulative hazard function over the interval $(0, 5]$.

3.4.2 Simulation Results

Table 3.1 summaries the mean values and standard deviations, of the sampling distributions of the estimators of β_1, β_2, and ν for each combination of $\gamma \in \{1.1, 0.9\}$, $\nu \in \{2, 5\}$, and $n \in \{100, 200, 300\}$. In the discussion of the simulation results we will examine the effect of changing n, changing ν and changing γ on the distributional properties of the estimators of β, ν, and biases and RMSEs of the estimator of the baseline survivor function $\bar{F}_0$.

Table 3.1: Summary of Simulation Results

n	ν	γ	$\hat{\beta}_1$	$\hat{\sigma}_{\hat{\beta}_1}$	95%CP	$\hat{\beta}_2$	$\hat{\sigma}_{\hat{\beta}_2}$	95%CP	$\hat{\nu}$	$\hat{\sigma}_{\hat{\nu}}$	95%CP
100	2	1.1	-0.97	0.23	0.84	0.95	0.20	0.87	2.74	0.83	0.93
200	2	1.1	-0.97	0.13	0.89	0.96	0.16	0.88	2.30	0.56	0.96
300	2	1.1	-0.99	0.11	0.94	0.98	0.10	0.94	2.13	0.35	0.98
100	5	1.1	-0.98	0.17	0.88	0.98	0.18	0.88	5.16	1.31	0.98
200	5	1.1	-1.00	0.13	0.90	0.98	0.10	0.94	5.51	1.36	0.99
300	5	1.1	-1.01	0.10	0.92	0.99	0.09	0.95	5.38	1.40	0.96
100	2	0.9	-0.98	0.19	0.90	0.95	0.19	0.93	2.69	0.93	0.94
200	2	0.9	-0.99	0.15	0.93	0.97	0.14	0.93	2.35	0.68	0.94
300	2	0.9	-0.99	0.10	0.95	0.98	0.13	0.95	2.22	0.45	0.95
100	5	0.9	-0.98	0.18	0.95	0.95	0.18	0.91	5.31	1.48	0.94
200	5	0.9	-0.99	0.12	0.97	0.97	0.11	0.92	5.23	1.37	0.94
300	5	0.9	-1.00	0.10	0.99	0.98	0.10	0.94	5.20	1.30	0.95

As is expected the standard deviation decreases when the sample size increases for all combination of the parameters $\nu \in \{2, 5\}$ and $\gamma \in \{1.1, 0.9\}$. For $n = 100$ with $\nu = 2$ there is a considerable bias for the estimate of ν. Larger values of ν (e.g., $\nu = 5$), resulting into a weaker correlation between gap-times, contributing to a smaller standard deviation for the parameter estimates of β_1 and β_2. However, for $\nu = 2$, its standard deviation increases as the sample size increases. For all combination of parameters values with sample size 200 and 300 the estimates are close to the true parameters.

We also investigate the point-wise biases and RMSEs of the baseline survivor function for different combination of parameter values and the resulting graphs are presented in Figure 3.2. As the sample size increases the biases and RMSEs decreases. While generating the gap-times we consider both increasing failure rate (IFR) and decreasing failure rate (DFR) Weibull baseline distribution with shape parameter $\gamma = 1.1$ and $\gamma = 0.9$, respectively. However, here we report the result for $\gamma = 1.1$ with different sample sizes. Convergence of MCMC algorithm are checked by "CODA"

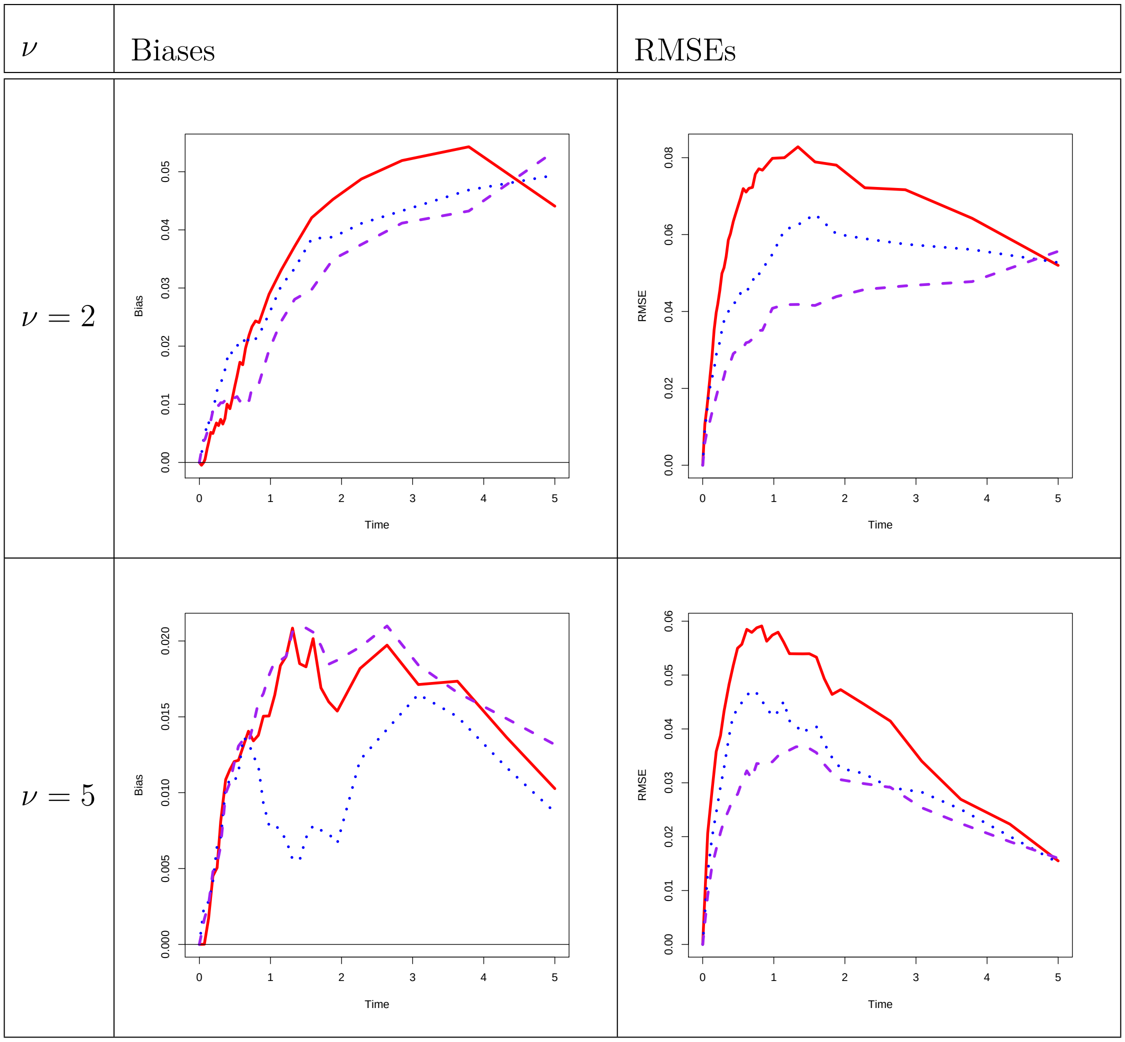

Figure 3.2: Simulated biases and root mean-squared errors (RMSE) for the estimator of the baseline survivor function $\bar{F}_0$ as the sample size increases [n=100 is red and solid; n=200 is blue and dotted; n=300 is purple and dashed]. Weibull shape $\gamma = 1.1$.

(Plumber et al. 2012, http://cran.r-project.org/web/packages/coda/coda.pdf) package available in R. As the sample size increases the estimators converge to the true parameters at a faster rate.

3.5 Data Analysis

We illustrate our methodology using the mammary tumors data given in Gail et al. (1980) and the bladder cancer data available in Wei et al. (1989). First, we will analyze mammary tumors data where the gap-times are the time between tumor occurrences in 48 rats in that 23 of them are in the treatment group and the other 25 are in the control group. The only covariate is the treatment. That is $X = 1$ indicates that the subject is in the treatment group and $X = 0$ means that the subject is in the control group. The main objective of their analysis is whether the treatment is effective in reducing the hazard of tumor occurrences. Gail et al. (1980) analyze each gap-time separately (e.g. first gap-times for all units and so on) ignoring the recurrent nature of event occurrences and the possibility of correlation between gap-times. More details of the data set and a description of the clinical trials are given in Gail et al. (1980).

Table 3.2: Summary estimates for the mammary tumors data

Parameter	Estimate	Std	95% Credible Interval
β (Treatment)	-0.660	0.17	(-0.996, -0.333)
ν	4.586	1.88	(2.071, 9.651)

In our Bayesian analysis we consider two priors for the baseline cumulative hazard $\Lambda_0(\cdot)$ such as Gamma process prior with known mean intensity $\Lambda_0^*(t) = t^\gamma, \gamma = 1.1$ with $c = 0.1$ and the empirical prior given in (3.16). We assume $\pi(\beta) \equiv N(-1, 1)$ and $\pi(\nu) \equiv Ga(2, .5)$. We choose distinct ordered censored observations as partition points and the end of study time is $t_M = 122$. However, the length between two successive partition points can be very small such as one unit. The choice of partition length is not sensitive to the estimates of the parameters. To obtain estimates of the parameters of interest we follow the algorithm given in MCMC sampling section. Note that for one dimensional binary covariate with non-informative flat prior the conditional posterior distribution of $exp(\beta) = \beta'$ is the Gamma distribution with

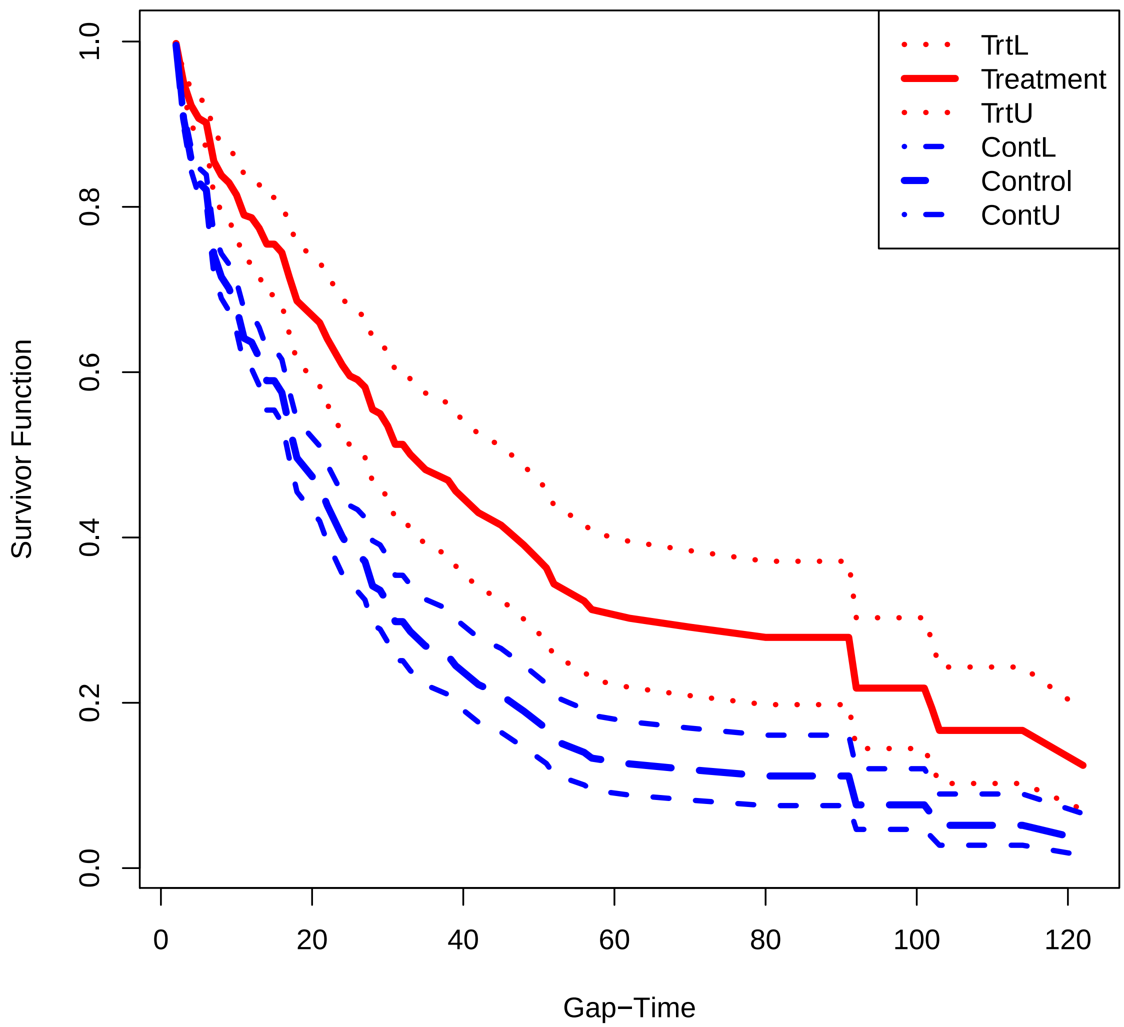

Figure 3.3: Survivor function of the treatment group (solid line) and control Group (long dash line) with credible intervals. Dotted lines are credible interval for the treatment group and dashed line are credible intervals for the control group.

shape parameter K_{trt} = total number of tumor occurrences for treatment group and the rate parameter $r_{trt} = \sum_{i \in \{trt\}} \left[W_i \int_0^\infty Y_i(v) \Lambda_0(dv) \right]$. Adaptive rejection sampling also yields identical result.

The parameter estimates associated with the treatment and the frailty random variable with standard deviation and credible intervals are presented in Table 3.2.

We obtained $\hat{\beta} = -0.66$, $\hat{\sigma}_\beta = 0.17$ and the 95% credible interval is $(-0.99, -0.33)$. The estimate of β with its credible interval indicates that the treatment is effective in reducing the hazard rate of tumor occurrences. The estimate of ν suggests that gap-times are not strongly correlated. Survivor function estimates with credible intervals are presented in Figure 3.3, where upper curve (solid line) is the survivor function with 95% credible intervals (dotted lines) for the treatment group, while lower curve (long dashed line) is the survivor function with 95% credible intervals (dashed line) for the control group. Figure 3.3 demonstrates that there is no clear evidence that the treatment group has a higher survival rate than the control group as the credible intervals are overlapped.

Table 3.3: Summary estimates for the bladder cancer data

Variable	Parameter	Estimate	Std	95% Credible Interval
rx	β_1	-0.36	0.19	(-0.740, 0.003)
size	β_2	-0.01	0.06	(-0.130, 0.100)
number	β_3	0.14	0.05	(0.040, 0.240)
frailty	ν	6.30	3.00	(2.210, 13.800)

The second biomedical application is the bladder cancer data analyzed by Wei et al. (1989), which is also available in *survival* package (Therneau and Lumley (2009)) in the *R library*. These data include the times to re-occurrences of tumors of bladder cancer patient for $n = 85$ subjects. Three covariates of interest are X_1, the treatment indicator (0= placebo; 1= thiotepa); X_2, the size (in centimeter) of the largest initial tumor; and X_3, the number of initial tumors. For our estimation purpose we assume similar priors for the parameters β and ν such as Normal and Gamma distributions, respectively.

The summary of the parameter estimates (bladder cancer) is presented in Table

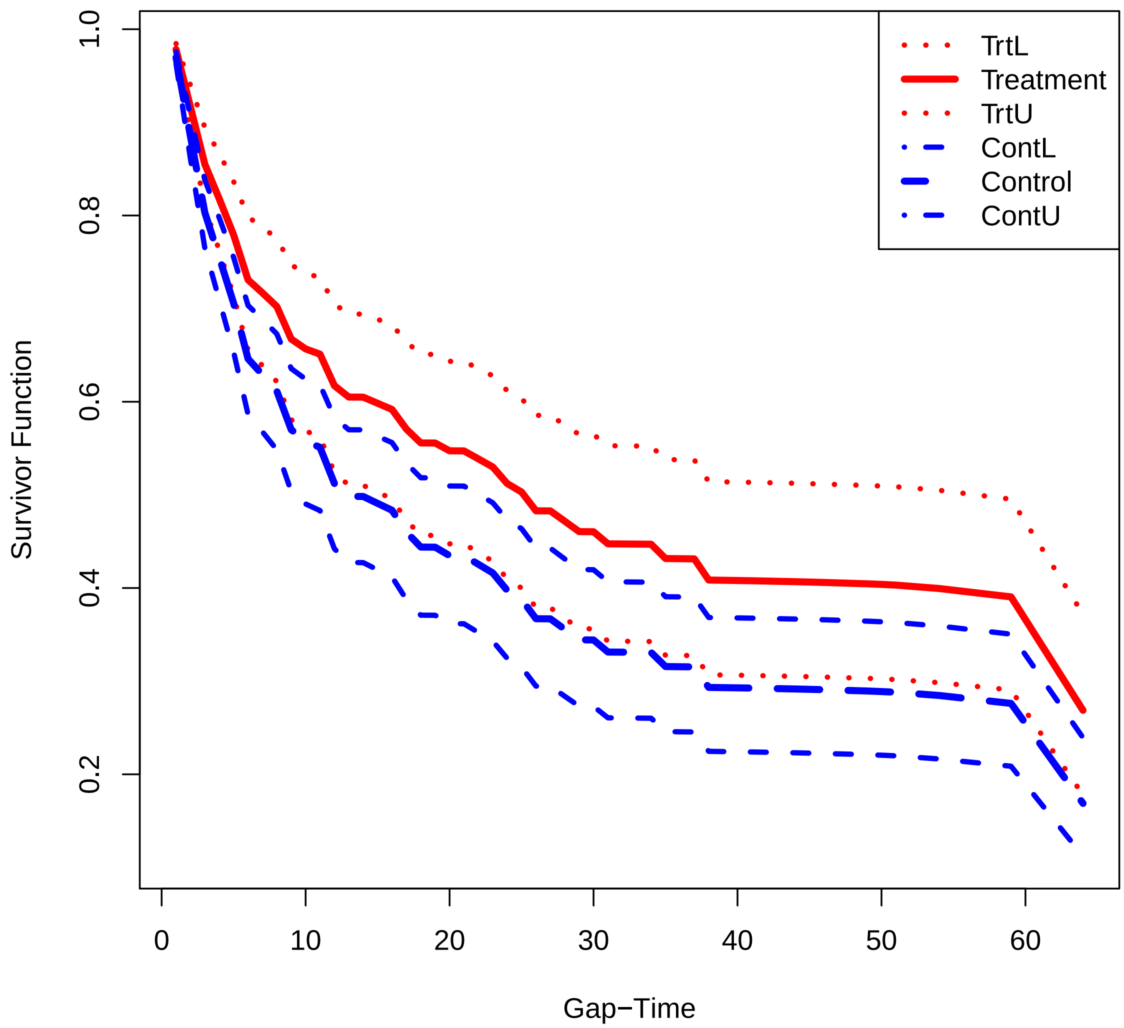

Figure 3.4: Survivor function of the treatment group (solid line) and control Group (long dash line) with credible intervals. Dotted lines are credible interval for the treatment group and dashed line are credible intervals for the control group.

