UNIVERSIDADE FEDERAL DE SÃO CARLOS

CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA

PROGRAMA INTERINSTITUCIONAL DE PÓS-GRADUAÇÃO EM ESTATÍSTICA UFSCar-USP

MARCO POLLO ALMEIDA

STATISTICAL INFERENCE FOR NON-HOMOGENEOUS POISSON PROCESS WITH COMPETING RISKS: A REPAIRABLE SYSTEMS APPROACH UNDER POWER-LAW PROCESS

Doctoral dissertation submitted to the Department of Statistics - Des/UFSCar and to the Institute of Mathematics and Computer Sciences – ICMC-USP, in partial fulfillment of the requirements for the PhD degree Statistics – Interagency Program Graduate in Statistics UFSCar-USP.

Advisor: Prof. Dr. Vera Lúcia Damasceno Tomazella

Co-advisor: Prof. Dr. Gustavo Leonel Gilardoni Avalle

São Carlos

September 2019

UNIVERSIDADE FEDERAL DE SÃO CARLOS

CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA

PROGRAMA INTERINSTITUCIONAL DE PÓS-GRADUAÇÃO EM ESTATÍSTICA UFSCar-USP

MARCO POLLO ALMEIDA

INFERÊNCIA ESTATÍSTICA PARA PROCESSO DE POISSON NÃO-HOMOGÊNEO COM RISCOS COMPETITIVOS: UMA ABORDAGEM DE SISTEMAS REPARÁVEIS SOB PROCESSO DE LEI DE POTÊNCIA

Tese apresentada ao Departamento de Estatística - Des/UFSCar e ao Instituto de Ciências Matemáticas e de Computação - ICMC-USP, como parte dos requisitos para obtenção do título de Doutor em Estatística – Programa Interinstitucional de Pós-graduação em Estatística UFSCar-USP.

Orientadora: Prof. Dr. Vera Lúcia Damasceno Tomazella

Coorientador: Prof. Dr. Gustavo Leonel Gilardoni Avalle

São Carlos

Setembro de 2019

UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa Interinstitucional de Pós-Graduação em Estatística

Folha de Aprovação

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Tese de Doutorado do candidato Marco Pollo Almeida, realizada em 30/08/2019:

Profa. Dra. Vera Lucia Damasceno Tomazella
UFSCar

Prof. Dr. Ricardo Sandes Ehlers
USP

Prof. Dr. Paulo Henrique Ferreira da Silva
UFBA

Prof. Dr. Jeremias da Silva Leão
UFAM

Prof. Dr. Gilberto Francisco Martha de Souza
POLI

Certifico que a defesa realizou-se com a participação à distância do(s) membro(s) Jeremias da Silva Leão e, depois das arguições e deliberações realizadas, o(s) participante(s) à distância está(ao) de acordo com o conteúdo do parecer da banca examinadora redigido neste relatório de defesa.

Profa. Dra. Vera Lucia Damasceno Tomazella

To my beloved son, João Pedro and my wife, Carla.

ACKNOWLEDGEMENTS

I wish to acknowledge Empresa de Assistência Técnica e Extensão Rural do Estado Pará (EMATER-PARÁ) for giving me the opportunity to complete the doctorate degree and providing the financial support.

I am very grateful to my advisor, Dr. Vera L. D. Tomazella, for his guidance, knowledge, support, friendship and encouragement prior to and throughout the preparation of this thesis.

To my co-advisor, Dr. Gustavo L. Gilardoni Avalle, for the ideas, contributions and discussions.

To all the teachers and staff of the USP-ICMC Statistics Department and the UFSCar Statistics Department.

To my USP and UFSCar colleagues, especially Rafael Paixão, Marcelo Hartmann and Carlos Hirth, for the stimulating discussion with me and friendship.

I would like to express my gratitude to professor Francisco Louzada Neto, Pedro Ramos and Amanda D'Andrea for all the contribution in this research work.

Special thanks are also extended to my committee members, Dr. Ricardo S. Ehlers, Dr. Gilberto F. M. de Souza, Dr. Jeremias da S. Leão, and Dr. Paulo H. F. da Silva, for their valuable comments and suggestions.

I am thankful to my family, especially Maurin, Clara, Evandro, Barriga, Claúdio, Rosangela, Ester, Léa, Mauren, Elmer and Marceonila (in memoriam), for their constant support, conversations and guidance.

Finally, my deepest gratitude goes to my wife, my son for their love, companionship, patience and for always believe in me.

ABSTRACT

ALMEIDA, M. P. **Statistical inference for non-homogeneous Poisson process with competing risks: a repairable systems approach under power-law process**. 2019. 111 p. Doctoral dissertation (Doctorate Candidate joint Graduate Program in Statistics DEs-UFSCar/ICMC-USP) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2019.

In this thesis, the main objective is to study certain aspects of modeling failure time data of repairable systems under a competing risks framework. We consider two different models and propose more efficient Bayesian methods for estimating the parameters. In the first model, we discuss inferential procedures based on an objective Bayesian approach for analyzing failures from a single repairable system under independent competing risks. We examined the scenario where a minimal repair is performed at each failure, thereby resulting in that each failure mode appropriately follows a power-law intensity. Besides, it is proposed that the power-law intensity is reparametrized in terms of orthogonal parameters. Then, we derived two objective priors known as the Jeffreys prior and reference prior. Moreover, posterior distributions based on these priors will be obtained in order to find properties which may be optimal in the sense that, for some cases, we prove that these posterior distributions are proper and are also matching priors. In addition, in some cases, unbiased Bayesian estimators of simple closed-form expressions are derived. In the second model, we analyze data from multiple repairable systems under the presence of dependent competing risks. In order to model this dependence structure, we adopted the well-known shared frailty model. This model provides a suitable theoretical basis for generating dependence between the components' failure times in the dependent competing risks model. It is known that the dependence effect in this scenario influences the estimates of the model parameters. Hence, under the assumption that the cause-specific intensities follow a PLP, we propose a frailty-induced dependence approach to incorporate the dependence among the cause-specific recurrent processes. Moreover, the misspecification of the frailty distribution may lead to errors when estimating the parameters of interest. Because of this, we considered a Bayesian nonparametric approach to model the frailty density in order to offer more flexibility and to provide consistent estimates for the PLP model, as well as insights about heterogeneity among the systems. Both simulation studies and real case studies are provided to illustrate the proposed approaches and demonstrate their validity.

Keywords: competing risks, power-law process, non-homogeneous Poisson process, Bayesian inference, repairable system.

RESUMO

ALMEIDA, M. P. **Inferência estatística para processo de Poisson não-homogêneo com riscos competitivos: uma abordagem de sistemas reparáveis sob processo de lei de potência**. 2019. 111 p. Doctoral dissertation (Doctorate Candidate joint Graduate Program in Statistics DEs-UFSCar/ICMC-USP) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2019.

Nesta tese, o objetivo principal é estudar certos aspectos da modelagem de dados de tempo de falha de sistemas reparáveis sob uma estrutura de riscos competitivos. Consideramos dois modelos diferentes e propomos métodos Bayesianos mais eficientes para estimar os parâmetros. No primeiro modelo, discutimos procedimentos inferenciais baseados em uma abordagem Bayesiana objetiva para analisar falhas de um único sistema reparável sob riscos competitivos independentes. Examinamos o cenário em que um reparo mínimo é realizado em cada falha, resultando em que cada modo de falha segue adequadamente uma intensidade de lei de potência. Além disso, propõe-se que a intensidade da lei de potência seja reparametrizada em termos de parâmetros ortogonais. Então, derivamos duas prioris objetivas conhecidas como priori de *Jeffreys* e priori de referência. Além disso, distribuições posteriores baseadas nessas prioris serão obtidas a fim de encontrar propriedades que podem ser ótimas no sentido de que, em alguns casos, provamos que essas distribuições posteriores são próprias e que também são *matching priors*. Além disso, em alguns casos, estimadores Bayesianos não-viesados de forma fechada são derivados. No segundo modelo, analisamos dados de múltiplos sistemas reparáveis sob a presença de riscos competitivos dependentes. Para modelar essa estrutura de dependência, adotamos o conhecido modelo de fragilidade compartilhada. Esse modelo fornece uma base teórica adequada para gerar dependência entre os tempos de falha dos componentes no modelo de riscos competitivos dependentes. Sabe-se que o efeito de dependência neste cenário influencia as estimativas dos parâmetros do modelo. Assim, sob o pressuposto de que as intensidades específicas de causa seguem um PLP, propomos uma abordagem de dependência induzida pela fragilidade para incorporar a dependência entre os processos recorrentes específicos da causa. Além disso, a especificação incorreta da distribuição de fragilidade pode levar a erros na estimativa dos parâmetros de interesse. Por isso, consideramos uma abordagem Bayesiana não paramétrica para modelar a densidade da fragilidade, a fim de oferecer mais flexibilidade e fornecer estimativas consistentes para o modelo PLP, bem como *insights* sobre a heterogeneidade entre os sistemas. São fornecidos estudos de simulação e estudos de casos reais para ilustrar as abordagens propostas e demonstrar sua validade.

Palavras-chave: riscos competitivos, processo de lei de potência, processo de Poisson não-homogêneo, inferência Bayesiana, sistemas reparáveis.

LIST OF FIGURES

LIST OF ALGORITHMS

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

ABAO	as bad as old
AGAN	as good as new
CI	credibility (credible) interval or confidence interval
CMLE	conditionally unbiased estimator
CP	coverage probability
diag	diagonal matrix
DP	Dirichlet process
DPM	Dirichlet process mixture
HMC	Hamiltonian Monte Carlo
HPP	homogeneous Poisson process
iid	independent and identically distributed
MAE	mean absolute error
MAP	maximum a posteriori
MCMC	Markov chain Monte Carlo
ML	maximum likelihood
MLEs	maximum likelihood estimators
MSE	mean square error
NHPP	non-homogenous Poisson process
PDF	probability density function
PLP	power-law process
ROCOF	rate of occurrences of failures
SD	standard deviation

CONTENTS

CHAPTER

1

INTRODUCTION

1.1 Introduction and bibliographical review

Studying recurrent event data is important in many areas such as engineering, social and political sciences and in the public health setting. In all these fields of study, the event of interest occurs on a recurring basis. For example, failure of a mechanical or electrical component may occur more than once; the recurrence of bugs over time in a software system that is under development; successive tumors in cancer studies; myocardial infarction and epileptic seizure in patients, to name but a few.

In particular, in reliability analysis, interest is usually centered on failure data from complex repairable systems (ASCHER; FEINGOLD, 1984). Monitoring the status of a repairable system leads to a recurrent events framework, where events correspond to failures of a system. A system is defined as repairable when it receives any corrective measure (other than replacing the whole system) in order to restore its components when they have failed and can be returned to the satisfactory operation state where it is able to perform all its functions. On the other hand, a nonrepairable system is a system that is discarded when the first failure occurs (RIGDON; BASU, 2000). However, we will just focus on repairable system case.

The primary challenge when modeling repairable systems data is how to account for the effect of a repair action performed immediately after a failure. In general, one assumes that repair actions are instantaneous and repair time is negligible. The most explored assumptions are either minimal repair and perfect repair. In the former, it is supposed that the repair action, after a failure, returns the system to the exact condition it was immediately before it failed. In the latter, the repair action leaves the system as if it were new. In the engineering literature, these types of repair or corrective maintenance are usually called: *as bad as old* (ABAO) and *as good as new* (AGAN) (BARLOW; HUNTER, 1960; AVEN, 1983; AVEN; JENSEN, 2000; FINKELSTEIN, 2004; MAZZUCHI; SOYER, 1996). More sophisticated models which account for repair action that leave the system somewhere between the ABAO and AGAN conditions are possible, although they will not be considered here; see for instance, Doyen and Gaudoin (2006).

Statistical modeling of the occurrence of failures is done using point processes, particularly, as we will see later, counting processes. In this framework, the model is completely characterized by its failure intensity function. The failure history of a repairable system, under a minimal repair strategy, is usually modeled according to a non-homogenous Poisson process (NHPP). In the repairable system literature, one of the most important and well-known parametric forms for the NHPP model is the power-law process (PLP). The PLP process is convenient because it is easy to implement, it is flexible and the parameters have good interpretations. Regarding classical inference for the PLP, see, for instance, Ascher and Feingold (1984) or Rigdon and Basu (2000). Bayesian inference has been considered among others by Bar-Lev, Lavi and Reiser (1992), Guida, Calabria and Pulcini (1989), Pievatolo and Ruggeri (2004) and Ruggeri (2006).

Additionally, in this thesis, we emphasize an alternative specification of the PLP, which is obtained by using a simple operational definition of its parameters, making them orthogonal to each other. This formulation is considered by Oliveira, Colosimo and Gilardoni (2012) motivated by ideas from Guida, Calabria and Pulcini (1989) and Sen (2002). The former authors show that this reparametrization leads to some advantageous results such as orthogonality among parameters, the likelihood function becomes proportional to a product of gamma densities and the expected Fisher information matrix is diagonal. The model we discuss here is based on such reparametrization because it results in mathematical and computational simplifications for our research.

In reliability theory, the most common system configurations are series systems, parallel systems, and series-parallel systems. In a series system, components are connected in series, in such a way that the failure of a single component results in system failure. The same setting may be expressed in an alternative way by a repairable system in which components can perform different operations, and thus be subject to different types of failures. Traditionally, models with this characteristic are known as competing risks. In complex systems, such as supercomputers, aircraft generators, industrial plants, jet engines, and cars, the presence of multiple types (or causes) of failure is common. From an economic perspective, such systems are commonly repaired rather than replacing the system with a new one after failure. Thus, this model can also be called a repairable competing risks system. As we pointed out already, commonly used methodologies for analyzing multi-type recurrent event data are based on multivariate counting processes and cause-specific intensity functions (ANDERSEN *et al.*, 2012; COOK; LAWLESS, 2007).

It is worth noticing that the existing literature on competing risks in reliability is extensive and focuses particularly on analysis for nonrepairable systems, e.g., Crowder (2001), Lawless (2011), Crowder *et al.* (1994) to cite a few. On the other hand, a number of authors have considered modeling competing risks in a repairable systems framework. For example, some authors have mainly been interested in questions concerning maintenance analysis (LANGSETH; LINDQVIST, 2006; LINDQVIST, 2006; DOYEN; GAUDOIN, 2006). Others have highlighted the relevance of failure analysis of the components of the system based on cause-specific intensity function, such as Liu and Tang (2010), Fu, Tang and Guan (2014), Somboonsavatdee and Sen (2015a), Somboonsavatdee and Sen (2015b).

After considering the scenarios mentioned above, we can highlight two remarkable aspects which pose as major statistical challenges of a model for competing risks systems with minimal repair policy:

Firstly, by considering the setting where a single complex repairable system is under the action of independent competing risks whose failures behavior is based on intensity functions using PLP, one knows that most of the literature considers this problem from a classical approach viewpoint, i.e. the statistical inference for the PLP is generally based on the maximum likelihood theory (SOMBOONSAVATDEE; SEN, 2015b), which may require a large sample size in order to reach good estimators for each cause of failure. However, in practice, Corset, Doyen and Gaudoin (2012) argue that complex systems are required to be highly reliable and for this reason only a few failures occur. Therefore, using asymptotic results on MLE should be avoided. In this case, one may propose a Bayesian analysis of this model as an advantageous procedure to usual classical methods. Additionally, when there is no expert opinion or there is little knowledge about the parameters (e.g., higher-dimensional parameter spaces or complex dependence structure between parameters), it is not feasible to elicit prior distributions (CONSONNI *et al.*, 2018).

To circumvent these problems, we propose an objective Bayesian analysis (BERNARDO, 1979; BERNARDO, 2005). The objective Bayesian method has several attractive features. In addition to solving the problem of few observations, the most worthwhile feature is the use of objective (or noninformative) priors. In this context, the data play a predominant role in obtaining of the posterior distribution. In other words, the inference is based only on information from the data and with the minimum (or absence) of subjective prior information; see Migon, Gamerman and Louzada (2014). There are few articles studying objective priors for competing risks, mainly in the engineering field. Related work (FU; TANG; GUAN, 2014) has provided a discussion on reference priors in order to estimate the cause-specific intensity functions taking into account the HPP. On the other hand, there is a large number of published studies that describe appropriate procedures for the formulation of objective priors; see, e.g., Bernardo and Smith (2009), Kass and Wasserman (1996), and references therein.

Secondly, in the field of reliability engineering, much of the current literature on competing risks pays particular attention to the hypothesis that the components' (causes) failures are independent from each other (LIU; TANG, 2010; HONG; MEEKER, 2010; MEEKER; ESCOBAR, 2014; SOMBOONSAVATDEE; SEN, 2015b). However, this assumption is restrictive in some real situations because there are many ways of dependence between components.

We can call this case dependent competing risks. Moreover, it is important to point out that neglecting existing dependence can lead to estimation errors and bad predictions of system behavior (ZHANG; YANG, 2015). A seminal study in this area is the work of Moeschberger (1974). Wu, Shi and Zhang (2017) and Zhang and Wilson (2017) give an extensive discussion on copula theory in order to model the dependence between components (competing failure modes) in particular settings. Zhang and Yang (2015) discuss an optimal maintenance planning for dependent competing risks systems. Liu (2012) mentions a particular situation where the components within a system are physically, logically, or functionally connected, as an example of dependent failure causes. It means that the condition of a component influences or induces the failure of other components and vice-versa. This author works with the dependence framework based on a gamma frailty model. An interesting perspective has been explored by Lindley and Singpurwalla (1986), who argue that dependence can be induced by the environment the system is subjected to, i.e., the situation where the components of the system (or cluster) share the same environmental stress. Along the same lines, Somboonsavatdee and Sen (2015a) assert that in a repairable systems context such clustering arises naturally across the recurrent failures of a system. The approach proposed by Somboonsavatdee and Sen (2015a) for modeling dependence is also based on frailty.

These examples demonstrate the importance of the theme and, therefore, the need to develop new analysis methodologies. However, very few articles address the dependence, particularly in the setting of recurrent competing risks in repairable systems with PLP. Based on these reasons, we propose a shared frailty (WIENKE, 2003; HOUGAARD, 2012) model using a (multivariate) counting process framework whose intensity function is that of reparametrized PLP. Specifically, the intensity is multiplied by a frailty (or random effect) term, which follows a suitable distribution for a positive random variable. This model provides a suitable theoretical basis for generating dependence between the components' failure times. In other words, the components belonging to a cluster (or system) share a common factor (frailty term), which generates such dependence. The assignment of a probability distribution to frailty plays an important role in the analysis of models with random effects. However, in order to avoid making incorrect model specifications when there is uncertainty about some inherent characteristics of a distribution (e.g., multimodality, skewness, and heavy tails) (WALKER; MALLICK, 1997), we propose a nonparametric approach to model the frailty density (density estimation) (FERGUSON, 1973; FERGUSON *et al.*, 1974; MÜLLER; QUINTANA, 2004). Our approach to these problems is fully Bayesian and based on both MCMC methods (for frailty) and closed-form Bayesian estimators (for PLP parameters) for estimation.

The main contributions of the proposed research include: (i) in the objective Bayesian context, we propose noninformative priors whose resulting posterior distributions are proper. Furthermore, the Bayes estimators have simple closed-form expressions and returned marginal posterior intervals with accurate coverage in the frequentist sense. Our findings outperform those of classical inference, especially in real situations where systems have high reliability and failures are rare; (ii) Our modeling of the dependence effect on multi-component systems, based on multivariate recurrent processes and the shared frailty model, is advantageous because we can perform an individual posterior analysis of the quantities of interest, i.e., we estimate the interest parameters of the PLP (our main focus) separately from nuisance parameter of frailty distribution (variance). Regarding PLP parameters, we consider noninformative priors so that the posterior distributions are proper. With respect to frailty, our proposal avoids making incorrect specifications of the frailty distribution when there is uncertainty about some inherent characteristics of distribution. In this case, we use nonparametric Bayesian inference. Besides, a particular novelty is our hybrid MCMC algorithm for computing the posterior estimates with respect to the frailty distribution.

1.2 Objectives

In this thesis, the main objective is to study certain aspects of modeling failure time data of repairable systems under a competing risks framework. We consider two different models and propose more efficient Bayesian methods for estimating the model parameters. Thus, we can list some specific objectives:

- to consider an orthogonal parametrization for the PLP model parameters (*which is common for all chapters*) such that the likelihood function becomes proportional to a product of gamma densities and the expected Fisher information matrix is diagonal;

- in the independent competing risks setting, to derive overall objective priors (Jeffreys prior and reference prior) for the PLP model parameters that allow one to obtain proper posterior distributions with advantageous properties. In particular, to obtain unbiased estimators that have simple closed-form expressions, as well as marginal posterior intervals with accurate coverage in the frequentist sense (Chapter 4);

- in the dependent competing risks setting, to propose a frailty-induced dependence approach to incorporate the dependence among the cause-specific recurrent processes. Besides, to consider a nonparametric approach to model the frailty density using a Dirichlet process mixture (DPM) prior. Additionally, to propose a hybrid Markov chain Monte Carlo (MCMC) sampler algorithm composed by Hamiltonian Monte Carlo (HMC) and Gibbs sampling to compute the posterior estimates with respect to the frailty distribution. Regarding PLP parameters, to propose a class of noninformative priors whose resulting posterior distributions are proper and to obtain closed-form Bayesian estimators (Chapter 5).

1.3 Overview of the chapters

This thesis consists of three main chapters (based on an overview of two papers), which are somewhat related, but each is in itself an independent portion and is a new contribution for research. The chapters are explored around a central theme that is repairable systems under a competing risks framework and minimal repair assumption with PLP. Each chapter is independent of the others, in the sense that intrinsic notations and comments are introduced to each one of them. In Chapter 4, results from Chapter 3 are cited, but the function of each of these chapters is to be independent.

Therefore, this thesis is organized as follows. In Chapter 2, we present notions and definitions that are necessary for our discussions later. First, we give a short overview of fundamental concepts regarding counting processes, repairable systems, competing risks, frailty and Bayesian inference. In Chapter 3, we initially present a dataset in order to motivate our discussion about a single complex repairable system subject to independent competing risks. We explore an alternative parametrization for the PLP parameters based on orthogonal parameters. Results of classical inference are presented and used in the dataset. Chapter 4 begins with a previously unpublished dataset about recurrent failures history of a sugarcane harvester, as a motivating example. Moreover, we present an objective Bayesian methodology to analyze a repairable system subject to several independent failure modes. A simulation study was carried out to analyze the efficiency of the methods proposed and compare both classical and Bayesian approaches. We apply our proposed methodology in the dataset analysis described initially. Finally, we present the conclusion and make some concluding remarks. Chapter 5 begins with the fundamentals of multiple repairable systems, dependent competing risks and a shared frailty model. The Bayesian framework (parametric and nonparametric) is developed with a discussion on the choice of prior distributions for the proposed model and the computation of posterior distributions. An extensive simulation study is described in order to evaluate the efficiency of the proposed Bayesian estimators, and uses them to analyze a real data set that comprises the failure history for a fleet of cars under warranty. Moreover, we conclude the chapter with final remarks. In Chapter 6, general comments and extension possibilities of this current research are presented.

1.4 Products of the thesis

- Almeida, M.P.; Tomazella, V.L.D.; Gilardoni, G.L.A; Ramos, P.L.; Louzada, F.; Nicola, M., (2019). *Objective Bayesian Inference for a Repairable System Subject to Competing Risks*. This paper is currently under review (IEEE Transactions on Reliability).

- Almeida, M.P.; Paixão, R.S; Ramos, P.L.; Tomazella, V.L.D.; Ehlers, R.S.; Louzada, F., (2019). *Multiple repairable systems under dependent competing risks with nonparametric Frailty*. This paper is currently under review (Technometrics).

CHAPTER 2

BACKGROUND

In this chapter, basic concepts and literature review of some important topics that are covered throughout this thesis will be described.

2.1 Counting process

A univariate point process can be described by an increasing sequence of positive random variables $0 < T_1 \leq T_2 \leq \ldots$, or by its corresponding counting process $N(t) = N_t(\omega)$ with $t \in \mathbb{R}_+$, this being a random variable that counts the number of points T_n that occur until the time t, formally written as

$$N_t(\omega) = \sum_{k \geq 1} \mathbb{I}(T_k(\omega) \leq t), \tag{2.1}$$

which is, for each realization ω, a non-decreasing, right-continuous step function with size jump 1. Let $\mathbb{I}(\cdot)$ be the indicator function that represents the number of events up to time t . Thus, the specifications $N(t)$ and T_n are equivalent and carry the same information.

Definition 2.1.1 (Counting Process)**.** $\{N(t) : t \geq 0\}$ is a random variable that denotes the random number of points (i.e. the events of interest) that have occurred in the time interval $[0,t]$ and must satisfies the following conditions:

1. $N(t) \geq 0$ and $N(0) = 0$;

2. $N(t)$ is integer valued;

3. $t \to N(t)$ is right-continuous;

4. $\Delta N(t) = \lim_{h \to 0} N(t) - N(t-h) = 0$.

Definition 2.1.2. A counting process has stationary increments if for all k,

$$P(N(t,t+s] = k) \tag{2.2}$$

is independent of t.

Definition 2.1.3. A counting process $\{N(t) : t \geq 0\}$ has independent increments if for all n and for all $r_1 < s_1 \leq r_2 < s_2 \leq \cdots < r_n \leq s_n$, the random variables $N(r_1,s_1], N(r_2,s_2], \ldots, N(r_n,s_n]$ are independent, i.e.,

$$P(N(r_1,s_1] = k_1, \ldots, N(r_n,s_n] = k_n) = \prod_{i=1}^{n} P(N(r_i,s_i] = k_i). \tag{2.3}$$

The assumption of independent increments means that the number of events in nonoverlapping intervals are independent.

The process of counting is governed by its intensity $\lambda(t)$, and takes into account the history $\mathbb{H}_t = \{t_i : i = 1,2,\ldots,N(t)\}$ from the occurrence of events of a process up until t. According to Andersen *et al.* (2012) and Krit (2014), all the random variables $N(t)$, are defined in the same probability space $(\Omega, \mathbb{A}, \mathscr{P})$. A filtration $\{\mathbb{H}_t : t \geq 0\}$ is an increasing sequence of sub-σ algebras of $\mathbb{A}$, i.e., $s < t \Rightarrow \mathbb{H}_s \subset \mathbb{H}_t$. $\{N(t) : t \geq 0\}$ is $\mathbb{H}_t$-adapted if and only if for all $t \geq 0$, $N(t)$ is $\mathbb{H}_t$-mensurable. So, one can say that the filtration $\mathbb{H}_t$ consists of all information of the history at time t. Thus, this "history", in turn, influences $N(t)$. As $\{N(t) : t \geq 0\}$ is a non-decreasing, right-continuous step function that alters its values only at the times $T_1, T_2, \ldots, T_n$, its history at time t is fully known by the number and the times of events occurred in $[0,t]$. Hence, $\mathbb{H}_t$ is the σ-algebra generated by the history of the process at time t, i.e., $\mathbb{H}_t = \sigma(N(t), T_1, \ldots, T_{N(t)})$.

Definition 2.1.4. The complete intensity function is specified by

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P(N(t,t+\Delta t] \geq 1 | \mathbb{H}_t)}{\Delta t}, \tag{2.4}$$

which represents the instantaneous probability of an event occurring, conditioned to the process history.

The counting process that has the property of independent increments has its intensity function given by

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P(N(t,t+\Delta t] \geq 1)}{\Delta t}, \tag{2.5}$$

and its cumulative intensity function is

$$\Lambda(t) = \int_0^t \lambda(u)du. \tag{2.6}$$

Multivariate counting processes

A collection of n univariate counting processes $\{N_i(t) : t \geq 0, i = 1, \ldots, n\}$ constitutes a multivariate counting process, denoted $\{\mathbf{N}(t) : t \geq 0\}$ or $\mathbf{N}(t) = (N_1(t), \ldots, N_n(t))^T$, with the additional assumption that no two components jump simultaneously (with probability one). Associated with a multivariate counting process $\{\mathbf{N}(t) : t \geq 0\}$ is an intensity process $\{\lambda(t) : t \geq$

$0\}$ where $\lambda(t) = (\lambda_1(t), \ldots, \lambda_n(t))^T$ is a collection of n intensity processes, where $\lambda_i(t)$ is the intensity process of N_i. This process is adapted to a filtration (or history) $\{\mathbb{H}_t\}$. The smallest filtration that makes $\mathbf{N}(t)$ adapted is the filtration generated by the process itself (ANDERSEN *et al.*, 2012; AALEN; BORGAN; GJESSING, 2008).

2.1.1 Poisson process

Poisson processes are a particular case of point processes and can be used to model occurrences (and counts) of rare events over time, when they are not affected by the past. In particular, they are applied to describe and predict the failures of a given system.

In many situations in reliability, medicine and social sciences, we need to count events to a specific point in time in such a way that we can describe them by a Poisson process. A Poisson process is a particular case of the Markov process in continuous time, where the only possible jumps are to the next higher state. A Poisson process can also be seen as a counting process that has good properties. Technical details of the subject can be found in Kingman (1993), Aalen, Borgan and Gjessing (2008), Rigdon and Basu (2000).

One can define the Poisson process based on an infinitesimal representation of the distribution of points in small intervals. Toward this end, we use *little-oh* notation. One define $f(h) = o(h)$ to mean that $\lim_{h\to 0} \frac{f(h)}{h} = 0$.

Definition 2.1.5. A counting process $\{N(t) : t \geq 0\}$ is said to be a Poisson process with rate λ, $\lambda > 0$, if:

(i) $N(0) = 0$;

(ii) the process has independent and stationary increments;

(iii) $P\{N(h) = 1\} = \lambda h + o(h)$;

(iv) $P\{N(h) \geq 2\} = o(h)$.

It is worth pointing out that, a Poisson process in which the rate does not depend on time is called the homogeneous Poisson process (HPP).

Definition 2.1.6. The counting process $\{N(t) : t \geq 0\}$ is a NHPP with intensity function $\lambda(t)$, $t \geq 0$, if:

(i) $N(0) = 0$;

(ii) the process has independent increments;

(iii) $P\{N(t+h) - N(t) = 1\} = \lambda h + o(h)$;

(iv) $P\{N(t+h) - N(t) \geq 2\} = o(h)$.

The importance of the NHPP lies in the fact that it does not require the condition of stationary increments. Furthermore, the independent increments property of the Poisson process mandates that $\lambda(t)$ does not depend on the history of the process up until t.

From Definitions 2.1.1 to 2.1.5, it is possible to deduce the property that the number of events (e.g., failures) in a interval is a random variable with Poisson distribution.

Theorem 2.1.7. As a consequence of Definitions 2.1.1 to 2.1.5, for $n \in \mathbb{Z}^+$,

$$P\{N(t) = n\} = \frac{1}{n!}\left(\int_0^t \lambda(s)ds\right)^n \exp\left(-\int_0^t \lambda(s)ds\right) \quad \text{with} \quad n = 0, 1, 2, \dots \tag{2.7}$$

Proof. See Rigdon and Basu (2000). □

In particular, for a HPP with rate λ, we have that $N(a,b] \sim Poisson(\lambda(b-a))$, and the increments are stationary, since their distribution does not depend on the starting point of the interval, but only on its length, i.e., the process rate will be λ for all times.

Definition 2.1.8. The mean function of a Poisson process is given by $\Lambda(t) = \mathbb{E}[N(t)] = \int_0^t \lambda(s)ds$, $t \geq 0$.

The expected number of events in the interval $(a,b]$ is $\Lambda(a,b] = \Lambda(b) - \Lambda(a)$. For the HPP case, we have that the mean number of points within the interval $(a,b]$ is $\Lambda(a,b] = \int_a^b \lambda ds = \lambda(b-a)$ and for $(0,t]$ is $\Lambda(t) = \int_0^t \lambda ds = \lambda t$. The constant λ is interpreted as the average rate of process points. With regard to NHPP, we have that $\Lambda(t) = \int_0^t \lambda(s)ds$ and $\Lambda(a,b] = \int_a^b \lambda(s)ds$. In terms of differentiable Λ, the rate of occurrences of failures (ROCOF) is determined by $\frac{d}{dt}\Lambda(t)$.

From Equation (2.7), one can define the reliability of a system through of $P(T > t) = P\{N(0,t] = 0\} = e^{-\lambda t}$, which is the probability of not finding points (events/failures) in a interval of length t. According to Meeker and Escobar (2014), Rigdon and Basu (2000), Ascher and Feingold (1984), the NHPP intensity function can be seen as a reliability measure because it is capable of modeling systems which wear out or improve over time. For instance, if the intensity is increasing, the times between failures are decreasing over time, and the system is deteriorating; and if the intensity is decreasing, the times between failures are increasing, and therefore, the system is improving.

2.2 Repairable system

In line with the definition given by Rigdon and Basu (2000), a repairable system is understood to be a system which, after failure, can be restored to an operating condition by some repair action other than replacement of the entire system. On the other hand, a nonrepairable system is one that is discarded after the first and only failure.

As we have pointed out already, the effect of a failure and subsequent repair action on the performance of a repairable system is what differentiates many models for repairable systems, as these repairs can have different effects on the reliability of a system. For example, if the repair returns the system to the exact condition it was in immediately before it failed, then the repair is said to be minimal or ABAO. This leads to the NHPP. By contrast, if a system is replaced with a new one (or the repair action leaves the system as if it were new) then the times between failures are independent and identically distributed. Such a repair action is called a renewal process or AGAN. Imperfect repair models are ones in which repair actions leave the system somewhere between the ABAO and AGAN. As we mentioned earlier, only the minimal repair model will be treated in this thesis.

2.2.1 A single repairable system

In order to be consistent with the current literature, we will borrow some definitions and notations from Reis *et al.* (2019). Failures times of a repairable system are described by a stochastic point process $T_1, T_2, \ldots$, where T_i denotes the i-th failure time, measured from the instant at which the system was first put into operation, $T = 0$. Time is not necessarily the calendar time, but it can be the length of a crack, operation time, total mileage of a vehicle, number of cycles, etc. When a failure occurs, a repair action is taken to bring the system back into operation. We make the assumption that repair times are negligible. The failure times generate a counting process $\{N(t) : t \geq 0\}$, where $N(t)$ counts the failures in the interval $(0,t]$.

In general, there are two ways to observe data from a repairable system. When the system observation window terminates after a predetermined number of failures, the data is said to be failure truncated. On the other hand, the data is said to be time truncated when the system observation period ends at a pre-specified time τ. Analyzing data from a time truncated system, both the number of failures $N(\tau)$ and the failure times $(T_1, T_2, \ldots, T_{N(\tau)})$ are random. The following theorem, due to Rigdon and Basu (2000), provides conditions concerning specifying the joint distribution of $(N(\tau); T_1, T_2, \ldots, T_{N(\tau)})$.

Theorem 2.2.1. If an NHPP with intensity function $\lambda(t)$ is observed until time τ, and if the failure times are $T_1 < T_2 < \cdots < T_{N(\tau)}$ where $N(\tau)$ is the random number of failures in the interval $(0, \tau]$, then conditioned on $N(\tau) = n$, the random variables $T_1 < T_2 < \cdots < T_n$ are distributed as n order statistics from the distribution with a cumulative distribution function

$$G(t) = \begin{cases} 0, & t \leq 0; \\ \Lambda(t)/\Lambda(\tau), & 0 < t \leq \tau; \\ 1, & t > \tau. \end{cases}$$

Hence, the aforementioned joint distribution is given by

$$\begin{aligned}P(N(\tau)=n;T_1=t_1,\ldots,T_n=t_n) &= P(N(\tau)=n)P(T_1=t_1,\ldots,T_n=t_n|N(\tau)=n)\\ &= \frac{1}{n!}\left(\int_0^\tau \lambda(s)ds\right)^n \exp\left\{-\int_0^\tau \lambda(s)ds\right\}\times n!\left\{\prod_{j=1}^n G^{'}(t_j)\right\}\\ &= \left\{\prod_{j=1}^n \lambda(t_j)\right\}\exp\{-\Lambda(\tau)\},\quad 0<t_1<t_2<\cdots<t_n<\tau.\end{aligned}$$

Proof. See Rigdon and Basu (2000). □

2.2.2 Multiple repairable systems

In this section, we present a brief overview to analyze data from multiple repairable systems, but we refer the reader to Rigdon and Basu (2000) and Oliveira, Colosimo and Gilardoni (2012) for details and proofs. Here, we highlight just two important assumptions in this context. The first is to assume that all systems are identical or different. The second is to assume that all systems have the same truncation time τ or, otherwise, have different truncations at τ_j. However, for the sake of simplicity and brevity of exposition, we assumed the observation lengths, τ, for each system to be equal. Moreover, in this thesis, we assume all systems to be identical, i.e., the systems are specified as m independent realizations of the same process, with intensity function λ.

If the multivariate counting processes $N_1(t),\ldots,N_m(t)$ are all observed at the same time τ, the NHPP resulting from the superposition of NHPPs is given by $N(t)=\sum_{j=1}^m N_j(t)$ and has an intensity function given by $\lambda(t)=m\lambda(t)$; e.g., overlapping realizations of a PLP. Therefore, inferences in models proposed for this framework can be made through the following likelihood function

$$L(\lambda)=\left(\prod_{j=1}^m\prod_{i=1}^{n_j}\lambda(t_{ji})\right)\exp\left(-\sum_{j=1}^m\int_0^\tau \lambda(s)ds\right).$$

Unobserved heterogeneity between multiple systems

The m systems are considered to be identical, and therefore have the same intensity function and thus we would have a random sample of systems. On the other hand, this assumption may not be true. That is, in many real-world reliability applications there may be some heterogeneity between "apparently identical" repairable systems. In this case, it is necessary to propose a statistical model capable of capturing this heterogeneity. Cha and Finkelstein (2014), Asfaw and Lindqvist (2015), Slimacek and Lindqvist (2016) and Slimacek and Lindqvist (2017) discuss frailty models for modeling and analyzing repairable systems data with unobserved heterogeneity.

2.2.3 The minimal repair model

Before turning to formal definitions, we provide an intuitive and real example. We said earlier that a minimal repair policy is enough to make the system operational again. For example, if the water pump fails on a car, the minimal repair consists only of repairing or replacing the water pump. As we said before, the purpose is to bring the car (system) back to operation as soon as possible. From an economic perspective, complex systems are commonly repaired rather than replacing the system with a new one after failure.

Recalling that NHPP is completely specified by its intensity function, then when parametric models are adopted for the intensity function of an NHPP, we are interested in making inferences about the parameters of this function. In addition, one knows that the NHPP forms a class of models that naturally applies to a "minimal repair", i.e., the repair brings the system back into the same state it was in just prior to the failure. One of the most important and used functional forms is the PLP.

The parametric form for the PLP intensity is given by

$$\lambda(t) = (\beta/\mu)(t/\mu)^{\beta-1}, \tag{2.8}$$

where $\mu, \beta > 0$. Its mean function is

$$\Lambda(t) = \mathbb{E}[N(t)] = \int_0^t \lambda(s)ds = (t/\mu)^{\beta}. \tag{2.9}$$

The scale parameter μ is the time for which we expect to observe a single event, while β is the elasticity of the mean number of events with respect to time (OLIVEIRA; COLOSIMO; GILARDONI, 2012).

Since (2.8) increases (decreases) in t for $\beta > 1$ ($\beta < 1$), the PLP can accommodate both systems that deteriorate or improve over time. Of course, when $\beta = 1$, the intensity (2.8) is constant and hence the PLP becomes an HPP.

Under minimal repair, the failure history of a repairable system is modeled as an NHPP. As mentioned above, the PLP (2.8) provides a flexible parametric form for the intensity of the process. Under the *time truncation* design, i.e. when failure data is collected up to time T, the likelihood becomes

$$L(\beta, \mu \mid n, \mathbf{t}) = \frac{\beta^n}{\mu^{n\beta}} \left(\prod_{i=1}^{n} t_i \right)^{\beta-1} \exp\left[-\left(\frac{T}{\mu} \right)^{\beta} \right], \tag{2.10}$$

where we assume that $n \geq 1$ failures at times $t_1 < t_2 < \ldots < t_n < T$ were observed, $i = 1, \ldots, n$ (RIGDON; BASU, 2000). The maximum likelihood estimators (MLEs) of the parameters are given by

$$\hat{\beta} = n / \sum_{i=1}^{n} \log(T/t_i) \quad \text{and} \quad \hat{\mu} = T/n^{1/\hat{\beta}}. \tag{2.11}$$

2.2.4 Criticisms of classical inference in the scenario of repairable systems with PLP model

The obtained MLEs in (2.11) are biased. Some studies have attempted to overcome the shortcomings encountered in this approach in terms of bias and confidence intervals for the model parameters. For instance, Bain and Englehardt (1991) and Rigdon and Basu (2000) discussed unbiased estimators for the shape parameter β (except for the scale parameter μ). The latter authors presented the following estimator

$$\tilde{\beta} = \frac{n-1}{n}\hat{\beta}, \tag{2.12}$$

which is called the conditionally unbiased estimator (CMLE) (hereafter referred to as CMLE), i.e.,

$$\mathbb{E}\left[\left(\frac{n-1}{n}\right)\hat{\beta}\middle| n\right] = \beta. \tag{2.13}$$

The proof of (2.13) can be found in the references mentioned above. It is important to note that the authors do not present a natural way to obtain such unbiased estimator (for β). However, as we will see in Chapter 4, we overcome this problem by proving that this CMLE is the Bayes estimator under maximum a posteriori (MAP) estimator using objective priors. Regarding the μ parameter, we will also discuss a way to obtain an unbiased estimator (in the same Bayesian framework) taking into account an advantageous reparametrization.

Another potential concern is regarding the procedures to derive confidence intervals under the classical inference. Rigdon and Basu (2000) argued that the confidence interval for the scale parameter has no simple interpretation and, the usual methodologies return extremely wide intervals. Moreover, in some situations, the pivotal quantity used to obtain such intervals does not exist or it is difficult to be derived. Bain and Englehardt (1991, ch. 9) were emphatic about the difficulties of obtaining confidence intervals for the scale parameter because of the non-existence of the pivotal quantity in the context of time-truncated data. Recent studies have been presented to derive confidence intervals for the scale parameter. However, the majority has presented limitations. For instance, Gaudoin, Yang and Xie (2006) investigated the interval estimation for the scale parameter from the PLP model. They used the Fisher information matrix to obtain asymptotic confidence intervals. Nevertheless, their results have many restrictions. Wang, Xie and Zhou (2013) considered a procedure to derive a generalized confidence interval for the scale parameter under some general conditions. However, the authors advertise that the proposed procedure is more complex than the asymptotic confidence interval from the findings of Gaudoin, Yang and Xie (2006). Somboonsavatdee and Sen (2015b) provided methods for deriving confidence intervals for the scale parameter for a system failing due to competing risks from a frequentist perspective. The study showed that the use of the large-sample confidence intervals for scale parameters are wide.

2.2.5 Reparametrized PLP

Oliveira, Colosimo and Gilardoni (2012) suggest reparametrizing the model (2.8) in terms of β and α, where the latter is given by

$$\alpha = \mathbb{E}[N(T)] = (T/\mu)^{\beta}, \tag{2.14}$$

so that the likelihood (2.10) becomes

$$\begin{aligned} L(\beta,\alpha|n,\mathbf{t}) &= c\left(\beta^{n} e^{-n\beta/\hat{\beta}}\right)\left(\alpha^{n} e^{-\alpha}\right) \\ &\propto \gamma(\beta|n+1, n/\hat{\beta})\gamma(\alpha|n+1,1)\,, \end{aligned} \tag{2.15}$$

where $c = \prod_{i=1}^{n} t_i^{-1}$, $\hat{\beta} = n/\sum_{i=1}^{n}\log(T/t_i)$ is the MLE of β and $\gamma(x|a,b) = b^a x^{a-1} e^{-bx}/\Gamma(a)$ is the probability density function (PDF) of the gamma distribution with shape and scale parameters a and b, respectively. It is important to point out that β and α are orthogonal parameters. For the advantages of having orthogonal parameters, see Cox and Reid (1987).

2.3 Competing risks

In reliability theory, the most common system configurations are the series systems, parallel systems, and series-parallel systems. In a series system, components are connected in such a way that the failure of a single component results in system failure. Such a system is depicted in Figure 1. A series system is also referred to as a competing risks system since the failure of a system can be classified as one of the p possible risks (components) that compete for the failure of the system. In general, the observations of a competing risks model consist of the

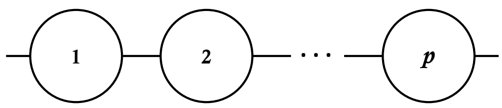

Figure 1 – Diagram for a competing risks system (i.e., series system) with p risks (components).

pair (t, δ), where $t \geq 0$ represents the time of failure and δ is the indicator of the component which failed. An example follows to illustrate the failure history data for this kind of framework.

Example 2.3.1. Suppose a repairable system, and let $0 < T_1 < T_2 < T_3 < \cdots < T_{N(\tau)} < \tau$ be the failure times of the system observed until a pre-fixed time τ. Moreover, there are two ($p = 2$) recurrent causes of failure, and at the i-th failure time T_i, we also observe $\delta \in \{1,2\}$, which is the cause of the failure related to the i-th failure (see Figure 2).

Basically, we could say that in most of the literature, there are two main approaches when analyzing failure times with competing risks: independent and dependent competing failure

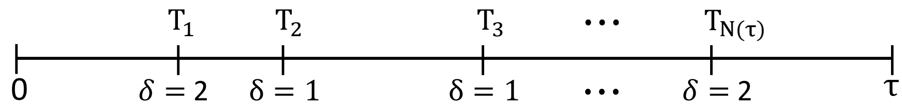

Figure 2 – Observable quantities from failure history of a repairable competing risks system with two recurrent causes of failure.

modes. For reliability models under competing risks, most research has been carried out considering statistical independence of component failure (CROWDER *et al.*, 1994; HØYLAND; RAUSAND, 2009; SOMBOONSAVATDEE; SEN, 2015b; TODINOV, 2015; WU; SCARF, 2017). Thus, one assumes that independent risks are equivalent to independent causes of failure. However, in some particular contexts (for instance, the existence of interactions between components in complex systems), the assumption of independent risks may lead to seriously misleading conclusions. To overcome this issue, some important and general approaches have been presented in the literature for modeling dependent competing risks data (DIJOUX; GAUDOIN, 2009; ZHANG; YANG, 2015; ZHANG; WILSON, 2017).

Considering Daniel Bernoulli's attempt in the 18th century to separate the risk of dying due to smallpox from other causes (BERNOULLI, 1760; BRADLEY; BRADLEY, 1971), the competing risks methodology has disseminated through various fields of science such as demography, statistics, actuarial sciences, medicine and reliability analysis. Therefore, one knows that both the theory and application of competing risks is too broad to cite here, but for an overview of the basic foundations, please see Pintilie (2006), Crowder (2001), Crowder (2012). For repairable systems failing due to competing risks, we refer the reader to Langseth and Lindqvist (2006), Doyen and Gaudoin (2006), Somboonsavatdee and Sen (2015b). Particularly, this thesis responds directly to the application in repairable systems under a recurrent data structure based on stochastic processes, which is the most natural way to describe the recurrence of multiple event types that occur over time.

2.3.1 Recurrent competing risks model for a single repairable system

The assumption of the repairable system under examination is that the components can perform different operations, and thus be subject to different types of failures. Hence, in our model there are K causes of failure. If n failures have been observed in $(0,T]$, then we observe the data $(t_1,\delta_1),\dots,(t_n,\delta_n)$, where $0 < t_1 < \cdots < t_n < T$ are the system failure times and $\delta(t_i) = \delta_i = q$ represents the q-th associated failure cause with i-th failure time, $i = 1,\dots,n$ and $q = 1,\dots,K$.

One can introduce a counting process $\{N_q(t); t \geq 0\}$ whose behavior is associated with the cause-specific intensity function

$$\lambda_q(t) = \lim_{\Delta t \to 0} \frac{P(\delta(t) = q, N(t+\Delta t] - N(t) = 1 \mid N(s), 0 \leq s \leq t)}{\Delta t}. \tag{2.16}$$

Let N_q be the cumulative number of observed failures for the q-th cause of failure and $N(t) = \sum_{q=1}^{K} N_q(t)$ be the cumulative number of failures of the system. Thus, $N(t)$ is a superposition of NHPPs and its intensity function is given by

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P(N(t+\Delta t] - N(t) = 1 \mid N(s), 0 \leq s \leq t)}{\Delta t} = \sum_{q=1}^{K} \lambda_q(t). \tag{2.17}$$

The cause-specific and the system cumulative intensities are, respectively,

$$\Lambda_q(t) = \int_0^t \lambda_q(u) du \quad \text{and} \quad \Lambda(t) = \sum_{q=1}^{K} \Lambda_q(t). \tag{2.18}$$

Under minimal repair, the failure history of a repairable system is modeled as an NHPP. We give special attention to functional form for the cause-specific intensity according to the PLP, as follow

$$\lambda_q(t) = \frac{\beta_q}{\psi_q} \left(\frac{t}{\psi_q} \right)^{\beta_q - 1}, \tag{2.19}$$

with $t \geq 0$, $\psi_q > 0$, $\beta_q > 0$ and for $q = 1, \ldots, K$. The model is quite flexible because it can accommodate both decay ($\beta_q < 1$) and growth ($\beta_q > 1$) in reliability. The corresponding mean function considering time-truncated scenario (with fixed time T) is

$$\mathbb{E}[N_q(T)] = \Lambda_q(T) = \left(\frac{T}{\psi_q} \right)^{\beta_q}. \tag{2.20}$$

If we reparametrize (2.19) in terms of β_q and α_q, where the latter is given by

$$\alpha_q = \left(\frac{T}{\psi_q} \right)^{\beta_q}, \tag{2.21}$$

one obtains the following advantageous likelihood function

$$\begin{aligned} L(\theta | \mathbf{t}, \delta) &= \left\{ \prod_{i=1}^{n} \prod_{q=1}^{K} \left[\beta_q \alpha_q t_i^{\beta_q - 1} T^{-\beta_q} \right]^{\mathbb{I}(\delta_i = q)} \right\} \exp \left\{ \sum_{q=1}^{K} \alpha_q \right\} \\ &\propto \prod_{q=1}^{K} \gamma(\beta_q | n_q + 1, n_q / \hat{\beta}_q) \prod_{q=1}^{K} \gamma(\alpha_q | n_q + 1, 1), \end{aligned} \tag{2.22}$$

where $n = \sum_{q=1}^{K} n_q$; $n_q = \sum_{i=1}^{n} \mathbb{I}(\delta_i = q)$; $\theta = (\beta, \alpha)$ with $\beta = (\beta_1, \ldots, \beta_K)$ and $\alpha = (\alpha_1, \ldots, \alpha_K)$; $\hat{\beta}_q = n_q / \sum_{i=1}^{n_q} \log(T/t_i) = n_q / \sum_{i=1}^{n} \log(T/t_i) \mathbb{I}(\delta_i = q)$.

2.4 Frailty model

Frailty models are generalizations of the well-known Cox model (COX, 1972), introduced by Vaupel, Manton and Stallard (1979). Over the past decades, most research in frailty has emphasized the analysis of medical and reliability data that present heterogeneity, which cannot be adequately explained by the Cox model. To be more precise, it can be said that, the frailty term is a random effect that acts multiplicatively on the hazard function of the Cox model. This random effect could represent misspecified or omitted covariates (unknown or unmeasured effects). Thus, one can say that such a term (frailty) is an unobservable (latent) quantity. In addition, the frailty methodology is very effective to account for dependency in event times that result from unknown sources of heterogeneity. For more details on general frailty theory, see Hougaard (2012), Hanagal (2011).

Considering recurrent event data, several approaches of the Cox model with a frailty factor have been discussed in the literature (TOMAZELLA, 2003; BERNARDO; TOMAZELLA, 2011). Additional results on frailty in the counting process context are given in Aalen, Borgan and Gjessing (2008) and Andersen *et al.* (2012). In the reliability field, the frailty model is commonly used to model heterogeneous repairable systems (CHA; FINKELSTEIN, 2014; ASFAW; LINDQVIST, 2015). Such heterogeneity is generated because some units have a higher (or lower) event rate than other units due to unobserved or unknown effects (e.g., instability of production processes, environmental factors, etc.). On the other hand, Somboonsavatdee and Sen (2015a) present a classical inference for repairable systems under dependent competing risks where the frailty is considered to model the dependence between the components arranged in series.

The many approaches differ in the modeling of the baseline hazard or in the distribution of the frailty. There is a vast amount of published studies describing fully parametric approaches. Regarding the probability distribution that should be assigned to the frailty term (random effect), in general, it follows a distribution appropriate for a positive random variable. Parametric frailty models are standard in the literature (AALEN; BORGAN; GJESSING, 2008; HOUGAARD, 2012) and the so-called gamma frailty model, in which the unobserved effects are assumed to be gamma distributed, is probably the most popular choice. Various frailty distributions are presented in Hougaard (2012) and the references therein such as the gamma, inverse Gaussian, log-normal or the positive stable frailty. Other distributions include the power variance frailty (AALEN; BORGAN; GJESSING, 2008) and the threshold frailty (LINDLEY; SINGPURWALLA, 1986).

Extensive research has been carried out on frailty distributions, as cited above, and it is well known that, generally, such distributions are primarily used by mathematical convenience. Furthermore, in general, such distributions do not encompass a range of possible features including skewness and multimodality. Furthermore, because the frailty variable is an unobservable quantity, it cannot be tested to verify whether or not it satisfies the distributional assumption (FERREIRA; GARCIA, 2001). It is known that the misspecification of this distribution can lead to several types of errors, including, for example, poor parameter estimates (WALKER; MALLICK, 1997). A more flexible and robust approach would be to estimate such a density using the nonparametric Bayesian methodology. In this thesis, this solution will be explored in Chapter 5.

Finally, a suitable choice of the distribution of unobserved effects can provide interesting general results, but generally the main quantity of interest is the variance of the unobserved effects. Usually, a significant variance may indicate high dependence (WIENKE, 2003; SOMBOONSAVATDEE; SEN, 2015a).

2.4.1 Shared frailty

In order to emphasize the subject matter of the repairable systems under dependent competing risks, we introduce the shared frailty model using the multivariate counting processes framework based on cause-specific intensity functions.

The referred dependency between competing risks may be modeled through a frailty variable, say Z, in such a way that, when the frailty is shared among several units in a cluster, it leads to dependence among the event times of the units (WIENKE, 2003; TOMAZELLA, 2003; HOUGAARD, 2012). Suppose that m clusters (or systems) are under observation, where each cluster is composed by K units (or components). The intensity function of the j-th cluster ($j = 1, \ldots, m$) of a shared frailty model is that of the Cox model multiplied by a frailty term Z_j (multiplicative random effect model). More specifically, for each individual counting process, $\{N_j(t) : t \geq 0\}$, their intensity function, conditionally on the frailty Z_j, is given by

$$\lambda_j(t \mid Z_j) = Z_j \lambda(t), \tag{2.23}$$

where $\lambda(t)$ is the basic intensity function and $j = 1, 2, \ldots, m$. The intensity function (2.23) describes the recurrent failure process on the j-th cluster and the intensity associated to the q-th component from the j-th cluster is defined as

$$\lambda_{jq}(t \mid Z_j) = Z_j \lambda_q(t), \tag{2.24}$$

where $\lambda_q(t)$ is the basic intensity function from the q-th component (cause-specific intensity function), $q = 1, 2, \ldots, K$. Note that intensities (2.23) and (2.24) follow the relation $\lambda(t) = \sum_{q=1}^{K} \lambda_q(t)$ (ANDERSEN *et al.*, 2012). Henceforth, we will omit the subscript j from λ_{jq} in (2.24) since we are assuming that the systems are identical, therefore $\lambda_q(t \mid Z_j) = Z_j \lambda_q^{'}(t)$. Note also that $\lambda_q^{'}(t)$ is referred to as the (basic) intensity function for type q events (e.g., PLP). Let $\mathbf{Z} = (Z_1, \ldots, Z_m)$ denote the vector of Z_js, which we assume arises from density $f_Z(\cdot)$, where each Z_j is independent and identically distributed (iid). These are typically parametrized so that both $E(Z)$ and $Var(Z)$ are finite, for $j = 1, 2, ..., m$. It is worth pointing out that the Z_js are assumed to be stochastically independent of the failure process $\lambda_q^{'}(t)$ (ANDERSEN *et al.*, 2012; SOMBOONSAVATDEE; SEN, 2015a).

The term of frailty in (2.23) aims to control the unobserved heterogeneity among systems. If we consider the situation where the dataset is divided into clusters (multiple units in a cluster), this term evaluates the dependence between the units that share the frailty Z_j. Thus, units from heterogeneous populations can be considered independent and homogeneous, conditionally to the terms of frailty (Z_js) attributed to the units or cluster of units.

The evaluation of the influence of unobserved heterogeneity in this type of data is made on the basis of the variability of the frailty distribution. In addition, it is worth pointing out that higher values of $Var(Z)$ mean greater heterogeneity among units and more dependence between the event times for the same unit. In general, in the literature, it is common to specify a distribution for the frailty variable with mean 1 and variance, say $Var(Z) = \eta$, in order to obtain two main advantages: (1) the model parameters become identifiable, and (2) it is possible to obtain an easily understandable interpretation of the model, because, as previously argued, η acts as a dependence parameter, meaning that, if the frailty variance is zero, it implies that we have independence between the event times in the clusters (since it is assumed that the mean is 1).

We emphasize here that the main focus of Chapter 5 of the present work will be to analyze data from multiple repairable systems under the presence of dependent competing risks. Thus, to estimate the model parameters considering shared frailty. In this sense, naturally, we can consider the dependence between the components as a nuisance parameter via frailty.

2.5 Bayesian inference

Because the chapters of this thesis present two different approaches to the Bayesian framework (parametric and nonparametric), it is important to define some differences between them. Then, to facilitate the exposure without the intention of exhausting the subject, we will present a brief definition of the tools used in each approach.

According to currently accepted theories (COX; HINKLEY, 1979; PAWITAN, 2001; CASELLA; BERGER, 2002; MÜLLER; QUINTANA, 2004; SCHERVISH, 2012), the fundamental problem to which the study of Statistics is addressed is that in which randomness is present. The statistical methodology to deal with the resulting uncertainty is based on the construction of probabilistic models that represent, or approximate, the generating mechanism of a random phenomenon under study. Specifically, data are conceived as realizations of a collection of random variables $y_1, \ldots, y_n$, where y_i itself could be a vector of random variables corresponding to data that are collected on the i-th observational unit in a sample of n units from some population of interest. A usual assumption is that the y_is are drawn independently from some underlying probability distribution F. The statistical problem begins when there is uncertainty about F. Let f denote the PDF of F. A statistical model arises when f is known to be a member f_θ from a family $\mathscr{F}^* = \{f_\theta : \theta \in \Theta\}$ labeled by a set of parameters θ from an index set Θ.

In addition, it is known that probabilistic models that are specified through a vector θ of a finite number of, usually, real values are referred to as finite-dimensional or parametric models. Parametric models can be described as $\mathscr{F}^* = \{f_\theta : \theta \in \Theta \subset \mathbb{R}^p\}$. The purpose of the analysis is then to use the observed sample to account for a plausible value for θ, or at least to determine a subset of Θ, which plausibly contains θ. In many situations, however, constraining inference to a specific parametric form may limit the scope and type of inferences that can be drawn from such models. Therefore, one can relax parametric assumptions to allow greater modeling flexibility and robustness against misspecification of a parametric statistical model. In these cases, one may want to consider models where a richer and larger class of densities which can no longer be indexed by a finite-dimensional parameter θ, and one can therefore require parameters θ in an infinite-dimensional space. Such a procedure is the central idea of traditional nonparametric analysis. In particular, to conduct Bayesian inference in a nonparametric framework, it is necessary to complete the probability model with a prior distribution on the infinite-dimensional parameter. Such priors are known as nonparametric Bayesian priors (MÜLLER; QUINTANA, 2004).