3.3. The estimate of β_1 (treatment) indicates that there is no clear evidence that the treatment is effective in reducing the bladder tumor occurrences since the credible interval includes zero. The initial size of the largest tumor is not a significant factor for reducing or increasing tumor recurrences. Estimates of β_3 with its credible interval suggest that initial number of tumor occurrences is an indication of increasing number of tumor re-occurrences. The estimates of survivor function for the treatment group and control group are computed at the mean values of X_2 and X_3 and Figure 3.4 is the resulting survivor function graph. Figure 3.4 seems to indicate that the thiotepa (treatment) group has a higher survival rate than the control group.

3.6 Concluding Remarks

In this chapter, we developed a semiparametric Bayesian inference procedure of the gap-time distribution with recurrent event data. This is a more general framework in the sense that the Breslow-Aalen type estimator of the baseline cumulative hazard function is a limiting case of our estimator by letting precision parameter tends to zero. In our procedure, we are able to assess the effects of treatment on event occurrences as well as estimate the baseline survivor function which is often an interest in reliability engineering applications. The proposed estimation procedure can easily be implemented as the baseline cumulative hazard function has a closed form estimator and the estimates of the parameter associated with covariates and frailty are easily implemented by *ars*() and by *Metro-Hastings*() package in *R*, respectively. Credible intervals are immediately available from the posterior samples. Through the procedure we are able to obtain the individual frailties. A frailty greater than one indicates that the individual is more prone to event occurrences and vice-versa. A larger value of $\hat{\nu}$ is an indication of weaker correlation between gap-times. Simulation studies demonstrate that parameter estimates are robust to some misspecification of the prior distribution.

In our procedure we assumed that censoring is independent of recurrent events. In many situations, however, there exists a terminal events (e.g. death), which prevent the occurrences of future events. Moreover it is often the case that the terminal event is strongly correlated with recurrent event occurrences. In particular, increasing the number of event occurrences could potentially increase the risk of death. Thus noninformative termination assumption may not be appropriate in some situations. Miloslavsky et al. (2004b) showed that regression parameter estimates in the recurrent event analysis are biased when dependent termination is ignored. One way of accommodating the dependent termination is to consider the joint modeling of recurrent event and terminal events. One of my future research goals is to a develop joint Bayesian inference procedure of recurrent and terminal events.

Chapter 4

Nonparametric Bayes Estimation of Reliability of a Coherent System [1]

Abstract

Simultaneous estimation of system and components reliabilities is considered when independent partition-based Dirichlet (PBD) priors are assigned on component lifetime distributions. Denote the lifetime of component j in the i-th system by $\{T_{ij}, j = 1, 2, 3, \ldots, K\}$ and the end of system monitoring time by $\{\tau_i, i = 1, 2, \ldots, n\}$. Assume that $\{T_{ij}, i = 1, 2, 3, \ldots, n\}$ and $\{\tau_i, i = 1, 2, \ldots, n\}$ are IID with distribution F_j and G, respectively, and with $\{T_{ij}\}$s and $\{\tau_i\}$s mutually independent and T_{ij} and T_{il} also independent for $j \neq l$. In our nonparametric Bayesian approach we assign independent partition-based Dirichlet (PBD) priors, $\mathscr{D}(\alpha_j)$, on $F_j, j = 1, 2, \ldots, K$, with the parameter α_j being a non-null finite measure on $\Re_+$. We derive the nonparametric Bayes estimator of component reliabilities, $\bar{F}_j = 1 - F_j$ for $j = 1, 2, \ldots, K$, and an estimator of the system reliability function $\bar{F}_\phi(t) = h_\phi(\bar{F}_1(t), \bar{F}_1(t), \ldots, \bar{F}_K(t))$, where ϕ is the structure function of the system. The estimator of the system reliability function presented in Doss et al. (Annals of Statistics, 1989) is a special case of our estimator, obtained by letting $\alpha_j(\Re_+) \to 0$ for $j = 1, 2, \ldots, K$. Through simulation studies, we demonstrate that the nonparametric estimator has smaller bias, but higher root-mean-squared errors (RMSE) than our proposed estimator. Even when the prior mean functions do not coincide with the

[1] A.K.M. Fazlur Rahman and Edsel A. Peña. To be submitted.

true distribution functions, the Bayes estimator has smaller or equal RMSE than the nonparametric estimator with a smaller value of precision parameter, indicating robustness of our estimator. In addition, our proposed estimator is smoother in some sense than the Doss et al. (1989) estimator.

4.1 Introduction

The lifetime of a sold/deployed system is a common interest for both the client and manufacturer as it is a meaningful measure of the quality of a system. Therefore it has been of interest to assess the risk of a failure and the reliability of systems in many settings, for instance in the mechanical, engineering, biomedical, military, and business areas. Precise and reliable knowledge of the performance of deployed systems enables the informed assessment of the risk and failure of the system that could potentially save life, wealth, and prevent destruction. It is therefore imperative to have probabilistic and statistical inferential methods to assess the risk and reliability of systems.

We begin with a brief review of relevant published literature, followed by some notations and definitions relevant to reliability of coherent systems. Meilijson (1981) considered estimation of F based on system failure times together with components failure times (autopsy statistics), whereas Moeschberger and David (1971) considered estimation of F under IID assumptions in the competing risk framework. Estimation of F under rank set sampling when systems on the test are k-out-of-K systems is considered by Kvam and Samaniego (1994) and Stokes and Sager (1988). Estimation of load sharing properties in a dynamic reliability system is considered by Kvam and Peña (2005). Joint estimation of components and system reliabilities is considered by Doss et al. (1989). In addition, inferential problems to assess the risk and reliability of systems has, among others, been considered by Barlow and Hunter (1960), Barlow and Marshall (1967), Barlow and Proschan (1969, 1986), Barlow (1984, 1985, 1986), Boyles et al. (1985), El-Neweihi et al. (1978), Esary and Proschan (1963), Esary et al. (1971, 1970), Hollander and Proschan (1984), Hollander and Peña (1995, 1996a,b, 2004), Langberg et al. (1981), and Peña and Hollander (2004). However, most of these works considered frequentist parametric/semi-parametric or nonparametric inference of system reliability.

The classical form of making inference is an ideal situation when the parametric assumptions of the distribution function match with the true distribution function but it becomes worse in the case of a misspecified parametric family of distributions. However, in the nonparametric Bayesian framework we can incorporate our prior knowledge in developing a robust estimation procedure to assess the reliability of a coherent system, which is still in development stage in the reliability literature. Thus it is worthwhile to develop Bayesian inferential methods to assess the risk and reliability of components and coherent systems and to compare their performance with existing methods. In our nonparametric Bayesian framework, we assign a partition-based Dirichlet prior (Sethuraman and Hollander (2009)) on F and the resulting posterior measure is also a partition-based Dirichlet measure. We develop a robust estimation procedure where our estimator of $\bar{F}$ is closed form and exact, and is a linear combination of the prior mean function and a corresponding nonparametric estimator. We define some notation and recall some definitions associated with coherent structure and system reliability based on Barlow and Proschan (1981).

A reliability system is composed of a finite number of components, with each component possibly a subsystem itself. For a reliability system with K components, denote the state vector of components by $\mathbf{x} = (x_1, x_2, \ldots, x_K)$, with $x_j \in \{0, 1\}$ such that $x_j = 1$ if component j is functioning and $x_j = 0$ if component j is not functioning. The structure function of a reliability system is defined by $\phi : \{0, 1\}^K \to \{0, 1\}$ such that $\phi(\mathbf{x})$ indicates whether the system is in a functioning state ($\phi(\mathbf{x}) = 1$) or is in a failed state ($\phi(\mathbf{x}) = 0$). A reliability system is said to be a coherent system if the structure function $\phi(\mathbf{x})$ satisfies the two conditions that (i) it is nondecreasing in each argument, that is a change of state of one and only one component from a failed state to a working state should not cause the system to change from a working state to a failed state, and (ii) each component is relevant in the sense that, for each $j \in \{1, 2, \ldots, K\}$ there exists an $\mathbf{x} \in \{0, 1\}^K$ such that $0 = \phi(\mathbf{x}, 0_j) < \phi(\mathbf{x}, 1_j) = 1$, where $(\mathbf{x}, 0_j) = (x_1, \ldots, x_{j-1}, 0, x_{j+1}, \ldots, x_K)$ and $(\mathbf{x}, 1_j) = (x_1, \ldots, x_{j-1}, 1, x_{j+1}, \ldots, x_K)$. See Barlow and Proschan (1981) for a comprehensive treatment of coherent system reliability.

A series system and a parallel system are two common examples of coherent reliability systems with respective structure functions

$$\phi_{ser}(\mathbf{x}) = min\{x_1, x_2, \ldots, x_K\} = \prod_{j=1}^{K} x_j$$

and

$$\phi_{par}(\mathbf{x}) = max\{x_1, x_2, \ldots, x_K\} = \coprod_{j=1}^{K} x_j = 1 - \prod_{j=1}^{K}(1 - x_j).$$

The coherent structure function of a more general k-out-of-K system is $\phi_{k:K}(x) = I(\sum_{j=1}^{K} x_j \geq k)$ with $I(A) = 1$ or 0 depending on whether event A does or does not hold, respectively. Clearly, a series system (K-out-of-K system) and a parallel system (1-out-of-K) are two extreme cases of k-out-of-K systems. A simple example of another coherent structure is a 3-component series-parallel system (Figure 4.1) with the structure function, $\phi_{sp}(x_1, x_2, x_3) = min\{x_1, max\{x_2, x_3\}\}$. This system functions as long as component 1 and at least one of component 2 and 3 are functioning.

Let X_k be the random variable indicating whether component k is in a functioning state or not, and let $p_j = Pr\{X_j = 1\}, j = 1, 2, \ldots, K$. Let $(X_1, X_2, \ldots, X_K)$ be independent random variables, and define $\mathbf{X} = (X_1, X_2, \ldots, X_K)$ and $\mathbf{p} = (p_1, p_2, \ldots, p_K)$. The reliability function associated with a structure function ϕ is defined by

$$h_\phi(\mathbf{p}) = E\{\phi(\mathbf{X})\} = Pr\{\phi(\mathbf{X}) = 1\}.$$

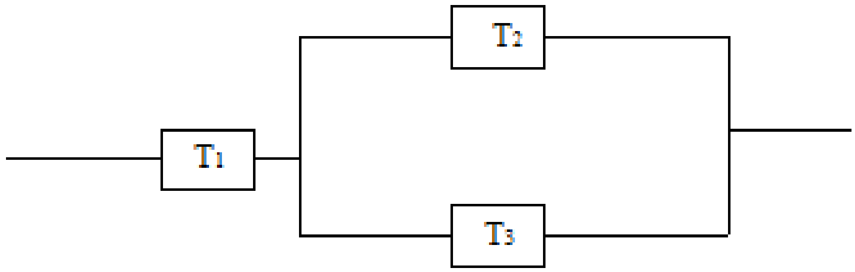

Figure 4.1: Three Component Series-Parallel System

The reliability function for a series system is $h_{ser}(\mathbf{p}) = \prod_{j=1}^{K} p_j$, while for a parallel system, it is $h_{par}(\mathbf{p}) = 1 - \prod_{j=1}^{K}(1 - p_j)$. On the other hand, the reliability function for the series-parallel system (Figure 4.1) is $h_{sp}(p_1, p_2, p_3) = p_1[1 - (1 - p_2)(1 - p_3)]$. The reliability function for the more general k-out-of K system is

$$h_{k:K}(\mathbf{p}) = \sum_{\{(x_1,x_2,\ldots,x_K)\in\{0,1\}^K;\sum_{j=1}^{K} x_j \geq k\}} \prod_{j=1}^{K} p_j^{x_j}.$$

These reliability functions represent the probabilities that the systems are functioning as a function of component reliabilities. Therefore it is useful to consider the component and system lifetimes. Let $\mathbf{T} = (T_1, T_2, \ldots, T_K)$ be the vector of component lifetimes and S be the system lifetime. For a given time t, the state vector of components is denoted by $\mathbf{X}(t) = (I(T_1 > t), I(T_2 > t), \ldots, I(T_K > t))$, hence the state of the system at time t is given by $\phi(\mathbf{X}(t))$. As a result $\{S > t\} = \{\phi(\mathbf{X}(t)) = 1\}$. Therefore the system lifetime survivor function is given by

$$\bar{F}_\phi(t) = Pr\{S > t\} = Pr\{\phi(\mathbf{X}(t) = 1\} = E\{\phi(\mathbf{X}(t))\}.$$

Denote the component lifetime survivor functions by $\bar{F}_j(t) = Pr\{T_j > t\}, j = 1, 2, \ldots, K$. If the component lifetimes are independent, then the system survivor function could be expressed in terms of the system's reliability function via

$$\bar{F}_\phi(t) = h_\phi(\bar{F}_1(t), \bar{F}_2(t), \ldots, \bar{F}_K(t)). \tag{4.1}$$

Assume n identical systems with the same structure function ϕ are in the study. We monitor the ith system over a period of time $[0, \tau_i]$, where τ_i could be some administrative time or some other time not related to the component failures. Denote by $S_1, S_2, \ldots, S_n$ the system lifetimes and by $\tau_1, \tau_2, \ldots, \tau_n$ the end of monitoring times. However, we may not observe all of the system failure times as the monitoring period of the i-th system terminated at τ_i, rather we only know that lifetime exceeded τ_i if the system is still functioning at time τ_i. The random observable for the n systems will therefore be

$$(\mathbf{V}, \epsilon) = ((V_1, \epsilon_1), (V_2, \epsilon_2), \ldots, (V_n, \epsilon_n)), \tag{4.2}$$

where $V_i = \min\{S_i, \tau_i\}$ and $\epsilon_i = I\{S_i \leq \tau_i\}$. A nonparametric estimator of system reliability function based on right-censored system lifetimes $(\mathbf{V}, \epsilon)$ is the Kaplan and Meier (1958) estimator also known as the Product-Limit estimator (PLE) given by

$$\widehat{R}_{PLE}(t) = \prod_{s \leq t} \left[1 - \frac{\Delta N(s)}{Y(s)}\right] = \prod_{s \leq t} \left[\frac{Y^+(s)}{Y(s)}\right], t \in \Re, \tag{4.3}$$

where $\prod$ means product-integral, and the processes $N = \{N(s) : s \in \Re\}$, $Y = \{Y(s) : s \in \Re\}$, and $Y^+ = \{Y^+(s) : s \in \Re\}$ are defined via

$$N(s) = \sum_{i=1}^{n} I\{V_i \leq s; \epsilon_i = 1\}, \quad Y(s) = \sum_{i=1}^{n} I\{V_i \geq s\}, \quad \text{and} \quad Y^+(s) = \sum_{i=1}^{n} I\{V_i > s\}. \tag{4.4}$$

Doss et al. (1989) developed a PL-type estimator of the system reliability function when component failure times are available by exploiting the relationship between the system reliability function and the component reliabilities, given in (4.1). The idea implemented here is to use the j-th component's right-censored data to estimate $\bar{F}_j, j = 1, 2, \ldots, K$, and then plug in these estimates in (4.1). In Doss et al. (1989)'s approach, each system is monitored until it failed so that they observed all the system failure times. Note, however, that any component's failure time could be right-censored. In our approach, the monitoring times could be fixed or random, and hence we may not observe all the system failure times. Thus in our settings with autopsy statistics of component lifetimes, the random observables for both the system and component failure times consist of complete and right-censored times.

Let T_{ij} denote the lifetime of component j for the i-th system. Define $Z_{ij} = \min\{T_{ij}, S^*_{ij}, \tau_i\}$ and $\delta_{ij} = I(T_{ij} \leq \min\{S^*_{ij}, \tau_i\})$. The right-censoring variable for T_{ij} involves S^*_{ij}, where S^*_{ij} is the lifetime of the original system with j-th component functioning uninterruptedly. Note that S^*_{ij} is independent of T_{ij} but depends on the structure function. We can view the random variable S^*_{ij} from the simple three-component series-parallel system as shown in Figure 4.1 where right-censoring variable for T_1 is $S^*_1 = \max\{T_2, T_3\}$, which is assumed independent of T_1. Similarly, the right-censoring variables for T_2 and T_3 are $S^*_2 = \min\{T_1, T_3\}$ and $S^*_3 = \min\{T_1, T_2\}$, respectively.

Denote by F the distribution of system lifetimes and assume F has partition-based Dirichlet (PBD) prior measure (Sethuraman and Hollander (2009)) which is defined formally in Section 4. The posterior distribution of F given the right-censored observations of system lifetime is also a PBD. Under an integrated squared error loss function, the Bayes estimator of F is the posterior mean. In a similar fashion, assigning an independent PBD prior measure on each component distribution function we estimate the distribution/reliability function for each component. Therefore joint estimation of component reliability and system reliability is obtained by plugging in the estimates of component reliability in (4.1).