With this clear and succinct overview given above, we can describe the Bayesian tools we used. First, we briefly present the parametric case with objective Bayesian inference (required in Chapter 4). Second, we present the nonparametric case in which we describe the general and concise idea of nonparametric Bayesian inference (required in Chapter 5).

2.5.1 Objective Bayesian inference

In this context, the prior distribution used to obtain the posterior quantities is of primary concern. Historical data or expert knowledge can be used to obtain a prior distribution. However, the elicitation process may be difficult and time-consuming. An alternative is to consider objective priors. In this case, we want to select a prior distribution in which subjective information obfuscates the information provided by the data. The formal rules to obtain such priors are presented next (BERNARDO; SMITH, 2009; KASS; WASSERMAN, 1996).

2.5.1.1 Jeffreys Prior

Jeffreys (1946) proposed a rule for deriving a non-informative prior which is invariant to any one-to-one reparametrization. The Jeffreys prior is one of the most popular objective priors and (in the multiparameter case) its density is proportional to the square root of the determinant of the expected Fisher information matrix $H(\theta)$, i.e.,

$$\pi^J(\theta) \propto |H(\theta)|^{1/2},$$

where the elements of $H(\theta)$ can be obtained through the $H_{i,j}(\theta) = -\mathbb{E}_\theta\left[\frac{\partial^2 \ell(\theta|\mathbf{t})}{\partial\theta_i \partial\theta_j}\right]$, $i, j = 1, \ldots, k$, $\theta = (\theta_1, \ldots, \theta_k)$ and $|\cdot|$ denotes a determinant.

Although Jeffreys prior performs satisfactorily in one-parameter cases, Jeffreys himself noticed that it might be reasonable for the multi-parameter case. The same author also argues that Bayesian estimators directly obtained from the Jeffreys prior usually have excellent frequentist properties.

2.5.1.2 Reference Prior

Reference priors can be seen as a collection of formal consensus prior functions which can be used as standards for scientific communication (BERNARDO, 2005). Bernardo (1979) introduced a class of objective priors known as reference priors. This class of priors maximizes the expected Kullback-Leibler divergence between the posterior distribution and the prior. The reference prior has minimal influence in a precise information-theoretical sense that separates the parameters into the parameters of interest and nuisance parameters. To derive the reference prior function, the parameters need to be set according to their order of inferential importance. See for instance, Bernardo (1979), Bernardo (2005).

However, the main problem is that different orderings of the parameters return different priors and selecting the more adequate prior may be quite challenging. To overcome this problem, Berger *et al.* (2015) discussed different procedures to construct the overall reference prior for all parameters. Additionally, under certain conditions, such a prior is unique in the sense of being the same regardless of the ordering of the parameters. The expected Fisher information matrix must have a diagonal structure to obtain this prior. The following result can be used to obtain the overall reference prior.

Theorem 2.5.1. [Berger *et al.* (2015)] Consider the unknown vector of parameters $\theta = (\theta_1, \ldots, \theta_k)$. If the Fisher information matrix $H(\theta)$ is of the form

$$H(\theta) = diag(f_1(\theta_1)g_1(\theta_{-1}), \ldots, f_k(\theta_k)g_k(\theta_{-k})),$$

where $\theta_{-i} = (\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_k)$, $diag$ is a diagonal matrix, $f_i(\cdot)$ and $g_i(\cdot)$ are positive functions of θ_i, for $i = 1, \ldots, k$, then the one-at-a-time reference prior, for any chosen parameter of interest and any ordering of the nuisance parameters in the derivation, hereafter, the overall reference prior is given by

$$\pi^R(\theta) \propto \sqrt{f_1(\theta_1)\ldots f_k(\theta_k)}. \tag{2.25}$$

Proof. See Berger *et al.* (2015). □

In this context, nuisance parameters are those that are not of primary interest but must be included in the analysis. On the other hand, in this scenario, all the parameters are of interest without a specific ordering of inferential interest.

The reference posterior distribution has desirable theoretical properties such as invariance under one-to-one transformations of the parameters, consistency under marginalization and consistent sampling properties.

2.5.1.3 Matching Priors

Researchers attempted to evaluate inferential procedures with good coverage errors for the parameters. While the frequentist methods usually rely on asymptotic confidence intervals, under the Bayesian approach, formal rules are proposed to derive such estimators. Tibshirani (1989) discussed sufficient conditions to derive a class of non-informative priors $\pi(\theta_1,\theta_2)$, where θ_1 is the parameter of interest so that the credible interval for θ_1 has a coverage error $O(n^{-1})$ in the frequentist sense, i.e.,

$$P\left[\theta_1 \leq \theta_1^{1-a}(\pi;t)|(\theta_1,\theta_2)\right] = 1-a-O(n^{-1}), \tag{2.26}$$

where $\theta_1^{1-a}(\pi;t)|(\theta_1,\theta_2)$ denote the $(1-a)$th quantile of the posterior distribution of θ_1. The class of priors satisfying (2.26) are known as matching priors (DATTA; MUKERJEE, 2012).

To obtain such priors, Tibshirani (1989) proposed to reparametrize the model in terms of the orthogonal parameters (ω,ζ), in the sense discussed by Cox and Reid (1987). That is, $I_{\omega,\zeta}(\omega,\zeta)=0$ for all (ω,ζ), where ω is the parameter of interest and ζ is the orthogonal nuisance parameter and $I_{\omega,\zeta}(\omega,\zeta)=0$ is the element of the Fisher information matrix given by

$$I(\omega,\zeta) = \begin{bmatrix} I_{\omega,\omega}(\omega,\zeta) & I_{\omega,\zeta}(\omega,\zeta) \\ I_{\omega,\zeta}(\omega,\zeta) & I_{\zeta,\zeta}(\omega,\zeta) \end{bmatrix}.$$

In this case, the matching priors are all priors of the form

$$\pi(\omega,\zeta) = g(\zeta)\sqrt{I_{\omega\omega}(\omega,\zeta)}, \tag{2.27}$$

where $g(\zeta)> 0$ is an arbitrary function and $I_{\omega\omega}(\omega,\ \zeta)$ is the ω diagonal entry of the Fisher information matrix. The same idea is applied to derive priors when there is a vector of nuisance parameters.

Bayesian point estimators

There are different types of Bayesian estimators. The three most commonly used are the posterior mean (quadratic loss function), the posterior mode (0-1 loss function) and the posterior median (absolute error loss). For more technical details, see Schervish (2012), O'Hagan (1994). In this thesis, we will consider both posterior mean (Chapters 4 and 5) and MAP (only in Chapter 4) estimators, that are usually referred to as the Bayes estimators. The MAP estimator, $\hat{\theta}^{MAP}$, is obtained by maximizing the posterior distribution, $\pi(\theta|t)$, as follows

$$\begin{aligned}
\hat{\theta}^{MAP} &= \arg\max_{\theta} \pi(\theta|t) \\
&= \arg\max_{\theta} \prod_{i=1}^{n} f(t_i|\theta)\pi(\theta) \\
&= \arg\max_{\theta} \left(\log(\pi(\theta)) + \sum_{i=1}^{n} \log(f(t_i|\theta)) \right).
\end{aligned} \tag{2.28}$$

It is more suitable to derive the $\arg\max_\theta$ of the log of the MAP function. Two recent studies proposed by Ramos, Louzada and Ramos (2016), Ramos, Louzada and Ramos (2018) investigate MAP estimators for the Nakagami-m distribution parameters. The authors showed that the MAP estimates have a closed-form expression and can be rewritten as a bias corrected MLE.

2.5.2 Bayesian nonparametric inference

Traditionally, the key idea of the Bayesian nonparametric methods is to obtain inference on an unknown distribution (density estimation). In this context, we can highlight two arguments for using this approach: (1) allowing model flexibility and (2) avoiding making incorrect model specifications when there is uncertainty about some inherent characteristics of a distribution and such distributional assumptions are untestable, (e.g., multimodality, skewness, and heavy tails). Before presenting the formal definitions on this topic, firstly, we will present a simple comparative example between the parametric and nonparametric Bayesian approaches, for ease of understanding.

In the parametric Bayesian approach, given a specific dataset, a family of models is chosen. In the gamma family of distributions case, denoted by $\mathscr{G}(a,b)$, this relates to shape and scale parameters. We then specify prior distributions for such parameters. This procedure induces a prior on the family of distributions. This same procedure is done in the nonparametric case, however, instead of having a finite number of parameters that identify the family of distributions for the data, we will have an infinite number of parameters. Since we have limitations in dealing with infinite quantities, both analytically and computationally, then in practice these models must be truncated to have a possibly large but finite number of parameters (CARVALHO, 2016). Technically this means that nonparametric Bayesian modeling requires a prior distribution on the infinite-dimensional parameters, similarly, this is equivalent to placing a prior distribution on the space of all distribution functions. This leads to models on function spaces.

Example 2.5.2. Let $z_i \in \Omega$, $z_i \mid F \sim \text{iid } F$, $F \in \mathscr{F}^*$, $\mathscr{F}^* = \{\mathscr{G}(a,b) : \ a,b \in \mathbb{R}^+\}$. In this parametric specification a prior on $\mathscr{F}^*$ is equivalent to a prior on (a,b). Nevertheless, $\mathscr{F}^*$ is very small relative to $\mathscr{F}$ = {the entirety distributions on Ω}; see Figure 3 (KOTTAS, 2018). Therefore, the nonparametric Bayesian approach requires priors on much larger subsets of $\mathscr{F}$ (infinite-dimensional spaces).

In order to introduce a few new terms and keep consistent with the current literature, in this section we summarize basic definitions and notations. We will mostly follow the definitions and notations from Escobar and West (1995), Müller and Quintana (2004), Kvam and Vidakovic (2007) .

The Dirichlet process (DP), with precursors in the work of Freedman *et al.* (1963), was formally developed by Ferguson (1973), Ferguson *et al.* (1974). The idea is to use the DP as a

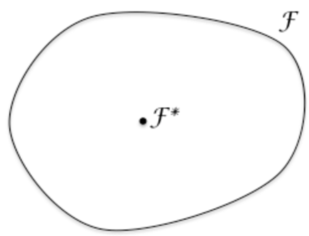

Figure 3 – Diagram for a simple comparison between the parametric and nonparametric Bayesian approaches.

prior on a set of probability measures in order to consider certain nonparametric problems from a Bayesian approach. It is the first prior developed for spaces of distribution functions. The DP is a probability measure (distribution) on the space of probability measures (distributions) defined on a common probability space χ. Hence, a realization of DP is a random distribution function.

Formally, the DP is described by two parameters: Q_0, a specific probability measure on χ (or equivalently, G_0 a specified distribution function on χ); c, a positive scalar parameter.

Definition 2.5.3. [Ferguson (1973)] The DP generates random probability measures (random distributions) Q on χ such that, for any finite partition $B_1,\ldots,B_k$ of χ,

$$(Q(B_1),\ldots,Q(B_K)) \sim \mathscr{D}\,(cQ_0(B_1),\ldots,cQ_0(B_K)),$$

where $Q(B_i)$ (a random variable) and $Q_0(B_i)$ (a constant) denote the probability of set B_i under Q and Q_0, respectively. Thus, for any B,

$$Q(B) \sim \mathscr{B}(cQ_0(B), c(1-Q_0(B)))$$

and

$$\mathbb{E}(Q(B)) = Q_0(B).$$

Let $\mathscr{B}$ denote the beta distribution and let $\mathscr{D}$ denote the Dirichlet distribution. The probability measure Q_0 plays the role of the center of the DP, while c can be viewed as a precision parameter. Large c implies small variability of DP with respect to its center Q_0. The above can be expressed in terms of distribution function, rather than in terms of probabilities. For $B=(-\infty,x]$, the probability $Q(B)=Q((-\infty,x])=G(x)$ is a distribution function. As a result, we can write $G(x)\sim\mathscr{B}(cG_0(x),c(1-G_0(x)))$ and $\mathbb{E}(G(x))=G_0(x)$; $Var(G(x))=G_0(x)(1-G_0(x))/c+1$. The notation $G\sim$ DP(cG_0) indicates that the DP prior is placed on the distribution G.

An often useful constructive definition of a DP is given by Sethuraman (1994).

Definition 2.5.4. Let $U_i\sim\mathscr{B}(1,c)$, $i=1,2,\ldots$, and $V_i\sim G_0$, $i=1,2,\ldots$, be two independent sequences of iid random variables. Define weights $\omega_i=U_i$ and $\omega_i=U_i\prod_{j=1}^{i-1}(1-U_j)$, $i>1$. Then,

$$G=\textstyle\sum_{k=1}^{\infty}\omega_k\delta(V_k)\sim \mathrm{DP}(c,G_0),$$

where $\delta(V_k)$ is a point mass at V_k.

However, limiting the prior to discrete distributions may not be appropriate for some applications. A simple extension to remove the constraint of discrete measures is to use a convoluted DP: $X|F \sim F$; $F(x) = \int f(x|\theta)dG(\theta)$; $G \sim DP(c, G_0)$. This model is called DPM, because the mixing is done by the DP. Posterior inference for DPM models is based on MCMC posterior simulation. Efficient MCMC simulation for general DPM models is discussed, among others, in Escobar (1994), Escobar and West (1995).

CHAPTER 3

CLASSICAL INFERENCE FOR A REPAIRABLE SYSTEM UNDER INDEPENDENT COMPETING RISKS WITH REPARAMETRIZED PLP

This chapter presents classical inference results for a specific scenario where we analyze failure data from a single repairable system with independent competing risks. The results presented here will be used in the following chapter for the purpose of comparing them with the results of the objective Bayesian approach.

3.1 Motivating situation

This section begins with an illustration of a single repairable system data under recurrent competing risks. We reanalyze the data extracted from Somboonsavatdee and Sen (2015b), which consist of warranty claims of a automobile fleet (see Table 1). The data describe the cumulative mileage at failure along with the associated cause of failure. The cars were observed from 0 to a maximum of 3000 mileage (truncated time). The authors state that this data set is consistent with the framework of a single repairable system under competing risks and they assume that the failures for each mode are according to a NHPP with a intensity function PLP.

The recurrence of failure causes can be seen in Figure 4. The histogram (Figure 5) simply shows the number of failures in each 250-mileage interval. Overall, there were 99 failures attributed to failure cause 1, 118 to failure cause 2 and 155 attributed to failure cause 3.

Figure 5 suggests that as the mileage increases, the number of failures is decreasing. We suppose that framework is that of a competing risks. This is done by considering a multicomponent system connected in series.

Table 1 – Automobile fleet warranty data.

mileage	cause	mileage	cause	mileage	cause	mileage	cause	mileage	cause
1.21	1	30.00	1	174.00	1	514.00	2	879.00	3
1.48	2	38.00	3	175.00	2	515.00	3	887.00	1
1.53	2	40.05	2	182.00	2	519.00	2	888.26	3
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1.66	2	40.35	3	189.00	2	520.00	1	888.91	3

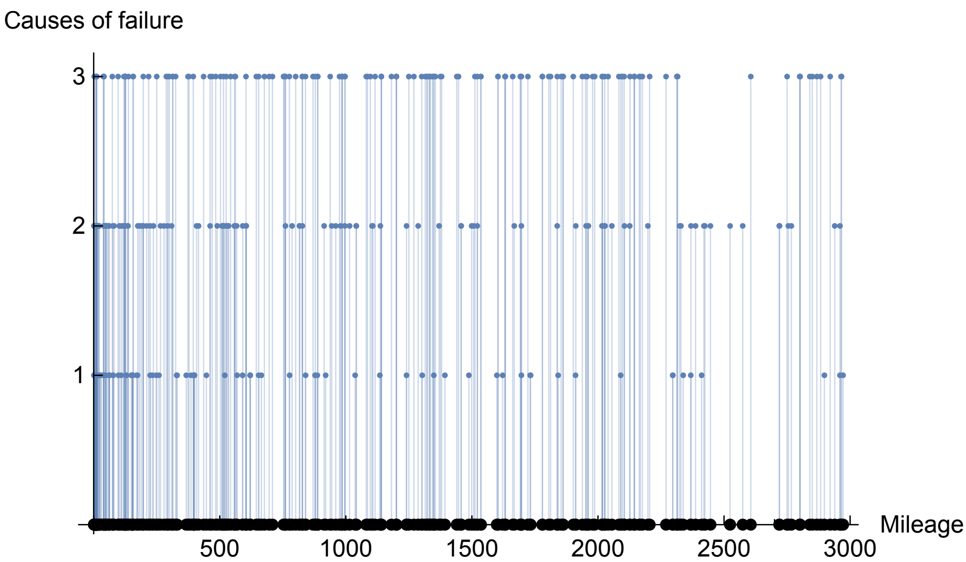

Figure 4 – Recurrence of failures by cause and mileage. The black points on the x-axis indicate the system (fleet) failures.

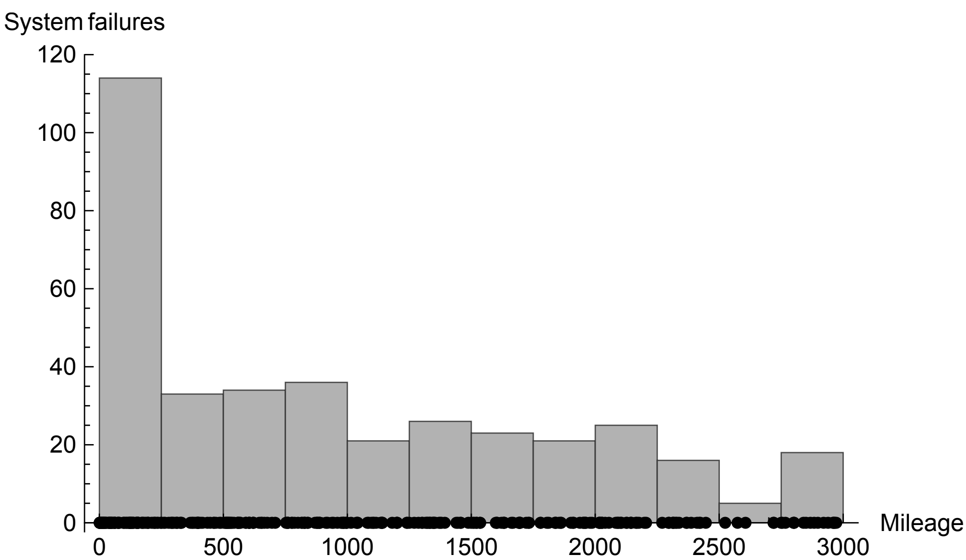

Figure 5 – Histogram of car failures. The black points on the x-axis indicate the system (fleet) failures; the mileage axis is divided into 250-mileage interval.

3.2 Modeling minimal repair under competing risks

Following Somboonsavatdee and Sen (2015b), let us assume the scenario where a minimal repair is performed at each failure thereby resulting in the NHPP as the suitable model. Moreover, we emphasize the use of the parametric form PLP for the intensity function of the NHPP. In particular, we present throughout this thesis a modified version of the PLP model, as given in Section 2.3.1.

In addition, we will use the definitions given in Section 2.3.1, with some modifications in the notation of quantities of interest.

Notation:

j: cause of failure index, $j = 1, \ldots, p$.

i: recurrent event index, $i = 1, \ldots, n_j$.

$(0, T]$: observation period of the data truncated at T.

$\delta_i = j$: indicator of cause of failure, the j-th associated failure cause with i-th failure time.

$n_j = \sum_{i=1}^{n} \mathbb{I}(\delta_i = j)$: number of failures due to cause-j.

$n = \sum_{j=1}^{p} n_j$: number of system failures.

$\lambda_j(t)$: cause-specific intensity function associated to cause-j.

$\Lambda_j(T)$: cause-specific cumulative intensity associated to cause-j.

$\lambda(t) = \sum_{j=1}^{p} \lambda_j(t)$: intensity function of the system.

$\Lambda(T) = \sum_{j=1}^{p} \Lambda_j(T)$: system cumulative intensity function.

$\lambda_j(t) = \beta_j \alpha_j t^{\beta_j - 1} T^{-\beta_j}$: reparametrized PLP.

$\mathbb{E}[N_j(T)] = \left(\frac{T}{\mu_j}\right)^{\beta_j}$: cause-specific mean function.

We assume that repair times are negligible. Then, the observable quantities for a system under competing risks and minimal repair can be summarized by Figure 6.

3.2.1 Maximum likelihood estimation for the reparametrized model

Recalling that causes of failure act independently and they are mutually exclusive, the likelihood contribution from the j-th cause is

$$L_j(\theta|\mathbf{t}, \delta) = \prod_{i=1}^{n} [\lambda_j(t_i)]^{\mathbb{I}(\delta_i = j)} \exp[-\Lambda_j(T)], \tag{3.1}$$

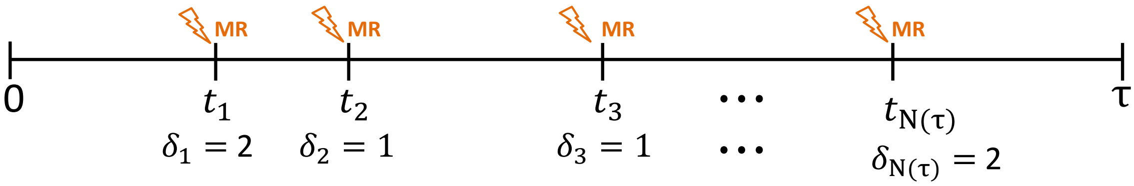

Figure 6 – Occurrence of minimal repair (MR) at each failure, for a repairable system under competing risks with two causes of failure (or components).

where $\mathbb{I}(\delta_i = j)$ represents the indicator function of the cause j associated with $i-$th time of failure and $\theta = (\mu_1, \beta_1, \mu_2, \beta_2, \ldots, \mu_p, \beta_p)$. Thus, the full likelihood function $L(\theta|\mathbf{t}, \delta) = \prod_{j=1}^{p} L_j(\theta|\mathbf{t}, \delta)$ can be written as

$$L(\theta|\mathbf{t}, \delta) = \prod_{i=1}^{n}\prod_{j=1}^{p}[\lambda_j(t_i)]^{\mathbb{I}(\delta_i=j)} \exp\left[-\sum_{j=1}^{p}\Lambda_j(T)\right]. \tag{3.2}$$

To achieve the MLEs, firstly we considered the scenario where $p = 2$ and the shape parameters are equal, i.e., $\beta_1 = \beta_2$. Then, we obtain the estimators with no restriction in the shape parameters, i.e., $\beta_1 \neq \beta_2$. Finally, we extend our analysis to p causes. Besides, hereafter we will assume that the system is observed until time T and the model PLP is reparametrized in terms of β_j and

$$\alpha_j = \mathbb{E}[N_j(T)] = (T/\mu_j)^{\beta_j}. \tag{3.3}$$

MLEs for case $p = 2$ with $\beta_1 = \beta_2$

Denote by β the common value of β_1 and β_2. The full likelihood function considering the proposed reparametrization is given by

$$\begin{aligned} L(\theta|\mathbf{t}, \delta) &= \beta^n \left(T\alpha_1^{-\frac{1}{\beta}}\right)^{-n_1\beta} \left(T\alpha_2^{-\frac{1}{\beta}}\right)^{-n_2\beta} \times \\ &\quad \left[\prod^{n_1} t_i \prod^{n_2} t_i\right]^{\beta-1} e^{-\alpha_1-\alpha_2} \\ &\propto \gamma(\beta|n+1, n/\hat{\beta}) \prod_{j=1}^{2}\gamma(\alpha_j|n_j+1, 1), \end{aligned} \tag{3.4}$$

where $\prod^{n_j} t_i = \prod_{i=1}^{n} t_i^{\mathbb{I}(\delta_i=j)}$, $\sum_{i=1}^{n}\mathbb{I}(\delta_i = j) = n_j$, $n = \sum_{j=1}^{p} n_j$, $\hat{\beta} = n/\sum_{i=1}^{n}\log(T/t_i)$ is the MLE of β and $\theta = (\beta, \alpha_1, \alpha_2)$. The factorization in (3.4) implies that β and α_j are orthogonal. A detailed explanation to derive the likelihood (3.4) is described in Appendix A.

The log-likelihood $\ell(\theta|\mathbf{t}, \delta) = \log(L(\theta|\mathbf{t}, \delta)$ is given by

$$\begin{aligned} \ell(\theta|\mathbf{t}, \delta) &= c + n\log(\beta) - n\beta\log(T) + \beta\sum_{i=1}^{n}\log(t_i) + \\ &\quad n_1\log(\alpha_1) + n_2\log(\alpha_2) - \alpha_1 - \alpha_2, \end{aligned} \tag{3.5}$$

where $c = -\log\left(\prod^{n_1} t_i \prod^{n_2} t_i\right)$.

By maximizing (3.5), the MLEs of β and α_j are found to be

$$\hat{\beta}^{MLE} = \frac{n}{\sum_{i=1}^{n} \log(T/t_i)}, \tag{3.6}$$

and

$$\hat{\alpha}_j^{MLE} = n_j. \tag{3.7}$$

Note that the MLEs only exist if $n_j \geq 1$ for $j = 1,2$. Note also that $\hat{\alpha}^{MLE}$ is unbiased. Hereafter, it is assumed that at least one failure from each cause should occur.

The Fisher information matrix is also presented as it is an important metric that has many properties. It is also used to derive asymptotic intervals and it is a key tool in objective Bayesian framework to achieve objective priors. The Fisher information matrix of θ is defined as

$$H(\theta) = -\mathbb{E}_\theta\left[\frac{\partial^2 \ell(\theta|\mathbf{t},\delta)}{\partial\theta_j \partial\theta_{j'}}\right], \quad j, j' = 1, \ldots, p. \tag{3.8}$$

To compute the Fisher information matrix, note that the partial derivatives are

$$\frac{\partial \ell(\theta|\mathbf{t},\delta)}{\partial\beta} = n/\beta - n\log(T) + \sum_{i=1}^{n} \log(t_i),$$

$$\frac{\partial \ell(\theta|\mathbf{t},\delta)}{\partial\alpha_j} = \frac{n_j}{\alpha_j} - 1,$$

$$\frac{\partial^2 \ell(\theta|\mathbf{t},\delta)}{\partial\beta^2} = -n\beta^{-2},$$

$$\frac{\partial^2 \ell(\theta|\mathbf{t},\delta)}{\partial\alpha_j^2} = \frac{-n_j}{\alpha_j^2} \quad \text{and}$$

$$\frac{\partial^2 \ell(\theta|\mathbf{t},\delta)}{\partial\beta\partial\alpha_1} = \frac{\partial^2 \ell(\theta|\mathbf{t},\delta)}{\partial\beta\partial\alpha_2} = \frac{\partial^2 \ell(\theta|\mathbf{t},\delta)}{\partial\alpha_1\partial\alpha_2} = 0$$

(since we are considering time truncation, both $n = \sum_{j=1}^{p} n_j$ and the n_js are random). Hence, the expectation of the second derivatives above are given by

$$-\mathbb{E}\left[\frac{\partial^2 \ell(\theta|\mathbf{t},\delta)}{\partial\beta^2}\right] = \frac{(\alpha_1+\alpha_2)}{\beta^2},$$

and

$$-\mathbb{E}\left[\frac{\partial^2 \ell(\theta|\mathbf{t},\delta)}{\partial\alpha_j^2}\right] = \frac{1}{\alpha_j}.$$

It follows that the Fisher information matrix is diagonal and given by

$$H(\theta) = \begin{bmatrix} (\alpha_1+\alpha_2)/\beta^2 & 0 & 0 \\ 0 & 1/\alpha_1 & 0 \\ 0 & 0 & 1/\alpha_2 \end{bmatrix}, \tag{3.9}$$

and the asymptotic covariance matrix of $\hat{\theta}$ is $Var(\hat{\theta}) \approx diag(\beta^2/(\alpha_1+\alpha_2), \alpha_1, \alpha_2)$.

To build confidence intervals for θ we can consider the asymptotic theory where $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N_p(0, H^{-1}(\theta))$ when $n \to \infty$. The Delta method may be necessary when there is interest in a function that depends on θ. For the latter, the reader is referred to Somboonsavatdee and Sen (2015b) for a detailed discussion. See Bain and Englehardt (1991) and Lehmann (2004) for more technical details.

MLEs for case $p = 2$ with $\beta_1 \neq \beta_2$

For the case of the different shape parameters, $\beta_1 \neq \beta_2$, the likelihood function is given by

$$\begin{aligned} L(\theta|\mathbf{t},\delta) &= c\left[\beta_1^{n_1} e^{-n_1\beta_1/\hat{\beta}_1}\right]\left[\beta_2^{n_2} e^{-n_2\beta_2/\hat{\beta}_2}\right] \times \\ &\quad \left[e^{-\alpha_1}\alpha_1^{n_1}\right]\left[e^{-\alpha_2}\alpha_2^{n_2}\right] \\ &\propto \gamma(\beta_1|n_1+1, n_1/\hat{\beta}_1)\gamma(\beta_2|n_2+1, n_2/\hat{\beta}_2)\times \\ &\quad \gamma(\alpha_1|n_1+1,1)\gamma(\alpha_2|n_2+1,1) \\ &\propto \prod_{j=1}^{2}\gamma(\beta_j|n_j+1, n_j/\hat{\beta}_j)\gamma(\alpha_j|n_j+1,1), \end{aligned} \tag{3.10}$$

where $\theta = (\beta_1, \beta_2, \alpha_1, \alpha_2)$, $c = (\prod^{n_1} t_i \prod^{n_2} t_i)^{-1}$ and $\hat{\beta}_j$ is the MLE of β_j given below.