We outline the contents of this Chapter. In Section 2 we review definitions and results for the PBD prior and also develop some results to estimate the component and system reliabilities. In Section 3 we drive the nonparametric Bayes estimator of system reliability and jointly estimate the component reliability and system reliability. We also explore the relationship between our proposed estimator and corresponding PL-type estimator. Section 4 includes simultaneous studies for correctly specified priors and misspecified priors and compares between proposed and PL-type estimators in terms of bias and RMSE function. Section 5 is an illustration of the developed estimators with a randomly generated data set. An appendix gathers the technical proofs.

4.2 Partition Based Dirichlet Prior and Posterior Measure

The partition based prior is a general class of priors introduced by Sethuraman and Hollander (2009) in the context of repair models. However, this prior can be used/extended as a nonparametric prior of the unknown distribution function F in different data settings including single event with right-censored, interval-censored, and truncated data settings. To define formally the notion of the partition based (PB) prior and some relevant results, let $(\mathscr{X}, \mathscr{A})$ be a measurable space where $\mathscr{A}$ is the σ-field of subsets of the space $\mathscr{X}$ and let $\mathscr{P}$ be the class of all probability measures (p.m.'s) on $(\mathscr{X}, \mathscr{A})$. Let $\mathscr{H}$ be the class of all pm's on $(\mathscr{P}, \mathscr{S})$ where $\mathscr{S} = \sigma(\{P : P(A) \leq r, A \in \mathscr{A}, 0 \leq r \leq 1\})$.

Dirichlet Process Probability Measure: Following Ferguson (1973) or Sethuraman (1994), let $\alpha(\cdot)$ be a non-null finite measure on $(\mathscr{X}, \mathscr{A})$. A random probability measure (p.m.) on $(\mathscr{P}, \mathscr{S})$ is said to be a Dirichlet probability measure if for any measurable partition, $\mathscr{B} = \{B_1, B_2, \ldots, B_m\}$ of $\mathscr{X}$, the distribution of $P(\mathscr{B}) \equiv (P(B_1), P(B_2), \ldots, P(B_m))$ is the finite dimensional Dirichlet distribution $\mathscr{D}(\alpha(B_1), \ldots, \alpha(B_m))$. Such a p.m. will be denoted by $\mathscr{D}_\alpha$.

Partitioned Based (PB) Measure: The partition based prior is a random measure restricted to a partition. Following Sethuraman and Hollander (2009), let $\mathscr{B} = (B_1, B_2, \ldots, B_m)$ be a measurable partition of the sample space $\mathscr{X}$. For $H \in \mathscr{H}$, and $P \sim H$, the law of total probability gives

$$P(\cdot) = \sum_{l=1}^{m} P_{B_l}(\cdot) P(B_l) = \sum_{l=1}^{m} P(\cdot \mid B_l) P(B_l)$$

where P_B is the restriction of P to B defined by $P_B(A) = P(A \cap B)/P(B)$ for all A, $B \in \mathscr{A}$. P_B is a restricted probability measure such that $P_B(B) = 1$.

We call $H(\mathscr{B}, h, \mathscr{G})$ a partition based (PB) distribution if the following conditions hold:

(a) The real random vector $P(\mathscr{B}) \equiv (P(B_1), P(B_2), \ldots, P(B_m))$ and restricted random p.m.'s $P_{B_l}, l = 1, 2, \ldots, m$, are all independent.

(b) $(P(B_1), P(B_2), \ldots, P(B_m))$ has the pdf $ch(\mathbf{y})$ on the simplex $R^m = \{\mathbf{y} : y_l \geq 0, l = 1, \ldots, m; \sum_{l=1}^m y_l = 1\}$, where c is a normalizing constant and $P(B_l) \equiv y_l$.

(c) $\mathscr{G} = G_1 \times \ldots \times G_m; P_{B_l} \sim G_l, l = 1, 2, \ldots, m$.

Partition Based Dirichlet (PBD) Measure: Let $\alpha(\cdot)$ be a finite measure on $(\mathscr{X}, \mathscr{A})$. For any non-empty set $B \in \mathscr{A}$ with $\alpha(B) > 0$, let $\alpha_B(A) = \alpha(A \cap B)$ for all $A \in \mathscr{A}$. If $\mathscr{G} = \mathscr{D}_{\alpha_{B_1}} \times \ldots \times \mathscr{D}_{\alpha_{B_m}}$, then $H(\mathscr{B}, h, \mathscr{G}) \equiv \mathscr{D}(\mathscr{B}, h, \alpha)$ will be referred to as a PB Dirichlet (PBD) measure. In this case,

$$h(\mathbf{y}) = \frac{\Gamma(\sum_{l=1}^m \alpha_l)}{\prod_{l=1}^m \Gamma(\alpha_l)} \prod_{l=1}^m {y_l}^{\alpha_l - 1} \tag{4.5}$$

is an m-dimensional multivariate Dirichlet distribution with parameter $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m)$ where $\alpha_l = \alpha(B_l)$.

The following result is a restatement of Lemma 1 of Sethuraman and Hollander (2009) regarding $\mathscr{D}_\alpha$ representation as a PBD distribution, $\mathscr{D}(\mathscr{B}, h, \alpha)$.

Result 1. Let α be a finite measure on $(\mathscr{X}, \mathscr{A})$ and let $\mathscr{B} = (B_1, B_2, \ldots, B_m)$ be a measurable partition of $\mathscr{X}$. Let $\alpha_l = \alpha(B_l)$ for $l = 1, 2, \ldots, m$, and $h(\mathbf{y})$ is as defined in (4.5). Then $\mathscr{D}_\alpha = \mathscr{D}(\mathscr{B}, h, \alpha)$.

Let $\mathscr{B} = (B_1, B_2, \ldots, B_m)$ be a measurable partition of $\mathscr{X}$ and let

$$\mathscr{B}_* = (B_1, \ldots, B_{r-1}, B_{r_1}, B_{r_2}, B_{r+1}, \ldots, B_m)$$

be a finer partition of $\mathscr{X}$. Then $\mathscr{D}(\mathscr{B}, h, \alpha)$ can be expressed as the PBD, $\mathscr{D}(\mathscr{B}_*, h_*, \alpha)$ with an explicit expression of h_* as shown in Theorem 1 of Sethuraman and Hollander (2009). We restate the theorem as Result 2 here without proof. Consequently the results derived here holds true for any arbitrary partition. Thus we can choose any partition of $\mathscr{X}$, but in practice we choose the partition by looking at the right-censored and/or left-truncated data for computational simplicity.

Result 2. Let $\mathscr{B} = (B_1, B_2, \ldots, B_m)$ be a measurable partition of $\mathscr{X}$ and consider the PBD $\mathscr{D}(\mathscr{B}, h, \alpha)$ where h is a pdf defined in (4.5) on the simplex R^m and α is a finite measure on $(\mathscr{X}, \mathscr{A})$ with $\alpha(B_r) > 0$ for $r = 1, 2, \ldots, m$. Split the set B_r as $B_{r_1} \cup B_{r_2}$ with $\alpha(B_{r_1}) > 0, \alpha(B_{r_2}) > 0$ such that $\mathscr{B}_* = (B_1, B_2, \ldots, B_{r-1}, B_{r_1}, B_{r_2}, B_{r+1}, \ldots, B_m)$ is a measurable partition of $\mathscr{X}$. Let

$$\begin{aligned} &h_*(y_1, \ldots, y_{r-1}, y_{r_1}, y_{r_2}, y_{r+1}, \ldots, y_m) \\ &\qquad = c_* h(y_1, \ldots, y_{r-1}, y_r, y_{r+1}, \ldots, y_m) \frac{y_{r_1}^{\alpha(B_{r_1})-1} y_{r_2}^{\alpha(B_{r_2})-1}}{y_r^{\alpha(B_r)-1}}, \end{aligned}$$

where $y_r = y_{r_1} + y_{r_2}$ and c_* is a normalizing constant. Then, $\mathscr{D}(\mathscr{B}, h, \alpha) = \mathscr{D}(\mathscr{B}_*, h_*, \alpha)$.

Suppose that T_1 is a complete observation and $P \mid T_1 \sim P$. Then the posterior measure of P given T_1 is given as Theorem 1.

Theorem 1. Let P have a PBD prior measure, $\mathscr{D}(\mathscr{B}, h, \alpha)$ and T_1 be a sample from P such that $T_1 \mid P \sim P$. Let $\mathscr{B} = (B_1, \ldots, B_m)$ be a measurable partition of $\mathscr{X}$ and r be an index of $\{1, 2, \ldots, m\}$ such that $T_1 \in B_r$. Then the posterior distribution of P, given T_1, is also PBD, $\mathscr{D}(\mathscr{B}, h^*, \alpha^*)$, where $h^*(\mathbf{y}) \propto h(\mathbf{y}) y_r$ and $\alpha^* = \alpha + \delta_{T_1}$.

Proof: The proof follows from the Theorem 3 of Sethuraman and Hollander (2009) with $T_1 \mid P \sim P = P_{\mathscr{X}}$, that is, $P(\mathscr{X}) \equiv y_1 + \ldots + y_m = 1$. Thus, the denominator of the Theorem 3 of Sethuraman and Hollander (2009) becomes 1. Hence the theorem is proved. □

Suppose that T_2 is a right-censored observation and we only know that $T_2 \in A$ and A is the union of some sets in the partition $\mathscr{B} = (B_1, B_2, \ldots, B_m)$, so that A satisfies

$$A = \cup_{l \in E_A} B_l \quad \text{for some} \quad E_A \subseteq \{1, 2, \ldots, m\}. \tag{4.6}$$

Theorem 2 follows from the Theorem 1 of Grego et al. (2013), which shows that when the prior is a PBD, given the right-censored observations, the posterior is also a PBD.

Theorem 2. Let P have a PBD measure $\mathscr{D}(\mathscr{B}, h, \alpha)$ and let T_2 be a sample of size one from P such that $T_2 \mid P \sim P$ and $T_2 \in A$, with A as defined in (4.6). Then the posterior measure of P, given $T_2 \in A$, is also PBD, $\mathscr{D}(\mathscr{B}, h^{**}, \alpha)$ where $h^{**}(\mathbf{y}) \propto h(\mathbf{y})y_A$ and $y_A = \sum_{l \in E_A} y_l$. Hence the theorem is proved. $\square$

Theorem 2 is obtained when A is a union of sets in the partition $\mathscr{B}$. When A is not necessarily a union of sets in the partition $\mathscr{B}$, we can form a larger partition $\mathscr{B}_{**}$ with restriction set A to the initial partition $\mathscr{B}$ which will ensure that A is a union of sets in the partition $\mathscr{B}_{**}$. It follows from the Result 2 that $\mathscr{D}(\mathscr{B}, h, \alpha) = \mathscr{D}(\mathscr{B}_{**}, h_{**}, \alpha_{**})$. Therefore, whether A is the union of sets of the partition $\mathscr{B}$ or not Theorem 2 holds.

Let $T_1, T_2, \ldots, T_n$ be a random sample from P such that $(T_i \mid P) \sim P, i = 1, 2, \ldots, n$. Without loss of generality, assume that $T_1, T_2 \ldots, T_{n-m}$, are the uncensored (complete) observations and $T^*_{(1)}, T^*_{(2)}, \ldots, T^*_{(m)}$ are m distinct right-censored observations, where $T^*_i = T_{n-m+i}, i = 1, 2, \ldots, m$. Consider $T^*_{(1)}, T^*_{(2)}, \ldots, T^*_{(m)}$ be the partition boundaries such that

$$\mathscr{B} = (B_1, B_2, \ldots, B_m, B_{m+1}) \tag{4.7}$$

is a measurable partition of $(0, \infty)$, where $B_l = (T^*_{(l-1)}, T^*_{(l)}], l = 1, 2, 3, \ldots, m$, and $B_{m+1} = (T^*_{(m)}, T^*_{(m+1)})$ with $T^*_{(0)} = 0$ and $T^*_{(m+1)} = \infty$. The posterior measure of P given all the censored and uncensored observations is given as Theorem 3.

Theorem 3. Let P have a PBD prior measure, $\mathscr{D}(\mathscr{B}, h, \alpha)$, where $\mathscr{B}$ is as defined in (4.7). Let $T_1, T_2 \ldots, T_{n-m}, T^*_{(1)}, T^*_{(2)}, \ldots, T^*_{(m)}$ be a sample from P with $T_1, T_2 \ldots, T_{n-m}$ being the ordered uncensored observations and $T^*_{(1)}, T^*_{(2)}, \ldots, T^*_{(m)}$ being the m right-censored observations. Then the posterior measure, $P \mid (T_1, T_2 \ldots, T_{n-m}, T^*_{(1)}, T^*_{(2)}, \ldots,$

$T^*_{(m)}$) is the PBD, $\mathscr{D}(\mathscr{B}, h^*, \alpha^*)$, where

$$h^*(\mathbf{y}) \;\propto\; h(\mathbf{y}) \left[\prod_{l=1}^{m+1} y_l^{\sum_{i=1}^{n-m} I(T_i \in B_l)} \right] \prod_{l=1}^{m} \left[\sum_{j=l+1}^{m+1} y_j \right] \;\propto\; \left[\prod_{l=1}^{m+1} y_l^{\alpha_l^* - 1} \right] \prod_{l=1}^{m} \left[\sum_{j=l+1}^{m+1} y_j \right],$$

and $\alpha^* = \alpha + \sum_{i=1}^{n-m} \delta_{T_i}$.

Proof: Repeated application of Theorem 1 for complete observations and repeated application of Theorem 2 for right-censored observations together yield the desired result. □

4.3 Nonparametric Bayes Inference of System Reliability

4.3.1 System Reliability Based on System Data

Recall that the random observables for the system lifetimes are $((V_1, \epsilon_1), (V_2, \epsilon_2), \ldots, (V_n, \epsilon_n))$ where $V_i = \min(S_i, \tau_i)$ and $\epsilon_i = I(S_i \le \tau_i)$ for $i = 1, 2, \ldots, n$. Assume that there are m right-censored observations and $n-m$ complete observations. To distinguish right-censored observations from complete observations, let $\{V_i', i = 1, 2, \ldots, n-m\}$ be the $n-m$ complete observations ($\epsilon_i = 1$) and $\{V_j^*, j = 1, 2, \ldots, m\}$ be the m right-censored ($\epsilon_i = 0$) observations. Let $V^*_{(1)}, V^*_{(2)}, \ldots, V^*_{(m)}$ be the partition boundaries such that

$$\mathscr{B} = (B_1, B_2, \ldots, B_m, B_{m+1})$$

is a measurable partition of $(0, \infty)$, where $B_l = (V^*_{(l-1)}, V^*_{(l)}], l = 1, 2, 3, \ldots, m$, and $B_{m+1} = (V^*_{(m)}, V^*_{(m+1)})$ with $V^*_{(0)} = 0$ and $V^*_{(m+1)} = \infty$. Our goal is to develop a nonparametric Bayes estimator of $F(t) \equiv P((0, t]) \equiv P$ when a PBD process prior is assigned on F and the right-censored system lifetimes are available. Let P have PBD prior measure, $\mathscr{D}(\mathscr{B}, h, \alpha)$, where $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m, \alpha_{m+1})$ and $\alpha_l = \alpha(B_l), l = 1, 2, \ldots, m+1$. Then the posterior measure of P is given in Corollary 1, follows directly from the Theorem 3.

Corollary 1. Let P have PBD prior $\mathscr{D}(\mathscr{B}, h, \alpha)$ and $\mathscr{V} = (V_1', V_1', \ldots, V'_{n-m}, V^*_{(1)}, V^*_{(2)},$

$\ldots, V^*_{(m)})$ is a random sample from P with $V^*_{(1)}, V^*_{(2)}, \ldots, V^*_{(m)}$ being the m right-censored observations. Then the posterior measure, $P^* \equiv (P \mid \mathscr{V})$, is also a PBD measure, $\mathscr{D}(\mathscr{B}, h^*, \alpha^*)$, where

$$h^*(\mathbf{y}) \propto h(\mathbf{y}) \prod_{l=1}^{m+1} \left[y_l^{\sum_{i=1}^{n-m} I(V'_i \in B_l)} \right] \prod_{l=1}^{m} \left[\sum_{j=l+1}^{m+1} y_j \right] \propto \prod_{l=1}^{m+1} \left[y_l^{\alpha^*_l - 1} \right] \prod_{l=1}^{m} \left[\sum_{j=l+1}^{m+1} y_j \right]$$

and $\alpha^* = \alpha + \sum_{i=1}^{n-m} \delta_{V'_i}$.

We obtain a nonparametric Bayes estimator of F, given a vector of random observables $\mathscr{V}$, under the integrated squared-error loss function

$$L(\widehat{F}, F) = \int [\widehat{F}(s) - F(s)]^2 ds, \tag{4.8}$$

where $\widehat{F}(s)$ is an estimator of $F(s)$.

Suppose that we want to estimate $F(t)$ where $t \in B_l, (l = 1, 2, \ldots, m, m+1)$. Note that $P(\cdot) = \sum_{i=1}^{m+1} P_{B_i}(\cdot) P(B_i)$ where $P(\mathscr{B})$ and $\{P_{B_l}\}$ are independent. Under the integrated squared-error loss function (4.8), the nonparametric Bayes estimator of $F(t)$ is the posterior mean which is

$$\hat{F}(t) \equiv E[P^*((0,t])] = \sum_{j=1}^{l-1} E(Y_j) + E(Y_l) \left[\alpha^*(B_l \cap (0,t]) / \alpha^*(B_l) \right] \tag{4.9}$$

where the expectations of $\mathbf{Y}$ are under the pdf proportional to $h^*(\mathbf{y})$. Therefore, an estimate of the system reliability function is $\widehat{\bar{F}}(t) = 1 - \widehat{F}(t)$. For $j = 1, 2, \ldots, m, m+1$, $E(Y_j)$ is given by

$$E(Y_j) = \frac{\int y_j h^*(\mathbf{y}) d\mathbf{y}}{\int h^*(\mathbf{y}) d\mathbf{y}}. \tag{4.10}$$

To see $E(Y_j), j = 1, 2, \ldots, m, m+1$, and thus an estimator of $\bar{F}$ in closed form we also define the following processes

$$N_j = \sum_{i=1}^{n} I(V_i \in B_j, \epsilon_i = 1) \quad \text{and} \quad \lambda_j = \sum_{l=1}^{m} I(V^*_l = V^*_{(j)}), \tag{4.11}$$

$j = 1, 2, \ldots, m, m+1$, with $\lambda_{m+1} = 0$.

Lemma 1. Let $Y(t), Y^+(t)$, and $N_j, \lambda_j, j = 1, 2, \ldots, m, m+1$, be the processes defined in (4.4) and (4.11), respectively. Then,

$$
\begin{aligned}
E(Y_j) = \frac{\int y_j h^*(\mathbf{y}) d\mathbf{y}}{\int h^*(\mathbf{y}) d\mathbf{y}} &= \frac{C_j(\alpha_l^* + I(l=j), \lambda_l, l = 1, 2, \ldots, m+1)}{C_j(\alpha_l^*, \lambda_l, l = 1, 2, \ldots, m, m+1)} \\
&= \left[\frac{\alpha(B_j) + N(B_j)}{\alpha(\Re_+) + n}\right] \left[\prod_{i=1}^{j-1} \left\{\frac{\alpha(V_{(i)}^*, \infty) + Y^+(V_{(i)}^*) + \lambda_i}{\alpha(V_{(i)}^*, \infty) + Y^+(V_{(i)}^*)}\right\}\right],
\end{aligned}
$$

where

$$
\begin{aligned}
&C_j(\alpha_l^* + I(l=j), \lambda_l, l = 1, 2, \ldots, m+1) \\
&\qquad = \prod_{l=1}^{m} \left[B(\alpha^*(B_l) + I(l=j), \alpha^*[V_{(l)}^*, \infty) + \sum_{r=l+1}^{m} \lambda_r)\right]
\end{aligned}
$$

The nonparametric Bayes estimator of system reliability function $\bar{F}(t) = 1 - F(t)$ is given as Theorem 4, which follows from the expression (4.9) using Lemma 1.

Theorem 4. Let P have PBD prior, $\mathscr{D}(\mathscr{B}, h, \alpha)$, and $\mathscr{V} = (V_1', \ldots, V_{n-m}', V_{(1)}^*, V_{(2)}^*, \ldots, V_{(m)}^*)$ be a vector of observables. Then for any $t \in B_l \subseteq \mathscr{B}$, the nonparametric Bayes estimator of $F(t)$ is given by

$$
\begin{aligned}
\widehat{F}_{NPB}(t) &= \sum_{j=1}^{l-1} E(Y_j) + E(Y_l)\alpha^*(B_l \cap (0,t])/\alpha^*(B_l) \\
&= \sum_{j=1}^{l-1} E(Y_j) + E(Y_l)\left[\frac{\alpha(B_l \cap (0,t]) + N(B_l \cap (0,t])}{\alpha(B_l) + N(B_l)}\right],
\end{aligned}
$$

where $E(Y_l), l = 1, 2, \ldots, m, m+1$, is given in Lemma 1 and $\mathscr{V}$ as defined in Corollary 1.

Equivalently,

$$
\widehat{\bar{F}}_{NPB}(t) = \sum_{j=l+1}^{m+1} E(Y_j) + E(Y_l)\left[\frac{\alpha(B_l \cap (t,\infty)) + N(B_l \cap (t,\infty))}{\alpha(B_l) + N(B_l)}\right]. \tag{4.12}
$$

The PL-type estimator in (4.3) is a limiting case of the proposed nonparametric Bayes estimator $\widehat{\bar{F}}_{NPB}(t)$ when prior measure $\alpha \to 0$, as given in Theorem 5.

Theorem 5.

$$
\widehat{\bar{F}}_{NPB}(t) \xrightarrow{\alpha(R^+) \to 0} \widehat{R}_{PLE}(t).
$$

4.3.2 Pointwise Credible Intervals

To construct pointwise credible intervals for $\bar{F}(t)$, we take sample from the posterior measure given in Corollary 1. The density function associated with the posterior measure is given by

$$h^*(\mathbf{y}) \propto h(\mathbf{y}) \prod_{l=1}^{m+1} \left[y_l^{\sum_{i=1}^{n-m} I(V_i' \in B_l)} \right] \prod_{l=1}^{m} \left[\sum_{i=l+1}^{m+1} y_i \right] \propto y_{m+1}^{\alpha_{m+1}^*-1} \prod_{l=1}^{m} \left[y_l^{\alpha_l^*-1} \left(1 - \sum_{j=1}^{l} y_j \right)^{\lambda_l} \right], \tag{4.13}$$

where $\alpha^* = \alpha + \sum_{i=1}^{n-m} \delta_{V_i'}$, which is also proportional to the so-called generalized Dirichlet distribution (see Connor and Mosimann (1969)). To sample from the posterior measure, consider a well known transformations

$$X_l = Y_l + Y_{l+1} + \ldots + Y_{m+1}, l = 1, 2, \ldots, m+1.$$

Define,

$$W_l = \frac{X_{l+1}}{X_l}, l = 1, 2, \ldots, m.$$

Simplification yields that

$$Y_1 = 1 - W_1, Y_2 = W_1(1 - W_2), \ldots, Y_m = (1 - W_m) \prod_{j=1}^{m-1} W_j, \;\; Y_{m+1} = \prod_{j=1}^{m} W_j. \tag{4.14}$$

Straight-forward derivations show that $W_1, W_2, \ldots, W_m$ have independent beta distributions with

$$W_1 \sim Beta(A_1, \alpha_1^*), W_2 \sim Beta(A_2, \alpha_2^*), \ldots, W_m \sim Beta(A_m, \alpha_m^*),$$

where $A_j = \alpha^*[V_{(j)}^*, \infty) + \sum_{j=1}^{m} \lambda_j = \alpha[V_{(j)}^*, \infty) + N([V_{(j)}^*, \infty)) + \sum_{j=1}^{m} \lambda_j, j = 1, 2, \ldots, m$. One may now take samples of $W_1, W_2, \ldots, W_m$, and then obtain $Y_1, Y_2, \ldots, Y_{m+1}$ using (4.14). An approximate posterior mean and thus nonparametric Bayes estimate and point-wise credible intervals of $\bar{F}(t)$ follows from the posterior samples.

4.3.3 Joint Estimation of System and Components Reliabilities

We now consider the joint estimation of component and system reliabilities when n identical systems each with K components are under study. Denote the lifetime of component j in the i-th system by $\{T_{ij}\}$ and let $(0, \tau_i]$ be the monitoring period for the i-th system. Assume that $\{T_{ij}, i = 1, 2 \ldots, n\}$ are IID with distribution F_j, and $\{T_{ij}\}$ and $\{\tau_i\}$ are independent, T_{ij} and T_{il} are also independent for $j \neq l$. Recall that the random observables for the j-th component are

$$\{(Z_{ij}, \delta_{ij}), i = 1, 2, \ldots, n\}, j = 1, 2, \ldots, K.$$

We assign independent PBD priors on $F_j, j = 1, 2, \ldots, K$ and obtain nonparametric Bayes estimators of $F_j, j = 1, 2, \ldots, K$, given the random observable $\{(Z_{ij}, \delta_{ij}), i = 1, 2, \ldots, n\}$ for the j-th component. Without loss of generality assume that first $n - m_j$ are the complete observations and last m_j are the right-censored observations for the j-th component. To distinguish the censored observations from the complete observations, define $Z'_{ij} = Z_{ij}$ if $\delta_{ij} = 1$ and $Z^*_{ij} = Z_{ij}$ if $\delta_{ij} = 0$. Then the observables for the j-th components are $\{Z'_{1j}, Z'_{2j}, \ldots, Z'_{n-m_j j}, Z^*_{1j}, Z^*_{2j}, \ldots, Z^*_{m_j j}\}$.

To specify a measurable partition of $(0, \infty)$ for F_j, let $Z^*_{(1)j}, Z^*_{(2)j}, \ldots, Z^*_{(m_j)j}$ be the partition boundaries so that

$$\mathscr{B}_j = (B_{1j}, B_{2j}, \ldots, B_{m_j j}, B_{m_j+1j}) \tag{4.15}$$

is a measurable partition of $(0, \infty)$, where

$$B_{lj} = (Z^*_{(l-1)j}, Z^*_{(l)j}], l = 1, 2, \ldots, m_j, \text{and } B_{(m_j+1)} = (Z^*_{(m_j)j}, Z^*_{(m_j+1)j})$$

with $Z^*_{(0)j} = 0$ and $Z^*_{(m_j+1)j} = \infty$. Define $F_j(t) = P_j((0, t])$ and $P_j(B_{lj}) = Y_{lj}$. We assume that the random probability measure P_j has a PBD prior measure, $\mathscr{D}(\mathscr{B}_j, h_j, \alpha_j)$, where

$$\alpha_j = (\alpha_{1j}, \alpha_{2j}, \ldots, \alpha_{m_j j}, \alpha_{m_j+1j}), \text{ with } \quad \alpha_{lj} = \alpha_j(B_{lj}), l = 1, 2, \ldots, m_j + 1, \tag{4.16}$$

and

$$h_j \equiv h(\mathbf{y}_j) = c_j \prod_{l=1}^{m_j+1} y_{lj}^{\alpha_{lj}-1}, \quad 0 \le y_{lj} \le 1, \quad l = 1, \ldots, m_j, m_j+1, \quad \sum_{l=1}^{m_j+1} y_{lj} = 1 \tag{4.17}$$

with normalizing constant c_j. Then the posterior measure given in Corollary 2 follows from the Theorem 3.

Corollary 2. Let P_j have PBD prior, $\mathscr{D}(\mathscr{B}_j, h_j, \alpha_j)$ and $\{Z'_{1j}, Z'_{2j}, \ldots, Z'_{n-m_j j}, Z^*_{(1)j}, Z^*_{(2)j}, \ldots, Z^*_{(m_j)j}\}$ be a random observable from P_j with $\{Z^*_{(1)j}, Z^*_{(2)j}, \ldots, Z^*_{(m_j)j}\}$ are right-censored observations. Then, the posterior measure, $P^*_j \equiv (P_j \mid (Z'_{1j}, Z'_{2j}, \ldots, Z'_{n-m_j j}, Z^*_{(1)j}, Z^*_{(2)j}, \ldots, Z^*_{(m_j)j}))$, is a PBD measure, $\mathscr{D}(\mathscr{B}_j, h^*_j, \alpha^*_j)$, where

$$h^*_j \equiv h^*(\mathbf{y}_j) \propto h_j(\mathbf{y}_j) \left[\prod_{l=1}^{m_j+1} y_{lj}{}^{\sum_{i=1}^{n-m_j} I(Z'_{ij} \in B_{lj})} \right] \prod_{l=1}^{m_j} \left[\sum_{l'=l+1}^{m_j+1} y_{l'j} \right]$$

and

$$\alpha^*_j = \alpha_j + \sum_{i=1}^{n-m_j} \delta_{z'_{ij}}.$$

□

Under the integrated squared-error loss function (4.8), a nonparametric Bayes estimator of the j-th component distribution function is

$$\widehat{F}_j(t) \equiv E[P^*_j((0,t])] = \sum_{r=1}^{l-1} E(y_{rj}) + E(y_{lj}) \left[\alpha^*_j(B_{lj} \cap (0,t]) / \alpha^*_j(B_{lj}) \right],$$

where the expectation of $\mathbf{y}_j$ is under the pdf proportional to $h^*(\mathbf{y}_j)$ and for $l = 1, 2, \ldots, m_j, m_j+1$ is given by

$$E(Y_{lj}) = \frac{\int y_{lj} h^*(\mathbf{y}_j) d\mathbf{y}_j}{\int h^*(\mathbf{y}_j) d\mathbf{y}_j}. \tag{4.18}$$

Therefore an estimate of the j-th components reliability function is given by $\hat{\bar{F}}_j(t) = 1 - \widehat{F}_j(t)$. Applying Lemma 1 and Theorem 4, we can then obtain a closed form estimate of the reliability function for each component and denote those by $\hat{\bar{F}}_j, j = 1, 2, \ldots, K$. Therefore an estimate of the system reliability function can be expressed in terms of the components reliability function using (4.1) is

$$\hat{\bar{F}}_\phi(t) = h_\phi(\hat{\bar{F}}_1(t), \hat{\bar{F}}_2(t), \ldots, \hat{\bar{F}}_K(t)). \tag{4.19}$$

Denote the corresponding Doss et al. (1989) PL-type estimator of system reliability function by

$$\widehat{R}_\phi(t) = h_\phi(\widehat{R}_{1,PLE}(t), \widehat{R}_{2,PLE}(t), \ldots, \widehat{R}_{K,PLE}(t)), \tag{4.20}$$

where $\widehat{R}_{j,PLE}(t), j = 1, 2, \ldots, K$, are the PL-type estimators of the components reliability function as defined in (4.3). The Doss et al. (1989) estimator (4.20) is a limiting case of our proposed estimator (4.19) when the prior measure $\alpha_j \to 0$ for each $j = 1, 2, \ldots, K$, as given in Theorem 7. The proof of Theorem 7 follows from Theorem 5.

Theorem 7.

$$\widehat{\bar{\bar{F}}}_\phi(t) \xrightarrow{\alpha_j(R^+)\to 0} \widehat{R}_\phi(t).$$

4.4 Monte Carlo Studies

4.4.1 Simulation Studies I: Prior Mean Function Coincide with True F

Simulation studies were carried out to examine the biases and root-mean-squared errors (RMSEs) of the proposed nonparametric Bayes estimator of the system reliability function based on system lifetime data, denoted by $\widehat{\bar{\bar{F}}}_{NPB}(t)$ (labeled BayesSys), and components lifetimes data, denoted by $\widehat{\bar{\bar{F}}}_\phi(t)$ (labeled BayesPhi), as well as corresponding nonparametric (PL-type) estimators denoted by $\widehat{R}_{PLE}(t)$ (labeled PLESys) and $\widehat{R}_\phi(t)$ (labeled PLEPhi). We consider the three component series-parallel system (Figure 4.1) with component lifetimes, $T_{ij} \sim \text{Exp}(\theta_j)$, j=1, 2, 3; $\theta = (1, 2, 1.5)$, and monitoring time $\tau_i \sim \text{Exp}(1)$. Simulated biases and RMSEs are obtained at the 5th, 10th , . . . , 95th percentile of the true data generating distribution based on 1000 replications for n=30, and compared with the corresponding nonparametric estimators. To compute $\widehat{\bar{\bar{F}}}_\phi(t)$ we assign independent prior measures $\alpha_j(t, \infty) = \beta_j \exp(-\theta_j t)$ with $\theta_1 = 1, \theta_2 = 2, \theta_3 = 1.5$ and $\beta_j = 10$, j=1, 2, 3, on the component distribution

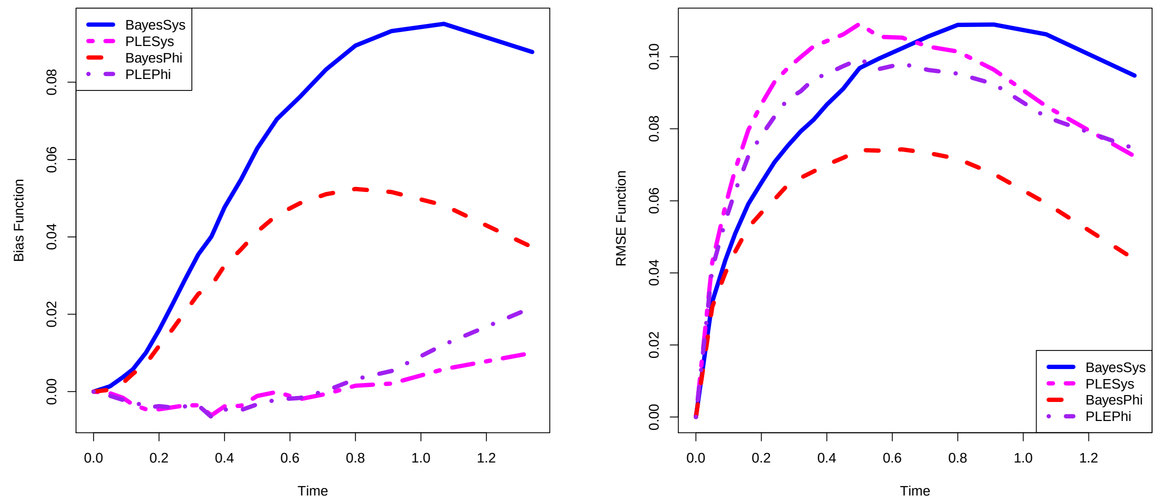

Figure 4.2: Simulated biases and RMSEs of the estimators $\hat{\bar{\bar{F}}}_{NPB}(t)$ (labeled BayesSys), $\hat{\bar{\bar{F}}}_{\phi}(t)$ (labeled BayesPhi), $\widehat{R}_{PLE}(t)$ (labeled PLESys), and $\widehat{R}_{\phi}(t)$ (labeled PLEPhi). Simulation parameters are $n = 30$, lifetimes (3-component series-parallel system) from Exponential distribution with rate $\theta = (1, 2, 1.5)$, and with 1,000 replications.

functions $F_j, j = 1, 2, 3$, respectively. That is, each prior measure is β_j times an exponential survivor function with parameter θ_j, where β_j may be viewed as a precision of the prior measure. Note that in this case the prior mean functions, $\bar{\alpha}_j$ coincide with component true reliability functions $\bar{F}_j, j = 1, 2, 3$. To compute $\hat{\bar{\bar{F}}}_{NPB}(t)$ we also assign a similar prior measure with $\theta = 1$ and $\beta = 1$. In this case the prior mean function does not coincide with the true distribution of system reliability. Even for the simple case, when components distributions are all exponential, the system lifetime distribution is no longer exponential distribution.