The MLEs have explicit solutions

$$\hat{\beta}_j^{MLE} = \frac{n_j}{\sum_{i=1}^{n}\log(T/t_i)\mathbb{I}(\delta_i = j)}, \tag{3.11}$$

and

$$\hat{\alpha}_j^{MLE} = n_j \quad j = 1, 2. \tag{3.12}$$

Since $\mathbb{E}[N_j(T)] = \alpha_j$, for $j = 1, 2$, the Fisher information matrix is

$$H(\theta) = \begin{bmatrix} \alpha_1\beta_1^{-2} & 0 & 0 & 0 \\ 0 & \alpha_2\beta_2^{-2} & 0 & 0 \\ 0 & 0 & \alpha_1^{-1} & 0 \\ 0 & 0 & 0 & \alpha_2^{-1} \end{bmatrix}. \tag{3.13}$$

MLEs for case $p > 2$, at least two β_js are different

The likelihood function in this case is

$$L(\theta|\mathbf{t},\delta) \propto \prod_{j=1}^{p} \gamma(\beta_j|n_j+1, n_j/\hat{\beta}_j)\gamma(\alpha_j|n_j+1,1), \tag{3.14}$$

where $\theta = (\beta_1,\ldots,\beta_p,\alpha_1,\ldots,\alpha_p)$.

The MLEs for p causes are the same as (3.11) and (3.12), but for $j = 1,\ldots p$. The Fisher information matrix is

$$H(\theta) = \begin{bmatrix} \alpha_1\beta_1^{-2} & 0 & \ldots & \ldots & \ldots & 0 \\ 0 & \ddots & 0 & \ldots & \ldots & 0 \\ \vdots & 0 & \alpha_p\beta_p^{-2} & 0 & \ldots & 0 \\ \vdots & \vdots & 0 & \alpha_1^{-1} & 0 & 0 \\ \vdots & \vdots & \vdots & 0 & \ddots & 0 \\ 0 & \ldots & 0 & \ldots & 0 & \alpha_p^{-1} \end{bmatrix}. \tag{3.15}$$

The MLEs are asymptotically normally distributed with a multivarite normal distribution given by

$$\hat{\theta}^{MLE} \sim N_p[\theta, H^{-1}(\theta)] \text{ for } n \to \infty,$$

where $\hat{\theta}^{MLE} = (\hat{\beta}_1^{MLE},\ldots,\hat{\beta}_p^{MLE},\hat{\alpha}_1^{MLE},\ldots,\hat{\alpha}_p^{MLE})$.

3.3 A real data application

We apply our proposed methodology to real-data example described in Section 3.1, under the assumption of minimal repair performed at each failure and considering the situation where such a system was monitored from 0 to a maximum of 3000 mileage (time truncated case).

Following Somboonsavatdee and Sen (2015b), we assessed the adequacy of the PLP for each cause of failure with the help of a Duane plot; see Rigdon and Basu (2000). Figure 7 shows plots of the logarithm of number of failures $N_j(t)$ against the logarithm of accumulated mileage at failure, for $j = 1,2,3$. Since the three plots exhibit reasonable linearity, they suggest that the PLP model is adequate.

The MLEs are presented in Table 2, along with the standard deviation (SD) and 95% confidence interval (CI) .

The results suggest that the reliability of all components are improving, since $\hat{\beta}_1 = 0.329$, $\hat{\beta}_2 = 0.451$ and $\hat{\beta}_3 = 0.749$ are smaller than 1. We remark that this information can provide important insights to the maintenance crew.

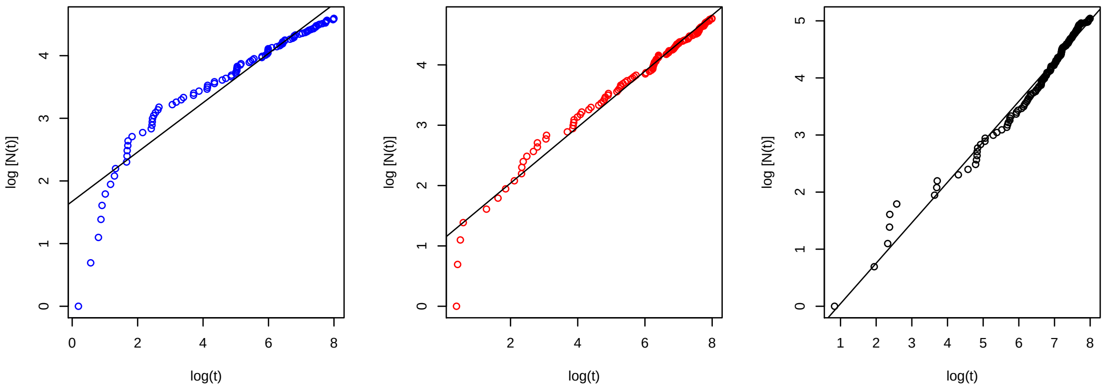

Figure 7 – Duane plots: cause 1 (blue), cause 2 (red), cause 3 (black).

Table 2 – Point and interval estimates for the PLP parameters for the warranty claims data.

Parameter	MLE	SD	95% CI
β_1	0.329	0.033	(0.264, 0.393)
β_2	0.451	0.042	(0.370, 0.533)
β_3	0.749	0.060	(0.631, 0.867)
α_1	99	9.950	(79.498, 118.502)
α_2	118	10.863	(96.709, 139.291)
α_3	155	12.450	(130.598, 179.402)

3.4 Conclusions

In this chapter, we provided a study of classical statistical inference for failure data arising from a single repairable system under independent competing risks (a complex multicomponent system where the components are connected in series). At failures, we use the minimal repair concept. In particular, we presented the parametric framework of a PLP reparametrized in terms of orthogonal parameters. Here, our major concern lies on the inference of the model parameters.

The insights gained from this study may be of assistance to evaluate progress in developing the reliability of a repairable system. Technically, reliability growth is an iterative process of testing, identifying problems, analyzing their causes, designing solutions, and implementing them in the system being tested (RIGDON; BASU, 2000; HØYLAND; RAUSAND, 2009; TODINOV, 2015).

It is worth mentioning that some results of this chapter will be used in Chapter 4, and they will be used in an extensive simulation study in the mentioned chapter.

CHAPTER

4

OBJECTIVE BAYESIAN INFERENCE FOR A REPAIRABLE SYSTEM SUBJECT TO COMPETING RISKS

In this chapter, we discuss inferential procedures based on an objective Bayesian approach for analyzing failures from a single repairable system under independent competing risks. We examined the scenario where a minimal repair is performed at each failure, thereby resulting in that each failure mode appropriately follows a power-law intensity. Besides, we derived two objective priors known as Jeffreys prior and reference prior. Moreover, posterior distributions based on these priors will be obtained in order to find advantageous properties. We prove that these posterior distributions are proper and are also matching priors. In addition, in some cases, unbiased Bayesian estimators of simple closed-form expressions are derived. To the best of our knowledge, no studies have investigated the referred setting based on objective Bayesian reasoning. As Somboonsavatdee and Sen (2015b) state: *"... However, no systematic treatment of statistical inference of PLP subject to multiple failure modes has been documented in the literature."*. Therefore, the main message of this part is that one can contribute with the advantages of the objective Bayesian philosophy in order to fill this gap.

4.1 Motivating data

This section begins with a real example of a single repairable system data under a recurrent competing risks framework in order to motivate our proposed methodology.

Table 3 shows failure times and causes for a sugarcane harvester during a crop. This machine harvests an average of 20 tons of sugarcane per hour, and its malfunction can lead to significant losses. It can fail either due to malfunction of electrical components, the engine or the elevator, which are denoted as cause 1, 2 and 3, respectively, in Table 3. There are 10, 24 and 14

Table 3 – Failure data for a sugarcane harvester.

Time	Cause	Time	Cause	Time	Cause	Time	Cause
4.987	1	7.374	1	15.716	1	15.850	2
20.776	2	27.476	3	29.913	1	42.747	1
47.774	2	52.722	2	58.501	2	65.258	1
71.590	2	79.108	2	79.688	1	79.794	3
80.886	3	85.526	2	91.878	2	93.541	3
94.209	3	96.234	2	101.606	3	103.567	2
117.981	2	120.442	1	120.769	3	123.322	3
124.158	2	126.097	2	137.071	2	142.037	3
150.342	2	150.467	2	161.743	2	161.950	2
162.399	3	185.381	1	193.435	3	205.935	1
206.310	2	210.767	3	212.982	2	216.284	2
219.019	2	222.831	2	233.826	3	234.641	3

failures for each of these causes. During the harvest time that consists of 256 days, the machine operates on a 7×24 regime. Therefore, we assume that each repair is minimal (i.e. it leaves the machine at precisely the same condition it was in before it failed) and that data collection is time truncated at $T = 256$ days. Figure 8 shows the recurrent failure process of the machine. Note that this system had very few failures. It seems possible that this occurred due to a highly

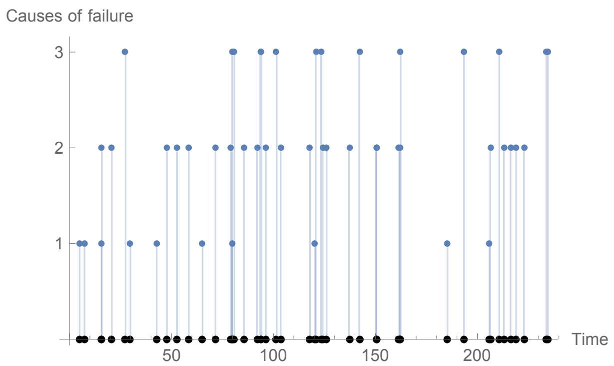

Figure 8 – Recurrence of failures by cause and time. The black points on the x-axis indicate the system (sugarcane harvester) failures.

reliable system. In this case, one knows that the frequentist paradigm has serious drawbacks with small sample sizes. For this reason, we propose an objective Bayesian approach in order to obtain efficient estimates for model parameters even for small sample sizes.

4.1.1 Objective Bayesian inference for the model

This subsection will rely on the same scenarios (the two cases: equal and unequal shape parameters) described in Chapter 3, as well as notation, MLEs and the Fisher information matrix. We will provide the Jeffreys and overall reference prior distributions for the parameters of the model for both these cases aforementioned. Moreover, we include a discussion about good properties of the resulting posterior distributions related to unbiased Bayesian estimators and accurate credibility intervals. Furthermore, we use the following notation conventions: $\hat{\theta}^{MAP} = (\hat{\alpha}_j^{MAP}, \hat{\beta}_j^{MAP})$ for the MAP estimators and $\hat{\theta}^{MEAN} = (\hat{\alpha}_j^{MEAN}, \hat{\beta}_j^{MEAN})$ for the posterior mean estimators, $j = 1, 2, \ldots, p$.

Case $p = 2$, $\beta_1 = \beta_2$

Jeffreys Prior

Let β be the common value of β_1 and β_2. It follows from (3.9) that the Jeffreys prior distribution is given by

$$\pi^J(\theta) \propto \frac{1}{\beta}\sqrt{\frac{\alpha_1+\alpha_2}{\alpha_1\alpha_2}}. \tag{4.1}$$

Proposition 4.1.1. The Jeffreys prior (4.1) is a matching prior for β.

Proof. Let β be the parameter of interest and denote by $\zeta = (\alpha_1, \alpha_2)$ the nuisance parameter. First, since the information matrix is diagonal, the Jeffreys prior can be written in the form (2.27). □

The joint posterior distribution for β and α_j produced by Jeffreys prior is proportional to the product of the likelihood function (3.4) and the prior distribution (4.1), resulting in

$$\begin{aligned}\pi^J(\beta,\alpha_1,\alpha_2|\mathbf{t},\delta) \propto \sqrt{\alpha_1+\alpha_2}\left[\beta^{n-1}e^{-n\beta/\hat{\beta}}\right] \times \\ \left[\alpha_1^{n_1-1/2}e^{-\alpha_1}\right]\left[\alpha_2^{n_2-1/2}e^{-\alpha_2}\right].\end{aligned} \tag{4.2}$$

The posterior (4.2) does not have a closed-form and this implies that it may be improper, which is undesirable. Moreover, to obtain the necessary credible intervals, we would have to resort to Monte Carlo methods.

Reference prior

From (3.9), note that, if in Theorem 2.5.1 we take $f_1(\beta) = \beta^{-2}$, $g_1(\alpha_1,\alpha_2) = (\alpha_1+\alpha_2)$, $f_2(\alpha_1) = \alpha_1^{-1}$, $g_2(\beta,\alpha_2) = 1$, $f_3(\alpha_2) = \alpha_2^{-1}$ and $g_3(\beta,\alpha_1) = 1$, the overall reference prior is

$$\pi^R(\beta,\alpha) \propto \frac{1}{\beta}\sqrt{\frac{1}{\alpha_1\alpha_2}}. \tag{4.3}$$

Proposition 4.1.2. The overall reference prior (4.3) is a matching prior for all the parameters.

Proof. If β is the parameter of interest and $\zeta = (\alpha_1, \alpha_2)$, then the proof is analogous to that for the Jeffreys' prior above but considering $g(\zeta) = \frac{1}{\sqrt{(\alpha_1+\alpha_2)\alpha_1\alpha_2}}$. If α_1 is the parameter of interest and $\zeta = (\beta, \alpha_2)$ are the nuisance parameters. Then, as $H_{\alpha_1,\alpha_1}(\alpha_1,\zeta) = \dfrac{1}{\alpha_1}$ and $g(\zeta) = \frac{1}{\beta\sqrt{\alpha_2}}$. Hence, the overall reference prior (4.3) can be written in the form (2.27). The case that α_2 is the parameter of interest is similar. □

Thus, the joint posterior distribution for $(\beta, \alpha_1, \alpha_2)$ is proportional to the product of the likelihood function (3.4) and the prior distribution (4.3), resulting in

$$\pi^R(\beta, \alpha_1, \alpha_2 | \mathbf{t}, \delta) \propto \left[\beta^{n-1} e^{-n\beta/\hat{\beta}}\right] \left[\alpha_1^{n_1 - 1/2} e^{-\alpha_1}\right] \times \left[\alpha_2^{n_2 - 1/2} e^{-\alpha_2}\right], \tag{4.4}$$

which can be recognized as

$$\pi^R(\beta, \alpha_1, \alpha_2 | \mathbf{t}, \delta) \propto \gamma(\beta | n, n/\hat{\beta}) \prod_{j=1}^{2} \gamma(\alpha_j | n_j + 1/2, 1), \tag{4.5}$$

which is the product of independent gamma densities and $\hat{\beta}$ as given earlier in (3.6). Note that if there is at least one failure for each cause, this posterior is proper. Then, the marginal posterior distributions are given by

$$\pi^R(\beta | \mathbf{t}, \delta) \propto \left[\beta^{n-1} e^{-n\beta/\hat{\beta}}\right] \sim \gamma(\beta | n, n/\hat{\beta}) \tag{4.6}$$

and

$$\pi^R(\alpha_j | \mathbf{t}, \delta) \propto \left[\alpha_j^{n_j - 1/2} e^{-\alpha_j}\right] \sim \gamma(\alpha_j | n_j + 1/2, 1), \tag{4.7}$$

for $j = 1, 2$.

From (4.6) and (4.7), one can obtain the Bayesian estimators and the CIs for β and α_js, respectively. Regarding β, we calculated the MAP estimator using (2.28) and (4.6), which is given by

$$\hat{\beta}^{MAP} = \left(\frac{n-1}{n}\right) \hat{\beta}^{MLE}. \tag{4.8}$$

A detailed explanation to derive (4.8) is discussed in Appendix A. Next, the Bayesian estimator using the posterior mean for β is given by

$$\mathbb{E}(\beta | \mathbf{t}, \delta) = \hat{\beta}^{MEAN} = \hat{\beta}^{MLE}. \tag{4.9}$$

Note that the estimators (4.8) and (4.9) present simple closed-form expressions, but just the MAP estimator is unbiased.

Now, with respect to α_j (for $j = 1, 2$), the Bayesian estimators are obtained from (4.7). Using also (2.28), we calculated the MAP estimator, which is given by

$$\hat{\alpha}_j^{MAP} = n_j - \frac{1}{2}. \tag{4.10}$$

The posterior mean is given by

$$\mathbb{E}(\alpha_j | \mathbf{t}, \delta) = \hat{\alpha}_j^{MEAN} = n_j + \frac{1}{2}. \tag{4.11}$$

In this case, the estimators (4.10) and (4.11) also present simple closed-form expressions, however, are biased.

Credibility intervals for β and α_js can also be obtained directly from (4.6) and (4.7), respectively, considering the 2.5% and 97.5% percentiles posteriors. Note that, as was proved in Proposition 4.1.2, the marginal posterior intervals have accurate frequentist coverage for all parameters.

Case p=2, $\beta_1 \neq \beta_2$

Jeffreys Prior

It follows from (3.13) that the Jeffreys prior distribution is given by

$$\pi^J(\theta) \propto \frac{1}{\beta_1\beta_2}. \tag{4.12}$$

Proposition 4.1.3. The Jeffreys prior (4.12) is a matching prior for β_1 and β_2.

Proof. Let β_1 be the parameter of interest and $\lambda = (\beta_2, \alpha_1, \alpha_2)$ be the nuisance parameters. Since the information matrix is diagonal and $H_{\beta_1,\beta_1}(\beta_1, \lambda) = \dfrac{\alpha_1}{\beta_1^2}$, taking $g(\lambda) = \frac{1}{\beta_2\sqrt{\alpha_1}}$, (4.12) can be written in the form (2.27). The case when β_2 is the parameter of interest is similar. □

The joint posterior distribution obtained using the Jeffreys prior (4.12) and the likelihood function (3.10) is given by

$$\pi^J(\theta|\mathbf{t},\delta) \propto \prod_{j=1}^{2} \gamma(\beta_j|n_j, n_j/\hat{\beta}_j)\gamma(\alpha_j|n_j+1, 1), \tag{4.13}$$

where $\hat{\beta}_j$ is the same as given in (3.11). From this it follows that the marginal posterior distributions for β_j and α_j $(j = 1, 2)$ are given by

$$\pi^J(\beta_j|\mathbf{t},\delta) \propto \gamma(\beta_j|n_j, n_j/\hat{\beta}_j) \tag{4.14}$$

and

$$\pi^J(\alpha_j|\mathbf{t},\delta) \propto \gamma(\alpha_j|n_j+1, 1). \tag{4.15}$$

Thus, the Bayesian estimators for β_j using the MAP and the posterior mean are given, respectively, by

$$\hat{\beta}_j^{MAP} = \left(\frac{n_j - 1}{n_j}\right)\hat{\beta}_j^{MLE} \tag{4.16}$$

and

$$\mathbb{E}(\beta_j|\mathbf{t},\delta) = \hat{\beta}_j^{MEAN} = \hat{\beta}_j^{MLE}. \tag{4.17}$$

Note that these estimators have simple closed-form expressions and only the MAP estimator is unbiased.

In terms of α_j, the Bayesian estimators are given by

$$\hat{\alpha}_j^{MAP} = n_j \tag{4.18}$$

and

$$\mathbb{E}(\alpha_j|\mathbf{t},\delta) = \hat{\alpha}_j^{MEAN} = n_j + 1. \tag{4.19}$$

In this case, the estimators (4.18) and (4.19) present closed-form expressions and only the MAP estimator is unbiased.

Credibility intervals for β_js and α_js are obtained directly from (4.14) and (4.15), respectively, considering the 2.5% and 97.5% percentiles posteriors. Note that, as was proved in Proposition 4.1.3, the marginal posterior intervals have accurate frequentist coverage only for β_js.

Reference prior

On the other hand, considering (3.13) and Theorem 2.5.1, the overall prior distribution is

$$\pi^R(\theta) \propto \frac{1}{\beta_1\beta_2}\frac{1}{\sqrt{\alpha_1\alpha_2}}. \tag{4.20}$$

Proposition 4.1.4. The overall reference prior (4.20) is a matching prior for all parameters.

Proof. The proofs for β_1 and β_2 follow the same steps as in the proof of Proposition 4.1.3. The cases of α_1 and of α_2 follow directly from Proposition 4.1.2. □

From the product of the likelihood function (3.10) and the overall prior distribution (4.20), the joint reference posterior distribution for θ is given by

$$\pi^R(\theta|\mathbf{t},\delta) \propto \prod_{j=1}^{2}\gamma(\beta_j|n_j,n_j/\hat{\beta}_j)\gamma(\alpha_j|n_j+\tfrac{1}{2},1). \tag{4.21}$$

Thus, the marginal posterior distributions for β_j and α_j are given, respectively, by

$$\pi^R(\beta_j|\mathbf{t},\delta) \propto \gamma(\beta_j|n_j,n_j/\hat{\beta}_j), \tag{4.22}$$

and

$$\pi^R(\alpha_j|\mathbf{t},\delta) \propto \gamma(\alpha_j|n_j+\tfrac{1}{2},1). \tag{4.23}$$

Therefore, the Bayesian estimators for β_j using the MAP and the posterior mean are given, respectively, by

$$\hat{\beta}_j^{MAP} = \left(\frac{n_j-1}{n_j}\right)\hat{\beta}_j^{MLE} \tag{4.24}$$

and

$$\mathbb{E}(\beta_j|\mathbf{t},\delta) = \hat{\beta}_j^{MEAN} = \hat{\beta}_j^{MLE}. \tag{4.25}$$

Note that these estimators have closed-form expressions and only MAP estimator is unbiased.

With regard to α_j, the Bayesian estimators are given by

$$\hat{\alpha}_j^{MAP} = n_j - \frac{1}{2} \tag{4.26}$$

and

$$\mathbb{E}(\alpha_j|\mathbf{t},\delta) = \hat{\alpha}_j^{MEAN} = n_j + \frac{1}{2}. \quad (4.27)$$

In this case, the estimators (4.26) and (4.27) present simple closed-form expressions, however, are biased.

Regarding CIs for β_js and α_js using the overall reference prior (4.20), note that, as was proved in Proposition 4.1.4, the marginal posterior intervals have accurate frequentist coverage for all parameters.

The following is a brief description of the extended model for $p > 2$. As the results are the same as those we just presented, we will omit some comments.

Case $p > 2$, at least two β_js are different

Jeffreys prior

From (3.15) the Jeffreys' prior is given by

$$\pi^J(\theta) \propto \prod_{j=1}^{p} \frac{1}{\beta_j}, \quad (4.28)$$

which gives the posterior distribution

$$\pi^J(\theta|\mathbf{t},\delta) \propto \prod_{j=1}^{p} \gamma(\beta_j|n_j, n_j/\hat{\beta}_j)\gamma(\alpha_j|n_j+1,1). \quad (4.29)$$

To prove that (4.28) is a matching prior only for $\beta_j, j = 1,\ldots,p$, we can consider the same steps of the proof of Proposition 4.1.3.

The Bayes estimators for β_j using the MAP and the posterior mean are the same as given in (4.16) and (4.17). In terms of α_j, the Bayesian estimators are the same as given in (4.18) and (4.19).

Reference prior

The overall reference prior using Theorem 2.5.1 and (3.15) is given by

$$\pi^R(\theta) \propto \prod_{j=1}^{p} \beta_j^{-1}\alpha_j^{-1/2}. \quad (4.30)$$

Proposition 4.1.5. The overall reference prior (4.30) is a matching prior for all parameters.

Proof. The proof is essentially the same as that of Proposition 4.1.4. □

The joint reference posterior distribution for θ, produced by the overall prior distribution, is proportional to the product of (3.14) and (4.30), resulting in

$$\pi^R(\theta|\mathbf{t},\delta) \propto \prod_{j=1}^{p} \gamma(\beta_j|n_j, n_j/\hat{\beta}_j)\gamma(\alpha_j|n_j+\tfrac{1}{2},1). \quad (4.31)$$

As discussed above, the Bayes estimators for β_j using the MAP and the posterior mean are the same as given in (4.24) and (4.25). With respect to α_j, the Bayesian estimators are the same as given in (4.26) and (4.27).

4.2 Simulation study

In this section, we present a simulation study to compare the Bayes estimators and the MLEs (from Chapter 3). The simulation design proposed is consistent with the setup described so far. The R software has been used for simulation (R Core Team, 2016).

We assume that there is a single system observed on the fixed time interval $(0,T]$, with T=20. For ease of presentation, we consider two and three independent causes of failure with distinct parameters for each cause $\theta = (\alpha_j, \beta_j)$, for $j = 1,2,3$. The parameter values were selected in order to obtain different sample sizes. Besides, among the many parameter choices made, we provide details of the findings for five scenarios due to the lack of space:

- Scenario 1: $\beta_1 = 1.2, \alpha_1 = 18.21, \beta_2 = 0.6, \alpha_2 = 16.28$;
- Scenario 2: $\beta_1 = 0.25, \alpha_1 = 4.23, \beta_2 = 1.12, \alpha_2 = 6.75$;
- Scenario 3: $\beta_1 = 1.5, \alpha_1 = 44.72, \beta_2 = 0.7, \alpha_2 = 16.28$, $\beta_3 = 0.6, \alpha_3 = 14.50$;
- Scenario 4: $\beta_1 = 1.5, \alpha_1 = 26.83, \beta_2 = 0.9, \alpha_2 = 44.46$, $\beta_3 = 1.2, \alpha_3 = 29.12$;
- Scenario 5: $\beta_1 = 1.2, \alpha_1 = 7.28, \beta_2 = 0.7, \alpha_2 = 16.28$, $\beta_3 = 1.0, \alpha_3 = 8.6$.

For each setup of parameters, we obtain the mean number of failures (18.21, 16.28), (4.23, 6.75), (44.72, 16.28, 14.50), (26.83, 44.46, 29.12) and (7.28, 16.28, 8.6), respectively. In the first two scenarios, the setting is a two-component system where each component supplies almost the same mean number of failures. On the other hand, the last three scenarios refer to the case where the failures of one component predominate more than other system's components. It is worth noting that the obtained results are similar for other parameter combinations and can be extended to more causes, i.e., $p > 3$.

Using the fact that the causes are independent and also using the known results from the literature about NHPPs (RIGDON; BASU, 2000), in each Monte Carlo replication the failure times and indicators of the cause of failure were generated as shown in Algorithm 1.

We used three criteria to evaluate the estimators' behavior. (i) The bias given by $Bias_{\hat{\theta}_i} = \frac{1}{M}\sum_{j=1}^{M} \hat{\theta}_{i,j} - \theta_i$; (ii) the mean absolute error given by $MAE_{\hat{\theta}_i} = \frac{1}{M}\sum_{j=1}^{M} |\hat{\theta}_{i,j} - \theta_i|$ and (iii) the mean square error given by $MSE_{\hat{\theta}_i} = \sum_{j=1}^{M} \frac{(\hat{\theta}_{i,j}-\theta_i)^2}{M}$, where M is the number of estimates (i.e. the Monte Carlo size), which we take $M = 50,000$ throughout the section, and $\theta = (\theta_1, \ldots, \theta_p)$ is the vector of parameters.

Algorithm 1 – Algorithm for generating random data from a single system with PLPs under competing risks.

1: For each cause of failure, generate random numbers $n_j \sim Poisson(\Lambda_j)$ $(j = 1, \ldots, p)$.

2: For each cause of failure, let the failure times be $t_{1,j}, \ldots, t_{n_j,j}$, where $t_{i,j} = T\, U_{i,j}^{1/\beta_j}$ and $U_{1,j}, \ldots, U_{n_j,j}$ are the order statistics of a size n_j random sample from the standard uniform distribution.

3: Finally, to obtain the data in the form (t_i, δ_i), let the t_is be the set of ordered failure times and set δ_i equal to j according to the corresponding cause of failure (i.e., set $\delta_i = 1$ if $t_i = t_{h,1}$ for some h or $\delta_i = j$ depending on the cause of failure).

Additionally, for the objective Bayesian credible intervals and the asymptotic maximum likelihood based confidence intervals for β_1, β_2, α_1 and α_2, we computed the 95% coverage probability, denoted by $CP_{95\%}$. Good estimators should have Bias, MAE and MSE close to zero and adequate intervals should be short while showing $CP_{95\%}$ close to 0.95. The Bias and MSE are widely used to measure the performance evaluation (SOMBOONSAVATDEE; SEN, 2015b), while some recent studies have put forward considering the MAE as well; see for instance, Willmott and Matsuura (2005).

The Bayes MAP estimators for β_j and α_j using the Jeffreys prior (MAP Jeffreys) are computed from (4.16) and (4.18). While the Bayes MAP estimators obtained using the reference prior (MAP Reference) are computed from (4.24) and (4.26). In these cases, no MCMC was needed as the estimators have closed-form expressions. Additionally, we consider MCMC techniques, as well as the posterior mean since this estimator is obtained when considering squared error loss or Kullback-Leibler as loss in the risk. Hence, the estimators called Mean Jeffreys and Mean Reference were obtained by considering the posterior mean of θ using the Jeffreys and reference prior and MCMC techniques. Since the marginal distributions for the parameters are gamma distributions, we can sample directly from the target distribution. In this sense, for each simulated data set, 10,000 iterations were performed using MCMC methods. These values were used to compute the posterior means and the credibility intervals obtained by the 2.5% and 97.5% quantile values. We also considered confidence intervals based on the CMLE (2.12) in which the asymptotic variances are estimated from the Fisher information matrix, similar to what is done to obtain the maximum likelihood (ML) intervals.

Tables 4 - 7 present the results of our simulation study. It is important to point out that, in Table 4 the estimators for β_j and $\alpha_j, j = 1, 2$, are the same for the Jeffreys prior and the CMLE. However, they are different when computing the credibility/confidence intervals and the CPs. The CIs using the Jeffreys prior are obtained from the quantile of gamma distributions, the CIs of the CMLE are obtained from the asymptotic theory as discussed above.

We note, from Tables 4 and 6, that the MAP Jeffreys estimator returned improved

Table 4 – The Bias, MAE, MSE from the estimates considering different values of θ, with $M = 50,000$ simulated samples using the different estimation methods.