Figure 4.2 demonstrates that both $\hat{\bar{\bar{F}}}_{NPB}(t)$ and $\hat{\bar{\bar{F}}}_{\phi}(t)$ possess larger biases but smaller RMSEs than the nonparametric estimators $\widehat{R}_{PLE}$ and $\widehat{R}_{\phi}$, respectively. By examining Figure 4.2, it is evident that $\hat{\bar{\bar{F}}}_{\phi}$ demonstrates smaller biases and RMSEs than $\hat{\bar{\bar{F}}}_{NPB}(t)$ and $\widehat{R}_{PLE}$ exhibits slightly larger bias and RMSE than $\widehat{R}_{\phi}$. Among all the four estimators $\hat{\bar{\bar{F}}}_{\phi}(t)$ demonstrates smallest RMSEs. Therefore the Bayes estimator of system reliability function, $\hat{\bar{\bar{F}}}_\phi$ based on the component lifetimes data and thus components reliability functions outperforms other estimators. However, in practice it is unlikely that our prior measure will coincide with the true distribution function. So we carried out other simulation studies when prior measures are misspecified.

4.4.2 Simulation Studies II: Prior Mean Functions do not Coincide with True $\bar{F}$

We also investigated the biases and RMSEs in the case of misspecified prior measures, that is, when $\bar{\alpha}_j$ differs from the component true lifetime generating distributions $\bar{F}_j, j = 1, 2, 3$. In particular, for each component, IID lifetimes are generated from the Weibull distribution with scale parameter θ_j and shape parameter γ_j with $\theta = (1, 1, 1), \gamma = (2, 1.5, 1.2)$ and random monitoring $\tau_i \sim \text{Exp}(1)$. However, we assign prior measure $\alpha_j(t, \infty) = \beta_{\mathbf{j}} \exp[-\theta_j t]$ instead of $\alpha_j(t, \infty) = \beta_{\mathbf{j}} \exp[-(\theta_j t)^{\gamma_j}]$ with $\theta = (1, 1, 1)$, and $\beta =$ (1,1,1) such that $\bar{\alpha}_j$ is proportional to an exponential survivor function and thus the prior mean function differs from the true data generating distribution for each component. We also choose $\beta =$(10,10,10) and (20,20,20) to examine the effects of precision parameter β when priors are misspecified.

From Figure 4.3, it is obvious that for a smaller value of the precision parameter, namely $\beta = (1, 1, 1)$, $\hat{\bar{\bar{F}}}_\phi(t)$ has smaller RMSE than $\widehat{R}_\phi(t)$ even in the case of misspecified prior measures. As the precision of the prior measure β increases, $\hat{\bar{\bar{F}}}_\phi(t)$ demonstrates higher biases but smaller RMSE's than $\widehat{R}_\phi(t)$ except for smaller values of t. Figure 4.3 therefore indicates that nonparametric Bayes estimator $\hat{\bar{\bar{F}}}_\phi(t)$ is robust in the sense that it does not suffer significantly due to a misspecification of the prior measures. When prior measures are misspecified, a larger magnitude of the precision parameter β produces larger biases and RMSE's for the Bayes estimators $\hat{\bar{\bar{F}}}_\phi$ and $\hat{\bar{\bar{F}}}_{NPB}$. The effect of misspecification can be restrained by choosing smaller values of the precision parameters.

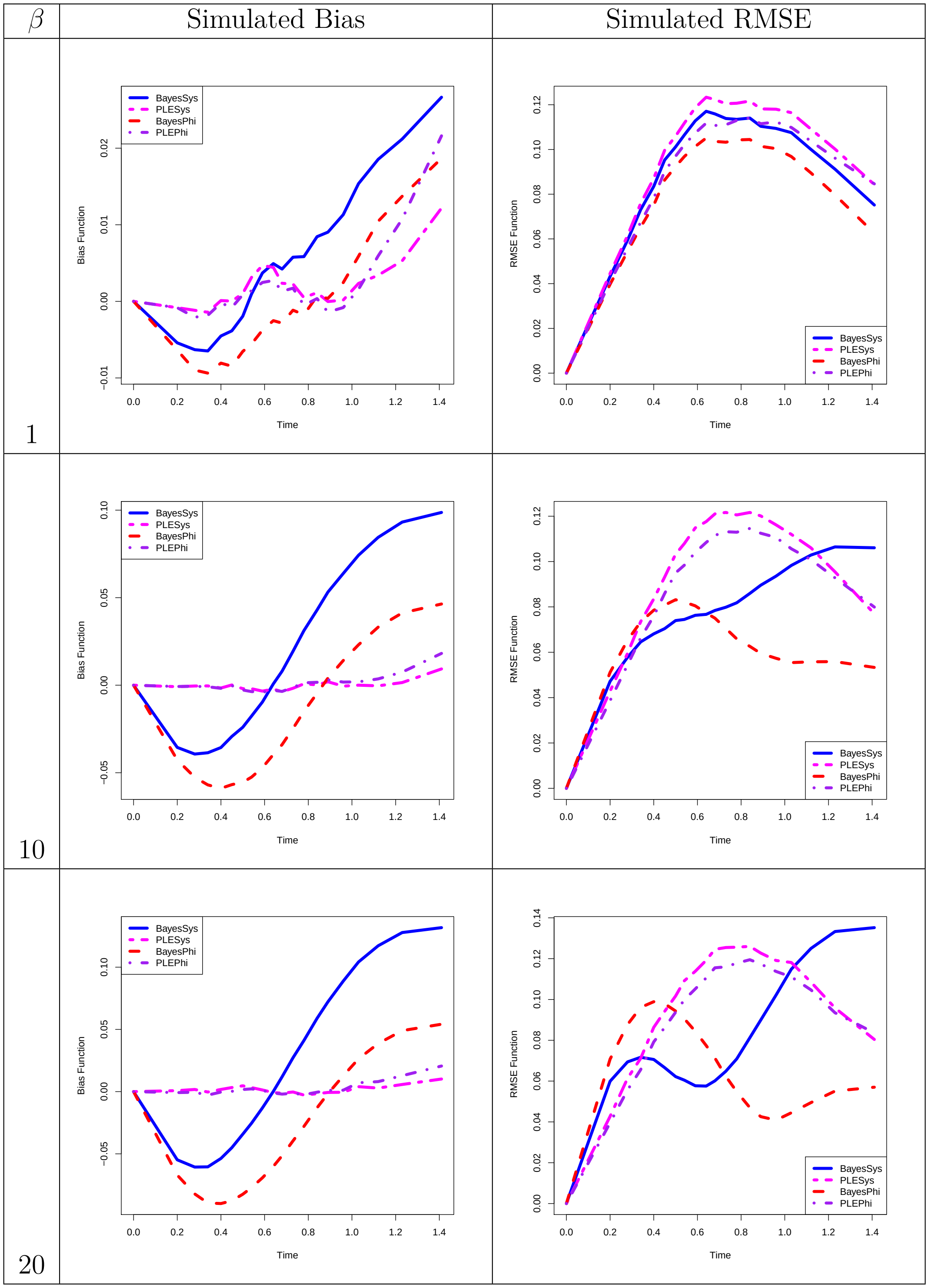

Figure 4.3: Simulated biases and RMSEs of the estimators $\widehat{\widehat{F}}_{NPB}(t)$ (labeled BayesSys), $\widehat{\widehat{F}}_{\phi}(t)$ (labeled BayesPhi), $\widehat{R}_{PLE}(t)$ (labeled PLESys), and $\widehat{R}_{\phi}(t)$ (labeled PLEPhi). Simulation parameters are $n = 30$, $\theta = (1, 1, 1)$ and $\gamma = (2, 1.5, 1.2)$ (Weibull(θ, γ)) with 1000 replications. Mis-specified prior measures $\alpha(u, \infty) = \beta \exp(-\theta u)$, with $\theta = (1, 1, 1)$ and $\beta = $ (1, 1, 1), (10, 10, 10), (20,20,20).

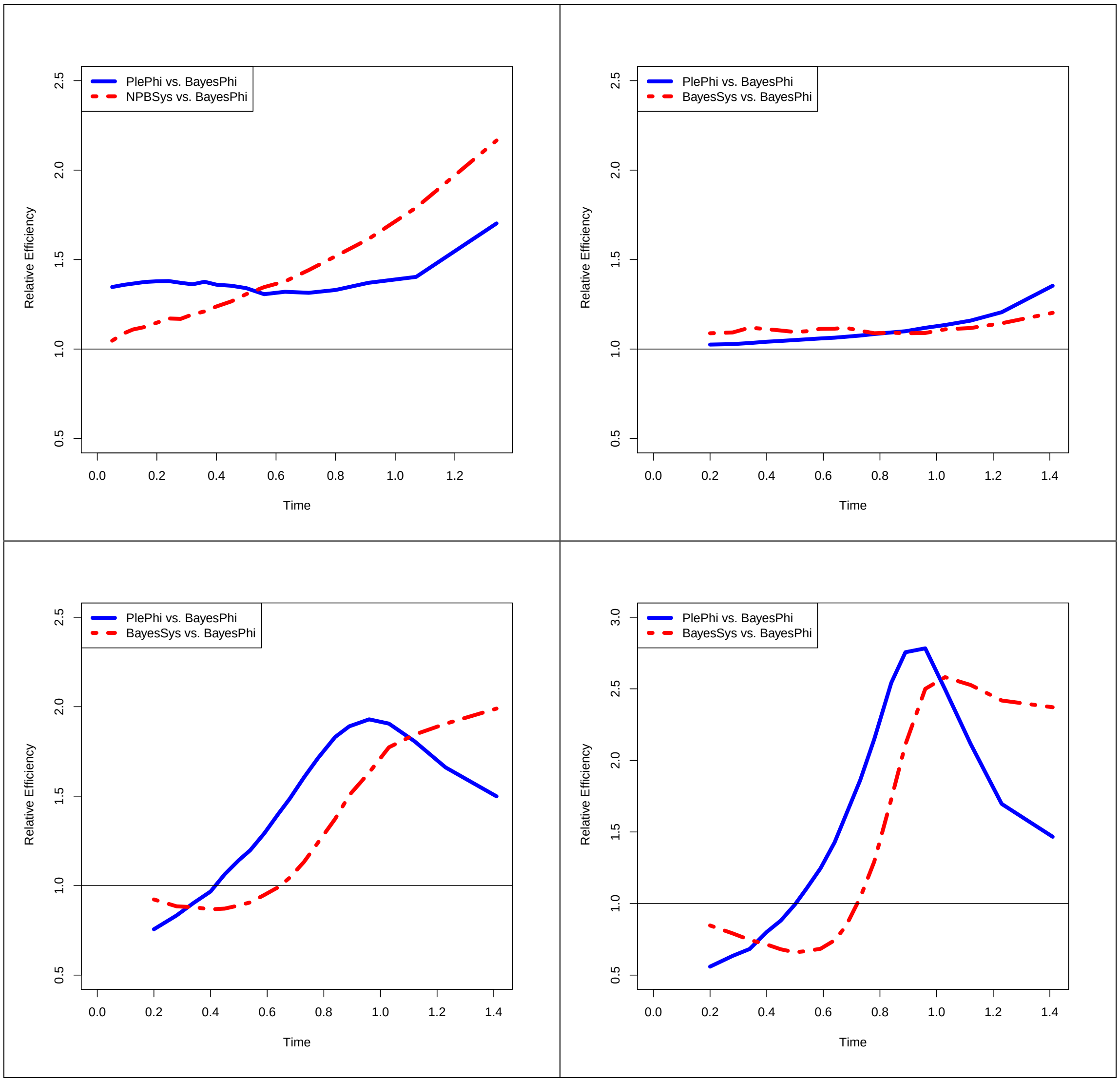

Figure 4.4: Asymptotic relative efficiency: true prior mean (top left), misspecified prior mean with β =1 (top right), β =10 (bottom left), and β = 20 (bottom right).

Relative efficiencies of different types of estimators of the system reliability function are obtained. Relative efficiency of $\bar{F}_{NPB}(t)$ and $\bar{F}_{\phi}(t)$ is defined by $\frac{Var(\bar{F}_{NPB}(t))}{Var(\bar{F}_{\phi}(t))}$. Similarly relative efficiency of $R_{\phi}(t)$ and $\bar{F}_{\phi}(t)$ is defined by $\frac{Var(\bar{F}_{\phi}(t))}{Var(R_{\phi}(t))}$. However, Figure 4.4 is based on RMSEs since Bayes estimators are biased. From Figure 4.4 (top left and right), it is evident that Bayes estimator $\bar{F}_{\phi}(t)$ is more efficient than the nonparametric estimator $R_{\phi}(t)$ both in the case of correctly specified prior measures and misspecified prior measures with smaller values of precision parameters.

4.5 Illustrative Example

We illustrate the proposed estimators with a randomly generated data set and compare it with nonparametric (PL-type) estimators. Again, we consider the three component series-parallel system (Figure 4.1) for data generation purposes. Assume that $T_{ij} \sim$ Weibull$(\theta_j,\ \gamma_j)$, $\theta = (1,1,1), \gamma = (2,1.5,1.2)$, and monitoring times, $\tau_i \sim$ Exp(1). Assign prior measure $\alpha_j(t,\infty) = \beta_j \exp[-(\theta_j t)^\gamma]$ with $\theta = (1,1,1)$, $\gamma = (1,1,1)$ and $\beta =$ (1,1,1) (right figure), that is, prior measures are misspecified. With $\theta = (1,1,1)$, $\gamma = (2,1.5,1.2)$ and $\beta =$ (10,10,10) (left figure) prior measures are correctly specified. Figure 4.5 demonstrates that nonparametric Bayes estimators of system reliability function, in particular, $\bar{F}_\phi(t)$, perform better than other estimators and is closer to the true reliability function than the other estimators. The right panel of Figure 4.5 indicates that Bayes estimators are robust in the sense that the effect of misspecification is not severe with smaller values of parameters β_j .

4.6 Concluding Remarks

The nonparametric Bayes estimator $\widehat{\bar{F}}_\phi(t)$ developed here served as a Bayesian counter part of the Doss et al. (1989) estimator $\widehat{R}_\phi(t)$. The Doss et al. (1989) estimator is a limiting case of our proposed estimator $\widehat{\bar{F}}_\phi(t)$. Bayes estimators of system reliability function are smoother in some sense than the corresponding nonparametric estimators. Simulation studies demonstrate that $\widehat{\bar{F}}_\phi(t)$ yields smaller RMSEs than $\widehat{R}_\phi(t)$. Simulation studies further demonstrate that (Figure 4.2 and Figure 4.3) $\widehat{\bar{F}}_\phi(t)$ and $\widehat{R}_\phi(t)$ perform better than $\bar{F}_{NPB}(t)$ and $\widehat{R}_{PLE}(t)$, respectively, in terms of RMSEs. Nonparametric Bayes estimators are robust in the sense that the effect of misspecification of prior measures is not severe with smaller values of precision parameters.

The PBD prior is an elegant nonparametric prior which provides succinct posterior calculation. Given the observations (left-truncated, interval-censored, and right-

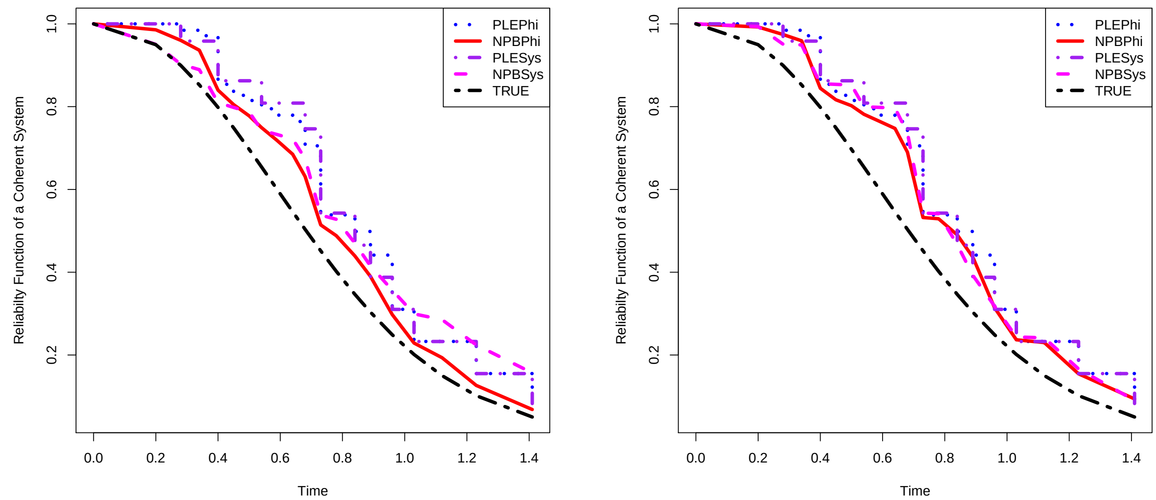

Figure 4.5: Reliability function of $\widehat{\widehat{F}}_{NPB}(t)$ (labeled NPBSys), $\widehat{\widehat{F}}_{\phi}(t)$ (labeled NPBPhi), $\widehat{R}_{PLE}(t)$ (labeled PLESys), and $\widehat{R}_{\phi}(t)$ (labeled PLEPhi) and the true distribution (labeled True). Priors are $\alpha(u,\infty) = \beta\exp\{-(u/\theta)^{\gamma}\}$ with θ = (1, 1, 1), γ = (2, 1.5, 1.2), β = (10,10,10) (left side graph), $\theta = (1,1,1), \gamma = (1,1,1), \beta = (1,1,1)$ (right side graph)

censored), the posterior measure is also a PBD measure when we assign a PBD prior. We derived closed form estimators as well as developed a procedure to sample from the posterior measure. Moreover, a PBD prior can conveniently handle left-truncated, interval-censored, and right-censored data. However, we did not consider here the case when data are left truncated as well as right-censored.

Bibliography

Aalen, O. O. and Husebye, E. (1991). Statistical analysis of repeated events forming renewal processes. *Statistics in Medicine*, 10(8):1227–1240.

Adekpedjou, A., Peña, E. A., and Quiton, J. (2010). Estimation and efficiency with recurrent event data under informative monitoring. *J. Statist. Plann. Inference*, 140(3):597–615.

Albert, P. S. (1991). A two-state markov mixture model for a time series of epileptic seizure counts. *Biometrics*, 47(4):pp. 1371–1381.

Andersen, P. K., Borgan, Ø., Gill, R. D., and Keiding, N. (1993). *Statistical models based on counting processes*. Springer Series in Statistics. Springer-Verlag, New York.

Andersen, P. K. and Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. *Ann. Statist.*, 10(4):1100–1120.

Barlow, R. E. (1984). Reliability calculation and inductive inference: an overview. In *Reliability theory and models (Charlotte, N.C., 1983)*, volume 10 of *Notes Rep. Comput. Sci. Appl. Math.*, pages 221–229. Academic Press, Orlando, FL.

Barlow, R. E. (1985). Combining component and system information in system reliability calculation. In *Probabilistic methods in the mechanics of solids and structures (Stockholm, 1984)*, pages 375–383. Springer, Berlin.

Barlow, R. E. (1986). Mathematical theory of reliability. Historical perspectives. In *Theory of reliability (Varenna, 1984)*, Proc. Internat. School of Phys. Enrico Fermi, XCIV, pages 3–11. North-Holland, Amsterdam.

Barlow, R. E. and Hunter, L. C. (1960). System efficiency and reliability. *Technometrics*, 2:43–53.

Barlow, R. E. and Marshall, A. W. (1967). Bounds on interval probabilities for restricted families of distribution. In *Proc. Fifth Berkeley Sympos. Mathematical Statistics and Probability (Berkeley, Calif., 1965/66), Vol. III: Physical Sciences*, pages 229–257. Univ. California Press, Berkeley, Calif.

Barlow, R. E. and Proschan, F. (1969). A note on tests for monotone failure rate based on incomplete data. *Ann. Math. Statist.*, 40:595–600.

Barlow, R. E. and Proschan, F. (1981). *Statistical theory of reliability and life testing: Probability Models*. Silver Spring, MD: To Begin With.

Barlow, R. E. and Proschan, F. (1986). Distributions with monotone failure rate. In *Theory of reliability (Varenna, 1984)*, Proc. Internat. School of Phys. Enrico Fermi, XCIV, pages 12–22. North-Holland, Amsterdam.

Boyles, R. A., Marshall, A. W., and Proschan, F. (1985). Inconsistency of the maximum likelihood estimator of a distribution having increasing failure rate average. *Ann. Statist.*, 13(1):413–417.

Breslow, N. and Crowley, J. (1974). A large sample study of the life table and product limit estimates under random censorship. *Ann. Statist.*, 2:437–453.

Byar, D. P. (1980). The veterans administration study of chemoprophylaxis for recurrent stage 1 bladder tumors: Comparisons of placebo, pyridoxine, and topical thiotepa. *In Bladder Tumors and Other Topics in Urological Oncology (M. Pavone-Macaluso, P. H. Smith and F. Edsmyr, eds.)*, pages 363–370.

Cantelli, F. P. (1933). Sulla determinazione empirica delle leggi di probabilita. *Giorn. Ist. Ita*, 4:421–424.

Casella, G. (1985). An introduction to empirical Bayes data analysis. *Amer. Statist.*, 39(2):83–87.

Connor, R. J. and Mosimann, J. E. (1969). Concepts of independence for proportions with a generalization of the Dirichlet distribution. *J. Amer. Statist. Assoc.*, 64:194–206.

Cook, R. J. and Lawless, J. F. (2002). Analysis of repeated events. *Statistical Methods in Medical Research*, 11(1):141–166.

Donsker, M. D. (1952). Justification and extension of Doob's heuristic approach to the Komogorov-Smirnov theorems. *Ann. Math. Statistics*, 23:277–281.

Doss, H., Freitag, S., and Proschan, F. (1989). Estimating jointly system and component reliabilities using a mutual censorship approach. *Ann. Statist.*, 17(2):764–782.

Efron, B. (1967). The two sample problem with censored data. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1965–1966, vol. 4*, pages 831–853, Berkeley and Los Angeles. University of California Press.

Efron, B. and Morris, C. (1972). Limiting the risk of Bayes and empirical Bayes estimators. II. The empirical Bayes case. *J. Amer. Statist. Assoc.*, 67:130–139.

Efron, B. and Morris, C. (1973). Stein's estimation rule and its competitors—an empirical Bayes approach. *J. Amer. Statist. Assoc.*, 68:117–130.

Efron, B. and Morris, C. (1976). Multivariate empirical Bayes and estimation of covariance matrices. *Ann. Statist.*, 4(1):22–32.

El-Neweihi, E., Proschan, F., and Sethuraman, J. (1978). Multistate coherent systems. *J. Appl. Probab.*, 15(4):675–688.

Esary, J. D., Marshall, A. W., and Proschan, F. (1970). Some reliability applications of the hazard transform. *SIAM J. Appl. Math.*, 18:849–860.

Esary, J. D., Marshall, A. W., and Proschan, F. (1971). Determining an approximate constant failure rate for a system whose components have constant failure rates. In *Operations research and reliability (Proc. NATO Conf., Turin, 1969)*, pages 195–210. Gordon and Breach, New York. With discussion by H. Reiche, M. Marcovici, Esary, Marshall and Proschan.

Esary, J. D. and Proschan, F. (1963). Coherent structures of non-identical components. *Technometrics*, 5:191–209.

Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.*, 1:209–230.

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.

Ghosh, D. and Lin, D. Y. (2000). Nonparametric analysis of recurrent events and death. *Biometrics*, 56(2):554–562.

Gilks, W. R. and Wild, P. (1992). Adaptive rejection sampling for gibbs sampling. *Applied Statistics*, 41(2):337–348.

Gill, R. D. (1980). Nonparametric estimation based on censored observations of a Markov renewal process. *Z. Wahrsch. Verw. Gebiete*, 53(1):97–116.

Gjessing, H. K., Røysland, K., Peña, E. A., and Aalen, O. O. (2010). Recurrent events and the exploding Cox model. *Lifetime Data Anal.*, 16(4):525–546.

Glivenko, V. (1933). Sulla determinazione empirica delle leggi di probabilita. *Giorn. Ist. Ita*, 4:92–99.

Grego, J., Li, S., Lynch, J., and Sethuraman, J. (2013). Partition based priors and multiple event censoring: An analysis of rosen's fibrous composite experiment. *Technometrics*.

Hollander, M. and Peña, E. A. (1995). Dynamic reliability models with conditional proportional hazards. *Lifetime Data Anal.*, 1(4):377–401.

Hollander, M. and Peña, E. A. (1996a). Dynamic reliability models. In *Lifetime data: models in reliability and survival analysis (Cambridge, MA, 1994)*, pages 131–140. Kluwer Acad. Publ., Dordrecht.

Hollander, M. and Peña, E. A. (1996b). Reliability models and inference for systems operating in different environments. *Naval Res. Logist.*, 43(8):1079–1108.

Hollander, M. and Peña, E. A. (2004). Nonparametric methods in reliability. *Statist. Sci.*, 19(4):644–651.

Hollander, M. and Proschan, F. (1984). Nonparametric concepts and methods in reliability. In *Nonparametric methods*, volume 4 of *Handbook of Statist.*, pages 613–655. North-Holland, Amsterdam.

Husebye, E., Skar, V., Aalen, O., and Osnes, M. (1990). Digital ambulatory manometry of the small intestine in healthy adults: Estimation of variation within and between individuals and statistical management of aborted mmc periods. *Digestive Diseases and Sciences.*, 35.

Jacod, J. (1975). Multivariate point processes: predictable projection, Radon-Nikodým derivatives, representation of martingales. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 31:235–253.

Kalbfleisch, J. D. (1978). Non-parametric Bayesian analysis of survival time data. *J. Roy. Statist. Soc. Ser. B*, 40(2):214–221.

Kalbfleisch, J. D., Lawless, J. F., and Robinson, J. A. (1991). Methods for the analysis and prediction of warranty claims. *Technometrics*, 33(3):p 273–285.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.*, 53:457–481.

Kelly, P. J. and Lim, L. L. Y. (2000). Survival analysis for recurrent event data: an application to childhood infectious diseases. *Statistics in Medicine*, 19(1):13–33.

Kolmogorov, A. (1933). Sulla determinazione empirica delle leggi di probabilita. *Giorn. Ist. Ita*, 4:92–99.

Korwar, R. M. and Hollander, M. (1976). Empirical Bayes estimation of a distribution function. *Ann. Statist.*, 4(3):581–588.

Kvam, P. H. and Peña, E. A. (2005). Estimating load-sharing properties in a dynamic reliability system. *J. Amer. Statist. Assoc.*, 100(469):262–272.

Kvam, P. H. and Samaniego, F. J. (1994). Nonparametric maximum likelihood estimation based on ranked set samples. *J. Amer. Statist. Assoc.*, 89(426):526–537.

Langberg, N., Proschan, F., and Quinzi, A. J. (1981). Estimating dependent life lengths, with applications to the theory of competing risks. *Ann. Statist.*, 9(1):157–167.

Lawless, J. F. and Nadeau, C. (1995a). Some simple robust methods for the analysis of recurrent events. *Technometrics*, 37(2):158–168.

Lawless, J. F. and Nadeau, C. (1995b). Some simple robust methods for the analysis of recurrent events. *Technometrics*, 37(2):158–168.

Lawless, J. F., Wigg, M. B., Tuli, S., Drake, J., and Lamberti-Pasculli, M. (2001). Analysis of repeated failures or durations, with application to shunt failures for patients with paediatric hydrocephalus. *J. Roy. Statist. Soc. Ser. C*, 50(4):449–465.

Lin, D. Y., Wei, L. J., Yang, I., and Ying, Z. (2000). Semiparametric regression for the mean and rate functions of recurrent events. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 62(4):711–730.

Lindqvist, B. H. (2006). On the statistical modeling and analysis of repairable systems. *Statist. Sci.*, 21(4):532–551.

McClean, S. and Devine, C. (1995). A nonparametric maximum likelihood estimator for incomplete renewal data. *Biometrika*, 82(4):791–803.

Meilijson, I. (1981). Estimation of the lifetime distribution of the parts from the autopsy statistics of the machine. *J. Appl. Probab.*, 18(4):829–838.

Miloslavsky, M., Keleş, S., van der Laan, M. J., and Butler, S. (2004a). Recurrent events analysis in the presence of time-dependent covariates and dependent censoring. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 66(1):239–257.

Miloslavsky, M., Keleş, S., van der Laan, M. J., and Butler, S. (2004b). Recurrent events analysis in the presence of time-dependent covariates and dependent censoring. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 66(1):239–257.

Moeschberger, M. L. and David, H. (1971). Life tests under competing cause of failure and the theory of competing risk. *Biometrics*, 27:909–944.

Nelson, W. B. (2003). *Recurrent events data analysis for product repairs, disease recurrences, and other applications*, volume 10 of *ASA-SIAM Series on Statistics and Applied Probability*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

Ouyang, B., Sinha, D., Slate, E. H., and Van Bakel, A. B. (2013). Bayesian analysis of recurrent event with dependent termination: an application to a heart transplant study. *Stat. Med.*, 32(15):2629–2642.

Peña, E. A. and Hollander, M. (2004). Models for recurrent events in reliability and survival analysis. In *Mathematical reliability: an expository perspective*, volume 67 of *Internat. Ser. Oper. Res. Management Sci.*, pages 105–123. Kluwer Acad. Publ., Boston, MA.

Peña, E. A., Slate, E. H., and González, J. R. (2007). Semiparametric inference for a general class of models for recurrent events. *J. Statist. Plann. Inference*, 137(6):1727–1747.

Peña, E. A., Strawderman, R. L., and Hollander, M. (2000). A weak convergence result relevant in recurrent and renewal models. In *Recent advances in reliability theory (Bordeaux, 2000)*, Stat. Ind. Technol., pages 493–514. Birkhäuser Boston, Boston, MA.

Peña, E. A., Strawderman, R. L., and Hollander, M. (2001). Nonparametric estimation with recurrent event data. *J. Amer. Statist. Assoc.*, 96(456):1299–1315.

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.

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.

Proschan, F. (1963). Theoretical explanation of observed decreasing failure rate. *Technometrics*, 5(3):pp. 375–383.

Rahman, A. F., Lynch, J., and Pena, E. A. (2014). Nonparametric Bayes estimation of gap-time distribution with recurrent event data. *Journal of Nonparametric Statistics*.

Rausand, M. and Høyland, A. (2004). *System reliability theory*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, second edition. Models, statistical methods, and applications.

Robbins, H. (1956). An empirical Bayes approach to statistics. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. I*, pages 157–163, Berkeley and Los Angeles. University of California Press.

Sellke, T. (1988). Weak convergence of the Aalen estimator for a censored renewal process. In *Statistical decision theory and related topics, IV, Vol. 2 (West Lafayette, Ind., 1986)*, pages 183–194. Springer, New York.

Sethuraman, J. (1994). A constructive definition of Dirichlet priors. *Statist. Sinica*, 4(2):639–650.

Sethuraman, J. (2008). Bayesian Nonparametrics Foundations and Applications: Notes to Stat 718 (spring 2008), University of South Carolina, Columbia SC 29208.

Sethuraman, J. and Hollander, M. (2009). Nonparametric Bayes estimation in repair models. *J. Statist. Plann. Inference*, 139(5):1722–1733.

Sinha, D. (1993). Semiparametric Bayesian analysis of multiple event time data. *J. Amer. Statist. Assoc.*, 88(423):979–983.

Smirnov, N. (1948). Table for estimating the goodness of fit of empirical distributions. *Ann. Math. Statistics*, 19:279–281.

Soon, G. and Woodroofe, M. (1996). Nonparametric estimation and consistency for renewal processes. *J. Statist. Plann. Inference*, 53(2):171–195.

Stocker, R. S. and Peña, E. A. (2007). A general class of parametric models for recurrent event data. *Technometrics*, 49(2):210–220.

Stokes, S. L. and Sager, T. W. (1988). Characterization of a ranked-set sample with application to estimating distribution functions. *J. Amer. Statist. Assoc.*, 83(402):374–381.

Susarla, V. and Van Ryzin, J. (1976). Nonparametric Bayesian estimation of survival curves from incomplete observations. *J. Amer. Statist. Assoc.*, 71(356):897–902.

Susarla, V. and Van Ryzin, J. (1978). Empirical Bayes estimation of a distribution (survival) function from right-censored observations. *Ann. Statist.*, 6(4):740–754.

Therneau, T. M. and Grambsch, P. M. (2000). *Modeling survival data: extending the Cox model*. Statistics for Biology and Health. Springer-Verlag, New York.

Therneau, T. M. and Hamilton, S. A. (1997). rhDNase as an example of recurrent event analysis. *Statistics in Medicine*, 16(18):2029–2047.

Therneau, T. M. and Lumley, T. (2009). The survival package. http://cran.r-project.org : The comprehensive r archive network.

Vardi, Y. (1982). Nonparametric estimation in the presence of length bias. *Ann. Statist.*, 10(2):616–620.

Wang, M.-C. and Chang, S.-H. (1999). Nonparametric estimation of a recurrent survival function. *J. Amer. Statist. Assoc.*, 94(445):146–153.

Wei, L. J., Lin, D. Y., and Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *J. Amer. Statist. Assoc.*, 84(408):1065–1073.

Appendix A

Chapter 2 Proofs and Copyright Permission

A.1 Proofs

Proof of Lemma 1: For any $C \in \sigma(\mathscr{Z})$,

$$Q_{Z|Y}(Z \in C \mid Y \in B_2) = \frac{\int_C Q_{Y|Z}(Y \in B_2 \mid z) Q_Z(dz)}{\int Q_{Y|Z}(Y \in B_2 \mid z) Q_Z(dz)}. \tag{A.1}$$

$$\therefore \mathbf{E}_{Z \sim Q_Z}\left[h(Z) \mid Y \in B_2\right] = \frac{\int h(z) Q_{Y|Z}(Y \in B_2 \mid z) Q_Z(dz)}{\int Q_{Y|Z}(Y \in B_2 \mid z) Q_Z(dz)} \tag{A.2}$$

$$\begin{aligned}
&Q_{Z|(X,Y)}\left[C \mid (X,Y) \in (B_1 \times B_2)\right] \\
&= \frac{Q_{(Z,X,Y)}(Z \in C, X \in B_1, Y \in B_2)}{Q_{(X,Y)}(X \in B_1, Y \in B_2)} \\
&= \frac{\int_C Q_{(X,Y)|Z}(X \in B_1, Y \in B_2 \mid Z) Q_Z(dz)}{\int Q_{(X,Y)|Z}(X \in B_1, Y \in B_2 \mid Z) Q_Z(dz)} \\
&= \frac{\int_C \int_{B_1} Q_{X|Z}(dx \mid z) Q_{Y|(Z,X)}(Y \in B_2 \mid z, x) Q_Z(dz)}{\int \int_{B_1} Q_{X|Z}(dx \mid z) Q_{Y|(Z,X)}(Y \in B_2 \mid z, x) Q_Z(dz)} \\
&= \frac{\int_C Q_{Y|(Z,X)}(Y \in B_2 \mid z, x) Q_{X|Z}(X \in B_1 \mid z) Q_Z(dz)}{\int Q_{Y|(Z,X)}(Y \in B_2 \mid z, x) Q_{X|Z}(X \in B_1 \mid z) Q_Z(dz)} \\
&= \frac{\int_C Q_{Y|(Z,X)}(Y \in B_2 \mid z, x) \left[\frac{Q_{X|Z}(X \in B_1|z) Q_Z(dz)}{\int Q_{X|Z}(X \in B_1|w) Q_Z(dw)}\right]}{\int Q_{Y|(Z,X)}(Y \in B_2 \mid z, x) \left[\frac{Q_{X|Z}(X \in B_1|z) Q_Z(dz)}{\int Q_{X|Z}(X \in B_1|w) Q_Z(dw)}\right]} \\
&= \frac{\int_C Q_{Y|(Z,X)}(Y \in B_2 \mid z, x) Q_{Z|X}(dz \mid X \in B_1)}{\int Q_{Y|(Z,X)}(Y \in B_2 \mid z, x) Q_{Z|X}(dz \mid X \in B_1)}
\end{aligned}$$

Therefore, $Q_{Z|(X,Y)}\left[C \mid (X,Y) \in (B_1 \times B_2)\right]$

$$= \frac{\int_C Q_{Y|(Z,X)}(Y \in B_2 \mid z, x) Q_{Z|X}(dz \mid X \in B_1)}{\int Q_{Y|(Z,X)}(Y \in B_2 \mid z, x) Q_{Z|X}(dz \mid X \in B_1)} \tag{A.3}$$

$$\therefore \mathbf{E}_{\substack{Z\sim Q_Z(\cdot)\\(X,Y)|Z\sim Q_{(X,Y)|Z}(\cdot,\cdot|z)}}[h(Z)\mid (X,Y)\in(B_1\times B_2)]$$
$$= \frac{\int h(z)Q_{Y|(Z,X)}(Y\in B_2\mid z,x)Q_{Z|X}(dz\mid X\in B_1)}{\int Q_{Y|(Z,X)}(Y\in B_2\mid z,x)Q_{Z|X}(dz\mid X\in B_1)} \quad \text{(A.4)}$$

Comparing (A.3) with (A.1) one can write

$Q_{Z|(X,Y)}\left[C \mid X\in B_1, Y\in B_2\right] = Q^*_{Z|Y}\left(\cdot \mid B_2\right)$ where $Q^*_Z \equiv Z\sim Q_{Z|X}\left(\cdot \mid B_1\right)$.