		Scenario 1			Scenario 2		
Parameter	Method	Bias	MAE	MSE	Bias	MAE	MSE
β_1	MLE	0.0760	0.2493	0.1180	0.1013	0.1587	0.2204
	MAP Jef.	0.0011	0.2366	0.0977	0.0009	0.1170	0.0692
	MAP Ref.	0.0011	0.2366	0.0977	0.0009	0.1170	0.0692
	Mean Jef.	0.0760	0.2493	0.1180	0.1013	0.1587	0.2198
	Mean Ref.	0.1509	0.2758	0.1515	0.2017	0.2299	0.4808
β_2	MLE	0.0608	0.1616	0.0581	0.2337	0.4556	0.8856
	MAP Jef.	-0.0006	0.1484	0.0418	0.0012	0.3822	0.3913
	MAP Ref.	-0.0006	0.1484	0.0418	0.0012	0.3822	0.3913
	Mean Jef.	0.0608	0.1616	0.0581	0.2337	0.4557	0.8883
	Mean Ref.	0.1222	0.1887	0.0857	0.4662	0.5928	1.7400
α_1	MLE	-0.0020	3.4207	18.3446	0.2874	1.4991	3.6142
	MAP Jef.	-0.0020	3.4207	18.3446	0.2874	1.4991	3.6142
	MAP Ref.	-0.5020	3.4612	18.5966	-0.2126	1.5464	3.5769
	Mean Jef.	0.9982	3.4861	19.3437	1.2873	1.7248	5.1892
	Mean Ref.	0.4979	3.4346	18.5932	0.7875	1.5644	4.1525
α_2	MLE	-0.0263	2.7671	12.1555	0.0522	2.0908	6.8925
	MAP Jef.	-0.0263	2.7671	12.1555	0.0522	2.0908	6.8925
	MAP Ref.	-0.5263	2.8380	12.4318	-0.4478	2.1612	7.0903
	Mean Jef.	0.9735	2.8404	13.1046	1.0522	2.2040	7.9984
	Mean Ref.	0.4739	2.7958	12.3834	0.5521	2.1226	7.1953

estimates when compared with the MLE across the different scenarios, since the Bias, MAE and MSE are closer to zero. This Bayes estimator also returned better results for both β_j and α_j parameters when compared to the posterior mean of the Jeffreys and reference priors obtained through MCMC, specially for α_j. These results are expected since the marginal posterior distribution with Jeffreys prior is given by $\gamma(\alpha_j|n_j+1,1)$, therefore posterior mean will be $\alpha_j^{MEAN} = n_j+1$ and the expected bias will be 1 since n_j is an unbiased estimator for α_j. Considering the marginal posterior distributions with reference prior we have the marginals given by $\gamma(\alpha_j|n_j+\frac{1}{2},1)$, then the expected bias will be 0.5 using the posterior mean and -0.5 when we considered the MAP as an estimator. In fact, from our simulation study we observed that the bias for α_js were close to theoretical bias.

Regarding the coverage probabilities in Tables 5 and 7, we note that for both the MLE and the CMLE CIs, the CPs are far from the assumed levels, especially for the scale parameters α_j. On the other hand, the CIs of the Bayes MAP estimators using the reference prior returned accurate coverage probabilities when compared with the other estimators. In some cases the values for the CPs are very close to the assumed levels and some of the other methods may have a closer value for a specific value of θ. However, taking the average of the CPs the reference prior was the one that returned closer values of 0.95 when compared with the other estimators. It is worth mentioning that the MAP Jeffreys estimator returned CIs very close to the ones obtained

Table 5 – Coverage probabilities from the estimates considering different scenarios, with $M = 50,000$ simulated samples and different estimation methods.

θ	Method	Scenario 1	Scenario 2
β_1	MLE	0.9532	0.9531
	CMLE	0.9279	0.8268
	MAP Jef.	0.9490	0.9501
	MAP Ref.	0.9490	0.9501
	Mean Jef.	0.9489	0.9499
	Mean Ref.	0.9394	0.9071
β_2	MLE	0.9545	0.9556
	CMLE	0.9161	0.8796
	MAP Jef.	0.9501	0.9496
	MAP Ref.	0.9501	0.9496
	Mean Jef.	0.9497	0.9490
	Mean Ref.	0.9373	0.9222
α_1	MLE	0.9366	0.9958
	CMLE	0.9366	0.9958
	MAP Jef.	0.9535	0.9697
	MAP Ref.	0.9515	0.9697
	Mean Jef.	0.9526	0.9693
	Mean Ref.	0.9460	0.9697
α_2	MLE	0.9476	0.9075
	CMLE	0.9476	0.9075
	MAP Jef.	0.9402	0.9551
	MAP Ref.	0.9565	0.9518
	Mean Jef.	0.9408	0.9551
	Mean Ref.	0.9564	0.9518

using the reference prior. In short, considering the MAP with the reference prior we obtained accurate confidence intervals but the obtained punctual estimates have a systematic Bias which is undesirable. On the other hand, considering the MAP Jeffreys estimator, we obtained unbiased estimators for the parameters but only with matching priors for β_j. Since our simulation study showed that the CIs of the MAP using the Jeffreys prior are also satisfactory and taking into the account the bias and CPs, we suggest the use of the closed-form MAP estimator with the Jeffreys prior to perform inference on the unknown parameters of the PLP model with competing risks. Notice that while we suggested using the MAP Jeffreys estimator, some concerns must be taken into account when selecting one of the two proposed MAP estimators. If the analyst is only interested in interval estimates, the MAP reference estimator should be used.

Table 6 – The Bias, MAE, MSE from the estimates considering different values of $\boldsymbol{\theta}$, with $M = 50,000$ simulated samples using the different estimation methods.

		Scenario 3			Scenario 4			Scenario 5		
Parameter	Method	Bias	MAE	MSE	Bias	MAE	MSE	Bias	MAE	MSE
β_1	MLE	0.0336	0.1857	0.0576	0.0608	0.2466	0.1065	0.2440	0.4834	1.0720
	MAP Jef.	-0.0015	0.1822	0.0539	0.0001	0.2385	0.0945	-0.0015	0.4060	0.4609
	MAP Ref.	-0.0015	0.1822	0.0539	0.0001	0.2385	0.0945	-0.0015	0.4060	0.4609
	Mean Jef.	0.0336	0.1857	0.0576	0.0607	0.2466	0.1065	0.2441	0.4834	1.0724
	Mean Ref.	0.0687	0.1932	0.0640	0.1214	0.2641	0.1266	0.4897	0.6272	2.1310
β_2	MLE	0.0499	0.1542	0.0467	0.0202	0.1120	0.0210	0.0488	0.1546	0.0476
	MAP Jef.	0.0003	0.1452	0.0375	-0.0010	0.1099	0.0196	-0.0008	0.1458	0.0379
	MAP Ref.	0.0003	0.1452	0.0375	-0.0010	0.1099	0.0196	-0.0008	0.1458	0.0379
	Mean Jef.	0.0500	0.1542	0.0467	0.0202	0.1120	0.0210	0.0488	0.1546	0.0476
	Mean Ref.	0.0996	0.1728	0.0620	0.0414	0.1167	0.0233	0.0984	0.1730	0.0636
β_3	MLE	0.1482	0.2766	0.3388	0.0398	0.1814	0.0562	0.1610	0.3521	0.7219
	MAP Jef.	0.0013	0.2268	0.1453	-0.0017	0.1765	0.0508	-0.0009	0.3062	0.2893
	MAP Ref.	0.0013	0.2268	0.1453	-0.0017	0.1765	0.0508	-0.0009	0.3062	0.2893
	Mean Jef.	0.1482	0.2766	0.3389	0.0398	0.1814	0.0562	0.1610	0.3521	0.7261
	Mean Ref.	0.2951	0.3669	0.6755	0.0814	0.1921	0.0654	0.3230	0.4387	1.4535
α_1	MLE	-0.0185	5.3300	44.5825	-0.0165	4.1286	26.7974	0.0129	2.1272	7.0329
	MAP Jef.	-0.0185	5.3300	44.5825	-0.0165	4.1286	26.7974	0.0129	2.1272	7.0329
	MAP Ref.	-0.5185	5.3533	44.8510	-0.5165	4.1673	27.0639	-0.4871	2.1838	7.2699
	Mean Jef.	0.9810	5.3706	45.5466	0.9833	4.1766	27.7611	1.0130	2.2279	8.0599
	Mean Ref.	0.4818	5.3328	44.8099	0.4834	4.1401	27.0339	0.5129	2.1353	7.2958
α_2	MLE	0.0311	3.2289	16.4019	0.0038	5.3473	44.9001	0.0053	3.2180	16.3156
	MAP Jef.	0.0311	3.2289	16.4019	0.0038	5.3473	44.9001	0.0053	3.2180	16.3156
	MAP Ref.	-0.4689	3.2646	16.6207	-0.4962	5.3570	45.1463	-0.4947	3.2567	16.5603
	Mean Jef.	1.0309	3.3004	17.4651	1.0034	5.3891	45.9142	1.0054	3.2868	17.3306
	Mean Ref.	0.5311	3.2361	16.6842	0.5039	5.3435	45.1633	0.5052	3.2236	16.5738
α_3	MLE	0.0609	2.0181	6.2823	-0.0095	4.4370	31.1655	0.0100	2.3392	8.5058
	MAP Jef.	0.0609	2.0181	6.2823	-0.0095	4.4370	31.1655	0.0100	2.3392	8.5058
	MAP Ref.	-0.4391	2.0528	6.4714	-0.5095	4.4830	31.4250	-0.4900	2.3747	8.7457
	Mean Jef.	1.0610	2.1348	7.4049	0.9905	4.4846	32.1481	1.0098	2.4301	9.5257
	Mean Ref.	0.5610	2.0141	6.5949	0.4910	4.4572	31.4110	0.5101	2.3301	8.7659

Table 7 – Coverage probabilities from the estimates considering different scenarios, with $M = 50,000$ simulated samples and different estimation methods.

θ	Method	Scenario 3	Scenario 4	Scenario 5
β_1	MLE	0.9503	0.9522	0.9559
	CMLE	0.9409	0.9359	0.8843
	MAP Jef.	0.9493	0.9494	0.9504
	MAP Ref.	0.9493	0.9494	0.9504
	Mean Jef.	0.9491	0.9492	0.9502
	Mean Ref.	0.9471	0.9429	0.9247
β_2	MLE	0.9558	0.9502	0.9536
	CMLE	0.9274	0.9399	0.9261
	MAP Jef.	0.9524	0.9489	0.9497
	MAP Ref.	0.9524	0.9489	0.9497
	Mean Jef.	0.9521	0.9491	0.9493
	Mean Ref.	0.9411	0.9465	0.9391
β_3	MLE	0.9565	0.9529	0.9549
	CMLE	0.8785	0.9398	0.8962
	MAP Jef.	0.9503	0.9508	0.9498
	MAP Ref.	0.9503	0.9508	0.9498
	Mean Jef.	0.9500	0.9507	0.9493
	Mean Ref.	0.9210	0.9461	0.9285
α_1	MLE	0.9417	0.9413	0.9290
	CMLE	0.9417	0.9413	0.9290
	MAP Jef.	0.9491	0.9466	0.9456
	MAP Ref.	0.9484	0.9466	0.9456
	Mean Jef.	0.9500	0.9505	0.9506
	Mean Ref.	0.9499	0.9472	0.9535
α_2	MLE	0.9225	0.9387	0.9232
	CMLE	0.9225	0.9387	0.9232
	MAP Jef.	0.9533	0.9478	0.9529
	MAP Ref.	0.9533	0.9478	0.9529
	Mean Jef.	0.9505	0.9478	0.9504
	Mean Ref.	0.9511	0.9494	0.9505
α_3	MLE	0.8976	0.9444	0.9247
	CMLE	0.8976	0.9444	0.9247
	MAP Jef.	0.9630	0.9407	0.9633
	MAP Ref.	0.9528	0.9516	0.9442
	Mean Jef.	0.9630	0.9450	0.9626
	Mean Ref.	0.9481	0.9510	0.9442

4.3 A real data application

We illustrate the proposed methodology by applying it to the real-world reliability dataset described in Section 4.1. This dataset is related to a sugarcane harvester machine. There were 10 failures attributed to cause 1, 24 to cause 2 and 14 to cause 3. Figure 9 shows a histogram of the evolution of the number of failures of the sugarcane harvester in an observation window of approximately 8 months. The harvester processes 20 tons of cane per hour, and with the purpose of avoiding losses, a corrective action (minimal repair) is performed at each failure to make the machine operational again. Although the machine is used intensively, very few failures occur.

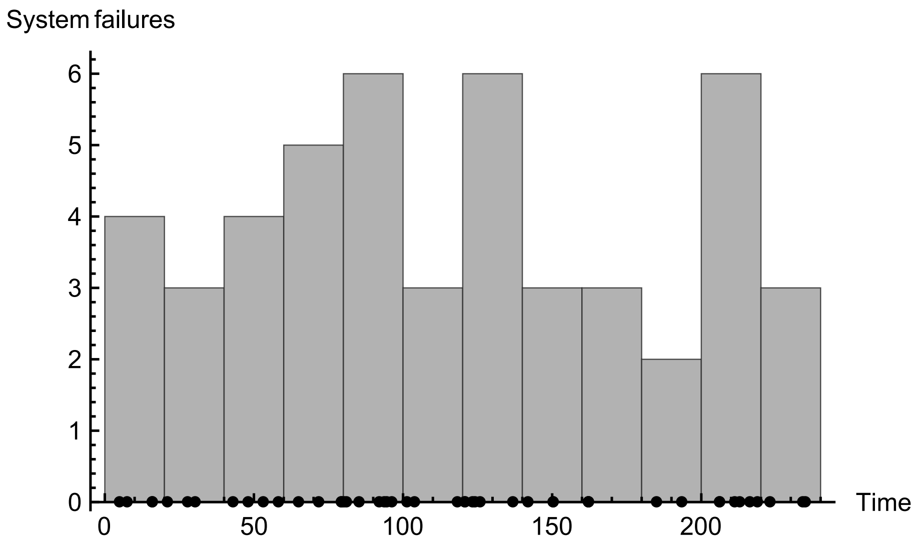

Figure 9 – Histogram of sugarcane harvester failures. The black points on the x-axis indicates the system (sugarcane harvester) failures; the time axis is divided into 20-day intervals.

Following Somboonsavatdee and Sen (2015b), we assessed the adequacy of the PLP for each cause of failure with the help of a Duane plot; see Rigdon and Basu (2000). Figure 10 shows plots of the logarithm of number of failures $N_j(t)$ against the logarithm of accumulated mileage at failure, for $j = 1,2,3$. Since the three plots exhibit reasonable linearity, they suggest that the PLP model is adequate.

We summarize here the results concerning the objective Bayesian inference using the Jeffreys prior (4.28). Table 8 shows the Bayes estimates, as well as the corresponding marginal posterior SD and CI. We remark that this posterior summary does not require a stochastic simulation to be obtained. For instance, the credible intervals can be calculated directly from the posterior quantiles from (4.29).

Table 8 – Bayesian estimates for sugarcane harvester machine dataset.

Parameter	Bayes	SD	CI (95%)
β_1	0.499	0.175	[0.266 ; 0.947]
β_2	1.038	0.221	[0.694 ; 1.558]
β_3	1.220	0.351	[0.718 ; 2.086]
α_1	10.000	3.317	[5.491 ; 18.390]
α_2	24.000	5.000	[16.179 ; 35.710]
α_3	14.000	3.873	[8.395 ; 23.490]

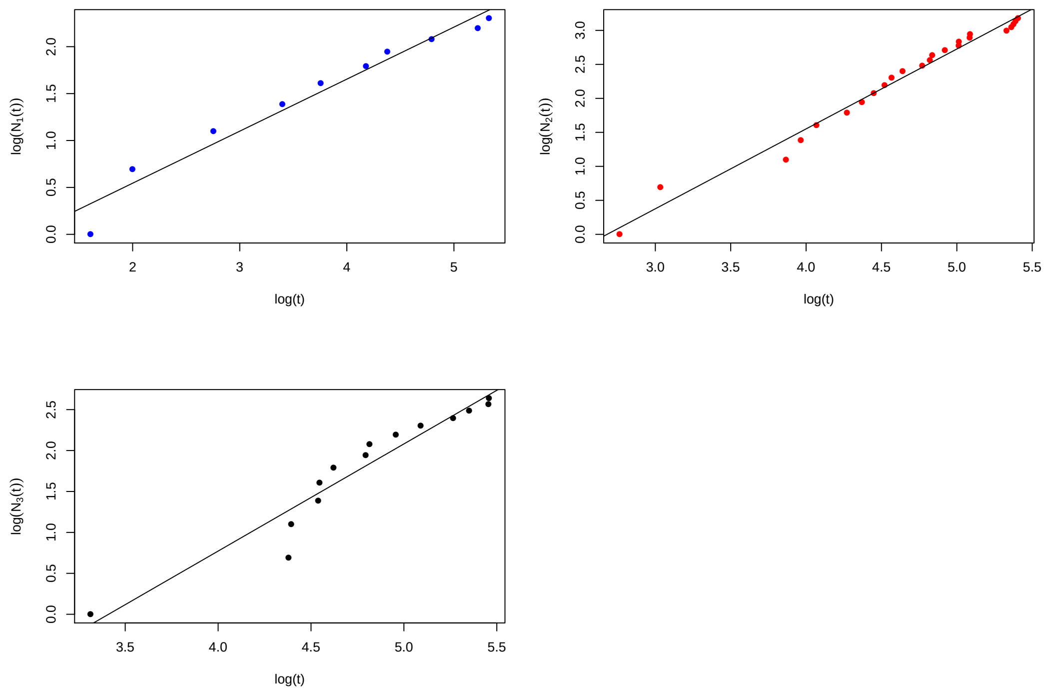

Figure 10 – Duane plots: Cause of failure 1 depicted by blue points; Cause of failure 2 depicted by red points; Cause of failure 3 depicted by black points.

Considering the point estimates of Table 8, the results suggest that the reliability of the electrical components (cause 1) is improving over time, since the corresponding $\hat{\beta}_1 = 0.499 < 1$, while the reliability of the elevator is decreasing ($\hat{\beta}_3 = 1.220 > 1$). The engine shows an intermediate behavior since $\hat{\beta}_2 = 1.038$ is slightly greater than one. We remark that this information can provide important insights into the maintenance crew.

4.4 Conclusions

In this chapter, we discussed inferential procedures based on an objective Bayesian approach for analyzing failures from a repairable system under competing risks. Besides, we assume that the multiple causes of failure are stochastically independent of each other. The competing risks approach may be advantageous in the engineering field because it may lead to a better understanding of the various causes of failure of a system, and hence design strategies to improve the overall reliability.

Since the Bayesian methods have been not well explored in this context, we proposed Bayesian estimators for the parameters of the failure intensity assuming the PLP model. We derived two objective priors known as Jeffreys prior and reference prior. The obtained posteriors are proper distributions and have interesting properties, such as one-to-one invariance and consistent marginalization. In addition, the Bayes MAP estimates have closed-form expressions. The MAP Jeffreys estimator is naturally unbiased for all the parameters, while the MAP Reference returned marginal posterior intervals with accurate coverage in the frequentist sense. A simulation study suggests confirming the theoretical results showing that the MAP Jeffreys estimator behaves consistently better than the other estimators across the different scenarios, since the Bias, MAE and MSE are closer to zero. On the other hand, the CIs of the Bayes MAP estimators using the reference prior returned improved coverage probabilities when compared with the other estimators. From a simulation study, we observed that the CIs of the MAP using the Jeffreys prior are also satisfactory and taking into the account the Bias and CPs, we suggest using such an estimator to perform inference on the parameters of the model.

The importance and originality of this study come from using orthogonal reparametrization that enabled us to obtain posterior distributions that return accurate estimates concerning bias and credibility intervals for the parameters of the PLP intensities under competing risks. The proposed methodology was applied for an original dataset regarding failures of a sugarcane harvester classified according to three possible causes. Since the data contain few failures, classical CIs based on the asymptotic ML theory could be inadequate in this case. Although in the application we considered three causes of failure, the proposed methodology can be applied to multiple causes.

Our findings can be applied in real data sets based on following assumptions. To consider a single multi-component repairable system whose components (causes of failures) are arranged in series. The causes of failures must be independent of each other. The recurrent data structure (failure history) should be based on cause-specific intensity functions with PLP. The data sampling scheme (system observational period) is the time truncated case. The dataset should be structured as Tables 1 and 3, containing: time at failure and the exact cause of failure (two columns is enough). It is worth pointing out that it is not feasible to consider the issue of dependence between components in the context proposed in this chapter because of the inestimability of the frailty parameter in the lack of multiple systems; see Somboonsavatdee and Sen (2015b).

There may be some interesting extensions of this work. One can consider that more significant challenge could be to investigate data where the cause of failure is unknown (masked failure data). Another possible extension of this work is to consider different truncation times.

CHAPTER 5

MULTIPLE REPAIRABLE SYSTEMS UNDER DEPENDENT COMPETING RISKS WITH NONPARAMETRIC FRAILTY

In this chapter, we propose a model to analyze data arising from multiple repairable systems under the presence of dependent competing risks. It is known that the dependence effect in this scenario influences the estimates of the model parameters. Hence, under the assumption that the cause-specific intensities follow a PLP, we propose a shared frailty model to incorporate the dependence among the cause-specific recurrent processes. Our approach allows us to carry out an individual posterior analysis of the quantities of interest, i.e., we estimate the interest parameters of the PLP (our main focus) separately from nuisance parameters of frailty distribution (variance). Regarding PLP parameters, we consider noninformative priors so that the posterior distributions are proper. With respect to frailty, our proposal avoids making incorrect specifications of the frailty distribution when there is uncertainty about some inherent characteristics of distribution. Thus, we considered a nonparametric approach to model the frailty density using a DPM prior. Besides, a particular novelty is our hybrid MCMC algorithm composed by HMC algorithm and Gibbs sampler. This algorithm was built for computing the posterior estimates with respect to the frailty distribution. Moreover, our model can provide information on unobserved heterogeneity among systems. This is meaningful information even when "identical" systems are considered.

5.1 Introduction

As mentioned before, the nonparametric frailty distribution takes into account a flexible class of distributions. In particular, we used the DPM model to describe the frailty distribution due to its flexibility in modeling unknown distributions. Many approaches on nonparametric Bayesian models have been explored in the literature related to reliability, for instance, Salinas-Torres, Pereira and Tiwari (2002) give a comprehensive theoretical exposition on Bayesian nonparametric estimation for survival functions arising from observed failures of a competing risks model (or a series system). Li *et al.* (2014) provide a flexible Bayesian nonparametric framework to modeling recurrent events in a repairable system to test the minimal repair assumption. Bayesian nonparametric inference for NHPPs is considered by Kuo and Ghosh (1997), who employed several classes of nonparametric priors. As mentioned above, the idea of our approach is to apply a Bayesian nonparametric prior (i.e., DPM prior) to modeling uncertainty in the distribution of shared frailty. Although this model has infinite parameters, due to the infinite mixture model, it is a flexible mixture, parsimonious and simple to sample. We chose the stick-breaking representation of the DP prior (SETHURAMAN, 1994), because of a simple implementation to build the algorithm. To obtain the posterior distribution, we created a hybrid MCMC algorithm (KALLI; GRIFFIN; WALKER, 2011), using the Gibbs sampler (SMITH; ROBERTS, 1993) and the HMC method (NEAL *et al.*, 2011). It is important to point out that no studies have been found which explore the use of DPM for frailty density in the context of multiple repairable systems under the action of dependent competing risks.

This research highlights the importance of modeling the dependence structure among competing causes of failure by using a more flexible distribution for unknown frailty density in order to provide good estimates of the model parameters. As stated before, our primary inference goal is to estimate PLP parameters. To this end, firstly, we model the dependence effect with shared frailty, and secondly, we consider the frailty distribution nonparametrically using a DPM. Regarding frailty, the advantage is that one obtains more flexibility at the level of density estimation and providing insights in terms of heterogeneity among systems.

5.2 Multiple repairable systems subject to multiple causes of failure

Here, we highlight the used notations in the multivariate counting process context. Hereafter, random variables are denoted by capital letters (e.g., Z_j, N_{jq}), while their realizations are denoted by the lowercase (e.g., z_j, n_{jq}).

Consider a sample of m identical systems in which each system is under the action of K different types of recurrent causes of failure. Let $N_{jq}(t) = \sum_{i=1}^{n_j} \mathbb{I}(\delta_{ji} = q)$ be the cumulative number of type q failures occurring over the interval $[0,t]$ for the j-th system ($j = 1,\ldots,m$; $q = 1,\ldots,K$ and $i = 1,2,\ldots,n_j$), where $\{N_{jq}(t) : t \geq 0\}$ is a counting process. Note that $N_{j\bullet}(t) = \sum_{q=1}^{K} N_{jq}(t)$ represents the cumulative number of failures of system j taking into account all failures arising from all components from the j-th system. Let $N_{\bullet q}(t) = \sum_{j=1}^{m} N_{jq}(t)$ denote the number of failures of cause q for all systems.

Suppose that each system is under observation for all types of events over the same period of time, i.e., $[0, T]$. Thus, let t_{ji}, $i = 1, 2, \ldots, n_j$, be the observed failure times for system j, satisfying $0 < t_{j1} < t_{j2} < \cdots < t_{jn_j} < T$. Besides, denote that $\delta_{ji} = q$ is the failure mode (or component) that caused the system failure. Hence, the observed data is $D_j = \{(t_{ji}, \delta_{ji} = q), i = 1, 2, \ldots, n_j; q = 1, \ldots, K\}$. The complete data is given by $D = (D_1, \ldots, D_m)$.

As mentioned earlier, our focus is mainly on the component level failure process which conforms to a PLP, therefore the cause-specific intensity function that governs the counting process $N_{\bullet q}(T)$, taking into account a orthogonal parametrization in terms of α_q and β_q, is defined as

$$\lambda_q(t) = \beta_q \alpha_q t^{\beta_q - 1} T^{-\beta_q}, \; q = 1, \ldots, K, \tag{5.1}$$

where α_q is the mean function given by

$$\alpha_q = \mathbb{E}\left[N_{\bullet q}(T)\right] = \Lambda_q(T) = \int_0^T \lambda_q(s) ds. \tag{5.2}$$

5.3 The shared frailty model for the PLP

It is worth pointing out that the main quantity of interest in the shared frailty methodology adopted here is the variance of the frailty (although it is considered as a nuisance parameter, because our major interest is to estimate the PLP parameters). This parameter should be estimated using information of multiple systems. Somboonsavatdee and Sen (2015b) state that in the single system setting there are limitations. Therefore, our approach requires multiple systems as presented so far.

We specify the model (2.24) in terms of (5.1) in order to present the likelihood function with a special form. To achieve this purpose, suppose a minimal repair is undertaken at each failure, thus the NHPP is the model of choice. Specifically, the failures from each component follow an NHPP, with PLP intensity function given in (5.1). Furthermore, let us consider that a realization $z_j \sim f_Z$ acts on all the cause-specific intensities (2.24) belonging to the j-th system. Thus, conditioning on the frailty term, the model is expressed as

$$\lambda_q(t|Z_j) = Z_j \beta_q \alpha_q t_{ji}^{\beta_q - 1} T^{-\beta_q} \tag{5.3}$$

and the mean function is given by

$$\Lambda_q(T|Z_j) = Z_j \alpha_q. \tag{5.4}$$

It is important to point out that, hereafter, our analysis relies on the constraint $\bar{\mathbf{Z}} = \frac{1}{m}\sum_{j=1}^{m} Z_j = 1$.

5.3.1 Likelihood function

To simplify notation in this section, we will drop the subscript $\bullet$ and refer to $n_{j\bullet}$ and $n_{\bullet q}$ as n_j and n_q, respectively. The likelihood contribution from the j-th system based on (5.3) is given by

$$L_j(\theta, Z_j|D_j) = \left[\prod_{i=1}^{n_j}\prod_{q=1}^{K}[\lambda_q(t_{ji}|Z_j)]^{\mathbb{I}(\delta_{ji}=q)}\right]\exp\left[-\sum_{q=1}^{K}\Lambda_q(T|Z_j)\right], \tag{5.5}$$

where $\mathbb{I}(\delta_{ji}=q)$ represents the indicator function aforementioned and $\theta = (\beta, \alpha)$ with $\beta = (\beta_1, \ldots, \beta_K)$ and $\alpha = (\alpha_1, \ldots, \alpha_K)$; for $i = 1, \ldots, n_j$; $j = 1, \ldots, m$ and $q = 1, \ldots, K$. Thus, the overall likelihood function is represented by

$$\begin{aligned} L(\theta, \mathbf{Z} \mid D) &= \prod_{j=1}^{m} L_j(\theta, Z_j|D_j) \\ &= c\prod_{j=1}^{m} Z_j^{n_j}\prod_{q=1}^{K}\left[\beta_q^{n_{jq}}\alpha_q^{n_{jq}\beta_q}T^{-n_{jq}\beta_q}\prod_{i=1}^{n_{jq}}\left(t_{ji}^{\beta_q}\right)\right]\exp\left[-Z_j\sum_{q=1}^{K}\alpha_q\right] \\ &\propto \prod_{j=1}^{m} Z_j^{n_j}\prod_{q=1}^{K}\gamma(\beta_q \mid n_q+1, n_q\hat{\beta}_q^{-1})\gamma(\alpha_q \mid n_q+1, m), \end{aligned} \tag{5.6}$$

where $n_j = \sum_{q=1}^{K} n_{jq}$; $n_q = \sum_{j=1}^{m} n_{jq}$; $n_{jq} = \sum_{i=1}^{n_j}\mathbb{I}(\delta_{ji}=q)$; $\prod_{i=1}^{n_{jq}}(\cdot) = \prod_{i=1}^{n_j}(\cdot)^{\mathbb{I}(\delta_{ji}=q)}$; $c = \prod_{j=1}^{m}\prod_{i=1}^{n_{jq}} t_{ji}^{-1}$. In addition,

$$\hat{\beta}_q = n_q / \sum_{j=1}^{m}\sum_{i=1}^{n_{jq}}\log(T/t_{ji}) \tag{5.7}$$

is the MLE for β_q.

As indicated previously, the overall likelihood function (5.6) may be factored as a product of three quantities, as follows:

$$L(\theta, \mathbf{Z} \mid D) = L_1(\mathbf{Z} \mid \mathbf{D})L_2(\beta \mid \mathbf{D})L_3(\alpha \mid \mathbf{D}), \tag{5.8}$$

where $L_1(\mathbf{Z} \mid \mathbf{D}) = \prod_{j=1}^{m} Z_j^{n_j}$; $L_2(\beta \mid \mathbf{D}) = \prod_{q=1}^{k}\gamma(\beta_q \mid n_q+1, n_q\hat{\beta}_q^{-1})$ and $L_3(\alpha \mid \mathbf{D}) = \prod_{q=1}^{k}\gamma(\alpha_q \mid n_q+1, m)$ and it will be used later in our posterior analysis.