Comparing (A.2) and (A.4) we conclude that

$$\begin{aligned}&\mathbf{E}_{\substack{Z\sim Q_Z(\cdot)\\(X,Y)|Z\sim Q_{(X,Y)|Z}(\cdot,\cdot|z)}}[h(Z)\mid (X,Y)\in(B_1\times B_2)]\\ &= \mathbf{E}_{\substack{Z\sim Q_{Z|X}(\cdot|B_1)\\Y|(Z,X)\sim Q_{Y|(Z,X)}(\cdot|z,B_1)}}\left[h(Z)\mid Y\in B_2\right],\end{aligned}$$

This completes the proof. □

Proof of Proposition 1: It suffices to prove the Proposition for $n=1$ and $K_1=k_1$, since the Proposition would then follow upon repeated application of the case with $n=1$. Note that $\{\tau_1, K_1=k_1, T_{11}=t_{11}, T_{12}=t_{12},\ldots,T_{1K_1}=t_{1k_1}\} = \{T_{11}=t_{11},\ldots,T_{1k_1}=t_{1k_1}, S_{1k_1}\le\tau_1<S_{1k_1+1}\} = \{T_{11}=t_{11},\ldots,T_{1k_1}=t_{1k_1}, S_{1k_1}\le\tau_1, T_{1k_1+1}\in[\tau_1-S_{1k_1},\infty)\}$. Once we observed that $K_1=k_1$ on $[0,\tau_1]$, then $S_{1k_1}\le\tau_1$ holds. It is also sufficient to prove the Proposition for $k_1=1$. For $t_{11}\le\tau_1$,

$$\begin{aligned}&\mathbf{E}_{P\sim\mathscr{D}(\alpha)}[h(P)\mid(\tau_1,k_1=1,T_{11}=t_{11})]\\ &=\mathbf{E}_{P\sim\mathscr{D}(\alpha)}[h(P)\mid(T_{11}=t_{11},T_{12}>\tau_1-t_{11})]\\ &=\mathbf{E}_{P\sim\mathscr{D}(\alpha+\delta_{t_{11}})}[h(P)\mid T_{12}>\tau_1-t_{11}]\quad\text{[using Corollary 2]}\end{aligned}$$

It follows that, for $t^*_1=\tau_1-S_{1k_1}\ge 0$,

$$\begin{aligned}&\mathbf{E}_{P\sim\mathscr{D}(\alpha)}[h(P)\mid(\tau_1,K_1=k_1,T_{11}=t_{11},\ldots,T_{1k_1}=t_{1k_1})]\\ &\qquad=\mathbf{E}_{P\sim\mathscr{D}(\alpha+\sum_{j=1}^{k_1}\delta_{t_{1j}})}[h(P)\mid T_{1k_1+1}\in[t^*_1,\infty)].\end{aligned}$$

By repeated application of the above result, it then follows that

$$\begin{aligned}&E_{P\sim\mathscr{D}(\alpha)}[h(P)\mid(\tau_i,K_i=k_i,T_{i1}=t_{i1},T_{i2}=t_{i2},\ldots,T_{ik_i}=t_{ik_i},i=1,2,\ldots,n)]\\ &\qquad=E_{P\sim\mathscr{D}(\alpha^*)}[h(P)\mid(T_{ik_i+1}\in[t^*_i,\infty),i=1,2,\ldots,n)]I(s_{ik_i}\le\tau_i).\end{aligned}$$

This completes the proof of the Proposition. $\Box$

Proof of Proposition 2: The proof of this result is motivated by the Lemma 2 of Susarla and Van Ryzin (1976). Consider the partition $[0, T_1^*), [T_1^*, T_2^*), \ldots, [T_m^*, \infty)$ on $\Re_+$. Define $V_i = P[T_{j-1}^*, T_j^*)$ for $j = 1, 2, \ldots, m, m+1$, then $(V_1, V_2, \ldots, V_m) \sim \mathfrak{D}(\beta_1, \ldots, \beta_{m+1})$ where $\beta_j = \alpha^*[T_{j-1}^*, T_j^*)$ for j = 1, 2, ..., m+1 and $\sum_{j=1}^{m+1} V_j = 1$. Clearly $P[T_1^*, \infty) = 1 - V_1$, $P[T_2^*, \infty) = 1 - V_1 - V_2, \ldots$, $P[T_m^*, \infty) = 1 - \sum_{j=1}^{m} V_j$.

$$
\begin{aligned}
& c\ \mathbf{E}\left[\prod_{j=1}^{m} (P[T_j^*, \infty))^{\lambda_j}\right] \\
& = c\ \mathbf{E}\left[\prod_{j=1}^{m} (1 - \sum_{j=1}^{m} V_j)^{\lambda_j}\right] \\
& = \int\int \ldots \int_0^{1-\sum_{j=1}^{m-1} v_j} \prod_{j=1}^{m} (1 - \sum_{j=1}^{m} v_j)^{\lambda_j} \prod_{j=1}^{m} [v_j^{\beta_j - 1}](1 - \sum_{j=1}^{m} v_j)^{\beta_{m+1}-1} dv_1 dv_2 \ldots dv_m
\end{aligned}
$$

where m-tuple integration is carried over the simplex $\{(v_1, v_2, \ldots, v_m)\colon 0 \leq v_j \leq 1, \sum_{j=1}^{m} v_j \leq 1\}$. Integrating first w.r.to v_m then $v_{m-1}, \ldots, v_1$, respectively, and at each stage using the result $\int_0^a t^{\gamma-1}(a-t)^{\eta-1} dw = a^{\gamma+\eta-1} B(\gamma, \eta)$, for $0 \leq a \leq 1$ and $\gamma, \eta \geq 0$, we obtain

$$
\begin{aligned}
& c\ \mathbf{E}[\prod_{j=1}^{m} (P[T_j^*, \infty))^{\lambda_j}] \\
& = \int \ldots \int_0^{1-\sum_{j=1}^{m-1} v_j} v_1^{\beta_m - 1} \ldots v_m^{\beta_m - 1} (1 - \sum_{j=1}^{m} v_j)^{\beta_{m+1}+\lambda_m - 1} dv_m \ldots dv_2 dv_1 \\
& = B(\beta_m, \beta_{m+1} + \lambda_m). \\
& \quad \int \ldots \int_0^{1-\sum_{j=1}^{m-2} v_j} v_1^{\beta_1 - 1} \ldots v_{m-1}^{\beta_m - 1} (1 - \sum_{j=1}^{m-1} v_j)^{\beta_{m+1}+\beta_m+\lambda_m - 1} dv_{m-1} \ldots dv_2 dv_1 \\
& = \prod_{j=1}^{m} B(\beta_j, \sum_{r=j+1}^{m+1} (\beta_r + \lambda_{r-1})).
\end{aligned}
$$

Proof of Proposition 3: The proof of this result is analogous to Lemma 1 of Susarla and Van Ryzin (1976).

$$\begin{aligned}
&\mathscr{P}\{\bar{F}(u) \geq w | T_{iK_i+1} \in [T_i^*, \infty), i = 1, 2, \ldots, n\} \\
&= \frac{\mathscr{P}\{\bar{F}(u) \geq w, T_{iK_i+1} \in [T_i^*, \infty), i = 1, 2, ..., n\}}{P\{T_{iK_i+1} \in [T_i^*, \infty), i = 1, 2, ..., n\}} \\
&= \frac{\mathbf{E}[\mathscr{P}\{\bar{F}(u) \geq w, T_{iK_i+1} \in [T_i^*, \infty) | I_{[\bar{F}(u) \geq w]} P[T_i^*, \infty), i = 1, \ldots, n\}]}{\mathbf{E}[\mathscr{P}\{T_{iK_i+1} \in [T_i^*, \infty) | P[T_i^*, \infty), i = 1, 2, ..., n\}]} \\
&= \frac{\mathbf{E}[I_{[\bar{F}(u) \geq w]} \prod_{i=1}^n P[T_i^*, \infty)]}{\mathbf{E}[\prod_{i=1}^n P[T_i^*, \infty)]} \quad \text{[using definition 2]} \\
&= \frac{\mathbf{E}[I_{[F(u) \leq 1-w]} \prod_{i=1}^n P[T_i^*, \infty)]}{\mathbf{E}[\prod_{i=1}^n P[T_i^*, \infty)]}
\end{aligned}$$

Proof of Proposition 4: The proof of this Proposition is analogous to Corollary 1 of Susarla and Van Ryzin (1976). We assume that $T_{(j)}^*$, $j = 1, 2, \ldots, m+1$, are the distinct ordered censoring event time of $T_i^*, i = 1, 2, \ldots, n$. For any random variable $\bar{F} \in [0, 1]$, we have

$$\begin{aligned}
&\mathbf{E}\{(\bar{F}(u))^\nu | T_{iK_i+1} \in [T_i^*, \infty), i = 1, 2, \ldots, n\} \\
&= \int_0^1 \mathscr{P}\{\bar{F}(u) \geq w | T_{iK_i+1} \in [T_i^*, \infty), i = 1, 2, \ldots, n\} dw^\nu \quad \text{(using Fubini's theorem)} \\
&= \frac{\int_0^1 \mathbf{E}\left[I_{[F(u) \leq 1-w]} \prod_{i=1}^n P[T_i^*, \infty)\right] dw^\nu}{\mathbf{E}\left[\prod_{i=1}^n P[T_i^*, \infty)\right]} \\
&= \frac{\mathbf{E}\left[\int_0^1 I_{[F(u) \leq 1-w]} dw^\nu \prod_{i=1}^n P[T_i^*, \infty)\right]}{\mathbf{E}\left[\prod_{i=1}^n P[T_i^*, \infty)\right]} \\
&= \frac{\mathbf{E}\left[\int_0^{1-F(u)} dw^\nu \prod_{j=1}^n P[T_j^*, \infty)\right]}{\mathbf{E}\left[\prod_{i=1}^n P[T_i^*, \infty)\right]} \\
&= \frac{\mathbf{E}\left[(P[u, \infty))^\nu \prod_{i=1}^n P[T_i^*, \infty)\right]}{\mathbf{E}\left[\prod_{i=1}^n P[T_i^*, \infty)\right]}.
\end{aligned}$$

Therefore,

$$\mathbf{E}\left\{(\bar{F}(u))^\nu | T_{iK_i+1} \in [T_i^*, \infty), i = 1, 2, \ldots, n\right\} = \frac{\mathbf{E}\left[(P[u, \infty))^\nu \prod_{j=1}^m (P[T_{(j)}^*, \infty))^{\lambda_j}\right]}{\mathbf{E}\left[\prod_{j=1}^m (P[T_{(j)}^*, \infty))^{\lambda_j}\right]}, \tag{A.5}$$

where m is the number of distinct censoring events and $\lambda_j, j = 1, 2, \ldots, m$, are the number of events censored at times $T_{(j)}^*$. Using Proposition 2, we have

$$\prod_{j=1}^{m+1} \Gamma(\beta_j) \mathbf{E}\left[\prod_{j=1}^m (P[T_j^*, \infty))^{\lambda_j}\right] = \Gamma(\alpha^*(\Re_+)) \prod_{j=1}^m B(\beta_j, \sum_{r=j+1}^{m+1} (\beta_r + \lambda_{r-1})).$$

Replacing β_j by $\alpha^*[T^*_{(j-1)}, T^*_{(j)})$ for $j = 1, 2, \ldots, m+1$, in the above we get

$$\prod_{j=1}^{m+1} \Gamma(\alpha^*[T^*_{(j-1)}, T^*_{(j)}))\mathbf{E}\left[\prod_{j=1}^{m}(P[T^*_{(j)}, \infty))^{\lambda_j}\right]$$
$$= \Gamma(\alpha^*(\Re_+)) \prod_{j=1}^{m} \left\{ B(\alpha^*[T^*_{(j-1)}, T^*_{(j)}), \alpha^*[T^*_{(j)}, \infty) + \sum_{r=j}^{m} \lambda_r) \right\}. \quad \text{(A.6)}$$

Again applying Proposition 2 with the partition points $0 = T^*_{(0)} < T^*_{(1)}, \ldots, < T^*_{(l)} < u < T^*_{(l+1)}, \ldots, < T^*_{(m)} < T^*_{(m+1)} = \infty$ on $\Re_+$ and associated exponents $\lambda_1, \ldots, \lambda_l, \nu, \lambda_{l+1}, \ldots, \lambda_m$, respectively, we get

$$\prod_{\substack{j=1 \\ j\neq l+1}}^{m+1} \Gamma(\alpha^*[T^*_{(j-1)}, T^*_{(j)}))\Gamma(\alpha^*[T^*_{(l)}, u))\Gamma(\alpha^*[u, T^*_{(l+1)}))\mathbf{E}\left[(P[u,\infty))^{\nu}\prod_{j=1}^{m}(P[T^*_{(j)}, \infty))^{\lambda_j}\right]$$
$$= \Gamma(\alpha^*(\Re_+)) \prod_{j=1}^{l} \left\{ B(\alpha^*[T^*_{(j-1)}, T^*_{(j)}), \alpha^*[T^*_{(j)}, \infty) + \nu + \sum_{r=j}^{m} \lambda_r) \right\}$$
$$\prod_{j=l+2}^{m} \left\{ B(\alpha^*[T^*_{(j-1)}, T^*_{(j)}), \alpha^*[T^*_{(j)}, \infty) + \sum_{r=j}^{m} \lambda_r) \right\}$$
$$\left\{ B(\alpha^*[T^*_{(l)}, u), \alpha^*[u, \infty) + \nu + \sum_{r=l+1}^{m} \lambda_r) \right\}$$
$$\left\{ B(\alpha^*[u, T^*_{(l+1)}), \alpha^*[T^*_{(l+1)}, \infty) + \sum_{r=l+1}^{m} \lambda_r) \right\}. \quad \text{(A.7)}$$

Using the results from (A.6), (A.7) and $B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ and canceling common terms, we obtain from right hand side of (A.5)

$$\frac{\mathbf{E}\left[(P[u,\infty))^{\nu}\prod_{j=1}^{m}(P[T^*_{(j)}, \infty))^{\lambda_j}\right]}{\mathbf{E}\left[\prod_{j=1}^{m}(P[T^*_{(j)}, \infty))^{\lambda_j}\right]}$$
$$= \left\{ \frac{\Gamma(\alpha^*[u,\infty) + \nu + \sum_{r=l+1}^{m}\lambda_r)}{\Gamma(\alpha^*[u,\infty) + \sum_{r=l+1}^{m}\lambda_r)} \frac{\Gamma(\alpha^*[T^*_{(l)},\infty) + \sum_{r=l+1}^{m}\lambda_r)}{\Gamma(\alpha^*[T^*_{(l)},\infty) + \nu + \sum_{r=l+1}^{m}\lambda_r)} \right\}$$
$$\prod_{j=1}^{l} \left\{ \frac{\Gamma(\alpha^*[T^*_{(j)},\infty) + \nu + \sum_{r=j}^{m}\lambda_r)}{\Gamma(\alpha^*[T^*_{(j)},\infty) + \sum_{r=j}^{m}\lambda_r)} . \frac{\Gamma(\alpha^*[T^*_{(j)},\infty) + \sum_{r=j}^{m}\lambda_r)}{\Gamma(\alpha^*[T^*_{(j-1)},\infty) + \nu + \sum_{r=j}^{m}\lambda_r)} \right\}.$$

Using the result $\Gamma(a+\nu) = \Gamma(a)\prod_{s=0}^{\nu-1}(a+s)$ for $a > 0$ in the above equation and canceling terms yields, L.H.S. of (A.5)

$$= \prod_{s=0}^{\nu-1} \left\{ \frac{\alpha^*(u,\infty) + s + \sum_{j=l+1}^m \lambda_j}{\alpha^*[T^*_{(l)},\infty) + s + \sum_{j=l+1}^m \lambda_j} \prod_{j=1}^{l} \left\{ \frac{\alpha^*[T^*_{(j)},\infty) + s + \sum_{r=j}^m \lambda_j}{\alpha^*[T^*_{(j-1)},\infty) + s + \sum_{r=j}^m \lambda_r} \right\} \right\}$$

$$= \prod_{s=0}^{\nu-1} \left\{ \left[\frac{\alpha(u,\infty) + s + Y^+(u)}{\alpha[T^*_{(l)},\infty) + s + Y(T^*_{(l)}) - \lambda_l} \right] \prod_{j=1}^{l} \left\{ \frac{\alpha[T^*_{(j)},\infty) + s + Y(T^*_{(j)})}{\alpha[T^*_{(j-1)},\infty) + s + Y(T^*_{(j-1)})} \right\} \right\},$$

with $\alpha^* = \alpha + \sum_{i=1}^n \sum_{j=1}^{k_i} \delta_{T_{ij}}$ and the definition of Y^+, Y, and N. Rearranging the terms of the denominators in the R.H.S. gives the result that

$$\mathbf{E}\left[(\bar{F}(u))^\nu \mid T_{iK_i+1} \in [T^*_i,\infty), i = 1,2,\ldots,n\right]$$
$$= \prod_{s=0}^{\nu-1} \left\{ \left[\frac{\alpha(u,\infty) + s + Y^+(u)}{\alpha(\Re_+) + s + N} \right] \prod_{j=1}^{l} \left\{ \frac{\alpha[T^*_{(j)},\infty) + s + Y(T^*_{(j)})}{\alpha[T^*_{(j)},\infty) + s + Y(T^*_{(j)}) - \lambda_j} \right\} \right\}.$$