5.4 Bayesian analysis

This section, in turn, is divided into two parts. In the first, we present the choice of the prior distributions for β_q and α_q ($q = 1, \ldots, k$) in the PLP model. In this case, we consider a similar approach according to the study of Bar-Lev, Lavi and Reiser (1992). In the second, we discuss a Bayesian nonparametric approach to model the uncertainty about the distribution of shared frailty. As we will see in this section, we can carry out an individual posterior analysis of the quantities of interest due to the orthogonality among α_q and β_q and the assumption that Z_js are stochastically independent of the failure processes λ_qs.

5.4.1 Prior specification for α and β

Selecting an adequate prior distribution using formal rules has been widely discussed in the literature (KASS; WASSERMAN, 1996). In the repairable systems context, Bar-Lev, Lavi and Reiser (1992) considered the following class of prior for the PLP model

$$\pi(\alpha,\beta) \propto \alpha^{-1}\beta^{-\zeta}, \tag{5.9}$$

where $\zeta > 0$ is a known hyperparameter. Following these authors, we apply their main results in the setting of repairable systems under competing risks using the particular parametric formulation of PLP (5.3). Thus, we propose the prior distribution for the referred context as follows:

$$\pi(\alpha,\beta) \propto \prod_{q=1}^{K} \alpha_q^{-1}\beta_q^{-\zeta}. \tag{5.10}$$

This class of prior distributions includes the invariant Jeffreys' prior when $\zeta = 1$. Moreover, it reduces to (5.9) when $q = 1$. Further, we will discuss the chosen value for ζ, and necessary conditions for the obtained posterior to be proper.

Note that, due to (5.8) and the assumption that Z_js are stochastically independent of the failure processes λ_qs, the joint posterior distribution of (5.10) is proper. Note also that, the marginal distributions $\pi(\beta \mid \mathbf{D})$ and $\pi(\alpha \mid \mathbf{D})$ are proper since they are independent, as follows:

$$\pi(\beta \mid \mathbf{D}) = \prod_{q=1}^{k} \gamma\left(\beta_q \mid n_q + 1 - \zeta, n_q\hat{\beta}_q^{-1}\right) \text{ and } \pi(\alpha \mid \mathbf{D}) = \prod_{q=1}^{k} \gamma(\alpha_q \mid n_q, m). \tag{5.11}$$

Since $\pi(\alpha \mid \mathbf{D})$ is the product of independent gamma distributions, then the marginal joint distribution $\pi(\alpha \mid \mathbf{D})$ is proper. Using the same idea, $\pi(\beta \mid \mathbf{D})$ is the product of independent gamma distributions if $n_q > \zeta$ and, therefore, is a proper marginal posterior distribution.

This work adopts the quadratic loss function, hence the Bayes estimator is the posterior mean which has optimality under Kullback-Leibler divergence. It is worth pointing out that, in this chapter, the notation adopted for posterior mean will be $\hat{\alpha}_q^{Bayes}$ and $\hat{\beta}_q^{Bayes}$. Therefore,

$$\begin{aligned} \hat{\alpha}_q^{Bayes} &= \mathbb{E}(\alpha_q \mid \mathbf{D}) = \frac{n_q}{m} \\ \hat{\beta}_q^{Bayes} &= \mathbb{E}(\beta_q \mid \mathbf{D}) = \frac{(n_q + 1 - \zeta)}{n_q}\hat{\beta}_q. \end{aligned} \tag{5.12}$$

Besides the good properties mentioned above, we have that

$$\begin{aligned} &\mathbb{E}\left[\hat{\alpha}_q^{Bayes}\right] = \alpha_q \text{ and} \\ &\mathbb{E}\left[\hat{\beta}_q^{Bayes}\right] = \mathbb{E}\left[\frac{(n_q + 1 - \zeta)}{n_q}\hat{\beta}_q\right] = \beta_q \quad \text{if} \quad \zeta = 2. \end{aligned} \tag{5.13}$$

Therefore, assuming that $\zeta = 2$ we have that both $\hat{\alpha}_q^{Bayes}$ and $\hat{\beta}_q^{Bayes}$ are unbiased estimators for α_q and β_q.

5.4.2 Bayesian nonparametric approach for frailty distribution

This work presents the frailty distribution as an unknown distribution, therefore we will apply the Bayesian nonparametric methodology. Traditionally, the key idea of the Bayesian nonparametric approach is to obtain inference on an unknown distribution function using process priors on the spaces of densities. According to a definition provided by Sethuraman (1994), the nonparametric Bayesian model involves infinitely many parameters. To better understand the technical definition of Bayesian nonparametric models in a broad way, please see Dey, MüIler and Sinha (2012), Antoniak (1974), for example. There are many methods that specify more flexible density such as finite mixtures, DP, DPM, and mixture of Polya trees. Here, we considered DPM for logarithm of the frailty $\mathbf{W} = \log(Z)$, represented by

$$\begin{aligned} W_1, \ldots, W_m &\sim F \\ F &\sim \mathscr{D}(c, F_0), \end{aligned} \tag{5.14}$$

where $\mathscr{D}$ is the DP prior with base distribution F_0; c is the concentration parameter and $\mathbf{W} = (W_1, \ldots, W_m)'$. c can also be interpreted as a precision parameter that indicates how close the F distribution is to the base distribution F_0 (ESCOBAR; WEST, 1995).

Using the stick-breaking representation discussed in Sethuraman (1994), a DPM of Gaussian distribution can be represented as infinite mixtures of Gaussian, which is an extension of the finite mixture model. Therefore, a density function of W can be represented by

$$f_W(W) = f_W(W \mid \Omega) = \sum_{l=1}^{\infty} \rho_l \mathscr{N}(w \mid \mu_l, \tau_l^{-1}), \tag{5.15}$$

where $\mathscr{N}(\cdot \mid \mu, \tau^{-1})$ denotes a normal density function with parameters (μ, τ^{-1}); $\Omega = \{\rho, \mu, \tau\}$ is the infinite-dimensional parameter vector describing the mixture distribution for W; $\rho = \{\rho_l\}_{l=1}^{\infty}$ represents the vector of weights, $\mu = \{\mu_l\}_{l=1}^{\infty}$ is the vector of means and $\tau = \{\tau_l\}_{l=1}^{\infty}$ is the vector of precision, for $l = 1, 2, \ldots$. Note that the density function of Z can be calculated as follows:

$$f_Z(Z) = f_Z(Z \mid \Omega) \;\; = \;\; \sum_{l=1}^{\infty} \rho_l \mathscr{L}\mathscr{N}(z \mid \mu_l, \tau_l^{-1}), \tag{5.16}$$

where $\mathscr{L}\mathscr{N}(\cdot \mid \mu, \tau^{-1})$ denotes log-normal density functions with parameters μ and τ^{-1}. Therefore, Z can be represented as the infinite mixture log-normal. Note that the base distributions of Z and W are a log-normal and a normal distribution, respectively.

Prior specification for Ω

As shown before, Ω represents a collection of all unknown parameters in (5.15) and (5.16). Based on this, we specified a prior distribution for Ω as follows. Firstly, we specify a prior for ρ.

Using the stick-breaking representation for prior distribution of ρ, denoted by $\pi(\rho)$, parameter vector ρ is reparameterized as follows:

$$\begin{aligned}\rho_1 &= v_1,\\ \rho_l &= \prod_{o=1}^{l-1}(1-v_o)v_l,\ \forall\ o=2,3,\ldots,\end{aligned} \tag{5.17}$$

where the prior distribution of the vector $v=\{v_l\}_{l=1}^{\infty}$ is independent and identically distributed with beta distribution denoted by

$$v \sim \mathscr{B}(1,c), \tag{5.18}$$

and the hyper-prior distribution of c is

$$\pi(c) \sim \mathscr{G}(ac_0, bc_0), \tag{5.19}$$

where $\mathscr{G}(\cdot,\cdot)$ represents the gamma distribution (ESCOBAR; WEST, 1995). Besides, we chose a normal-gamma distribution as the prior of $(\mu_l,\tau_l) \sim \mathscr{N}\mathscr{G}(m_0,s_0,d_0p_0,d_0)$, for $l=1,2,\ldots$, due to the fact that this prior is conjugate to the normal distribution, where

$$\begin{aligned}&\mu_l \mid \tau_l \sim \mathscr{N}(m_0,(s_0\tau_l)^{-1}),\\ &\tau_l \sim \mathscr{G}(d_0,d_0p_0).\end{aligned}$$

Thus, joint prior density of Ω can be expressed as

$$\pi(\Omega) = \pi(c)\pi(\rho)\pi(\mu,\tau). \tag{5.20}$$

For our Bayesian estimation scheme, the joint posterior distribution of $\mathbf{Z}$ and all the unknown parameters in Ω are reached by joining all the prior information (5.16), (5.20) and the likelihood function (5.8), as follows:

$$\pi(\mathbf{Z},\Omega \mid \mathbf{D}) \propto L_1(Z|\mathbf{D})f_Z(\mathbf{Z} \mid \Omega)\pi(\Omega). \tag{5.21}$$

However, it is easy to see that (5.21) does not have a closed form. Besides, the marginal posterior of $\mathbf{Z}$ is intractable and it is therefore necessary to use MCMC algorithms, as we will see next. Recalling that one of our primary goals is to estimate Z_js, thus, the Bayes estimator of $\mathbf{Z}$ is given by

$$\hat{\mathbf{Z}}^{Bayes} = \sum_{i=1}^{L}\frac{\mathbf{Z}^{(i)}}{L}, \tag{5.22}$$

where $\mathbf{Z}^{(i)}$ is the i-th iteration and L is the total number of iterations of the MCMC chain.

MCMC algorithm

This section describes an MCMC algorithm to sample from the posterior distribution of Z. Our algorithm is based on Kalli, Griffin and Walker (2011), and its main characteristic is to estimate infinite parameters by introducing latent variables. We introduce a finite set of latent variables with uniform distribution with parameters 0 and 1, denoted by $U \sim \text{Uniform}[0,1]$. Therefore, applying the variable U in (5.16) follows the joint density of (Z,U)

$$f_{Z,U}(z,u \mid \Omega) = \sum_{l=1}^{\infty} \mathscr{LN}(z \mid \mu_l, \tau_l^{-1})\mathbb{I}(u < \rho_l), \tag{5.23}$$

where $\mathbb{I}(\cdot)$ is an indicator function. Note that there is a finite number of elements in ρ which are greater than u, denoted as $A_\rho(u) = \{j : \rho_j > u\}$. Therefore, the representation in (5.23) is similar to

$$f_{Z,U}(z,u \mid \Omega) = \sum_{l \in A_\rho} \mathscr{LN}(z \mid \mu_l, \tau_l^{-1}), \tag{5.24}$$

so that, given $\mathbf{U}$, the number of mixture components is finite for $\mathbf{Z}$.

In order to simplify the likelihood, we introduce a new discrete latent variable Y which indicates the mixture component that Z comes from

$$f_{Z,U,Y}(z,u,Y = l \mid \Omega) = \mathscr{LN}(z \mid \mu_l, \tau_l^{-1})\mathbb{I}(l \in A_\rho(u)). \tag{5.25}$$

Note that $Pr(Y = l \mid \Omega) = \rho_l$, $\forall l = 1,2,\ldots$, therefore the conditional distribution of $Z \mid U, Y = l$ is log-normal with parameters μ_l and τ_l^{-1}, so $W \mid U, Y = l \sim \mathscr{N}(\mu_l, \tau_l^{-1})$. Hence, the complete posterior distribution of $\mathbf{Z}, \Omega$ with the latent variables $\mathbf{U}$ and $\mathbf{Y}$ is given by

$$\pi(\mathbf{Z}, \Omega, \mathbf{U}, \mathbf{Y} \mid \mathbf{D}) \propto L_1(Z|\mathbf{D}) f_{Z,U,Y}(\mathbf{Z} \mid \Omega, U, Y) f_U(U) Pr(Y \mid \Omega)\pi(\Omega), \tag{5.26}$$

where $U = \{U_j\}_{j=1}^m$ and $Y = \{Y_j\}_{j=1}^m$ are latent variables.

Hybrid MCMC - computational strategy

Using the latent variables presented above, we now construct the following MCMC algorithm which is a combination of the Gibbs sampler with the HMC method. For more details on the HMC method, see Neal *et al.* (2011). We chose the HMC algorithm because it generates samples with less dependence with a high probability of acceptance between state if compared with the Random Walk Metropolis-Hastings algorithm. The Gibbs algorithm requires knowledge of complete conditional distributions in order to be able to sample from them. For further details, see Kalli, Griffin and Walker (2011) and Escobar and West (1995). The complete conditional distributions are listed below.

1. **Conditional Distribution of c**

Escobar and West (1995) shows that given $\mathbf{Y}$, the parameter is independent of all other parameters and the conditional distribution of c is given by

$$\pi(c \mid \mathbf{Y}) \propto (c+m)c^{y^*-1}\mathscr{G}(c \mid ac_0, bc_0)\mathbb{B}(c+1,m)\mathbb{I}(c>0), \tag{5.27}$$

where $y^* = \max(\mathbf{Y})$ and $\mathbb{B}(\cdot,\cdot)$ is the Beta function. Using the definition of the Beta function we can create an auxiliary variable ξ with the joint distribution for which the marginal distribution is (5.27) and is given by

$$\pi(c,\xi \mid \mathbf{Y}) \propto (c+m)c^{y^*-1}\xi\mathscr{G}(c \mid ac_0, bc_0)\,\xi^c(1-\xi)^{m-1}\mathbb{I}(c>0)\mathbb{I}(0<\xi<1). \tag{5.28}$$

Hence, it follows that the conditional posteriors of ξ and c are given by

$$\xi \mid c, \mathbf{Y} \sim \mathscr{B}(c+1,m) \tag{5.29}$$

and

$$c \mid \xi, \mathbf{Y} \sim p_\xi \mathscr{G}(a_1^*, b_1^*) + (1-p_\xi)\mathscr{G}(a_2^*, b_1^*), \tag{5.30}$$

where $a_1^* = a_0 + y^*$, $a_2^* = a_1^* + 1$, $b_1^* = b_0 - \log(\xi)$ and $p_\xi = (a_0 + y^* - 1)/(a_0 + z^* - 1 + m(b_0 - \log(\xi)$. Therefore, c can be sampled using the auxiliary ξ with equations (5.29) and (5.30).

2. **Conditional Distribution of ν**

   Note that by equations (5.25) and (5.26), ν depends on $\mathbf{Y}$, $\mathbf{U}$ and c, therefore the conditional distribution of ν is

$$\nu_l \mid \mathbf{Y}, \mathbf{U}, c \sim \begin{cases} \mathscr{B}(n_l + 1, m + \sum_{o=1}^{l} n_o + c) & ,\forall l = 1,\ldots,y^* \\ \mathscr{B}(1,c) & ,\forall l = y^*+1, y^*+2,\ldots, \end{cases} \tag{5.31}$$

   where n_l is the number of observations in the l-th component. It is worth noting that in order to sample ρ it is enough to simulate ν calculated by equation (5.17).

3. **Conditional Distribution of U**

   The latent variable U depends only on ρ, and the conditional distribution of $\mathbf{U}$ is

$$U_j \mid \rho \sim \text{Uniform}[0,\rho_j]\ \forall j = 1,2,\ldots,m. \tag{5.32}$$

4. **Conditional Distribution of μ and τ**

   The μ and τ parameters of each component are independent and adding the fact that the Normal-Gamma is conjugated from the Normal distribution, the conditional distribution of μ and τ is given by

$$\mu_l, \tau_l \mid \mathbf{Y} \sim \begin{cases} \mathscr{N}\mathscr{G}(m_l, s_l, d_l p_l, d_l) & ,\forall l = 1,\ldots,y^* \\ \mathscr{N}\mathscr{G}(m_0, s_0, d_0 p_0, d_0) & ,\forall l = y^*+1, y^*+2,\ldots, \end{cases} \tag{5.33}$$

where

$$\begin{aligned}
m_l &= \frac{s_0 m_0 + n_l \bar{w}}{s_0 + n_l}, \\
s_l &= s_0 + n_l, \\
d_l p_l &= d_0 p_0 + \sum_{j:y_j=l} (w_j - \bar{w})^2 + \frac{s_0 n_l}{s_0 + n_l}(m_0 - \bar{w})^2, \\
d_l &= d_0 + n_l, \\
\bar{w} &= \sum_{j:y_j=l} \frac{w_j}{n_l}.
\end{aligned}$$

5. **Conditional Distribution of Y**

The latent variable Y is discrete, therefore using equations (5.25) and (5.26) the conditional distribution of **Y** is

$$Pr(Y_j = l \mid \Omega, \mathbf{W}, \mathbf{U}, \mathbf{D}) \propto \mathcal{N}(w \mid \mu_l, \tau_l^{-1}) \mathbb{I}(l \in A_\rho). \tag{5.34}$$

6. **Conditional Distribution of Z**

The conditional distribution of **Z** is given by

$$\pi(\mathbf{Z} \mid \Omega, U, Y, D) \propto \prod_{j=1}^{m} \mathcal{LN}(z_j \mid \mu_{Y_j}, \tau_{Y_j}^{-1}) L_1(Z|\mathbf{D}), \tag{5.35}$$

with restriction $\bar{\mathbf{Z}} = 1$. Different from the previous parameters and latent variable, we simulate them using the HMC algorithm. However, the HMC algorithm requires that the support random variable is unrestricted. Therefore, we transform the variable **Z** to a variable with unrestricted support as explained below.

Let $\mathbf{Z}^*$ be a random vector with $m-1$ elements and unrestricted support. We define the following variables:

$$\begin{aligned}
B_j &= \text{logit}^{-1}(Z_j^* - \log(m-j)), \\
A_j &= \left(1 - \sum_{j'=1}^{j-1} A_{j'}\right) B_J \;\; \forall j = 1, 2, \ldots, m-1, \\
A_m &= 1 - \sum_{j'=1}^{m-1} A_{j'},
\end{aligned} \tag{5.36}$$

where $logit^{-1}$ is an inverse function of $logit$. Note that the functions of transformed variables are bijection, $B_j \in (0,1)$ and sum(**A**) = 1. Naturally, we assume that $\mathbf{Z} = m\mathbf{A}$. Therefore, the determinant of the Jacobian matrix is given by,

$$\mid J(\mathbf{z}^*) \mid = \prod_{j=1}^{m-1} \left(b_j(1-b_j) \left(1 - \sum_{j'=1}^{j-1} a_{j'} \right) \right).$$

Therefore, the conditional distribution of $\mathbf{Z}^*$ is given by

$$\pi(\mathbf{Z}^* \mid \Omega, U, Y, D) \propto \mid J(\mathbf{z}^*) \mid \mathcal{LN}(z_j \mid \mu_{Y_j}, \tau_{Y_j}^{-1}) L_1(Z|\mathbf{D}). \tag{5.37}$$

Thus, we constructed a Hybrid MCMC algorithm that combines Gibbs sampling with HMC sampling to sample $\mathbf{Z}$ and Ω; see Algorithm 2.

Algorithm 2 – Hybrid MCMC algorithm.

1: Initialize $c^{(0)}$, $\mathbf{Z}^{*(0)}$ and $\mathbf{Y}^{(0)}$.

2: Calculate $\mathbf{Z}^{(0)}$ of Equation (5.36) and $\mathbf{W}^{(0)} = \log(\mathbf{Z}^{(0)})$.

3: Draw $\xi^{(i)}$ from $\pi(\xi \mid c^{(i-1)}, \mathbf{Y}^{(i-1)})$ of Equation (5.29).

4: Draw $c^{(i)}$ from $\pi(c \mid \xi^{(i)}, \mathbf{Y}^{(i-1)})$ of Equation (5.30).

5: Draw $v_l^{(i)}$ from $\pi(v_l \mid \mathbf{Y}^{(i-1)}, c^{(i)})$ of Equation (5.31), $\forall l = 1, 2, \ldots, y^*$.

6: Calculate $\rho_l^{(i)}$ of Equation (5.17) $\forall l = 1, 2, \ldots, y^*$.

7: Draw $U_j^{(i)}$ from $\pi(U_j \mid \rho^{(i)})$ of Equation (5.32) $\forall j = 1, 2, \ldots, m$.

8: Find the smallest l^* such that $\sum_{l=1}^{l^*} \rho_l > (1 - \min(\mathbf{U}^{(i)}))$ and draw $v_l^{(i)}$ from $\pi(v_l \mid \mathbf{Y}^{(i-1)}, c^{(i)})$, $\forall l = y^* + 1, \ldots, l^*$.

9: Draw $\mu_l^{(i)}$ and $\tau_l^{(i)}$ from $\pi(\mu_l, \tau_l \mid \mathbf{Y}^{(i-1)})$ of Equation (5.33) $\forall l = 1, 2, \ldots, l^*$.

10: Draw $Y_j^{(i)}$ from $Pr(Y_j \mid \mu^{(i)}, \tau^{(i)}, \mathbf{W}^{(i-1)}, \mathbf{U}^{(i)}, \mathbf{D})$ of Equation (5.34) $\forall j = 1, 2, \ldots, m$.

11: Draw $\mathbf{Z}^{*(i)}$ from $\pi(\mathbf{Z}^* \mid \mu^{(i)}, \tau^{(i)}, U^{(i)}, Y^{(i)})$ of Equation (5.37).

12: Calculate $\mathbf{Z}^{(i)}$ of Equation (5.36) and $\mathbf{W}^{(i)} = \log(\mathbf{Z}^{(i)})$.

13: Set $i = i + 1$ and go to Step #3.

In this scheme, the HMC sampler is applied in Step #11. The algorithm was developed in the C++ language using the RccpArmadilho library (EDDELBUETTEL; SANDERSON, 2014). Its main advantages are processing speed and interaction with the R program (R Core Team, 2016). This code was used both in the generation of posterior sampling and in the simulation study presented in the following section.

5.5 Simulation study

In this section, a simulation study is performed to evaluate the efficiency of the Bayesian estimators via the Monte Carlo method. To make our presentation easier, we consider two causes of failure with distinct parameters for each cause $\theta = (\beta_1, \alpha_1, \beta_2, \alpha_2)$. The proposed simulation design is consistent with the following setup: (i) there are $m = (10, 50, 100)$ systems, each observed on the fixed time interval from $(0, 20]$; (ii) the failure process for each component follows a power-law NHPPs with intensity (5.3); (iii) among the many possible parameter choices, we provide details for ($\beta_1 = 1.2, \alpha_1 = 5, \beta_2 = 0.7, \alpha_2 = 13.33$) and ($\beta_1 = 0.75, \alpha_1 = 9.46, \beta_2 = 1.25, \alpha_2 = 12.69$); and (iv) we generate each random observation z_j, $j = 1, \ldots, m$, *iid* with mean one and variance η, according to a gamma distribution. In addition, we consider a set of values for variance of Z, $\eta = (0.5, 1, 5)$, indicating low, middle and high dependence degrees, respectively. For each setup of parameters, we obtain the mean number of failures (5, 13.3), (9.5, 12.7), respectively. In the first simulated scenario, the mean number of failures of one of the components is predominant over the other component. In the last scenario, the mean number of failures of each component are almost equal to each other. It is worth noting that the obtained results are similar for other parameter combinations and can be extended to more causes, i.e. $p > 2$. Using the fact that the causes are dependent due to frailty term Z_j and also using the known results from the literature about NHPPs (RIGDON; BASU, 2000), in each Monte Carlo replication the failure times and indicators of the cause of failure were generated as shown in Algorithm 3.

Algorithm 3 – Algorithm for generating random data from multiple systems with PLPs under competing risks.

1: Generate iid $z_j \sim \gamma(\eta, 1/\eta)$ for $j = 1, 2, \ldots, m$, with mean one and variance η.

2: For each cause of failure, generate random numbers n_{j1} and n_{j2}, $j = 1, \ldots, m$, both from a Poisson distribution with mean $z_j \alpha_q$, for $q = 1, 2$, respectively.

3: For the q-th cause of failure from j-th system, let the failure times be $t_{j,1,q}, \ldots, t_{j,n_j,q}$, where $t_{j,i,q} = T\, U_{j,i,q}^{1/\beta_{jq}}$ and $U_{j,1,q}, \ldots, U_{j,n_j,q}$ are the order statistics of a size n_j random sample from the standard uniform distribution.

4: Finally, to obtain the data in the form (t_i, δ_i), let the t_is be the set of ordered failure times and set δ_i equal to j according to the corresponding cause of failure (i.e., set $\delta_i = 1$ if $t_i = t_{h,1}$ for some h or $\delta_i = j$ depending on the cause of failure).

Software R was used to implement this simulation study (R Core Team, 2016). We considered two criteria to evaluate the estimators' behaviour: the Bias, given by $Bias_{\hat{\theta}_i} = \sum_{j=1}^{M} (\hat{\theta}_{i,j} - \theta_i)/M$ and the MSE, given by $MSE_{\hat{\theta}_i} = \sum_{j=1}^{M} (\hat{\theta}_{i,j} - \theta_i)^2/M$, where M is the number of estimates (i.e. the Monte Carlo size), where we take $M = 50{,}000$ throughout the section, and $\theta = (\theta_1, \ldots, \theta_p)$ is the vector of parameters. Additionally, we computed the $CP_{95\%}$. Good estimators should have Bias, MSE close to zero and adequate intervals should be short while showing $CP_{95\%}$ close to 0.95. The Bias and MSE are widely used to measure the performance evaluation.

The Bayes estimators for β_j and α_j were obtained using independent marginal posteriors according to gamma distributions given in (5.11). Since the marginal posterior distributions for the parameters β_j and α_j follow gamma distributions, we can obtain closed-form expressions for the posterior means and obtain the credibility intervals based on the 2.5% and 97.5% percentile posteriors. Hence, no MCMC was needed to obtain the estimates for these parameters. On the other hand, to obtain the estimates of the Z_js, $j = 1,\ldots,m$, we considered the HMC described in Section 5.4.2. For each simulated data set, 10,000 iterations were performed using the MCMC methods. As a burn-in, the first 5,000 initial values were discarded. The Geweke criterion (GEWEKE, 1992) was considered to check the convergence of the obtained chains under a 95% confidence level. In addition, trace and autocorrelation plots of the generated sampled values of each Z_j showed that they converged to the target distribution. The remaining 5,000 were used for posterior inference. Specifically, these values were used to compute the posterior means of Z_js. Table 9 presents the Bias, the MSE and coverage probability with a 95% confidence level of the Bayes estimates for $\alpha_1, \alpha_2, \beta_1, \beta_2$ and the variance of Z.

As shown in Tables 9 and 10, the biases of the Bayes estimator are very close to zero for all the parameters, while both Bias and MSE tend to zero as m increases. Hence, in terms of Bias and MSE, the Bayes estimators provided accurate inferences for the parameters of the PLP

Table 9 – The Bias, MSE, CP(95%) from the estimates considering different values for variance of Z and number of systems (m) with scenario θ=(1.2, 5, 0.7, 13.3).

η	Parameter	m	α_1	α_2	β_1	β_2	η
0.5	Bias	10	-0.0041	0.0083	-0.0001	0.0007	0.0784
		50	0.0011	-0.0055	-0.0001	0.0002	0.0276
		100	-0.0001	0.0074	0.0000	-0.0001	0.0165
	MSE	10	0.7035	1.1696	0.0729	0.0882	0.2709
		50	0.3182	0.5093	0.0321	0.0390	0.1359
		100	0.2230	0.3650	0.0225	0.0276	0.0957
	CP(95%)	10	0.9427	0.9443	0.9449	0.9501	0.8395
		50	0.9483	0.9500	0.9459	0.9506	0.9366
		100	0.9502	0.9496	0.9512	0.9488	0.9444
1	Bias	10	0.0084	-0.0105	0.0003	-0.0023	0.0307
		50	0.0020	-0.0007	-0.0006	0.0001	0.0253
		100	0.0010	0.0039	0.0000	0.0000	0.0158
	MSE	10	0.6996	1.1449	0.0735	0.0879	0.5395
		50	0.3120	0.5185	0.0316	0.0393	0.2891
		100	0.2231	0.3690	0.0226	0.0275	0.2015
	CP(95%)	10	0.9444	0.9517	0.9432	0.9477	0.9423
		50	0.9532	0.9488	0.9477	0.9477	0.9544
		100	0.9477	0.9472	0.9489	0.9492	0.9478
5	Bias	10	0.0174	-0.0120	-0.0005	0.0009	-0.1693
		50	-0.0017	-0.0085	-0.0009	0.0000	0.0453
		100	0.0000	0.0005	0.0001	-0.0004	-0.0425
	MSE	10	0.7156	1.1508	0.0723	0.0873	2.2234
		50	0.3179	0.5141	0.0317	0.0390	2.1358
		100	0.2239	0.3711	0.0223	0.0276	1.5038
	CP(95%)	10	0.9419	0.9508	0.9460	0.9483	0.9340
		50	0.9476	0.9504	0.9500	0.9472	0.9473
		100	0.9470	0.9462	0.9505	0.9476	0.9426

Table 10 – The Bias, MSE, CP(95%) from the estimates considering different values for variance of Z and number of systems (m) with scenario θ=(0.75, 9.5, 1.25, 12.7)

η	Parameter	m	α_1	α_2	β_1	β_2	η
0.5	Bias	10	0.0215	-0.0189	0.0004	0.0000	0.0745
		50	0.0017	-0.0003	0.0004	-0.0004	0.0218
		100	0.0025	-0.0005	0.0003	0.0001	0.0155
	MSE	10	0.9632	1.1248	0.0787	0.1120	0.2691
		50	0.4341	0.4992	0.0347	0.0495	0.1312
		100	0.3085	0.3569	0.0243	0.0350	0.0946
	CP(95%)	10	0.9498	0.9477	0.9467	0.9470	0.8346
		50	0.9506	0.9516	0.9497	0.9508	0.9377
		100	0.9471	0.9462	0.9509	0.9482	0.9417
1	Bias	10	-0.0025	-0.0013	0.0003	0.0003	0.0233
		50	0.0005	0.0039	-0.0004	0.0003	0.0155
		100	0.0024	-0.0013	0.0001	-0.0003	0.0087
	MSE	10	0.9678	1.1311	0.0780	0.1138	0.5279
		50	0.4340	0.5065	0.0346	0.0497	0.2808
		100	0.3087	0.3592	0.0246	0.0355	0.1987
	CP(95%)	10	0.9497	0.9503	0.9478	0.9471	0.9407
		50	0.9495	0.9465	0.9477	0.9515	0.9510
		100	0.9456	0.9470	0.9463	0.9482	0.9495
5	Bias	10	-0.0148	-0.0076	-0.0008	-0.0002	-0.1558
		50	-0.0027	-0.0039	0.0007	-0.0001	0.0577
		100	0.0040	-0.0040	-0.0003	0.0002	-0.1141
	MSE	10	0.9663	1.1197	0.0785	0.1119	2.2178
		50	0.4366	0.5044	0.0348	0.0497	2.0724
		100	0.3075	0.3592	0.0246	0.0347	1.4197
	CP(95%)	10	0.9517	0.9512	0.9493	0.9491	0.9354
		50	0.9491	0.9471	0.9499	0.9522	0.9479
		100	0.9489	0.9467	0.9497	0.9551	0.9485

model. In terms of coverage probabilities, we observed that using our Bayes estimators returned accurate credibility intervals even for a small number of system m. This result may be explained by the fact that our proposed Bayes estimators do not depend on asymptotic results to obtain the credibility intervals, which leads to accurate results for small sample sizes.

5.6 Application to the warranty repair data

The dataset considered in this section comprises the recurrent failure history of a fleet of identical automobiles obtained from a warranty claim database presented in Somboonsavatdee and Sen (2015a). For the sake of clarity, our graphics present only the cars that presented failures in the observation period. Figure 11 shows the recurrence of failures of the 172 cars according to the cause of failure and the car mileage at each failure. The x-axis indicates the mileage. It is worth noting that the process of data collection has truncated time, where the observation period is 3000 miles for all cars. Each car from the fleet is represented by a horizontal line, where the cause of failure 1 is identified by the green circle, the cause of failure 2 by the red triangle and the cause of failure 3 by the blue square. We suppose that maintenance policy is minimal repair.

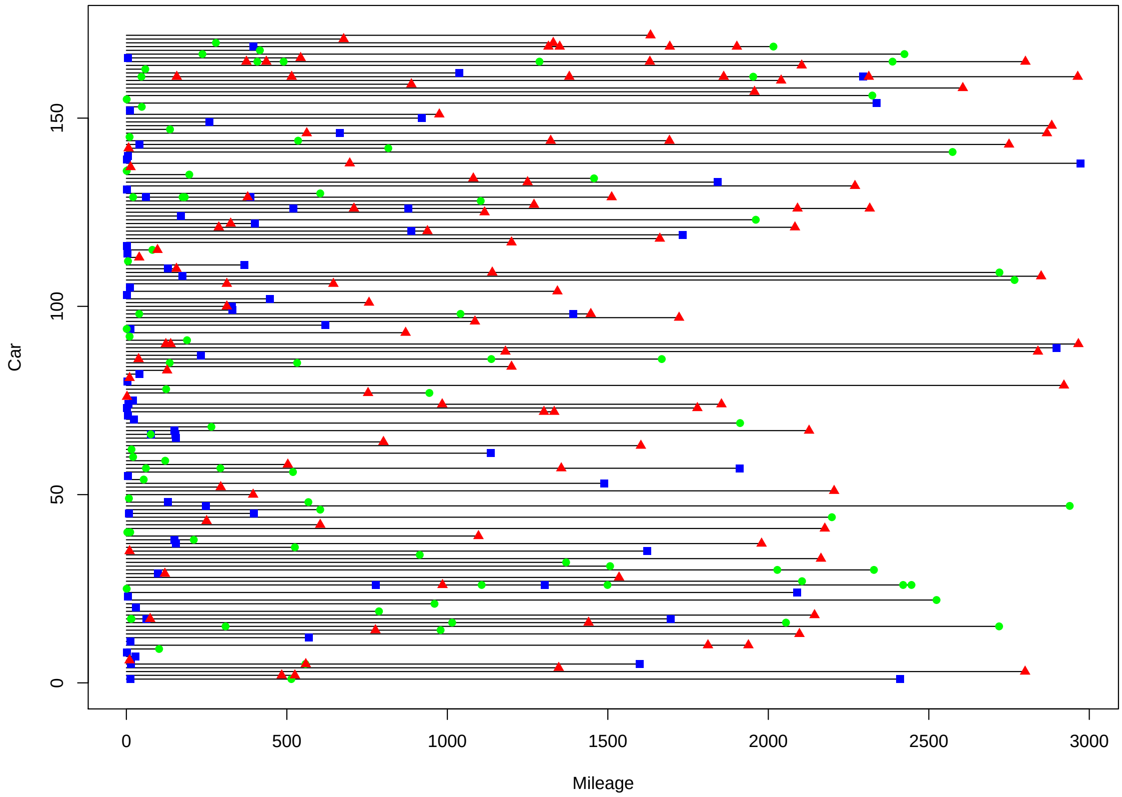

Figure 11 – Recurrences of three causes of failure for 172 cars from warranty claims data. The green circle represents the cause of failure 1, the red triangle represents the cause of failure 2 and the blue square represents the cause of failure 3.

The main authors make only a table available (omitted here) containing the mileage to repeated failures of 172 vehicles, as well as the associated cause of failure. There were 76 failures related to the cause of failure 1, 87 related to the cause of failure 2 and 111 related to the cause of failure 3. They also pointed out that there were 267 cars that did not fail during the observation period. However, following the correct methodology, we consider 439 automobiles in our analysis.

Following Somboonsavatdee and Sen (2015a), Somboonsavatdee and Sen (2015b), we assessed the adequacy of the PLP for each cause of failure using the Duane plot (DUANE, 1964; CROW, 1974; RIGDON; BASU, 2000). Figure 12 shows plots of logarithm of the number of failures $N_q(t)$ (for $q = 1, 2, 3$) against the logarithm of the accumulated mileage at failure. Since the three plots exhibit reasonable linearity, the PLP model seems to be adequate.

Since the PLP is adequate we consider our proposed approach to fit the data. As presented

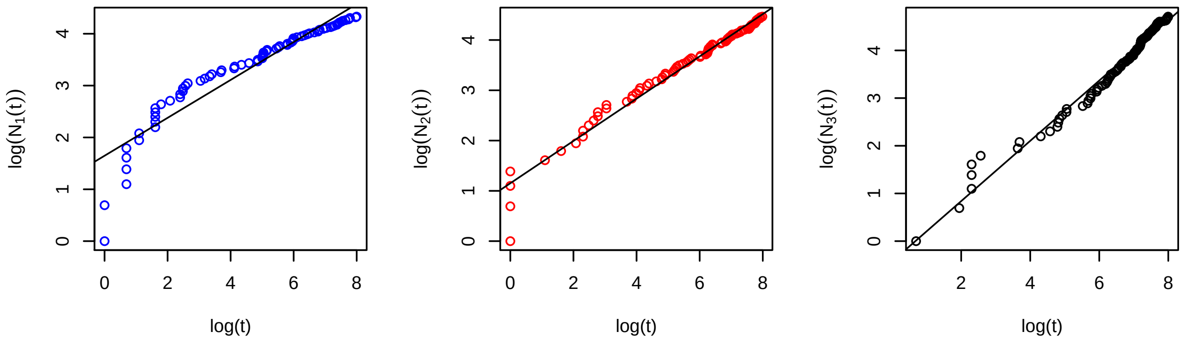

Figure 12 – The plot shows a fairly linear pattern for the three causes of failure indicating the fit according to the PLP model: cause 1 (blue circles); cause 2 (red circles) and cause 3 (black circles)

in Section 5.4, we assume the prior distribution (5.10) for parameters α_q and β_q ($q = 1,2,3$) and, consequently, the marginal posterior distributions (5.11). On the basis of the latter consideration, the posterior mean estimates are computed in closed-form and the CIs are obtained directly from the gamma distribution. The results of the analysis are presented in Table 11, which show Bayes estimates along with the corresponding SDs and CIs. According to these data, the estimates of the shape parameters $\left(\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3\right)$ are smaller than 1; see Table 11. This clearly indicates improvement in reliability.

Table 11 – Parameter estimates for warranty claim dataset

Parameter	Bayes	SD	CI (95%)
β_1	0.300	0.035	[0.236 ; 0.372]
β_2	0.409	0.044	[0.327 ; 0.500]
β_3	0.698	0.067	[0.574 ; 0.835]
α_1	0.173	0.020	[0.136 ; 0.214]
α_2	0.198	0.021	[0.159 ; 0.242]
α_3	0.253	0.024	[0.208 ; 0.302]
$Var(Z)$	1.755	0.438	[1.050 ; 2.777]

The hybrid MCMC sampler algorithm presented in Section 5.4.2 was used to obtain a sample from the joint posterior distribution related to the frailty distribution. The initial values to start the sample of the chains for the DPM were random. For the MCMC chain, we considered 10,000 iterations initially, where the first 5,000 were discarded as burn-in samples and the last 5,000 iterations were used to compute the posterior estimates of $Var(Z)$ (at the bottom of the Table 11) and the individual values of Z_js, as presented in Figure 13. The convergence was monitored for the Geweke test assuming a 95% confidence level (see Figure 16 in Appendix B). For completeness, we also present MCMC diagnostic plots, such as traces and autocorrelations for the HMC algorithm; see Appendix B.

It is worth pointing out that higher values of $Var(Z)$ signify greater heterogeneity among systems and more dependence between the times of the causes of failure for the same system. Therefore, as Table 11 shows, the posterior mean of $Var(Z)$ provides evidence of a meaningful dependence between the times of the causes of failure within a system.

5.6.1 Insights on the unobserved heterogeneity

As shown in Table 11, the estimate of $Var(Z)$ shows that there is strong posterior evidence of a meaningful degree of heterogeneity in the population of systems. Table 12 (Appendix B) shows the estimated posterior means and the corresponding standard deviations of the $\hat{z}_j$s.

Figure 13 shows the individual frailty estimates (posterior means) of $\hat{z}_j, j = 1, \ldots, 172$. As mentioned earlier, each Z_j acts in a multiplicative way in the specific-cause intensities. Thus it follows that values of Z_j equal to or very close to 1 (red line) do not significantly affect such intensities. On the other hand, values larger than 1 indicate increased intensity. It is apparent that some cars have values of Z_j greater than 2. These cars are probably subject to environmental stress variations or other unobserved issues, which make them more vulnerable than those with Z_j values closer to or less than 1.

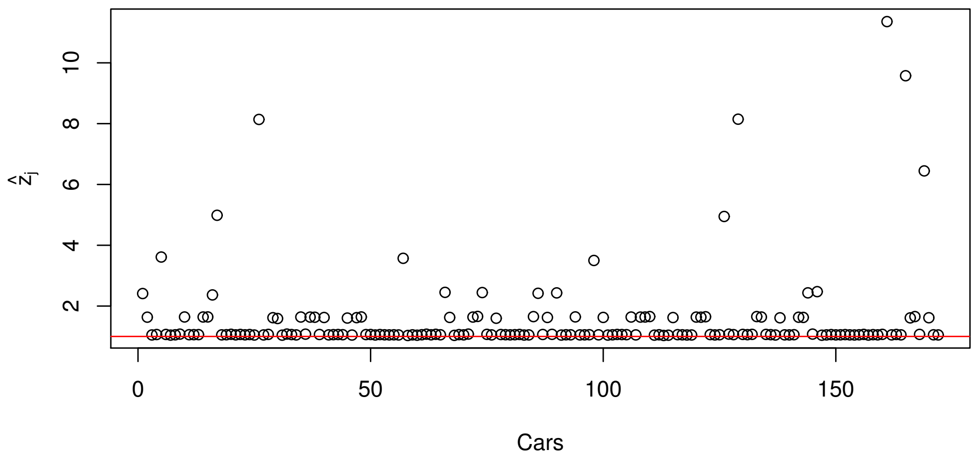

Figure 13 – The individual frailty estimates, $\hat{z}_j$'s. The red line highlights value 1 in the y-axis.

Figure 14 indicates that the estimated frailties are overall larger for cars that had a failure early than those who had a failure later. We also note that a system with a large value of $\hat{z}_j$ experienced more failures than a system with a smaller value of $\hat{z}_j$ (see Figure 15).

These outcomes indicate that neglecting these effects can result in an underestimation of the parameters. Overall, the multiplicative shared frailty model is appropriate for modeling this effect accurately.

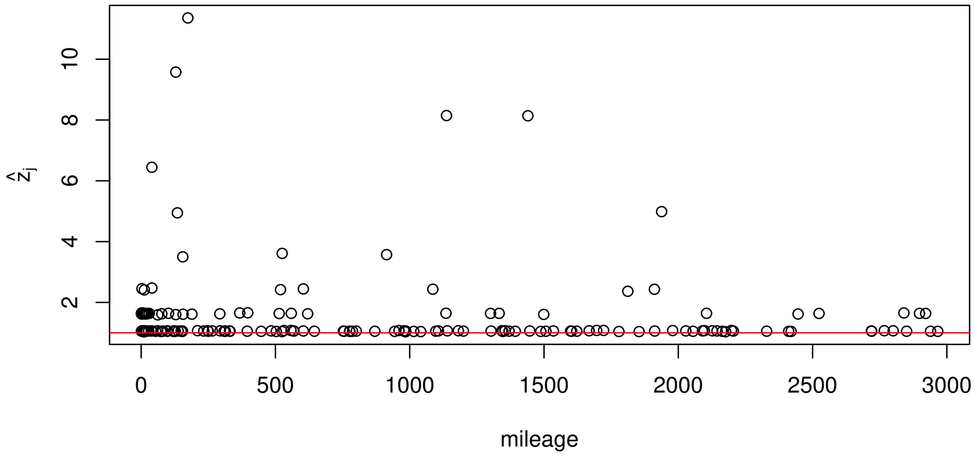

Figure 14 – Estimated frailty versus mileage observed at failure for each car in automobile warranty data. The red line highlights value 1 in the y-axis. The reasoning is that cars that are more frail failed earlier than ones that are less frail.

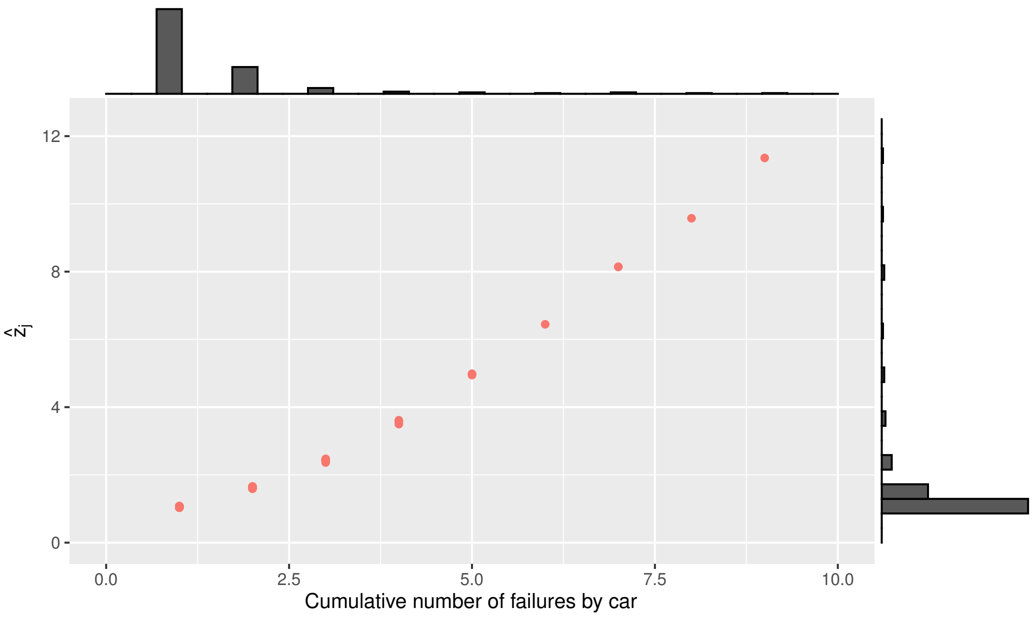

Figure 15 – Scatterplot of individual estimates $\hat{z}_j$ against cumulative number of failures by car. Note that systems with a large value of $\hat{z}_j$ experienced more failures than a system with a smaller value of $\hat{z}_j$.

5.7 Conclusions

In this chapter, we proposed a new approach to analyzing multiple repairable systems data under the action of dependent competing risks. We have shown how to model the frailty-induced dependence nonparametrically using a DPM which does not make restrictive assumptions about the density of the frailty variable. Although some research has been carried out on nonparametric frailty in the reliability field (SLIMACEK; LINDQVIST, 2016; SLIMACEK; LINDQVIST, 2017), to the best of our knowledge, the proposed approach is the first for this competing risks setup. The main focus of this chapter was to provide estimates for the PLP model taking into account the dependence effect among component failures of the system. Such a dependence effect influences the statistical inferences of the model parameters, thus the misspecification of the frailty distribution may lead to errors when estimating the quantities of interest.

An orthogonal parametrization for the cause-specific intensity PLP parameters was presented, which allowed us to consider a generalized version of Bar-Lev, Lavi and Reiser (1992) prior distribution for the parameters of the model. Assuming the quadratic loss function as the risk function, we obtained the posterior mean for the parameters in closed-form expression. Moreover, since the marginal posterior distributions for the PLP parameters follow gamma distributions, we obtained the credibility intervals directly for the quantile function. Assuming a specific value for ζ, we obtained unbiased estimators for the cited parameters. A simulation study was conducted to confirm our theoretical results, as well as to measure if the variability of the frailty distributions was correctly computed. This study returned excellent results that confirmed that our Bayes estimators are robust in terms of Bias, MSE and coverage probabilities.

Using nonparametric Bayesian methods with a mixture prior distribution enabled us to increase the amount of information beyond the parameter estimates. We considered a Bayesian nonparametric prior to describing the frailty distribution due to its flexibility in modeling unknown distributions. Although this model has infinite parameters, it is a flexible mixture model, parsimonious and straightforward to sample from. In this case, we chose the stick-breaking representation of the DP prior because of a simple implementation to build the algorithm. Hence, we proposed a hybrid MCMC algorithm that comprises a mixture of the Gibbs sampler and the HMC method, thus generating a chain with little dependence.

The results of this investigation show that we can obtain more precise parameter estimations by considering the high flexibility due to nonparametric Bayesian prior density for Z. It also enables us to obtain insights into the heterogeneity between the systems by individually estimating Z_js, as presented in Section 5.6.1. The methodology proposed in this study may be of assistance to industrial applications and also where the interest may be in the phases of developmental programs of prototypes with purposes to predict the reliability, for example.

Our findings can be applied in real data sets based on the following assumptions. The proposed model requires m identical repairable systems subjected to K competing risks (assuming dependece). Minimal repair policy is assumed. The recurrent data structure (failure history) should be based on cause-specific intensity functions with PLP. The data sampling scheme (system observational period) is the time truncated case. Consider the shared frailty model to incorporate the dependence among the cause-specific recurrent processes. Finally, the dataset should be structured as Table 13 in Appendix C.

More flexible modeling can be further proposed by extending our approach to model the intensity function of failures of the NHPP nonparametrically since that the PLP intensity cannot capture non-monotonic behaviors. This extension would make the model more robust and flexible. In this case, we would have a fully nonparametric approach. The proposed study can also be further adapted under other types of repair such as perfect or imperfect. Our approach should be investigated further in these contexts.

CHAPTER 6

COMMENTS AND FURTHER DEVELOPMENT

6.1 Comments

In this thesis, we studied certain aspects of modeling failure time data of repairable systems under a competing risks framework. Furthermore, we paid our attention to minimal repairs. The minimal repair concept is suitable for complex systems (multi-components) when the purpose is to bring the system back to operation as soon as possible and also from an economic viewpoint, where such systems are commonly repaired rather than replacing the system with a new one after failure. In this case, we specified the parametric framework of a power-law process. The failure history of each component is governed by a PLP. Throughout all the chapters we used an advantageous parametrization for the specific-cause intensity. In accordance with this framework, we considered two different models and proposed more efficient Bayesian methods for estimating the model parameters.

The first model refers to failure data arising from a single repairable system under independent competing risks. Using the Jeffreys prior and reference prior, we obtained proper posterior distributions with interesting properties such as one-to-one invariance and consistent marginalization. In addition, the Bayes MAP estimates have closed-form expressions. Besides, in some cases, the marginal posterior intervals have accurate frequentist coverage for all parameters. In addition to the theoretical proofs, an extensive simulation study is presented, which confirms that the resulting Bayes estimates are more accurate than the estimates obtained from the classical approach regarding bias and accurate credibility intervals.

In the second model, we explore a new methodology for analyzing failures from multiple repairable systems under the action of dependent competing risks. We specify a joint prior distribution for the PLP parameters that returned closed-form estimators for the posterior mean. Besides, we have shown how to model the frailty-induced dependence nonparametrically using a

DPM which does not make restrictive assumptions about the density of the frailty variable. The main focus here is to provide estimates for the PLP model taking into account the dependence effect among component failures of the system. The results of this investigation show that we can obtain more precise parameter estimation by considering the high flexibility due to nonparametric Bayesian prior density for Z. It also enables us to obtain insights into the heterogeneity between the systems by individually estimating Z_js.

In addition, the background of applications of the methodologies described in this thesis is expected to be advantageous in reliability analysis, engineering applications and other fields where the setup is equivalent. To this end, we also present at the end of each chapter the assumptions that must be considered for the correct use of each model.

6.2 Further developments

More flexible modeling can be further proposed by extending our approach to model the intensity function of failures of the NHPP nonparametrically since the PLP intensity cannot capture non-monotonic behaviors. This extension would make the model more robust and flexible. In this case, we would have a fully nonparametric approach. The proposed study can also be further adapted under other types of repair such as perfect or imperfect. Our approach should be investigated further in these contexts.

BIBLIOGRAPHY

AALEN, O.; BORGAN, O.; GJESSING, H. **Survival and event history analysis: a process point of view**. [S.l.]: Springer Science & Business Media, 2008. Citations on pages 31 e 40.

ANDERSEN, P. K.; BORGAN, O.; GILL, R. D.; KEIDING, N. **Statistical models based on counting processes**. [S.l.]: Springer Science & Business Media, 2012. Citations on pages 24, 30, 31, 40 e 41.

ANTONIAK, C. E. Mixtures of dirichlet processes with applications to bayesian nonparametric problems. **The annals of statistics**, JSTOR, p. 1152–1174, 1974. Citation on page 78.

ASCHER, H.; FEINGOLD, H. **Repairable system reliability: Modeling, Inference, Misconceptions and their cause**. [S.l.]: Marcel Dekker, New York, 1984. Citations on pages 23, 24 e 32.

ASFAW, Z. G.; LINDQVIST, B. H. Unobserved heterogeneity in the power law nonhomogeneous poisson process. **Reliability Engineering & System Safety**, Elsevier, v. 134, p. 59–65, 2015. Citations on pages 34 e 40.

AVEN, T. Optimal replacement under a minimal repair strategy, a general failure model. **Advances in Applied Probability**, Cambridge Univ Press, v. 15, n. 01, p. 198–211, 1983. Citation on page 23.

AVEN, T.; JENSEN, U. A general minimal repair model. **journal of applied probability**, Cambridge Univ Press, v. 37, n. 01, p. 187–197, 2000. Citation on page 23.

BAIN, L.; ENGLEHARDT, M. **Statistical Analysis of Reliability and Life-Testing Models: Theory and Methods**. [S.l.]: CRC Press, 1991. Citations on pages 36 e 54.

BAR-LEV, S. K.; LAVI, I.; REISER, B. Bayesian inference for the power law process. **Annals of the Institute of Statistical Mathematics**, Springer, v. 44, n. 4, p. 623–639, 1992. Citations on pages 24, 76, 77 e 91.

BARLOW, R.; HUNTER, L. Optimum preventive maintenance policies. **Operations research**, INFORMS, v. 8, n. 1, p. 90–100, 1960. Citation on page 23.

BERGER, J. O.; BERNARDO, J. M.; SUN, D. *et al.* Overall objective priors. **Bayesian Analysis**, International Society for Bayesian Analysis, v. 10, n. 1, p. 189–221, 2015. Citation on page 44.

BERNARDO, J. M. Reference posterior distributions for bayesian inference. **Journal of the Royal Statistical Society. Series B (Methodological)**, JSTOR, p. 113–147, 1979. Citations on pages 25 e 44.

____. Reference analysis. **Handbook of statistics**, Elsevier, v. 25, p. 17–90, 2005. Citations on pages 25 e 44.

BERNARDO, J. M.; SMITH, A. F. **Bayesian theory**. [S.l.]: John Wiley & Sons, 2009. Citations on pages 25 e 43.

BERNARDO, J. M.; TOMAZELLA, V. L. Objective bayesian reference analysis for the poisson process model in presence of recurrent events data. **Test**, Springer, v. 20, n. 1, p. 204–221, 2011. Citation on page 40.

BERNOULLI, D. Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prévenir. **Histoire de l'Acad., Roy. Sci.(Paris) avec Mem**, Des Math. And Phis., Mem, p. 1–45, 1760. Citation on page 38.

BRADLEY, L.; BRADLEY, L. **Smallpox inoculation: an eighteenth century mathematical controversy**. [S.l.]: University of Nottingham, Department of Adult Education, 1971. Citation on page 38.

CARVALHO, V. I. de. **Introduction to nonparametric Bayesian inference**. 2016. Available: <https://www.ime.usp.br/~lsanchez/ebeb2016/slides2vanda.pdf>. Accessed: 29/03/2019. Citation on page 46.

CASELLA, G.; BERGER, R. L. **Statistical inference**. [S.l.]: Duxbury Pacific Grove, CA, 2002. Citation on page 42.

CHA, J. H.; FINKELSTEIN, M. Some notes on unobserved parameters (frailties) in reliability modeling. **Reliability Engineering & System Safety**, Elsevier, v. 123, p. 99–103, 2014. Citations on pages 34 e 40.

CONSONNI, G.; FOUSKAKIS, D.; LISEO, B.; NTZOUFRAS, I. *et al.* Prior distributions for objective bayesian analysis. **Bayesian Analysis**, International Society for Bayesian Analysis, v. 13, n. 2, p. 627–679, 2018. Citation on page 25.

COOK, R. J.; LAWLESS, J. **The statistical analysis of recurrent events**. [S.l.]: Springer Science & Business Media, 2007. Citation on page 24.

CORSET, F.; DOYEN, L.; GAUDOIN, O. Bayesian analysis of ara imperfect repair models. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 41, n. 21, p. 3915–3941, 2012. Citation on page 25.

COX, D. R. Regression models and life-tables. **Journal of the Royal Statistical Society: Series B (Methodological)**, Wiley Online Library, v. 34, n. 2, p. 187–202, 1972. Citation on page 40.

COX, D. R.; HINKLEY, D. V. **Theoretical statistics**. [S.l.]: Chapman and Hall/CRC, 1979. Citation on page 42.

COX, D. R.; REID, N. Parameter orthogonality and approximate conditional inference. **Journal of the Royal Statistical Society. Series B (Methodological)**, JSTOR, p. 1–39, 1987. Citations on pages 37 e 45.

CROW, L. Reliability analysis for complex repairable systems. **Reliability and Biometry**, SIAM, 1974. Citation on page 87.

CROWDER, M. J. **Classical competing risks**. [S.l.]: CRC Press, 2001. Citations on pages 24 e 38.

____. **Multivariate survival analysis and competing risks**. [S.l.]: Chapman and Hall/CRC, 2012. Citation on page 38.

CROWDER, M. J.; KIMBER, A.; SWEETING, T.; SMITH, R. **Statistical analysis of reliability data**. [S.l.]: CRC Press, 1994. Citations on pages 24 e 38.

DATTA, G. S.; MUKERJEE, R. **Probability matching priors: higher order asymptotics**. [S.l.]: Springer Science & Business Media, 2012. Citation on page 45.

DEY, D. D.; MÜILER, P.; SINHA, D. **Practical nonparametric and semiparametric Bayesian statistics**. [S.l.]: Springer Science & Business Media, 2012. Citation on page 78.

DIJOUX, Y.; GAUDOIN, O. The alert-delay competing risks model for maintenance analysis. **Journal of Statistical Planning and inference**, Elsevier, v. 139, n. 5, p. 1587–1603, 2009. Citation on page 38.

DOYEN, L.; GAUDOIN, O. Imperfect maintenance in a generalized competing risks framework. **Journal of Applied Probability**, Cambridge University Press, v. 43, n. 3, p. 825–839, 2006. Citations on pages 24, 25 e 38.

DUANE, J. Learning curve approach to reliability monitoring. **IEEE transactions on Aerospace**, IEEE, v. 2, n. 2, p. 563–566, 1964. Citation on page 87.

EDDELBUETTEL, D.; SANDERSON, C. Rcpparmadillo: Accelerating r with high-performance c++ linear algebra. **Computational Statistics and Data Analysis**, v. 71, p. 1054–1063, March 2014. Available: <http://dx.doi.org/10.1016/j.csda.2013.02.005>. Citation on page 83.

ESCOBAR, M. D. Estimating normal means with a dirichlet process prior. **Journal of the American Statistical Association**, Taylor & Francis, v. 89, n. 425, p. 268–277, 1994. Citation on page 48.

ESCOBAR, M. D.; WEST, M. Bayesian density estimation and inference using mixtures. **Journal of the american statistical association**, Taylor & Francis, v. 90, n. 430, p. 577–588, 1995. Citations on pages 46, 48, 78, 79, 80 e 81.

FERGUSON, T. S. A bayesian analysis of some nonparametric problems. **The annals of statistics**, JSTOR, p. 209–230, 1973. Citations on pages 26, 46 e 47.