Proof of Theorem 2: The posterior variance of $\bar{F}$ is

$$\sigma^2_{NPB}(u) = \mathbf{E}[(\bar{F}(u))^2 \mid T_{ij} = t_{ij}, T_{iK_i+1} \in [T^*_i,\infty), i = 1,\ldots,n; j = 1,\ldots,K_i] -$$
$$\left[\mathbf{E}[\bar{F}(u) \mid T_{ij} = t_{ij}, T_{iK_i+1} \in [T^*_i,\infty), i = 1,\ldots,n; j = 1,\ldots,K_i)]\right]^2.$$

Plugging $\nu = 2$ and $\nu = 1$ in Proposition 4, we obtain the expressions for the posterior second moment and first moment of $\bar{F}(u)$, respectively. Therefore the posterior variance of $\bar{F}(u)$ is

$$\sigma^2_{NPB}(u)$$
$$= \hat{\bar{F}}_{NPB}(u) \left[\frac{\alpha(u,\infty) + Y^+(u) + 1}{\alpha(\Re_+) + N + 1} \prod_{j=1}^{l} \left\{ \frac{\alpha[T^*_{(j)},\infty) + Y(T^*_{(j)}) + 1}{\alpha[T^*_{(j)},\infty) + Y(T^*_{(j)}) + 1 - \lambda_j} \right\} \right] - \hat{\bar{F}}^2_{NPB}(u)$$
$$= \hat{\bar{F}}_{NPB}(u) \left[\frac{\alpha(u,\infty) + Y^+(u) + 1}{\alpha(\Re_+) + N + 1} \prod_{j=1}^{l} \left\{ \frac{\alpha[T^*_{(j)},\infty) + Y(T^*_{(j)}) + 1}{\alpha[T^*_{(j)},\infty) + Y(T^*_{(j)}) + 1 - \lambda_j} \right\} - \hat{\bar{F}}_{NPB}(u) \right].$$

This completes the proof. □

Proof of Theorem 3: It follows from Proposition 1 that

$$P \mid (T_{ij} = t_{ij}, i = 1, 2, \ldots, n; j = 1, 2, \ldots, k_i) \sim \mathscr{D}(\alpha^*)$$

where $\alpha^* = \alpha + \sum_{i=1}^{n} \sum_{j=1}^{k_i} \delta_{t_{ij}}$. Without loss of generality assume that $T_{1k_1} \in [t^*_{(1)}, \infty)$. Then,

$$\begin{aligned}
\mathscr{P}\{P(\mathscr{B}^*) \in \mathbf{B} \mid T_{1k_1} \in [t^*_{(1)}, \infty)\} &\propto \mathscr{P}\{P(\mathscr{B}^*) \in \mathbf{B}, T_{1k_1} \in [t^*_{(1)}, \infty)\} \\
&= \mathbf{E}\left[\mathscr{P}\left(P(\mathscr{B}^*) \in \mathbf{B}, T_{1k_1} \in [t^*_{(1)}, \infty) \mid P\right)\right] \\
&= \mathbf{E}\left[\mathscr{P}\{T_{1k_1} \in [t^*_{(1)}, \infty) \mid P\} I(P(\mathscr{B}^*) \in \mathbf{B})\right] \\
&= \mathbf{E}\left[P([t^*_{(1)}, \infty)) I(P(\mathscr{B}^*) \in \mathbf{B})\right] \\
&\propto \int_{\mathbf{B}} \left[\prod_{l=1}^{m+1} \left[y_l^{\alpha_l^* - 1}\right] (1 - y_1)\right] d\mathbf{y},
\end{aligned}$$

since $\sum_{l=1}^{m+1} y_l = 1$. Repeating the above procedure for all right-censored observations we obtain

$$\begin{aligned}
&\mathscr{P}\left\{P(\mathscr{B}^*) \in \mathbf{B} \mid T_{ij} = t_{ij}, T_{ik_i+1} \in [t^*_i, \infty), i = 1, 2, \ldots, n\right\} \\
&\qquad\qquad \propto \int_{\mathbf{B}} y_{m+1}^{\alpha^*_{m+1} - 1} \prod_{l=1}^{m} \left[y_l^{\alpha_l^* - 1} \left(1 - \sum_{j=1}^{l} y_j\right)\right] d\mathbf{y}.
\end{aligned}$$

This completes the proof. □

Proof of Theorem 4: When $\alpha(\Re_+) \to 0$, then the nonparametric Bayes estimator

$$\begin{aligned}
\hat{\bar{F}}_{NPB}(u) &= \left\{\frac{\alpha(u, \infty) + Y^+(u)}{\alpha(\Re_+) + N} \prod_{j=1}^{l} \left\{\frac{\alpha(T^*_{(j)}, \infty) + Y(T^*_{(j)})}{\alpha(T^*_{(j)}, \infty) + Y(T^*_{(j)}) - \lambda_j}\right\}\right\} \\
&\to \left\{\frac{Y^+(u)}{N} \prod_{\substack{j=1 \\ \delta_j = 0}}^{l} \left\{\frac{Y(T^*_{(j)})}{Y(T^*_{(j)}) - \lambda_j}\right\}\right\}. \quad \text{(A.8)}
\end{aligned}$$

Let $i(u)$ be the largest integer such that $T'_{i(u)} \leq u$. Then

$$\frac{Y^+(u)}{N} = \prod_{j \leq i(u)} \left\{\frac{Y^+(T'_j)}{Y(T'_j)}\right\} = \left\{\prod_{\substack{j \leq i(u) \\ \delta_j = 1}} \left\{\frac{Y^+(T'_j)}{Y(T'_j)}\right\}\right\} \left\{\prod_{\substack{j \leq i(u) \\ \delta_j = 0}} \left\{\frac{Y^+(T'_j)}{Y(T'_j)}\right\}\right\}.$$

Now, using the above results in the right-hand side of (A.8) and replacing $T^*_{(j)}$ by $T'_{(j)}$ for $\delta_j = 0$, we get

$$\left\{ \frac{Y^+(u)}{N} \prod_{\substack{j=1 \\ \delta_j=0}}^{l} \left\{ \frac{Y(T^*_{(j)})}{Y(T^*_{(j)}) - \lambda_j} \right\} \right\}$$

$$= \prod_{\substack{j\leq i(u) \\ \delta_j=1}} \left\{ \frac{Y^+(T'_j)}{Y(T'_j)} \right\} \prod_{\substack{j\leq i(u) \\ \delta_j=0}} \left\{ \frac{Y^+(T'_j)}{Y(T'_j)} \right\} \prod_{\substack{j=1 \\ \delta_j=0}}^{l} \left\{ \frac{Y(T'_{(j)})}{Y(T'_{(j)}) - \lambda_j} \right\} = \prod_{\substack{j\leq i(u) \\ \delta_j=1}} \left\{ \frac{Y^+(T'_j)}{Y(T'_j)} \right\}.$$

Therefore, as $\alpha(\Re_+) \to 0$,

$$\hat{\bar{\bar{F}}}_{NPB}(u) \to \left\{ \frac{Y^+(u)}{N} \prod_{\substack{j=1 \\ \delta_j=0}}^{l} \left\{ \frac{Y(T'_j)}{Y(T'_j) - \lambda_j} \right\} \right\} = \prod_{\substack{j\leq i(u) \\ \delta_j=1}} \left\{ \frac{Y^+(T'_j)}{Y(T'_j)} \right\} = \hat{\bar{\bar{F}}}_{PLE}(u).$$

This completes the proof of the Theorem. □

A.2 Copyright Permission to Reprint Chapter 2

Our Ref: DE/GNST/P836

23rd April 2014

Dear A.K.M. Fazlur Rahman,

Thank you for your correspondence requesting permission to reproduce the following article published in our journal in your printed thesis and to be posted in your university's repository.

Nonparametric Bayes estimation of gap-time distribution with recurrent event data-Journal of Nonparametric Statistics -Author's - A. K. M. Fazlur Rahman, James D. Lynch & Edsel A. Pena.

We will be pleased to grant permission on the sole condition that you acknowledge the original source of publication and insert a reference to the article on the Journals website:

This is the authors accepted manuscript of an article published as the version of record in Journal of Nonparametric Statistics date of publication 17th April 2014 - http://www.tandfonline.com/ http://www.tandfonline.com/doi/full/10.1080/10485252.2014.906744

Please note that this license does not allow you to post our content on any third party websites or repositories. Thank you for your interest in our Journal.

Yours sincerely

Deborah East

Permissions Administrator

Taylor Francis Group

Taylor Francis Group is a trading name of Informa UK Limited, registered in England under no. 1072954

Appendix B

Chapter 4 Proofs

B.1 Proofs

Proof of Lemma 1:

$$E(Y_j) = \frac{\int y_j h^*(\mathbf{y})d\mathbf{y}}{\int h^*(\mathbf{y})d\mathbf{y}} = \frac{C_j(\alpha_l^* + I(l=j), \lambda_l, l = 1, 2, \ldots, m, m+1)}{C_j(\alpha_l^*, \lambda_l, l = 1, 2, \ldots, m, m+1)},$$

where $C_j(\alpha_l^* + I(l=j), \lambda_l, l = 1, 2, \ldots, m, m+1)$

$$= \int y_j \prod_{l=1}^{m+1} \left[y_l^{\alpha_l^*-1}\right] \prod_{l=1}^{m} \left[\sum_{r=l+1}^{m+1} y_r\right]^{\lambda_l} dy_1 dy_2 \ldots dy_m$$

$$= \int y_j \prod_{l=1}^{m} \left[y_l^{\alpha_l^*-1}\right] \left(1 - \sum_{r=1}^{m} y_r\right)^{\alpha_{m+1}^*-1} \prod_{l=1}^{m} \left[1 - \sum_{r=1}^{l} y_r\right]^{\lambda_l} dy_1 dy_2 \ldots dy_m$$

$$= \int \ldots \int y_1^{\alpha_1^*-1} y_2^{\alpha_2^*-1} \ldots y_j^{\alpha_j^*+1-1} \ldots y_{m-1}^{\alpha_{m-1}^*-1} \prod_{l=1}^{m-1} \left[1 - \sum_{r=1}^{l} y_r\right]^{\lambda_l} dy_1 dy_2 \ldots dy_{m-1}$$

$$\int_0^{1-\sum_{l=1}^{m-1} y_l} y_m^{\alpha_m^*-1} \left[1 - \sum_{r=1}^{m} y_r\right]^{\alpha_{m+1}^*+\lambda_m-1} dy_m$$

$$\left[\text{integrate first w.r.t. } y_m \text{ over the simplex } \{(y_1, y_2, \ldots, y_m)\colon 0 \le y_j \le 1, \sum_{j=1}^{m} y_j \le 1\}\right]$$

$$= \mathrm{B}(\alpha_m^*, \alpha_{m+1}^* + \lambda_m) \int \ldots \int \prod_{l=1}^{m-2} y_l^{\alpha_l^*-1} \prod_{l=1}^{m-2} \left[1 - \sum_{r=1}^{l} y_r\right]^{\lambda_l} dy_1 dy_2 \ldots dy_{m-2}$$

$$\int_0^{1-\sum_{l=1}^{m-1} y_l} y_m^{\alpha_{m-1}^*-1} \left[1 - \sum_{r=1}^{m-1} y_r\right]^{\alpha_{m+1}^*+\lambda_m-1} dy_{m-1}$$

Then integrating w.r.to $y_{m-1}, \ldots, y_1$, respectively, and at each stage using the result $\int_0^a t^{\gamma-1}(a-t)^{\eta-1}dt = a^{\gamma+\eta-1}B(\gamma, \eta)$, for $0 \le a \le 1$ and $\gamma, \eta \ge 0$, and $B(\gamma, \eta) = \frac{\Gamma(\gamma)\Gamma(\eta)}{\Gamma(\gamma+\eta)}$, we obtain

$$C_j(\alpha_l^* + I(l=j), \lambda_l, l = 1, 2, \ldots, m, m+1) = \prod_{l=1}^{m} \left[\mathrm{B}\left(\alpha_l^* + I(l=j), \sum_{r=l+1}^{m+1} [\alpha_r^* + \lambda_{r-1}]\right)\right]$$

$$
\begin{aligned}
E(Y_j) &= \frac{C_j\left(\alpha_l^* + I(l=j), \lambda_l, l = 1, 2, \ldots, m, m+1\right)}{C_j\left(\alpha_l^*, \lambda_l, l = 1, 2, \ldots, m, m+1\right)} \\
&= \frac{\prod_{l=1}^{m}\left[\mathrm{B}\left(\alpha_l^* + I(l=j), \sum_{r=l+1}^{m+1}[\alpha_r^* + \lambda_{r-1}]\right)\right]}{\prod_{l=1}^{m}\left[\mathrm{B}\left(\alpha_l^*, \sum_{r=l+1}^{m+1}[\alpha_r^* + \lambda_{r-1}]\right)\right]} \\
&= \left[\frac{\alpha(B_j) + N(B_j)}{\alpha(\Re_+) + n}\right]\left[\prod_{l=1}^{j-1}\left\{\frac{\alpha(V_{(l)}^*, \infty) + Y^+(V_{(l)}^*) + \lambda_l}{\alpha(V_{(l)}^*, \infty) + Y^+(V_{(l)}^*)}\right\}\right].
\end{aligned}
$$

Proof of Theorem 5: We can recover the PL-type estimator from our nonparametric Bayes (NPB) estimator $\hat{\hat{\bar{F}}}_{NPB}$. To this end, let $V_i^{**}, i = 1, 2, \ldots, n$, denote the ordered (increasing magnitude) observed values of $V_i, i = 1, 2, \ldots, n$, so that $0 \leq V_1^{**} \leq V_2^{**} \leq, \ldots, \leq V_n^{**}$. Let $N^\dagger(w) = \sum_{r=1}^{n} I(V_r^{**} \leq w, \epsilon_r = 1)$, where $\epsilon_r = 1$ if V_r^{**} is an uncensored (complete) observation, and 0 otherwise. Then the PL-type estimator

$$
\widehat{R}_{PLE}(t) = \prod_{w \leq t}\left\{1 - \frac{\Delta N^\dagger(w)}{Y(w)}\right\} = \prod_{w \leq t}\left\{\frac{Y^+(w)}{Y(w)}\right\}, \tag{B.1}
$$

where $Y^+(w)$ and $Y(w)$ are as defined (4.4)

When

$$
\alpha \to 0, \hat{\hat{\bar{F}}}_{NPB} \to \left\{\frac{N((0,t) \cap B_l)}{n} \prod_{\substack{j=1 \\ \epsilon_j = 0}}^{l-1}\left\{\frac{Y(V_{(j)}^{**}) + \lambda_j}{Y^+(V_{(j)}^{**})}\right\}\right\}.
$$

Therefore when $\alpha \to 0$ jump size at time t is

$$
F(t) - F(t-) = \left[\frac{\Delta N^\dagger(t)}{n}\right]\left[\prod_{\substack{j=1 \\ \epsilon_j = 0}}^{l-1}\left\{\frac{Y(V_{(j)}^{**}) + \lambda_j}{Y^+(V_{(j)}^{**})}\right\}\right]. \tag{B.2}
$$

Let $i(t)$ be the largest integer such that $V_{i(t)}^{**} \leq t$. Then

$$
\begin{aligned}
\frac{\Delta N^\dagger(t)}{n} &= \prod_{j \leq i(t)}\left\{\frac{Y^+(V_{(j)}^{**})}{Y(V_{(j)}^{**})}\right\}\left[\frac{\Delta N^\dagger(t)}{Y^+(t)}\right] \\
&= \left\{\prod_{\substack{j \leq i(t) \\ \epsilon_j = 1}}\left\{\frac{Y^+(V_{(j)}^{**})}{Y(V_{(j)}^{**})}\right\}\right\}\left\{\prod_{\substack{j \leq i(t) \\ \epsilon_j = 0}}\left\{\frac{Y^+(V_{(j)}^{**})}{Y(V_{(j)}^{**})}\right\}\right\}\left[\frac{\Delta N^\dagger(t)}{Y^+(t)}\right].
\end{aligned}
$$

Now (B.2) implies that

$$\left\{\prod_{\substack{j\le i(t)\\ \epsilon_j=1}}\left\{\frac{Y^+(V_j^{**})}{Y(V_j^{**})}\right\}\right\}\left\{\prod_{\substack{j\le i(t)\\ \epsilon_j=0}}\left\{\frac{Y^+(V_{(j)}^{**})}{Y(V_{(j)}^{**})}\right\}\right\}\left[\frac{\Delta N^\dagger(t)}{Y^+(t)}\right]\prod_{\substack{j\le i(t)\\ \epsilon_j=0}}\left\{\frac{Y(V_{(j)}^{**})}{Y^+(V_{(j)}^{**})}\right\}$$

$$=\left\{\prod_{\substack{j\le i(t)\\ \epsilon_j=1}}\left\{\frac{Y^+(V_{(j)}^{**})}{Y(V_{(j)}^{**})}\right\}\right\}\left[\frac{\Delta N^\dagger(t)}{Y^+(t)}\right]=\frac{\Delta N^\dagger(t)}{Y(t)}$$

This implies that the jump is only at the complete observations and the jump size is $\frac{\Delta N(t)}{Y(t)}$ which is exactly same as the PL-type estimator. Therefore as $\alpha(\Re_+)\to 0$, PL-type estimator is a limiting case of the proposed nonparametric Bayes estimator. This completes the proof of the Theorem. $\square$