FERGUSON, T. S. *et al.* Prior distributions on spaces of probability measures. **The annals of statistics**, Institute of Mathematical Statistics, v. 2, n. 4, p. 615–629, 1974. Citations on pages 26 e 46.

FERREIRA, A.; GARCIA, N. L. Simulation study for misspecifications on a frailty model. **Brazilian Journal of Probability and Statistics**, JSTOR, p. 121–134, 2001. Citation on page 40.

FINKELSTEIN, M. Minimal repair in heterogeneous populations. **Journal of Applied Probability**, Cambridge Univ Press, v. 41, n. 01, p. 281–286, 2004. Citation on page 23.

FREEDMAN, D. A. *et al.* On the asymptotic behavior of bayes' estimates in the discrete case. **The Annals of Mathematical Statistics**, Institute of Mathematical Statistics, v. 34, n. 4, p. 1386–1403, 1963. Citation on page 46.

FU, J.; TANG, Y.; GUAN, Q. Objective bayesian analysis for recurrent events in presence of competing risks. **Quality Technology & Quantitative Management**, Taylor & Francis, v. 11, n. 3, p. 265–279, 2014. Citation on page 25.

GAUDOIN, O.; YANG, B.; XIE, M. Confidence intervals for the scale parameter of the power-law process. **Communications in Statistics—Theory and Methods**, Taylor & Francis, v. 35, n. 8, p. 1525–1538, 2006. Citation on page 36.

GEWEKE, J. Evaluating the accuracy of sampling-based approaches to the calculations of posterior moments. **Bayesian statistics**, Clarendon Press, v. 4, p. 641–649, 1992. Citation on page 85.

GUIDA, M.; CALABRIA, R.; PULCINI, G. Bayes inference for a non-homogeneous poisson process with power intensity law (reliability). **IEEE Transactions on Reliability**, IEEE, v. 38, n. 5, p. 603–609, 1989. Citation on page 24.

HANAGAL, D. D. **Modeling survival data using frailty models**. [S.l.]: Chapman and Hall/CRC, 2011. Citation on page 40.

HONG, Y.; MEEKER, W. Q. Field-failure and warranty prediction based on auxiliary use-rate information. **Technometrics**, Taylor & Francis, v. 52, n. 2, p. 148–159, 2010. Citation on page 25.

HOUGAARD, P. **Analysis of multivariate survival data**. [S.l.]: Springer Science & Business Media, 2012. Citations on pages 26, 40 e 41.

HØYLAND, A.; RAUSAND, M. **System reliability theory: models and statistical methods**. [S.l.]: John Wiley & Sons, 2009. Citations on pages 38 e 56.

JEFFREYS, H. An invariant form for the prior probability in estimation problems. In: THE ROYAL SOCIETY. **Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences**. [S.l.], 1946. v. 186, n. 1007, p. 453–461. Citation on page 43.

KALLI, M.; GRIFFIN, J. E.; WALKER, S. G. Slice sampling mixture models. **Statistics and computing**, Springer, v. 21, n. 1, p. 93–105, 2011. Citations on pages 74 e 80.

KASS, R. E.; WASSERMAN, L. The selection of prior distributions by formal rules. **Journal of the American Statistical Association**, Taylor & Francis, v. 91, n. 435, p. 1343–1370, 1996. Citations on pages 25, 43 e 76.

KINGMAN, J. F. C. **Poisson processes**. [S.l.]: Wiley Online Library, 1993. Citation on page 31.

KOTTAS, A. **Bayesian nonparametric methods**. 2018. Available: <https://ams241-spring18-01.courses.soe.ucsc.edu/system/files/attachments/notes-1_1.pdf>. Accessed: 29/03/2019. Citation on page 46.

KRIT, M. **Goodness-of-fit tests in reliability: Weibull distribution and imperfect maintenance models**. Phd Thesis (PhD Thesis) — Université de Grenoble, 2014. Citation on page 30.

KUO, L.; GHOSH, S. K. Bayesian nonparametric inference for nonhomogeneous poisson processes. In: **Technical Report 9718**. [S.l.]: University of Connecticut Department of Statistics, 1997. Citation on page 74.

KVAM, P. H.; VIDAKOVIC, B. **Nonparametric statistics with applications to science and engineering**. [S.l.]: John Wiley & Sons, 2007. Citation on page 46.

LANGSETH, H.; LINDQVIST, B. H. Competing risks for repairable systems: A data study. **Journal of Statistical Planning and Inference**, Elsevier, v. 136, n. 5, p. 1687–1700, 2006. Citations on pages 25 e 38.

LAWLESS, J. F. **Statistical models and methods for lifetime data**. [S.l.]: John Wiley & Sons, 2011. Citation on page 24.

LEHMANN, E. L. **Elements of large-sample theory**. [S.l.]: Springer Science & Business Media, 2004. Citation on page 54.

LI, L.; HANSON, T.; DAMIEN, P.; POPOVA, E. A bayesian nonparametric test for minimal repair. **Technometrics**, Taylor & Francis, v. 56, n. 3, p. 393–406, 2014. Citation on page 74.

LINDLEY, D. V.; SINGPURWALLA, N. D. Multivariate distributions for the life lengths of components of a system sharing a common environment. **Journal of Applied Probability**, Cambridge University Press, v. 23, n. 2, p. 418–431, 1986. Citations on pages 26 e 40.

LINDQVIST, B. H. On the statistical modeling and analysis of repairable systems. **Statistical science**, JSTOR, p. 532–551, 2006. Citation on page 25.

LIU, X. Planning of accelerated life tests with dependent failure modes based on a gamma frailty model. **Technometrics**, Taylor & Francis, v. 54, n. 4, p. 398–409, 2012. Citation on page 26.

LIU, X.; TANG, L.-C. Accelerated life test plans for repairable systems with multiple independent risks. **IEEE Transactions on Reliability**, IEEE, v. 59, n. 1, p. 115–127, 2010. Citation on page 25.

MAZZUCHI, T. A.; SOYER, R. A bayesian perspective on some replacement strategies. **Reliability Engineering & System Safety**, Elsevier, v. 51, n. 3, p. 295–303, 1996. Citation on page 23.

MEEKER, W. Q.; ESCOBAR, L. A. **Statistical methods for reliability data**. [S.l.]: John Wiley & Sons, 2014. Citations on pages 25 e 32.

MIGON, H. S.; GAMERMAN, D.; LOUZADA, F. **Statistical inference: an integrated approach**. [S.l.]: Chapman and Hall/CRC, 2014. Citation on page 25.

MOESCHBERGER, M. L. Life tests under dependent competing causes of failure. **Technometrics**, Taylor & Francis Group, v. 16, n. 1, p. 39–47, 1974. Citation on page 26.

MÜLLER, P.; QUINTANA, F. A. Nonparametric bayesian data analysis. **Statistical science**, JSTOR, p. 95–110, 2004. Citations on pages 26, 42, 43 e 46.

NEAL, R. M. *et al.* Mcmc using hamiltonian dynamics. **Handbook of Markov Chain Monte Carlo**, v. 2, n. 11, p. 2, 2011. Citations on pages 74 e 80.

O'HAGAN, A. **Kendall's Advanced Theory of Statistics, Vol. 2b: Bayesian Statistics**. [S.l.]: John Wiley & Sons, 1994. Citation on page 45.

OLIVEIRA, M. D. de; COLOSIMO, E. A.; GILARDONI, G. L. Bayesian inference for power law processes with applications in repairable systems. **Journal of Statistical Planning and Inference**, Elsevier, v. 142, n. 5, p. 1151–1160, 2012. Citations on pages 24, 34, 35 e 37.

PAWITAN, Y. **In all likelihood: statistical modelling and inference using likelihood**. [S.l.]: Oxford University Press, 2001. Citation on page 42.

PIEVATOLO, A.; RUGGERI, F. Bayesian reliability analysis of complex repairable systems. **Applied Stochastic Models in Business and Industry**, Wiley Online Library, v. 20, n. 3, p. 253–264, 2004. Citation on page 24.

PINTILIE, M. **Competing risks: a practical perspective**. [S.l.]: John Wiley & Sons, 2006. Citation on page 38.

R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2016. Available: <https://www.R-project.org>. Citations on pages 64, 83 e 84.

RAMOS, P. L.; LOUZADA, F.; RAMOS, E. An efficient, closed-form map estimator for nakagami-m fading parameter. **IEEE Communications Letters**, IEEE, v. 20, n. 11, p. 2328–2331, 2016. Citation on page 46.

____. Posterior properties of the nakagami-m distribution using noninformative priors and applications in reliability. **IEEE Transactions on Reliability**, IEEE, v. 67, n. 1, p. 105–117, 2018. Citation on page 46.

REIS, R. C. P. dos; COLOSIMO, E. A.; GILARDONI, G. L. *et al.* Hierarchical modelling of power law processes for the analysis of repairable systems with different truncation times: An empirical bayes approach. **Brazilian Journal of Probability and Statistics**, Brazilian Statistical Association, v. 33, n. 2, p. 374–396, 2019. Citation on page 33.

RIGDON, S. E.; BASU, A. P. **Statistical methods for the reliability of repairable systems**. [S.l.]: Wiley New York, 2000. Citations on pages 23, 24, 31, 32, 33, 34, 35, 36, 55, 56, 64, 70, 84 e 87.

RUGGERI, F. On the reliability of repairable systems: methods and applications. In: **Progress in Industrial Mathematics at ECMI 2004**. [S.l.]: Springer, 2006. p. 535–553. Citation on page 24.

SALINAS-TORRES, V.; PEREIRA, C.; TIWARI, R. Bayesian nonparametric estimation in a series system or a competing-risks model. **Journal of Nonparametric Statistics**, Taylor & Francis, v. 14, n. 4, p. 449–458, 2002. Citation on page 74.

SCHERVISH, M. J. **Theory of statistics**. [S.l.]: Springer Science & Business Media, 2012. Citations on pages 42 e 45.

SEN, A. Bayesian estimation and prediction of the intensity of the power law process. **Journal of Statistical computation and simulation**, Taylor & Francis, v. 72, n. 8, p. 613–631, 2002. Citation on page 24.

SETHURAMAN, J. A constructive definition of dirichlet priors. **Statistica sinica**, JSTOR, p. 639–650, 1994. Citations on pages 47, 74 e 78.

SLIMACEK, V.; LINDQVIST, B. H. Nonhomogeneous poisson process with nonparametric frailty. **Reliability Engineering & System Safety**, Elsevier, v. 149, p. 14–23, 2016. Citations on pages 34, 90 e 91.

____. Nonhomogeneous poisson process with nonparametric frailty and covariates. **Reliability Engineering & System Safety**, Elsevier, v. 167, p. 75–83, 2017. Citations on pages 34, 90 e 91.

SMITH, A. F.; ROBERTS, G. O. Bayesian computation via the gibbs sampler and related markov chain monte carlo methods. **Journal of the Royal Statistical Society: Series B (Methodological)**, Wiley Online Library, v. 55, n. 1, p. 3–23, 1993. Citation on page 74.

SOMBOONSAVATDEE, A.; SEN, A. Parametric inference for multiple repairable systems under dependent competing risks. **Applied Stochastic Models in Business and Industry**, Wiley Online Library, v. 31, n. 5, p. 706–720, 2015. Citations on pages 25, 26, 40, 41, 86 e 87.

____. Statistical inference for power-law process with competing risks. **Technometrics**, Taylor & Francis, v. 57, n. 1, p. 112–122, 2015. Citations on pages 25, 36, 38, 49, 51, 54, 55, 57, 65, 70, 72, 75 e 87.

TIBSHIRANI, R. Noninformative priors for one parameter of many. **Biometrika**, Biometrika Trust, v. 76, n. 3, p. 604–608, 1989. Citation on page 45.

TODINOV, M. **Reliability and risk models: setting reliability requirements**. [S.l.]: John Wiley & Sons, 2015. Citations on pages 38 e 56.

TOMAZELLA, V. L. D. **Modelagem de dados de eventos recorrentes via processo de Poisson com termo de fragilidade.** Phd Thesis (PhD Thesis) — Universidade de São Paulo, 2003. Citations on pages 40 e 41.

VAUPEL, J. W.; MANTON, K. G.; STALLARD, E. The impact of heterogeneity in individual frailty on the dynamics of mortality. **Demography**, Springer, v. 16, n. 3, p. 439–454, 1979. Citation on page 40.

WALKER, S. G.; MALLICK, B. K. Hierarchical generalized linear models and frailty models with bayesian nonparametric mixing. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, Wiley Online Library, v. 59, n. 4, p. 845–860, 1997. Citations on pages 26 e 41.

WANG, B. X.; XIE, M.; ZHOU, J. X. Generalized confidence interval for the scale parameter of the power-law process. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 42, n. 5, p. 898–906, 2013. Citation on page 36.

WIENKE, A. Frailty models. **Max Planck Institute for Demographic Research, MPIDR Working Papers: No. WP-2003-32**, 2003. Citations on pages 26 e 41.

WILLMOTT, C. J.; MATSUURA, K. Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. **Climate research**, v. 30, n. 1, p. 79–82, 2005. Citation on page 65.

WU, M.; SHI, Y.; ZHANG, C. Statistical analysis of dependent competing risks model in accelerated life testing under progressively hybrid censoring using copula function. **Communications in Statistics-Simulation and Computation**, Taylor & Francis, v. 46, n. 5, p. 4004–4017, 2017. Citation on page 26.

WU, S.; SCARF, P. Two new stochastic models of the failure process of a series system. **European Journal of Operational Research**, Elsevier, v. 257, n. 3, p. 763–772, 2017. Citation on page 38.

ZHANG, N.; YANG, Q. Optimal maintenance planning for repairable multi-component systems subject to dependent competing risks. **IIE Transactions**, Taylor & Francis, v. 47, n. 5, p. 521–532, 2015. Citations on pages 26 e 38.

ZHANG, X.; WILSON, A. System reliability and component importance under dependence: A copula approach. **Technometrics**, Taylor & Francis, v. 59, n. 2, p. 215–224, 2017. Citations on pages 26 e 38.

APPENDIX

A

TECHNICAL DETAILS

We present the deductions for some results from Chapter 3.

MAPs

The MAP estimate is the mode of the "posterior" distribution for θ (roughly speaking, one should choose the value for our parameters that is the most likely given the data). The MAP estimator using Jeffreys Prior is given by:

The likelihood function

$$\begin{aligned} L(\theta) = c \left[\beta_1^{n_1} e^{-n_1\beta_1/\hat{\beta}_1}\right] \left[\beta_2^{n_2} e^{-n_2\beta_2/\hat{\beta}_2}\right] \times \\ \left[e^{-\alpha_1}\alpha_1^{n_1}\right]\left[e^{-\alpha_2}\alpha_2^{n_2}\right]. \end{aligned}$$

The log-likelihood

$$\begin{aligned} \ell(\theta) = n_1 \log\beta_1 + n_2\log\beta_2 - \frac{n_1}{\hat{\beta}_1}\beta_1 - \frac{n_2}{\hat{\beta}_2}\beta_2 \\ + n_1\log\alpha_1 + n_2\log\alpha_2 - \alpha_1 - \alpha_2. \end{aligned}$$

The MLEs

$$\hat{\beta}_j^{MLE} = \frac{n_j}{\sum_{i=1}^n \log(T/t_i)(\mathbb{I}(\delta_i = j))} \quad \text{and} \quad \hat{\alpha}_j^{MLE} = n_j.$$

Jeffreys Prior

$$\pi^J(\theta) \propto \frac{1}{\beta_1\beta_2} \quad \xrightarrow{\log} \quad \log\pi^J(\theta) \propto -\log\beta_1 - \log\beta_2.$$

Definition of MAP estimator: It will be more convenient to find the argmax of the log of the MAP function

$$\begin{aligned} \hat{\theta}^{MAP} &= \arg\max_{\theta} \pi(\theta|t,\delta) \\ &= \arg\max_{\theta} L(\theta|t,\delta)\pi(\theta) \\ &= \arg\max_{\theta} \left(\log(\pi(\theta)) + \ell(\theta|t,\delta)\right). \end{aligned}$$

Applying the definition of MAP above, we obtain the MAP estimator of β_j for $j = 1,2$

$$\hat{\beta}_j^{MAP} = \arg\max_{\theta} \sum_{j=1}^{2} (n_j - 1) \log \beta_j - \sum_{j=1}^{2} \frac{n_j}{\hat{\beta}_j^{MLE}} \beta_j + \sum_{j=1}^{2} n_j \log \alpha_j - \sum_{j=1}^{2} \alpha_j.$$

The Score equations for β_j can be obtained

$$\frac{\partial \ell}{\partial \beta_j} = \frac{n_j - 1}{\hat{\beta}_j^{MAP}} - \frac{n_j}{\hat{\beta}_j^{MLE}} = 0.$$

Thus, the MAP estimator for β_j is given by

$$\hat{\beta}_j^{MAP} = \frac{n_j - 1}{n_j} \hat{\beta}_j^{MLE}.$$

For α_j follow the same steps

$$\frac{\partial \ell}{\partial \alpha_j} = \frac{n_j}{\hat{\alpha}_j^{MAP}} - 1 = 0.$$

Thus, the MAP estimator for α_j is given by

$$\hat{\alpha}_j^{MAP} = n_j.$$

Likelihood kernel under reparametrization

In the following, we will show how to obtain the kernel of full likelihood function (3.4). For the sake of brevity and simplicity, we present only 2 causes of failure $j = 1,2$, $\beta_1 = \beta_2 = \beta$ and $\alpha_j = \left(\frac{T}{\mu_j}\right)^{\beta}$:

$$\begin{aligned}
L(\beta, \alpha | t, \delta) =& \beta^n \left(T \alpha_1^{-\frac{1}{\beta}}\right)^{-n_1 \beta} \left(T \alpha_2^{-\frac{1}{\beta}}\right)^{-n_2 \beta} \times \left[\prod^{n_1} t_i \prod^{n_2} t_i\right]^{\beta - 1} e^{-\alpha_1 - \alpha_2} \\
=& \beta^n \frac{T^{-n_1 \beta - n_2 \beta}}{\alpha_1^{-n_1} \alpha_2^{-n_2}} \left[\prod^{n_1} t_i \prod^{n_2} t_i\right]^{\beta - 1} e^{-\alpha_1} e^{-\alpha_2} \\
=& \beta^n T^{-\beta n} \alpha_1^{n_1} \alpha_2^{n_2} \frac{\left[\prod_{i=1}^{n_1} t_i \prod_{j=1}^{n_2} t_j\right]^{\beta}}{\left[\prod_{i=1}^{n_1} t_i \prod_{j=1}^{n_2} t_j\right]} e^{-\alpha_1} e^{-\alpha_2} \\
=& c \beta^n \left[T^{-n} \prod_{i=1}^{n_1} t_i \prod_{j=1}^{n_2} t_j\right]^{\beta} \left[\alpha_1^{n_1} e^{-\alpha_1}\right] \left[\alpha_2^{n_2} e^{-\alpha_2}\right],
\end{aligned}$$

where $c = \left[\prod_{i=1}^{n_1} t_i \prod_{j=1}^{n_2} t_j\right]^{-1}$ and where $\hat{\beta} = n/\sum_{i=1}^{n} \log(T/t_i) = n/\left\{n(\log(T)) - \sum_{i=1}^{n} \log(t_i)\right\}$. Note that

$$\begin{aligned}
\left[T^{-n}\prod_{i=1}^{n_1} t_i \prod_{j=1}^{n_2} t_j\right]^{\beta} &= \exp\left\{\log(\left[T^{-n}\prod_{i=1}^{n_1} t_i \prod_{j=1}^{n_2} t_j\right]^{\beta}\right\} \\
&= \exp\left\{\beta\left[-n\log(T) + \sum_{i=1}^{n_1}\log(t_i) + \sum_{j=1}^{n_2}\log(t_j)\right]\right\} \\
&= \exp\left\{-\beta\left[n\log(T) - \sum_{i=1}^{n}\log(t_i)\right]\right\} \\
&= \exp\left\{-n\beta/\hat{\beta}\right\} \\
&= e^{-n_1\beta/\hat{\beta}} e^{-n_2\beta/\hat{\beta}}.
\end{aligned}$$

Hence, it follows that the likelihood kernel is proportional to a product of independent gamma densities.

$$\begin{aligned}
L(\beta,\alpha|t,\delta) =& c[\beta^{n_1} e^{-n_1\beta/\hat{\beta}}][\beta^{n_2} e^{-n_2\beta/\hat{\beta}}]\left[\alpha_1^{n_1} e^{-\alpha_1}\right]\left[\alpha_2^{n_2} e^{-\alpha_2}\right] \\
=& c\left[\beta^{n} e^{-n\beta/\hat{\beta}}\right]\left[\alpha_1^{n_1} e^{-\alpha_1}\right]\left[\alpha_2^{n_2} e^{-\alpha_2}\right] \\
\propto& \gamma(\beta|n+1, n/\hat{\beta})\prod_{j=1}^{2}\gamma(\alpha_j|n_j+1,1).
\end{aligned}$$

APPENDIX

B

TECHNICAL DETAILS

In this appendix, we presented estimates of some Z_j's associated to cars 1, 17, 26, 161, 165 and 169, according to Figure 13 (these are the estimates that presented the highest values). For completeness, we also present here the Geweke diagnostic test for checking the convergence of the chains, as well as MCMC diagnostic plots, such as trace and autocorrelations for the HMC algorithm of some Z_js.

Table 12 – Bayesian estimates of some Z_js with their SD.

Z_j	Bayes	SD
Z_1	2.469	1.711
$\vdots$	$\vdots$	$\vdots$
Z_{17}	5.1	2.87
$\vdots$	$\vdots$	$\vdots$
Z_{26}	8.269	3.669
$\vdots$	$\vdots$	$\vdots$
Z_{161}	11.519	4.359
$\vdots$	$\vdots$	$\vdots$
Z_{165}	9.941	4.038
$\vdots$	$\vdots$	$\vdots$
Z_{169}	6.615	3.354

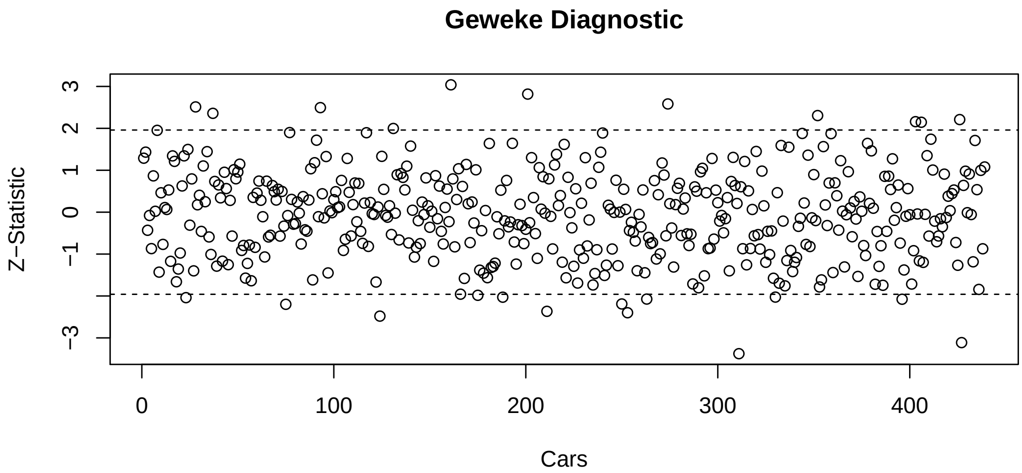

Figure 16 – Geweke diagnostic test - implemented using CODA package in R software.

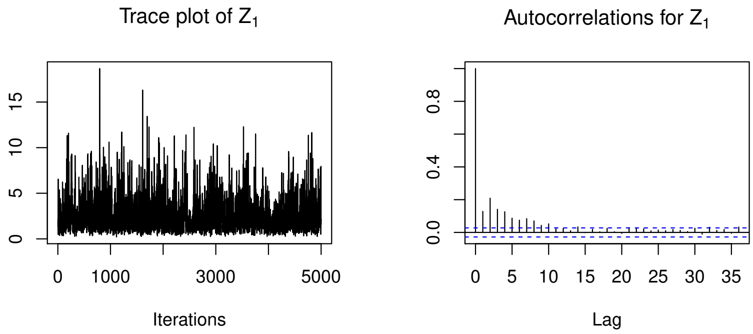

Figure 17 – Markov chain and autocorrelation plots for the HMC algorithm - Z_1.

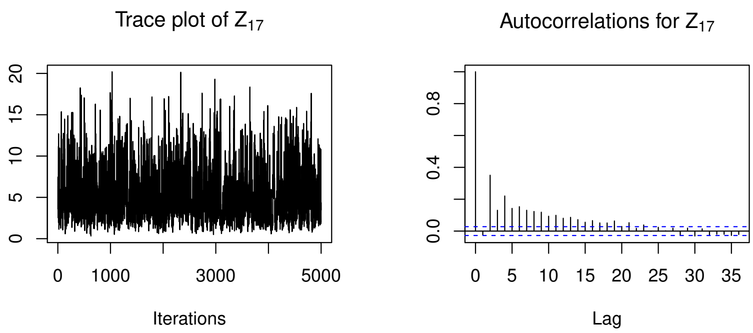

Figure 18 – Markov chain and autocorrelation plots for the HMC algorithm - Z_{17}.

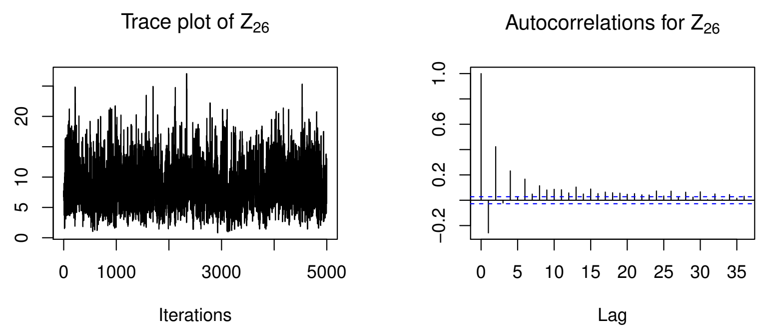

Figure 19 – Markov chain and autocorrelation plots for the HMC algorithm - Z_{26}.

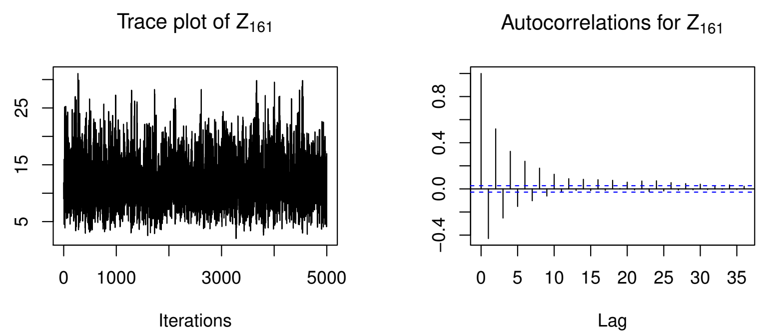

Figure 20 – Markov chain and autocorrelation plots for the HMC algorithm - Z_{161}.

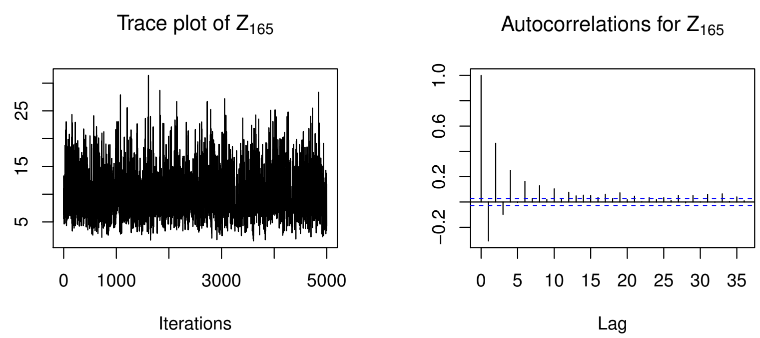

Figure 21 – Markov chain and autocorrelation plots for the HMC algorithm - Z_{165}.

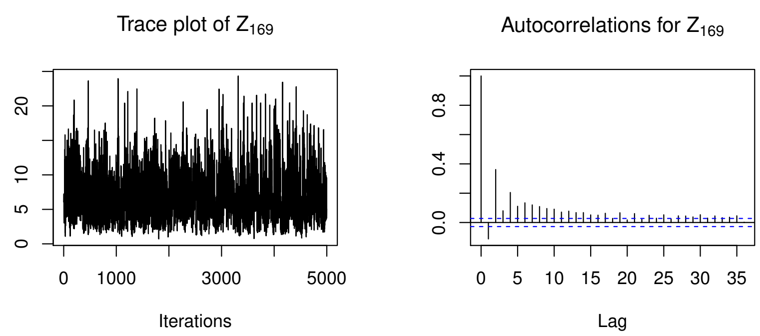

Figure 22 – Markov chain and autocorrelation plots for the HMC algorithm - Z_{169}.

APPENDIX

C

TECHNICAL DETAILS

Table 13 – Data structure - Observations for m systems with K competing risks.

System	Competing Risks (δ)	Failure times (t_{ji})	Number of failures (n_{jq})
	1	$t_{11}, t_{12}, \ldots, t_{1n_{11}}$	n_{11}
1	2	$t_{11}, t_{12}, \ldots, t_{1n_{12}}$	n_{12}
	$\vdots$	$\vdots$	$\vdots$
	K	$t_{11}, t_{12}, \ldots, t_{1n_{1K}}$	n_{1K}
	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$\vdots$	$\vdots$	$\vdots$
	$\vdots$	$\vdots$	$\vdots$
	1	$t_{m1}, t_{m2}, \ldots, t_{mn_{m1}}$	n_{m1}
m	2	$t_{m1}, t_{m2}, \ldots, t_{mn_{m2}}$	n_{m2}
	$\vdots$	$\vdots$	$\vdots$
	K	$t_{m1}, t_{m2}, \ldots, t_{mn_{mK}}$	n_{mK}